

# A phase transition for the biased tree-builder random walk

Arthur Blanc-Renaudie\*, Camille Cazaux†  
Guillaume Conchon-Kerjan‡, Tanguy Lions§, Arvind Singh¶

December 30, 2024

## Abstract

We consider a recent model of random walk that recursively grows the network on which it evolves, namely the Tree Builder Random Walk (TBRW). We introduce a bias  $\rho \in (0, \infty)$  towards the root, and exhibit a phase transition for transience/recurrence at a critical threshold  $\rho_c = 1 + 2\bar{\nu}$ , where  $\bar{\nu}$  is the (possibly infinite) expected number of new leaves attached to the walker's position at each step. This generalizes previously known results, which focused on the unbiased case  $\rho = 1$ .

The proofs rely on a recursive analysis of the local times of the walk at each vertex of the tree, after a given number of returns to the root.

We moreover characterize the strength of the transience (law of large numbers and central limit theorem with positive speed) via standard arguments, establish recurrence at  $\rho_c$ , and show a condensation phenomenon in the non-critical recurrent case.

**Keywords:** random walk, random trees, multi-type branching processes, phase transition

**AMS Classification 2020:** Primary 60K35; 60K37 Secondary 05C80

## 1 Introduction

The Tree Builder Random Walk (TBRW), introduced in full generality in [10], is a prime instance of a random walk that modifies the local structure of the network along its trajectory, see e.g. [1] for a broad formulation and [9] for a partly intersecting model. In simple words, the TBRW is a random walk on a tree such that at time-step  $n$ , a random number  $\xi_n$  of leaves are attached to the current position of the walker. When  $(\xi_n)_{n \geq 1}$  is i.i.d., the TBRW is known to be transient and even ballistic [7, 10], with a positive speed [7, 14] and a renewal structure [14].

Hence, it is natural to seek to increase the chances of returning to the root. One way to do it is to decrease the probability to grow a new vertex at the  $n$ -th time step, i.e.  $\xi_n \sim \text{Ber}(p_n)$ , which makes the TBRW recurrent if  $p_n = n^{-\gamma}$  with  $\gamma > 1/2$  [6]. When  $\gamma \in (2/3, 1]$ , the tree on which the random walk evolves grows like a Barabási-Albert preferential attachment model [5].

---

\*CNRS and LMRS, Université de Rouen. Email: [ablancrenaudiepro@gmail.com](mailto:ablancrenaudiepro@gmail.com)

†LPSM, Sorbonne Université. Email: [ccazaux@lpsm.paris](mailto:ccazaux@lpsm.paris)

‡Department of Mathematics, King's College London. Email: [guillaume.conchon-kerjan@kcl.ac.uk](mailto:guillaume.conchon-kerjan@kcl.ac.uk)

§ENS Lyon. Email: [tanguy.lions@ens-lyon.fr](mailto:tanguy.lions@ens-lyon.fr)

¶CNRS and LMO, Université Paris-Saclay. Email: [arvind.singh@universite-paris-saclay.fr](mailto:arvind.singh@universite-paris-saclay.fr) Partially supported by ANR 19-CE40-0025 "ProGraM"

Note that the case  $\gamma \leq 1/2$  is still open, with the walk being conjectured to be transient for  $\gamma < 1/2$ .

Another intuitive option is to add a bias towards the root for the walker's trajectory, and this is the purpose of the present paper. Due to the local structure changing around the walker, this model is noticeably different from usual biased random walks, e.g. on Galton-Watson trees (see [12] and the rich literature that followed).

We consider the following model which constructs a growing sequence  $(\mathcal{T}_n, n \geq 0)$  of random trees together with a random walk  $(S_n, n \geq 0)$  moving on those trees.

Fix  $\rho > 0$  and a probability distribution  $\nu$  on  $\mathbb{N} = \{0, 1, \dots\}$  which are the two parameters of the model. In the following, we always assume that  $\nu(0) < 1$  (otherwise the model is trivial). At time 0, the tree  $\mathcal{T}_0$  consists of a single vertex  $o$  (the root) and the position of the walk is  $S_0 = o$ . Then, by induction, given  $\mathcal{T}_n$  and  $S_n$ :

- First, independently of everything else, we grow  $\mathcal{T}_{n+1}$  from  $\mathcal{T}_n$  by adding a random number  $\xi_n$  of leaves drawn according to  $\nu$ . All these leaves are attached to the vertex  $S_n$ . In particular, with probability  $\nu(0)$ , we have  $\mathcal{T}_{n+1} = \mathcal{T}_n$ .
- Second, we define  $S_{n+1}$  as a neighbor of  $S_n$  in  $\mathcal{T}_{n+1}$  chosen with bias  $\rho$  toward the root. More precisely, if  $S_n = v \in \mathcal{T}_n$  and if we denote by  $v^0$  the father of  $v$  and by  $v^1, \dots, v^k$  the children of  $v$  in  $\mathcal{T}_{n+1}$ , then

$$\mathbf{P}(S_{n+1} = v^i \mid \mathcal{T}_{n+1}, S_n) = \begin{cases} \frac{\rho}{k+\rho} & \text{if } i = 0, \\ \frac{1}{k+\rho} & \text{if } 1 \leq i \leq k. \end{cases}$$

By convention, the father of the root  $o$  is the root itself so that we do not have to treat this case separately. This means that when the walker is at the root, it has positive probability of staying at the root at the next time step - this amounts to adding a loop around  $o$  that the walker can take with the corresponding probability.

When  $\rho = 1$ ,  $S_n$  performs a simple random walk on the current tree, and this corresponds to the original model of the TBRW with  $s = 1$  and  $(\xi_n)$  i.i.d., where  $s$  is the number of consecutive steps taken by the walk between two additions of new leaves. In this setting, it is known that for any distribution  $\nu$ , the walk is transient and ballistic [10, 14].

In this paper, we demonstrate how varying the bias  $\rho$  towards the root affects the geometry of the tree  $\mathcal{T}_n$  and the behavior of the walk  $S_n$ . Our main result is the existence of a non-trivial phase transition for this model depending on  $(\rho, \nu)$ .

**Theorem 1.1.** *Let  $\bar{\nu}$  denote the (possibly infinite) expectation of  $\nu$ :*

$$\bar{\nu} = \sum_{k=0}^{\infty} k\nu(k) \in (0, \infty].$$

*For any  $\rho > 0$ , the following holds.*

- *If  $\rho < 1 + 2\bar{\nu}$ , then walk  $S_n$  is transient, i.e. each vertex created is visited only finitely many times by the walk a.s. In particular, the walk returns only finitely many times to the root a.s.*

- If  $\rho \geq 1 + 2\bar{\nu}$ , then the walk  $S_n$  is recurrent, i.e. every vertex is visited infinitely often a.s. by the walk. Furthermore,
  - If  $\rho = 1 + 2\bar{\nu}$ , then the walk is “null recurrent” in the sense that all return times to the root have infinite expectation.
  - If  $\rho > 1 + 2\bar{\nu}$ , then the walk is “positive recurrent” in the sense that all return times to the root have finite expectation.

**Remark 1.2.** • When  $\rho = 1$ , Theorem 1.1 recovers the fact that the walk is always transient, irrespectively of the distribution  $\nu$  (see [10, Theorem 1.3], which also holds for  $s$  odd and  $(\xi_n)$  uniformly elliptic - interestingly, [10, Theorem 1.2] proves that for  $s$  even and under mild assumptions on  $(\xi_n)$ , the walker will be recurrent).

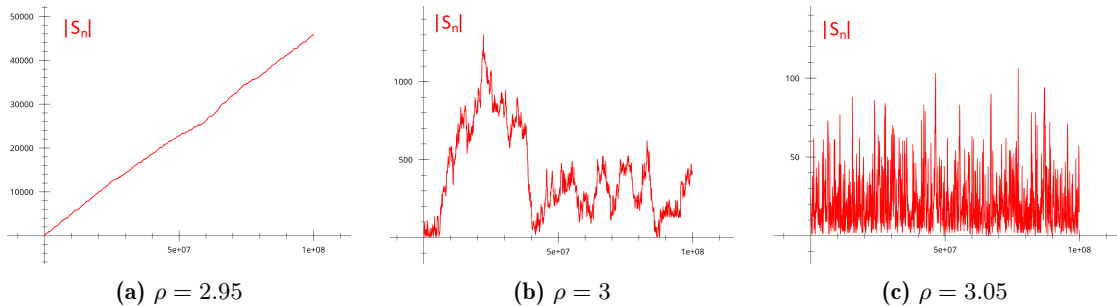
- Conversely, for any distribution  $\nu$  with finite expectation, the walk becomes recurrent whenever  $\rho$  is chosen large enough. This may seem surprising at first glance because, if we assume the walk to be recurrent, then the degree of every given vertex of the tree must ultimately increase to infinity, inducing local bias against the root which, intuitively, should in turn make the walk transient. A rough explanation for this apparent contradiction comes from the fact that, in the recurrent setting, even though all vertices have their degrees increasing to infinity, because the walk creates new vertices all the time, most of the vertices present at any given time have a current degree which is smaller than  $\rho$ , hence they induce a local bias toward the root which counterbalances the opposite bias created by the (few) vertices of large degree.
- When the walk is recurrent, every vertex created is visited infinitely often a.s. By a straightforward application of the Borel-Cantelli Lemma, this implies that every vertex has as a degree which increase to infinity so the sequence of tree  $\mathcal{T}_n$  converge locally to an infinite tree where every vertex has a (countably) infinite number of children, and yet, all of its vertices are visited infinitely often.
- We point out that the walk  $S_n$  is not a Markov process (but the pair  $(\mathcal{T}_n, S_n)$  is Markov) so the meaning of “positive recurrence” when  $\rho > 1 + 2\bar{\nu}$  differs from that for usual Markov chains. For instance, the excursions of the tree-builder walk  $S_n$  between two returns to the root are not i.i.d. as in the Markov case and the expectation of the  $k$ -th return time to the root grows faster than linearly (in fact exponentially in  $k$ , c.f. Lemma 3.2 and Lemma 4.5) and

$$\frac{1}{n} \sum_{k=0}^n \mathbf{1}_{\{S_n=o\}} \xrightarrow[n \rightarrow \infty]{} 0 \quad a.s.$$

In particular, the ergodic theorem for Markov chain does not apply here and the distribution of  $S_n$  does not converge to a non-degenerate distribution.

By definition, the number of vertices added to the tree at each step is given by a sequence of i.i.d. r.v. distributed as  $\nu$ . Therefore, by the strong law of large numbers (assuming  $\nu$  has finite expectation), the number of vertices  $\#\mathcal{T}_n$  of  $\mathcal{T}_n$  satisfies:

$$\#\mathcal{T}_n \underset{n \rightarrow \infty}{\sim} \bar{\nu}n \quad a.s.$$



**Figure 1:** Simulation of the first  $10^8$  steps of the height  $|S_n|$  of the tree-builder random walk when  $\nu(k) = \mathbf{1}_{k=1}$  (*i.e.* a new leaf is created at each step) in the transient (a), null recurrent (b) and positive recurrent (c) regime. Here, the phase transition occurs at the critical parameter  $\rho_c = 1 + 2\bar{\nu} = 3$ .

The asymptotic geometry of the tree  $\mathcal{T}_n$  differs depending on the transience/recurrence regime considered. In the transient case, we prove that the walk is ballistic, hence the tree is asymptotically a line onto which a grafted finite random trees. We use the notation  $|v|$  for the height of a vertex (*i.e.* its distance from the root) and define the maximum height of the tree  $\mathcal{T}_n$  by

$$|\mathcal{T}_n| := \max_{v \in \mathcal{T}_n} |v|.$$

**Theorem 1.3.** *Suppose that  $\rho < 1 + 2\bar{\nu}$ . There exist positive constants  $v = v(\rho, \nu)$  and  $\sigma = \sigma(\rho, \nu)$  such that*

$$\frac{|S_n|}{n} \xrightarrow[n \rightarrow \infty]{a.s.} v \quad \text{and} \quad \frac{|S_n| - v}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{law} \mathcal{N}(0, \sigma^2). \quad (1)$$

*The same results also hold with  $|\mathcal{T}_n|$  in place of  $|S_n|$  in (1).*

This generalizes [7, Theorem 1.1] and [14, Theorems 3 and 4(a)].

On the other hand, in the positive recurrent case, we observe a condensation of the tree whose height now grows logarithmically in  $n$ .

**Theorem 1.4.** *Suppose that  $\rho > 1 + 2\bar{\nu}$ . There exists positive constants  $c_1 = c_1(\rho, \nu)$  and  $c_2 = c_2(\rho, \nu)$  such that*

$$c_1 \log n \leq |\mathcal{T}_n| \leq c_2 \log n \quad \text{for all } n \text{ large enough a.s.}$$

In this paper, we leave out the study of the asymptotical behavior of the walk and the tree in the critical case.

**Question 1.5.** What is the typical height  $|S_n|$  and diameter  $|\mathcal{T}_n|$  in the critical case  $\rho = 1 + 2\bar{\nu}$ ? Numerical simulation suggest that these quantities should be of order  $\sqrt{n}$  as  $n$  tend to infinity, *c.f.* Figure 1b.

The rest of the paper is organized as follows. In section 2, we first set some notations and then prove a monotonicity property (Proposition 2.1) for the recurrence/transience of the TBRW with respect to its parameter  $(\rho, \nu)$  via a coupling argument. This allows us, later on,

to prove the recurrence/transience criterion under the assumption that  $\nu$  has finite expectation and subsequently extend it, by comparison, to the case where  $\bar{\nu} = \infty$ .

Next, we introduce two new processes: an urn process (section 2.3) and a branching Markov chain (section 2.4). These processes are closely related to the local times of the tree-builder walk (Proposition 2.2 and Proposition 2.7) and reduce the study of the recurrence/transience of the TBRW to the survival of a particular “infinite-type” branching process. Section 3 is devoted to proving Theorem 1.1 which is done by computing a (strikingly explicit) quasi-stationary distribution for the multi-type branching process, *c.f.* Corollary 3.5. Then, in section 4, we recall the classical method of cut/regeneration times and apply it to prove Theorem 1.3 and then carry on proving Theorem 1.4.

## 2 An urn process and a branching Markov chain related to the TBRW.

### 2.1 Notations

In the rest of the paper,  $(S_n, \mathcal{T}_n)$  always denotes the TBRW and associated tree at step  $n$ . Since the sequence  $(\mathcal{T}_n, n \geq 0)$  of random trees is non-decreasing (we do not remove vertices), it makes sense to consider the inductive limit

$$\mathcal{T}_\infty := \bigcup_n \mathcal{T}_n.$$

We point out that  $\mathcal{T}_\infty$  is a (random) rooted tree but it needs not be locally finite: a vertex may have an infinite number of children. By the Borel-Cantelli Lemma, this happens i.f.f.  $S_n$  visits this vertex infinitely many times since, at each visit and independently of everything happening before, it has probability  $1 - \nu(0) > 0$  of creating at least one new child.

The trees  $\mathcal{T}_n$  (and also  $\mathcal{T}_\infty$ ) have a natural planar embedding, i.e. an ordering of the children of any vertex, given by their order of creation. Let us set some notation. For a vertex  $v \in \mathcal{T}_n$ , we define  $\#_n v$  to be the number of children of  $v$  in  $\mathcal{T}_n$ , and we denote  $v^1, v^2, \dots$  the children of  $v$  (ordered by time of creation) and by  $v^0$  the father of  $v$ . By convention, we set  $o^0 = o$ . Note that by compatibility of the trees  $\mathcal{T}_n$ , we do not need to add a subscript  $n$  to the children (or the father) of  $v$  as they are the same in all trees where they exist. Given a vertex  $v$ , we write  $|v|$  for its height defined as the distance between this vertex and the root  $o$ . Once again, this definition does not depend on the tree considered (provided  $v$  exists in it).

Given a rooted tree  $\mathcal{T}$  and  $d \in \mathbb{N}$ , we define  $\mathcal{T}^{\leq d}$  the subtree composed of all vertices of height at most  $d$ . Finally, we denote by  $\mathcal{T}_{\text{full}}$  the “full” planar tree where every vertex  $v$  has an infinite number of children  $v^1, v^2, \dots$ . Any planar tree  $\mathcal{T}$  can therefore be seen as a subtree of  $\mathcal{T}_{\text{full}}$ .

For two probability distributions  $\nu, \tilde{\nu}$  (on  $\mathbb{N}$ ), we write  $\nu \prec \tilde{\nu}$ , meaning that  $\tilde{\nu}$  dominates  $\nu$  for the stochastic partial order, if  $\tilde{\nu}([k, \infty)) \geq \nu([k, \infty))$  for all  $k \geq 0$ .

### 2.2 Coupling of walks for different values of the parameters

One should expect the tree-builder walks to become “more recurrent” when  $\rho$  increases or when the distribution  $\nu$  of the number of children created at each step decreases (stochastically). In order to formalize this idea, we construct couplings between processes with different parameters.

**Proposition 2.1.** Fix  $\rho \geq \tilde{\rho} > 0$  and two distributions  $\nu \prec \tilde{\nu}$ . If the  $(\tilde{\rho}, \tilde{\nu})$ -TBRW goes through the root loop infinitely often a.s., then the  $(\rho, \nu)$ -TBRW does the same a.s.

*Proof.* We construct two TBRWs  $(S_n, \mathcal{T}_n)$  and  $(\tilde{S}_n, \tilde{\mathcal{T}}_n)$  with respective parameters  $(\rho, \nu)$  and  $(\tilde{\rho}, \tilde{\nu})$  on the same probability space, by a natural extension of the coupling described in [14, Lemma 1] for  $\rho = \tilde{\rho}$  and  $\nu = \text{Ber}(1 - \tilde{\nu}(0))$ .

Roughly speaking, we decompose the trajectory of  $\tilde{S}$  where it evolves together with  $S$ , and excursions in subtrees above a given position of  $S$ . This can happen either when  $S$  moves towards the root while  $\tilde{S}$  moves away, coupling their jumps while using that  $\rho > \tilde{\rho}$ , or when  $\tilde{S}$  moves to a child of  $\tilde{\mathcal{T}} \setminus \mathcal{T}$ , using that  $\nu \prec \tilde{\nu}$ . For any such excursion, either it never ends, in which case  $\tilde{S}$  is transient, or it comes back to the position where we froze  $S$ . Then the two walks resume their joint evolution, until a next time when  $\tilde{S}$  moves away from the root and  $S$  towards it, and so on. We note that while  $S$  is frozen, the excursions of  $\tilde{S}$  will create new vertices on the tree  $\tilde{\mathcal{T}}_n$  (compared to  $\mathcal{T}_n$ ) but these additional vertices only make the walk  $\tilde{S}$  more transient compared to  $S$  so that they are not problematic.

We consider two i.i.d. sequences  $(\xi_{k,v}, v \in \mathcal{T}_{\text{full}}, k \in \mathbb{N}^*)$  and  $(\tilde{\xi}_{k,v}, v \in \mathcal{T}_{\text{full}}, k \in \mathbb{N}^*)$  such that  $\xi_{k,v}$  (resp.  $\tilde{\xi}_{k,v}$ ) has distribution  $\nu$  (resp.  $\tilde{\nu}$ ) and such that

$$\xi_{k,v} \leq \tilde{\xi}_{k,v} \text{ a.s.}, \quad (2)$$

which is possible since  $\nu \prec \tilde{\nu}$ . We will impose that at the  $k$ -th visit of  $v$  by  $S$  (resp.  $\tilde{S}$ ), we add  $\xi_{k,v}$  (resp.  $\tilde{\xi}_{k,v}$ ) neighbours to  $v$  in the corresponding TBRW trees.

Construct  $\tilde{S}$  according to the  $(\tilde{\rho}, \tilde{\nu})$ -TBRW dynamics. We now construct  $S$ . Initially, we start with  $S_0 = \tilde{S}_0 = o$  and  $\mathcal{T}_1$  is  $o$  with  $\xi_{1,o}$  leaves attached to it. For  $n \geq 0$ , let  $t_n$  be the first time  $t > t_{n-1}$  such that  $\tilde{S}_t = S_n =: v$ ,  $\mathcal{T}_{n+1} \subseteq \tilde{\mathcal{T}}_{t+1}$ , and  $\tilde{S}_{t+1}$  is either  $v^0$ , or a child of  $v$  in  $\mathcal{T}_{n+1}$ , with the possibility that  $t_n = \infty$ .

If  $t_n < \infty$ , there are two cases. If  $\tilde{S}_{t+1} = v^0$ , we let  $S_{n+1} = v^0$ . And if  $\tilde{S}_{t+1} =: v'$  is a child of  $v$  in  $\mathcal{T}_{n+1}$ , we let  $x := \xi_{1,v} + \dots + \xi_{k,v}$ ,  $S_n$  being the  $k$ -th visit of  $v$  by  $S$ . With probability  $b := \frac{\rho}{\rho+x} - \frac{\tilde{\rho}}{\tilde{\rho}+x}$  we let  $S_{n+1} = v^0$ , and with probability  $1 - b$  we set  $S_{n+1} = v'$ . Note that  $b \geq 0$  since  $\rho \geq \tilde{\rho}$ .

By induction on  $n$ , the first  $t > t_{n-1}$  such that  $\tilde{S}_t = S_n =: v$  and  $\tilde{S}_{t+1}$  is either  $v^0$ , or a child of  $v$  in  $\mathcal{T}_n$  also satisfies that  $\mathcal{T}_{n+1} \subseteq \tilde{\mathcal{T}}_{t+1}$ . Indeed, at any vertex of  $\mathcal{T}_{\text{full}}$  the local time  $k$  of  $(S_j)_{j \leq n}$  is no greater than that  $\tilde{k}$  of  $(\tilde{S}_j)_{j \leq t}$  since  $S$  always waits for  $\tilde{S}$  to catch it before moving again. By (2), we have thus  $\xi_{1,v} + \dots + \xi_{k,v} \leq \tilde{\xi}_{1,v} + \dots + \tilde{\xi}_{k,v} \leq \tilde{\xi}_{1,v} + \dots + \tilde{\xi}_{\tilde{k},v}$ . Hence,  $t$  is in fact  $t_n$ .

Moreover, in our construction, the first case happens with probability  $\frac{\tilde{\rho}}{\tilde{\rho}+x} = \frac{\rho}{\rho+x} - b$  so that  $S_{n+1}$  moves to  $v^0$  with total probability  $\frac{\rho}{\rho+x}$ , and to any given child of  $v$  in  $\mathcal{T}_{n+1}$  with probability  $x^{-1}(1 - \rho/(\rho+x)) = 1/(\rho+x)$ . Hence,  $S$  performs a  $(\rho, \nu)$ -TBRW. Furthermore, one has easily by induction on  $n$  that  $S_n$  is always an ancestor of  $\tilde{S}_t$  for  $t_{n-1} + 1 \leq t \leq t_n$ , and that if  $\tilde{S}_t = S_n$  for some  $t > t_{n-1}$ , then  $\tilde{S}$  only moves towards  $o$  if so does  $S$ . Thus

$$(S_j)_{j \leq n} \text{ visits } o \text{ at least as many times as } (\tilde{S}_j)_{j \leq t_n}. \quad (3)$$

Taking  $n = \min\{j, t_j = \infty\} - 1$  if this minimum exists or  $n \rightarrow \infty$  else yields the conclusion.  $\square$

### 2.3 The urn process associated with the local time at a vertex

We now introduce an auxiliary urn process, the *TBRW urn*, which captures the local behavior of the TBRW around a fixed vertex and whose detailed analysis will play a critical role in the rest of the paper. Initially, we start with the urn containing  $\rho$  balls<sup>1</sup> of color 0. At each time step  $n \geq 1$ , we make the following operation.

- (a) First, independently of everything else, we add a random number of new balls sampled according to the distribution  $\nu$ . Each added ball is given a new color, whose index is the smallest color that has not yet been attributed (so if the colors already present in the urn are  $0, 1, \dots, k$  and we add  $j$  new balls, then they are given colors  $k+1, k+2, \dots, k+j$ ).
- (b) Second, we draw a uniform random ball from the urn, we record its color  $B_n$ , and we put it back in the urn.

We construct in this way a random sequence  $B_1, B_2 \dots$  of integers representing the colors of the balls drawn from the urn at each step.

This urn process and the TBRW are related in the following way: fix a vertex  $v \in \mathcal{T}_{\text{full}}$  and look at the times when the walk  $S$  is at  $v$  (if ever). The crucial observation is that when  $S$  is at  $v$ , the displacement rule is local, meaning that it only depends on the current number of children of  $v$  and on an independent additional randomness (to make one step given this local environment). But it does not depend on what happens to the walk when it is not at  $v$ . Moreover, conditionally on having  $k$  children at  $v$ , the probability of jumping to  $v^i$  for  $i \in \{0, 1, \dots, k\}$  is equal to the probability of drawing the ball of color  $i$  in the urn currently containing colors  $0, 1, \dots, k$ . Putting these arguments together, it is immediate to check that the TBRW can be constructed from a family of independent TBRW urns as follows:

**Proposition 2.2** (Construction of the TBRW from an i.i.d. family of urns). *The TBRW  $S$  can be constructed in the following way:*

1. We initially put an urn  $U_v$  for every  $v \in \mathcal{T}_{\text{full}}$ , containing  $\rho$  balls of color 0.
2. We start the walk at  $S_0 = o$ .
3. When  $S$  is at a vertex  $v$ , we perform (a) and (b) above on the urn  $U_v$  (independently of all past operations).  $S$  then moves to the vertex  $v^B$  where  $B$  is the color of the ball drawn.

We insist on the fact that the urns are independent. In order to study the recurrence or transience of the walk, it is therefore crucial to understand the asymptotic behavior of the TBRW urn and, in particular, to check whether a given color is picked infinitely often a.s.

**Lemma 2.3.** *Suppose that*

$$\mathbf{P}(\text{color 0 is drawn a finite number of times from the TBRW urn}) > 0. \quad (4)$$

*Then,*

$$\mathbf{P}(\text{color 0 is never drawn from the TBRW urn}) > 0.$$

---

<sup>1</sup> $\rho$  need not be an integer, so that the urn can contain a non-integer number of balls, but this does not matter from the mathematical point of view.

*Proof.* We can construct the urn process by first prescribing the number of balls added at each step according to an i.i.d. sequence  $\xi = (\xi_i)_{i \geq 1}$  with distribution  $\nu$  and then, conditionally on  $\xi$ , by performing the draws  $B_1, B_2, \dots$  independently. We can also assume without loss of generality that  $\xi_1 > 0$  since we can discard all the draws (which are necessarily of color 0) until a non-zero color ball is added to the urn. Then the process after that time is the same as the initial urn process conditioned on  $\xi_1 > 0$ .

Now, if (4) holds, there must exist  $k \geq 1$  and indices  $i_1, i_2, \dots, i_k$  such that, with positive probability, color 0 is drawn from the urn at times  $i_1, \dots, i_k$  and at no other time. For each draw  $B_{i_j}$ , we have  $\mathbf{P}(B_{i_j} = 0 | \xi) \mathbf{P}(B_{i_j} \neq 0 | \xi) > 0$  (since the urn always contains a non-zero color). This means that we can change the color  $B_{i_j}$ , for  $j = 1, \dots, k$ , to non-zero values while still keeping an event of positive probability. On that new event, color 0 is never drawn.  $\square$

**Corollary 2.4** (Dichotomy for the recurrence/transience of the TBRW). *For any vertex  $v \in \mathcal{T}_{\text{full}}$ , we have:*

$$\mathbf{P}(S \text{ visits } v \text{ i.o. but jumps from } v \text{ to } v^0 \text{ finitely many times}) = 0. \quad (5)$$

As a consequence, we have

$$\mathbf{P}\left(S \text{ crosses the root loop infinitely many times}\right) + \mathbf{P}\left(|S_n| \xrightarrow[n \rightarrow \infty]{} \infty\right) = 1$$

*Proof.* On the event that  $v$  is visited infinitely often but  $v^0$  only finitely many times, the urn  $U_v$  at vertex  $v$  (defined in Proposition 2.2) draws color 0 only finitely many times. If this event has positive probability and since the urns at all the vertices of  $\mathcal{T}_{\text{full}}$  are i.i.d., it follows from Lemma 2.3 that each one has a positive probability to never draw 0. But for any  $w \in \mathcal{T}_{\text{full}}$ , if 0 is never drawn from  $U_w$ , this means that the walk gets trapped inside the subtree rooted at  $w$  when it enters it. Since the walk a.s. visits infinitely many vertices (the probability of staying on the same finite number of vertices while accumulating a diverging number of neighbors being zero), this must happen an infinite number of time a.s. Hence,  $|S_n| \rightarrow \infty$  a.s. However, this is incompatible with the assumption that  $v$  is visited infinitely often, hence that event must have null probability.

To see why the second statement holds true, we recall that  $\mathcal{T}_{\text{full}}$  is countable and observe that on the event  $\mathcal{E} := \{\text{there exists a vertex } v \in \mathcal{T}_{\text{full}} \text{ that is visited i.o.}\}$ , then, by a straightforward induction and thanks to (5), the root loop must also be crossed infinitely many times. Conversely, on the complementary event  $\mathcal{E}^c = \{\text{every } v \in \mathcal{T}_{\text{full}} \text{ is visited finitely many times}\}$ , the number of children of each vertex is necessarily finite, hence the subtree of vertices at distance at most  $\ell \geq 1$  from the root is also finite, which implies that  $|S_n| \geq \ell$  for  $n$  large enough since each vertex of this subtree is visited only finitely many times. This is true for all  $\ell$  so we conclude that  $|S_n| \rightarrow \infty$  on  $\mathcal{E}^c$ .  $\square$

In the rest of the paper, we will be mostly concerned with the case where  $\nu$  has finite expectation. In that setting the urn is “well-behaved” as the next result shows.

**Lemma 2.5.** *Suppose that  $\bar{\nu} < \infty$ . Then, the number of colors present in the basic urn increases to infinity and every color is drawn infinitely many times a.s., i.e.*

$$\sum_{n \geq 1} \mathbf{1}_{\{B_n = j\}} = \infty \quad \text{a.s. for all } j \geq 0.$$



*Proof.* By the law of large numbers, the number of balls in the urn at step  $n$  is a.s. asymptotically equivalent to  $\bar{\nu}n$  as  $n$  tend to infinity. Furthermore, the probability  $\mathbf{P}(B_n = 0)$  of picking up a ball of color 0 at step  $n$  is of order  $\rho/(\bar{\nu}n)$  so that by the Borel-Cantelli Lemma, a ball of color 0 is drawn from the urn infinitely many times a.s. Similarly,  $\mathbf{P}(B_n = j) \sim 1/(\bar{\nu}n)$  for  $j \geq 1$  and  $n$  larger than the time at which the color  $j$  was introduced in the urn, this said time being finite a.s. Therefore, any color is picked up infinitely many times a.s.  $\square$

We now let  $\Theta_0 := 0$  and, for  $k \geq 1$ , we define the stopping time at which color 0 is drawn from the urn for the  $k$ -th time:

$$\Theta_k := \inf \left\{ n \geq 1, \sum_{i=1}^n \mathbf{1}_{\{B_i=0\}} = k \right\}.$$

We also define  $N_0 := 0$  and for  $k \geq 1$  and  $j \geq 1$ ,

$$N_k := \begin{array}{l} \text{number of non-zero colors present in the urn at step } \Theta_k \\ \text{(i.e. when color 0 has been drawn } k \text{ times).} \end{array} \quad (6)$$

$$Y_k^j := \sum_{i=1}^{\Theta_k} \mathbf{1}_{\{B_i=j\}} = \begin{array}{l} \text{number of times color } j \text{ has been drawn} \\ \text{at the } k\text{-th draw of color 0.} \end{array} \quad (7)$$

Thanks to the previous lemma, all the quantities above are well-defined and finite a.s. as soon as  $\bar{\nu} < \infty$ . Thus, from now on and until section 4, we will always assume that

$$\bar{\nu} = \sum_k k\nu(k) < \infty. \quad (8)$$

## 2.4 The branching Markov chain $Z$

Proposition 2.2 show that we can construct the TBRW from urn processes but it does not really help to understand the trajectory of the walk  $S$  i.e. the order in which the urns are chosen. To this end, we now construct a branching Markov chain whose transition kernel  $\mathbf{K}$  is defined in terms of the urn process. We set

$$\begin{cases} \mathbf{K}(k, [y_1, \dots, y_j]) = \mathbf{P}(N_k = j \text{ and } Y_k^1 = y_1, \dots, Y_k^j = y_j) & \text{for } j \geq 1, \\ \mathbf{K}(k, []) = \mathbf{P}(N_k = 0). \end{cases}$$

In words,  $\mathbf{K}$  is the kernel which, given a non-negative integer  $k$ , returns a finite (possibly empty) list representing, for each non-zero color present in the urn, the number of times it has been drawn at the moment when color 0 is drawn for the  $k$ -th time. In particular, we have  $\mathbf{K}(0, []) = 1$  because no non-zero color is initially present in the urn. Note also that a color  $i$  can be present in the urn without having being drawn so that  $Y_k^i = 0$ . This is not the same thing as the color not yet being present in the urn.

We use this kernel to construct a tree-indexed particle system, where each particle has a type  $k \in \mathbb{N}$  and reproduces itself by creating offspring according to  $\mathbf{K}$ . More precisely, we define a branching Markov chain process  $Z = (Z(v), v \in \mathcal{T})$  with the following rules:

- (a)  $\mathcal{T}$  is the rooted planar tree representing the genealogy of the process. A vertex  $v \in \mathcal{T}$  is called a *particle* and  $Z(v) \in \mathbb{N}$  is its *type*. The *generation* of a particle  $v$  corresponds to its distance  $|v|$  from the root particle  $o$ .

- (b) We start initially from a single root particle  $o$  with type  $Z(o) \in \mathbb{N}$  (generation 0).
- (c) Suppose that we have constructed the process up to generation  $n$ . Then, all the particles at generation  $n$  evolve independently. Each particle  $v$  (with  $|v| = n$ ) gives birth to a random finite number of children  $v^1, \dots, v^j$  with respective types  $Z(v^1), \dots, Z(v^j)$ . The distribution of the number of children and their types is given by the transition kernel  $\mathbf{K}$ :

$$\begin{cases} \mathbf{P}(v \text{ has } j \text{ children with } Z(v^1) = z_1, \dots, Z(v^j) = z_j \mid Z(v) = z) = \mathbf{K}(z, [z_1, \dots, z_j]) \\ \mathbf{P}(v \text{ does not have any child} \mid Z(v) = z) = \mathbf{K}(z, \emptyset) \end{cases}$$

- (d) If there is no particle at generation  $n$ , then we say that the process *dies out* and we stop the construction.

The process  $Z$  can be interpreted as a multi-type branching process with infinitely many types (indexed by  $\mathbb{N}$ ). We observe that, when both  $\nu(0) > 0$  and  $\nu(1) > 0$ , then  $Z$  is irreducible in the sense that, for any  $z \geq 1$  and any  $z_1, \dots, z_j \geq 0$ , a particle of type  $z$  has positive probability to have exactly  $j$  children with respective types  $z_1, \dots, z_j$ . In the general case, additional particles of type 0 may be unavoidable, but this is not a problem because particles of type 0 play no role in the genealogy of the process.

**Proposition 2.6** (irreducibility of  $Z$  when omitting particles of type 0). *Particles of type 0 in the branching Markov chain  $Z$  never reproduce. For any  $z \geq 1$  and any  $z_1, \dots, z_j \geq 1$ , there exists  $n \geq 0$  (which may depend on  $z, z_1, \dots, z_j$ ) such that:*

$$\mathbf{P}\left(\begin{array}{l} v \text{ has } j+n \text{ children with } Z(v^1) = z_1, \dots, Z(v^j) = z_j \\ \text{and } Z(v^{j+1}) = 0, \dots, Z(v^{j+n}) = 0 \end{array} \mid Z(v) = z\right) > 0. \quad (9)$$

Furthermore, for all  $z \geq 1$ ,

$$\mathbf{P}\left(v \text{ has at least one child of type } 1 \mid Z(v) = z\right) \geq \frac{(1 - \nu(0))\rho}{(1 + \rho)^2} > 0. \quad (10)$$

*Proof.* The proposition is a direct consequence of the definition of the kernel  $\mathbf{K}$  in terms of the urn process. As stated above when defining  $\mathbf{K}$ , the urn starts without any ball of non-zero color implying  $\mathbf{K}(0, \emptyset) = 1$ , which means that a particle of type 0 has a.s. no progeny.

Now, given  $z \geq 1$ , because we assume  $\nu(0) < 1$ , it is possible to construct a sequence of draws in the urn that creates at least  $j$  new colors before drawing a ball of color 0 for the  $z$ -th time and such that the first  $j$  balls are picked respectively  $z_1, \dots, z_j$  times while all balls of color  $\geq j+1$  are not picked. This proves (9).

Finally, fix  $z \geq 1$  and consider the urn process just after a ball of color 0 has been drawn for the  $z-1$ -th time. Then, with probability  $1 - \nu(0)$ , at least one new color is added to the urn at that step and such a color is subsequently drawn a geometric number of times, with parameter  $\rho/(1+\rho)$ , before drawing a ball of color 0 for the  $z$ -th time. Therefore, the probability that this ball is drawn exactly once (yielding a child of type 1) is equal to  $\frac{\rho}{(1+\rho)^2}$ . This proves (10).  $\square$

Our interest for the branching Markov chain  $Z$  comes from its close relation to the local time of the TBRW stopped at a crossing time of the loop at the root. Let us define  $\tau_k$  as the

first time that the tree-builder walk  $S_n$  crosses the root loop for the  $k$ -th time. More precisely, we set  $\tau_0 := 0$  and for  $k \geq 1$ ,

$$\tau_k := \inf \left\{ n \in \mathbb{N} : \sum_{i=1}^n \mathbf{1}_{\{S_{i-1}=S_i=o\}} = k \right\} \quad (11)$$

with the usual convention that  $\inf \emptyset = +\infty$ . We also define the local time process  $L(\cdot, k)$  of the walk  $S$  at time  $\tau_k$  on the (oriented) edge of the tree  $\mathcal{T}_{\tau_k}$  constructed so far:

$$L(v, k) := \sum_{n=1}^{\tau_k} \mathbf{1}_{\{S_{n-1}=v \text{ and } S_n=v^0\}} \quad \text{for } v \in \mathcal{T}_{\tau_k} \quad (12)$$

The following ‘‘Ray-Knight type’’ result highlights the connection between the local time process  $L$  and the branching Markov chain  $Z$  defined previously and will enable us to reduce questions about the recurrence/transience of the TBRW to questions pertaining the survival of  $Z$ .

**Proposition 2.7.** *Fix  $k \geq 1$ . The following statements hold true.*

- (a) *We have  $\mathbf{P}(\tau_k < \infty) = \mathbf{P}(Z \text{ dies out} \mid Z(o) = k)$ .*
- (b) *The process  $L(\cdot, k)$  under the conditional law  $\mathbf{P}(\cdot \mid \tau_k < \infty)$  has the same distribution as the process  $Z$  under the conditional law  $\mathbf{P}(\cdot \mid Z(o) = k \text{ and } Z \text{ dies out})$ .*
- (c) *The r.v.  $\tau_k$  under the conditional law  $\mathbf{P}(\cdot \mid \tau_k < \infty)$  has the same distribution as the r.v.  $k + 2 \sum_{v \in \mathcal{T} \setminus \{o\}} Z(v)$  under the conditional law  $\mathbf{P}(\cdot \mid Z(o) = k \text{ and } Z \text{ dies out})$ .*

**Remark 2.8.** The arguments of the proof provided below follow those given in Section 3.1 of [2] to describe the local time process of a multi-excited random walk on a regular tree (again stopped at a crossing of the root loop) *c.f.* in particular Lemma 3.2 of [2]. Let us point out that the general idea that a non-Markovian walk on an acyclic graph with local interactions can have a local time process with a Markovian structure more amenable to analysis is not new. It traces back at least to the seminal paper of Kesten, Kozlov and Spitzer [11] to study random walks in random environments. It was subsequently used for different kinds of self-interacting processes, e.g. self-repelling walks [18, 19] or reinforced walks [3]. Although our setting here is slightly different due to the constant evolution of the state space, the general method still applies since the graph is always acyclic.

*Proof of Proposition 2.7.* We fix an integer  $d \geq 1$  and we consider a modification  $(S^{(d)}, \mathcal{T}^{(d)})$  of the TBRW reflected at height  $d$ , that is, when the walk is at a site  $v$  at distance  $|v| = d$  from the root, then it always moves back towards its father  $v^0$  at the next step without creating any new site neighboring  $v$ . Otherwise, if  $|v| < d$ , it behaves like the regular TBRW. This new walk constructs a tree of height  $d$  in which, ultimately, all vertices at height  $< d$  have infinite degree (and vertices at height  $d$  are leaves). This assertion follows from the same argument use in the proof of Lemma 2.5 to show that the urn process draws each color infinitely many times. Indeed, when the walk is at a given vertex  $v$  for the  $k$ -th time, by the strong law of large numbers,  $v$  has  $\sim k\bar{v}$  children so that the walk moves to the father  $v^0$  with probability  $\sim \rho/(\bar{v}k)$ , and to a given child with probability  $\sim 1/(\bar{v}k)$ . Hence, all those events happen infinitely often a.s. by the Borel-Cantelli Lemma.

In particular, the walk reflected at height  $d$  crosses the root loop infinitely often a.s. so the crossing times  $\tau_k^{(d)}$  defined as in (11) with  $S^{(d)}$  in place of  $S$  are well-defined and finite a.s. for every  $k \geq 1$ . We can similarly define the local time process  $L^{(d)}(\cdot, k)$  as in (12) but with  $S^{(d)}$  in place of  $S$ . In words,  $L^{(d)}(v, k)$  is the number of crossing for the reflected walk of the oriented edge  $v \rightarrow v^0$  before crossing the root loop  $k$  times. Now, the crucial observation is that, since we are on a tree and because the walk starts and ends at the root, then at time  $\tau_k$  every edge (apart from the root loop) has been crossed the same number of times in both directions. Indeed, when the walk crosses an edge  $v \rightarrow v^i$ , then it must necessarily return to  $v$  through the edge  $v^i \rightarrow v$  before returning to the root. Looking at the definition of the TBRW, it is now clear that the dynamics of the walk at any given vertex  $v$  is independent of what happens at others vertices (because the walk returns by the same edge it exited) and follows, for vertices at height  $< d$ , the dynamics of the urn process of Section 2.3. Putting all these arguments together, we conclude that the process  $L^{(d)}(\cdot, k)$  restricted to vertices with height  $< d$  and the branching Markov chain  $Z$  starting from a particle of type  $k$  and considered up to generation  $d - 1$  have the same law.

By trivial coupling, the processes  $(S_n, n \leq \tau_k)$  and  $(S_n^{(d)}, n \leq \tau_k^{(d)})$  coincide on the event  $\{\tau_k < \infty \text{ and } \max_{n \leq \tau_k} S_n < d\} = \{\max_{n \leq \tau_k^{(d)}} S_n^{(d)} < d\}$ . Therefore, the equality in distributions stated above shows that

$$\begin{aligned} \mathbf{P}(\tau_k < \infty \text{ and } \max_{n \leq \tau_k} S_n < d) &= \mathbf{P}(\max_{n \leq \tau_k^{(d)}} S_n^{(d)} < d) \\ &= \mathbf{P}(Z \text{ dies out before generation } d \mid Z(o) = k) \end{aligned}$$

which proves statement (a) of the proposition by letting  $d$  go to infinity. Statement (b) follows from the same argument since, a.s., the walk and its reflected modification coincide on the event  $\{\tau_k < \infty\}$  for  $d$  large enough. Finally, statement (c) is a consequence of the deterministic equality

$$\tau_k = k + 2 \sum_{v \in \mathcal{T}_{\tau_k} \setminus \{o\}} L(v, k)$$

which holds true when  $\tau_k < \infty$  and summarizes the facts that the walk traverse a single edge at each step and that every edge, apart from the root loop, is traversed the same number of times in both direction up to time  $\tau_k$ .  $\square$

We can now establish a 0 – 1 law for the TBRW.

**Corollary 2.9** (0-1 law for the recurrence/transience). *We are in exactly one of the two settings:*

1. *Either  $\mathbf{P}(Z \text{ dies out} \mid Z(o) = k) < 1$  for all  $k \geq 1$ . In that case, the infinite tree  $\mathcal{T}_\infty$  created by the tree-builder walk consists of a single infinite line on which are grafted finite trees. We have a.s. that  $\lim_{\infty} |S_n| = +\infty$  (in particular the root loop is visited only finitely many times): we say that the walk is transient.*
2. *Or  $\mathbf{P}(Z \text{ dies out} \mid Z(o) = k) = 1$  for all  $k \geq 1$ . In that case, the tree  $\mathcal{T}_\infty$  created by the walk is the full infinite tree a.s. (i.e. each vertex has an infinite number of children) and the walk  $(S_n)$  visits every vertex of  $\mathcal{T}_\infty$  (and the root loop, by Corollary 2.4) infinitely many times a.s: we say that the walk is recurrent.*

*Proof.* Let us first assume that  $\mathbf{P}(Z \text{ dies out} \mid Z(o) = k) = 1$  for all  $k \geq 1$ . Then, by Proposition 2.7, we have  $\tau_k < \infty$  a.s. for all  $k \geq 1$  hence the walk returns to the root infinitely often a.s. But, because we know that the TBRW urn ultimately contains infinitely many colors and that each one is drawn infinitely many times a.s. (Lemma 2.5), this implies, by induction on the height, that every vertex of the tree is visited infinitely many times and has infinitely many children which are themselves visited infinitely often a.s.

Conversely, let us assume that  $\mathbf{P}(Z \text{ dies out} \mid Z(o) = k_0) < 1$  for some  $k_0 \geq 1$ . Thanks to the irreducibility property (9) of  $Z$  of Proposition 2.6, this assumption must, in fact, hold true for all  $k \geq 1$ . Furthermore, the extinction probability of  $Z$  is non-increasing in  $k$  in view of (a) of Proposition 2.7 and the fact that  $\tau_{k+1} > \tau_k$ . Therefore, on the one hand, we have for all  $k \geq 1$ :

$$\mathbf{P}(Z \text{ dies out} \mid Z(o) = k) \leq \mathbf{P}(Z \text{ dies out} \mid Z(o) = 1) =: \alpha < 1.$$

On the other hand, recalling again that the TBRW urn creates infinitely many colors and that each one is drawn infinitely many times, we find that, for any fixed  $N$ ,

$$\lim_{k \rightarrow \infty} \mathbf{P}(o \text{ has at least } N \text{ children with non-zero type} \mid Z(o) = k) = 1$$

Therefore, using the branching property of  $Z$ , we have

$$\mathbf{P}(Z \text{ dies out} \mid Z(o) = k) \leq \mathbf{P}(o \text{ has less than } N \text{ children with non-zero type} \mid Z(o) = k) + \alpha^N$$

from which we deduce that (taking  $N$  arbitrary large and then  $k \rightarrow \infty$ )

$$\mathbf{P}(\tau_k < \infty) = \mathbf{P}(Z \text{ dies out} \mid Z(o) = k) \xrightarrow[k \rightarrow \infty]{} 0.$$

This means that the walk returns to the root only finitely many times a.s. By the same argument as before, it follows by induction that every vertex is visited a finite number of times (and has a finite number of children). Finally, since the walk  $(S_n)$  performs only nearest-neighbor moves, we must have  $\lim_n |S_n| = \infty$  a.s. and the tree  $\mathcal{T}_\infty$  must be one-ended.  $\square$

### 3 Proof of theorem 1.1

Thanks to Proposition 2.7, we can relate the transience of the TBRW to the survival of the branching Markov chain  $Z$ . In order to study whether  $Z$  dies out or not, we will make use of the following criterion for characterizing the survival of the multi-type branching process  $Z$  in terms of the spectral radius of its expectation matrix.

**Proposition 3.1.** *Consider the urn process of section 2.3 and recall the definitions of  $N_k$  and  $Y_k^i$  given in (6) and (7). Let  $f : \mathbb{N}^* \rightarrow \mathbb{R}_+$  be a non-negative and non-null function.*

1. *Suppose that there exists  $\lambda > 1$  such that*

$$\sum_{i=1}^{\infty} f(i) \mathbf{E} \left[ \sum_{\ell=1}^{N_i} \mathbf{1}_{\{Y_i^\ell = k\}} \right] \geq \lambda f(k) \quad \text{for all } k \geq 1. \quad (13)$$

*Then, the branching Markov chain  $Z$  is super-critical: it has positive probability to survive indefinitely starting from any non-empty initial configuration.*

2. Suppose that  $\sum f(k) < \infty$  and that there exists  $\lambda < 1$  such that

$$\sum_{i=1}^{\infty} f(i) \mathbf{E} \left[ \sum_{\ell=1}^{N_i} \mathbf{1}_{\{Y_i^\ell = k\}} \right] \leq \lambda f(k) \quad \text{for all } k \geq 1.$$

Then, the branching Markov chain  $Z$  is sub-critical: starting from any finite configuration of particles, it dies out almost surely and the total population over the entire lifetime of the process has finite expectation.

3. Suppose that  $\sum f(k) < \infty$  and

$$\sum_{i=1}^{\infty} f(i) \mathbf{E} \left[ \sum_{\ell=1}^{N_i} \mathbf{1}_{\{Y_i^\ell = k\}} \right] = f(k) \quad \text{for all } k \geq 1.$$

Then, the branching Markov chain  $Z$  is critical: starting from any finite configuration of particles, it dies out almost surely but the total population over the entire lifetime of the process has infinite expectation.

*Proof.* We first prove 1. Let  $M = (M_{i,j})_{i,j \geq 1}$  denote the expectation matrix for the branching Markov chain  $Z$  (ignoring particles of type 0 which do not reproduce):

$$M_{i,j} := \text{expected number of children of type } j \text{ for a particle of type } i = \mathbf{E} \left[ \sum_{\ell=1}^{N_i} \mathbf{1}_{\{Y_i^\ell = j\}} \right]$$

Let us first suppose that  $M_{i_0, j_0} = \infty$  for some  $i_0, j_0 \geq 1$ . Then, by irreducibility of  $Z$  (Proposition 2.6) we have  $M_{i_0, i_0}^2 \geq M_{i_0, j_0} M_{j_0, i_0} = \infty$  so the expected number of particles of type  $i_0$  created by a single particle of the same type after two generations is infinite. By comparison with a classical (single-type) branching process, this shows that  $Z$  is supercritical.

We now assume  $M_{i,j} < \infty$  for all  $i, j$ . Inequality (13) can be rewritten in matrix form  $fM \geq \lambda f$  so it shows that the spectral radius of  $M$  is smaller or equal to  $1/\lambda$ . For  $L \geq 1$ , let  $M|_L$  denote the finite matrix obtained by restricting  $M$  to  $\llbracket 1, L \rrbracket$ . Due to (9), the matrix  $M|_L$  is irreducible (all its coefficients are strictly positive), therefore, according to [15, Theorem 6.8], as  $L$  goes to infinity, the spectral radius of  $M|_L$  decreases to that of  $M$ . Since  $\lambda > 1$ , we can choose  $L_0$  large enough so that the spectral radius of  $M|_{L_0}$  is  $1/\lambda_0 < 1$ . Consider now the modified branching Markov chain  $Z|_{L_0}$  constructed from  $Z$  by instantaneously removing all particles of types outside  $\llbracket 1, L_0 \rrbracket$ . It is a classical, irreducible, multi-type branching process (with a finite number of types) with expectation matrix  $M|_{L_0}$  which has its Perron-Frobenius eigenvalue  $\lambda_0 > 1$ . Standard results about multi-type branching processes (see for instance [8, Theorem II.7.1]) ensure that  $Z|_{L_0}$  is super-critical and therefore  $Z$  also has a positive probability of survival.

We now establish 2. Since  $\sum f(k) < \infty$ , we can assume without loss of generality that  $\sum f(k) = 1$  (because  $f$  is assumed to be non-zero). Starting the branching Markov chain  $Z$  from a single particle with initial random type distributed as  $f$ , we expect at time  $n$ , for each  $k \geq 1$ , no more than  $\lambda^n f(k)$  particles of type  $k$  on average. Therefore, the expected total number of non-zero type particles that will ever exist in the system is  $\sum_n \sum_k f(k) \lambda^n = 1/(1-\lambda) < \infty$ . In particular, its expectation is finite so that the total number of particles is a fortiori an a.s.

finite random variable. Hence the process dies out a.s. By irreducibility, this result remains true when starting from any finite configuration of particles.

Finally, in case 3., the same argument as above shows that starting from a single particle with type distributed as  $f$ , the expected number of non-zero type particles at time  $n$  is always 1. Now, assume that for some  $k \geq 1$  such that  $f(k) > 0$ , there is a positive probability  $p_k$  that  $Z$  started from a single particle of type  $k$  survives indefinitely. On this event, there are a.s. infinitely many generations containing at least one vertex of type 1, by (10) and the Borel-Cantelli Lemma. Since  $Z$  is irreducible (hence a vertex of type 1 can have an arbitrarily large offspring of vertices of type  $k$ ), by the same argument, for every  $\ell \geq 1$  there is a.s. at least one generation containing  $\ell$  vertices of type  $k$ , each having then probability  $p_k$  to survive indefinitely. Therefore, starting from a single vertex of type  $k$  (resp. of type distributed as  $f$ ), there is a probability  $p_k > 0$  (resp.  $f(k)p_k > 0$ ) that the sizes of the successive generations of  $Z$  grow to infinity. But it is standard to check that this contradicts the fact that the expected size of each generation is 1. Hence we must have  $p_k = 0$ , and thus  $p_j = 0$  for all  $j \geq 1$  by irreducibility: the process  $Z$  dies a.s., regardless of the type of the initial particle.  $\square$

Recall the definition of the urn process of section 2.3.

**Lemma 3.2.** *For any  $\rho, \nu$  such that  $\rho > \bar{\nu}$ , we have:*

$$\mathbf{E}[N_k - N_{k-1}] = \bar{\nu} \left( \frac{\rho}{\rho - \bar{\nu}} \right)^k \quad (14)$$

*In particular, we have  $\mathbf{E}[N_k] < \infty$  for all  $k$ .*

*Proof.* We consider the following variant of the urn process defined in Section 2.3. First, we start with  $\rho$  balls of color 0 and a random number  $r$  of balls with colors  $1, \dots, r$ , where  $r$  is distributed as  $\nu$ . Second, whenever we pick a ball of color 0, we do not add any additional balls in the urn. The rest of the dynamics is unchanged. Let  $\bar{N}_k$  be the number of balls in this modified urn when we pick the ball 0 for the  $k$ -th time. We claim that

$$\mathbf{E}[\bar{N}_k] = \bar{\nu} \left( \frac{\rho}{\rho - \bar{\nu}} \right)^k < \infty, \quad (15)$$

which we will show at the end of the proof.

We compare this modified urn process with the original urn process from Section 2.3. Consider the following change of color attribution for the original urn: instead of naming the colors in  $\mathbb{N}$  we name them from  $\{0\} \cup (\mathbb{N}^* \times \mathbb{N}^*)$  such that:

- (a) when we add  $\ell$  new balls after picking the ball of color  $(i, j)$ , we give them the colors  $(i, k), \dots, (i, k + \ell - 1)$ , where  $k$  is the smallest positive integer such that  $(i, k)$  is not yet in the urn;
- (b) when we add  $\ell$  new balls after picking color 0 for the  $i$ -th time, we give them the colors  $(i, 1), \dots, (i, \ell)$ .

Now, we claim that for every  $k \geq 1$ ,

$$\mathbf{E}[N_k] = \sum_{j=1}^k \mathbf{E}[\bar{N}_j] \quad (16)$$

which thus implies (by subtracting the LHS for two consecutive values of  $k$ ) that for every  $k \geq 1$ ,

$$\mathbf{E}[N_k - N_{k-1}] = \mathbf{E}[\bar{N}_k]. \quad (17)$$

Together with (15), this concludes the proof of the lemma.

**Proof of (15).** In order to inject some independence into the modified urn process, instead of drawing the balls one by one, we put an exponential clock of rate 1 for each ball of color  $i \geq 1$  and an exponential clock of rate  $\rho$  for color 0 and pick them when their clock rings. Recall that we start at time 0 with  $\rho$  balls of color 0 and  $r$  ball of respective colors  $1, 2, \dots, r$  with  $r$  random and distributed as  $\nu$ . When color 0 rings, we do nothing. When a color  $i \geq 1$  rings, we add a random number  $r'$  of balls, where  $r'$  is again distributed as  $\nu$  and we give new colors to each of these balls.

We focus on  $C_t$ , the number of colors  $i \geq 1$  at time  $t$ . In a small interval  $[t, t + dt]$ , with probability  $C_t dt$  one of the clocks corresponding to a color  $i \geq 1$  rings and the expected number of new colors then added is  $\bar{\nu}$ . Hence,

$$\frac{d}{dt} \mathbf{E}[C_t] = \bar{\nu} \mathbf{E}[C_t].$$

Since  $\mathbf{E}[C_0] = \bar{\nu}$ , we find  $\mathbf{E}[C_t] = \bar{\nu} e^{\bar{\nu} t}$ .

Next we look at the stopping time  $T_k$  when 0 rings for the  $k$ -th time. By definition,  $T_k$  is a sum of  $k$  independent exponential random variables of parameter  $\rho$ . Hence,  $T_k$  has probability density  $\frac{\rho}{\Gamma(k)} (\rho x)^{k-1} e^{-\rho x}$ ,  $x \geq 0$ . Therefore, by independence of the clocks,

$$\mathbf{E}[\bar{N}_k] = \mathbf{E}[C_{T_k}] = \bar{\nu} \frac{\rho}{\Gamma(k)} \int_0^\infty e^{\bar{\nu} x} (\rho x)^{k-1} e^{-\rho x} dx = \bar{\nu} \left( \frac{\rho}{\rho - \bar{\nu}} \right)^k.$$

**Proof of (16).** For  $k \geq j \geq 1$ , let  $N_{k,j}$  be the number of balls in the original urn process at time  $\Theta_k$  (i.e. when we pick 0 for the  $k$ -th time) whose color is in  $j \times \mathbb{N}^*$ . Remark that it is enough to show the following statement: for all  $k \geq j \geq 1$ ,  $N_{k,j}$  is distributed as  $\bar{N}_{k-j}$ .

For fixed  $k \geq j \geq 1$ , when we pick color 0 for the  $j$ -th time, we then add the first  $r$  balls of color in  $j \times \mathbb{N}$ , where  $r \sim \nu$ . Since the balls with color  $\mathbb{N}^* \setminus \{j\} \times \mathbb{N}^*$  do not affect the probability that the next color picked among  $\{0\} \cup \{j \times \mathbb{N}^*\}$  is 0, we have that  $N_{k,j}$  is distributed as  $\bar{N}_{k-j}$ .  $\square$

**Lemma 3.3.** For any  $(\rho, \nu)$  such that  $\rho > \bar{\nu}$ , any non-negative function  $f$  and any  $k \geq 0$ , we have the identity

$$\mathbf{E} \left[ \sum_{i=1}^{N_k} f(Y_k^i) \right] = \bar{\nu} \sum_{j=1}^k \mathbf{E} [f(\zeta_1 + \dots + \zeta_j)] \left( \frac{\rho}{\rho - \bar{\nu}} \right)^{k+1-j}$$

where  $(\zeta_i)_{i \geq 1}$  is a sequence of i.i.d. random variables with geometric distribution with parameter  $\frac{\rho}{1+\rho}$  starting from 0 i.e.  $\mathbf{P}(\zeta_i = j) = \frac{\rho}{1+\rho} \left( \frac{1}{1+\rho} \right)^j$  for  $j \geq 0$ .

*Proof.* We have

$$\mathbf{E} \left[ \sum_{i=1}^{N_k} f(Y_k^i) \right] = \mathbf{E} \left[ \sum_{i=1}^{\infty} f(Y_k^i) \mathbf{1}_{\{i \leq N_k\}} \right] = \sum_{i=1}^{\infty} \sum_{\ell=1}^k \mathbf{E} [f(Y_k^i) \mathbf{1}_{\{N_{\ell-1} < i \leq N_\ell\}}]$$



Now, for any fixed  $i \geq 1$  and  $1 \leq \ell \leq k$ , conditionally on the event  $\{N_{\ell-1} < i \leq N_\ell\}$ , the random variable  $Y_k^i$  has the same law as  $\zeta_1 + \dots + \zeta_{k-\ell+1}$ . To see why this is true, notice that, in the urn process, on  $\{N_{\ell-1} < i \leq N_\ell\}$ , color  $i$  appears in the urn between the  $\ell - 1$  and the  $\ell$  draw of color 0 and therefore  $Y_k^i$  is the number of times the ball of color  $i$  will be drawn before drawing color 0 another  $k - \ell + 1$  times. Now, irrespectively of other balls being added to the urn, the probability of drawing a ball of color 0 is always  $\rho$  times larger than that of drawing a ball of color  $i$ . Therefore,  $Y_k^i$  follows the law of the number of failures before having  $k - \ell + 1$  successes in a sequence of independent Bernoulli trials with parameter  $\rho/(1 + \rho)$ . This is the negative binomial distribution which has the same law as  $\zeta_1 + \dots + \zeta_{k-\ell+1}$ . Therefore,

$$\begin{aligned}
\mathbf{E} \left[ \sum_{i=1}^{N_k} f(Y_k^i) \right] &= \sum_{i=1}^{\infty} \sum_{\ell=1}^k \mathbf{E} [f(\zeta_1 + \dots + \zeta_{k-\ell+1})] \mathbf{P}(N_{\ell-1} < i \leq N_\ell) \\
&= \sum_{\ell=1}^k \mathbf{E} [f(\zeta_1 + \dots + \zeta_{k-\ell+1})] \sum_{i=1}^{\infty} \mathbf{P}(N_{\ell-1} < i \leq N_\ell) \\
&= \sum_{\ell=1}^k \mathbf{E} [f(\zeta_1 + \dots + \zeta_{k-\ell+1})] \mathbf{E}[N_\ell - N_{\ell-1}] \\
&= \sum_{\ell=1}^k \mathbf{E} [f(\zeta_1 + \dots + \zeta_{k-\ell+1})] \bar{\nu} \left( \frac{\rho}{\rho - \bar{\nu}} \right)^\ell \\
&= \bar{\nu} \sum_{j=1}^k \mathbf{E} [f(\zeta_1 + \dots + \zeta_j)] \left( \frac{\rho}{\rho - \bar{\nu}} \right)^{k-j+1}
\end{aligned}$$

where we used (14) to compute  $\mathbf{E}[N_\ell - N_{\ell-1}]$ . □

**Proposition 3.4.** *For any  $(\rho, \nu)$  such that  $\rho > \bar{\nu}$  and  $0 < s < \frac{\rho - \bar{\nu}}{\rho}$ , we have, for all  $k \geq 1$ ,*

$$\sum_{n=1}^{\infty} s^n \mathbf{E} \left[ \sum_{\ell=1}^{N_n} \mathbf{1}_{\{Y_n^\ell = k\}} \right] = \frac{\bar{\nu} \rho^2 s}{(\rho - \bar{\nu} - \rho s)(\rho + 1 - \rho s)} \left( \frac{1}{\rho + 1 - \rho s} \right)^k$$

*Proof.* We simply compute, using Lemma 3.3 with  $f(x) = \mathbf{1}_{x=k}$  in the first line:

$$\begin{aligned}
\sum_{n=1}^{\infty} s^n \mathbf{E} \left[ \sum_{\ell=1}^{N_n} \mathbf{1}_{\{Y_n^\ell=k\}} \right] &= \bar{\nu} \sum_{n=1}^{\infty} s^n \sum_{j=1}^n \mathbf{E} \left[ \mathbf{1}_{\{\zeta_1+\dots+\zeta_j=k\}} \right] \left( \frac{\rho}{\rho-\bar{\nu}} \right)^{n-j+1} \\
&= \bar{\nu} \sum_{j=1}^{\infty} \mathbf{P}(\zeta_1 + \dots + \zeta_j = k) \left( \frac{\rho}{\rho-\bar{\nu}} \right)^{-j+1} \sum_{n \geq j} \left( \frac{s\rho}{\rho-\bar{\nu}} \right)^n \\
&= \frac{\bar{\nu}\rho}{\rho-\bar{\nu}-\rho s} \sum_{j=1}^{\infty} \mathbf{P}(\zeta_1 + \dots + \zeta_j = k) s^j \\
&= \frac{\bar{\nu}\rho}{\rho-\bar{\nu}-\rho s} \sum_{j=1}^{\infty} \left( \frac{\rho}{\rho+1} \right)^j \left( \frac{1}{\rho+1} \right)^k \binom{k+j-1}{j-1} s^j \\
&= \frac{\bar{\nu}\rho^2 s}{(\rho-\bar{\nu}-\rho s)(\rho+1)^{k+1}} \sum_{i=0}^{\infty} \left( \frac{\rho s}{\rho+1} \right)^i \binom{k+i}{i} \\
&= \frac{\bar{\nu}\rho^2 s}{(\rho-\bar{\nu}-\rho s)(\rho+1)^{k+1}} \left( \frac{1}{1-\frac{\rho s}{\rho+1}} \right)^{k+1} \\
&= \frac{\bar{\nu}\rho^2 s}{(\rho-\bar{\nu}-\rho s)(\rho+1-\rho s)} \left( \frac{1}{\rho+1-\rho s} \right)^k.
\end{aligned}$$

□

**Corollary 3.5.** Let  $(\rho, \nu)$  be such that  $\rho > 1 + \bar{\nu}$ , and define  $f(k) := \frac{1}{\rho^k}$ . We have, for all  $k \geq 1$ ,

$$\sum_{i=1}^{\infty} f(i) \mathbf{E} \left[ \sum_{\ell=1}^{N_i} \mathbf{1}_{\{Y_i^\ell=k\}} \right] = \frac{\bar{\nu}}{\rho-\bar{\nu}-1} f(k).$$

*Proof of Theorem 1.1.* Let us first assume that  $\rho > 1 + \bar{\nu}$ . We observe that

$$\frac{\bar{\nu}}{\rho-\bar{\nu}-1} \begin{cases} < 1 & \text{if } \rho > 2\bar{\nu} + 1, \\ = 1 & \text{if } \rho = 2\bar{\nu} + 1, \\ > 1 & \text{if } \rho < 2\bar{\nu} + 1. \end{cases}$$

Combining Corollary 3.5, Proposition 3.1 and Proposition 2.7, we conclude that the TBRW is recurrent for  $\rho \geq 1 + 2\bar{\nu}$  and transient for  $\rho \in (1 + \bar{\nu}, 1 + 2\bar{\nu})$ . Furthermore, for  $\rho > 1 + 2\bar{\nu}$ , Proposition 3.1 states that the expected total population of the branching Markov chain  $Z$  is finite, which in view of (c) of Proposition 2.7 shows that  $\mathbf{E}[\tau_k] < \infty$  for all  $k$  so the TBRW is positive recurrent. Conversely, when  $\rho = 1 + 2\bar{\nu}$ , the total population of  $Z$  has infinite expectation hence  $\mathbf{E}[\tau_k] = \infty$  for all  $k$  so the TBRW is null recurrent in that case. Finally, when  $\rho \leq 1 + \bar{\nu}$ , then the walk is transient thanks to Proposition 2.1 by comparison with a  $(\rho', \nu)$ -TBRW with  $\rho' \in (1 + \bar{\nu}, 1 + 2\bar{\nu})$ .

The proof of Theorem (1.1) will be complete once we lift assumption (8) that  $\nu$  has finite expectation. Thus, we now assume that  $\bar{\nu} = \infty$  and show that the TBRW is transient: by truncation, we can find another distribution  $\nu'$  such that  $\nu' \prec \nu$  that also satisfies  $\rho < 1 + 2\bar{\nu}' < \infty$  hence we know that the  $(\rho, \nu')$ -TBRW is transient. Combining Proposition 2.1 and Corollary 2.4, we conclude that the  $(\rho, \nu)$ -TBRW is also transient. □

## 4 Asymptotic behavior of the TBRW

We prove Theorem 1.3 in section 4.2 and Theorem 1.4 in section 4.3.

### 4.1 Cut times for the TBRW

The proof of the law of large numbers and the central limit theorem stated in Theorem 1.3 is based on the classical path decomposition of the walk into excursions between cut times (also called regeneration times). The method is rather generic and was popularized to prove ballistic behavior for random walks in random environments (see for instance [13, 16, 17, 20]) and then subsequently applied successfully to other models such as the excited random walk [2, 4] and, more recently, for the original model of the TBRW (without bias) [14]. We quickly recall here the definition of the cut times and how the finiteness of the first/second moment relates to the existence of a LLN/CLT for the TBRW.

Recall that  $(S_n, \mathcal{T}_n)$  denote the  $(\rho, \nu)$ -TBRW and that  $|S_n|$  is the height of the walk at time  $n$  (i.e. its distance from the root). The first cut time  $C_1$  is defined as

$$C_1 := \inf \{n > 0 : S_n \notin \{S_0, \dots, S_{n-1}\} \text{ and } |S_k| \geq |S_n| \text{ for all } k \geq n\}.$$

In words,  $C_1$  (if it exists) is the first time that the walk visits a new vertex  $v$  and then never leaves the subtree rooted at  $v$ . We point out that  $C_1$  is not a stopping time because its definition depends on the future of the walk, but it can be decomposed into a stopping time (the walk visits a new vertex  $v$ ) and an event that does not depend on the past of the walk (the walk never goes back to the father of  $v$  afterward). Moreover, when  $S$  is transient, the latter event has positive probability which does not depend on  $v$  and equals

$$\alpha := \mathbf{P}(\tau_1 = \infty) \in (0, 1). \quad (18)$$

Our interest in cut-times comes from the following standard result.

**Proposition 4.1.** *Suppose that the first cut time  $C_1$  for the TBRW  $S$  is well-defined and has finite expectation  $\mathbf{E}[C_1] < \infty$ . Then, there exists  $v > 0$  such*

$$\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{a.s.} v.$$

Furthermore, if  $\mathbf{E}[C_1^2] < \infty$ , then there exists  $\sigma > 0$  such that

$$\frac{|S_n| - nv}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{law} \mathcal{N}(0, \sigma^2).$$

*Proof.* The proof is classical. Under the assumption that  $C_1$  exists and because the event  $\{\tau_1 = \infty\}$  has positive probability  $\alpha$ , there exists a.s. an infinite number of cut times  $C_1 < C_2 < \dots$ . Furthermore, by definition of the cut times, the sequences  $(C_{i+1} - C_i)_{i \geq 1}$  and  $(|S_{C_{i+1}}| - |S_{C_i}|)_{i \geq 1}$  are i.i.d. and have the same distribution, respectively, as  $C_1$  and  $S_{C_1}$  conditioned on the event  $\{\tau_1 = \infty\}$ . Then, the LLN/CLT for the walk now follows from the LLN/CLT applied to classical i.i.d. sequences of random variables that have a first/second moment together with a change of time  $n \leftrightarrow C_n \sim n \mathbf{E}[C_1]$ . We leave out the details and we refer the reader to [16] and [17] for additional explanations.  $\square$

## 4.2 Proof of Theorem 1.3

In all this section, we assume that the  $(\rho, \nu)$ -TBRW is transient *i.e.*

$$\rho < 1 + 2\bar{\nu}.$$

We also do not require that  $\bar{\nu} < \infty$  anymore so the proof given here will hold in full generality. In view of Proposition 4.1, The theorem will be established once we show that the first cut time  $C_1$  is well-defined and has a finite second moment. We will prove the stronger result:

**Proposition 4.2.** *The first cut time  $C_1$  is finite a.s. and, for all  $n$  large enough, we have*

$$\mathbf{P}(C_1 > n) \leq e^{-n^{1/6+o(1)}}.$$

*In particular,  $C_1$  has moments of all orders.*

The proposition relies on two technical estimates whose proofs are specific to the model of the TBRW (Lemma 4.4 and Lemma 4.3) while the main argument to control the tail of  $C_1$  using a decomposition of the path into “failed excursions” until the first cut time is standard. We first show that the walk must visit new vertices often. Let

$$R_n := \#\{S_i, 0 \leq i \leq n\}$$

denote the range (*i.e.* number of distinct vertices visited) of  $S$  up to time  $n$ . We point out that this quantity is usually strictly smaller than the number of vertices of  $\mathcal{T}_n$ .

**Lemma 4.3.** *There exists  $\delta > 0$  such that for all  $n \geq 1$ :*

$$\mathbf{P}(R_n \leq \delta n^{1/2}) \leq e^{-n^{1/2+o(1)}}.$$

*Proof.* Let  $\delta > 0$ , and let  $v_i$  be the  $i$ -th vertex visited by  $S$ . Note that  $\{R_n \leq \delta n^{1/2}\} \subseteq \{\exists j \leq n, v_j \text{ is visited at least } \delta^{-1}n^{1/2} \text{ times by } (S_k)_{1 \leq k \leq n}\}$ . The number  $k$  of neighbors created during the first  $\lfloor \delta^{-1}n^{1/2}/2 \rfloor$  visits to any given vertex stochastically dominates a binomial  $\text{Bin}(\lfloor \delta^{-1}n^{1/2}/2 \rfloor, 1 - \nu(0))$  variable. Hence, choosing  $\delta < (1 - \nu(0))^2/4$ , by Azuma’s inequality (or standard large deviation bounds) and a union bound on  $j \leq n$ ,

$$\mathbf{P}(\exists j \leq n, v_j \text{ is visited } \lfloor \delta^{-1}n^{1/2}/2 \rfloor \text{ times, creating } \leq \delta^{-1/2}n^{1/2} \text{ neighbors}) \leq e^{-n^{1/2+o(1)}}.$$

But if a vertex has already at least  $\delta^{-1/2}n^{1/2}$  neighbors, for each subsequent visit of  $S$ , the probability to exit via a not yet visited vertex, if the range of the walk so far is  $\leq \delta n^{1/2}$ , is therefore at least  $1 - \delta n^{1/2}/(\delta^{-1/2}n^{1/2}) \geq 1 - \delta^{3/2} \geq 3/4$ . Thus, letting  $Y \sim \text{Bin}(\lfloor \delta^{-1}n^{1/2}/2 \rfloor, 3/4)$ ,

$$\begin{aligned} \mathbf{P}(R_n \leq \delta n^{1/2}) &\leq \mathbf{P}\left(\begin{array}{l} \exists j \leq n, v_j \text{ is visited } \delta^{-1/2}n^{1/2} \text{ times,} \\ S \text{ leaving } v_j \text{ via } \leq \delta n^{1/2} \text{ distinct neighbors} \end{array}\right) \\ &\leq e^{-n^{1/2+o(1)}} + n\mathbf{P}(Y \leq \delta n^{1/2}) \leq e^{-n^{1/2+o(1)}}, \end{aligned}$$

where the last inequality follows again, for  $\delta$  small enough, from large deviation bounds on binomial variables.  $\square$

We also need a control for the range of the walk at time  $\tau_1$  when it first crosses the root loop, conditionally on this return time being finite.

**Lemma 4.4.** *For all  $n$  large enough, we have*

$$\mathbf{P}(R_{\tau_1} > n, \tau_1 < \infty) \leq e^{-n^{1/2+o(1)}}.$$

*Proof.* We bound the range  $R_{\tau_1}$  by bounding separately the depth and the width of the subtree of vertices visited by the walk up to time  $\tau_1$ . We start by looking at the depth of this subtree: fix  $k \geq 1$  and consider the stopping time  $h_k := \inf(n \geq 0 : |S_n| = k)$ . Define the event  $\mathcal{E}_k := \{h_k < \tau_1 < \infty\}$ . On  $\mathcal{E}_k$ , the walk reaches height  $k$  at vertex  $v_k := S_{h_k}$ . Let  $o = v_0, v_1, \dots, v_k$  be the vertices along the ancestor line from the root  $o$  to  $v_k$ . After time  $h_k$  and because  $\tau_1$  is finite on  $\mathcal{E}_k$ , the walk must, at some time, jump from  $v_k$  to  $v_{k-1}$ , then later on it must jump from  $v_{k-1}$  to  $v_{k-2}$  and so on.

Now, for each  $i$ , when the walk goes back to  $v_i$  for the first time, it has probability  $1 - \nu(0)$  to create at least one new leaf. Furthermore, the probability that a given leaf is visited before crossing the edge from  $v_i$  to  $v_{i-1}$  is exactly  $\frac{1}{1+\rho}$ . Finally, when the leaf is visited for the first time, because the walk is transient, it has probability  $\alpha := \mathbf{P}(\tau_1 = \infty) > 0$  to never return to  $v_i$ . Putting all this together, we conclude that the probability that the walk starting from  $v_k$  crosses back the root loop at some point is bounded above by the product of the probability of crossing each edge  $(v_i, v_{i-1})$  successively, and therefore

$$\mathbf{P}(\mathcal{E}_k) \leq \left(1 - \frac{\alpha(1 - \nu(0))}{1 + \rho}\right)^k.$$

We now look at the width of the excursion. Fix  $\ell \geq 1$  and let  $G_\ell := \#\{S_k : k \leq \tau_1 \text{ and } |S_k| = \ell\}$  denote the number of vertices visited by the walk until time  $\tau_1$  which are at distance  $\ell$  from the root. Fix  $m \geq 1$  and consider  $\mathcal{G}_{m,\ell} := \{G_\ell \geq m \text{ and } \tau_1 < \infty\}$ . On this event, the walk visits  $m$  distinct sites at height  $\ell$  but, at each first visit of a site  $v$  at height  $\ell$ , it has probability  $\alpha$  never to go back to  $v^0$  - which is necessary to visit another site of the same height. Therefore, we find that

$$\mathbf{P}(\mathcal{G}_{m,\ell}) \leq \alpha^m$$

Finally, we decompose

$$\begin{aligned} \mathbf{P}(R_{\tau_1} > n, \tau_1 < \infty) &\leq \mathbf{P}(h_k < \tau_1 < \infty) + \mathbf{P}(\exists \ell < k, G_\ell \geq n/k, \tau_1 < \infty) \\ &\leq \mathbf{P}(\mathcal{E}_k) + \sum_{\ell=1}^k \mathbf{P}(\mathcal{G}_{\lfloor n/k \rfloor, \ell}) \\ &\leq \left(1 - \frac{\alpha(1 - \nu(0))}{1 + \rho}\right)^k + k\alpha^{\lfloor n/k \rfloor}, \end{aligned}$$

and the result follows choosing  $k = \lfloor \sqrt{n} \rfloor$  □

*Proof of Proposition 4.2 (hence of Theorem 1.3).* Now that we have obtained the critical estimates on  $R_n$  and  $R_{\tau_1}$ , the proof of the proposition is *mutatis mutandis* the same as that of Proposition 6.1 of [2] with Lemma 6.3 and 6.4 of [2] replaced respectively by Lemma 4.4 and 4.3. We provide below the main arguments for the sake of completeness.

Let  $(\mathcal{F}_n)_{n \geq 0}$  denote the natural filtration for  $S$ . We define by induction two interlaced sequences of stopping times:  $A_0 = B_0 = 0$  and for  $i \geq 1$ ,

$$A_i := \inf\{n > B_{i-1} : S_n \notin \{S_0, S_1, \dots, S_{n-1}\}\} \quad ; \quad B_i := \inf\{n > A_i, |S_n| < |S_{A_i}|\}.$$

In other words  $A_i$  is the first time after  $B_{i-1}$  when the walk visits a new vertex, and  $B_i$  is the first time after  $A_i$  that it goes back to the father of  $S_{A_i}$ . Note that we have  $B_i = \infty$  with positive probability since  $(S_i)_{i \geq 0}$  is a.s. transient.

We know that the walk visits infinitely many sites a.s., so that we have the equality of events  $\{B_{k-1} < \infty\} = \{A_k < \infty\}$  a.s. Furthermore, because the walk evolves on a new subtree during each excursion  $[A_k, B_k)$ , it follows that, conditionally on  $\{A_k < \infty\}$ , the random variable  $B_k - A_k$  is independent of  $\mathcal{F}_{A_k}$  and distributed as  $\tau_1$ . In particular, we have  $\mathbf{P}(B_k = \infty | A_k < \infty, \mathcal{F}_{A_k}) = \mathbf{P}(\tau_1 = \infty)$ . Therefore, by induction, it follows that the r.v.  $\theta := \inf\{i : B_i = \infty\}$  has geometric distribution with parameter  $\alpha := \mathbf{P}(\tau_1 = \infty) > 0$ . This proves that  $C_1 = A_\theta$  is finite a.s.

Now, we consider  $R_{C_1}$  the number of site visited by the walk at time  $C_1$ . Observing that  $S$  only visits new sites during the excursions  $[A_i, B_i)$ , we obtain the identity

$$R_{C_1} = 1 + \sum_{1 \leq i < \theta} (V_i + 1)$$

where  $V_i$  is the number of distinct sites visited during the excursion  $[A_i, B_i)$ . Now, using again that those excursions occur on distinct subtrees, we deduce that, conditionally on  $\theta$ , the sequence  $(V_1, \dots, V_{\theta-1})$  is i.i.d. with  $V_i$  having the same distribution as the r.v.  $R_{\tau_1}$  conditioned on  $\{\tau_1 < \infty\}$ . Therefore, we get by a union bound:

$$\begin{aligned} \mathbf{P}(V_1 + \dots + V_{\theta-1} > n) &\leq \mathbf{P}(\theta > \lfloor n^{1/3} \rfloor) + \mathbf{P}(V_1 + \dots + V_{\theta-1} > n, \theta \leq \lfloor n^{1/3} \rfloor) \\ &\leq \mathbf{P}(\theta > \lfloor n^{1/3} \rfloor) + \lfloor n^{1/3} \rfloor \mathbf{P}\left(R_{\tau_1} > \frac{n}{\lfloor n^{1/3} \rfloor} \mid \tau_1 < \infty\right) \\ &\leq (1 - \alpha)^{\lfloor n^{1/3} \rfloor} + \frac{\lfloor n^{1/3} \rfloor}{1 - \alpha} e^{-\left(\frac{n}{\lfloor n^{1/3} \rfloor}\right)^{1/2+o(1)}} \\ &\leq e^{-n^{1/3+o(1)}} \end{aligned}$$

where we used the Lemma 4.4 and the fact that  $\theta$  has geometric distribution with parameter  $\alpha$  to upper bound the probabilities on the second line. Finally, using Lemma 4.3, we conclude that

$$\begin{aligned} \mathbf{P}(C_1 > n) &\leq \mathbf{P}(C_1 > n, R_n > \delta n^{1/2}) + \mathbf{P}(R_n \leq \delta n^{1/2}) \\ &\leq \mathbf{P}(R_{C_1} > \delta n^{1/2}) + \mathbf{P}(R_n \leq \delta n^{1/2}) \\ &\leq e^{-[\delta n^{1/2} - 2]^{1/2+o(1)}} + e^{-n^{1/2+o(1)}} \\ &\leq e^{-n^{1/6+o(1)}}. \end{aligned}$$

□

### 4.3 Proof of Theorem 1.4

We assume in this last section that the  $(\rho, \nu)$ -TBRW is positive recurrent, *i.e.*

$$1 + 2\bar{\nu} < \rho.$$

We will use the following lemma which shows that the return times of the walk through the root loop increase exponentially fast.

**Lemma 4.5.** *There exists  $\delta = \delta(\rho, \nu) > 0$  such that*

$$\tau_k \geq e^{\delta k} \quad \text{for all } k \text{ large enough a.s.}$$

*Proof.* We crudely lower bound  $\tau_k$  by the number of jumps from the root  $o$  to a child of  $o$  before jumping through the root loop for the  $k$ -th time. Define the filtration  $\mathcal{H}_k := \sigma((S_i, \mathcal{T}_i)_{i \leq \tau_k})$  and let  $d_k := \deg_o(\mathcal{T}_{\tau_k})$  denote the number of children of the root  $o$  at time  $\tau_k$ . We also define  $\Delta_k := \sum_{i=\tau_k}^{\tau_{k+1}-1} \mathbf{1}_{\{S_i=o\}}$  the number of visits to the root vertex between two consecutive passages through the root loop. We observe that, conditionally on  $\mathcal{H}_k$ , the r.v.  $\Delta_k$  stochastically dominates a geometric random variable with parameter  $\rho/(d_k + \rho)$ . Moreover, every time we are at  $o$  between  $\tau_k$  and  $\tau_{k+1}$ , we have a positive probability  $\beta := 1 - \nu(0) > 0$  of creating at least 1 child. Thus, taking  $d_k$  large, we find that

$$\begin{aligned} \mathbf{P}(d_{k+1} - d_k \leq \beta d_k \mid \mathcal{H}_k) &\leq \mathbf{P}(\Delta_k < 2d_k \mid \mathcal{H}_k) + \mathbf{P}(d_{k+1} - d_k \leq \beta d_k, \Delta_k \geq 2d_k \mid \mathcal{H}_k) \\ &\leq \mathbf{P}\left(\text{Geom}\left(\frac{\rho}{\rho + d_k}\right) < 2d_k \mid \mathcal{H}_k\right) + \mathbf{P}\left(\text{Binom}(2d_k, \beta) \leq \beta d_k \mid \mathcal{H}_k\right) \\ &\xrightarrow{d_k \rightarrow \infty} 1 - e^{-2\rho} < 1. \end{aligned}$$

Therefore, for any  $p_0 < e^{-2\rho}$ , by comparison with a sequence of i.i.d. variables with  $\text{Ber}(p_0)$  distribution, the strong law of large numbers implies that

$$\liminf_n \frac{1}{n} \sum_{k \leq n} \mathbf{1}_{\{d_{k+1} - d_k > \beta d_k\}} \geq p_0 \quad \text{a.s.}$$

This in turn implies the existence of  $\delta > 0$  such that a.s.,  $d_k \geq e^{\delta k}$  for all  $k$  large enough. Finally, denoting  $(\xi_n)_{n \geq 0}$  the i.i.d. sequence with distribution  $\nu$  where  $\xi_n$  is the number of children added to the root at its  $n$ -th visit, we observe that

$$\begin{aligned} \mathbf{P}(\tau_k \leq e^{\delta k/3}, d_k \geq e^{\delta k}) &\leq \mathbf{P}(\xi_1 + \dots + \xi_{\lfloor e^{\delta k/3} \rfloor} \geq e^{\delta k}) \\ &\leq \left\lfloor e^{\delta k/3} \right\rfloor \mathbf{P}\left(\xi_i \geq \frac{e^{\delta k}}{\lfloor e^{\delta k/3} \rfloor}\right) \leq \frac{\bar{\nu} \lfloor e^{\delta k/3} \rfloor^2}{e^{\delta k}} = O(e^{-\delta k/3}), \end{aligned}$$

using Markov's inequality. This proves that the events  $\{\tau_k \leq e^{\delta k/3}, d_k \geq e^{\delta k}\}$  occur only finitely many times a.s. Hence, we conclude that we also have  $\tau_k \geq e^{\delta k/3}$  for all  $k$  large enough a.s.  $\square$

*Proof of Theorem 1.4.* We first establish the upper bound. We assume here that we are in the positive recurrent setting  $\rho > 1 + 2\bar{\nu}$ . Fix  $k \geq 1$  and consider  $f(i) := \frac{\rho^k}{\rho^i}$ . Corollary 3.5 states that  $f$  is a left eigenvector associated with the eigenvalue  $\lambda := \frac{\bar{\nu}}{\rho - \bar{\nu} - 1} \in (0, 1)$  for the expectation matrix of the multi-type branching process  $Z$ . This means that if we start  $Z$  from

an initial distribution of particles with, in average,  $f(i)$  particles of type  $i$  for all  $i \geq 1$ , then at generation  $\ell$ , there are, in average, exactly  $\lambda^\ell f(i)$  particles of type  $i$ . But since  $f(k) = 1$ , this means by positivity that starting  $Z$  from a single particle of type  $k$ , the average number of particle of type  $i \geq 1$  at generation  $n$  is at most  $\lambda^n f(i)$ . Therefore, we find that

$$\begin{aligned} \mathbf{P}(|\mathcal{T}_{\tau_k}| > \ell) &\leq \mathbf{P}\left(\text{there exists a particle of non-zero type at generation } \ell \mid Z(o) = k\right) \\ &\leq \mathbf{E}\left[\text{number of non-zero type particles at generation } \ell \mid Z(o) = k\right] \leq \sum_{i \geq 1} f(i) \lambda^\ell \leq \frac{\rho^k \lambda^\ell}{\rho - 1}. \end{aligned}$$

Let us now fix  $c > \frac{\log(\rho)}{\log(1/\lambda)}$ . It follows that  $\mathbf{P}(|\mathcal{T}_{\tau_k}| > ck)$  decreases exponentially fast, hence  $|\mathcal{T}_{\tau_k}| \leq ck$  for all  $k$  large enough a.s. On the other hand, Lemma 4.5 states that  $\tau_k \geq e^{\delta k}$  for all  $k$  large enough a.s. Combining these two facts, we conclude that we can choose  $c_2$  large enough such that  $|\mathcal{T}_n| \leq c_2 \log n$  for all  $n$  large enough a.s. This proves the upper bound.

In order to establish the lower bound, we compare  $(|S_n|, n \geq 0)$  with a biased random walk. We observe that at each step, the probability that the walk increases its height at the next step is always larger than the probability that the walk creates at least one new child (with probability  $1 - \nu(0)$ ) and then move to a child rather than the father at the next step (with probability at least  $1/(1 + \rho)$  since the vertex now has at least one child). Therefore,

$$\mathbf{P}(|S_{n+1}| - |S_n| = 1 \mid |S_0|, |S_1|, \dots, |S_n|) \geq q := \frac{1 - \nu(0)}{1 + \rho}.$$

This show that we can couple  $(|S_n|, n \geq 0)$  with a  $q$ -biased nearest-neighbor random walk  $(X_n, n \geq 0)$  reflected at 0 in such way that  $|S_n| \geq X_n$  for all  $n$ . We conclude using classical estimates on random walks that we can choose  $c_1 > 0$  small enough such that

$$|\mathcal{T}_n| \geq \sup_{i \leq n} |S_i| \geq \sup_{i \leq n} |X_i| \geq c_1 \log n \quad \text{for all } n \text{ large enough a.s.}$$

□

## Acknowledgement

This work is a collaborative project initiated during an open problem session at the conference *Processes on Graphs and Maps days* (also fondly known as *La Bardade au CIRM*), which took place in November 2024 and was funded by ANR 19-CE40-0025 ‘‘ProGraM’’. The authors would like to thank once again Eleanor Archer and Laurent Menard for organizing this great meeting.

We also warmly thank Alessandra Caraceni for stimulating discussions and for making simulations of the process during the early stage of this work.

## References

- [1] Gideon Amir, Itai Benjamini, Ori Gurel-Gurevich, and Gady Kozma. Random walk in changing environment. *Stochastic Processes and their Applications*, 130(12):7463–7482, 2020.



- [2] Anne-Laure Basdevant and Arvind Singh. Recurrence and transience of a multi-excited random walk on a regular tree. *Electronic Journal of Probability*, 14(none):1628 – 1669, 2009.
- [3] Anne-Laure Basdevant and Arvind Singh. Continuous-time vertex reinforced jump processes on Galton–Watson trees. *The Annals of Applied Probability*, 22(4):1728 – 1743, 2012.
- [4] Jean Bérard and Alejandro Ramírez. Central limit theorem for the excited random walk in dimension  $D \geq 2$ . *Electron. Comm. Probab.*, 12:303–314, 2007.
- [5] Janos Engländer, Giulio Iacobelli, Gábor Pete, and Rodrigo Ribeiro. Structural results for the tree builder random walk, 2023.
- [6] János Engländer, Giulio Iacobelli, and Rodrigo Ribeiro. Recurrence, transience and degree distribution for the tree builder random walk, 2023.
- [7] Daniel Figueiredo, Giulio Iacobelli, Roberto Oliveira, Bruce Reed, and Rodrigo Ribeiro. On a random walk that grows its own tree. *Electronic Journal of Probability*, 26(none):1 – 40, 2021.
- [8] Theodore E. Harris. *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin, 1963.
- [9] Giulio Iacobelli, Daniel R. Figueiredo, and Giovanni Neglia. Transient and slim versus recurrent and fat: Random walks and the trees they grow. *Journal of Applied Probability*, 56(3):769–786, 2019.
- [10] Giulio Iacobelli, Rodrigo Ribeiro, Glaucio Valle, and Leonel Zuaznábar. Tree builder random walk: Recurrence, transience and ballisticity. *Bernoulli*, 28, 02 2022.
- [11] H. Kesten, M. V. Kozlov, and F. Spitzer. A limit law for random walk in a random environment. *Compositio Mathematica*, 30(2):145–168, 1975.
- [12] Russell Lyons, Robin Pemantle, and Yuval Peres. Biased random walks on Galton–Watson trees. *Probability Theory and Related Fields*, 106(2):249–264, October 1996.
- [13] Didier Piau. Théorème central limite fonctionnel pour une marche au hasard en environnement aléatoire. *The Annals of Probability*, 26(3):1016 – 1040, 1998.
- [14] Rodrigo Ribeiro. Renewal structure of the tree builder random walk, 2024.
- [15] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer Series in Statistics. Springer New York, 2006.
- [16] Alain-Sol Sznitman. Slowdown estimates and central limit theorem for random walks in random environment. *Journal of the European Mathematical Society*, 002(2):93–143, 2000.
- [17] Alain-Sol Sznitman and Martin Zerner. A Law of Large Numbers for Random Walks in Random Environment. *The Annals of Probability*, 27(4):1851 – 1869, 1999.

- [18] Bálint Tóth. Generalized Ray-Knight theory and limit theorems for self-interacting random walks on  $\mathbb{Z}^1$ . *The Annals of Probability*, 24(3):1324 – 1367, 1996.
- [19] Bálint Tóth and Wendelin Werner. The true self-repelling motion. *Probab. Theory Related Fields*, 111(3):375–452, 1998.
- [20] Ofer Zeitouni. *Lectures on probability theory and statistics*, volume 1837 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2004. Lectures from the 31st Summer School on Probability Theory held in Saint-Flour, July 8–25, 2001.