Electronic Journal of Probability

Vol. 14 (2009), Paper no. 55, pages 1628–1669.

Journal URL
http://www.math.washington.edu/~ejpecp/

Recurrence and transience of a multi-excited random walk on a regular tree

Anne-Laure Basdevant* and Arvind Singh†‡

Abstract

We study a model of multi-excited random walk on a regular tree which generalizes the models of the once excited random walk and the digging random walk introduced by Volkov (2003). We show the existence of a phase transition and provide a criterion for the recurrence/transience property of the walk. In particular, we prove that the asymptotic behaviour of the walk depends on the order of the excitations, which contrasts with the one dimensional setting studied by Zerner (2005). We also consider the limiting speed of the walk in the transient regime and conjecture that it is not a monotonic function of the environment.

Key words: Multi-excited random walk, self-interacting random walk, branching Markov chain.

AMS 2000 Subject Classification: Primary 60F20, 60K35, 60J80.

Submitted to EJP on December 10, 2008, final version accepted June 15, 2009.

*Institut de Mathématiques de Toulouse, France,anne-laure.basdevant@math.univ-toulouse.fr

†Institut für Mathematik, Universität Zürich, Schweiz, arvind.singh@math.unizh.ch

‡Research supported by the Swiss Science Foundation, grant PDAM2-114536/1.

1 Introduction

The model of the excited random walk on $\mathbb{Z}^d$ was introduced by Benjamini and Wilson in [6] and studied in details in, for instance, [1; 7; 11; 12; 21; 22]. Roughly speaking, it describes a walk which receives a push in some specific direction each time it reaches a new vertex of $\mathbb{Z}^d$. Such a walk is recurrent for $d = 1$ and transient with linear speed for $d \geq 2$. In [25; 26], Zerner introduced a generalization of this model called multi-excited random walk (or cookie random walk) where the walk receives a push, not only on its first visit to a site, but also on some subsequent visits. This model has received particular attention in the one-dimensional setting (*c.f.* [2; 4; 5; 13; 17] and the references therein) and is relatively well understood. In particular, a one-dimensional multi-excited random walk can be recurrent or transient depending on the strength of the excitations and may exhibit sub-linear growth in the transient regime.

Concerning multi-excited random walks in higher dimensions, not much is known when one allows the excitations provided to the walk to point in different directions. For instance, as remarked in [13], for $d \geq 2$, when the excitations of a 2-cookies random walk push the walk in opposite directions, then there is, so far, no known criterion for the direction of transience. In this paper, we consider a similar model where the state space of the walk is a regular tree and we allow the excitations to point in opposite directions. Even in this setting simpler than $\mathbb{Z}^d$, the walk exhibits a complicated phase transition concerning its recurrence/transience behaviour.

Let us be a bit more precise about the model. We consider a rooted b-ary tree $\mathbb{T}$. At each vertex of the tree, we initially put a pile of $M \geq 1$ "cookies" with strengths $p_1, \ldots, p_M \in [0, 1)$. Let us also choose some other parameter $q \in (0, 1)$ representing the bias of the walk after excitation. Then, a cookie random walk on $\mathbb{T}$ is a nearest neighbor random walk $X = (X_n)_{n \geq 0}$, starting from the root of the tree and moving according to the following rules:

- If $X_n = x$ and there remain the cookies with strengths $p_j, p_{j+1}, \ldots, p_M$ at this vertex, then X eats the cookie with attached strength p_j and then jumps at time $n + 1$ to the father of x with probability $1 - p_j$ and to each son of x with probability p_j/b.
- If $X_n = x$ and there is no remaining cookie at site x, then X jumps at time $n + 1$ to the father of x with probability $1 - q$ and to each son of x with probability q/b.

In particular, the bias provided to the height process $|X|$ upon consuming a cookie of strength p is $2p - 1$. Therefore, a cookie pushes the walk toward the root when $p < 1/2$ and towards infinity when $p > 1/2$. The main question we address in this paper is to investigate, whether X is recurrent or transient *i.e.* does it return infinitely often to the origin or does it wander to infinity.

For the one dimensional cookie random walk, a remarkably simple criterion for the recurrence of the walk was obtained by Zerner [25] and generalized in [13]. This characterization shows that the behavior of the walks depends only on the sum of the strengths of the cookies, but not on their respective positions in the pile. However, in the tree setting considered here, as in the multi-dimensional setting, the order of the cookies does matter, meaning that inverting the position of two cookies in the pile may affect the asymptotic behaviour of the walk. We give here a criterion for recurrence from which we derive explicit formulas for particular types of cookie environments.

1.1 The model

Let us now give a rigorous definition of the transition probabilities of the walk and set some notations. In the remainder of this paper, $\mathbb{T}$ will always denote a rooted b-ary tree with $b \geq 2$. The root of the tree is denoted by o. Given $x \in \mathbb{T}$, let $\overleftarrow{x}$ stand for the father of x and $\overrightarrow{x}^1, \overrightarrow{x}^2, \ldots, \overrightarrow{x}^b$ stand for the sons of x. We also use the notation $|x|$ to denote the height of a vertex $x \in \mathbb{T}$. For convenience, we also add an additional edge from the root to itself and adopt the convention that the father of the root is the root itself ($\overleftarrow{o} = o$).

We call cookie environment a vector $\mathscr{C} = (p_1, p_2, \ldots, p_M\,; q) \in [0,1)^M \times (0,1)$, where $M \geq 1$ is the number of cookies. We put a semicolon before the last component of the vector to emphasize the particular role played by q. A $\mathscr{C}$ multi-excited (or cookie) random walk is a stochastic process $X = (X_n)_{n\geq 0}$ defined on some probability space $(\Omega, \mathscr{F}, \mathbf{P})$, taking values in $\mathbb{T}$ with transition probabilities given by

$$
\begin{aligned}
&\mathbf{P}\{X_0 = o\} = 1,\\
&\mathbf{P}\{X_{n+1} = \overrightarrow{X}_n^i \mid X_0, \ldots, X_n\} = \begin{cases} \frac{p_j}{b} & \text{if } j \leq M,\\ \frac{q}{b} & \text{if } j > M,\end{cases}\\
&\mathbf{P}\{X_{n+1} = \overleftarrow{X}_n \mid X_0, \ldots, X_n\} = \begin{cases} 1 - p_j & \text{if } j \leq M,\\ 1 - q & \text{if } j > M,\end{cases}
\end{aligned}
$$

where $i \in \{1, \ldots, b\}$ and $j \stackrel{\text{def}}{=} \sharp\{0 \leq k \leq n, X_k = X_n\}$ is the number of previous visits of the walk to its present position.

Remark 1.1.

1. We do not allow $q = 0$ in the definition of a cookie environment. This assumption is made to insure that a $0-1$ law holds for the walk. Yet, the method developed in this paper also enables to treat the case $q = 0$, *c.f.* Remark 8.1.
2. When $p_1 = p_2 = \ldots = p_M = q$, then X is a classical random walk on $\mathbb{T}$ and its height process is a drifted random walk on $\mathbb{Z}$. Therefore, the walk is recurrent for $q \leq \frac{1}{2}$ and transient for $q > \frac{1}{2}$. More generally, an easy coupling argument shows that, when all the p_i's and q are smaller than $\frac{1}{2}$ (resp. larger than $\frac{1}{2}$), the walk is recurrent (resp. transient). The interesting cases occur when at least one of the cookies pushes the walk in a direction opposite to the bias q of the walk after excitation.
3. This model was previously considered by Volkov [24] for the particular cookie environments:
   (a) $(p_1\,; \frac{b}{b+1})$ "once-excited random walk".
   (b) $(0, 0\,; \frac{b}{b+1})$ "two-digging random walk".

   In both cases, Volkov proved that the walk is transient with linear speed and conjectured that, more generally, any cookie random walk which moves, after excitation, like a simple random walk on the tree (*i.e.* $q = b/(b+1)$) is transient. Theorem 1.2 below shows that such is indeed the case.

Theorem 1.2 (Recurrence/Transience criterion).
Let $\mathscr{C} = (p_1, p_2, \ldots, p_M\,; q)$ be a cookie environment and let $P(\mathscr{C})$ denote its associated cookie environment matrix as in Definition 3.1. This matrix has only a finite number of irreducible classes. Let $\lambda(\mathscr{C})$ denote the largest spectral radius of theses irreducible sub-matrices (in the sense of Definition 5.1).

(a) *If* $q < \frac{b}{b+1}$ *and* $\lambda(\mathscr{C}) \leq \frac{1}{b}$, *then the walk in the cookie environment* $\mathscr{C}$ *is recurrent* i.e. *it hits any vertex of* $\mathbb{T}$ *infinitely often with probability* 1. *Furthermore, if* $\lambda(\mathscr{C}) < \frac{1}{b}$, *then the walk is positive recurrent* i.e. *all the return times to the root have finite expectation.*

(b) *If* $q \geq \frac{b}{b+1}$ *or* $\lambda(\mathscr{C}) > \frac{1}{b}$, *then the walk is transient* i.e. $\lim_{n\to\infty} |X_n| = +\infty$.

Moreover, if $\tilde{\mathscr{C}} = (\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_M\,; \tilde{q})$ *denotes another cookie environment such that* $\mathscr{C} \leq \tilde{\mathscr{C}}$ *for the canonical partial order, then the* $\tilde{\mathscr{C}}$ *cookie random walk is transient whenever the* $\mathscr{C}$ *cookie random walk is transient.*

The matrix $P(\mathscr{C})$ of the theorem is explicit. Its coefficients can be expressed as a rational function of the p_i's and q and its irreducible classes are described in Section 4.1. However, we do not know, except in particular cases, a simple formula for the spectral radius $\lambda(\mathscr{C})$.

Let us stress that the condition $\lambda(\mathscr{C}) \leq \frac{1}{b}$ does not, by itself, insure the recurrence of the walk. Indeed, when X a biased random walk on the tree ($p_1 = \ldots = p_M = q$), then $P(\mathscr{C})$ is the transition matrix of a Galton-Watson process with geometric reproduction law with parameter $\frac{q}{q+b(1-q)}$. According to [20], we have

$$\lambda(\mathscr{C}) = \begin{cases} \frac{q}{b(1-q)} & \text{for } q \leq \frac{b}{b+1}, \\ \frac{b(1-q)}{q} & \text{for } q > \frac{b}{b+1}. \end{cases}$$

Therefore, for q sufficiently close to 1, the walk is transient yet $\lambda(\mathscr{C}) < 1/b$.

Let us also remark that the monotonicity property of the walk with respect to the initial cookie environment stated in Theorem 1.2, although being quite natural, is not straightforward since there is no simple way to couple two walks with different cookie environments (in fact, we suspect that such a coupling does not exist in general, see the conjecture concerning the monotonicity of the speed below).

Theorem 1.3 (Speed and CLT when $\mathbf{p_i > 0}$).
Let $\mathscr{C} = (p_1, p_2, \dots, p_M\,; q)$ *be a cookie environment such that* $p_i > 0$ *for all* i. *If the* $\mathscr{C}$*-cookie random walk is transient, then it has a positive speed and a central limit theorem holds: there exist deterministic* $v = v(\mathscr{C}) > 0$ *and* $\sigma = \sigma(\mathscr{C}) > 0$ *such that*

$$\frac{|X_n|}{n} \xrightarrow[n\to\infty]{a.s.} v \quad \text{and} \quad \frac{|X_n| - nv}{\sqrt{n}} \xrightarrow[n\to\infty]{law} \mathscr{N}(0, \sigma^2).$$

The assumption that all cookies have positive strength cannot be removed. When some cookies have zero strength, it is possible to construct a transient walk with sub-linear growth, *c.f.* Proposition 1.9.

A natural question to address is the monotonicity of the speed. It is known that the speed of a one-dimensional cookie random walk is non decreasing with respect to the cookie environment. However, numerical simulations suggest that such is not the case for the model considered here (*c.f.* Figure 1). We believe this behaviour to be somewhat similar to that observed for a biased random walk on a Galton-Watson tree: the slowdown of the walk is due to the creation of "traps" where the walk spends a long time. When $p_2 = 0$, this is easily understood by the following heuristic argument: the walk returns to each visited site at least once (except on the boundary of its trace) and the length of an excursion of the walk away from the set of vertices it has already visited is a

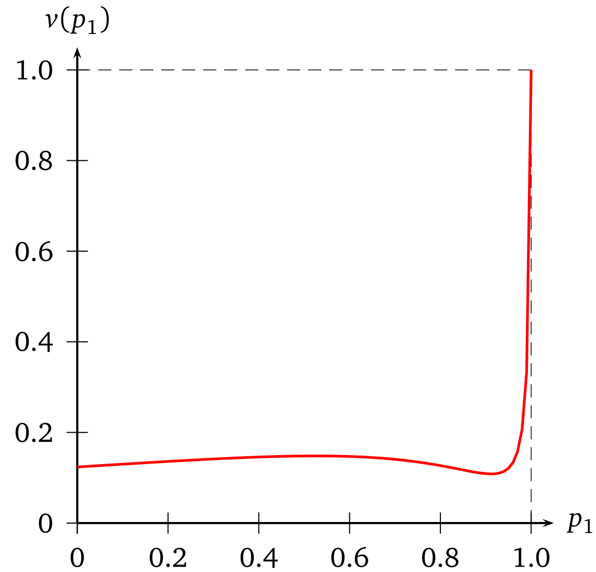

Figure 1: Speed of a $(p_1, 0.01\,; 0.95)$ cookie random walk on a binary tree obtained by Monte Carlo simulation.

geometric random variable with parameter p_1 (the first time the walk moves a step towards the root, it moves back all the way until it reaches a vertex visited at least twice). Therefore, as p_1 increases to 1, the expectation of the length of theses excursions goes to infinity so we can expect the speed of the walk to go to 0. What we find more surprising is that this slowdown also seems to hold true, to some extend, when p_2 is not zero, contrarily to the conjecture that the speed of a biased random walk on a Galton-Watson tree with no leaf is monotonic, *c.f.* Question 2.1 of [15].

1.2 Special cookie environments

The value of the critical parameter $\lambda(\mathscr{C})$ can be explicitly computed in some cases of interest.

Theorem 1.4. *Let $\mathscr{C} = (p_1, \ldots, p_M\,; q)$ denote a cookie environment such that*

$$p_i = 0 \quad \text{for all } i \leq \lfloor M/2 \rfloor \tag{1}$$

where $\lfloor x \rfloor$ denotes the integer part of x. Define

$$\lambda_{sym}(\mathscr{C}) \stackrel{\text{def}}{=} \frac{q}{b(1-q)} \prod_{i=1}^{M} \left((1-p_i)\left(\frac{q}{b(1-q)}\right) + \frac{(b-1)p_i}{b} + \frac{p_i}{b}\left(\frac{q}{b(1-q)}\right)^{-1} \right).$$

For $q < \frac{b}{b+1}$, it holds that

$$\lambda(\mathscr{C}) = \lambda_{sym}(\mathscr{C}).$$

Remark 1.5. For any cookie environment, we have $\lambda(\mathscr{C}) \leq 1$ (it is the maximal spectral radius of sub-stochastic matrices). Moreover, when $\lfloor M/2 \rfloor$ cookies have strength 0, the function $q \mapsto \lambda_{sym}(p_1, \ldots, p_M\,; q)$ is strictly increasing and $\lambda_{sym}(p_1, \ldots, p_M\,; \frac{b}{b+1}) = 1$. Thus, $\lambda(\mathscr{C}) \leq 1 < \lambda_{sym}(\mathscr{C})$ for all $q > \frac{b}{b+1}$.

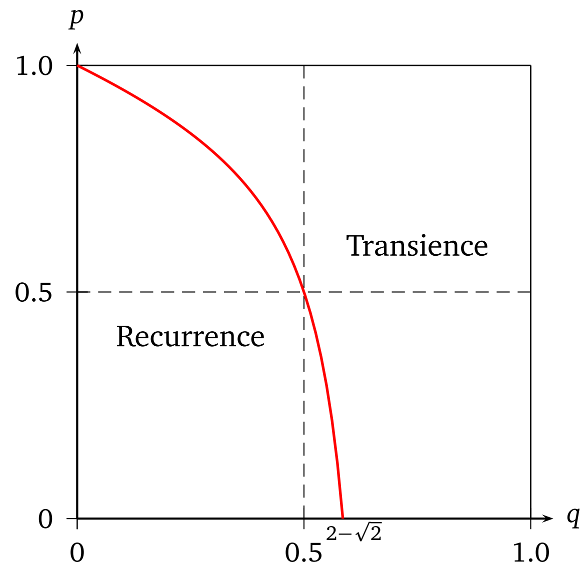

Figure 2: Phase transition of a $(p\,;q)$ cookie random walk on a binary tree.

Let us also note that, under Assumption (1), in order to reach some vertex x, the walk has to visit every vertex on the path $[o, \overleftarrow{x})$ at least M times. Therefore, for such a walk, except on the boundary of its trace, every vertex of the tree is visited either 0 or more than M times. This justifies $\lambda(\mathscr{C})$ being, in this case, a symmetric function of the p_i's.

The combination of Theorem 1.2, Theorem 1.4 and Remark 1.5 directly yields particularly simple criterions for the model of the once excited and the digging random walk.

Corollary 1.6 (Once excited random walk).
Let X denote a $(p\,;q)$ cookie random walk (i.e. $M = 1$) and define

$$\lambda_1 \stackrel{\text{def}}{=} (1-p)\left(\frac{q}{b(1-q)}\right)^2 + \frac{(b-1)p}{b}\left(\frac{q}{b(1-q)}\right) + \frac{p}{b}.$$

Then X is recurrent if and only if $\lambda_1 \leq 1/b$.

In particular, the phase transition of the once excited random walk is non trivial in both cases $p < \frac{1}{2} < q$ and $q < \frac{1}{2} < p$ (*c.f.* Figure 2).

Corollary 1.7 (M-digging random walk).
Let X denote a $\mathscr{C} = (\underbrace{0,\ldots,0}_{M \text{ times}};q)$ cookie random walk and define

$$\lambda_{dig} \stackrel{\text{def}}{=} \left(\frac{q}{b(1-q)}\right)^{M+1}.$$

Then X is recurrent if and only if $\lambda_{dig} \leq 1/b$.

Recall that, according to Theorem 1.2, the condition $q \geq b/(b+1)$ is sufficient to insure the transience of the walk. Corollary 1.7 shows that this condition is also necessary to insure transience independently of $p_1, \dots, p_M$: for any $q < b/(b+1)$, the M digging random walk is recurrent when M is chosen large enough.

We now consider another class of cookie environment to show that, contrarily to the one dimensional case, the order of the cookies in the pile does matter in general.

Proposition 1.8. *Let X be a $\mathscr{C} = (p_1, p_2, \underbrace{0, \dots, 0}_{K \text{ times}}; q)$ cookie random walk with $K \geq 2$. Define $\nu(p_1, p_2)$ to be the largest positive eigenvalue of the matrix*

$$\begin{pmatrix} \frac{p_1}{b} + \frac{p_1 p_2}{b} - \frac{2p_1p_2}{b^2} & \frac{p_1 p_2}{b^2} \\ \frac{p_1+p_2}{b} - \frac{2p_1p_2}{b^2} & \frac{p_1p_2}{b^2} \end{pmatrix},$$

namely

$$\nu(p_1, p_2) = \frac{1}{2b^2}\left((b-1)p_1p_2 + bp_1 + \sqrt{(b^2-6b+1)p_1^2p_2^2 + 2b(b-1)p_1^2p_2 + b^2p_1^2 + 4bp_1p_2^2}\right).$$

Recall the definition of $\lambda_{\text{sym}}(\mathscr{C})$ given in Theorem 1.4 and set

$$\tilde{\lambda} \;=\; \max\left(\lambda_{\text{sym}}(\mathscr{C}), \nu(p_1, p_2)\right).$$

The walk X is recurrent if and only if $\tilde{\lambda} \leq \frac{1}{b}$.

Since ν is not symmetric in (p_1, p_2), Proposition 1.8 confirms that it is possible to construct a recurrent cookie random walk such that the inversion of the first two cookies yields a transient random walk. For $b = 2$, one can choose, for example, $p_1 = \frac{1}{2}$, $p_2 = \frac{4}{5}$ and $q \leq \frac{1}{2}$.
Proposition 1.8 also enables to construct a transient cookie random walk with sub-linear growth.

Proposition 1.9. *Let X be a $\mathscr{C} = (p_1, p_2, 0, 0; q)$ cookie random walk with $q \geq b/(b+1)$ and $\nu(p_1, p_2) = 1/b$. Then X is transient yet*

$$\liminf_{n\to\infty} \frac{|X_n|}{n} = 0.$$

We do not know whether the liminf above is, in fact, a limit.

The remainder of this paper is organized as follows. In the next section, we prove a $0-1$ law for the cookie random walk. In section 3, we introduce a branching Markov chain L (or equivalently a multi-type branching process with infinitely many types) associated with the local time of the walk. We show that the walk is recurrent if and only if this process dies out almost surely. We also prove some monotonicity properties of the process L which imply the monotonicity property of the cookie random walk stated in Theorem 1.2. In section 4, we study the decomposition of the transition matrix P of L and provide some results concerning the evolution of a tagged particle. Section 5 is devoted to completing the proof of Theorem 1.2. In section 6, we prove the law of large numbers and C.L.T. of Theorem 1.3 and Proposition 1.9. In section 7, we compute the value of the critical parameter $\lambda(\mathscr{C})$ for the special cookie environments mentioned above and prove Theorem 1.4 and Proposition 1.8. Finally, in the last section, we discuss some possible extensions of the model.

2 The 0 - 1 law

In the remainder of the paper, X will always denote a $\mathscr{C} = (p_1, \dots, p_M\,; q)$ cookie random walk on a b-ary tree $\mathbb{T}$. We denote by $\mathbb{T}^x$ the sub-tree of $\mathbb{T}$ rooted at x. For $n \in \mathbb{N}$, we also use the notation $\mathbb{T}_n$ (resp. $\mathbb{T}_{\le n}$, $\mathbb{T}_{<n}$) to denote the set of vertices which are at height n (resp. at height $\le n$ and $< n$) from the root . We introduce the sequence $(\tau_o^k)_{k\ge 0}$ of return times to the root.

$$\begin{cases} \tau_o^0 \stackrel{\text{def}}{=} 0, \\ \tau_o^{k+1} \stackrel{\text{def}}{=} \min\{i > \tau_o^k,\, X_i = o\}, \end{cases}$$

with the convention $\min \emptyset = \infty$. The following result shows that, although a cookie random walk is not a Markov process, a $0-1$ law holds (recall that we assume $q \ne 0$ in the definition of a cookie environment).

Lemma 2.1 ($\mathbf{0-1}$ **law).** *Let X be a $\mathscr{C}$ cookie random walk.*

1. *If there exists $k \ge 1$ such that $\mathbf{P}\{\tau_o^k = \infty\} > 0$, then $\lim_{n\to\infty} |X_n| = \infty$ $\mathbf{P}$-a.s.*
2. *Otherwise, the walk visits any vertex infinitely often $\mathbf{P}$-a.s.*

Proof. Let us first assume that $\mathbf{P}\{\tau_o^k < \infty\} = 1$ for all k *i.e.* the walk returns infinitely often to the origin almost surely. Since there are no cookies left after the M^{th} visit of the root, the walk will visit every vertex of height 1 infinitely often with probability 1. By induction, we conclude that the walk visits every vertex of $\mathbb{T}$ infinitely often almost surely.

We now prove the transience part of the proposition. We assume that $\mathbf{P}\{\tau_o^{k_0} < \infty\} < 1$ for some $k_0 \in \mathbb{N}$. Let Ω_1 denote the event

$$\Omega_1 \stackrel{\text{def}}{=} \left\{ \lim_{i\to\infty} |X_i| = \infty \right\}^c.$$

Given $N \in \mathbb{N}$, let $\tilde{X}^N$ denote a multi-excited random walk on $\mathbb{T}$ reflected at height N (*i.e.* a process with the same transition rule as X but which always goes back to its father when it reaches a vertex of height N). This process takes values in the finite state space $\mathbb{T}_{\le N}$ and thus visits any site of $\mathbb{T}_{\le N}$ infinitely often almost surely. For $x \in \mathbb{T}_{<N}$, let $\tilde{\tau}_x^{k_0}$ be the time of the k_0^{th} return of $\tilde{X}^N$ to the vertex x. For $n < N$, let also $\tilde{\tau}_n^{k_0} = \sup_{x\in\mathbb{T}_n} \tilde{\tau}_x^{k_0}$ be the first time when all the vertices of height n have been visited at least k_0 times. We consider the family of events $(A_{n,N})_{n<N}$ defined by:

$$A_{n,N} \stackrel{\text{def}}{=} \{\tilde{X}^N \text{does not reach height } N \text{ before } \tilde{\tau}_n^{k_0}\}.$$

Let us note that, on $A_{n,N}$, the processes X and $\tilde{X}^N$ are equal up to time $\tilde{\tau}_n^{k_0}$. Moreover, given $n \in \mathbb{N}$ and $\omega \in \Omega_1$, we can always find $N > n$ such that $\omega \in A_{n,N}$. Hence,

$$\Omega_1 \subset \bigcap_{n\ge 1} \bigcup_{N>n} A_{n,N}.$$

In particular, for any fixed $n \ge 1$, we get

$$\mathbf{P}\{\Omega_1\} \le \sup_{N>n} \mathbf{P}\{A_{n,N}\}. \tag{2}$$

It remains to bound $\mathbf{P}\{A_{n,N}\}$. For $x \in \mathbb{T}_n$, we consider the subsets of indices:

$$I_x \stackrel{\text{def}}{=} \{0 \le i \le \tilde{\tau}_n^{k_0},\ \tilde{X}_i^N \in \mathbb{T}^x\}.$$
$$I'_x \stackrel{\text{def}}{=} \{0 \le i \le \tilde{\tau}_x^{k_0},\ \tilde{X}_i^N \in \mathbb{T}^x\} \subset I_x.$$

With these notations, we have

$$\begin{aligned}\mathbf{P}\{A_{n,N}\} &= \mathbf{P}\{\forall x \in \mathbb{T}_n,\ (\tilde{X}_i^N, i \in I_x) \text{ does not reach height } N\}\\ &\le \mathbf{P}\{\forall x \in \mathbb{T}_n,\ (\tilde{X}_i^N, i \in I'_x) \text{ does not reach height } N\}.\end{aligned}$$

Since the multi-excited random walk evolves independently in distinct subtrees, up to a translation, the stochastic processes $(\tilde{X}_i^N, i \in I'_x)_{x\in\mathbb{T}_n}$ are i.i.d. and have the law of the multi-excited random walk X starting from the root o, reflected at height $N-n$ and killed at its k_0^{th} return to the root. Thus,

$$\mathbf{P}\{A_{n,N}\} \le \mathbf{P}\Big\{(\tilde{X}_i^{N-n}, i \le \tilde{\tau}_o^{k_0}) \text{ does not reach height } N-n\Big\}^{b^n} \le \mathbf{P}\{\tau_o^{k_0} < \infty\}^{b^n}. \tag{3}$$

Putting (2) and (3) together, we conclude that

$$\mathbf{P}\{\Omega_1\} \le \mathbf{P}\{\tau_o^{k_0} < \infty\}^{b^n}$$

and we complete the proof of the lemma by letting n tend to infinity. □

3 The branching Markov chain L

3.1 Construction of L

In this section, we construct a branching Markov chain which coincides with the local time process of the walk in the recurrent setting and show that the survival of this process characterizes the transience of the walk.

Recall that $\tilde{X}^N$ denotes the cookie random walk X reflected at height N. Fix $k_0 > 0$. Let σ_{k_0} denote the time of the k_0^{th} crossing of the edge joining the root of the tree to itself:

$$\sigma_{k_0} \stackrel{\text{def}}{=} \inf\Big\{i > 0,\ \sum_{j=1}^{i} \mathbb{1}_{\{\tilde{X}_j^N = \tilde{X}_{j-1}^N = o\}} = k_0\Big\}.$$

Since the reflected walk $\tilde{X}^N$ returns to the root infinitely often, we have $\sigma_{k_0} < \infty$ almost surely. Let now $\ell^N(x)$ denote the number of jumps of $\tilde{X}^N$ from $\overleftarrow{x}$ to x before time σ_{k_0} *i.e.*

$$\ell^N(x) \stackrel{\text{def}}{=} \sharp\{0 \le i < \sigma_{k_0},\ \tilde{X}_i^N = \overleftarrow{x} \text{ and } \tilde{X}_{i+1}^N = x\} \qquad \text{for all } x \in \mathbb{T}_{\le N}.$$

We consider the $(N+1)$-step process $L^N = (L_0^N, L_1^N, \dots, L_N^N)$ where

$$L_n^N \stackrel{\text{def}}{=} (\ell^N(x),\ x \in \mathbb{T}_n) \in \mathbb{N}^{\mathbb{T}_n}.$$

Since the quantities L^N, ℓ^N depend on k_0, we should rigourously write L^{N,k_0}, ℓ^{N,k_0}. Similarly, we should write $\sigma_{k_0}^N$ instead of σ_{k_0}. Yet, in the whole paper, for the sake of clarity, as we try to keep the notations as simple as possible, we only add a subscript to emphasize the dependency upon some parameter when we feel that it is really necessary. In particular, the dependency upon the cookie environment $\mathscr{C}$ is usually implicit.

The process L^N is Markovian, in order to compute its transition probabilities we need to introduce some notations which we will extensively use in the rest of the paper.

Definition 3.1.

- *Given a cookie environment $\mathscr{C} = (p_1, \ldots, p_M\,; q)$, we denote by $(\xi_i)_{i \geq 1}$ a sequence of independent random variables taking values in $\{0, 1, \ldots, b\}$, with distribution:*

$$\begin{aligned} \mathbf{P}\{\xi_i = 0\} &= \begin{cases} 1 - p_i & \text{if } i \leq M, \\ 1 - q & \text{if } i > M, \end{cases} \\ \mathbf{P}\{\xi_i = 1\} = \ldots = \mathbf{P}\{\xi_i = b\} &= \begin{cases} \frac{p_i}{b} & \text{if } i \leq M, \\ \frac{q}{b} & \text{if } i > M. \end{cases} \end{aligned}$$

  We say that ξ_i is a "failure" when $\xi_i = 0$.

- *We call "cookie environment matrix" the non-negative matrix $P = (p(i,j))_{i,j \geq 0}$ whose coefficients are given by $p(0, j) = \mathbb{1}_{\{j=0\}}$ and, for $i \geq 1$,*

$$p(i,j) \stackrel{\text{def}}{=} \mathbf{P}\Big\{ \sum_{k=1}^{\gamma_i} \mathbb{1}_{\{\xi_k = 1\}} = j \Big\} \quad \text{where} \quad \gamma_i \stackrel{\text{def}}{=} \inf \Big\{ n, \sum_{k=1}^{n} \mathbb{1}_{\{\xi_k = 0\}} = i \Big\}.$$

  Thus, $p(i,j)$ is the probability that there are exactly j random variables taking value 1 before the i^{th} failure in the sequence $(\xi_1, \xi_2, \ldots)$.

The following lemma characterizes the law of L^N.

Lemma 3.2. *The process $L^N = (L_0^N, L_1^N, \ldots, L_N^N)$ is a Markov process on $\bigcup_{n=0}^{N} \mathbb{N}^{\mathbb{T}_n}$. Its transition probabilities can be described as follows:*

(a) $L_0 = (k_0)$ i.e. $\ell(o) = k_0$.

(b) *For $1 \leq n \leq N$ and $x_1, \ldots, x_k \in \mathbb{T}_n$ with distinct fathers, conditionally on L_{n-1}^N, the random variables $\ell^N(x_1), \ldots, \ell^N(x_k)$ are independent.*

(c) *For $x \in \mathbb{T}_n$ with children $\overrightarrow{x}^1, \ldots, \overrightarrow{x}^b$, the law of $\big(\ell^N(\overrightarrow{x}^1), \ldots, \ell^N(\overrightarrow{x}^b)\big)$, conditionally on L_n^N, depends only on $\ell^N(x)$ and is given by:*

$$\begin{aligned} &\mathbf{P}\Big\{ \ell^N(\overrightarrow{x}^1) = 0, \ldots, \ell^N(\overrightarrow{x}^b) = 0 \,\Big|\, \ell^N(x) = 0 \Big\} = 1 \\ &\mathbf{P}\Big\{ \ell^N(\overrightarrow{x}^1) = j_1, \ldots, \ell^N(\overrightarrow{x}^b) = j_b \,\Big|\, \ell^N(x) = j_0 > 0 \Big\} \\ &\qquad\qquad = \mathbf{P}\Big\{ \forall k \in [0,b],\ \sharp\{1 \leq i \leq j_0 + \ldots + j_b,\ \xi_i = k\} = j_k \text{ and } \xi_{j_0 + \ldots + j_b} = 0 \Big\}. \end{aligned}$$

In particular, conditionally on $\ell^N(x) = j_0$, the random variable $\ell^N(\overrightarrow{x}^k)$ is distributed as the number of ξ_i's taking value k before the j_0^{th} failure. By symmetry, this distribution does not depend on k and, with the notation of Definition 3.1, we have

$$\mathbf{P}\Big\{ \ell^N(\overrightarrow{x}^k) = j \,\Big|\, \ell^N(x) = j_0 \Big\} = p(j_0, j).$$

Proof. (a) is a direct consequence of the definition of σ_{k_0}. Let $x \in \mathbb{T}_{\leq N}$. Since the walk $\tilde{X}^N$ is at the root of the tree at times 0 and σ_{k_0}, the number of jumps $\ell^N(x)$ from $\overleftarrow{x}$ to x is equal to the number of jumps from x to $\overleftarrow{x}$. Moreover, the walk can only enter and leave the subtree $\mathbb{T}^x \cap \mathbb{T}_{\leq N}$ by crossing the edge $(x, \overleftarrow{x})$. Therefore, conditionally on $\ell^N(x)$, the families of random variables $(\ell^N(y),\ y \in \mathbb{T}^x \cap \mathbb{T}_{\leq N})$ and $(\ell^N(y),\ y \in \mathbb{T}_{\leq N} \backslash \mathbb{T}^x)$ are independent. This fact implies (b) and the Markov property of L. Finally, (c) follows readily from the definition of the transition probabilities of a cookie random walk and the construction of the sequence $(\xi_i)_{i \geq 1}$ in terms of the same cookie environment. □

In view of the previous lemma, it is clear that for all $x \in \mathbb{T}_{\leq N}$, the distribution of the random variables $\ell^N(x)$ does not, in fact, depend on N. More precisely, for all $N' > N$, the $(N+1)$ first steps $(L_0^{N'}, \ldots, L_N^{N'})$ of the process $L^{N'}$ have the same distribution as $(L_0^N, \ldots, L_N^N)$. Therefore, we can consider a Markov process L on the state space $\bigcup_{n=0}^{\infty} \mathbb{N}^{\mathbb{T}_n}$:

$$L = (L_n,\ n \geq 0) \quad \text{with} \quad L_n = (\ell(x),\ x \in \mathbb{T}_n) \in \mathbb{N}^{\mathbb{T}_n}$$

where, for each N, the family $(\ell(x),\ x \in \mathbb{T}_{\leq N})$ is distributed as $(\ell^N(x),\ x \in \mathbb{T}_{\leq N})$. We can interpret L as a branching Markov chain (or equivalently a multi-type branching process with infinitely many types) where the particles alive at time n are indexed by the vertices of $\mathbb{T}_n$:

- The process starts at time 0 with one particle o located at $\ell(o) = k_0$.
- At time n, there are b^n particles in the system indexed by $\mathbb{T}_n$. The position (in $\mathbb{N}$) of a particle x is $\ell(x)$.
- At time $n+1$, each particle $x \in \mathbb{T}_n$ evolves independently: it splits into b particles $\overrightarrow{x}^1, \ldots, \overrightarrow{x}^b$. The positions $\ell(\overrightarrow{x}^1), \ldots, \ell(\overrightarrow{x}^b)$ of these new particles, conditionally on $\ell(x)$, are given by the transition kernel described in (c) of the previous lemma.

Remark 3.3.

(1) Changing the value of k_0 only affects the position $\ell(o)$ of the initial particle but does not change the transition probabilities of the Markov process L. Thus, we shall denote by $\mathbf{P}_k$ the probability where the process L starts from one particle located at $\ell(o) = k$. The notation $\mathbf{E}_k$ will be used for the expectation under $\mathbf{P}_k$.

(2) The state 0 is absorbing for the branching Markov chain L: if a particle is at 0, then all its descendants remain at 0 (if the walk never crosses an edge $(\overleftarrow{x}, x)$, then, *a fortiori*, it never crosses any edge of the subtree $\mathbb{T}^x$).

(3) Let us stress that, given $\ell(x)$, the positions of the b children $\ell(\overrightarrow{x}^1), \ldots, \ell(\overrightarrow{x}^b)$ are not independent. However, for two distinct particles, the evolution of their progeny is independent *c.f.* (b) of Lemma 3.2.

(4) When the cookie random walk X is recurrent, the process L coincides with the local time process of the walk and one can directly construct L from X without reflecting the walk at height N and taking the limit. However, when the walk is transient, one cannot directly construct L with $N = \infty$. In this case, the local time process of the walk, stopped at its k_0^{th} jump from the root to itself (possibly ∞), is not a Markov process.

Since 0 is an absorbing state for the Markov process L, we say that L dies out when there exists a time such that all the particles are at 0. The following proposition characterizes the transience of the cookie random walk in terms of the survival of L.

Proposition 3.4. *The cookie random walk is recurrent if and only if, for any choice of k, the process L, under $\mathbf{P}_k$ (i.e. starting from one particle located at $\ell(o) = k$), dies out almost surely.*

Proof. Let us assume that, for any k, the process L starting from k dies out almost surely. Then, k being fixed, we can find N large enough such that L dies out before time N with probability c arbitrarily close to 1. Looking at the definition of L, this means that the walk X crosses at least k times the edge $(o, \overleftarrow{o})$ before reaching level N with probability c. Letting c tend to 1, we conclude that X returns to the root at least k times almost surely. Thus, the walk is recurrent.

Conversely, if, for some k, the process L starting from k has probability $c > 0$ never to die out, then the walk X crosses the edge $(o, \overleftarrow{o})$ less than k times with probability c. This implies that X returns to the root only a finite number of times with strictly positive probability. According to Lemma 2.1, the walk is transient. □

Recall that, in the definition of a cookie environment, we do not allow the strengths of the cookies p_i to be equal to 1. This assumption insures that, for a particle x located at $\ell(x) > M$, the distribution $(\ell(\overrightarrow{x}^1), \ldots, \ell(\overrightarrow{x}^b))$ of the position of its b children has a positive density everywhere on $\mathbb{N}^b$. Indeed, for any $j_1, \ldots, j_n \in \mathbb{N}$, the probability

$$\mathbf{P}\Big\{\ell(\overrightarrow{x}^1) = j_1, \ldots, \ell(\overrightarrow{x}^b) = j_b \mid \ell(x) = i > M\Big\} \tag{4}$$

is larger that the probability of the $i + j_1 + \ldots + j_b$ first terms of the sequence $(\xi_k)_{k\geq 1}$ being

$$\underbrace{0, \ldots, 0}_{i-1 \text{ times}}, \underbrace{1, \ldots, 1}_{j_1 \text{ times}}, \ldots, \underbrace{b, \ldots, b}_{j_b \text{ times}}, 0$$

which is non zero. Therefore, we get the simpler criterion:

Corollary 3.5. *The cookie random walk is recurrent if and only if L under P_{M+1} dies out almost surely.*

3.2 Monotonicity property of L

The particular structure of the transition probabilities of L in terms of successes and failures in the sequence (ξ_k) yields useful monotonicity properties for this process.

Given two branching Markov chains L and $\tilde{L}$, we say that L is stochastically dominated by $\tilde{L}$ if we can construct both processes on the same probability space in such way that

$$\ell(x) \leq \tilde{\ell}(x) \quad \text{for all } x \in \mathbb{T}, \text{ almost surely.}$$

Proposition 3.6 (monotonicity w.r.t. the initial position). *For any $0 \leq i \leq j$, the process L under $\mathbf{P}_i$ is stochastically dominated by L under $\mathbf{P}_j$.*

Proof. Since each particle in L reproduces independently, we just need to prove that $L_1 = (\ell(\overrightarrow{o}^1), \dots, \ell(\overrightarrow{o}^b))$ under $\mathbf{P}_i$ is stochastically dominated by L_1 under $\mathbf{P}_j$ and the result will follows by induction. Recalling that, under $\mathbf{P}_i$ (resp. $\mathbf{P}_j$), $\ell(\overrightarrow{o}^k)$ is given by the number of random variables ξ taking value k before the i^{th} failure (resp. j^{th} failure) in the sequence (ξ_n), we conclude that, when $i \leq j$, we can indeed create such a coupling by using the same sequence (ξ_n) for both processes. □

Proposition 3.7 (monotonicity w.r.t. the cookie environment).
Let $\mathscr{C} = (p_1, \dots, p_M\,; q)$ and $\tilde{\mathscr{C}} = (\tilde{p}_1, \dots, \tilde{p}_M\,; \tilde{q})$ denote two cookies environments such that $\mathscr{C} \leq \tilde{\mathscr{C}}$ for the canonical partial order. Let L (resp. $\tilde{L}$) denote the branching Markov chain associated with the cookie environment $\mathscr{C}$ (resp. $\tilde{\mathscr{C}}$). Then, for any $i \geq 0$, under $\mathbf{P}_i$, the process $\tilde{L}$ stochastically dominates L.

Proof. Keeping in mind Proposition 3.6 and using again an induction argument, we just need to prove the result for the first step of the process *i.e.* prove that we can construct L_1 and $\tilde{L}_1$ such that, under $\mathbf{P}_i$,

$$\ell(\overrightarrow{o}^k) \leq \tilde{\ell}(\overrightarrow{o}^k) \quad \text{for all } k \in \{1, \dots, b\}. \tag{5}$$

Let (ξ_n) denote a sequence of random variables as in Definition 3.1 associated with the cookie environment $\mathscr{C}$. Similarly, let $(\tilde{\xi}_n)$ denote a sequence associated with $\tilde{\mathscr{C}}$. When $\mathscr{C} \leq \tilde{\mathscr{C}}$, we have $\mathbf{P}\{\xi_n = 0\} \geq \mathbf{P}\{\tilde{\xi}_n = 0\}$ and $\mathbf{P}\{\xi_n = k\} \leq \mathbf{P}\{\tilde{\xi}_n = k\}$ for all $k \in \{1, \dots, b\}$. Moreover, the random variables $(\xi_n)_{n\geq 1}$ (resp. $(\tilde{\xi}_n)_{n\geq 1}$) are independent. Thus, we can construct the two sequences (ξ_n) and $(\tilde{\xi}_n)$ on the same probability space in such way that for all $n \geq 1$ and all $k \in \{1, \dots, b\}$,

$$\begin{aligned} \tilde{\xi}_n = 0 \quad &\text{implies} \quad \xi_n = 0, \\ \xi_n = k \quad &\text{implies} \quad \tilde{\xi}_n = k. \end{aligned}$$

Defining now, for each k, the random variable $\ell(\overrightarrow{o}^k)$ (resp. $\tilde{\ell}(\overrightarrow{o}^k)$) to be the number of random variables taking value k in the sequence (ξ_n) (resp. $(\tilde{\xi}_n)$) before the i^{th} failure, it is clear that (5) holds. □

The monotonicity of the recurrence/transience behaviour of the cookie walk with respect to the initial cookie environment stated in Theorem 1.2 now follows directly from the combination of Corollary 3.5 and Proposition 3.7:

Corollary 3.8. *Let $\mathscr{C} = (p_1, p_2, \dots, p_M\,; q)$ and $\tilde{\mathscr{C}} = (\tilde{p}_1, \tilde{p_2}, \dots, \tilde{p}_M\,; \tilde{q})$ denote two cookie environments such that $\mathscr{C} \leq \tilde{\mathscr{C}}$. The $\tilde{\mathscr{C}}$ cookie random walk is transient whenever the $\mathscr{C}$ cookie random walk is transient.*

4 The Matrix P and the process Z

4.1 Irreducible classes of P

The matrix P plays a key role in the study of L. Since we allow the strength of a cookie to be zero, the transition matrix P need not be irreducible (a matrix is said to be irreducible if, for any i, j, there exists n such that $p^{(n)}(i, j) > 0$, where $p^{(n)}(i, j)$ denotes the (i, j) coefficient of P^n).

For $i, j \in \mathbb{N}$, we use the classical notations

- $i \to j$ if $p^{(n)}(i,j) > 0$ for some $n \geq 1$.
- $i \longleftrightarrow j$ if $i \to j$ and $j \to i$.

Lemma 4.1. *For any $i, j \in \mathbb{N}$, we have*

(a) *If $p(i,j) > 0$ then $p(i,k) > 0$ for all $k \leq j$ and $p(k,j) > 0$ for all $k \geq i$.*

(b) *If $i \to j$ then $i \to k$ for all $k \leq j$ and $k \to j$ for all $k \geq i$.*

Proof. Recall the specific form of the coefficients of P: $p(i,j)$ is the probability of having j times 1 in the sequence $(\xi_n)_{n\geq 1}$ before the i^{th} failure. Let us also note that we can always transform a realization of $(\xi_n)_{n\geq 1}$ contributing to $p(i,j)$ into a realization contributing to $p(i,k)$ for $k \leq j$ (resp. for $p(k,j)$ for $k \geq i$) by inserting additional failures in the sequence. Since no cookie has strength 1, for any $n \geq 1$, $\mathbf{P}\{\xi_n = 0\} > 0$. Therefore, adding a finite number of failures still yields, when $p(i,j) > 0$, a positive probability for these new realizations of the sequence (ξ_n). This entails (a).

We have $i \to j$ if and only if there exists a path $i = n_0, n_1, \ldots, n_{m-1}, n_m = j$ such that $p(n_{t-1}, n_t) > 0$. Using (a), we also have, for $k \leq j$, $p(n_{m-1}, k) > 0$ (resp. for $k \geq i$, $p(k, n_1) > 0$). Hence $i, n_1, \ldots, n_{m-1}, k$ (resp. $k, n_1, \ldots, n_{m-1}, j$) is a path from i to k (resp. from k to j). This proves (b). $\square$

Lemma 4.2. *Let $a \leq b$ such that $a \longleftrightarrow b$. The finite sub-matrix $(p(i,j))_{a\leq i,j\leq b}$ is irreducible.*

Proof. Let $i, j \in [a,b]$. In view of (b) of Lemma 4.1, $a \to b$ implies $i \to b$ and $a \to j$. Therefore $i \to b \to a \to j$ so that $i \to j$. Thus, there exists a path in $\mathbb{N}$:

$$i = n_0, n_1, \ldots, n_m = j \tag{6}$$

such that $p(n_{t-1}, n_t) > 0$ for all t. It remains to show that this path may be chosen in $[a,b]$. We separate the two cases $i \leq j$ and $i > j$.

Case i ≤ j. In this case, the path (6) from i to j may be chosen non decreasing (*i.e.* $n_{t-1} \leq n_t$). Indeed, if there exists $0 < t < m$ such that $n_{t-1} > n_t$, then, according to (a) of Lemma 4.1, $p(n_t, n_{t+1}) > 0$ implies that $p(n_{t-1}, n_{t+1}) > 0$. Therefore, n_t can be removed from the path. Concerning the last index, note that, if $n_{m-1} > n_m$, then we can remove n_{m-1} from the path since $p(n_{m-2}, n_m) > 0$.

Case i > j. According to the previous case, there exists a non decreasing path from i to i. This implies $p(i,i) > 0$ and therefore $p(i,j) > 0$ whenever $j < i$. Thus, there exists a path (of length 1) from i to j contained in $[a,b]$. $\square$

We now define

$$I \stackrel{\text{def}}{=} \{i \geq 0,\ p(i,i) > 0\} = \{i \geq 0,\ i \longleftrightarrow i\}.$$

On I, the relation $\longleftrightarrow$ is an equivalence relation. In view of the previous lemma, we see that the equivalence classes for this relation must be intervals of $\mathbb{N}$. Note that $\{0\}$ is always an equivalence class since 0 is absorbent. Moreover, we have already noticed that, for $i, j \geq M+1$, $p(i,j) > 0$ *c.f.* (4). Therefore, there is exactly one infinite class of the form $[a, \infty)$ for some $a \leq M+1$. In particular, there are only a finite number of equivalence classes. We summarize these results in the following definition.

Definition 4.3. *Let $K+1$ be the number of equivalence classes of $\longleftrightarrow$ on I. We denote by $(l_i)_{1\le i\le K}$ and $(r_i)_{1\le i\le K}$ the left (resp. right) endpoints of the equivalence classes:*

- *The equivalence classes of $\longleftrightarrow$ on I are $\{0\}, [l_1, r_1], \dots, [l_{K-1}, r_{K-1}], [l_K, r_K)$.*
- $0 < l_1 \le r_1 < l_2 \le r_2 < \dots \le r_{K-1} < l_K < r_K = \infty$.
- *We have $l_K \le M+1$.*

We denote by $(P_k, 1 \le k \le K)$ the sub-matrices of P defined by $P_k \stackrel{\text{def}}{=} (p(i,j))_{l_k \le i,j \le r_k}$. By construction, the (P_k) are irreducible sub-stochastic matrices and P has the form

$$P = \begin{pmatrix}
\boxed{1} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
* & \begin{smallmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ * & \cdots & 0 \end{smallmatrix} & \ddots & & & & & \vdots \\
\vdots & \ddots & \boxed{P_1} & \ddots & & \mathbf{0} & & \vdots \\
\vdots & & \ddots & \begin{smallmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ * & \cdots & 0 \end{smallmatrix} & \ddots & & & \vdots \\
\vdots & & \ast & \ddots & \boxed{P_2} & \ddots & & \vdots \\
\vdots & & & & \ddots & \begin{smallmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ * & \cdots & 0 \end{smallmatrix} & 0 \\
* & \cdots & \cdots & \cdots & \cdots & \cdots & * & \begin{array}{c} P_K \\ \text{(infinite class)} \end{array}
\end{pmatrix}.$$

Remark 4.4. The sequences $(l_i)_{1\le i\le K}$ and $(r_i)_{1\le i\le K-1}$ can be explicitly expressed in terms of the positions of the zeros in the vector $(p_1, \dots, p_M)$. By construction, we have

$$\begin{aligned}
\{l_i,\ 1 \le i \le K\} &= \{n \ge 1, p(n,n) > 0 \text{ and } p(n-1,n) = 0\} \\
\{r_i,\ 1 \le i \le K-1\} &= \{n \ge 1, p(n,n) > 0 \text{ and } p(n,n+1) = 0\},
\end{aligned}$$

which we may rewrite in terms of the cookie vector:

$$\begin{aligned}
\{l_i,\ 1 \le i \le K\} &= \{n \ge 1, \sharp\{1 \le j \le 2n-1, p_j = 0\} = n-1 \text{ and } p_{2n-1} \ne 0\} \\
\{r_i,\ 1 \le i \le K-1\} &= \{n \ge 1, \sharp\{1 \le j \le 2n-1, p_j = 0\} = n-1 \text{ and } p_{2n} = 0\}.
\end{aligned}$$

For example, if there is no cookie with strength 0, then $K = 1$ and $l_1 = 1$. Conversely, if all the p_i's have strength 0 (the digging random walk case), then $K = 1$ and $l_1 = M+1$.

4.2 The process Z

In order to study the branching Markov chain L introduced in the previous section, it is convenient to keep track of the typical evolution of a particle of L: fix a deterministic sequence $(j_i)_{i\ge 0} \in \{1, \dots, b\}^{\mathbb{N}}$ and set

$$\begin{cases} x_0 \stackrel{\text{def}}{=} o, \\ x_{i+1} \stackrel{\text{def}}{=} \overset{\rightarrow j_i}{x_i} & \text{for } i \ge 0. \end{cases}$$

Define the process $Z = (Z_n)_{n\geq 0}$ by

$$Z_n \stackrel{\text{def}}{=} \ell(x_n).$$

According to (c) of Lemma 3.2, given a particle x located at $\ell(x)$, the positions of its b children have the same law. Therefore, the law of Z does not depend on the choice of the sequence $(j_i)_{i\geq 0}$. Moreover, Lemma 3.2 yields:

Lemma 4.5. *Under $\mathbf{P}_i$, the process Z is a Markov chain starting from i, with transition matrix P given in Definition 3.1.*

Let us note that, if Z_n is in some irreducible class $[l_k, r_k]$, it follows from Lemma 4.1 that $Z_m \leq r_k$ for all $m \geq n$. Thus, Z can only move from an irreducible class $[l_k, r_k]$ to another class $[l_{k'}, r_{k'}]$ where $k' < k$. Recall also that $\{0\}$ is always an irreducible class (it is the unique absorbing state for Z). We introduce the absorption time

$$T_0 \stackrel{\text{def}}{=} \inf\{k \geq 0,\ Z_k = 0\}. \tag{7}$$

Lemma 4.6. *Assume that the cookie environment is such that $q < b/(b+1)$. Let $i_0 \in \mathbb{N}$, we have*

(a) $T_0 < \infty$ $\mathbf{P}_{i_0}$*-a.s.*

(b) *For any $\alpha > 0$,* $\sup_n \mathbf{E}_{i_0}[Z_n^\alpha] < \infty$.

Proof. The proof of the lemma is based on a coupling argument. Recall Definition 3.1 and notice that the sequence $(\xi_k)_{k\geq M+1}$ is i.i.d. Thus, for any stopping time τ such that $\tau \geq M+1$ a.s., the number of random variables in the sub-sequence $(\xi_k)_{k>\tau}$ taking value 1 before the first failure in this sub-sequence has a geometric distribution with parameter

$$s \stackrel{\text{def}}{=} \mathbf{P}\{\xi_{M+1} = 1 \mid \xi_{M+1} \in \{0,1\}\} = \frac{q}{q + b(1-q)}.$$

It follows that, for any i, the number of random variables in the sequence $(\xi_k)_{k\geq 1}$ taking value 1 before the i^{th} failure is stochastically dominated by $M + \mathcal{G}_1 + \ldots + \mathcal{G}_i$ where $(\mathcal{G}_k)_{k\geq 1}$ denotes a sequence of i.i.d. random variables with geometric distribution *i.e.*

$$\mathbf{P}\{\mathcal{G}_k = n\} = (1-s)s^n \quad \text{for } n \geq 0.$$

This exactly means that, conditionally on $Z_n = i$, the distribution of Z_{n+1} is stochastically dominated by $\mathcal{G}_1 + \ldots + \mathcal{G}_i + M$. Let us therefore introduce a new Markov chain $\tilde{Z}$ with transition probabilities

$$\mathbf{P}\{\tilde{Z}_{n+1} = j \mid \tilde{Z}_n = i\} = \mathbf{P}\{\mathcal{G}_1 + \ldots + \mathcal{G}_i + M = j\},$$

It follows from the stochastic domination stated above that we can construct both processes Z and $\tilde{Z}$ on the same space in such way that, under $\mathbf{P}_{i_0}$, almost surely,

$$Z_0 = \tilde{Z}_0 = i_0 \quad \text{and} \quad Z_n \leq \tilde{Z}_n \text{ for all } n \geq 1. \tag{8}$$

The process $\tilde{Z}$ is a branching process with geometric reproduction and with M immigrants at each generation. Setting

$$c \stackrel{\text{def}}{=} \frac{q}{b(1-q)} = \mathbf{E}[\mathcal{G}_1],$$

we get

$$\mathbf{E}[\tilde{Z}_{n+1} \mid \tilde{Z}_n] = c\tilde{Z}_n + M. \tag{9}$$

When $q < b/(b+1)$, we have $c < 1$ so that $\tilde{Z}_n \geq M/(1-c)$ implies $\mathbf{E}[\tilde{Z}_{n+1} \mid \tilde{Z}_n] \leq \tilde{Z}_n$. Therefore, the process $\tilde{Z}$ stopped at its first hitting time of $[0, M/(1-c)]$ is a positive super-martingale which converges almost surely. Since no state in $(M/(1-c), \infty)$ is absorbent for $\tilde{Z}$, we deduce that $\tilde{Z}$ hits the set $[0, M/(1-c)]$ in finite time. Using the Markov property of $\tilde{Z}$, it follows that $\tilde{Z}$ returns below $M/(1-c)$ infinitely often, almost surely. Since $Z \leq \tilde{Z}$, the same result also holds for Z. Furthermore, the process Z has a strictly positive probability of reaching 0 from any $i \leq M/(1-c)$ in one step (because no cookie has strength 1). Thus Z reaches 0 in finite time. This entails (a).

Concerning assertion (b), it suffices to prove the result for the process $\tilde{Z}$ when α is an integer. We prove the result by induction on α. For $\alpha = 1$, equation (9) implies $\mathbf{E}[\tilde{Z}_{n+1}] = c\mathbf{E}_{i_0}[\tilde{Z}_n] + M$ so that

$$\sup_n \mathbf{E}_{i_0}[\tilde{Z}_n] \leq \max(i_0, M/(1-c)).$$

Let us now assume that, for any $\beta \leq \alpha$, $\mathbf{E}_{i_0}[\tilde{Z}_n^\beta]$ is uniformly bounded in n. We have

$$\begin{aligned}\mathbf{E}_{i_0}[Z_{n+1}^{\alpha+1}] &= \mathbf{E}_{i_0}[\mathbf{E}[(\mathcal{G}_1 + \ldots + \mathcal{G}_{Z_n} + M)^{\alpha+1} | Z_n]] \\ &= c^{\alpha+1}\mathbf{E}_{i_0}[Z_n^{\alpha+1}] + \mathbf{E}_{i_0}[Q(Z_n)] \end{aligned} \tag{10}$$

where Q is a polynomial of degree at most α. Therefore the induction hypothesis yields $\sup_n |\mathbf{E}_{i_0}[Q(Z_n)]| < \infty$. In view of (10), we conclude that $\sup_n \mathbf{E}_{i_0}[Z_n^{\alpha+1}] < \infty$. □

The following lemma roughly states that Z does not reach 0 with a "big jump".

Lemma 4.7. *Assume that the cookie environment is such that $q < b/(b+1)$. Recall that $[l_K, \infty)$ denotes the unique infinite irreducible class of Z. We have*

$$\inf_{j \geq l_K} \mathbf{P}_j\{\exists n \geq 0, Z_n = l_K\} > 0.$$

Proof. We introduce the stopping time

$$\sigma \stackrel{\text{def}}{=} \inf\{n > 0,\ Z_n \leq M+1\}.$$

We are going to prove that

$$\inf_{j > M+1} \mathbf{P}_j\{Z_\sigma = M+1\} > 0. \tag{11}$$

This will entail the lemma since $\mathbf{P}_{M+1}\{Z_1 = l_K\} > 0$ (recall that $l_K \leq M+1$). According to (a) of Lemma 4.6, σ is almost surely finite from any starting point j so we can write, for $j > M+1$,

$$\begin{aligned} 1 &= \sum_{k=0}^{M+1} \sum_{i=M+2}^{\infty} \mathbf{P}_j\{Z_{\sigma-1} = i \text{ and } Z_\sigma = k\} \\ &= \sum_{k=0}^{M+1} \sum_{i=M+2}^{\infty} \mathbf{P}_j\{Z_{\sigma-1} = i\} \frac{p(i,k)}{\sum_{m=0}^{M+1} p(i,m)}. \end{aligned} \tag{12}$$

Let us for the time being admit that, for $i > M+1$ and $k \in \{0,\dots,M+1\}$,

$$p(i,k) \le \left(\frac{b}{q}\right)^{M+1} p(i,M+1). \tag{13}$$

Then, combining (12) and (13), we get

$$\begin{aligned} 1 &\le \left(\frac{b}{q}\right)^{M+1} (M+2) \sum_{i=M+2}^{\infty} \mathbf{P}_j\{Z_{\sigma-1} = i\} \frac{p(i,M+1)}{\sum_{m=0}^{M+1} p(i,m)} \\ &= \left(\frac{b}{q}\right)^{M+1} (M+2) \mathbf{P}_j\{Z_\sigma = M+1\}, \end{aligned}$$

which yields (11). It remains to prove (13). Recalling Definition 3.1, we have

$$p(i,k) = \sum_{n=M}^{\infty} \sum_{\substack{e_1,\dots,e_n \text{ s.t.} \\ \sharp\{j\le n, e_j=1\}=k \\ \sharp\{j\le n, e_j=0\}=i-1}} \mathbf{P}\{\xi_1 = e_1,\dots,\xi_n = e_n\}\mathbf{P}\{\xi_{n+1} = 0\}.$$

Keeping in mind that $(\xi_j)_{j\ge M+1}$ are i.i.d. with $\mathbf{P}(\xi_j = 1) = q/b$, we get, for $n \ge M$,

$$\mathbf{P}\{\xi_{n+1} = 0\} = \left(\frac{b}{q}\right)^{M+1-k} \mathbf{P}\{\xi_{n+1} = 1,\dots,\xi_{n+M+1-k} = 1\}\mathbf{P}\{\xi_{n+M+2-k} = 0\}.$$

Thus,

$$\begin{aligned} p(i,k) &\le \left(\frac{b}{q}\right)^{M+1-k} \sum_{\tilde{n}=M}^{\infty} \sum_{\substack{e_1,\dots,e_{\tilde{n}} \text{ s.t.} \\ \sharp\{j\le \tilde{n}, e_j=1\}=M+1 \\ \sharp\{j\le \tilde{n}, e_j=0\}=i-1}} \mathbf{P}\{\xi_1 = e_1,\dots,\xi_{\tilde{n}} = e_{\tilde{n}}\}\mathbf{P}\{\xi_{\tilde{n}+1} = 0\} \\ &\le \left(\frac{b}{q}\right)^{M+1} p(i,M+1). \end{aligned}$$

□

5 Proof of Theorem 1.2

The monotonicity result of Theorem 1.2 was proved in Corollary 3.8. It remains to prove the recurrence/transience criterion. The proof is split into four propositions: Proposition 5.2, 5.4, 5.5 and 5.6.

Definition 5.1. *Given an irreducible non negative matrix Q, its spectral radius is defined as:*

$$\lambda = \lim_{n\to\infty} \left(q^{(n)}(i,j)\right)^{\frac{1}{n}},$$

where $q^{(n)}(i,j)$ denotes the (i,j) coefficient of the matrix Q^n. According to Vere-Jones [23], this quantity is well defined and is independent of i and j.

When Q is a finite matrix, it follows from the classical Perron-Frobenius theory that λ is the largest positive eigenvalue of Q. In particular, there exist left and right λ-eigenvectors with positive coefficients. However, when Q is infinite, the situation is more complicated. In this case, one cannot ensure, without additional assumptions, the existence of left and right eigenvectors associated with the value λ. Yet, we have the following characterization of λ in terms of right sub-invariant vectors (*c.f.* [23], p372):

- λ is the smallest value for which there exists a vector Y with strictly positive coefficients such that $QY \leq \lambda Y$.

By symmetry, we have a similar characterization with left sub-invariant vectors. Let us stress that, contrarily to the finite dimensional case, this characterization does not apply to super-invariant vectors: there may exist a strictly positive vector Y such that $QY \geq \lambda' Y$ for some $\lambda' > \lambda$. For more details, one can refer to [19; 23].

Recall that, according to Definition 4.3, $P_1, \ldots, P_K$ denote the irreducible sub-matrices of P. Let $\lambda_1, \ldots, \lambda_K$ stand for their associated spectral radii. We denote by λ the largest spectral radius of these sub-matrices:

$$\lambda \stackrel{\text{def}}{=} \max(\lambda_1, \ldots, \lambda_K). \tag{14}$$

5.1 Proof of recurrence

Proposition 5.2. *Assume that the cookie environment $\mathscr{C} = (p_1, \ldots, p_M\,; q)$ is such that*

$$q < \frac{b}{b+1} \quad \textit{and} \quad \lambda \leq \frac{1}{b}.$$

Then, the cookie random walk is recurrent.

The proposition is based on the following lemma.

Lemma 5.3. *Let $k \in \{1, \ldots, K\}$ and assume that $\lambda_k \leq 1/b$. Then, for any starting point $\ell(o) = i \in [l_k, r_k]$ and for any $j \in [l_k, r_k]$, we have*

$$\sharp\{x \in \mathbb{T},\ \ell(x) = j\} < \infty \quad \mathbf{P}_i\text{-a.s.}$$

Proof of Proposition 5.2. We assume that $\lambda \leq 1/b$ and $q < b/(b+1)$. For $k < K$, the irreducible class $[l_k, r_k]$ is finite. Thus, Lemma 5.3 insures that, for any $i \in [l_k, r_k]$,

$$\sharp\{x \in \mathbb{T},\ \ell(x) \in [l_k, r_k]\} < \infty \quad \mathbf{P}_i\text{-a.s.} \tag{15}$$

We now show that this result also holds for the infinite class $[l_K, \infty)$ by using a contradiction argument. Let us suppose that, for some starting point $\ell(o) = i$,

$$\mathbf{P}_i\{\sharp\{x \in \mathbb{T},\ \ell(x) \geq l_K\} = \infty\} = c > 0.$$

Then, for any n,

$$\mathbf{P}_i\{\exists x \in \mathbb{T},\ |x| \geq n \text{ and } \ell(x) \geq l_K\} \geq c. \tag{16}$$

According to Lemma 4.7, given a particle x located at $\ell(x) = j \geq l_K$, the probability that one of its descendants reaches level l_K is bounded away from 0 uniformly in j. In view of (16), we deduce that, for some constant $c' > 0$, uniformly in n,

$$\mathbf{P}_i\{\exists x \in \mathbb{T},\ |x| \geq n \text{ and } \ell(x) = l_K\} \geq c'.$$

This contradicts Lemma 5.3 stating that

$$\sharp\{x \in \mathbb{T},\ \ell(x) = l_K\} < \infty \quad \mathbf{P}_i\text{-a.s.}$$

Thus (15) holds also for the infinite class.

We can now complete the proof of the proposition. According to Corollary 3.5, we just need to prove that the branching Markov chain L starting from $\ell(o) = M+1$ dies out almost surely. In view of (15), the stopping time $N = \inf\{n,\ \forall x \in \mathbb{T}_n\ \ell(x) < l_K\}$ where all the particle are located strictly below l_K is finite almost surely. Moreover, if a particle x is located at $\ell(x) = i \in (r_{K-1}, l_K)$ (*i.e.* its position does not belong to an irreducible class), then, the positions of all its children $\ell(\overrightarrow{x}^1), \ldots, \ell(\overrightarrow{x}^b)$ are strictly below i. Thus, at time $N' = N + (l_K - r_{K-1} - 1)$, all the particles in the system are located in $[0, r_{K-1}]$. We can now repeat the same procedure with the irreducible class $[l_{K-1}, r_{K-1}]$. Since there are only a finite number of irreducible classes, we conclude, by induction, that all the particles of L are at zero in finite time with probability 1. □

Proof of Lemma 5.3. Fix $k \leq K$ and $j_0 \in [l_k, r_k]$. By irreducibility, if suffices to prove that

$$\sharp\{x \in \mathbb{T},\ \ell(x) = j_0\} < \infty \quad \mathbf{P}_{j_0}\text{-a.s.} \tag{17}$$

Let us note that, when $k \neq K$, the class $[l_k, r_k]$ is finite. Thus, the process L restricted to $[l_k, r_k]$ (*i.e.* the process where all the particles leaving this class vanish) is a multi-type branching process with only a finite number of types. Using Theorem 7.1, Chapter II of [10], it follows that this process is subcritical (it has parameter $\rho = \lambda_k b \leq 1$ with the notation of [10] and is clearly positive regular and non-singular) and thus it dies out almost surely, which implies (17). However, this argument does not apply when $k = K$. We now provide an argument working for any k.

As already mentioned, Criterion I of Corollary 4.1 of [23] states that λ_k is the smallest value for which there exists a vector $Y_k = (y_{l_k}, y_{l_k+1}, \ldots)$, with strictly positive coefficients such that

$$P_k Y_k \leq \lambda_k Y_k.$$

For $k \neq K$, the inequality above is, in fact, an equality. Since $\lambda_k \leq 1/b$, we get

$$P_k Y_k \leq \frac{1}{b} Y_k. \tag{18}$$

Define the function $f : \mathbb{N} \mapsto \mathbb{N}$ by

$$f(i) \stackrel{\text{def}}{=} \begin{cases} y_i & \text{for } l_k \leq i \leq r_k \\ 0 & \text{otherwise.} \end{cases}$$

Recall the definition of the Markov chain Z, with transition matrix P, introduced in the previous section. It follows from (18) that, for any $i \in [0, r_k]$,

$$\mathbf{E}[f(Z_1)\,|Z_0 = i] \leq \frac{1}{b} f(i). \tag{19}$$

We now consider a process $\tilde{L} = (\tilde{L}_n, n \geq 0)$ obtained by a slight modification of the process L:

- $\tilde{L}_0 = L_0$ *i.e.* $\tilde{\ell}(o) = \ell(o) = j_0$.
- $\tilde{L}_1 = L_1$.
- For $n \geq 1$, $\tilde{L}_n$ is a branching Markov chain with the same transition probabilities as L except at point j_0 which becomes an absorbing state without branching *i.e* when a particle x is located at $\tilde{\ell}(x) = j_0$, then $\tilde{\ell}(\overrightarrow{x}^1) = j_0$ and $\tilde{\ell}(\overrightarrow{x}^2) = \ldots = \tilde{\ell}(\overrightarrow{x}^b) = 0$.

Following [16], we consider the process

$$\tilde{\mathscr{M}}_n = \sum_{x \in \mathbb{T}_n} f(\tilde{\ell}(x))$$

together with the filtration $\mathscr{F}_n = \sigma(\tilde{\ell}(x),\ x \in \mathbb{T}_{\leq n})$. Using (19), we have

$$\begin{aligned}
\mathbf{E}_{j_0}[\tilde{\mathscr{M}}_{n+1}|\mathscr{F}_n] &= \sum_{x \in \mathbb{T}_n,\, \tilde{\ell}(x) \neq j_0} \mathbf{E}[f(\tilde{\ell}(\overrightarrow{x}^1)) + \ldots + f(\tilde{\ell}(\overrightarrow{x}^b)) \mid \tilde{\ell}(x)] + \sum_{x \in \mathbb{T}_n,\, \tilde{\ell}(x) = j_0} f(\tilde{\ell}(x)) \\
&= b \sum_{x \in \mathbb{T}_n,\, \tilde{\ell}(x) = k \neq j_0} \mathbf{E}[f(Z_1)|Z_0 = k] + \sum_{x \in \mathbb{T}_n,\, \tilde{\ell}(x) = j_0} f(\tilde{\ell}(x)) \\
&\leq \sum_{x \in \mathbb{T}_n,\, \tilde{\ell}(x) \neq j_0} f(\tilde{\ell}(x)) + \sum_{x \in \mathbb{T}_n,\, \tilde{\ell}(x) = j_0} f(\tilde{\ell}(x)) \\
&= \tilde{\mathscr{M}}_n.
\end{aligned}$$

Thus, $\tilde{\mathscr{M}}_n$ is a non-negative super-martingale which converges almost surely towards some random variable $\tilde{\mathscr{M}}_\infty$ with

$$\mathbf{E}_{j_0}[\tilde{\mathscr{M}}_\infty] \leq \mathbf{E}_{j_0}[\tilde{\mathscr{M}}_0] = f(j_0).$$

Let $\tilde{N}(n)$ denote the number of particles of $\tilde{L}$ located at site j_0 at time n. Since j_0 is an absorbing state for the branching Markov chain $\tilde{L}$, the sequence $\tilde{N}(n)$ is non-decreasing and thus converges almost surely to some random variable $\tilde{N}(\infty)$. Moreover, we have $\tilde{N}(n)f(j_0) \leq \tilde{\mathscr{M}}_n$ so that $\tilde{N}(\infty)f(j_0) \leq \tilde{\mathscr{M}}_\infty$. This shows that $\tilde{N}_\infty$ is almost surely finite and

$$\mathbf{E}_{j_0}[\tilde{N}(\infty)] \leq 1.$$

We can now complete the proof of the lemma. The random variable $\tilde{N}(\infty)$ represents the total number of particles reaching level j_0 for the branching Markov chain $\tilde{L}$ (where the particles returning at j_0 are frozen). Thus, the total number of particles reaching j_0 for the original branching Markov chain L, starting from one particle located at $\ell(o) = j_0$, has the same law as the total progeny of a Galton-Watson process $W = (W_n)_{n \geq 0}$ with $W_0 = 1$ and with reproduction law $\tilde{N}(\infty)$ (this corresponds to running the process $\tilde{L}$, then unfreezing all the particles at j_0 and then repeating this procedure). Thus, we get the following equality in law for the total number of particles located at j_0 for the original process L starting from one particle located at j_0:

$$\sharp\{x \in \mathbb{T}, \ell(x) = j_0\} \stackrel{\text{law}}{=} \sum_{n=0}^{\infty} W_n.$$

Since $\mathbf{E}_{j_0}[\tilde{N}(\infty)] \leq 1$ and $\mathbf{P}_{j_0}\{\tilde{N}(\infty) = 1\} < 1$, the Galton-Watson process W dies out almost surely. This proves

$$\sharp\{x \in \mathbb{T}, \ell(x) = j_0\} < \infty \quad \mathbf{P}_{j_0}\text{-a.s.}$$

□

5.2 Proof of positive recurrence

Proposition 5.4. *Assume that the cookie environment $\mathscr{C} = (p_1, \dots, p_M\,; q)$ is such that*

$$q < \frac{b}{b+1} \quad \text{and} \quad \lambda < \frac{1}{b}.$$

Then, all the return times of the walk to the root of the tree have finite expectation.

Proof. Let σ_i denote the time of the i^{th} crossing of the edge joining the root of the tree to itself for the cookie random walk:

$$\sigma_i \stackrel{\text{def}}{=} \inf\Big\{n > 0,\ \sum_{j=1}^{n} \mathbb{1}_{\{X_j = X_{j-1} = o\}} = i\Big\}.$$

We prove that $\mathbf{E}[\sigma_i] < \infty$ for all i. Recalling the construction of the branching Markov chain L in section 3.1 and the definition of Z, we have

$$\mathbf{E}[\sigma_i] = i + 2\mathbf{E}_i\Big[\sum_{x \in \mathbb{T}\setminus\{o\}} \ell(x)\Big] = i + 2\sum_{n=1}^{\infty} b^n \mathbf{E}_i[Z_n].$$

Let us for the time being admit that

$$\limsup_{n\to\infty} \mathbf{P}_i\{Z_n > 0\}^{1/n} \le \lambda \quad \text{ for any } i. \tag{20}$$

Then, using Hölder's inequality and (b) of Lemma 4.6, choosing $\alpha, \beta, \tilde{\lambda}$ such that $\tilde{\lambda} > \lambda$, $b\tilde{\lambda}^{1/\alpha} < 1$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, we get

$$\sum_{n=1}^{\infty} b^n \mathbf{E}_i[Z_n] \le \sum_{n=1}^{\infty} b^n \mathbf{P}_i\{Z_n > 0\}^{1/\alpha} \mathbf{E}_i[Z_n^{\beta}]^{1/\beta} \le C_\beta \sum_{n=1}^{\infty} (b\tilde{\lambda}^{1/\alpha})^n < \infty.$$

It remains to prove (20). Recall that $\{0\}, [l_1, r_1], \dots, [l_k, \infty)$ denote the irreducible classes of P and that Z can only move from a class $[l_k, r_k]$ to another class $[l_{k'}, r_{k'}]$ with $k' < k$. Thus, for $i \in [l_k, r_k]$, we have

$$\mathbf{P}_i\{Z_n \ge l_k\} = \mathbf{P}_i\{Z_n \in [l_k, r_k]\} = \sum_{j=l_k}^{r_k} \mathbf{P}_i\{Z_n = j\} = \sum_{j=l_k}^{r_k} p^{(n)}(i,j).$$

For $k < K$, the sum above is taken over a finite set. Recalling the definition of λ_k, we get

$$\lim_{n\to\infty} \mathbf{P}_i\{Z_n \ge l_k\}^{1/n} = \lambda_k \quad \text{for all } i \in [l_k, r_k].$$

Using the Markov property of Z, we conclude by induction that, for any $i < l_K$,

$$\limsup_{n\to\infty} \mathbf{P}_i\{Z_n > 0\}^{1/n} \le \max(\lambda_1, \dots, \lambda_{K-1}) \le \lambda. \tag{21}$$

It remains to prove the result for $i \ge l_K$. In view of (21) and using the Markov property of Z, it is sufficient to show that, for $i \ge l_K$,

$$\limsup_{n\to\infty} \mathbf{P}_i\{Z_n \ge l_K\}^{1/n} \le \lambda_K. \tag{22}$$

Let us fix $i \geq l_K$. We write

$$\mathbf{P}_i\{Z_n \geq l_K\} = \mathbf{P}_i\{\exists m \geq n,\ Z_m = l_K\} + \sum_{j=l_K}^{\infty} \mathbf{P}_i\{Z_n = j\}\mathbf{P}_j\{\nexists m \geq 0, Z_m = l_K\}.$$

According to lemma 4.7, there exists $c > 0$ such that, for all $j \geq l_K$, $\mathbf{P}_j\{\nexists m \geq 0,\ Z_m = l_K\} \leq 1 - c$. Therefore, we deduce that

$$\mathbf{P}_i\{Z_n \geq l_K\} \leq \frac{1}{c}\mathbf{P}_i\{\exists m \geq n,\ Z_m = l_K\} \leq \frac{1}{c}\sum_{m=n}^{\infty} p^{(m)}(i, l_K). \tag{23}$$

Moreover, we have $\lim_{m\to\infty}(p^{(m)}(i, l_K))^{1/m} = \lambda_K < 1$ hence

$$\lim_{n\to\infty}\left(\sum_{m=n}^{\infty} p^{(m)}(i, l_K)\right)^{1/n} = \lambda_K. \tag{24}$$

The combination of (23) and (24) yields (22) which completes the proof of the proposition. □

5.3 Proof of transience when $\lambda > 1/b$

Proposition 5.5. *Assume that the cookie environment $\mathscr{C} = (p_1, \ldots, p_M\,; q)$ is such that*

$$\lambda > \frac{1}{b}.$$

Then, the cookie random walk is transient.

Proof. The proof uses the idea of "seed" as explained in [18]: we can find a restriction $\tilde{L}$ of L to a finite interval $[l, r]$ which already has a non zero probability of survival.

To this end, let us first note that we can always find a finite irreducible sub-matrix $Q = (p(i, j))_{l\leq i,j\leq r}$ of P with spectral radius $\tilde{\lambda}$ strictly larger than $1/b$. Indeed, by definition of λ, either

- There exists $k \leq K - 1$ such that $\lambda_k > 1/b$ in which case we set $l \stackrel{\text{def}}{=} l_k$ and $r \stackrel{\text{def}}{=} r_k$.
- Otherwise $\lambda_K > 1/b$. In this case, we choose $l = l_K$ and $r > l$. Lemma 4.2 insures that the sub-matrix $Q \stackrel{\text{def}}{=} (p(i, j))_{l\leq i,j\leq r}$ is irreducible. Moreover, as r goes to infinity, the spectral radius of Q tends to λ_K (*c.f.* Theorem 6.8 of [19]). Thus, we can choose r large enough such that the spectral radius $\tilde{\lambda}$ of Q is strictly larger than $1/b$.

We now consider the process $\tilde{L}$ obtained from L by removing all the particles x whose position $\ell(x)$ is not in $[l, r]$ (we also remove from the process all the descendants of such a particle). The process $\tilde{L}$ obtained in this way is a multi-type branching process with only finite number of types indexed by $[l, r]$. It follows from the irreducibility of Q that, with the terminology of [10], this process is positive regular. It is also clearly non singular. Moreover, the matrix $\mathbf{M}$ defined in Definition 4.1, Chapter II of [10], is, in our setting, equal to bQ so that the critical parameter ρ of Theorem 7.1, Chapter II of [10] is given by $\rho = b\tilde{\lambda} > 1$. Thus, Theorem 7.1 states that there exists $i \in [l, r]$ such that the process $\tilde{L}$ starting from one particle located at position (*i.e.* with type) i has a non zero probability of survival. *A fortiori*, this implies that L also has a positive probability of survival. Thus the cookie random walk in transient. □

5.4 Proof of transience when $\mathbf{q \geq b/(b+1)}$

Proposition 5.6. *Assume that the cookie environment $\mathscr{C} = (p_1, \ldots, p_M\,; q)$ is such that*

$$q \geq \frac{b}{b+1}.$$

Then, the cookie random walk is transient.

Remark 5.7. Under the stronger assumption $q > b/(b+1)$, one can prove, using a similar coupling argument as in the proof of Lemma 4.6, that the absorbtion time T_0 of Z defined in (7) is infinite with strictly positive probability. This fact implies the transience of the cookie random walk. However, when $q = b/(b+1)$, the absorbtion time T_0 may, or may not, depending on the cookie environment, be finite almost surely. Yet, Proposition 5.6 states that the walk is still transient in both cases.

Proof of Proposition 5.6. In view of the monotonicity property of the walk w.r.t. the cookie environment stated in Corollary 3.8, we just need to prove that, for any M, we can find $\tilde{q} < b/(b+1)$ such that the walk in the cookie environment

$$\tilde{\mathscr{C}} = (\underbrace{0, \ldots, 0}_{M \text{ times}}; \tilde{q}) \tag{25}$$

is transient. It easily checked that the irreducible classes of the matrix $\tilde{P}$ associated to a cookie environment of the form (25) are $\{0\}$ and $[M+1, \infty[$ (see, for instance, Remark 4.4). Moreover, for such a cookie environment, the coefficients of $\tilde{P}$ have a particularly simple form. Indeed, recalling Definition 3.1, a few line of elementary calculus yields, for $i, j \geq M+1$,

$$\tilde{p}(i,j) = \binom{j+i-M-1}{j} s^j (1-s)^{i-M} \quad \text{where } s \stackrel{\text{def}}{=} \frac{\tilde{q}}{\tilde{q} + (1-\tilde{q})b} \tag{26}$$

(see Lemma 7.3 for a proof of (26) in a more general setting). Therefore, the polynomial vector $U \stackrel{\text{def}}{=} (i(i-1)\ldots(i-M))_{i \geq M+1}$ is a right eigenvector of the irreducible sub-matrix $\tilde{P}_1 \stackrel{\text{def}}{=} (\tilde{p}(i,j))_{i,j \geq M+1}$ associated with the eigenvalue

$$\tilde{\lambda} \stackrel{\text{def}}{=} \left(\frac{s}{1-s}\right)^{M+1}.$$

i.e. $\tilde{P}_1 U = \tilde{\lambda} U$. Similarly, setting $V \stackrel{\text{def}}{=} \left((s/(1-s))^{i-1}\right)_{i \geq M+1}$, it also follows from (26) that V is a left eigenvector of $\tilde{P}_1$ associated with the same eigenvalue $\tilde{\lambda}$ *i.e.* ${}^t V \tilde{P}_1 = \tilde{\lambda}\, {}^t V$. Moreover, the inner product ${}^t V U$ is finite. Thus, according to Criterion III p375 of [23], the spectral radius of $\tilde{P}_1$ is equal to $\tilde{\lambda}$. Since $\tilde{\lambda}$ tends to 1 as $\tilde{q}$ increases to $b/(b+1)$, we can find $\tilde{q} < b/(b+1)$ such that $\tilde{\lambda} > 1/b$. Proposition 5.5 insures that, for this choice of $\tilde{q}$, the cookie random walk is transient. □

6 Rate of growth of the walk.

6.1 Law of large numbers and central limit theorem

We now prove Theorem 1.3. Thus, in rest of this section, we assume that X is a transient cookie random walk in an environment $\mathscr{C} = (p_1, \ldots, p_M\,; q)$ such that

$$p_i > 0 \quad \text{for all } i \in \{1, \ldots, M\}. \tag{27}$$

The proof is based on the classical decomposition of the walk using the regeneration structure provided by the existence of cut times for the walk. Recall that $\mathbb{T}^x$ denotes the sub-tree of $\mathbb{T}$ rooted at site x. We say that (random) time $C > 0$ is a cut time for the cookie random walk X if it is such that:

$$\begin{cases} X_i \notin \mathbb{T}^{X_C} & \text{for all } i < C, \\ X_i \in \mathbb{T}^{X_C} & \text{for all } i \geq C. \end{cases}$$

i.e. C is a time where the walk first enters a new subtree of $\mathbb{T}$ and never exits it. Let now $(C_n)_{n\geq 1}$ denote the increasing enumeration of these cut times:

$$\begin{cases} C_1 \stackrel{\text{def}}{=} \inf\{k > 0,\ k \text{ is a cut time}\}, \\ C_{n+1} \stackrel{\text{def}}{=} \inf\{k > C_n,\ k \text{ is a cut time}\}, \end{cases}$$

with the convention that $\inf\{\emptyset\} = \infty$ and $C_{n+1} = \infty$ when $C_n = \infty$.

Proposition 6.1. *Suppose that the sequence of cut times $(C_n)_{n\geq 1}$ is well defined (*i.e. *finite a.s.). Suppose further that $\mathbf{E}[C_1^2] < \infty$. Then, there exist deterministic $v, \sigma > 0$ such that*

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

Proof. Let us first note that the event $A \stackrel{\text{def}}{=} \{X \text{ never crosses the edge from } o \text{ to } o\}$ has non zero probability since the walk is transient and no cookies have strength 0 (in this case, the irreducible classes for the matrix P are $\{0\}$ and $[1, \infty)$). Recalling that the walk evolves independently on distinct subtrees, it is easily seen that the sequence $(C_{n+1} - C_n, |X_{C_{n+1}}| - |X_{C_n}|)_{n\geq 1}$ is i.i.d. and distributed as $(C_1, |X_{C_1}|)$ under the conditional measure $\mathbf{P}\{\cdot|A\}$ (*c.f.* for instance [7; 13] for details). Since $\mathbf{P}\{A\} > 0$ and the walk X is nearest neighbor, we get $\mathbf{E}[(C_{n+1} - C_n)^2] = \mathbf{E}[C_1^2|A] < \infty$ and $\mathbf{E}[(|X_{C_{n+1}}| - |X_{C_n}|)^2] = \mathbf{E}[|X_{C_1}|^2\ |A] < \infty$. Thus, we have

$$\frac{C_n}{n} \xrightarrow[n\to\infty]{\text{a.s.}} \mathbf{E}[C_1|A], \quad \frac{|X_{C_n}|}{n} \xrightarrow[n\to\infty]{\text{a.s.}} \mathbf{E}[|X_{C_1}||A], \quad \frac{|X_{C_n}| - \mathbf{E}[|X_{C_1}||A]n}{\sqrt{n}} \xrightarrow[n\to\infty]{\text{law}} \mathcal{N}(0, \mathbf{E}[|X_{C_1}|^2|A]),$$

and the proposition follows from a change of time, *c.f.* [7; 13] for details. □

Theorem 1.3 will now follow from Proposition 6.1 once we have shown that the cut times of the walk are well defined and have a finite second moment. We shall, in fact, prove the stronger result:

Proposition 6.2. *The cut times of the walk are well defined and, for all $\beta > 0$, $\mathbf{E}[C_1^\beta] < \infty$.*

The proof of this result relies on the following two lemmas whose proofs are provided after the proof of the proposition.

Lemma 6.3. *Recall the definition of the branching Markov chain L. Let U denote the total number of particles not located at 0 for the entire lifetime of the process* i.e.

$$U \stackrel{\text{def}}{=} \sharp\{x \in \mathbb{T},\ \ell(x) > 0\}.$$

There exists $c_1 > 0$ such that, for all n,

$$\mathbf{P}_1\{U > n \mid L \text{ dies out}\} \leq c_1 e^{-n^{1/3}}.$$

Lemma 6.4. *Let $(\gamma_n)_{n\geq 0}$ denote the increasing sequence of times where the walk visits a new site:*

$$\begin{cases} \gamma_0 \stackrel{\text{def}}{=} 0 \\ \gamma_{n+1} \stackrel{\text{def}}{=} \inf\{k > \gamma_n,\ X_k \neq X_i \text{ for all } i < k\}. \end{cases}$$

There exist $\nu, c_2 > 0$ such that, for all n,

$$\mathbf{P}\{\gamma_n > n^\nu\} \leq c_2 e^{-n}.$$

Proof of Proposition 6.2. We need to introduce some notation. We define two interlaced sequences $(S_i)_{i\geq 0}$ and $(D_i)_{i\geq 0}$ by

$$\begin{cases} S_0 = \gamma_1, \\ D_0 = \inf\{n > S_0,\ X_n = \overleftarrow{X}_{S_0} = o\}, \end{cases}$$

and by induction, for $k \geq 1$,

$$\begin{cases} S_k = \inf\{\gamma_n,\ \gamma_n > D_{k-1}\}, \\ D_k = \inf\{n > S_k,\ X_n = \overleftarrow{X}_{S_k}\}. \end{cases}$$

with the convention that, if $D_k = \infty$, then $D_j, S_j = \infty$ for all $j \geq k$. Let us set

$$\chi \stackrel{\text{def}}{=} \inf\{k \geq 0, D_k = \infty\}.$$

Since the walk visits infinitely many distinct vertices, we have $S_k < \infty$ whenever $D_{k-1} < \infty$ so that these two interlaced sequences have the form

$$S_0 < D_0 < S_1 < D_1 < \ldots < S_\chi < D_\chi = \infty.$$

The interval $[S_k, D_k)$ represents the times where the walk performs an excursion away from the set of vertices it has already visited before time S_k. With these notations, the first cut time is given by

$$C_1 = S_\chi.$$

For n, m such that $X_m \in \mathbb{T}^{X_n}$, we use the slight abuse of notation $X_m - X_n$ to denote the position of X_m shifted by X_n *i.e.* the position of X_m with respect to the subtree $\mathbb{T}^{X_n}$. Using the Markov property for the stopping times S_k, D_k and noticing that the walk evolves on distinct subtrees on the time intervals $[S_k, D_k)$, it follows that (compare with Lemma 3 of [7] for details):

(a) Conditionally on $D_k < \infty$ (*i.e.* $\chi > k$), the sequences $((X_{S_j+i} - X_{S_j})_{0\leq i<D_j-S_j},\ j \leq k)$ are i.i.d. and distributed as $(X_i)_{i<D}$ under the conditional measure $\mathbf{P}\{\cdot | D < \infty\}$ with $D = \inf\{k \geq 1, X_{k-1} = X_k = o\}$.

(b) Conditionally on $D_k < \infty$, the random variable $D_{k+1} - S_{k+1}$ has the same distribution as $D_0 - S_0$. In particular, $\mathbf{P}\{D_{k+1} < \infty \mid D_k < \infty\} = \mathbf{P}\{D_0 < \infty\}$. Thus, χ has a geometric distribution with parameter $r \stackrel{\text{def}}{=} \mathbf{P}\{D_0 < \infty\} = \mathbf{P}_1\{L \text{ dies out}\} > 0$:

$$\mathbf{P}\{\chi = k\} = (1 - r)r^k \quad \text{for } k \geq 0.$$

Fact (b) implies, in particular, that the first cut time $C_1 = S_\chi$ (and thus all cut times) is finite almost surely. It remains to bound the moments of C_1. We write

$$\begin{aligned}\mathbf{P}\{S_\chi > n^\nu\} &= \mathbf{P}\{S_\chi > n^\nu \text{ and } \chi \geq \alpha \ln n\} + \mathbf{P}\{S_\chi > n^\nu \text{ and } \chi < \alpha \ln n\} \\ &\leq r^{\alpha \ln n} + \mathbf{P}\{S_\chi > n^\nu \text{ and } \chi < \alpha \ln n\}\end{aligned}$$

where $\alpha > 0$ and where ν is the constant of Lemma 6.4. Let $\beta > 0$ be fixed, we can choose α large enough so that

$$\mathbf{P}\{S_\chi > n^\nu\} \leq \frac{1}{n^{(\beta+1)\nu}} + \mathbf{P}\{S_\chi > n^\nu \text{ and } \chi < \alpha \ln n\}. \tag{28}$$

It remains to find an upper bound for the second term. Let us first note that

$$\mathbf{P}\{S_\chi > n^\nu \text{ and } \chi < \alpha \ln n\} \leq \sum_{k=0}^{\alpha \ln n} \mathbf{P}\{S_k > n^\nu \text{ and } \chi \geq k\}. \tag{29}$$

We introduce the sequence $(V_k)_{k\geq 0}$ defined by

$$V_k \stackrel{\text{def}}{=} \text{number of distinct vertices visited by the walk during the excursion } [S_k, D_k),$$

with the convention that $V_k = \infty$ when $D_k = \infty$. By definition of S_k, D_k, the total number of distinct vertices other than the root visited by the walk up to time S_k is exactly the sum of the number of vertices visited in each excursion $[S_i, D_i)$ $(i < k)$ which is $V_0 + \ldots + V_{k-1}$. Thus, S_k is the time where the walk visits its $(V_0 + \ldots + V_{k-1} + 2)^{\text{th}}$ new vertex. This yields the identity

$$S_k = \gamma_{V_0 + \ldots + V_{k-1} + 2}$$

which holds for all k with the convention $\gamma_\infty = \infty$. Thus, we can rewrite the r.h.s. of (29) as

$$\sum_{k=0}^{\alpha \ln n} \mathbf{P}\{S_k > n^\nu \text{ and } \chi \geq k\} = \sum_{k=0}^{\alpha \ln n} \mathbf{P}\{\gamma_{V_0 + \ldots + V_{k-1} + 2} > n^\nu \text{ and } V_1 + \ldots + V_{k-1} < \infty\}. \tag{30}$$

Each term on the r.h.s. of (30) is bounded by

$$\begin{aligned}&\mathbf{P}\{\gamma_{V_0+\ldots+V_{k-1}+2} > n^\nu \text{ and } V_1 + \ldots + V_{k-1} < \infty\} \\ &\qquad = \mathbf{P}\{\gamma_{V_0+\ldots+V_{k-1}+2} > n^\nu \text{ and } n < V_0 + \ldots + V_{k-1} + 2 < \infty\} \\ &\qquad\quad + \mathbf{P}\{\gamma_{V_0+\ldots+V_{k-1}+2} > n^\nu \text{ and } V_0 + \ldots + V_{k-1} + 2 \leq n\} \\ &\qquad \leq \mathbf{P}\{n - 2 < V_0 + \ldots + V_{k-1} < \infty\} + \mathbf{P}\{\gamma_n > n^\nu\} \\ &\qquad \leq \mathbf{P}\{n - 2 < V_0 + \ldots + V_{k-1} < \infty\} + c_2 e^{-n}\end{aligned} \tag{31}$$

where we used Lemma 6.4 for the last inequality. Let us note that, according to Fact (a), conditionally on $\{V_0 + \ldots + V_{k-1} < \infty\} = \{\chi \geq k\}$, the random variables $(V_0, V_1, \ldots, V_{k-1})$ are i.i.d. and have the same law as the number of vertices visited by the walk before the time D of its first jump from the root to the root under the conditional measure $\mathbf{P}\{\cdot \mid D < \infty\}$. Recalling the construction of the branching Markov chain L described in Section 3, we see that this distribution is exactly that of the random variable U of Lemma 6.3 under the measure $\tilde{\mathbf{P}} \stackrel{\text{def}}{=} \mathbf{P}_1\{\cdot \mid \text{L dies out}\}$. Let now $(U_i)_{i\geq 0}$ denote a sequence of i.i.d. random variables with the same distribution as U under $\tilde{\mathbf{P}}$. For $k \leq \alpha \ln n$, we get

$$\begin{aligned}
\mathbf{P}\{n-2 < V_0 + \ldots + V_{k-1} < \infty\} &\leq \mathbf{P}\{V_0 + \ldots + V_{k-1} > n-2 \mid V_0 + \ldots + V_{k-1} < \infty\} \\
&= \tilde{\mathbf{P}}\{U_0 + \ldots + U_{k-1} > n-2\} \\
&\leq (\alpha \ln n)\tilde{\mathbf{P}}\left\{U > \frac{n-2}{\alpha \ln n}\right\} \\
&\leq c_1(\alpha \ln n)\exp\left(-\left(\frac{n-2}{\alpha \ln n}\right)^{\frac{1}{3}}\right) \qquad (32)
\end{aligned}$$

where we used Lemma 6.3 for the last inequality. Combining (28)-(32), we conclude that

$$\mathbf{P}\{S_\chi > n^\nu\} \leq \frac{1}{n^{(\beta+1)\nu}} + c_2(\alpha \ln n)e^{-n} + c_1(\alpha \ln n)^2 \exp\left(-\left(\frac{n-2}{\alpha \ln n}\right)^{\frac{1}{3}}\right) \leq \frac{2}{n^{(\beta+1)\nu}}$$

for all n large enough. This yields $\mathbf{E}[S_\chi^\beta] < \infty$. □

We now provide the proof of the lemmas.

Proof of Lemma 6.3. Let $\sharp L_n$ denote the number of particles not located at 0 at time n:

$$\sharp L_n \stackrel{\text{def}}{=} \sharp\{x \in \mathbb{T}_n,\ \ell(x) > 0\}.$$

Let also Θ stand for the lifetime of L:

$$\Theta \stackrel{\text{def}}{=} \inf\{n,\ \sharp L_n = 0\}$$

with the convention $\Theta = \infty$ when L does not die out. Since no cookie has strength 0, the irreducible classes of P are $\{0\}$ and $[1,\infty)$. Thus, the transience of the walk implies $\mathbf{P}_1\{\Theta < \infty\} \in (0,1)$. Let H denote the maximal number of particles alive at the same time for the process L:

$$H \stackrel{\text{def}}{=} \sup_n \sharp L_n.$$

It follows from the inequality $U \leq H\Theta$ that

$$\mathbf{P}_1\{U \geq n, \Theta < \infty\} \leq \mathbf{P}_1\{H \geq \sqrt{n}, \Theta < \infty\} + \mathbf{P}_1\{H < \sqrt{n}, \Theta \geq \sqrt{n}\}. \qquad (33)$$

The first term on the r.h.s of (33) is easy to bound. Recalling the monotonicity property of Proposition 3.6, we have $\mathbf{P}_j\{\Theta < \infty\} \leq \mathbf{P}_1\{\Theta < \infty\}$ for any $j \geq 1$. Therefore, using the Markov property of L with the stopping time $\zeta \stackrel{\text{def}}{=} \inf\{k,\ \sharp\{x \in \mathbb{T}_k,\ \ell(x) > 0\} \geq \sqrt{n}\}$, we get, with obvious notation,

$$\begin{aligned}
\mathbf{P}_1\{H \geq \sqrt{n}, \Theta < \infty\} &= \mathbf{E}_1[\mathbb{1}_{\{\zeta<\infty\}}\mathbf{P}_{L_\zeta}\{\Theta < \infty\}] \\
&\leq \mathbf{P}_{\lfloor\sqrt{n}\rfloor \text{ particles loc. at } 1}\{\Theta < \infty\} \\
&= \mathbf{P}_1\{\Theta < \infty\}^{\lfloor\sqrt{n}\rfloor} \\
&\leq e^{-n^{1/3}} \qquad (34)
\end{aligned}$$

where the last inequality holds for n large enough. We now compute an upper bound for the second term on the r.h.s. of (33). Given $k < \sqrt{n}$, it follows again from Proposition 3.6 that,

$$\mathbf{P}_1\{H < \sqrt{n}, \Theta \geq \sqrt{n}\} \leq \mathbf{P}_1\{\sharp L_k < \sqrt{n}\}^{\lfloor n/(k+1) \rfloor}. \tag{35}$$

(this bound is obtained by considering the process where all the particles at time $(k+1), 2(k+1), \ldots$ are replaced by a single particle located at 1). Let us for the time being admit that there exist $\rho > 1$ and $\alpha > 0$ such that,

$$a \stackrel{\text{def}}{=} \liminf_{i \to \infty} \mathbf{P}_1\{\sharp L_i \geq \alpha \rho^i\} > 0. \tag{36}$$

Then, choosing $k = \lfloor \ln n / \ln \rho \rfloor$, the combination of (35) and (36) yields, for all n large enough,

$$\mathbf{P}_1\{H < \sqrt{n}, \Theta \geq \sqrt{n}\} \leq \mathbf{P}_1\{\sharp L_k < \alpha \rho^k\}^{\lfloor n/(k+1) \rfloor} \leq (1-a)^{\lfloor n/(k+1) \rfloor} \leq e^{-n^{1/3}}. \tag{37}$$

Putting (33),(34) and (37) together, we conclude that,

$$\mathbf{P}_1\{U \geq n \mid L \text{ dies out}\} = \frac{\mathbf{P}_1\{U \geq n, \Theta < \infty\}}{\mathbf{P}_1\{\Theta < \infty\}} \leq c_1 e^{-n^{1/3}}$$

which is the claim of the Lemma. It remains to prove (36). Recall that q represents the bias of the walk when all the cookie have been eaten. We consider separately the two cases $q < b/(b+1)$ and $q \geq b/(b+1)$.

(a) $\mathbf{q} < \mathbf{b}/(\mathbf{b}+\mathbf{1})$. Since the walk is transient, the spectral radius of the irreducible class $[1, \infty)$ of the matrix P is necessarily strictly larger than $1/b$ (otherwise the walk would be recurrent according to Proposition 5.2). Using exactly the same arguments as in the proof of Proposition 5.5, we can find r large enough such that the finite sub-matrix $(p_{i,j})_{1 \leq i,j \leq r}$ is irreducible with spectral radius $\tilde{\lambda}$ strictly larger than $1/b$. We consider again the process $\tilde{L}$ obtained from L by removing all the particles x (along with their progeny) whose position $\ell(x)$ is not in $[1, r]$. As already noticed in the proof of Proposition 5.5, the process $\tilde{L}$ is a positive regular, non singular, multi-type branching process with a finite number of types and with parameter $\rho = b\tilde{\lambda} > 1$. Therefore, Theorem 1 p192 of [3] implies that, for $\alpha > 0$ small enough,

$$\lim_{i \to \infty} \mathbf{P}_1\{\sharp \tilde{L}_i \geq \alpha \rho^i\} > 0$$

which, in turn, implies (36).

(b) $\mathbf{q} \geq \mathbf{b}/(\mathbf{b}+\mathbf{1})$. The spectral radius λ of the irreducible class $[1, \infty)$ may, in this case, be strictly smaller that $1/b$ (see the remark below the statement of Theorem 1.2). However, as shown during the proof of Proposition 5.6, we can always find $\hat{q} < b/(b+1) < q$ such that the walk in the cookie environment $(0, \ldots, 0; \hat{q})$ is transient. Therefore, the walk in the cookie environment $\hat{\mathscr{C}} = (p_1, \ldots, p_M; \hat{q}) \leq \mathscr{C}$ is also transient. Denoting by $\hat{L}$ the branching Markov chain associated with $\hat{\mathscr{C}}$, it follows from the previous case (a) combined with Proposition 3.7 that, for some $\rho > 1$, $\alpha > 0$,

$$\liminf_{i \to \infty} \mathbf{P}_1\{\sharp L_i \geq \alpha \rho^i\} \geq \liminf_{i \to \infty} \mathbf{P}_1\{\sharp \hat{L}_i \geq \alpha \rho^i\} > 0.$$

□

Proof of Lemma 6.4. Recall that, given $x \in \mathbb{T}$ and $i \in \{0,\dots,b\}$, we denote by $\overrightarrow{x}^{i}$ the i^{th} child of x (with the convention $\overrightarrow{x}^{0} = \overleftarrow{x}$). We call (un-rooted) path of length k an element $[v_1,\dots,v_k] \in \{0,\dots,b\}^k$. Such a path is said to be increasing if $v_i \neq 0$ for all $1 \le i \le k$. The sub-paths of v are the paths of the form $[v_i,\dots,v_j]$ with $0 \le i \le j \le k$. Given $x \in \mathbb{T}$, we use the notation $x[v_1,\dots,v_k]$ to denote the endpoint of the path rooted at site x *i.e.*

$$x[\emptyset] \overset{\text{def}}{=} x \quad \text{and} \quad x[v_1,\dots,v_k] \overset{\text{def}}{=} \overrightarrow{x[v_1,\dots,v_{k-1}]}^{v_k}.$$

The proof of the lemma is based on the following observation: given two increasing paths v, w with same length k such that $[v_1,\dots,v_k] \neq [w_1,\dots,w_k]$, we have

$$x[v_1,\dots,v_k] \neq y[w_1,\dots,w_k] \quad \text{for any } x, y \in \mathbb{T}.$$

Let $(u_k)_{k\ge1}$ denote the sequence of random variables taking values in $\{0,\dots,b\}$ defined by $X_n = \overrightarrow{X}_{n-1}^{u_n}$. With the previous notation, we have, for any $m \le n$,

$$X_n = X_m[u_{m+1},\dots,u_n].$$

It follows from the previous remark that, for any fixed $k \le n$, the number of distinct vertices visited by the walk X up to time n is larger than the number of distinct increasing sub-paths of length k in the random path $[u_1,\dots,u_n]$. We get a lower bound for the number of such sub-paths using a coupling argument. Recall that no cookie has strength 0 and set $\eta = \min\{\frac{p_1}{b},\dots,\frac{p_M}{b},\frac{q}{b}\} > 0$. It is clear from the definition of the transition probabilities of the cookie random walk X that

$$\begin{cases} \mathbf{P}\{u_n = i \mid u_1,\dots,u_n\} \ge \eta & \text{for } i \in \{1,\dots,b\}, \\ \mathbf{P}\{u_n = 0 \mid u_1,\dots,u_n\} \le 1 - b\eta. \end{cases}$$

Therefore, we can construct on the same probability space a sequence of i.i.d random variables $(\tilde{u}_n)_{n\ge1}$ with distribution:

$$\begin{cases} \mathbf{P}\{\tilde{u}_n = i\} = \eta & \text{for } i \in \{1,\dots,b\} \\ \mathbf{P}\{\tilde{u}_n = 0\} = 1 - b\eta, \end{cases}$$

in such way that

$$\tilde{u}_n = i \neq 0 \quad \text{implies} \quad u_n = i.$$

With this construction, any increasing sub-path of $[\tilde{u}_1,\dots,\tilde{u}_n]$ is also an increasing sub-path of $[u_1,\dots,u_n]$. Moreover, since the sequence $(\tilde{u}_n)_{n\ge1}$ is i.i.d., we have, for any fixed increasing path $[v_1,\dots,v_k]$,

$$\begin{aligned} &\mathbf{P}\left\{[\tilde{u}_1,\dots,\tilde{u}_n] \text{ does not contain the sub-path } [v_1,\dots,v_k]\right\} \\ &\qquad\le \prod_{j=1}^{\lfloor n/k \rfloor} \mathbf{P}\left\{[\tilde{u}_{(j-1)k+1},\dots,\tilde{u}_{jk}] \neq [v_1,\dots,v_k]\right\} = (1-\eta^k)^{\lfloor n/k \rfloor}. \end{aligned} \tag{38}$$

We now choose $k \stackrel{\text{def}}{=} \lfloor c \ln n \rfloor + 1$ with $c \stackrel{\text{def}}{=} \frac{1}{3\ln(1/\eta)}$ and set $\delta \stackrel{\text{def}}{=} c \ln b$. Since there are $b^k > n^\delta$ increasing paths of length k, we get, for n large enough,

$$\begin{aligned} &\mathbf{P}\left\{[\tilde{u}_1, \dots, \tilde{u}_n] \text{ contains less than } n^\delta \text{ distinct increasing sub-paths of same length}\right\} \\ &\qquad \leq \mathbf{P}\left\{[\tilde{u}_1, \dots, \tilde{u}_n] \text{ does not contain all increasing sub-paths of length } k\right\}\\ &\qquad \leq b^k(1-\eta^k)^{\lfloor n/k \rfloor} \\ &\qquad \leq e^{-\sqrt{n}}. \end{aligned}$$

Thus, if $\mathscr{V}_n$ denotes the number of distinct vertices visited by the cookie random walk X up to time n, we have proved the lower bound:

$$\mathbf{P}\{\mathscr{V}_n \leq n^\delta\} \leq e^{-\sqrt{n}}.$$

Choosing $\nu > \max(1/\delta, 2)$, we conclude that, for all n large enough,

$$\mathbf{P}\{\gamma_n \geq n^\nu\} = \mathbf{P}\{\mathscr{V}_{\lfloor n^\nu \rfloor} \leq n\} \leq \mathbf{P}\{\mathscr{V}_{\lfloor n^\nu \rfloor} \leq \lfloor n^\nu \rfloor^\delta\} \leq e^{-\sqrt{\lfloor n^\nu \rfloor}} \leq e^{-n}.$$

□

6.2 Example of a transient walk with sub-linear growth

In this section, we prove Proposition 1.9 whose statement is repeated below.

Proposition 6.5. *Let X be a $\mathscr{C} = (p_1, p_2, 0, 0\,; q)$ cookie random walk with $q \geq b/(b+1)$ and $p_1, p_2 > 0$ such that the largest positive eigenvalue of the matrix*

$$P_1 \stackrel{\text{def}}{=} \begin{pmatrix} \frac{p_1}{b} + \frac{p_1 p_2}{b} - \frac{2 p_1 p_2}{b^2} & \frac{p_1 p_2}{b^2} \\ \frac{p_1 + p_2}{b} - \frac{2 p_1 p_2}{b^2} & \frac{p_1 p_2}{b^2} \end{pmatrix}$$

is equal to $1/b$ (such a choice of p_1, p_2 exists for any $b \geq 2$). Then, X is transient (since $q \geq b/(b+1)$) yet

$$\liminf_{n\to\infty} \frac{|X_n|}{n} = 0.$$

Proof. For this particular cookie environment, it is easily seen that the cookie environment matrix P has three irreducible classes $\{0\}, [1,2], [3,\infty)$ and takes the form

$$P = \begin{pmatrix} \boxed{1} & & \\ & \boxed{P_1} & 0 \\ & * & \underset{\text{(infinite class)}}{*} \end{pmatrix}.$$

where P_1 is the matric given in the proposition. By hypothesis, the spectral radius of the irreducible class $[1,2]$ is $1/b$, therefore, the branching Markov chain L starting from $\ell(o) \in [1,2]$ dies out almost surely (the restriction of L to $[1,2]$ is simply a critical 2-type branching process where each particle gives birth to, at most, 2 children). In particular, the quantity

$$\Lambda \stackrel{\text{def}}{=} \sum_{x \in \mathbb{T}} \ell(x)$$

is $\mathbf{P}_i$ almost surely finite for $i \in \{1,2\}$. Moreover, one can exactly compute the generating functions $\mathbf{E}_i[s^\Lambda]$ for $i \in \{1,2\}$ using the recursion relation given by the branching structure of L. After a few lines of elementary (but tedious) calculus and using a classical Tauberian theorem, we get the following estimate on the tail distribution of Λ:

$$\mathbf{P}_1\{\Lambda > x\} \sim \frac{C}{\sqrt{x}},$$

for some constant $C > 0$ depending on p_1, p_2 (alternatively, one can invoke Theorem 1 of [9] for the total progeny of a general critical multi-type branching process combined with the characterization of the domain of attraction to a stable law).

As in the previous section, let $(\gamma_n)_{n \geq 0}$ denote the increasing sequence of times where the walk visits a new site (as defined in Lemma 6.4) and define two interlaced sequences $(S_i)_{i \geq 0}$ and $(D_i)_{i \geq 0}$ in a similar way as in the proof of Proposition 6.2 (only the initialization changes):

$$\begin{cases} S_0 = 0 \\ D_0 = \inf\{n > 0,\ X_n = X_{n-1} = o\}, \end{cases}$$

and by induction, for $k \geq 1$,

$$\begin{cases} S_k = \inf\{\gamma_n,\ \gamma_n > D_{k-1}\}, \\ D_k = \inf\{n > S_k,\ X_n = \overleftarrow{X_{S_k}}\}. \end{cases}$$

Since, L starting from $\ell(0) = 1$ dies out almost surely, the walk crosses the edge from the root to the root at least once almost surely. Therefore, D_0 is almost surely finite. Using the independence of the cookie random walk on distinct subtrees, it follows that the random variables sequences $(S_i)_{i \geq 0}$, $(D_i)_{i \geq 0}$ are all finite almost surely. Moreover, recalling the construction of L, it also follows that the sequence of excursion lengths $(R_i)_{i \geq 0} \stackrel{\text{def}}{=} (D_i - S_i)_{i \geq 0}$ is a sequence of i.i.d. random variables, distributed as the random variable $2\Lambda - 1$ under $\mathbf{P}_1$. We also have the trivial facts:

- For all $k > 0$, $|X_{D_k}| = |X_{S_k}| - 1$.
- The walk only visits new vertices of the tree during the time intervals $([S_k, D_k])_{k \geq 0}$. Thus, the number of vertices visited by the walk at time S_k is smaller than $1 + \sum_{i=0}^{k-1} R_i$. In particular, we have $|X_{S_k}| \leq 1 + \sum_{i=0}^{k-1} R_i$.
- For all $k \geq 0$, $D_k \geq \sum_{i=0}^{k} R_i$.

Combining these three points, we deduce that, for $k \geq 1$,

$$\frac{|X_{D_k}|}{D_k} \leq \frac{\sum_{i=0}^{k-1} R_i}{\sum_{i=0}^{k} R_i} = 1 - \frac{R_k}{\sum_{i=0}^{k} R_i}.$$

Since $(R_i)_{i\geq 0}$ is a sequence of i.i.d random variables in the domain of normal attraction of a positive stable distribution of index $1/2$, it is well known that

$$\limsup_{k\to\infty} \frac{R_k}{\sum_{i=0}^k R_i} = 1 \quad a.s.$$

(this result follows, for instance, from Exercise 6, p.99 of [8] by considering a subordinator without drift and with Lévy measure Λ). We conclude that

$$\liminf_{n\to\infty} \frac{|X_n|}{n} \leq \liminf_{k\to\infty} \frac{|X_{D_k}|}{D_k} = 0 \quad a.s.$$

□

7 Computation of the spectral radius

In this section, we prove Theorem 1.4 and Proposition 1.8 by computing the maximal spectral radius λ of the cookie environment matrix P. Recall that the irreducible classes of P are $\{0\}, [l_1, r_1], \ldots, [r_K, \infty)$ and that P_k denotes the restriction of P to $[l_k, r_k]$ ($[l_k, \infty)$ for $k = K$). Denoting by λ_k the spectral radius of P_k, we have, by definition:

$$\lambda = \max(\lambda_1, \ldots, \lambda_K).$$

Since the non negative matrices $P_1, \ldots, P_{K-1}$ are finite, their spectral radii are equal to their largest eigenvalue. Finding the spectral radius of the infinite matrix P_K is more complicated. We shall make use on the following result.

Proposition 7.1. *Let $Q = (q(i,j))_{i,j\geq 1}$ be an infinite irreducible non negative matrix. Suppose that there exists a non-negative left eigenvector $Y = (y_i)_{i\geq 1}$ of Q associated with some eigenvalue $\nu > 0$* i.e.

$$^tYQ = \nu\, ^tY. \tag{39}$$

Assume further that, for all $\varepsilon > 0$, there exists $N \geq 1$ such that the finite sub-matrix $Q_N = (q(i,j))_{1\leq i,j\leq N}$ is irreducible and the sub-vector $Y_N = (y_i)_{1\leq i\leq N}$ is $\nu - \varepsilon$ super-invariant i.e

$$^tY_N Q_N \geq (\nu - \varepsilon)\, ^tY_N. \tag{40}$$

Then, the spectral radius of Q is equal to ν.

Remark 7.2. By symmetry, the proposition above remains unchanged if one considers a right eigenvector in place of a left eigenvector. Let us also note that Proposition 7.1 does not cover all possible cases. Indeed, contrarily to the finite case, there exist infinite non negative irreducible matrices for which there is no eigenvector Y satisfying Proposition 7.1.

Proof. On the one hand, according to Criterion I of Corollary 4.1 of [23], the spectral radius λ_Q of Q is the smallest value for which there exists a non negative vector $Y \neq \mathbf{0}$ such that

$$^tYQ \leq \lambda_Q\, ^tY.$$

Therefore, we deduce from (39) that

$$\nu \geq \lambda_Q.$$

On the other hand, the matrix Q_N is finite so that, according to the Perron-Frobenius Theorem, its spectral radius is equal to its largest eigenvalue λ_{Q_N} and is given by the formula

$$\lambda_{Q_N} = \sup_{(x_1,\dots,x_N)} \min_j \frac{\sum_{i=1}^N x_i q(i,j)}{x_j}$$

where the supremum is taken over all N-dimensional vectors with strictly positive coefficients (*c.f.* (1.1) p.4 of [19]). In view of (40), we deduce that $\lambda_{Q_N} \geq \nu - \varepsilon$.

Furthermore, when Q_N is irreducible, Theorem 6.8 of [19] states that $\lambda_{Q_N} \leq \lambda_Q$. We conclude that

$$\lambda_Q \leq \nu \leq \lambda_Q + \varepsilon.$$

□

7.1 Preliminaries

Recall the construction of the random variables $(\xi_i)_{i\geq 1}$ given in Definition 3.1 and set

$$\begin{aligned}
\mathscr{E}_{m,n} &\stackrel{\text{def}}{=} \{\text{in the finite sequence } (\xi_1,\xi_2,\dots,\xi_M), \text{ there are at least } m \text{ terms equal to } 0 \\
&\qquad\qquad \text{and exactly } n \text{ terms are equal to 1 before the } m^{\text{th}}\ 0\} \\
\mathscr{E}'_{m,n} &\stackrel{\text{def}}{=} \{\text{in the finite sequence } (\xi_1,\xi_2,\dots,\xi_M), \text{ there are exactly } m \text{ terms equal to } 0 \\
&\qquad\qquad \text{and exactly } n \text{ terms equal to } 1\}.
\end{aligned}$$

Let us note that, for $n+m > M$,

$$\mathbf{P}\{\mathscr{E}_{m,n}\} = \mathbf{P}\{\mathscr{E}'_{m,n}\} = 0. \tag{41}$$

In the rest of this section, we use the notation

$$s \stackrel{\text{def}}{=} \frac{q}{q+(1-q)b} = \mathbf{P}\{\xi_{M+1} = 1 \mid \xi_{M+1} \in \{0,1\}\}.$$

Lemma 7.3. *For $i,j \geq 1$, the coefficient $p(i,j)$ of the matrix P associated with the cookie environment $\mathscr{C} = (p_1,\dots,p_M\,;q)$ is given by*

$$p(i,j) \;=\; \mathbf{P}\{\mathscr{E}_{i,j}\} \;+ \sum_{\substack{0\leq n\leq j\\ 0\leq m\leq i-1}} \mathbf{P}\{\mathscr{E}'_{m,n}\}\binom{j+i-m-n-1}{j-n} s^{j-n}(1-s)^{i-m}.$$

Proof. Recall that $p(i,j)$ is equal to the probability of having j times 1 in the sequence $(\xi_l)_{l\geq 1}$ before the i^{th} 0. We decompose this event according to the number of 0's and 1's in the subsequence $(\xi_l)_{l\leq M}$. Let $\mathscr{F}_{m,n}$ be the event

$$\mathscr{F}_{m,n} \stackrel{\text{def}}{=} \{\text{in the sub-sequence } (\xi_i)_{i>M},\ n \text{ terms equal to 1 before the } m^{\text{th}} \text{ failure}\}.$$

Thus we have

$$p(i,j) = \mathbf{P}\{\mathscr{E}_{i,j}\} + \sum_{\substack{0\le n\le j\\ 0\le m\le i-1}} \mathbf{P}\{\mathscr{E}'_{m,n}\}\mathbf{P}\{\mathscr{F}_{i-m,j-n}\} \tag{42}$$

(the first term of the r.h.s. of the equation comes from the case $m = i$ which cannot be included in the sum). Since the sequence $(\xi_i)_{i>M}$ is a sequence of i.i.d. random variables, it is easy to compute $\mathbf{P}\{\mathscr{F}_{m,n}\}$. Indeed, noticing that,

$$\mathbf{P}\{\mathscr{F}_{m,n}\} = \mathbf{P}\{\mathscr{F}_{m,n} \mid \xi_l \in \{0,1\} \text{ for all } l \in [M, M+n+m]\},$$

we get

$$\mathbf{P}\{\mathscr{F}_{m,n}\} = \binom{n+m-1}{n} s^n(1-s)^m. \tag{43}$$

The combination of (42) and (43) completes the proof of the lemma. □

We can now compute the image ${}^tYP$ of the exponential vector $Y = ((s/(1-s))^{i-1})_{i\ge 1}$. Let us first recall the notation

$$\lambda_{\text{sym}} \overset{\text{def}}{=} \frac{q}{b(1-q)} \prod_{i=1}^{M} \left((1-p_i)\left(\frac{q}{b(1-q)}\right) + \frac{(b-1)p_i}{b} + \frac{p_i}{b}\left(\frac{q}{b(1-q)}\right)^{-1}\right).$$

We use the convention that $\sum_u^v = 0$ when $u > v$.

Lemma 7.4. *We have*

$$\sum_{i=1}^{\infty} p(i,j)\left(\frac{s}{1-s}\right)^{i-1} = \lambda_{\text{sym}}\left(\frac{s}{1-s}\right)^{j-1} + A(j),$$

with

$$A(j) \overset{\text{def}}{=} \sum_{i=1}^{M-j} \mathbf{P}\{\mathscr{E}_{i,j}\}\left(\frac{s}{1-s}\right)^{i-1} - \sum_{n=j+1}^{M}\sum_{m=0}^{M-n} \mathbf{P}\{\mathscr{E}'_{m,n}\}\left(\frac{s}{1-s}\right)^{j+m-n}.$$

In particular, $A(j) = 0$ for $j \ge M$.

Proof. With the help of Lemma 7.3, and in view of (41), we have

$$\begin{aligned}
&\sum_{i=1}^{\infty} p(i,j)\left(\frac{s}{1-s}\right)^{i-1}\\
&\quad = \sum_{i=1}^{\infty} \mathbf{P}\{\mathscr{E}_{i,j}\}\left(\frac{s}{1-s}\right)^{i-1} + \sum_{i=1}^{\infty}\sum_{\substack{0\le n\le j\\ 0\le m\le i-1}} \mathbf{P}\{\mathscr{E}'_{m,n}\}\binom{j+i-m-n-1}{j-n} s^{j+i-n-1}(1-s)^{1-m}\\
&\quad = \sum_{i=1}^{M-j} \mathbf{P}\{\mathscr{E}_{i,j}\}\left(\frac{s}{1-s}\right)^{i-1} + \sum_{\substack{0\le n\le j\wedge M\\ 0\le m\le M-n}} \mathbf{P}\{\mathscr{E}'_{m,n}\}\sum_{i=0}^{\infty}\binom{j+i-n}{i} s^{j+i+m-n}(1-s)^{1-m}.
\end{aligned}$$

Using the relation

$$\sum_{i=0}^{\infty}\binom{j+i-n}{i} s^{j+i+m-n}(1-s)^{1-m} = \left(\frac{s}{1-s}\right)^{j+m-n},$$

we deduce that

$$\sum_{i=1}^{\infty} p(i,j)\left(\frac{s}{1-s}\right)^{i-1} = \sum_{i=1}^{M-j} \mathbf{P}\{\mathscr{E}_{i,j}\}\left(\frac{s}{1-s}\right)^{i-1} + \sum_{n=0}^{j\wedge M}\sum_{m=0}^{M-n} \mathbf{P}\{\mathscr{E}'_{m,n}\}\left(\frac{s}{1-s}\right)^{j+m-n}.$$
$$= \sum_{n=0}^{M}\sum_{m=0}^{M-n} \mathbf{P}\{\mathscr{E}'_{m,n}\}\left(\frac{s}{1-s}\right)^{j+m-n} + A(j).$$

It simply remains to show that

$$\lambda_{\text{sym}} = \sum_{n=0}^{M}\sum_{m=0}^{M-n} \mathbf{P}\{\mathscr{E}'_{m,n}\}\left(\frac{s}{1-s}\right)^{m-n+1}. \tag{44}$$

Let us note that,

$$\lambda_{\text{sym}} = \frac{s}{1-s}\prod_{l=1}^{M}\left(\frac{s}{1-s}\mathbf{P}\{\xi_l = 0\} + \mathbf{P}\{\xi_l \geq 2\} + \mathbf{P}\{\xi_l = 1\}\frac{1-s}{s}\right).$$

Expanding the r.h.s. of this equation and using the definition of $\mathscr{E}'_{m,n}$, we get (44) which concludes the proof of the lemma. □

We have already noticed that $A(j) = 0$ whenever $j \geq M$. In fact, if some cookies have strength 0, the lower bound on j can be improved. Let M_0 denote the number of cookies with strength 0:

$$M_0 \stackrel{\text{def}}{=} \sharp\{1 \leq i \leq M, p_i = 0\}.$$

Lemma 7.5. *Let $\mathscr{C} = (p_1, \dots, p_M\,; q)$ be a cookie environment with $p_M \neq 0$. We have,*

$$A(j) = 0 \quad \textit{for all } j \geq M - M_0.$$

Proof. Since M_0 cookies have strength 0, there are at most $M - M_0$ terms equal to 1 in the sequence $(\xi_1, \dots, \xi_M)$. Keeping in mind the definitions of $\mathscr{E}_{m,n}$ and $\mathscr{E}'_{m,n}$, we see that

$$\mathbf{P}\{\mathscr{E}_{m,n}\} = \mathbf{P}\{\mathscr{E}'_{m,n}\} = 0 \quad \text{for } n > M - M_0.$$

Moreover, recall that $p_M \neq 0$. Thus, if exactly $M - M_0$ terms are equal to 1, the last one, ξ_M, must also be equal to 1. Therefore, we have

$$\mathbf{P}\{\mathscr{E}_{m,M-M_0}\} = 0.$$

Let us now fix $j \geq M - M_0$, and look at the expression of $A(j)$.

$$A(j) \stackrel{\text{def}}{=} \sum_{i=1}^{M-j} \mathbf{P}\{\mathscr{E}_{i,j}\}\left(\frac{s}{1-s}\right)^{i-1} - \sum_{n=j+1}^{M}\sum_{m=0}^{M-n} \mathbf{P}\{\mathscr{E}'_{m,n}\}\left(\frac{s}{1-s}\right)^{j+m-n}.$$

The terms in the first sum $\sum_{i=1}^{M-j}$ are all zero since $j \geq M - M_0$. Similarly, all the terms in the sum $\sum_{m=0}^{M-n}$ are also zero since $n \geq j + 1 > M - M_0$. □

Proposition 7.6. *Let $\mathscr{C} = (p_1, \dots, p_M; q)$ be a cookie environment such that*

$$q < \frac{b}{b+1} \quad \text{and } M_0 \geq \left\lfloor \frac{M}{2} \right\rfloor.$$

If M is an odd integer, assume further that $p_M \neq 0$. Then, the spectral radius λ_K of the infinite irreducible sub-matrix $P_K = (p(i,j))_{i,j \geq l_K}$ is equal to λ_{sym}.

Proof. Let us note that, when M is an even integer and $p_M = 0$, we can consider $\mathscr{C}$ as the $M+1$ cookie environment $(p_1, \dots, p_M, q; q)$ and this $M+1$ cookie environment still possesses, at least, half of its cookies with zero strength because $\lfloor (M+1)/2 \rfloor = \lfloor M/2 \rfloor$. Thus, we can assume, without loss of generality that the cookie environment is such that $p_M \neq 0$. In order to prove the proposition, we shall prove that

$$Y \stackrel{\text{def}}{=} \left(\left(\frac{s}{1-s} \right)^{i-1} \right)_{i \geq l_K}$$

is a left eigenvector of P_K for the eigenvalue λ_{sym} fulfilling the assumptions of Proposition 7.1. Since the cookie environment has $M_0 \geq \lfloor M/2 \rfloor$ cookies with strength 0, there are, in the $2\lfloor M/2 \rfloor$ first cookies, at most $\lfloor M/2 \rfloor$ random variables taking value 1 in the sequence $(\xi_i)_{i \geq 1}$ before the $\lfloor M/2 \rfloor^{\text{th}}$ failure *i.e.*

$$p(i,j) = 0 \quad \text{for } i \leq \lfloor M/2 \rfloor < j. \tag{45}$$

This implies, in particular, that $l_K \geq \lfloor M/2 \rfloor + 1 \geq M - M_0$. Using Lemma 7.5, we deduce that

$$A(j) = 0 \quad \text{for all } j \geq l_k. \tag{46}$$

Combining (46) and Lemma 7.4, we conclude that Y is indeed a left eigenvector:

$${}^tY P_K = \lambda_{\text{sym}} {}^tY.$$

Let $\varepsilon > 0$. We consider the sub-vector

$$Y_N \stackrel{\text{def}}{=} \left(\left(\frac{s}{1-s} \right)^{i-1} \right)_{l_K \leq i < l_K + N}.$$

It remains to show that, for N large enough, ${}^tY_N P_{K,N} \geq (\lambda_K - \varepsilon) {}^tY_N$ *i.e.*

$$\sum_{i=l_K}^{l_K+N-1} p(i,j) \left(\frac{s}{1-s} \right)^{i-1} \geq (\lambda_K - \varepsilon) \left(\frac{s}{1-s} \right)^{j-1} \quad \text{for all } j \in \{l_K, \dots, l_K + N - 1\}. \tag{47}$$

Keeping in mind that, for $j \geq l_K$

$$\sum_{i=l_K}^{\infty} p(i,j) \left(\frac{s}{1-s} \right)^{i-1} = \lambda_K \left(\frac{s}{1-s} \right)^{j-1},$$

we see that (47) is equivalent to proving that,

$$\sum_{i=l_K+N}^{\infty} p(i,j) \left(\frac{s}{1-s} \right)^{i-1} \leq \varepsilon \left(\frac{s}{1-s} \right)^{j-1} \quad \text{for } j \in \{l_K, \dots, l_K + N - 1\}. \tag{48}$$

Choosing N such that $l_K + N \geq M + 1$, and using the expression of $p(i,j)$ stated in Lemma 7.3, we get, for any $j \in \{l_K, \dots, l_K + N - 1\}$,

$$\sum_{i=l_K+N}^{\infty} p(i,j)\left(\frac{s}{1-s}\right)^{i-1} = \sum_{\substack{0\leq n\leq j\\ 0\leq m\leq M}} \mathbf{P}\{\mathscr{E}'_{m,n}\} \sum_{i=l_K+N}^{\infty} \binom{j+i-m-n-1}{j-n} s^{j+i-1-n}(1-s)^{1-m}$$

where we used that $\mathbf{P}\{\mathscr{E}'_{m,n}\} = \mathbf{P}\{\mathscr{E}_{m,n}\} = 0$ when either n or m is strictly larger than M. We now write

$$\binom{j+i-m-n-1}{j-n} s^{j+i-1-n}(1-s)^{1-m} = \left(\frac{s}{1-s}\right)^{j+m-n} \binom{j+i-m-n-1}{i-m-1} s^{i-m-1}(1-s)^{j-n+1}$$

and we interpret the term

$$\binom{j+i-m-n-1}{i-m-1} s^{i-m-1}(1-s)^{j-n+1}$$

as the probability of having $(i-m-1)$ successes before having $(j-n+1)$ failures in a sequence $(B_r)_{r\geq 1}$ of i.i.d. Bernoulli random variables with distribution $\mathbf{P}\{B_r = 1\} = 1 - \mathbf{P}\{B_r = 0\} = s$. Therefore, we deduce that

$$\begin{aligned}\sum_{i=l_K+N}^{\infty} \binom{j+i-m-n-1}{i-m-1} s^{i-m-1}(1-s)^{j-n+1} &= \mathbf{P}\big\{\text{there are at least } l_K{+}N{-}m{-}1 \text{ successes before the } (j{-}n{+}1)^{\text{th}} \text{ failure in } (B_r)_{r\geq 1}\big\}\\ &\leq \mathbf{P}\big\{\text{there are at least } l_K{+}N{-}M{-}1 \text{ successes before the } (l_K{+}N{+}1)^{\text{th}} \text{ failure in } (B_r)_{r\geq 1}\big\}.\end{aligned}$$

Noticing that $s < 1/2$ since $q < b/(b+1)$, the law of large numbers for the biased Bernoulli sequence $(B_r)_{r\geq 1}$ implies that the above probability converges to 0 as N tends to infinity. Thus, for all $\varepsilon > 0$, we can find $N \geq 1$ such that (48) holds. □

7.2 Proofs of Theorem 1.4 and Proposition 1.8

Proof of Theorem 1.4. Consider a cookie environment $\mathscr{C} = (p_1, \dots, p_M\,; q)$ such that:

$$q < \frac{b}{b+1} \quad \text{and} \quad p_i = 0 \text{ for all } i \leq \lfloor M/2 \rfloor.$$

Recall that $M_0 \stackrel{\text{def}}{=} \sharp\{1 \leq i \leq M, p_i = 0\}$ stands for the number of cookies with strength 0. We simply need to check that the irreducible classes of P are $\{0\}$ and $[M_0 + 1, \infty)$ *i.e* P takes the form:

$$P = \begin{pmatrix} \boxed{1} & & & & 0 \\ & 0 & \dots & 0 & \\ & \vdots & \ddots & \vdots & \\ & * & \dots & 0 & \\ * & & & & \boxed{P_1} \end{pmatrix}$$

and it will follows from Proposition 7.6 that $\lambda = \lambda_1 = \lambda_{\text{sym}}$. Thus, we just need to check that:

(1) for all $1 \le i \le M_0$, $p(i,i) = 0$ (the index i does not belong to any irreducible class).

(2) for all $j \in \mathbb{N}$, $p(M_0+1, j) > 0$ ($M_0 + 1$ belongs to the infinite irreducible class).

The second assertion is straightforward since there are only M_0 cookies with strength 0. In order to see why (1) holds, we consider the two cases:

- $1 \le i \le \lfloor M/2 \rfloor$. Then, clearly $p(i,j) = \mathbb{1}_{\{j=0\}}$. In particular $p(i,i) = 0$.
- $\lfloor M/2 \rfloor + 1 \le i \le M_0$. In this case, there are $M_0 \ge i$ cookies with strength 0 in the first $2i - 1 \ge M$ cookies. Therefore, there cannot be i random variables ξ taking value 1 in the sequence $(\xi_k)_{k\ge1}$ before the i^{th} failure. This means that $p(i,i) = 0$.

□

Proof of Proposition 1.8. Let X be a $(p_1, p_2, \overbrace{0, \dots, 0}^{K times}; q)$ cookie random walk with $K \ge 2$. Using similar argument as before, it is easily checked that the irreducible classes of the cookie environment matrix P are, in this case, $\{0\}$, $[1,2]$ and $[K+1, \infty)$. Moreover, the matrix associated with the irreducible class $[1,2]$ is given, as in Proposition 6.5, by

$$\begin{pmatrix} \frac{p_1}{b} + \frac{p_1 p_2}{b} - \frac{2p_1p_2}{b^2} & \frac{p_1p_2}{b^2} \\ \frac{p_1+p_2}{b} - \frac{2p_1p_2}{b^2} & \frac{p_1p_2}{b^2} \end{pmatrix}. \tag{49}$$

Thus, denoting by ν the largest spectral radius of this matrix and using Proposition 7.6, we deduce that, for $q < \frac{b}{b+1}$, the maximal spectral radius of P is given by:

$$\lambda = \max(\nu, \lambda_{\text{sym}}).$$

We conclude the proof of the Proposition using Theorem 1.2 and the fact that $\lambda_{\text{sym}} > 1/b$ whenever $q \ge \frac{b}{b+1}$ (*c.f.* Remark 1.5). □

8 Other models

8.1 The case $q = 0$.

As stated in Proposition 1.8, a $(p_1, p_2, 0, 0; q)$ cookie random walk is transient as soon as the spectral radius $\nu(p_1, p_2)$ of the matrix given in (49) is strictly larger that $1/b$. Let us remark that this quantity does not depend on q. Therefore, when $\nu(p_1, p_2) > \frac{1}{b}$ the walk is transient for any arbitrarily small q. In fact, using similar arguments to those provided in this paper, one can deal with the case $q = 0$. The study of the walk is even simpler in this case since the cookie environment matrix P does not have an infinite class ($p(i,i) = 0$ for all $i \ge M$). Thus, the process L is, in this case, just a multi-type branching process with finitely many types.

However, when $q = 0$, a $0 - 1$ law does not hold for the walk anymore since it always has a strictly positive probability of getting stuck at o eventually (this probability is bounded below by $\prod(1 - p_i)$).

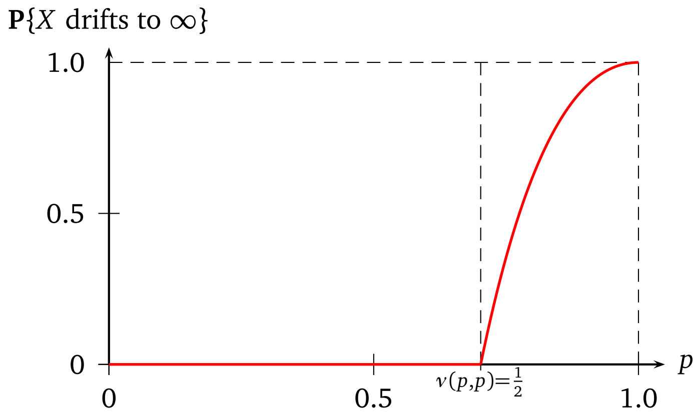

Figure 3: Phase transition of a $(p, p\,; 0)$ cookie random walk on a binary tree.

Therefore, the recurrence/transience criterion now translates to finding whether the walk eventually gets stuck at o with probability 1 or has a positive probability of drifting towards infinity.

For instance, an easy adaptation of Proposition 1.8 (the details are left over for the reader) shows that, for a two cookies environment $\mathscr{C} = (p_1, p_2\,; 0)$, the walk has a positive probability of drifting towards infinity if and only if $\nu(p_1, p_2) > \frac{1}{b}$. Moreover, the process L is, in this setting, a 2-type branching process and the probability that the walk gets stuck at o is equal to the probability that L, starting from one particle located at 2, dies out. This probability of extinction is obtained by computing the fixed point of the generating function of L (*c.f.* Theorem 2, p186 of [3]) and yields

$$\mathbf{P}\{X \text{ is stuck at } o \text{ eventually}\} = \begin{cases} 1 & \text{if } \nu(p_1,p_2) \leq \frac{1}{b}, \\ \frac{(1-p_1)(b+bp_2+p_1^2p_2^3-bp_1p_2^2-p_1p_2^3-bp_1p_2)}{p_1p_2(b-1)} & \text{if } \nu(p_1,p_2) > \frac{1}{b}. \end{cases}$$

An illustration of this phase transition is given in Figure 3 for the case of a binary tree.

8.2 Multi-excited random walks on Galton-Watson trees

In the paper, we assumed that the tree $\mathbb{T}$ is regular. Yet, one may also consider a cookie random walk on more general kinds of trees like, for instance, Galton-Watson trees. Recalling the classical model of biased random walk on Galton-Watson trees [14], a natural way to define the excited random walk on such a tree is as follows: a cookie random walk X on $\mathbb{T}$ in a cookie environment $\mathscr{B} = (\beta_1, \ldots, \beta_M\,; \alpha) \in [0,\infty)^M \times (0,\infty)$ is a stochastic process moving according the following rule:

- If $X_n = x$ is at a vertex with B children and there remain the cookies with strengths $\beta_j, \beta_{j+1}, \ldots, \beta_M$ at this vertex, then X eats the cookie with attached strength β_j and then jumps at time $n+1$ to the father of x with probability $\frac{1}{1+B\beta_j}$ and to each son of x with probability $\frac{\beta_j}{1+B\beta_j}$.
- If $X_n = x$ is at a vertex with B children and there is no remaining cookie at site x, then X jumps at time $n+1$ to the father of x with probability $\frac{1}{1+B\alpha}$ and to each son of x with probability $\frac{\alpha}{1+B\alpha}$.

In the case of a regular b-ary tree, this model coincides with the one studied in this paper with the transformation $p_j = \frac{b\beta_j}{1+b\beta_j}$ and $q = \frac{b\alpha}{b\alpha+1}$. In this new setting, one can still construct a Markov process L associated with the local time process of the walk and one can easily adapt the proof of Theorem 1.2 to show the following result:

Let X be a $\mathscr{B} = (\beta_1, \ldots, \beta_M\,; \alpha)$ cookie random walk on a Galton-Watson tree $\mathbb{T}$ with reproduction law B such that $\mathbf{P}\{B = 0\} = 0$ and $\mathbf{P}\{B = 1\} < 1$ and $\mathbf{E}[B] < \infty$. Fix $b \geq 2$ and let P be the matrix of Definition 3.1 associated with a cookie random walk on a regular b-ary tree in the cookie environment $\mathscr{C} = (p_1, \ldots, p_M\,; q)$, where $p_i \stackrel{\text{def}}{=} \frac{b\beta_i}{b\beta_i+1}$ and $q \stackrel{\text{def}}{=} \frac{b\alpha}{b\alpha+1}$ (this matrix does not, in fact, depend, on the choice of b). Then, the $\mathscr{B}$-cookie random walk on the Galton Watson tree $\mathbb{T}$ is transient if and only if

$$\alpha \geq 1 \quad \text{or} \quad \lambda(\mathscr{C}) > \frac{1}{\mathbf{E}[B]},$$

where $\lambda(\mathscr{C})$ denotes, as before, the largest spectral radius of the irreducible sub-matrices of P.

Acknowledgments. We would like to thank Gady Kozma for stimulating discussion concerning this model. We would also like to thank Sebastian Müller and Serguei Popov for showing us how to prove the recurrence of the walk in the critical case $\lambda(\mathscr{C}) = 1/b$.

References

[1] G. Amir, I. Benjamini and G. Kozma (2008) Excited random walk against a wall, *Probab. Theory Related Fields* **140**, no. 1-2, 83–102. MR2357671

[2] T. Antal and S. Redner (2005) The excited random walk in one dimension, *J. Phys. A* **38**, no. 12, 2555–2577. MR2132073

[3] K. B. Athreya and P. E. Ney (1972) *Branching processes.* Springer-Verlag, New York. Die Grundlehren der mathematischen Wissenschaften, Band 196. MR0373040

[4] A.-L. Basdevant and A. Singh (2008) On the speed of a cookie random walk, *Probab. Theory Related Fields* **141**, no. 3-4, 625–645. MR2391167

[5] A.-L. Basdevant and A. Singh (2008) Rate of growth of a transient cookie random walk, *Electron. J. Probab.* **13**, no. 26, 811–851 (electronic). MR2399297

[6] I. Benjamini and D. B. Wilson (2003) Excited random walk, *Electron. Comm. Probab.* **8**, 86–92 (electronic). MR1987097

[7] J. Bérard and A. Ramírez (2007) Central limit theorem for the excited random walk in dimension $D \geq 2$, *Electron. Comm. Probab.* **12**, 303–314 (electronic).

[8] J. Bertoin (1996) *Lévy processes.* Volume 121 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge. MR1406564

[9] Z. Chi (2004) Limit laws of estimators for critical multi-type Galton-Watson processes, *Ann. Appl. Probab.* **14**, no. 4, 1992–2015. MR2099660

[10] T. E. Harris (1963) *The theory of branching processes*, Springer, Berlin. MR0163361

[11] G. Kozma (2003) Excited random walk in three dimensions has positive speed. Preprint, math.PR/0310305 MR1987097

[12] G. Kozma (2005) Excited random walk in two dimensions has linear speed. Preprint, math.PR/0512535

[13] E. Kosygina and M. P. W. Zerner (2008) Positively and negatively excited random walks on integers, with branching processes, *Electron. J. Probab.* **13**, no. 64, 1952–1979 (electronic). MR2453552

[14] R. Lyons, R. Pemantle and Y. Peres (1996) Biased random walks on Galton-Watson trees, *Probab. Theory Related Fields* **106**, no. 2, 249–264. MR1410689

[15] R. Lyons, R. Pemantle and Y. Peres (1997) Unsolved problems concerning random walks on trees, in *Classical and modern branching processes (Minneapolis, MN, 1994)*, 223–237, Springer, New York. MR1601753

[16] M. V. Menshikov and S. E. Volkov (1997), Branching Markov chains: qualitative characteristics, *Markov Process. Related Fields* **3**, no. 2, 225–241. MR1468175

[17] T. Mountford, L. P. R. Pimentel and G. Valle (2006) On the speed of the one-dimensional excited random walk in the transient regime, *ALEA Lat. Am. J. Probab. Math. Stat.* **2** , 279–296 (electronic). MR2285733

[18] S. Müller (2008) Recurrence for branching Markov chains. *Electron. Comm. Probab.* **13**, 576–605 (electronic).

[19] E. Seneta (1981) *Nonnegative matrices and Markov chains*, Second edition, Springer, New York. MR0719544

[20] E. Seneta and D. Vere-Jones (1966) On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states, *J. Appl. Probability* **3**, 403–434. MR0207047

[21] R. van der Hofstad and M. Holmes (2007) An expansion for self-interacting random walks. Preprint, arxiv.org/abs/0706.0614v3

[22] R. van der Hofstad and M. Holmes (2009) Monotonicity for excited random walk in high dimensions. To appear in *Probab. Theory Related Fields*. arxiv.org/abs/0803.1881

[23] D. Vere-Jones (1967) Ergodic properties of nonnegative matrices. I, *Pacific J. Math.* **22**, 361–386. MR0214145

[24] S. Volkov (2003) Excited random walk on trees, *Electron. J. Probab.* **8**, no. 23, 15 pp. (electronic). MR2041824

[25] M. P. W. Zerner (2005) Multi-excited random walks on integers, *Probab. Theory Related Fields* **133**, no. 1, 98–122. MR2197139

[26] M. P. W. Zerner (2006) Recurrence and transience of excited random walks on $\mathbb{Z}^d$ and strips, *Electron. Comm. Probab.* **11**, 118–128 (electronic). MR2231739