Electronic Journal of Probability

Electron. J. Probab. **23** (2018), no. 43, 1–60.
ISSN: 1083-6489 https://doi.org/10.1214/18-EJP155

Excited random walk in a Markovian environment

Nicholas F. Travers*

Abstract

One dimensional excited random walk has been extensively studied for bounded, i.i.d. cookie environments. In this case, many important properties of the walk including transience or recurrence, positivity or non-positivity of the speed, and the limiting distribution of the position of the walker are all characterized by a single parameter δ, the total expected drift per site. In the more general case of stationary ergodic environments, things are not so well understood. If all cookies are positive then the same threshold for transience vs. recurrence holds, even if the cookie stacks are unbounded. However, it is unknown if the threshold for transience vs. recurrence extends to the case when cookies may be negative (even for bounded stacks), and moreover there are simple counterexamples to show that the threshold for positivity of the speed does not. It is thus natural to study the behavior of the model in the case of Markovian environments, which are intermediate between the i.i.d. and stationary ergodic cases. We show here that many of the important results from the i.i.d. setting, including the thresholds for transience and positivity of the speed, as well as the limiting distribution of the position of the walker, extend to a large class of Markovian environments. No assumptions are made about the positivity of the cookies.

Keywords: excited random walk; cookie random walk; self-interacting random walk; Markovian environment; random environment.
AMS MSC 2010: 60K35; 60K37; 60J80.
Submitted to EJP on May 17, 2017, final version accepted on February 28, 2018.

1 Introduction and statement of results

Let $\Omega = [0,1]^{\mathbb{Z}\times\mathbb{N}}$. An *excited random walk* (ERW) started from position x_0 in a *cookie environment* $\omega = (\omega(x,i))_{x\in\mathbb{Z},i\in\mathbb{N}} \in \Omega$ is an integer-valued stochastic process $(X_n)_{n\geq 0}$ with probability measure $P^{\omega}_{x_0}$ given by

$$\begin{aligned} &P^{\omega}_{x_0}(X_0 = x_0) = 1, \\ &P^{\omega}_{x_0}(X_{n+1} = X_n + 1 | X_0,\ldots,X_n) = \omega(X_n, I_n), \\ &P^{\omega}_{x_0}(X_{n+1} = X_n - 1 | X_0,\ldots,X_n) = 1 - \omega(X_n, I_n), \end{aligned}$$

*Indiana University, Bloomington. E-mail: travers.nicholas@gmail.com

where $I_n = |\{0 \le m \le n : X_m = X_n\}|$. The name cookie environment comes from the following informal interpretation first given in [22]. At each site $x \in \mathbb{Z}$ we initially place an infinite stack of "cookies". The strength of the i-th cookie at site x is $\omega(x,i)$. Upon the i-th visit to x the walker consumes the i-th cookie there, which biases it to jump right with probability $\omega(x,i)$ and left with probability $1-\omega(x,i)$.

A multi-dimensional form of this excited random walk model was first introduced in [4], but with only a single cookie per site. Subsequently the model was extended in [22] to the case of infinite cookie stacks and random environments, for the 1-dimensional setting, and has since been a topic of substantial interest [10, 2, 3, 8, 12, 1, 13, 9]. For a good and fairly recent survey the reader is referred to [11].

Henceforth, we will be concerned only with the case of 1-dimensional ERW when the environment ω is itself random. The setup is analogous to the case of classical random walk in random environment. We first sample $\omega \in \Omega$ according to some probability measure $\mathbf{P}$ on Ω. Then, conditional on selecting the particular environment ω, the random walk proceeds as described above. For an initial position $x \in \mathbb{Z}$, P_x^ω is called the *quenched measure* of the walk in the environment ω, and the *averaged* or *annealed* measure P_x is defined by

$$P_x(\cdot) \equiv \int P_x^\omega(\cdot)\mathbf{P}(d\omega) = \mathbf{E}[P_x^\omega(\cdot)]. \tag{1.1}$$

Of course, to say anything meaningful about the behavior of the walk under the averaged measure P_x one must put some assumptions on the probability measure $\mathbf{P}$ over cookie environments. The following assumptions have often been used in previous works:

(IID) The sequence of cookie stacks $(\omega(x,\cdot))_{x\in\mathbb{Z}}$ is i.i.d. under $\mathbf{P}$.

(SE) The sequence of cookie stacks $(\omega(x,\cdot))_{x\in\mathbb{Z}}$ is stationary and ergodic under $\mathbf{P}$.

(BD) The cookie stacks are of uniform bounded height. That is, there exists a deterministic $M \in \mathbb{N}$ such that $\omega(x,i) = 1/2$ a.s., for all $i > M$ and $x \in \mathbb{Z}$.

(POS) The cookies are positive (or right biased): $\omega(x,i) \ge 1/2$ a.s., for all $x \in \mathbb{Z}$ and $i \in \mathbb{N}$.

(ELL) The cookie environment is elliptic: $\omega(x,i) \in (0,1)$ a.s., for all $x \in \mathbb{Z}$ and $i \in \mathbb{N}$.

If the measure $\mathbf{P}$ satisfies either (IID) or (SE) and either (POS) or (BD) then the total expected drift per site

$$\delta \equiv \sum_{i=1}^{\infty} \Big(2\mathbf{E}[\omega(x,i)] - 1\Big) \tag{1.2}$$

is well defined (possibly $+\infty$) and independent of x, and it turns out that many key properties of the walk depend on this parameter δ. The (ELL) condition is only a technical assumption that is necessary to avoid certain trivialities, and some weaker forms have been used instead in certain instances.

1.1 Summary of known results

Here we list some important known results about the ERW model relevant to our work. Some of these results hold (and were initially stated) with a somewhat weaker form of ellipticity than the (ELL) condition given above, but these differences will not be important for us. The following terminology will be used in our statements, as well as in the statements of the new results that follow in Section 1.2.

- We say that the random walk (X_n) is *recurrent* if $P_0(X_n = 0, \text{ i.o.}) = 1$, *right transient* if $P_0(X_n \to +\infty) = 1$, and *left transient* if $P_0(X_n \to -\infty) = 1$.
- We say that the random walk (X_n) *satisfies a law of large numbers* with *velocity* $v \in \mathbb{R}$ if $\lim_{n\to\infty} X_n/n = v$, P_0 a.s. If $v \neq 0$, we say the random walk is *ballistic*.

Theorem 1.1. *(Zero-One Law for Directional Transience, Theorem 1.2 of [1])* *Assume **P** is (SE) and (ELL). Then* $P_0(X_n \to +\infty) \in \{0, 1\}$ *and* $P_0(X_n \to -\infty) \in \{0, 1\}$.

Theorem 1.2. *(Law of Large Numbers, Theorem 4.1 of [11])* *If **P** is (SE) and the conclusion of Theorem 1.1 holds, then* (X_n) *satisfies a law of large numbers with some velocity* $v \in [-1, 1]$. *In particular, if **P** is (SE) and (ELL), then* (X_n) *satisfies a law of large numbers.*

Theorem 1.3. *(Transience vs. Recurrence Threshold)*

*(i) (Theorem 1 of [10]). Assume **P** satisfies (IID), (BD), and (ELL) and let* δ *be as in* (1.2). *Then* (X_n) *is recurrent if* $\delta \in [-1, 1]$, *right transient if* $\delta > 1$, *and left transient if* $\delta < -1$.

*(ii) (Theorem 12 of [22]) Assume **P** satisfies (SE), (POS), and (ELL) and let* δ *be as in* (1.2). *Then* (X_n) *is recurrent if* $\delta \in [0, 1]$ *and right transient if* $\delta > 1$.

Theorem 1.4. *(Ballisticity Threshold, Theorem 2 of [10])* *Assume **P** satisfies (IID), (BD), and (ELL) and let* δ *be as in* (1.2). *Also, let* v *be the velocity from Theorem 1.2. Then* $v = 0$ *if* $\delta \in [-2, 2]$, $v > 0$ *if* $\delta > 2$, *and* $v < 0$ *if* $\delta < -2$.

Theorem 1.5. *(Limit Laws, Theorem 6.5 of [11]) Assume that **P** satisfies (IID), (BD), and (ELL) and let* δ *be as in* (1.2). *Also, let* v *be the velocity from Theorem 1.2. For* $\alpha \in (0, 2]$ *and* $b > 0$, *let* $Z_{\alpha,b}$ *be a random variable with totally asymmetric stable law of index* α, *defined by its characteristic function*

$$\phi_{\alpha,b}(t) \equiv E[e^{itZ_{\alpha,b}}] = \begin{cases} \exp\left[-b|t|^{\alpha}\left(1 - i\tan(\frac{\pi\alpha}{2})\text{sgn}(t)\right)\right], & \alpha \neq 1, \\ \exp\left[-b|t|\left(1 + \frac{2i}{\pi}\log|t|\text{sgn}(t)\right)\right], & \alpha = 1. \end{cases} \tag{1.3}$$

(Note that $Z_{2,b}$ *is simply a normal random variable with mean 0 and variance* $2b$*.)*

(i) If $\delta \in (1, 2)$ *then there is some* $b > 0$ *such that*

$$\frac{X_n}{n^{\delta/2}} \to (Z_{\delta/2,b})^{-\delta/2} \text{ in distribution, as } n \to \infty. \tag{1.4}$$

(ii) If $\delta = 2$ *then there exist constants* $a, b > 0$ *and a sequence* $\Gamma(n) \sim an/\log(n)$ *such that*

$$\frac{X_n - \Gamma(n)}{a^2 n/\log^2(n)} \to -Z_{1,b} \text{ in distribution, as } n \to \infty. \tag{1.5}$$

(iii) If $\delta \in (2, 4)$ *then there is some* $b > 0$ *such that*

$$\frac{X_n - vn}{n^{2/\delta}} \to -Z_{\delta/2,b} \text{ in distribution, as } n \to \infty. \tag{1.6}$$

(iv) If $\delta = 4$ *then there is some* $b > 0$ *such that*

$$\frac{X_n - vn}{\sqrt{n\log(n)}} \to Z_{2,b} \text{ in distribution, as } n \to \infty. \tag{1.7}$$

(v) If $\delta > 4$ *then there is some* $b > 0$ *such that*

$$\frac{X_n - vn}{\sqrt{n}} \to Z_{2,b} \text{ in distribution, as } n \to \infty. \tag{1.8}$$

Remark 1.6. Analogous results to Theorem 1.5 hold in the case of negative δ by symmetry.

Remark 1.7. Reference [11] is a review paper and many of the results there were proved earlier in other places (either partially or completely). Case (v) of Theorem 1.5 was originally proven in [10]. Cases (iii) and (iv) were originally proven in [8]. Case (i) under some stronger hypothesis was first shown in [3]. The extension to the (IID), (BD), (ELL) case follows from the same methods used in [8], as noted in that work.

Remark 1.8. If the cookie stacks are unbounded and negative cookies are allowed then δ is not necessarily well defined. However, recent work in [13] and [9] has considered extensions of the parameter δ to certain unbounded environments with both positive and negative cookies. Specifically, in [13] the authors consider deterministic, periodic cookie stacks, which are the same at each site x. Then, in [9] this model is extended to the case where the cookie stack $(\omega(x,i))_{i\in\mathbb{N}}$ at each site x is a finite state Markov chain, started from some distribution η that is the same for all x. In this latter case of Markovian cookie stacks, analogs of Theorems 1.3, 1.4, and 1.5 are all proved for the random walk (X_n), in terms of some parameters δ and $\widetilde{\delta}$, which are generalized versions of the δ given in (1.2) and its negative. This work extends many of the old results from (IID), (BD), and (ELL) environments by removing the (BD) assumption while maintaining the (IID) assumption. Our Assumption (A) presented in the following section will go in a somewhat different direction. We will maintain the (BD) assumption, but weaken the (IID) condition.

1.2 Statement of new results

In light of Theorems 1.3-(i), 1.4, and 1.5 we see that the behavior of the random walk (X_n) is fairly well understood in the case of (IID), (BD), and (ELL) environments. On the other hand, much less is known if the (IID) assumption is weakened to (SE). A zero-one law for directional transience and law of large numbers still hold, but it is generally not well understood when the walk will be transient or ballistic. In the case that the cookies are all positive, Theorem 1.3-(ii) implies that the same threshold for transience as the (IID) case holds, without the boundedness assumption on the cookie stacks. However, even in the case of bounded cookie stacks, it is still unknown whether the same transience/recurrence threshold is always valid for (SE) environments with both positive and negative cookies. Furthermore, there are simple counterexamples (given in Section 1.3) which indicate that the ballisticity threshold of Theorem 1.4 is not, in general, valid for (SE) environments, even if (BD), (ELL), and (POS) are all assumed.

It is thus reasonable to wonder if there are classes of non-(IID) environments for which an explicit characterization of the behavior of the random walk (X_n) is possible, similar to the (IID) case. A natural first step in this direction is to consider Markovian environments, and we will show that in fact Theorems 1.3-(i), 1.4, and 1.5 all extend to a large class of Markovian environments.

Definition 1.9. *Let $\mathcal{S}$ be a countable (either finite or countably infinite) set, and let (S_n) be a discrete time Markov chain on state space $\mathcal{S}$ with transition matrix $\mathcal{K} = \{\mathcal{K}(s,s')\}_{s,s'\in\mathcal{S}}$. The Markov chain (S_n) is said to be* ergodic *if it is irreducible, aperiodic, and positive recurrent. In this case (see [14, Theorem 21.14]), there exists a unique stationary distribution π on $\mathcal{S}$ satisfying $\pi = \pi\mathcal{K}$, and*

$$\lim_{n\to\infty} \|\mathcal{K}^n(s,\cdot) - \pi\|_{TV} = 0, \text{ for each } s \in \mathcal{S}, \tag{1.9}$$

where $\|\mu - \nu\|_{TV} \equiv \frac{1}{2}\|\mu - \nu\|_1$ is the total variational norm between two probability distributions μ and ν on $\mathcal{S}$. The Markov chain (S_n) is said to be uniformly ergodic *if the* *rate of convergence in* (1.9) *is uniform in the initial state. That is, if*

$$\lim_{n\to\infty}\left[\sup_{s\in\mathcal{S}}\|\mathcal{K}^n(s,\cdot)-\pi\|_{TV}\right]=0. \tag{1.10}$$

Of course, all irreducible, aperiodic Markov chains on a finite state space are uniformly ergodic, and many natural positive recurrent Markov chains on a countably infinite state space are also uniformly ergodic. The positive recurrent assumption is clearly necessary to give some asymptotic form of stationarity for the Markov chain (S_n), which, in attempt to partially extend from (IID) environments to (SE) environments, is what we will want to have.

For given $M\in\mathbb{N}$, we denote by $\mathcal{S}_M^*$ the set of all elliptic cookie stacks of height M:

$$\mathcal{S}_M^*=\{s=(s(i))_{i\in\mathbb{N}}: s(i)\in(0,1) \text{ for } i=1,\ldots,M \text{ and } s(i)=1/2 \text{ for } i>M\}.$$

If $\mathcal{S}\subset\mathcal{S}_M^*$ and $(S_k)_{k\in\mathbb{Z}}$ is a stochastic process taking values in $\mathcal{S}$, then we can define a bounded, elliptic cookie environment $\omega=(\omega(k,i))_{k\in\mathbb{Z},i\in\mathbb{N}}$ by

$$\omega(k,i)=S_k(i)\ ,\ k\in\mathbb{Z} \text{ and } i\in\mathbb{N}. \tag{1.11}$$

In the theorems below we will always make the following assumption on our cookie environments.

Assumption (A)
The probability measure $\mathbf{P}$ on cookie environments $\omega\in\Omega$ is the probability measure obtained when ω is defined by (1.11) and the process $(S_k)_{k\in\mathbb{Z}}$ is as follows: $M\in\mathbb{N}$ is a positive integer, $\mathcal{S}\subset\mathcal{S}_M^*$ is a countable set, and $(S_k)_{k\in\mathbb{Z}}$ and $(R_k)_{k\in\mathbb{Z}}$ are both uniformly ergodic Markov chains on the state space $\mathcal{S}$, where $R_k\equiv S_{-k}$, $k\in\mathbb{Z}$.

Remark 1.10. If (S_k) is an ergodic Markov chain on a countable state space $\mathcal{S}$, then the reversed process (R_k) defined by $R_k=S_{-k}$, $k\in\mathbb{Z}$, is also always an ergodic Markov chain. However, if (S_k) is a uniformly ergodic Markov chain, then the reversed process (R_k) is not necessarily uniformly ergodic. We assume explicitly that both the stack sequence (S_k) and its reversal (R_k) are uniformly ergodic Markov chains, so that our Assumption (A) is symmetric with respect to spatial directions of the model. This condition is satisfied in many natural cases, e.g. when the Markov chain (S_k) is uniformly ergodic and reversible, or when the state space $\mathcal{S}$ is finite and (S_k) is irreducible and aperiodic.

With our current methods of proof, this symmetric assumption of bi-directional uniform ergodicity is necessary for a complete extension of the results from the (IID) setting given in Theorems 1.3-(i), 1.4, and 1.5. If one assumes instead only that (S_k) is uniformly ergodic (or only that (R_k) is uniformly ergodic) then the transience/recurrence characterization of Theorem 1.3-(i) still extends fully (with only minor modifications of the proof given in this paper). However, the results of Theorem 1.4 on ballisticity and Theorem 1.5 on limiting distributions do not quite fully extend (only one-sided versions are available, for $\delta>0$ when (R_k) is uniformly ergodic, or for $\delta<0$ when (S_k) is uniformly ergodic).

In the case Assumption (A) is satisfied we will denote the transition matrix for the Markov chain $(S_k)_{k\in\mathbb{Z}}$ by $\mathcal{K}=\{\mathcal{K}(s,s')\}_{s,s'\in\mathcal{S}}$ and the marginal distribution of S_0 by $\phi=(\phi(s))_{s\in\mathcal{S}}$. Together the pair $(\mathcal{K},\phi)$ completely determines the law of the process (S_k), and, hence, the probability measure $\mathbf{P}$. If $\phi=\pi$ is the stationary distribution of the Markov chain, then $(S_k)_{k\in\mathbb{Z}}$ is a stationary and ergodic stochastic process. Thus, the (SE) assumption is satisfied for the environment ω, and the δ defined in (1.2) is, indeed, well defined. If $\phi \neq \pi$, then the process $(S_k)_{k\in\mathbb{Z}}$ is no longer stationary, so the definition (1.2) is no longer directly applicable, because it is not the same for each site x. Nevertheless, asymptotically, for $|k|$ large, S_k has distribution close to π, whatever the distribution ϕ on S_0 is. We, thus, extend the definition of δ as follows when Assumption (A) is satisfied:

$$\delta \equiv \sum_{s\in\mathcal{S}} \pi(s)\cdot\delta(s) \quad \text{where} \quad \delta(s) \equiv \sum_{i=1}^{\infty}\big(2s(i)-1\big) = \sum_{i=1}^{M}\big(2s(i)-1\big). \tag{1.12}$$

Our first theorem gives a threshold for transience vs. recurrence of the random walk (X_n), analogous to Theorem 1.3.

Theorem 1.11. *Assume that Assumption (A) is satisfied for the probability measure* $\boldsymbol{P}$ *on cookie environments, and let* δ *be as in* (1.12). *Then the ERW* $(X_n)_{n\geq 0}$ *is recurrent if* $\delta\in[-1,1]$*, right transient if* $\delta>1$*, and left transient if* $\delta<-1$*.*

Our next theorem gives a threshold for ballisticity of the random walk, analogous to Theorem 1.4.

Theorem 1.12. *Assume that Assumption (A) is satisfied for the probability measure* $\boldsymbol{P}$ *on cookie environments, and let* δ *be as in* (1.12). *Then there exists a deterministic* $v\in[-1,1]$ *such that* $X_n/n\to v$*,* P_0 *a.s. Moreover,* $v=0$ *if* $\delta\in[-2,2]$*,* $v>0$ *if* $\delta>2$*, and* $v<0$ *if* $\delta<-2$*.*

Our final theorem characterizes the limiting distribution of (X_n), analogous to Theorem 1.5.

Theorem 1.13. *Assume that Assumption (A) is satisfied for the probability measure* $\boldsymbol{P}$ *on cookie environments, and let* δ *be as in* (1.12). *Also, let* v *be the velocity of the random walk* (X_n) *as in Theorem 1.12. Then (i)-(v) of Theorem 1.5 all hold.*

Remark 1.14. Just as for Theorem 1.5 in the (IID) case, analogous results to Theorem 1.13 also hold in the case of negative δ. This follows from the fact that Assumption (A) is preserved when spatial directions of the model are interchanged (see Remark 3.4 below).

Remark 1.15. In fact, Assumption (A) in Theorems 1.11-1.13 can be weakened to the somewhat more general *hidden Markov* Assumption (B) given below. The proofs are essentially unchanged; it is only for convenience of notation and language that our proofs are written using Assumption (A) instead of (B).

Assumption (B)
The probability measure $\mathbf{P}$ on cookie environments $\omega\in\Omega$ is the probability measure obtained when ω is defined by (1.11) and the process $(S_k)_{k\in\mathbb{Z}}$ is as follows: $(Z_k)_{k\in\mathbb{Z}}$ is a uniformly ergodic Markov chain on a countable state space $\mathcal{Z}$ such that the reversed Markov chain $(Z_{-k})_{k\in\mathbb{Z}}$ is also uniformly ergodic, $M\in\mathbb{N}$ is a positive integer, $\mathcal{S}\subset\mathcal{S}_M^*$ is a countable set, and $S_k=f(Z_k)$ where $f:\mathcal{Z}\to\mathcal{S}$ is some observation function (f not necessarily 1-1).

1.3 Counterexamples

In this section we present three closely related counterexamples, which indicate that the threshold for positivity of the speed given in Theorem 1.4 (and also as a consequence the limiting distributions presented in Theorem 1.5 when $\delta>2$), do not extend as far as one might hope from the (IID), (BD), (ELL) case. The first example shows that the threshold for ballisticity does not in general hold for (POS) and (SE) environments. The second, which is a modification of the first, shows that the same conclusion is true if one adds (BD) and (ELL) to the (POS) and (SE) assumptions. These examples are known, see e.g. [11, Example 5.7] (and also [15, page 290] for an earlier and somewhat related example). The final example, which we present here for the first time, is a modification of the second. It shows that if one considers Markovian environments as in Assumption (A) with bounded and elliptic cookie stacks, but assumes only that the Markov chain $(S_k)_{k\in\mathbb{Z}}$ is ergodic (rather than uniformly ergodic), then again the threshold of Theorem 1.4 for ballisticity is not in general valid[1]. Thus, our uniformly ergodic assumption on the Markov chains (S_k) and (R_k), or at least some assumption beyond ergodicity, is necessary for the (IID) results to translate completely.

The following definition and lemma on monotonicity of the speed will be needed for our examples.

Definition 1.16. *If $\omega_1, \omega_2 \in \Omega$ are two cookie environments, we say that ω_1* dominates *ω_2 if $\omega_1(x,i) \geq \omega_2(x,i)$, for all $x \in \mathbb{Z}$ and $i \in \mathbb{N}$. If $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are two probability measures on Ω, we say that $\boldsymbol{P}_1$* dominates *$\boldsymbol{P}_2$ if there exists some joint probability measure $\overline{\boldsymbol{P}}$ on $\Omega \times \Omega$ with marginals $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ such that $\overline{\boldsymbol{P}}\big(\{(\omega_1,\omega_2) \in \Omega^2 : \omega_1 \text{ dominates } \omega_2\}\big) = 1$.*

Lemma 1.17 (Special case of Proposition 4.2 from [11])**.** *Let $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ be two (SE) and (ELL) probability measures on Ω such that $\boldsymbol{P}_1$* dominates *$\boldsymbol{P}_2$, and let v_1 and v_2 be the corresponding velocities of the associated ERWs (as in Theorem 1.2). Then $v_1 \geq v_2$.*

Example 1. Define cookie stacks s_0 and s_1 by $s_0(i) = 1$ for all $i \in \mathbb{N}$, $s_1(i) = 1/2$ for all $i \in \mathbb{N}$. That is, s_0 is an infinite stack of completely right biased cookies, and s_1 is an infinite stack of "placebo cookies" which induce no bias on the random walker. Let $\mathcal{S} = \{s_0, s_1\}$, and let $(S_k)_{k\in\mathbb{Z}}$ be a stationary and ergodic process taking values in $\mathcal{S}$ such that the intervals between consecutive occurrences of s_0 in the process (S_k) are i.i.d. Define the random environment $\omega \in \Omega$ by (1.11). Also, define $\tau_1 = \inf\{k > 0 : S_k = s_0\}$ and $\tau_{i+1} = \inf\{k > \tau_i : S_k = s_0\}$, for $i \geq 1$, and let $T_i = \inf\{n > 0 : X_n = \tau_i\}$. That is, T_i is the time at which the walker first reaches the i-th position to the right of 0 which has an s_0 stack. Assume that the distribution of the random variable $(\tau_2 - \tau_1)$, i.e. the distribution of the distance between consecutive occurrences of stack s_0 in the process (S_k), is such that $E(\tau_2 - \tau_1) < \infty$, but $E\left((\tau_2 - \tau_1)^2\right) = \infty$. Then the following all hold.

1. $\delta = \infty > 2$.
2. $E(T_2 - T_1) = \infty$, since $E(T_2 - T_1 | \tau_2 - \tau_1 = \ell) = \ell^2$.
3. $v \equiv \lim_{n\to\infty} X_n/n = \lim_{i\to\infty} \tau_i/T_i = \dfrac{E(\tau_2 - \tau_1)}{E(T_2 - T_1)} = 0$.

Example 2. Let $M \in \mathbb{N}$ and $p \in (1/2, 1)$, and modify Example 1 so that the stack s_0 is defined as follows: $s_0(i) = p$ for $i = 1, \ldots, M$ and $s_0(i) = 1/2$ for $i > M$. Then the natural coupling between environments shows that the probability measure $\mathbf{P}_1$ on cookie environments from Example 1 dominates the new probability measure $\mathbf{P}_2$. Thus, by Lemma 1.17, $v_2 \leq v_1 = 0$, where v_1 and v_2 are the associated velocities of the random walks. This construction works for any $M \in \mathbb{N}$, and so we may choose M sufficiently large that $\delta = M(2p-1)/E(\tau_2 - \tau_1) > 2$ in the new modified case. Then we have a (SE), (POS), (BD), (ELL) probability measure on cookie environments with velocity 0, but $\delta > 2$.

Example 3. Let s_0, s_1 and the process $(S_k)_{k\in\mathbb{Z}}$ be as in Example 2. Assume that M is sufficiently large that $\frac{M(2p-1)}{E(\tau_2-\tau_1)} > 3$, and that the distribution of $(\tau_2 - \tau_1)$ has support on all of $\mathbb{N}$. Let $\tilde{\mathcal{S}} = \{\tilde{s}_j : j \geq 0\}$ where the stacks $\tilde{s}_j$ are defined by $\tilde{s}_0 = s_0$ and, for $j \geq 1$,

$$\tilde{s}_j(1) = 1/2 - \frac{1}{3j} \quad \text{and} \quad \tilde{s}_j(i) = 1/2,\ i > 1.$$

[1] In fact, although this is not explicit in their descriptions, the first two examples are of the hidden Markov type as in Assumption (B), but where the underlying Markov chain $(Z_k)_{k\geq 0}$ is not uniformly ergodic.

Now, define a new process $(\widetilde{S}_k)_{k\in\mathbb{Z}}$ from the process $(S_k)_{k\in\mathbb{Z}}$ by the following projection:

$$\widetilde{S}_k = \tilde{s}_j, \text{ where } j = k - \ell \text{ and } \ell = \sup\{m \le k : S_m = s_0\}.$$

Thus, $\widetilde{S}_k$ is equal to $\tilde{s}_0$ if $S_k = s_0$, and otherwise $\widetilde{S}_k$ equals $\tilde{s}_j$ where j is the number of time steps since the last occurrence of s_0 for the process $(S_m)_{m\in\mathbb{Z}}$. With this construction $(\widetilde{S}_k)_{k\in\mathbb{Z}}$ is a Markov chain with transition probabilities

$$P(\widetilde{S}_{k+1} = \tilde{s}_0 | \widetilde{S}_k = \tilde{s}_j) = P(\tau_2 - \tau_1 = j+1 | \tau_2 - \tau_1 > j),$$
$$P(\widetilde{S}_{k+1} = \tilde{s}_{j+1} | \widetilde{S}_k = \tilde{s}_j) = P(\tau_2 - \tau_1 > j+1 | \tau_2 - \tau_1 > j).$$

Moreover, the state space $\tilde{\mathcal{S}}$ of this Markov chain has only bounded, elliptic stacks, and the Markov chain $(\widetilde{S}_k)$ itself is stationary (since (S_k) is), irreducible and aperiodic (since $(\tau_2 - \tau_1)$ has support on all of $\mathbb{N}$), and positive recurrent (since $E(\tau_2 - \tau_1) < \infty$). Finally, for the probability measure $\mathbf{P}_3$ on environments ω constructed from the $(\widetilde{S}_k)$ process and associated parameter δ the following hold:

1. $\mathbf{P}_3$ is dominated by the probability measure $\mathbf{P}_2$ from Example 2. Hence, by Lemma 1.17, $v_3 \le v_2 = 0$, where v_3 and v_2 are associated velocities.

2. Let $\delta(\tilde{s}_j) = \sum_{i=1}^{\infty}(2\tilde{s}_j(i) - 1)$ be the net drift in stack $\tilde{s}_j$, and let $\pi = (\pi_j)_{j\ge 0}$ be the stationary distribution of the Markov chain $(\widetilde{S}_k)$. Then $\delta(\tilde{s}_0) = M(2p-1)$ and $-2/3 = \delta(\tilde{s}_1) < \delta(\tilde{s}_2) < \delta(\tilde{s}_3) < \ldots$, so

$$\begin{aligned}\delta &= \sum_{j=0}^{\infty} \pi_j \delta(\tilde{s}_j) > \pi_0 \delta(\tilde{s}_0) + \sum_{j=1}^{\infty} \pi_j \delta(\tilde{s}_1) \\ &= \frac{1}{E(\tau_2 - \tau_1)} M(2p-1) + \left[1 - \frac{1}{E(\tau_2 - \tau_1)}\right](-2/3) > 2.\end{aligned}$$

In summary, the stack sequence $(\widetilde{S}_k)_{k\in\mathbb{Z}}$ (hence also the reversed sequence $(\widetilde{R}_k)_{k\in\mathbb{Z}}$) are stationary and ergodic Markov chains, and $\delta > 2$, but the velocity $v_3 \le 0$. So, the ballisticity threshold of Theorem 1.12 does not extend to the case when the sequence of cookies stacks is simply an ergodic Markov chain rather than a uniformly ergodic one. Indeed, to the best of our knowledge, it is not even known for this example whether the walk is transient or recurrent since the environment does not satisfy the (POS) condition (and, thus, Theorem 1.3-(ii) is not applicable).

1.4 Outline of paper

An outline of the remainder of the paper is as follows. In Section 1.5 we introduce some basic notation and conventions that will be used throughout. In Section 2 we review a well known connection between ERW and certain branching processes, called the *forward branching process* and *backward branching process*. We also introduce some related processes, which are easier to analyze, and give some concentration estimates and expectation and variance calculations for these related processes. In Section 3 we prove Theorem 1.11. The proof is based on the connection between the ERW and forward branching process, and follows the general approach used in [13]. In Section 4 we prove Theorems 1.12 and 1.13. The proofs are based on a connection between the ERW and backward branching process and follow the general approach used in [8], [9]. Central to these arguments is a diffusion approximation limit for the backward branching process introduced in [8]. The proofs of several technical results are deferred to the appendices.

1.5 Notation

The positive integers are denoted by $\mathbb{N}$ (as above), and the non-negative integers by $\mathbb{N}_0$. The infimum of the empty set is defined to be ∞, and $\sum_{i=j}^{k} z_i \equiv 0$, for any $j > k$ and sequence (z_i). For a stochastic process $Z = (Z_n)_{n\geq 0}$, $\tau_x^Z \equiv \inf\{n > 0 : Z_n = x\}$. The same notation is also used for a continuous time process $(Z(t))_{t\geq 0}$ with continuous sample paths. For sequences of real numbers $(a_n), (b_n)$ we write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$. Similarly, $f(x) \sim g(x)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$, for real valued functions f and g. Constants of the form c_i are assumed to carry over between the various propositions and lemmas throughout. Other constants C, c, C_i, K_i, ... etc. are particular only to the specific lemma or proposition where they are introduced.

Unless otherwise specified it is assumed in the remainder of the paper that Assumption (A) holds for the probability measure $\mathbf{P}$ on cookie environments. The stack height M, state set $\mathcal{S} \subset \mathcal{S}_M^*$, and transition matrix $\mathcal{K}$ for the Markov chain $(S_k)_{k\in\mathbb{Z}}$ are all assumed to be fixed. The marginal distribution of S_0 (according to $\mathbf{P}$) is denoted by ϕ, and the stationary distribution of the Markov chain (S_k) is denoted by π. Also, δ is given by (1.12). By a slight abuse of notation, we will use $\mathbf{P}$ as the probability measure for the Markov chain (S_k) itself, as well as the environment ω derived from it according to (1.11). The probability measures $\mathbf{P}_s$, $s \in \mathcal{S}$, and $\mathbf{P}_\pi$ are the modified probability measures for the Markov chain (S_k) (or equivalently for the environment ω) when $S_0 = s$ or $S_0 \sim \pi$:

$$\mathbf{P}_s(\cdot) \equiv \mathbf{P}(\cdot|S_0 = s) \quad \text{and} \quad \mathbf{P}_\pi(\cdot) \equiv \sum_{s\in\mathcal{S}} \pi(s)\mathbf{P}_s(\cdot).$$

Expectations with respect to $\mathbf{P}_s$ and $\mathbf{P}_\pi$ are denoted by $\mathbf{E}_s$ and $\mathbf{E}_\pi$, respectively, and the corresponding averaged measures for the random walk (X_n) started from position x are

$$P_{x,s}(\cdot) \equiv \mathbf{E}_s[P_x^\omega(\cdot)] \quad \text{and} \quad P_{x,\pi}(\cdot) \equiv \mathbf{E}_\pi[P_x^\omega(\cdot)].$$

The probability measure $\mathbb{P}$ and corresponding expectation operator $\mathbb{E}$ will be used generically for auxiliary random variables living on outside probability spaces, separate from those of the environment ω and random walk (X_n).

2 Branching processes

In this section we introduce our main tool in the analysis of the ERW, which is a connection with two related branching processes known as the forward branching process and backward branching process. The definition of the forward branching process is given in Section 2.1, and the definition of the backward branching process in Section 2.2. Some related branching processes which are easier to analyze are introduced in Section 2.3. Various concentration estimates and expectation and variance calculations for some of the related branching processes are given in Section 2.4.

2.1 The forward branching process

The construction of both the forward and backward branching processes is based on the *coin tossing construction* of the ERW introduced in [10]. For a fixed environment $\omega \in \Omega$, we initially flip an infinite sequence of coins at each site k, where the i-th coin at site k has probability $\omega(k, i)$ of landing heads. The walker begins its walk at some given site x, and if it ever reaches site k for the i-th time, then it jumps right if the i-th coin toss at site k was heads and left otherwise. More formally, let ξ_i^k, $k \in \mathbb{Z}$ and $i \in \mathbb{N}$, be independent random variables such that ξ_i^k has Bernoulli distribution with parameter $p = \omega(k, i)$. Then the random walk $(X_n)_{n\geq 0}$ started from position x in the given environment ω can be constructed from the ξ_i^k's as follows:

$$X_0 = x \;\text{ and }\; X_{n+1} = (2\xi_{I_n}^{X_n} - 1) + X_n \tag{2.1}$$

where $I_n = |\{0 \le m \le n : X_m = X_n\}|$. We will say that ξ_i^k is a *success* if $\xi_i^k = 1$ (i.e. heads) and a failure if $\xi_i^k = 0$ (i.e. tails). The *forward branching process* $(U_k)_{k\ge 0}$ started from level $u_0 \in \mathbb{N}_0$ is defined by

$$U_0 = u_0 \;\text{ and }\; U_{k+1} = \inf\left\{m : \sum_{i=1}^{m} \mathbb{1}\{\xi_i^{k+1} = 0\} = U_k\right\} - U_k. \tag{2.2}$$

That is, U_{k+1} is the number of successes in the sequence $(\xi_i^{k+1})_{i\in\mathbb{N}}$ before the U_k-th failure[2]. If we define G_i^k to be the number of successes in the sequence $(\xi_j^k)_{j\in\mathbb{N}}$ between the $(i-1)$-th and i-th failures then we have, for each $k \ge 0$,

$$U_{k+1} = \sum_{i=1}^{U_k} G_i^{k+1}, \text{ where } (G_i^k)_{i>M,k\ge 1} \text{ are i.i.d. Geo(1/2) random variables.} \tag{2.3}$$

Thus, the process $(U_k)_{k\ge 0}$ may be seen as a type of branching process with a time dependent migration term. More precisely, the k-th step of the process may be interpreted as a combination of the following 3 things:

- First, $U_k \wedge M$ individuals emigrate out of the population before reproducing.
- Then, all remaining individuals (if any) have a Geo(1/2) number of offspring independently.
- Finally, $\sum_{i=1}^{U_k\wedge M} G_i^{k+1}$ individuals immigrate into the population after reproduction.

Now, this construction for the process $U = (U_k)_{k\ge 0}$ has been for a fixed environment ω, but one can also consider the same process when the environment ω is first chosen randomly according to some probability measure. We will denote by $P_{u_0}^{U,\omega}$ the probability measure for the process U started from level u_0 in a fixed environment ω, as constructed above, and by $P_{u_0,s}^U$ the probability measure for the joint process $(U_k, S_k)_{k\ge 0}$ when $U_0 = u_0$ and $S_0 = s$. That is, we first sample $(S_k)_{k\in\mathbb{Z}}$ according to $\mathbf{P}_s$ to get an environment $\omega = (\omega(k,i))_{k\in\mathbb{Z},i\in\mathbb{N}} = (S_k(i))_{k\in\mathbb{Z},i\in\mathbb{N}}$, and then we sample $(U_k)_{k\ge 0}$ according to $P_{u_0}^{U,\omega}$. This two step procedure gives a joint measure on[3] $(U_k, S_k)_{k\ge 0}$ which is the measure $P_{u_0,s}^U$. The measures $P_{u_0,\pi}^U$ and $P_{u_0}^U$ are the averaged measures when S_0 is distributed according to π or ϕ, respectively:

$$P_{u_0,\pi}^U(\cdot) \equiv \sum_{s\in\mathcal{S}} \pi(s) P_{u_0,s}^U(\cdot) \quad \text{ and } \quad P_{u_0}^U(\cdot) \equiv \sum_{s\in\mathcal{S}} \phi(s) P_{u_0,s}^U(\cdot).$$

Under any of these measures $P_{u_0,s}^U$, $P_{u_0,\pi}^U$, and $P_{u_0}^U$ the joint process $(U_k, S_k)_{k\ge 0}$ is a time-homogeneous Markov chain with transition probabilities $p_{(u,r)(u',r')}^U \equiv \text{Prob}(U_{k+1} = u', S_{k+1} = r' | U_k = u, S_k = r)$ given by

$$p_{(u,r)(u',r')}^U = \mathcal{K}(r,r') P_u^{U,\omega_{r'}}(U_1 = u'), \tag{2.4}$$

where $\mathcal{K}$ is the transition matrix for the Markov chain (S_k) and $\omega_{r'}$ is the deterministic environment with stack r' at each site: $\omega_{r'}(x,i) = r'(i)$, for all $x \in \mathbb{Z}$ and $i \in \mathbb{N}$.

The main interest in the forward branching process is its connection to a related process $(U_k')_{k\ge 1}$ defined by

$$U_k' = |\{0 \le n < \tau_0^X : X_n = k, X_{n+1} = k+1\}|.$$

Clearly, survival of the process (U_k'), i.e. occurrence of the event $\{U_k' > 0, \forall k > 0\}$, is closely related to right transience of the random walk (X_n). The following lemma is standard, but we will provide a proof for the convenience of the reader.

[2] In the case $U_k = \infty$ (when (2.2) is no longer directly meaningful) we will extend this interpretation, so that U_{k+1} is then defined to be the total number of successes in the sequence $(\xi_i^{k+1})_{i\in\mathbb{N}}$.

[3] Note that the process $(U_k)_{k\ge 0}$ depends only on $S_k(i) = \omega(k,i)$, for $i, k \ge 1$. So, we do not need to completely specify ω to construct $(U_k)_{k\ge 0}$. It is sufficient to consider $(S_k)_{k\ge 0}$.

Lemma 2.1. *Assume the process* $(U_k)_{k\geq 0}$ *is started from level* $u_0 = 1$ *and that* $(X_n)_{n\geq 0}$ *is started from position* $X_0 = 1$*. Then, for any realization of the random variables* $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$*,* $U_k' \leq U_k$ *for all* $k \geq 1$*. Moreover, for any realization of the random variables* $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ *such that* $\tau_0^X < \infty$*,* $U_k' = U_k$ *for all* $k \geq 1$*.*

Note: The lemma does not specify anything about the probability measure on the environment ω. The relation between U_k and U_k' is a deterministic function of the values of $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$, from which both processes (U_k) and (U_k') are constructed.

Proof. First fix a realization $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ such that $\tau_0^X < \infty$. By definition U_1' is the number of right jumps from site 1 before time τ_0^X, which (if $\tau_0^X < \infty$) is simply the number of right jumps from site 1 before the first left jump from this site, or equivalently the number of successes in the sequence $(\xi_i^1)_{i\in\mathbb{N}}$ before the first failure. Since we assume $u_0 = 1$, the latter quantity is exactly U_1, so we have $U_1' = U_1$. Now suppose that $U_k = U_k' = m \geq 0$, for some $k \geq 1$. Then, by the definition of the (U_j') process, the random walk (X_n) must jump right from site k exactly m times prior to time τ_0^X. Thus, the walk must jump left from site $k+1$ exactly m times prior to time τ_0^X. Thus, the number of right jumps from site $k+1$ prior to time τ_0^X is exactly the number of right jumps from site $k+1$ before there are m left jumps from it, or equivalently the number of successes in the sequence $(\xi_i^{k+1})_{i\in\mathbb{N}}$ before the m-th failure. Since $U_k = m$, this shows that $U_{k+1}' = U_{k+1}$. It follows, by induction, that $U_k' = U_k$ for all $k \geq 1$.

Now, fix a realization $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ such that $\tau_0^X = \infty$. Then, for each $k \geq 1$, U_k' is simply the total number of right jumps from site k by the random walk (X_n). Because the walk never jumps left from site 1, the total number of right jumps from site 1 is at most the number of successes in the sequence $(\xi_i^1)_{i\in\mathbb{N}}$ before the first failure. So, we have $U_1' \leq U_1$. Now suppose that $U_k = \ell$ and $U_k' = m$ for some $0 \leq m \leq \ell \leq \infty$ and $k \geq 1$. If $m = 0$, then the walk never jumps right from site k, so it never reaches site $k+1$, so $U_{k+1}' = 0$, so $U_{k+1}' \leq U_{k+1}$. If $1 \leq m < \infty$, then the walk jumps right from site k exactly m times total, so either it jumps left from site $k+1$ exactly $(m-1)$ times or it jumps left from site $k+1$ exactly m times and never returns to site $k+1$ after the m-th left jump. Either way, the total number of right jumps from site $k+1$ can be at most the number of successes in $(\xi_i^{k+1})_{i\in\mathbb{N}}$ before the m-th failure, which is at most the number of successes in $(\xi_i^{k+1})_{i\in\mathbb{N}}$ before the ℓ-th failure. Hence, again, $U_{k+1}' \leq U_{k+1}$. Finally, if $m = \infty$ then the walk must jump right from site k infinitely many times, so it must jump left from site $k+1$ infinitely many times, so it must visit site $k+1$ infinitely often, so the total number of right jumps from site $k+1$ is simply the total number of successes in the sequence $(\xi_i^{k+1})_{i\in\mathbb{N}}$, which is equal to U_{k+1} (since $\ell = \infty$, with $m = \infty$ and $\ell \geq m$). Thus, in all possible cases $U_{k+1}' \leq U_{k+1}$, so it follows, by induction, that $U_k' \leq U_k$ for all $k \geq 1$. □

In Appendix A we will prove the following basic fact using a finite modification argument.

Lemma 2.2. *Define* $A^+ = \{X_n > 0, \forall n > 0 \text{ and } \lim_{n\to\infty} X_n = +\infty\}$*. Then, for each* $x \in \mathbb{N}$ *and* $s \in \mathcal{S}$*,* $P_{x,s}(A^+) > 0$ *if and only if* $P_{x,s}(X_n \to +\infty) > 0$*.*

Using this fact along with Lemma 2.1 and Theorem 1.1, we now establish an explicit criteria relating transience/recurrence of the random walk (X_n) to the forward branching process (U_k). This criteria will be used to prove Theorem 1.11 in Section 3. For the statement of the lemma recall that $P_{u_0,s}^U(\cdot)$ is the joint probability measure for $(U_k, S_k)_{k\geq 0}$ when $U_0 = u_0$ and $S_0 = s$.

Lemma 2.3. *The following hold:*

$$\text{If } \exists s \in \mathcal{S} \text{ such that } P_{1,s}^U(U_k > 0, \forall k > 0) > 0, \text{ then } P_0(X_n \to +\infty) = 1.$$
$$\text{If } \exists s \in \mathcal{S} \text{ such that } P_{1,s}^U(U_k > 0, \forall k > 0) = 0, \text{ then } P_0(X_n \to +\infty) = 0.$$

Proof. We will say that "U_k survives" if $U_k > 0$ for all $k > 0$, and similarly for U'_k. Also, we extend the probability measure $P_{x,s}$ for the random walk (X_n) to the process (U'_k) derived from it. By Theorem 1.1, $P_{1,\pi}(X_n \to +\infty) = P_{0,\pi}(X_n \to +\infty) \in \{0,1\}$. Since $\pi(s) > 0$ for all $s \in \mathcal{S}$, this implies that either

(a) $P_{1,s}(X_n \to +\infty) = P_{0,s}(X_n \to +\infty) = 1$, for all $s \in \mathcal{S}$, or

(b) $P_{1,s}(X_n \to +\infty) = P_{0,s}(X_n \to +\infty) = 0$, for all $s \in \mathcal{S}$.

We consider these two cases separately.

Case (a):
In this case, it follows from Lemma 2.2 that $P_{1,s}(A^+) > 0$, for each $s \in \mathcal{S}$. Hence, $P_{1,s}(U'_k \text{ survives}) > 0$, for each $s \in \mathcal{S}$. By Lemma 2.1 this implies $P^U_{1,s}(U_k \text{ survives}) > 0$, for each $s \in \mathcal{S}$.

Case (b):
Since $\omega(x,i) = 1/2$ for each $i > M$ and $x \in \mathbb{Z}$, $\mathbf{P}_s$ a.s., we have $P_{1,s}(\liminf_{n\to\infty} X_n = x) = 0$, for each $x \in \mathbb{Z}$. Thus, if $P_{1,s}(X_n \to +\infty) = 0$ for each $s \in \mathcal{S}$, then $P_{1,s}(\liminf_{n\to\infty} X_n = -\infty) = 1$ for each $s \in \mathcal{S}$, and in particular, $P_{1,s}(\tau_0^X < \infty) = 1$ for each $s \in \mathcal{S}$. By Lemma 2.1 and the definition of the (U'_k) process this implies $P^U_{1,s}(U_k \text{ survives}) = 0$ for each $s \in \mathcal{S}$.

Thus, we have established the following dichotomy: Either

(a') $P^U_{1,s}(U_k \text{ survives}) > 0$ for each $s \in \mathcal{S}$ and (a) holds, or

(b') $P^U_{1,s}(U_k \text{ survives}) = 0$ for each $s \in \mathcal{S}$ and (b) holds.

Since $P_0(X_n \to +\infty) = 1$ if (a) holds, and $P_0(X_n \to +\infty) = 0$ if (b) holds, this establishes the lemma. □

2.2 The backward branching process

Let the random variables $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ be as above in Section 2.1. We continue to assume the random walk (X_n) is constructed from these random variables using (2.1). Also, we recall that $(R_k)_{k\in\mathbb{Z}}$ is the spatial reversal of the stack sequence $(S_k)_{k\in\mathbb{Z}}$:

$$R_k = S_{-k}\ ,\ k \in \mathbb{Z}. \tag{2.5}$$

The *backward branching process* $(V_k)_{k\geq 0}$ started from level $v_0 \in \mathbb{N}_0$ is defined by

$$V_0 = v_0 \ \text{ and } \ V_{k+1} = \inf\Big\{m : \sum_{i=1}^{m} \mathbb{1}\{\xi_i^{-(k+1)} = 1\} = V_k + 1\Big\} - (V_k+1). \tag{2.6}$$

That is, V_{k+1} is the number of failures in the sequence $(\xi_i^{-(k+1)})_{i\in\mathbb{N}}$ (i.e. at stack R_{k+1}) before there are $V_k + 1$ successes. If we let H_i^k be the number of failures in the sequence $(\xi_j^{-k})_{j\in\mathbb{N}}$ between the $(i-1)$-th and i-th successes then

$$V_{k+1} = \sum_{i=1}^{V_k+1} H_i^{k+1}, \text{ where } (H_i^k)_{i>M,k\geq 1} \text{ are i.i.d. Geo(1/2) random variables.} \tag{2.7}$$

Thus, by similar reasoning as for the forward branching process $(U_k)_{k\geq 0}$, this process $(V_k)_{k\geq 0}$ may also be seen as a type of branching process with a time dependent migration term.

In fact, the definition of the backward branching process (V_k) is almost exactly symmetric to the definition of the forward branching process (U_k) with successes replaced by failures and S_k replaced by R_k, but there is one notable difference: In the backward process we count failures until V_k+1 successes, whereas in the forward process we only count successes until U_k failures. This "+1" is important, because it means that 0 is not an absorbing state for the process (V_k), as it is for the process (U_k).

Our interest in the backward branching process stems from the following lemma about down crossings. The analog in the case of (IID) environments is well known.

Lemma 2.4. *Assume that $\delta > 1$ and $X_0 = V_0 = 0$. For $n \in \mathbb{N}$ and $k \le n$, let*

$$D_{n,k} = |\{0 \le m < \tau_n^X : X_m = k, X_{m+1} = k-1\}|$$

be the number of down crossings of the edge $(k, k-1)$ by the random walk (X_m) up to time τ_n^X. Then $(V_0, V_1, \ldots, V_n)$ and $(D_{n,n}, D_{n,n-1}, \ldots, D_{n,0})$ have the same distribution if the environment ω is chosen according to the stationary measure $\mathbf{P}_\pi$.

Proof. Since we assume that $\delta > 1$, it follows from Theorem 1.11 that τ_n^X is $P_{0,\pi}$ a.s. finite[4] for each n. Fix $n \in \mathbb{N}$ and any realization of the random variables $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ such that τ_n^X is finite. Then, define $\widetilde{\xi}_i^k = \xi_i^{k+n}$, for $k \in \mathbb{Z}$ and $i \in \mathbb{N}$. Let $D_{n,k}$, $0 \le k \le n$, be as in the statement of the lemma when the random walk (X_m) is generated according to the specific fixed values of $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$, and let $(\widetilde{V}_k)_{k\ge 0}$ be defined as in (2.6) with $v_0 = 0$, but $\xi_i^{-(k+1)}$ replaced with $\widetilde{\xi}_i^{-(k+1)}$. We claim that, in this case, $D_{n,n-k} = \widetilde{V}_k$, for each $0 \le k \le n$.

The proof is by induction on k. For $k = 0$ we have $\widetilde{V}_0 = 0$, by assumption, and $D_{n,n} = 0$, since the walk cannot down cross the edge $(n, n-1)$ before first hitting site n. Now assume $D_{n,n-k} = \widetilde{V}_k = \ell$ for some $0 \le k < n$ and $\ell \ge 0$. Then the walk (X_m) must jump left from site $n-k$ exactly ℓ times prior to time τ_n^X. Thus, the walk must jump right from site $n-(k+1)$ exactly $\ell+1$ times prior to τ_n^X. Thus, the number of left jumps from site $n-(k+1)$ prior to time τ_n^X is exactly the number of left jumps from site $n-(k+1)$ before the $(\ell+1)-th$ right jump. So, we have:

$$\begin{aligned} D_{n,n-(k+1)} &= \text{\# left jumps from site } n-(k+1) \text{ prior to time } \tau_n^X \\ &= \text{\# left jumps from site } n-(k+1) \text{ before } (\ell+1)\text{-th right jump} \\ &= \text{\# failures in } (\xi_i^{n-(k+1)})_{i\in\mathbb{N}} \text{ before } (\ell+1)\text{-th success} \\ &= \text{\# failures in } (\widetilde{\xi}_i^{-(k+1)})_{i\in\mathbb{N}} \text{ before } (\ell+1)\text{-th success} \\ &= \text{\# failures in } (\widetilde{\xi}_i^{-(k+1)})_{i\in\mathbb{N}} \text{ before } (\widetilde{V}_k+1)\text{-th success} \\ &= \widetilde{V}_{k+1}. \end{aligned}$$

This completes the proof that $D_{n,n-k} = \widetilde{V}_k$, for each $0 \le k \le n$, using the specific fixed values of $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ and associated values of $\widetilde{\xi}_i^k = \xi_i^{k+n}$. The lemma now follows since $(S_k)_{k\in\mathbb{Z}}$ is stationary under $\mathbf{P}_\pi$, so the stochastic process $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ has the same distribution as the process $(\xi_i^{k+n})_{k\in\mathbb{Z},i\in\mathbb{N}}$. □

The importance of the down crossings is their relation to hitting times for the random walk (X_n). If $X_0 = 0$, then for each $n \in \mathbb{N}$

$$\tau_n^X = n + 2\sum_{k\le n} D_{n,k} = n + 2\sum_{k=0}^{n} D_{n,k} + 2\sum_{k<0} D_{n,k}. \tag{2.8}$$

[4] Theorem 1.11 will not be proved till later in Section 3, but the proof uses only the forward branching process described above, and is thus independent of the development in this section. The theorem is stated when ω is chosen according to $\mathbf{P} \equiv \mathbf{P}_\phi$, rather than $\mathbf{P}_\pi$, but ϕ is allowed to be any initial distribution on $\mathcal{S}$ in the theorem. So, in particular, the conclusion is valid when $\phi = \pi$.

If $\delta > 1$, so that the random walk is right transient, then $\lim_{n\to\infty} \sum_{k<0} D_{n,k}$ is a.s. finite, and thus the asymptotic distribution of τ_n^X is determined by the asymptotic distribution of $\sum_{k=0}^n D_{n,k}$. By Lemma 2.4, the latter sum has the same distribution as $\sum_{k=0}^n V_k$ (assuming that $\phi = \pi$). The general proof strategy for Theorems 1.12 and 1.13 is to analyze asymptotic properties of $\sum_{k=0}^n V_k$, relate these to asymptotic properties of the hitting times τ_n^X, and then relate those to asymptotic properties of the random walk (X_n) itself. This basic approach has been employed many times before in the study of excited random walks, e.g. [10], [3], [8], [9]. (See also [7], [20], [21], and [18] for similar uses of branching processes in analyzing one dimensional self-interacting random walks and random walk in random environment.)

In the sequel we will use the following notation for the backward branching process $V = (V_k)_{k\geq 0}$, similar to that for the forward branching process $(U_k)_{k\geq 0}$. $P_{v_0}^{V,\omega}$ is the probability measure for $(V_k)_{k\geq 0}$ started from level v_0 in a fixed environment ω, and, for $s \in \mathcal{S}$, $P_{v_0,s}^V$ is the probability measure for the joint process $(V_k, R_k)_{k\geq 0}$ when $V_0 = v_0$ and $R_0 = s$. The measures $P_{v_0,\pi}^V$ and $P_{v_0}^V$ are defined by

$$P_{v_0,\pi}^V(\cdot) \equiv \sum_{s\in\mathcal{S}} \pi(s) P_{v_0,s}^V(\cdot) \quad \text{ and } \quad P_{v_0}^V(\cdot) \equiv \sum_{s\in\mathcal{S}} \phi(s) P_{v_0,s}^V(\cdot).$$

Under any of these measures $P_{v_0,s}^V$, $P_{v_0,\pi}^V$, and $P_{v_0}^V$ the joint process $(V_k, R_k)_{k\geq 0}$ is a time-homogeneous Markov chain with transition probabilities $p_{(v,r)(v',r')}^V \equiv \text{Prob}(V_{k+1} = v', R_{k+1} = r' | V_k = v, R_k = r)$ given by

$$p_{(v,r)(v',r')}^V = \widetilde{\mathcal{K}}(r,r') P_v^{V,\omega_{r'}}(V_1 = v'), \tag{2.9}$$

where $\widetilde{\mathcal{K}}$ is the transition matrix for the Markov chain (R_k) given by $\widetilde{\mathcal{K}}(r,r') = \mathcal{K}(r',r) \cdot \frac{\pi(r')}{\pi(r)}$ and $\omega_{r'}$ is the deterministic environment with stack r' at each site: $\omega_{r'}(x,i) = r'(i)$, for all $x \in \mathbb{Z}$ and $i \in \mathbb{N}$.

2.3 Related processes

The branching processes (U_k) and (V_k) are difficult to analyze directly because their transition probabilities depend on the underlying environment ω, and therefore these processes are not Markovian when ω is chosen randomly according to $\mathbf{P}$ (or $\mathbf{P}_\pi$ or $\mathbf{P}_s$), and are not time-homogeneous in a fixed environment ω. In this section we introduce some simpler related processes, which are both Markovian and time-homogeneous and, thus, easier to analyze.

2.3.1 The processes $(\widehat{U}_k)_{k\geq 0}$ and $(\widehat{V}_k)_{k\geq 0}$

Throughout this section and the remainder of the paper $s \in \mathcal{S}$ is an arbitrary but fixed stack. We define stopping times $(\tau_k)_{k\geq 0}$ and $(\tau_k')_{k\geq 0}$ by

$$\tau_0 = \inf\{j \geq 0 : R_j = s\} \text{ and } \tau_{k+1} = \inf\{j > \tau_k : R_j = s\}\ ,\ k \geq 0; \tag{2.10}$$

$$\tau_0' = \inf\{j \geq 0 : S_j = s\} \text{ and } \tau_{k+1}' = \inf\{j > \tau_k' : S_j = s\}\ ,\ k \geq 0. \tag{2.11}$$

Then we define processes $(\widehat{V}_k)_{k\geq 0}$ and $(\widehat{U}_k)_{k\geq 0}$ by

$$\widehat{V}_k = V_{\tau_k} \quad \text{and} \quad \widehat{U}_k = U_{\tau_k'}. \tag{2.12}$$

In other words, the processes $(\widehat{V}_k)$ and $(\widehat{U}_k)$ are constructed from (V_k) and (U_k) by observing the latter only at times j when $R_j = s$ or $S_j = s$, respectively.

Since the process $(U_k, S_k)_{k\geq 0}$ is a time-homogeneous Markov chain (under P_x^U, $P_{x,\pi}^U$, and $P_{x,r}^U$, $r \in \mathcal{S}$) and the process $(V_k, R_k)_{k\geq 0}$ is a time-homogeneous Markov chain (under

P_x^V, $P_{x,\pi}^V$, and $P_{x,r}^V$), the processes $(\widehat{U}_k)_{k\geq 0}$ and $(\widehat{V}_k)_{k\geq 0}$ are also time homogeneous Markov chains (under these same measures) with transition probabilities[5]

$$P_{x,r}^U(\widehat{U}_{k+1}=y|\widehat{U}_k=z)=P_{x,\pi}^U(\widehat{U}_{k+1}=y|\widehat{U}_k=z)=P_x^U(\widehat{U}_{k+1}=y|\widehat{U}_k=z)=P_{z,s}^U(U_{\tau_s^S}=y), \tag{2.13}$$

$$P_{x,r}^V(\widehat{V}_{k+1}=y|\widehat{V}_k=z)=P_{x,\pi}^V(\widehat{V}_{k+1}=y|\widehat{V}_k=z)=P_x^V(\widehat{V}_{k+1}=y|\widehat{V}_k=z)=P_{z,s}^V(V_{\tau_s^R}=y) \tag{2.14}$$

for $x,y,z,k\in\mathbb{N}_0$, where (in accordance with our conventions in Section 1.5)

$$\tau_s^S\equiv\inf\{j>0:S_j=s\}\quad\text{and}\quad\tau_s^R\equiv\inf\{j>0:R_j=s\}.$$

In words, the probability of transitioning from z to y for the Markov chain $(\widehat{U}_k)$ is the probability the process (U_k) transitions from level z to level y during the time period that the process (S_k) makes one excursion from state s. Similarly, the probability of transitioning from z to y for the Markov chain $(\widehat{V}_k)$ is the probability the process (V_k) transitions from level z to level y during the time period that the process (R_k) makes one excursion from state s. Although the transition probabilities for these Markov chains are complicated because they depend on the random return times τ_s^S and τ_s^R, we will see in Section 2.4 that they can be analyzed reasonably well. By contrast, trying to analyze the processes (U_k) and (V_k) directly, under any of the above averaged measures, appears difficult, because they are not Markovian.

2.3.2 The dominating processes $(U_k^{\pm})_{k\geq 0}$ and $(V_k^{\pm})_{k\geq 0}$

To analyze the transition probabilities for the processes $(\widehat{U}_k)$ and $(\widehat{V}_k)$ it will be helpful to introduce some additional auxiliary processes, which dominate the processes (U_k) and (V_k), from both above and below. Recall that the forward branching process $(U_k)_{k\geq 0}$ and the backward branching process $(V_k)_{k\geq 0}$, started from level x, are defined in terms of the random variables $(\xi_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ by

$$U_0=x\ \text{ and }\ U_{k+1}=\inf\Big\{m:\sum_{i=1}^m \mathbb{1}\{\xi_i^{k+1}=0\}=U_k\Big\}-U_k,$$
$$V_0=x\ \text{ and }\ V_{k+1}=\inf\Big\{m:\sum_{i=1}^m \mathbb{1}\{\xi_i^{-(k+1)}=1\}=V_k+1\Big\}-(V_k+1).$$

That is, U_{k+1} is the number of successes in the sequence $(\xi_i^{k+1})_{i\in\mathbb{N}}$ prior to the U_k-th failure, and V_{k+1} is the number of failures in the sequence $(\xi_i^{-(k+1)})_{i\in\mathbb{N}}$ prior to the (V_k+1)-th success. Let us define modified processes $(U_k^+)_{k\geq 0}$, $(V_k^+)_{k\geq 0}$, $(U_k^-)_{k\geq 0}$, and

[5] Note that we extend here the probability measures P_x^U, $P_{x,\pi}^U$, $P_{x,r}^U$ for the joint process $(U_k,S_k)_{k\geq 0}$ to the process $(\widehat{U}_k)_{k\geq 0}$ derived from it. The initial value x is still for U_0. This is not, in general, the same as the initial value $\widehat{U}_0$, except under $P_{x,s}^U$ where $S_0=s$ deterministically. Similar remarks apply to the process $(\widehat{V}_k)$ with respect to the probability measures P_x^V, $P_{x,\pi}^V$, $P_{x,r}^V$.

$(V_k^-)_{k\geq 0}$, all started from level x as follows[6]:

$$U_0^+ = x \text{ and } U_{k+1}^+ = \inf\Big\{m \geq M : \sum_{i=M+1}^{m} \mathbb{1}\{\xi_i^{k+1}=0\} = U_k^+ + 1\Big\} - (U_k^+ + 1),$$

$$V_0^+ = x \text{ and } V_{k+1}^+ = \inf\Big\{m \geq M : \sum_{i=M+1}^{m} \mathbb{1}\{\xi_i^{-(k+1)}=1\} = V_k^+ + 1\Big\} - (V_k^+ + 1),$$

$$U_0^- = x \text{ and } U_{k+1}^- = \inf\Big\{m \geq M : \sum_{i=M+1}^{m} \mathbb{1}\{\xi_i^{k+1}=0\} = (U_k^- - M)^+\Big\} - (U_k^- - M)^+ - M,$$

$$V_0^- = x \text{ and } V_{k+1}^- = \inf\Big\{m \geq M : \sum_{i=M+1}^{m} \mathbb{1}\{\xi_i^{-(k+1)}=1\} = (V_k^- - M)^+\Big\} - (V_k^- - M)^+ - M.$$

In words, U_{k+1}^+ is the number of successes in the sequence $(\xi_i^{k+1})_{i\in\mathbb{N}}$ before the $(U_k^+ + 1)$-th failure, when we condition that $\xi_1^{k+1}, \ldots, \xi_M^{k+1}$ are all successes, and U_{k+1}^- is the number of successes in the sequence $(\xi_i^{k+1})_{i\in\mathbb{N}}$ before the U_k^--th failure, when we condition that $\xi_1^{k+1}, \ldots, \xi_M^{k+1}$ are all failures. The interpretations for V_{k+1}^+ and V_{k+1}^- are the same with "success" replaced by "failure" and " ξ_i^{k+1} " replaced by " $\xi_i^{-(k+1)}$ ".

By construction we have

$$U_k^- \leq U_k \leq U_k^+ \text{ , for all } k \quad \text{ and } \quad V_k^- \leq V_k \leq V_k^+ \text{ , for all } k \tag{2.15}$$

if all processes are started from the same level x. Also, since $(\xi_i^k)_{k\in\mathbb{Z},i>M}$ are i.i.d. Ber(1/2) random variables, for any values of the cookie stacks $(S_k)_{k\in\mathbb{Z}}$, each of the processes (U_k^+), (U_k^-), (V_k^+), (V_k^-) is a time-homogeneous Markov chain and

$$(U_k^-, U_k^+)_{k\geq 0} \perp (S_k)_{k\in\mathbb{Z}} \quad \text{and} \quad (V_k^-, V_k^+)_{k\geq 0} \perp (S_k)_{k\in\mathbb{Z}}. \tag{2.16}$$

These statements hold for any initial values $U_0^-, U_0^+, V_0^-, V_0^+$ and any marginal distribution ρ on S_0 (including ϕ, π, or a point mass at $r \in \mathcal{S}$). For the same reason (i.e. that $(\xi_i^k)_{k\in\mathbb{Z},i>M}$ are i.i.d. Ber(1/2)), we also have

$$(U_k^-, U_k^+)_{k\geq 0} \overset{law}{=} (V_k^-, V_k^+)_{k\geq 0} \tag{2.17}$$

when all processes are started from the same level x (again for any marginal distribution on S_0).

To analyze the processes $(U_k^{\pm})$ and $(V_k^{\pm})$ it will be helpful to represent them in a form similar to (2.3) and (2.7) for the processes (U_k) and (V_k). Define $\mathcal{G}_i^k$ to be the number of successes in the sequence $(\xi_j^k)_{j>M}$ between the $(i-1)$-th and i-th failures, and define $\mathcal{H}_i^k$ to be the number of failures in the sequence $(\xi_j^{-k})_{j>M}$ between the $(i-1)$-th and i-th successes. Then

$$U_{k+1}^+ = M + \sum_{i=1}^{U_k^+ + 1} \mathcal{G}_i^{k+1} \quad \text{and} \quad U_{k+1}^- = \sum_{i=1}^{(U_k^- - M)^+} \mathcal{G}_i^{k+1} \tag{2.18}$$

where $(\mathcal{G}_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ are i.i.d. Geo(1/2) random variables, and

$$V_{k+1}^+ = M + \sum_{i=1}^{V_k^+ + 1} \mathcal{H}_i^{k+1} \quad \text{and} \quad V_{k+1}^- = \sum_{i=1}^{(V_k^- - M)^+} \mathcal{H}_i^{k+1} \tag{2.19}$$

where $(\mathcal{H}_i^k)_{k\in\mathbb{Z},i\in\mathbb{N}}$ are i.i.d. Geo(1/2) random variables.

[6] Note that, by our conventions for empty sums, the infimum in the last two equations is defined to be M if $U_k^- \leq M$ or $V_k^- \leq M$. Thus, in these cases the whole right hand side is 0.

2.4 Expectation, variance, and concentration estimates

In this section we use the dominating processes $(U_k^{\pm})$ and $(V_k^{\pm})$ to analyze the transition probabilities (2.13) and (2.14) for the processes $(\widehat{U}_k)$ and $(\widehat{V}_k)$. We also prove, slightly more generally, concentration estimates for the processes (U_k) and (V_k) up to the random stopping times τ_s^S and τ_s^R, when (S_k) and (R_k), respectively, are started from an arbitrary initial state $r \in \mathcal{S}$, rather than s. Finally, we prove a type of "overshoot lemma" for the processes $(\widehat{U}_k)$ and $(\widehat{V}_k)$ analogous to Lemma 5.1 of [8].

Throughout it is assumed, when not otherwise specified, that the processes (U_k), (U_k^+), (U_k^-) are all started from the same level x, and the processes (V_k), (V_k^+), (V_k^-) are all started from the same level x. The probability measure $P_{x,r}^U$ will be used for all these "U-processes" started from level x when $S_0 = r$, and the probability measure $P_{x,r}^V$ will be used for all these "V-processes" started from level x when $R_0 = r$. The following general fact will be needed in our analysis of the "U-processes" and "V-processes" below, as well as in several other parts of the paper.

Lemma 2.5. *Let Z be a random variable with mean μ and exponential tails, and let $Z_1, Z_2, \dots$ be i.i.d. random variables distributed as Z. Then for any $\epsilon_0 \in (0, \infty)$ there exist constants $C_1(\epsilon_0), C_2(\epsilon_0) > 0$ such that the empirical means $\overline{Z}_n \equiv \frac{1}{n}\sum_{i=1}^n Z_i$ satisfy:*

$$\mathbb{P}(|\overline{Z}_n - \mu| \geq \epsilon) \leq C_1 \exp(-C_2 \epsilon^2 n)\ , \text{ for all } 0 < \epsilon \leq \epsilon_0 \text{ and } n \in \mathbb{N}. \tag{2.20}$$

$$\mathbb{P}(|\overline{Z}_n - \mu| \geq \epsilon) \leq C_1 \exp(-C_2 \epsilon n)\ , \text{ for all } \epsilon \geq \epsilon_0 \text{ and } n \in \mathbb{N}. \tag{2.21}$$

Proof. The exponential tails condition on the random variable Z implies there exist some positive constants b, c such that for all $\lambda \in [-b, b]$, $\mathbb{E}(e^{\lambda(Z-\mu)}) \leq e^{c\lambda^2}$ and $\mathbb{E}(e^{\lambda(\mu-Z)}) \leq e^{c\lambda^2}$. Thus, the lemma is a consequence of [17, Theorem III.15]. □

Using Lemma 2.5 along with (2.18) and (2.19) and a small bit of analysis one may obtain the following concentration estimates for the differences $(U_k^{\pm} - U_{k-1}^{\pm})$ and $(V_k^{\pm} - V_{k-1}^{\pm})$.

Lemma 2.6. *For each $\epsilon_0 \in (0, \infty)$, there exist constants $c_1(\epsilon_0)$, $c_2(\epsilon_0) > 0$ such that the following hold for all $r \in \mathcal{S}$, $x, y \in \mathbb{N}_0$, and $k \in \mathbb{N}$:*

$$P_{x,r}^U\left(|U_k^- - U_{k-1}^-| \geq \epsilon y \middle| U_{k-1}^- = y\right) \leq c_1 e^{-c_2\epsilon^2 y}\ , \quad \text{for } 0 < \epsilon \leq \epsilon_0. \tag{2.22}$$

$$P_{x,r}^U\left(|U_k^- - U_{k-1}^-| \geq \epsilon y \middle| U_{k-1}^- = y\right) \leq c_1 e^{-c_2\epsilon y}\ , \quad \text{for } \epsilon \geq \epsilon_0. \tag{2.23}$$

$$P_{x,r}^U\left(|U_k^+ - U_{k-1}^+| \geq \epsilon y \middle| U_{k-1}^+ = y\right) \leq c_1 e^{-c_2\epsilon^2 y}\ , \quad \text{for } 0 < \epsilon \leq \epsilon_0. \tag{2.24}$$

$$P_{x,r}^U\left(|U_k^+ - U_{k-1}^+| \geq \epsilon y \middle| U_{k-1}^+ = y\right) \leq c_1 e^{-c_2\epsilon y}\ , \quad \text{for } \epsilon \geq \epsilon_0. \tag{2.25}$$

Moreover, the equivalent statements also hold for (V_k^+) and (V_k^-) with the same constants c_1, c_2.

Remark 2.7. Note that since (U_k^+) and (U_k^-) are each time-homogeneous Markov chains independent of $(S_k)_{k\in\mathbb{Z}}$ the probabilities on the left hand side of these equations do not depend on x, r, or k. Similar statements also apply for the processes (V_k^+) and (V_k^-).

We wish now to extend these concentration estimates for the single time step differences in the processes $(U_k^{\pm})$ and $(V_k^{\pm})$ to concentration estimates for these processes up to the random stopping times τ_s^S and τ_s^R. This is where we will need the uniform ergodicity hypothesis on the Markov chains (S_k) and (R_k). Due to Lemma 2.8 below and uniform ergodicity of (S_k) and (R_k) there exist some constants $c_3, c_4 > 0$ such that:

$$\mathbf{P}_r(\tau_s^S \geq t) \leq c_3 e^{-c_4 t}\ , \quad \text{for all } r \in \mathcal{S} \text{ and } t \in [0, \infty).$$

$$\mathbf{P}_r(\tau_s^R \geq t) \leq c_3 e^{-c_4 t}\ , \quad \text{for all } r \in \mathcal{S} \text{ and } t \in [0, \infty). \tag{2.26}$$

Here, as described in Section 1.5, $\mathbf{P}_r$ is the probability measure for the process (S_k) itself (equivalently the process (R_k)) when $S_0 = R_0 = r$.

Lemma 2.8. *Let (Z_k) be a uniformly ergodic Markov chain on a countable state space $\mathcal{Z}$, and let $\mathbb{P}_x(\cdot)$ be the probability measure for the Markov chain (Z_k) started from $Z_0 = x$. Then, for each $z \in \mathcal{Z}$, there exist constants $C > 0$ and $0 < \alpha < 1$ such that*

$$\mathbb{P}_x(\tau_z^Z > n) \leq C\alpha^n \ , \ \ \textit{for all } x \in \mathcal{Z}, n \in \mathbb{N}_0.$$

Proof. Fix $z \in \mathcal{Z}$. Let $\rho = (\rho(x))_{x\in\mathcal{Z}}$ denote the stationary distribution of the Markov chain (Z_k), and let $\mathcal{M} = \{\mathcal{M}(x,y)\}_{x,y\in\mathcal{Z}}$ be its transition matrix. Also, let $\epsilon = \rho(z)$. By uniform ergodicity of (Z_k), there is some $\ell \in \mathbb{N}$ such that $\|\mathcal{M}^\ell(y,\cdot) - \rho(\cdot)\|_{TV} \leq \epsilon/2$, for all $y \in \mathcal{Z}$. This implies $\mathbb{P}_y(Z_\ell = z) = \mathcal{M}^\ell(y,z) \geq \epsilon/2$, for all $y \in \mathcal{Z}$. Thus, starting from any initial state x we have

$$\mathbb{P}_x(\tau_z^Z > n\cdot\ell) = \prod_{m=0}^{n-1} \mathbb{P}_x(\tau_z^Z > (m+1)\ell | \tau_z^Z > m\ell) \leq (1-\epsilon/2)^n = \alpha^{\ell n},$$

where $\alpha \equiv (1-\epsilon/2)^{1/\ell} \in (0,1)$. It follows that $\mathbb{P}_x(\tau_z^Z > n) \leq C\alpha^n$, for all $n \geq 0$ and $x \in \mathcal{Z}$, with $C \equiv 1/\alpha^\ell$. ☐

Lemma 2.9. *There exist constants $c_5, c_6 > 0$ such that the following hold for each $r \in \mathcal{S}$:*

$$P_{x,r}^U\left(\max_{0\leq k\leq \tau_s^S} |U_k^- - x| \geq \epsilon x\right) \leq c_5(1+\epsilon^{2/3}x^{1/3})e^{-c_6\epsilon^{2/3}x^{1/3}} \ , \ \textit{for all } x \in \mathbb{N}_0 \textit{ and } 0 < \epsilon \leq 1. \tag{2.27}$$

$$P_{x,r}^U\left(\max_{0\leq k\leq \tau_s^S} |U_k^- - x| \geq \epsilon x\right) \leq c_5(1+\epsilon^{1/3}x^{1/3})e^{-c_6\epsilon^{1/3}x^{1/3}} \ , \ \textit{for all } x \in \mathbb{N}_0 \textit{ and } \epsilon \geq 1. \tag{2.28}$$

$$P_{x,r}^U\left(\max_{0\leq k\leq \tau_s^S} |U_k^+ - x| \geq \epsilon x\right) \leq c_5(1+\epsilon^{2/3}x^{1/3})e^{-c_6\epsilon^{2/3}x^{1/3}} \ , \ \textit{for all } x \in \mathbb{N}_0 \textit{ and } 0 < \epsilon \leq 1. \tag{2.29}$$

$$P_{x,r}^U\left(\max_{0\leq k\leq \tau_s^S} |U_k^+ - x| \geq \epsilon x\right) \leq c_5(1+\epsilon^{1/3}x^{1/3})e^{-c_6\epsilon^{1/3}x^{1/3}} \ , \ \textit{for all } x \in \mathbb{N}_0 \textit{ and } \epsilon \geq 1. \tag{2.30}$$

Moreover, the equivalent statements (with τ_s^S replaced by τ_s^R) also hold for the processes (V_k^+) and (V_k^-) with the same constants c_5, c_6.

All statements are trivially true if $x = 0$ (taking any $c_5 \geq 1$ and any $c_6 > 0$), so we will assume $x \geq 1$. We will prove (2.27) and (2.28). The proof of (2.29) is identical to that of (2.27) with U_k^- replaced by U_k^+ line by line, and the proof of (2.30) is almost identical to that of (2.28) with U_k^- replaced by U_k^+ line by line[7]. The equivalent statements to (2.27)-(2.30) for the processes (V_k^-) and (V_k^+) are also proved exactly the same way; all the proofs use is Lemma 2.6 and (2.26), and these estimates are the same for $(V_k^\pm)$ and $(U_k^\pm)$ and for τ_s^R and τ_s^S. The proofs of (2.27) and (2.28) will be given separately and the constants c_5, c_6 obtained in the two cases will not be the same. To find a single c_5 and c_6 that hold in both cases simply take c_5 to be the maximum of the c_5's from the two proofs and c_6 to be the minimum of the c_6's from the two proofs.

[7] In the derivation of (2.44) the case $z = 0$ must be considered separately for both proofs. For the process (U_k^-) we have $U_k^- = 0$ (deterministically) if $U_{k-1}^- = z = 0$, as noted in the proof of (2.28) below. For the process (U_k^+) this is not the case. However, if we start with $U_0^+ = x \geq 1$, as we assume, then it is actually impossible that U_k^+ is 0 for any $k \geq 0$ (indeed, $U_k^+ \geq M$ for all $k \geq 1$). So we do not have this problem to deal with.

Proof of Lemma 2.9, Equation (2.27)*, with* $x \geq 1$. Fix $\epsilon \leq 1$ and denote $\mathcal{M} = \max_{0 \leq k \leq \tau_s^S} \left|U_k^- - x\right|$. Then

$$\begin{aligned}
P^U_{x,r}(\mathcal{M} \geq \epsilon x) &= P^U_{x,r}\Big(\tau_s^S \leq \frac{1}{2}\epsilon^{2/3}x^{1/3}\Big) P^U_{x,r}\Big(\mathcal{M} \geq \epsilon x \Big| \tau_s^S \leq \frac{1}{2}\epsilon^{2/3}x^{1/3}\Big) \\
&\quad + P^U_{x,r}\Big(\tau_s^S > \frac{1}{2}\epsilon^{2/3}x^{1/3}\Big) P^U_{x,r}\Big(\mathcal{M} \geq \epsilon x \Big| \tau_s^S > \frac{1}{2}\epsilon^{2/3}x^{1/3}\Big) \\
&\leq \max_{0 \leq n \leq \frac{1}{2}\epsilon^{2/3}x^{1/3}} P^U_{x,r}\Big(\mathcal{M} \geq \epsilon x \Big| \tau_s^S = n\Big) \; + \; P^U_{x,r}\Big(\tau_s^S > \frac{1}{2}\epsilon^{2/3}x^{1/3}\Big). \qquad (2.31)
\end{aligned}$$

By (2.26),

$$P^U_{x,r}\Big(\tau_s^S > \frac{1}{2}\epsilon^{2/3}x^{1/3}\Big) = \mathbf{P}_r\Big(\tau_s^S > \frac{1}{2}\epsilon^{2/3}x^{1/3}\Big) \leq c_3 e^{-c_4 \cdot \frac{1}{2}\epsilon^{2/3}x^{1/3}}. \qquad (2.32)$$

So, we need only bound $P^U_{x,r}(\mathcal{M} \geq \epsilon x | \tau_s^S = n)$, for $n \leq \frac{1}{2}\epsilon^{2/3}x^{1/3}$. Fix such an n and define events A_k, $k \geq 1$, by $A_k = \left\{\left|U_k^- - U_{k-1}^-\right| \leq \epsilon^{1/3}x^{2/3}\right\}$. Then, for $0 \leq k \leq n$, we have

$$U_k^- \leq x + n \cdot \epsilon^{1/3}x^{2/3} \leq x + x/2 < 2x \;, \quad \text{on the event } A_1 \cap \ldots \cap A_k \qquad (2.33)$$

and

$$U_k^- \geq x - n \cdot \epsilon^{1/3}x^{2/3} \geq x - x/2 = x/2 \;, \quad \text{on the event } A_1 \cap \ldots \cap A_k. \qquad (2.34)$$

Also, by Lemma 2.6, there exist some constants $c_1, c_2 > 0$ such that

$$P^U_{x,r}\Big(|U_k^- - U_{k-1}^-| > \widetilde{\epsilon} z \Big| U_{k-1}^- = z\Big) \leq c_1 \exp(-c_2 \cdot \widetilde{\epsilon}^2 \cdot z) \;, \quad \text{for each } k, z \text{ and } 0 < \widetilde{\epsilon} \leq 2. \qquad (2.35)$$

Since $\epsilon \leq 1$ and $x \geq 1$, $\epsilon^{1/3}x^{2/3}/z \leq 2$ for all $z \geq x/2$. Thus, for each $x/2 \leq z \leq 2x$, we have

$$\begin{aligned}
P^U_{x,r}(A_k^c | U_{k-1}^- = z) &= P^U_{x,r}\Big(|U_k^- - U_{k-1}^-| > \Big(\frac{\epsilon^{1/3}x^{2/3}}{z}\Big) \cdot z \Big| U_{k-1}^- = z\Big) \\
&\leq c_1 \exp\Big(-c_2 \frac{(\epsilon^{1/3}x^{2/3})^2}{2x}\Big) = c_1 \exp\Big(-\frac{c_2}{2}\epsilon^{2/3}x^{1/3}\Big). \qquad (2.36)
\end{aligned}$$

Combining (2.33), (2.34), and (2.36) and using the fact that (U_k^-) is a Markov chain shows

$$P^U_{x,r}(A_k^c | A_1, \ldots, A_{k-1}) \leq c_1 \exp\Big(-\frac{c_2}{2}\epsilon^{2/3}x^{1/3}\Big) \;, \quad \text{for } 1 \leq k \leq n.$$

Hence,

$$P^U_{x,r}\left(\bigcup_{k=1}^n A_k^c\right) \leq n \cdot c_1 \exp\Big(-\frac{c_2}{2}\epsilon^{2/3}x^{1/3}\Big) \leq \frac{1}{2}\epsilon^{2/3}x^{1/3} \cdot c_1 \exp\Big(-\frac{c_2}{2}\epsilon^{2/3}x^{1/3}\Big).$$

Now, when $\tau_s^S = n$, $\mathcal{M} \leq n \cdot \epsilon^{1/3}x^{2/3} \leq \frac{1}{2}\epsilon x$ on the event $\cap_{k=1}^n A_k$. So,

$$P^U_{x,r}(\mathcal{M} \geq \epsilon x | \tau_s^S = n) \leq \frac{1}{2}\epsilon^{2/3}x^{1/3} \cdot c_1 \exp\Big(-\frac{c_2}{2}\epsilon^{2/3}x^{1/3}\Big). \qquad (2.37)$$

Since, (2.37) is valid for each $n \leq \frac{1}{2}\epsilon^{2/3}x^{1/3}$ it follows from (2.31) and (2.32) that

$$P^U_{x,r}(\mathcal{M} \geq \epsilon x) \leq c_5(1 + \epsilon^{2/3}x^{1/3}) \exp(-c_6 \epsilon^{2/3}x^{1/3})$$

where $c_5 = \max\{\frac{c_1}{2}, c_3\}$ and $c_6 = \min\{\frac{c_2}{2}, \frac{c_4}{2}\}$. □

Remark 2.10. If we define the deterministic time $t_{\epsilon,x} = \lfloor \frac{1}{2}\epsilon^{2/3}x^{1/3} \rfloor$, for $0 < \epsilon \le 1$ and $x \in \mathbb{N}$, then the same exact steps used in the derivation of (2.37) for an arbitrary $n \le t_{\epsilon,x}$ show that

$$P^U_{x,r}\left(\max_{0\le k\le t_{\epsilon,x}} |U^-_k - x| \ge \epsilon x\right) \le c_1 t_{\epsilon,x} \exp\left(-c_2 t_{\epsilon,x}\right). \tag{2.38}$$

We isolate this observation, as it will be needed later in the proof of Lemma 2.12.

Proof of Lemma 2.9, Equation (2.28)*, with* $x \ge 1$*.* As above, let $\mathcal{M} = \max_{0\le k\le \tau^S_s} \left|U^-_k - x\right|$. We will assume that $\epsilon > 1$, as the case $\epsilon = 1$ follows from (2.27). Then

$$\begin{aligned} P^U_{x,r}(\mathcal{M} \ge \epsilon x) &= P^U_{x,r}\left(\tau^S_s < \epsilon^{1/3}x^{1/3}\right) P^U_{x,r}\left(\mathcal{M} \ge \epsilon x \middle| \tau^S_s < \epsilon^{1/3}x^{1/3}\right) \\ &\quad + P^U_{x,r}\left(\tau^S_s \ge \epsilon^{1/3}x^{1/3}\right) P^U_{x,r}\left(\mathcal{M} \ge \epsilon x \middle| \tau^S_s \ge \epsilon^{1/3}x^{1/3}\right) \\ &\le \max_{0\le n<\epsilon^{1/3}x^{1/3}} P^U_{x,r}\left(\mathcal{M} \ge \epsilon x \middle| \tau^S_s = n\right) \ + \ P^U_{x,r}\left(\tau^S_s \ge \epsilon^{1/3}x^{1/3}\right). \end{aligned} \tag{2.39}$$

By (2.26),

$$P^U_{x,r}(\tau^S_s \ge \epsilon^{1/3}x^{1/3}) = \mathbf{P}_r(\tau^S_s \ge \epsilon^{1/3}x^{1/3}) \le c_3 e^{-c_4\epsilon^{1/3}x^{1/3}}. \tag{2.40}$$

So, we need only bound $P^U_{x,r}(\mathcal{M} \ge \epsilon x | \tau^S_s = n)$, for $n < \epsilon^{1/3}x^{1/3}$. Fix such an n and define events A_k, $k \ge 1$, by $A_k = \left\{\left|U^-_k - U^-_{k-1}\right| \le \epsilon^{2/3}x^{2/3}\right\}$. Then, for $0 \le k \le n$, we have

$$U^-_k \le x + n\cdot\epsilon^{2/3}x^{2/3} \le x + \epsilon x \le 2\epsilon x\ , \ \text{ on the event } A_1\cap\ldots\cap A_k. \tag{2.41}$$

Also, by Lemma 2.6, there exist some constants $c_1, c_2 > 0$ such that:

$$\begin{aligned} &P^U_{x,r}\left(|U^-_k - U^-_{k-1}| > \widetilde{\epsilon} z \middle| U^-_{k-1} = z\right) \le c_1\exp(-c_2\cdot\widetilde{\epsilon}^2\cdot z)\ , \ \text{ for each } k, z \text{ and } 0 < \widetilde{\epsilon} \le 1. \\ &P^U_{x,r}\left(|U^-_k - U^-_{k-1}| > \widetilde{\epsilon} z \middle| U^-_{k-1} = z\right) \le c_1\exp(-c_2\cdot\widetilde{\epsilon}\cdot z)\ , \ \text{ for each } k, z \text{ and } \widetilde{\epsilon} \ge 1. \end{aligned} \tag{2.42}$$

Thus, for each $\epsilon^{2/3}x^{2/3} \le z \le 2\epsilon x$ and $1 \le k \le n$,

$$\begin{aligned} P^U_{x,r}(A^c_k | U^-_{k-1} = z) &= P^U_{x,r}\left(|U^-_k - U^-_{k-1}| > \Big(\frac{\epsilon^{2/3}x^{2/3}}{z}\Big)\cdot z \middle| U^-_{k-1} = z\right) \\ &\le c_1\exp\Big(-c_2\frac{(\epsilon^{2/3}x^{2/3})^2}{2\epsilon x}\Big) = c_1\exp\Big(-\frac{c_2}{2}\epsilon^{1/3}x^{1/3}\Big) \end{aligned} \tag{2.43}$$

and, for each $z < \epsilon^{2/3}x^{2/3}$ and $1 \le k \le n$,

$$P^U_{x,r}(A^c_k | U^-_{k-1} = z) = P^U_{x,r}\left(|U^-_k - U^-_{k-1}| > \Big(\frac{\epsilon^{2/3}x^{2/3}}{z}\Big)\cdot z \middle| U^-_{k-1} = z\right) \le c_1\exp\Big(-c_2\epsilon^{2/3}x^{2/3}\Big). \tag{2.44}$$

Note that the inequality (2.44) remains valid when $z = 0$ (even though the derivation above has an issue dividing by 0), since in this case $U^-_k = U^-_{k-1} = 0$, deterministically. Now, since $\epsilon, x \ge 1$, by assumption, and the process (U^-_k) is a Markov chain, it follows from (2.41), (2.43), and (2.44) that

$$P^U_{x,r}(A^c_k | A_1, \ldots, A_{k-1}) \le c_1\exp\left(-\frac{c_2}{2}\epsilon^{1/3}x^{1/3}\right)\ , \ \text{ for } 1 \le k \le n.$$

Thus,

$$P^U_{x,r}\left(\bigcup_{k=1}^n A^c_k\right) \le n\cdot c_1\exp\left(-\frac{c_2}{2}\epsilon^{1/3}x^{1/3}\right) < \epsilon^{1/3}x^{1/3}\cdot c_1\exp\left(-\frac{c_2}{2}\epsilon^{1/3}x^{1/3}\right).$$

Now, when $\tau_s^S = n$, $\mathcal{M} \le n \cdot \epsilon^{2/3} x^{2/3} < \epsilon^{1/3} x^{1/3} \cdot \epsilon^{2/3} x^{2/3} = \epsilon x$ on the event $\cap_{k=1}^n A_k$. So,

$$P_{x,r}^U(\mathcal{M} \ge \epsilon x | \tau_s^S = n) < \epsilon^{1/3} x^{1/3} \cdot c_1 \exp\left(-\frac{c_2}{2}\epsilon^{1/3} x^{1/3}\right). \tag{2.45}$$

Since (2.45) is valid for each $n < \epsilon^{1/3} x^{1/3}$ it follows from (2.39) and (2.40) that

$$P_{x,r}^U(\mathcal{M} \ge \epsilon x) \le c_5(1 + \epsilon^{1/3} x^{1/3}) \exp(-c_6 \epsilon^{1/3} x^{1/3})$$

where $c_5 = \max\{c_1, c_3\}$ and $c_6 = \min\{\frac{c_2}{2}, c_4\}$. □

Using (2.15) the concentration estimates for the processes $(U_k^\pm)$ and $(V_k^\pm)$ proven above in Lemma 2.9 yield the following concentration estimates for the processes (U_k) and (V_k).

Lemma 2.11. *Let $c_7 = 2c_5$. Then the following hold for each $r \in \mathcal{S}$:*

$$P_{x,r}^U\left(\max_{0 \le k \le \tau_s^S} |U_k - x| \ge \epsilon x\right) \le c_7(1 + \epsilon^{2/3} x^{1/3}) e^{-c_6 \epsilon^{2/3} x^{1/3}}, \text{ for all } x \in \mathbb{N}_0 \text{ and } 0 < \epsilon \le 1. \tag{2.46}$$

$$P_{x,r}^U\left(\max_{0 \le k \le \tau_s^S} |U_k - x| \ge \epsilon x\right) \le c_7(1 + \epsilon^{1/3} x^{1/3}) e^{-c_6 \epsilon^{1/3} x^{1/3}}, \text{ for all } x \in \mathbb{N}_0 \text{ and } \epsilon \ge 1. \tag{2.47}$$

$$P_{x,r}^V\left(\max_{0 \le k \le \tau_s^R} |V_k - x| \ge \epsilon x\right) \le c_7(1 + \epsilon^{2/3} x^{1/3}) e^{-c_6 \epsilon^{2/3} x^{1/3}}, \text{ for all } x \in \mathbb{N}_0 \text{ and } 0 < \epsilon \le 1. \tag{2.48}$$

$$P_{x,r}^V\left(\max_{0 \le k \le \tau_s^R} |V_k - x| \ge \epsilon x\right) \le c_7(1 + \epsilon^{1/3} x^{1/3}) e^{-c_6 \epsilon^{1/3} x^{1/3}}, \text{ for all } x \in \mathbb{N}_0 \text{ and } \epsilon \ge 1. \tag{2.49}$$

The next two lemmas give estimates for the expectation and variance of $U_{\tau_s^S}$ and $V_{\tau_s^R}$ in the case that $S_0 = R_0 = s$. For these lemmas, and the remainder of this section, $E_{x,s}^U$, $\text{Var}_{x,s}^U$, and $\text{Cov}_{x,s}^U$ are used to denote, respectively, expectation, variance, and covariance under the probability measure $P_{x,s}^U$. Similarly, $E_{x,s}^V$, $\text{Var}_{x,s}^V$, and $\text{Cov}_{x,s}^V$ are used to denote expectation, variance, and covariance with respect to $P_{x,s}^V$. Also, we define

$$\mu_s \equiv \mathbf{E}_s(\tau_s^S) = \mathbf{E}_s(\tau_s^R)$$

to be the mean return time to state s for the Markov chain (S_k), or equivalently for the Markov chain (R_k). Note that $E_{x,s}^U(\tau_s^S) = E_{x,s}^V(\tau_s^R) = \mu_s$, for any x.

Lemma 2.12. *As $x \to \infty$,*

$$E_{x,s}^U(U_{\tau_s^S}) = x + \delta \cdot \mu_s + O(e^{-x^{1/4}}) \quad \text{and} \quad E_{x,s}^V(V_{\tau_s^R}) = x + (1 - \delta) \cdot \mu_s + O(e^{-x^{1/4}}).$$

Proof. We will prove the statement about the expectation of $U_{\tau_s^S}$. The proof of the analogous claim for the expectation of $V_{\tau_s^R}$ is very similar.

Fix a realization $(s_k)_{k \in \mathbb{Z}}$ of the random variables $(S_k)_{k \in \mathbb{Z}}$ with $s_0 = s$, and let $\omega = (\omega(k,i))_{k \in \mathbb{Z}, i \in \mathbb{N}}$ be the corresponding cookie environment defined by $\omega(k,i) = s_k(i)$. Also, let $t_s^S = \inf\{k > 0 : s_k = s\}$ be the corresponding realization of the random variable τ_s^S. We will consider first the process $(U_k)_{k \ge 0}$ started from level x in this fixed environment ω. For $k \in \mathbb{Z}$, we define $\delta_k = \sum_{i=1}^M (2\omega(k,i) - 1)$ to be the net drift induced by consuming all cookies in stack s_k. Also, we denote by $E_x^{U,\omega}$ expectation with respect to the probability measure $P_x^{U,\omega}$ for the process $(U_k)_{k \ge 0}$ in this fixed environment ω.

We decompose $E_x^{U,\omega}(U_{t_s^S})$ as

$$E_x^{U,\omega}(U_{t_s^S}) = E_x^{U,\omega}(U_0) + \sum_{k=1}^{t_s^S} E_x^{U,\omega}(U_k - U_{k-1}) = x + \sum_{k=1}^{t_s^S} E_x^{U,\omega}(U_k - U_{k-1}). \tag{2.50}$$

Straightforward calculations show that

$$E_x^{U,\omega}(U_k - U_{k-1}|U_{k-1} = m) = \delta_k \text{ , for each } k \geq 1 \text{ and } m \geq M\text{, and}$$
$$\left|E_x^{U,\omega}(U_k - U_{k-1}|U_{k-1} = m)\right| \leq M \text{ , for each } k \geq 1 \text{ and } m \geq 0.$$

Thus,

$$\begin{aligned}\left|E_x^{U,\omega}(U_k - U_{k-1}) - \delta_k\right| &= \Big|P_x^{U,\omega}(U_{k-1} \geq M) \cdot E_x^{U,\omega}(U_k - U_{k-1}|U_{k-1} \geq M) \\ &\qquad + P_x^{U,\omega}(U_{k-1} < M) \cdot E_x^{U,\omega}(U_k - U_{k-1}|U_{k-1} < M) - \delta_k\Big| \\ &\leq \left|P_x^{U,\omega}(U_{k-1} \geq M) \cdot \delta_k - \delta_k\right| + P_x^{U,\omega}(U_{k-1} < M) \cdot M \\ &\leq 2M \cdot P_x^{U,\omega}(U_{k-1} < M).\end{aligned}$$

Plugging into (2.50) gives

$$\left|E_x^{U,\omega}(U_{t_s^S}) - x - \sum_{k=1}^{t_s^S} \delta_k\right| \leq 2M \sum_{k=1}^{t_s^S} P_x^{U,\omega}(U_{k-1} < M). \tag{2.51}$$

So far our analysis has been for a fixed environment ω. To prove the lemma we will need to take expectations with respect to the probability measure $\mathbf{P}_s$ on environments. Recall that, for $r \in \mathcal{S}$, $\delta(r) \equiv \sum_{i=1}^{M}(2r(i) - 1)$ is the net drift induced by all cookies in stack r. Taking expectation of the random variable $g(\omega) = \sum_{k=1}^{t_s^S(\omega)} \delta_k(\omega)$ with respect to $\mathbf{P}_s$ gives

$$\begin{aligned}\mathbf{E}_s\left[\sum_{k=1}^{t_s^S} \delta_k\right] &= \mathbf{E}_s\left[\sum_{k=1}^{t_s^S}\sum_{r\in\mathcal{S}} \delta(r) \cdot \mathbb{1}\{s_k = r\}\right] = \mathbf{E}_s\left[\sum_{r\in\mathcal{S}}\sum_{k=1}^{t_s^S} \delta(r) \cdot \mathbb{1}\{s_k = r\}\right] \\ &\overset{(*)}{=} \sum_{r\in\mathcal{S}} \mathbf{E}_s\left[\sum_{k=1}^{t_s^S} \delta(r)\mathbb{1}\{s_k = r\}\right] = \sum_{r\in\mathcal{S}} \delta(r) \cdot \left[\pi(r) \cdot \mu_s\right] = \delta \cdot \mu_s.\end{aligned} \tag{2.52}$$

In step (*) we have used Fubini's Theorem to interchange the sum with the expectation; this is applicable since $\mathbf{E}_s\left(\sum_{r\in\mathcal{S}}\left|\sum_{k=1}^{t_s^S} \delta(r) \cdot \mathbb{1}\{s_k = r\}\right|\right) \leq \mathbf{E}_s(M \cdot t_s^S) = M\mu_s < \infty$. Combining (2.51) and (2.52), and using the fact that $\mathbf{E}_s[E_x^{U,\omega}(U_{t_s^S})] = E_{x,s}^U(U_{\tau_s^S})$, gives

$$\left|E_{x,s}^U(U_{\tau_s^S}) - x - \delta \cdot \mu_s\right| = \left|\mathbf{E}_s\left(E_x^{U,\omega}(U_{t_s^S}) - x - \sum_{k=1}^{t_s^S} \delta_k\right)\right| \leq 2M \cdot \mathbf{E}_s\left(\sum_{k=1}^{t_s^S} P_x^{U,\omega}(U_{k-1} < M)\right). \tag{2.53}$$

Denote $p_{x,k} = P_{x,s}^U(U_k^- < M)$ and $q_n = \mathbf{P}_s(\tau_s^S = n)$. By construction of the process $(U_k^-)_{k\geq 0}$, $P_x^{U,\omega}(U_{k-1} < M) \leq P_x^{U,\omega}(U_{k-1}^- < M) = p_{x,k-1}$, for $\mathbf{P}_s$ a.e. ω. Thus, it follows from (2.53) that

$$\left|E_{x,s}^U(U_{\tau_s^S}) - x - \delta \cdot \mu_s\right| \leq 2M \sum_{n=1}^{\infty} q_n \sum_{k=1}^{n} p_{x,k-1}. \tag{2.54}$$

We define $n_0 = \lfloor (1/2)^{5/3} x^{1/3} \rfloor$. Splitting the sum at n_0, the right hand side of (2.54) may be bounded as follows:

$$\begin{aligned} 2M \sum_{n=1}^{\infty} q_n \sum_{k=1}^{n} p_{x,k-1} &\leq 2M \cdot \left[\sum_{n=1}^{n_0} q_n \sum_{k=1}^{n_0} p_{x,k-1} \; + \sum_{n>n_0} q_n \sum_{k=1}^{n} p_{x,k-1} \right] \\ &\leq 2M \cdot \left[\sum_{k=1}^{n_0} p_{x,k-1} \; + \sum_{n>n_0} q_n \cdot n \right]. \end{aligned} \tag{2.55}$$

By (2.26),

$$\sum_{n>n_0} q_n \cdot n \leq \sum_{n>n_0} c_3 e^{-c_4 n} \cdot n = O(e^{-x^{1/4}}). \tag{2.56}$$

Also, using (2.38) with $\epsilon = 1/2$ shows that, for all $x > 2M$ and $0 \leq k \leq n_0$,

$$p_{x,k} \leq P^U_{x,s} \left(|U^-_k - x| > \frac{1}{2} x \right) \leq P^U_{x,s} \left(\max_{0 \leq j \leq n_0} |U^-_j - x| > \frac{1}{2} x \right) \leq c_1 n_0 e^{-c_2 n_0},$$

for some constants $c_1, c_2 > 0$. Hence,

$$\sum_{k=1}^{n_0} p_{x,k-1} \leq n_0 \cdot c_1 n_0 e^{-c_2 n_0} = O(e^{-x^{1/4}}). \tag{2.57}$$

Combining (2.54)-(2.57) shows that $\left| E^U_{x,s}(U_{\tau^S_s}) - x - \delta \cdot \mu_s \right| = O(e^{-x^{1/4}})$, which proves the lemma. □

Lemma 2.13. *As* $x \to \infty$,

$$\mathit{Var}^U_{x,s}(U_{\tau^S_s}) = 2x \cdot \mu_s + O(x^{1/2}) \quad \textit{and} \quad \mathit{Var}^V_{x,s}(V_{\tau^R_s}) = 2x \cdot \mu_s + O(x^{1/2}).$$

Proof. We will prove the statement about the variance of $U_{\tau^S_s}$. The proof of the analogous statement for the variance of $V_{\tau^R_s}$ is again very similar. A central element of the proof is the following claim.

Claim: There exists a non-negative random variable Δ with finite variance (defined on some outside probability space, separate from the (U^+_k) and (U^-_k) processes) such that

$$U^+_{\tau^S_s} - U^-_{\tau^S_s} \overset{stoch}{\leq} \Delta \;, \quad \text{under } P^U_{x,s}, \text{ for any } x \in \mathbb{N}_0. \tag{2.58}$$

Note that although the distributions of $U^+_{\tau^S_s}$ and $U^-_{\tau^S_s}$ do depend on the initial value $U^+_0 = U^-_0 = x$, the random variable Δ does not.

The proof of the claim will be given after the main proof of the lemma. The basic idea for the proof of the lemma is to approximate the process $(U_k)_{k \geq 0}$ by the process $(U^*_k)_{k \geq 0}$ defined by

$$U^*_0 = U_0 = x \quad \text{and} \quad U^*_{k+1} = \sum_{i=1}^{U^*_k} \mathcal{G}^{k+1}_i, \; k \geq 0$$

where the random variables $\mathcal{G}^k_i$ are as in (2.18). This process $(U^*_k)_{k \geq 0}$ is a standard Galton-Watson branching process with Geo(1/2) offspring distribution, independent of (S_k), and satisfies

$$U^-_k \leq U^*_k \leq U^+_k \;, \quad \text{for all } k \geq 0. \tag{2.59}$$

We express $U_{\tau_s^S}$ as

$$U_{\tau_s^S} = U^*_{\tau_s^S} + \widetilde{\Delta}, \text{ where } \widetilde{\Delta} \equiv U_{\tau_s^S} - U^*_{\tau_s^S}. \tag{2.60}$$

By (2.15) and (2.59), along with the claim (2.58), $|\widetilde{\Delta}| \overset{stoch}{\leq} \Delta$, for a random variable Δ with finite variance (not depending on x). The first term $U^*_{\tau_s^S}$ can be analyzed exactly. Since $(U_k^*)_{k\geq 0}$ is a standard Geo(1/2) Galton-Watson branching processes started from level x, we have

$$E^U_{x,s}(U^*_k) = x \text{ and } \mathrm{Var}^U_{x,s}(U^*_k) = 2xk$$

for any fixed $k > 0$. Thus, since the stopping time τ_s^S is independent of the process (U_k^*) and a.s. finite,

$$\mathrm{Var}^U_{x,s}(U^*_{\tau_s^S}) = E^U_{x,s}[\mathrm{Var}^U_{x,s}(U^*_{\tau_s^S}|\tau_s^S)] + \mathrm{Var}^U_{x,s}[E^U_{x,s}(U^*_{\tau_s^S}|\tau_s^S)] \tag{2.61}$$

$$= E^U_{x,s}(2x \cdot \tau_s^S) + \mathrm{Var}^U_{x,s}(x) = 2x\mu_s. \tag{2.62}$$

Expanding (2.60) gives,

$$\mathrm{Var}^U_{x,s}(U_{\tau_s^S}) = \mathrm{Var}^U_{x,s}(U^*_{\tau_s^S}) + \mathrm{Var}^U_{x,s}(\widetilde{\Delta}) + 2\mathrm{Cov}^U_{x,s}(U^*_{\tau_s^S}, \widetilde{\Delta}). \tag{2.63}$$

By the calculation above, the first term on the right hand side of (2.63) is exactly equal to $2x\mu_s$. The second two terms may be bounded as follows:

$$|\mathrm{Var}^U_{x,s}(\widetilde{\Delta})| \leq E^U_{x,s}(\widetilde{\Delta}^2) \leq \mathbb{E}(\Delta^2) \equiv C < \infty. \tag{2.64}$$

$$|2\mathrm{Cov}^U_{x,s}(U^*_{\tau_s^S}, \widetilde{\Delta})| \leq 2\left[\mathrm{Var}^U_{x,s}(U^*_{\tau_s^S})\right]^{1/2}\left[\mathrm{Var}^U_{x,s}(\widetilde{\Delta})\right]^{1/2} \leq 2 \cdot (2x\mu_s)^{1/2} \cdot C^{1/2}. \tag{2.65}$$

Combining (2.61)-(2.65) shows that $\mathrm{Var}^U_{x,s}(U_{\tau_s^S}) = 2x\mu_s + O(x^{1/2})$. Thus, it remains only to prove the claim (2.58)

Proof of Claim: Fix any $x \in \mathbb{N}_0$ and assume throughout that $U_0^+ = U_0^- = x$. Let $(B_k)_{k\geq 0}$ be a standard Galton-Watson branching processes with Geo(1/2) offspring distribution, started from $B_0 = 1$. Also, let $(\beta_{k,i})_{k\geq 0, i\in\mathbb{N}}$ and $(\tilde{\beta}_{k,i})_{k\geq 0, i\in\mathbb{N}}$ all be independent random variables such that $\beta_{k,i} \overset{law}{=} \tilde{\beta}_{k,i} \overset{law}{=} B_k$. Finally, let T be a random time with the same distribution as τ_s^S (under $P^U_{x,s}$) which is defined on the same probability space as the β and $\tilde{\beta}$ random variables, but independently of them. We will denote the probability measure for this probability space by $\mathbb{P}$, and the corresponding expectation operator by $\mathbb{E}$. We claim that

$$U_n^+ - U_n^- \overset{stoch}{\leq} \Delta_n \equiv \sum_{k=1}^{n}\sum_{i=1}^{M+1}\beta_{k,i} + \sum_{k=0}^{n-1}\sum_{i=1}^{M}\tilde{\beta}_{k,i}, \text{ for all } n \in \mathbb{N}. \tag{2.66}$$

Since $(U_n^+, U_n^-)_{n\in\mathbb{N}} \perp \tau_s^S$, $(\Delta_n)_{n\in\mathbb{N}} \perp T$, and $\tau_s^S \overset{law}{=} T$ it follows from this that

$$U^+_{\tau_s^S} - U^-_{\tau_s^S} \overset{stoch}{\leq} \Delta \equiv \sum_{n=1}^{\infty}\Delta_n \cdot \mathbb{1}\{T = n\}.$$

Direct computations using $\mathbb{E}(\beta_{k,i}) = \mathbb{E}(\tilde{\beta}_{k,i}) = 1$ and $\mathrm{Var}(\beta_{k,i}) = \mathrm{Var}(\tilde{\beta}_{k,i}) = 2k$, along with independence, give

$$\mathbb{E}(\Delta_n) = (2M+1)n \ , \ \mathrm{Var}(\Delta_n) = 2Mn^2 + n^2 + n \ , \ \mathbb{E}(\Delta_n^2) = (2M+1)(2M+2)n^2 + n.$$

Thus, since T has an exponential tail and the random variables Δ_n are independent of T,

$$\mathbb{E}(\Delta^2) = \sum_{n=1}^{\infty} \mathbb{P}(T=n)\mathbb{E}(\Delta^2|T=n) = \sum_{n=1}^{\infty} \mathbb{P}(T=n)\cdot\mathbb{E}(\Delta_n^2) < \infty.$$

So, it remains only to show (2.66).

To this end, recall again that the processes $(U_k^+)_{k\geq 0}$ and $(U_k^-)_{k\geq 0}$ may be represented in terms of the independent Geo(1/2) random variables $\mathcal{G}_i^k$ according to (2.18). Let us define a new process $(Z_k)_{k\geq 0}$ as follows:

$$Z_0 = x \quad \text{and} \quad Z_{k+1} = M + \sum_{i=1}^{\ell_k+1} \mathcal{G}_i^{k+1} \quad \text{where} \quad \ell_k = (U_k^- - M)^+ \ + \ M \ + \ (Z_k - U_k^-).$$

We claim that $Z_k \geq U_k^+$, for all $k \geq 0$. The proof is by induction on k. For $k=0$, we have $U_0^+ = Z_0 = x$, by assumption. Now, assume that $Z_k \geq U_k^+$. By the definition $\ell_k \geq Z_k$, so it follows that $\ell_k \geq U_k^+$, and therefore $Z_{k+1} = M + \sum_{i=1}^{\ell_k+1} \mathcal{G}_i^{k+1} \geq M + \sum_{i=1}^{U_k^+ +1} \mathcal{G}_i^{k+1} = U_{k+1}^+$. This shows that $Z_k \geq U_k^+$, for all k. So, $U_k^+ - U_k^- \leq Z_k - U_k^- \equiv W_k$, for all k. So, to establish (2.66) it will suffice to show that

$$W_n \overset{law}{=} \Delta_n \ , \quad \text{for all } n. \tag{2.67}$$

To prove (2.67) observe that $W_0 = 0$ and, for all $k \geq 0$,

$$W_{k+1} \ = \ \left[M + \sum_{i=1}^{\ell_k+1} \mathcal{G}_i^{k+1}\right] - \left[\sum_{i=1}^{(U_k^- - M)^+} \mathcal{G}_i^{k+1}\right] \ = \ M \ + \sum_{i=(U_k^- - M)^+ \ + \ 1}^{(U_k^- - M)^+ \ + \ W_k \ + \ (M+1)} \mathcal{G}_i^{k+1}.$$

Thus, the process $(W_k)_{k\geq 0}$ has the same distribution as the process $(\widetilde{W}_k)_{k\geq 0}$ defined by

$$\widetilde{W}_0 = 0 \quad \text{and} \quad \widetilde{W}_{k+1} = M \ + \sum_{i=1}^{\widetilde{W}_k + (M+1)} \widetilde{\mathcal{G}}_i^{k+1} \tag{2.68}$$

where $(\widetilde{\mathcal{G}}_i^k)_{i,k\in\mathbb{N}}$ are i.i.d. Geo(1/2) random variables (living on some probability space). To see this note that, for each $\vec{u} = (u_0, \ldots, u_n) \in \mathbb{N}_0^{n+1}$, occurrence of the event $E(\vec{u}) \equiv \{U_0^- = u_0, \ldots, U_n^- = u_n\}$ is a deterministic function of the random variables $\mathcal{G}_i^k$ such that $1 \leq k \leq n$ and, for each such k, $1 \leq i \leq (u_{k-1} - M)^+$. Thus, the conditional law of the random variables $\mathcal{G}_i^k$ such that $1 \leq k \leq n$ and $i > (u_{k-1} - M)^+$ on the event $E(\vec{u})$ is still i.i.d. Geo(1/2). Thus, conditional on the event $E(\vec{u})$, $(W_k)_{k=0}^n$ has the same law as $(\widetilde{W}_k)_{k=0}^n$. And since that holds for all $n \in \mathbb{N}$ and $\vec{u} = (u_0, \ldots, u_n)$ with $P_{x,s}^U(E(\vec{u})) > 0$, it follows that the entire process $(W_k)_{k\geq 0}$ has the same (unconditional) law as $(\widetilde{W}_k)_{k\geq 0}$.

Now, the process $(\widetilde{W}_k)$ defined by (2.68) may be understood as a type of branching process with migration where, at each step k, in addition to all the "regular children" created by the branching mechanism we also introduce $M+1$ immigrants prior to the reproduction stage (the $M+1$ in the upper limit of the sum) along with M additional immigrants after the reproduction stage (the M in front of the sum). It follows that, for each n,

$$\widetilde{W}_n \overset{law}{=} \Delta_n \equiv \sum_{k=1}^{n} \sum_{i=1}^{M+1} \beta_{k,i} + \sum_{k=0}^{n-1} \sum_{i=1}^{M} \tilde{\beta}_{k,i}. \tag{2.69}$$

The $\tilde{\beta}_{k,i}$ sum corresponds to the descendants of the M immigrants coming in after the reproduction stage in each generation that are alive at time n, and the $\beta_{k,i}$ sum corresponds to the descendants of the $(M+1)$ immigrants coming in prior to the reproduction stage in each generation that are alive at time n. Since $W_n \stackrel{law}{=} \widetilde{W}_n$ for all n, (2.69) implies (2.67). □

Our final result of this section is an "overshoot lemma" which gives concentration estimates for $(\widehat{V}_k)$ when exiting certain intervals. Analogous statements also hold for the process $(\widehat{U}_k)$, but we will not need these.

Lemma 2.14. *Let $\tau^{\widehat{V}}_{x^+} = \inf\{k > 0 : \widehat{V}_k \geq x\}$ and $\tau^{\widehat{V}}_{x^-} = \inf\{k > 0 : \widehat{V}_k \leq x\}$. Then there exist constants $c_8, c_9 > 0$ and $N \in \mathbb{N}$ such that for all $x \geq N$ the following hold:*

(i) $\sup_{0\leq z<x} P^V_{z,s}(\widehat{V}_{\tau^{\widehat{V}}_{x^+}} > x+y | \tau^{\widehat{V}}_{x^+} < \tau^{\widehat{V}}_0) \leq \begin{cases} c_8(1+y^{2/3}x^{-1/3})e^{-c_9 y^{2/3}x^{-1/3}}, & \text{for } 0 \leq y \leq x \\ c_8(1+y^{1/3})e^{-c_9 y^{1/3}}, & \text{for } y \geq x. \end{cases}$

(ii) $\sup_{x<z<4x} P^V_{z,s}(\widehat{V}_{\tau^{\widehat{V}}_{x^-} \wedge \tau^{\widehat{V}}_{(4x)^+}} < x-y) \leq c_8(1+y^{2/3}x^{-1/3})e^{-c_9 y^{2/3}x^{-1/3}}$, *for* $0 \leq y \leq x$.

Before proceeding to the main proof we isolate an important piece as its own lemma. For the statement of this lemma recall that $P^{V,\omega}_{v_0}$ is the probability measure for the backward branching process $(V_k)_{k\geq 0}$ defined by (2.6) in the particular environment $\omega \in \Omega$, started from $V_0 = v_0$.

Lemma 2.15. *There exist constants $C, c > 0$ such that for any environment ω satisfying $\omega(k,i) = 1/2$ for all $k \in \mathbb{Z}$ and $i > M$,*

$$P^{V,\omega}_z(V_1 > x+y | V_1 \geq x) \leq C(e^{-cy^2/x} + e^{-cy}) \tag{2.70}$$

for each $y \geq 6M$, $x \geq 2M+1$, and $0 \leq z < x$.

Proof. This is implicit in the proof of Lemma 5.1 in [8]. That lemma is stated for the process (V_k) under the averaged measure on environments in an (IID) and (BD) setting, rather than for a fixed environment ω. But the proof of the equivalent statement to (2.70) in this setting (given within the proof of Lemma 5.1 in [8]) uses only the fact that all cookies stacks are of bounded height M. □

Proof of Lemma 2.14. We will prove (i) under the assumption that $y \geq 12M$ and $x \geq N \equiv 2M+1$. By increasing the constant c_8 if necessary, the desired inequalities in (i) will then hold for all $y \geq 0$. A similar proof gives (ii) with some different values of the constants.

Under the measure $P^V_{z,s}$ we have $\widehat{V}_0 = V_0 = z < x$, so $\tau^{\widehat{V}}_{x^+} > 0$. Define a random variable k_0 as follows: Given that $\tau^{\widehat{V}}_{x^+} = j+1$, for some $j \in \mathbb{N}_0$, let $k_0 = \inf\{k > \tau_j : V_k \geq x\}$. In other words, $(j+1)$ is the first time ℓ that the process $(\widehat{V}_\ell)$ reaches level x or above, and k_0 is the first time $k > \tau_j$ that the process (V_k) reaches level x or above. Since the process (V_k, R_k) is Markovian under $P^V_{z,s}$ it follows that, for each $y \in \mathbb{N}$,

$$\begin{aligned}
&P^V_{z,s}(\widehat{V}_{\tau^{\widehat{V}}_{x^+}} > x+y | \tau^{\widehat{V}}_{x^+} < \tau^{\widehat{V}}_0) \\
&\leq \sup_{k\in\mathbb{N}, 0\leq w<x, r\in\mathcal{S}} P^V_{z,s}(\widehat{V}_{\tau^{\widehat{V}}_{x^+}} > x+y | \tau^{\widehat{V}}_{x^+} < \tau^{\widehat{V}}_0, k_0 = k, R_{k-1} = r, V_{k-1} = w) \\
&= \sup_{0\leq w<x, r\in\mathcal{S}} P^V_{w,r}(V_{\tau^R_s} > x+y | V_1 \geq x, V_{\tau^R_s} \geq x) \\
&= \sup_{0\leq w<x, r\in\mathcal{S}} \frac{P^V_{w,r}(V_{\tau^R_s} > x+y | V_1 \geq x)}{P^V_{w,r}(V_{\tau^R_s} \geq x | V_1 \geq x)}.
\end{aligned}$$

Thus, it will suffice to show the following two claims to establish the lemma.

Claim 1: There exist $C_1, C_2 > 0$ such that for all $y \geq 12M$, $x \geq 2M+1$, $0 \leq w < x$, and $r \in \mathcal{S}$

$$P^V_{w,r}(V_{\tau^R_s} > x+y | V_1 \geq x) \leq \begin{cases} C_1(1+y^{2/3}x^{-1/3})e^{-C_2 y^{2/3}x^{-1/3}}\ , & \text{if } y \leq x \\ C_1(1+y^{1/3})e^{-C_2 y^{1/3}}\ , & \text{if } y \geq x. \end{cases}$$

Claim 2: There exists $C_3 > 0$ such that for each $x \geq 2M+1$, $0 \leq w < x$, and $r \in \mathcal{S}$

$$P^V_{w,r}(V_{\tau^R_s} \geq x | V_1 \geq x) \geq C_3.$$

Proof of Claim 1. Fix any $x \geq 2M+1$, $0 \leq w < x$, and $r \in \mathcal{S}$. For each $y \geq 12M$ we have

$$\begin{aligned} P^V_{w,r}(V_{\tau^R_s} > x+y | V_1 \geq x) &= \frac{P^V_{w,r}(V_{\tau^R_s} > x+y, V_1 \geq x)}{P^V_{w,r}(V_1 \geq x)} \\ &= \frac{P^V_{w,r}(V_{\tau^R_s} > x+y, x \leq V_1 \leq x+y/2)}{P^V_{w,r}(V_1 \geq x)} + \frac{P^V_{w,r}(V_{\tau^R_s} > x+y, V_1 > x+y/2)}{P^V_{w,r}(V_1 \geq x)} \\ &\leq \frac{P^V_{w,r}(V_{\tau^R_s} > x+y, x \leq V_1 \leq x+y/2)}{P^V_{w,r}(x \leq V_1 \leq x+y/2)} + \frac{P^V_{w,r}(V_1 > x+y/2)}{P^V_{w,r}(V_1 \geq x)} \\ &= P^V_{w,r}(V_{\tau^R_s} > x+y | x \leq V_1 \leq x+y/2) \ + \ P^V_{w,r}(V_1 > x+y/2 | V_1 \geq x). \end{aligned} \tag{2.71}$$

By Lemma 2.15,

$$P^V_{w,r}(V_1 > x+y/2 | V_1 \geq x) \leq C(e^{-c\lfloor y/2 \rfloor^2/x} + e^{-c\lfloor y/2 \rfloor}) \tag{2.72}$$

for some $C, c > 0$. So, we need only bound $P^V_{w,r}(V_{\tau^R_s} > x+y | x \leq V_1 \leq x+y/2)$. Define $w_0 = \lfloor x+y/2 \rfloor$. By monotonicity of (V_k) with respect to its initial condition and Lemma 2.11,

$$\begin{aligned} P^V_{w,r}(V_{\tau^R_s} > x+y | x \leq V_1 \leq x+y/2) &\leq \sup_{r' \in \mathcal{S}} P^V_{w_0,r'}(V_{\tau^R_s} > x+y) \\ &\leq \sup_{r' \in \mathcal{S}} P^V_{w_0,r'}(|V_{\tau^R_s} - w_0| > y/2) \leq c_7(1+(y/2)^{2/3}w_0^{-1/3})e^{-c_6(y/2)^{2/3}w_0^{-1/3}}. \end{aligned}$$

Now, for $x \leq y$ we have $\frac{1}{8}y < (y/2)^2 w_0^{-1} < y$, and for $x \geq y$ we have $\frac{1}{8}y^2/x < (y/2)^2 w_0^{-1} < y^2/x$. So, it follows that

$$P^V_{w,r}(V_{\tau^R_s} > x+y | x \leq V_1 \leq x+y/2) \leq \begin{cases} c_7(1+y^{2/3}x^{-1/3})e^{-\frac{1}{2}c_6 y^{2/3}x^{-1/3}}\ , & \text{if } y \leq x \\ c_7(1+y^{1/3})e^{-\frac{1}{2}c_6 y^{1/3}}\ , & \text{if } y \geq x. \end{cases} \tag{2.73}$$

Combining (2.71), (2.72), and (2.73) proves the claim.

Proof of Claim 2. Let $(\widetilde{\mathcal{H}}_i)_{i \in \mathbb{N}}$ be i.i.d. Geo(1/2) random variables, and let $q = \inf_{y \geq 2M+1} \mathbb{P}\left(\sum_{i=1}^{y-M} \widetilde{\mathcal{H}}_i \geq y\right)$. By the central limit theorem, $\lim_{y \to \infty} \mathbb{P}\left(\sum_{i=1}^{y-M} \widetilde{\mathcal{H}}_i \geq y\right) = 1/2$, so $q > 0$. Further, since $(V^-_k)_{k \geq 0}$ is a Markov chain independent of $(R_k)_{k \geq 0}$ it follows from (2.15) and (2.19) that, for any $(r_0, \ldots, r_m) \in \mathcal{S}^{m+1}$ and $n \geq x \geq 2M+1$,

$$\begin{aligned} P^V_{n,r_0}(V_m \geq x | R_0 = r_0, \ldots, R_m = r_m) &\geq P^V_{n,r_0}(V^-_m \geq x | R_0 = r_0, \ldots, R_m = r_m) \\ = P^V_{n,r_0}(V^-_m \geq x) &\geq \prod_{j=1}^{m} P^V_{n,r_0}(V^-_j \geq V^-_{j-1} | V^-_{j-1} \geq \ldots \geq V^-_1 \geq V^-_0) \geq q^m. \end{aligned}$$

To complete the proof we note that by (2.26) there exist some $m_0 \in \mathbb{N}$ and $p > 0$ such that $\mathbf{P}_r(\tau^R_s \leq m_0) \geq p$, for each $r \in \mathcal{S}$. Therefore, for any $x \geq 2M+1$, $0 \leq w < x$, and $r' \in \mathcal{S}$,

$$P^V_{w,r'}(V_{\tau^R_s} \geq x | V_1 \geq x) \geq \inf_{n \geq x, r \in \mathcal{S}} P^V_{n,r}(V_{\tau^R_s} \geq x) \geq p \cdot q^{m_0} \equiv C_3 > 0. \qquad \square$$

3 Proof of Theorem 1.11

The proof of Theorem 1.11 relies on the following lemma concerning transience vs. recurrence of integer-valued Markov chains, which will be proved in Appendix B.

Lemma 3.1. *Let $(Z_k)_{k\geq 0}$ be an irreducible, time-homogeneous Markov chain on state space $\mathbb{N}_0$. Let $\mathbb{P}_x$ be the probability measure for (Z_k) started from $Z_0 = x$, and let $\mathbb{E}_x$ be the corresponding expectation operator. Assume that there exist constants $C_1, C_2 > 0$ and $x_0 \in \mathbb{N}$ such that:*

$$\mathbb{P}_x(|Z_1 - x| \geq \epsilon x) \leq C_1(1 + \epsilon^{2/3}x^{1/3})e^{-C_2\epsilon^{2/3}x^{1/3}} \ , \ \text{ for all } \ 0 < \epsilon \leq 1 \text{ and } x \geq x_0.$$
$$\mathbb{P}_x(|Z_1 - x| \geq \epsilon x) \leq C_1(1 + \epsilon^{1/3}x^{1/3})e^{-C_2\epsilon^{1/3}x^{1/3}} \ , \ \text{ for all } \ \epsilon \geq 1 \text{ and } x \geq x_0. \tag{3.1}$$

Define $\rho(x)$, $\nu(x)$, and $\theta(x)$ by

$$\rho(x) = \mathbb{E}_x(Z_1 - x) \ \ , \ \ \nu(x) = \frac{\mathbb{E}_x[(Z_1 - x)^2]}{x} \ \ , \ \ \theta(x) = \frac{2\rho(x)}{\nu(x)} \tag{3.2}$$

and assume also that $\liminf_{x\to\infty} \nu(x) > 0$. Then the following hold.

(i) If there exists some $a \in (1, \infty)$ such that $\theta(x) \leq 1 + \dfrac{1}{a\ln(x)}$ for all sufficiently large x, then the Markov chain (Z_k) is recurrent.

(ii) If there exists some $a \in (1, \infty)$ such that $\theta(x) \geq 1 + \dfrac{2a}{\ln(x)}$ for all sufficiently large x, then the Markov chain (Z_k) is transient.

Corollary 3.2. *Let $(Z_k)_{k\geq 0}$ be a time-homogeneous Markov chain on state space $\mathbb{N}_0$ and define $\rho(x)$, $\nu(x)$, and $\theta(x)$ as in Lemma 3.1. Assume that the concentration condition* (3.1) *is satisfied and $\liminf_{x\to\infty} \nu(x) > 0$. Assume also that*

$$\mathbb{P}_x(Z_1 = y) > 0, \text{ for all } x \geq 1 \text{ and } y \geq 0. \tag{3.3}$$

Then the following hold.

(i) If there exists some $a \in (1, \infty)$ such that $\theta(x) \leq 1 + \frac{1}{a\ln(x)}$ for all sufficiently large x, then $\mathbb{P}_z(Z_k > 0, \forall k > 0) = 0$, for each $z \geq 1$.

(ii) If there exists some $a \in (1, \infty)$ such that $\theta(x) \geq 1 + \frac{2a}{\ln(x)}$ for all sufficiently large x, then $\mathbb{P}_z(Z_k > 0, \forall k > 0) > 0$, for each $z \geq 1$.

Remark 3.3. Note that we do not assume in the corollary that the Markov chain (Z_k) is irreducible. State 0 may be absorbing, and in fact when we apply the corollary in the proof of Theorem 1.11 below it will be for the Markov chain $(\widehat{U}_k)$, where state 0 is absorbing.

Proof of Corollary 3.2. The Markov chain $(Z_k)_{k\geq 0}$ is not necessarily irreducible since state 0 may be absorbing, but the modified Markov chain $(\widetilde{Z}_k)_{k\geq 0}$ with transition probabilities

$$\widetilde{\mathbb{P}}_0(\widetilde{Z}_1 = 1) = 1 \quad \text{and} \quad \widetilde{\mathbb{P}}_x(\widetilde{Z}_1 = y) = \mathbb{P}_x(Z_1 = y), \, x \geq 1 \text{ and } y \geq 0$$

is certainly irreducible. Due to (3.3), this modified chain $(\widetilde{Z}_k)$, like the original chain (Z_k), has positive transition probabilities from each state $x \geq 1$ to each state $y \geq 0$. Thus, if $(\widetilde{Z}_k)$ is transient it has positive probability never to hit state 0 started from any state $z \geq 1$. Applying Lemma 3.1 shows that (i) and (ii) of the corollary hold for the Markov chain $(\widetilde{Z}_k)$. This implies (i) and (ii) also hold for (Z_k), since (Z_k) and $(\widetilde{Z}_k)$ have the same transition probabilities from all non-zero states. ☐

Before proceeding to the proof of Theorem 1.11 we isolate one important observation in the following remark. This observation will also be necessary for the proof of Theorem 1.12 in Section 4.1.

Remark 3.4 (Invariance of Assumption (A) under interchange of spatial directions)**.** Assume that Assumption (A) is satisfied for the probability measure $\mathbf{P}$, and let (S_k) be the associated stack sequence and $R_k = S_{-k}$, $k \in \mathbb{Z}$. Construct the *spatially reversed probability measure* $\widetilde{\mathbf{P}}$ on environments $\widetilde{\omega}$ by the following coupling: $\widetilde{\omega}(k,i) = \widetilde{S}_k(i)$, for $k \in \mathbb{Z}$ and $i \in \mathbb{N}$, where the process $(\widetilde{S}_k)$ is given by $\widetilde{S}_k(i) = 1 - S_{-k}(i) = 1 - R_k(i)$. Thus, the probability of jumping right (left) on the i-th visit to site k for the random walk $(\widetilde{X}_n)$ in the spatially reversed environment $\widetilde{\omega}$ is the same as the probability of jumping left (right) on the i-th visit to site $-k$ for the original random walk (X_n). With this construction there is also a natural coupling between the random walks $(\widetilde{X}_n)$ and (X_n), when both are started from site 0, such that $(\widetilde{X}_n)$ jumps right whenever (X_n) jumps left, and vice versa. This implies $\widetilde{X}_n = -X_n$, for all n, and so we may consider $(\widetilde{X}_n)$ as the spatially reversed version of the random walk (X_n).

Moreover, since the original model with probability measure $\mathbf{P}$ is assumed to satisfy Assumption (A), we know that (R_k) is a uniformly ergodic Markov chain, which implies $(\widetilde{S}_k)$ is also a uniformly ergodic Markov chain. Similarly, since the original model is assumed to satisfy Assumption (A), we know that (S_k) is a uniformly ergodic Markov chain, which implies $(\widetilde{R}_k)$ is a uniformly ergodic Markov chain (where $\widetilde{R}_k \equiv \widetilde{S}_{-k}$). Thus, the spatially reversed probability measure $\widetilde{\mathbf{P}}$ also satisfies Assumption (A). Finally, note that $\widetilde{\delta} = -\delta$, where $\widetilde{\delta}$ and δ are the drift parameters for the probability measures $\widetilde{\mathbf{P}}$ and $\mathbf{P}$ on environments.

Proof of Theorem 1.11. Since $\omega(k,i) = 1/2$ for each $i > M$ and $k \in \mathbb{Z}$, $\mathbf{P}_\pi$ a.s., we have

$$P_{0,\pi}(\liminf_{n\to\infty} X_n = k) = P_{0,\pi}(\limsup_{n\to\infty} X_n = k) = 0 \ , \quad \text{for each } k \in \mathbb{Z}.$$

It follows from this and Theorem 1.1 that one of the following must hold:

$$\text{(a) } P_{0,\pi}(X_n \to +\infty) = 1, \ \ \text{(b) } P_{0,\pi}(X_n \to -\infty) = 1, \ \text{or (c) } P_{0,\pi}(X_n = 0, i.o.) = 1.$$

Since the probability measure π places positive probability on each states $r \in \mathcal{S}$ we must have the same trichotomy under the measure P_0. That is, either

$$\text{(a')} \ P_0(X_n \to +\infty) = 1, \ \ \text{(b')} \ P_0(X_n \to -\infty) = 1, \ \text{or (c')} \ P_0(X_n = 0, i.o.) = 1$$

where (a) holds if and only if (a') holds, (b) holds if and only if (b') holds, and (c) holds if and only if (c') holds. We will show below that (a') holds if $\delta > 1$, and (a') does not hold if $\delta \leq 1$. It follows from this by interchanging spatial directions of the model, as described in Remark 3.4, that (b') holds if $\delta < -1$, and (b') does not hold if $\delta \geq -1$. Together these facts imply (c') must hold when $\delta \in [-1,1]$. Thus, it remains only to show the claim about (a') to establish the lemma.

To do this we will apply Corollary 3.2 to the Markov chain $(\widehat{U}_k)_{k\geq 0}$ with transition probabilities given by (2.13). We consider this Markov chain under the family of measures $P^U_{x,s}$, $x \in \mathbb{N}_0$, so that $\widehat{U}_0 = U_0 = x$ deterministically. Thus, the measure $P^U_{x,s}$ for the Markov chain $(\widehat{U}_k)$ is the equivalent of the measure $\mathbb{P}_x$ for the Markov chain (Z_k) considered in the corollary. Condition (3.3) of the corollary is satisfied since the cookie stacks $r \in \mathcal{S}$ are all elliptic. Also, the concentration condition (3.1) is satisfied due to Lemma 2.11. Finally, by Lemmas 2.12 and 2.13,

$$\rho(x) = \delta \cdot \mu_s + O(e^{-x^{1/4}}) \quad \text{and} \quad \nu(x) = 2\mu_s + O(x^{-1/2}).$$

Thus, $\liminf_{x\to\infty} \nu(x) > 0$, as required by the corollary, and $\theta(x) = 2\rho(x)/\nu(x) = \delta + O(x^{-1/2})$. Applying the corollary with $z = 1$ we have

$$P^U_{1,s}(\widehat{U}_k > 0, \forall k > 0) > 0 \text{ , if } \delta > 1 \quad \text{ and } \quad P^U_{1,s}(\widehat{U}_k > 0, \forall k > 0) = 0 \text{ , if } \delta \le 1.$$

Since 0 is an absorbing state for the process $(U_k)_{k\ge0}$ it follows also that

$$P^U_{1,s}(U_k > 0, \forall k > 0) > 0 \text{ , if } \delta > 1 \quad \text{ and } \quad P^U_{1,s}(U_k > 0, \forall k > 0) = 0 \text{ , if } \delta \le 1.$$

Hence, by Lemma 2.3,

$$P_0(X_n \to +\infty) = 1 \text{ , if } \delta > 1 \quad \text{ and } \quad P_0(X_n \to +\infty) = 0 \text{ , if } \delta \le 1$$

which establishes the claim about (a’). □

4 Proof of Theorems 1.12 and 1.13

The proofs of Theorems 1.12 and 1.13 are based on an analysis of the backward branching process $(V_k)_{k\ge0}$, and follow the general strategy used in [8]. However, there is an additional complication due to our different probability measure **P** on environments. If **P** is (IID), as considered in [8], then the process $(V_k)_{k\ge0}$ is Markovian. However, if **P** is Markovian, as we consider, then the process $(V_k)_{k\ge0}$ is not. Thus, we will instead analyze the modified process $(\widehat{V}_k)_{k\ge0}$, which is Markovian, and then translate the results from the $(\widehat{V}_k)$ process back to the (V_k) process. The outline for our proofs and the remainder of this section is described below.

Step 1: In Appendix C we will establish the following two propositions which are analogous to Theorems 2.1 and 2.2 of [8]. The general methodology is very similar to that in [8], so we will provide only a general outline and reprove a few key lemmas from which everything else follows just as in [8].

Proposition 4.1. *If $\delta > 1$ then there exists some $c_{10} > 0$ such that* $P^V_{0,s}\left(\tau^{\widehat{V}}_0 > x\right) \sim c_{10} \cdot x^{-\delta}$.

Proposition 4.2. *If $\delta > 1$ then there exists some $c_{11} > 0$ such that* $P^V_{0,s}\left(\sum_{k=0}^{\tau^{\widehat{V}}_0} \widehat{V}_k > x\right) \sim c_{11} \cdot x^{-\delta/2}$.

Step 2: Define $\sigma^V_0 = \inf\{k > 0 : V_k = 0 \text{ and } R_k = s\}$. In Appendix D we will translate the above results for the process $(\widehat{V}_k)$ to the process (V_k) proving the following propositions.

Proposition 4.3. *If $\delta > 1$ then $P^V_{0,s}\left(\sigma^V_0 > x\right) \sim c_{12} \cdot x^{-\delta}$, where $c_{12} \equiv c_{10} \cdot \mu^\delta_s$.*

Proposition 4.4. *If $\delta > 1$ then $P^V_{0,s}\left(\sum_{k=0}^{\sigma^V_0} V_k > x\right) \sim c_{13} \cdot x^{-\delta/2}$, where $c_{13} \equiv c_{11} \cdot \mu^{\delta/2}_s$.*

Step 3: In Section 4.1 we will use Propositions 4.3 and 4.4 to prove Theorem 1.12.

Step 4: In Section 4.2 we will use Propositions 4.3 and 4.4 to prove Theorem 1.13.

For future reference we observe the following simple corollary of Proposition 4.3.

Corollary 4.5. *If $\delta > 1$ then, for each $x \in \mathbb{N}_0$ and $r \in \mathcal{S}$, $P^V_{x,r}\left(\sigma^V_0 < \infty\right) = 1$.*

Proof. The process $(V_k, R_k)_{k\ge0}$ is a time-homogeneous Markov chain under $P^V_{x,r}$, for any x, r. Furthermore since the Markov chain (R_k) is irreducible and the cookie stacks in $\mathcal{S}$ are elliptic this Markov chain on pairs (V_k, R_k) is also irreducible. By Propostion 4.3 the pair $(0, s)$ is a recurrent state for the Markov chain (V_k, R_k), so the Markov chain itself is recurrent, so the hitting time of state $(0, s)$ is a.s. finite starting from any initial state (x, r). □

4.1 Proof of Theorem 1.12

Define stopping times $(\sigma_i)_{i\geq 0}$ by

$$\sigma_0 = \inf\{k \geq 0 : V_k = 0, R_k = s\} \quad \text{and} \quad \sigma_{i+1} = \inf\{k > \sigma_i : V_k = 0, R_k = s\}\ ,\ i \geq 0.$$

Also, define $\Delta_{\sigma,i} = \sigma_i - \sigma_{i-1}$, for $i \geq 1$, and let

$$Q_0 = \sum_{k=0}^{\sigma_0} V_k \quad \text{and} \quad Q_i = \sum_{k=\sigma_{i-1}+1}^{\sigma_i} V_k\ ,\ i \geq 1.$$

Since $(V_k, R_k)_{k\geq 0}$ is a time-homogeneous Markov chain (under any of the measures P_x^V, $P_{x,\pi}^V$, or $P_{x,r}^V$), it follows from Corollary 4.5 that if $\delta > 1$ then (under any of these same measures)

$$\begin{aligned}&\text{The times } \sigma_i,\ i \geq 0, \text{ are all a.s. finite and}\\ &(Q_i)_{i\geq 1} \text{ and } (\Delta_{\sigma,i})_{i\geq 1} \text{ are each i.i.d sequences.}\end{aligned} \tag{4.1}$$

We denote the mean of the Q_i's by μ_Q and the mean of the $\Delta_{\sigma,i}$'s by μ_σ.

Proof of Theorem 1.12. By Theorem 1.2 there exists some deterministic $v \in [-1,1]$ such that $X_n/n \to v$, $P_{0,\pi}$ a.s. Since the probability measure π places positive probability on each states $r \in \mathcal{S}$, it follows that $X_n/n \to v$, $P_{0,r}$ a.s., for each $r \in \mathcal{S}$. Hence, also $X_n/n \to v$, P_0 a.s. By Theorem 1.11, the walk is recurrent for $\delta \in [-1,1]$, so we must have $v = 0$ in this case. We will show that

$$v > 0 \ \text{ if } \ \delta > 2 \quad \text{ and } \quad v = 0 \ \text{ if } \ \delta \in (1,2]. \tag{4.2}$$

It follows from this by interchanging spatial directions of the model (see Remark 3.4) that $v < 0$ if $\delta < -2$ and $v = 0$ if $\delta \in [-2,-1)$.

For the remainder of the proof we assume $\delta > 1$. For $n \in \mathbb{N}$, define i_n by

$$i_n = \begin{cases} \sup\{i \geq 0 : \sigma_i \leq n\}\ , & \text{if } n \geq \sigma_0 \\ -1\ , & \text{if } n < \sigma_0. \end{cases}$$

Observe that (with the convention $\sigma_{-1} \equiv 0$)

$$\sigma_{i_n} \leq n < \sigma_{i_n+1} \quad \text{and} \quad \sum_{j=1}^{i_n} Q_j \leq \sum_{k=0}^{n} V_k \leq \sum_{j=0}^{i_n+1} Q_j\ , \ \text{ for each } n \in \mathbb{N}. \tag{4.3}$$

Thus, by (4.1) and the strong law of large numbers,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} V_k \leq \limsup_{n\to\infty} \frac{i_n}{\sigma_{i_n}} \cdot \left[\frac{1}{i_n}\sum_{j=0}^{i_n+1} Q_j\right] = \limsup_{i\to\infty} \frac{i}{\sigma_i} \cdot \left[\frac{1}{i}\sum_{j=0}^{i+1} Q_j\right] = \frac{\mu_Q}{\mu_\sigma}\ ,\ P_{0,\pi}^V \text{ a.s.}$$

and

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} V_k &\geq \liminf_{n\to\infty} \frac{i_n+1}{\sigma_{i_n+1}} \cdot \left[\frac{1}{i_n+1}\sum_{j=1}^{i_n} Q_j\right] \\ &= \liminf_{i\to\infty} \frac{i+1}{\sigma_{i+1}} \cdot \left[\frac{1}{i+1}\sum_{j=1}^{i} Q_j\right] = \frac{\mu_Q}{\mu_\sigma}\ ,\ P_{0,\pi}^V \text{ a.s.}\end{aligned}$$

So, $\frac{1}{n}\sum_{k=0}^n V_k \to \frac{\mu_Q}{\mu_\sigma}$ a.s. (and in probability) under $P_{0,\pi}^V$. Hence, by Lemma 2.4, $\frac{1}{n}\sum_{k=0}^n D_{n,k} \to \frac{\mu_Q}{\mu_\sigma}$, in probability under $P_{0,\pi}$. Now, since the random walk (X_n) is right transient with $\delta > 1$, $\limsup_{n\to\infty}\sum_{k<0} D_{n,k} < \infty$, $P_{0,\pi}$ a.s. So, it follows from (2.8) that $\tau_n^X/n \to 1 + \frac{2\mu_Q}{\mu_\sigma}$ in probability under $P_{0,\pi}$. Since we know a priori that $X_n/n \to v$, $P_{0,\pi}$ a.s. (for some unknown $v \in [-1,1]$), this in fact implies $\tau_n^X/n \to 1 + \frac{2\mu_Q}{\mu_\sigma}$, $P_{0,\pi}$ a.s., and

$$v = \frac{1}{1 + 2\frac{\mu_Q}{\mu_\sigma}}. \tag{4.4}$$

Now, by Proposition 4.3, μ_σ is finite for all $\delta > 1$, and by Proposition 4.4, μ_Q is finite for $\delta > 2$ but infinite for $\delta \in (1,2]$. So, it follows that (4.2) holds. □

4.2 Proof of Theorem 1.13

Throughout Section 4.2 the random variables $Z_{\alpha,b}$ are as in (1.3), the random variables Q_i, σ_i, $\Delta_{\sigma,i}$ and i_n are as in Section 4.1, and $m_n \equiv \lfloor n/\mu_\sigma \rfloor$. The general proof strategy for Theorem 1.13 will be to first prove a limiting distribution for $\sum_{k=0}^n V_k$, then translate to a limiting distribution for the hitting times τ_n^X using (2.8) and Lemma 2.4, then translate to a limiting distribution for the walk (X_n) itself. This basic approach has been used before in [3, 8, 9], and our methods will be quite similar to these works, but the details differ a bit because the process (V_k) is not Markovian when the environment is not (IID). Thus, we must consider renewal times $(\sigma_i)_{i\geq 0}$, rather than simply successive times at which $V_k = 0$. Also, we have the additional minor complication of the Q_0 and σ_0 terms to deal with (which would be 0 in the (IID) case).

Unless other specified it is assumed throughout that $V_0 = 0$ and the probability measure on environments is $\mathbf{P}_\pi$ (rather than $\mathbf{P}$). Everything will be proved initially under the stationary measure $\mathbf{P}_\pi$, and then at the very end after proving Theorem 1.13 under the stationary measure $\mathbf{P}_\pi$ we will translate the result to the given measure $\mathbf{P}$.

To state our first lemma for the limiting distribution of $\sum_{k=0}^n V_k$ we first need to introduce a little notation. Let $\mu_Q(t)$ be the truncated expectation of the random variables Q_i: $\mu_Q(t) \equiv E_{0,\pi}^V[Q_i \cdot \mathbb{1}_{\{Q_i \leq t\}}]$, $i \geq 1$, where $E_{0,\pi}^V(\cdot)$ is expectation with respect to the probability measure $P_{0,\pi}^V$. Also, let $Z_{1,b,c}$ be a random variable with characteristic function

$$E[e^{itZ_{1,b,c}}] = \exp\left[itc - b|t|\left(1 + \frac{2i}{\pi}\log|t|\mathrm{sgn}(t)\right)\right], \tag{4.5}$$

for $b > 0$ and $c \in \mathbb{R}$. For future reference we note the following scaling relations hold for the stable random variables $Z_{\alpha,b}$ and $Z_{1,b,c}$:

$$aZ_{\alpha,b} \overset{d.}{=} Z_{\alpha,ba^\alpha}\ , \quad \text{for all } \alpha \in (0,1)\cup(1,2], b > 0, a > 0. \tag{4.6}$$

$$a_1 Z_{1,b,c} + a_2 \overset{d.}{=} Z_{1,ba_1,[ca_1 + a_2 - \frac{2}{\pi}ba_1\log(a_1)]}\ , \quad \text{for all } b > 0,\ c \in \mathbb{R},\ a_1 > 0,\ a_2 \in \mathbb{R}. \tag{4.7}$$

Also, we note that μ_σ is finite for all $\delta > 1$ by Propositions 4.3, and μ_Q is finite for all $\delta > 2$ by Proposition 4.4.

Lemma 4.6. *Under the probability measure $P_{0,\pi}^V$ for the process $(V_k)_{k\geq 0}$ the following hold:*

(i) If $\delta \in (1,2)$ then there is some $b > 0$ such that $\frac{\sum_{k=0}^n V_k}{n^{2/\delta}} \xrightarrow{d.} Z_{\delta/2,b}$, as $n \to \infty$.

(ii) If $\delta = 2$ then there are constants $b > 0$ and $c \in \mathbb{R}$ such that $\frac{\sum_{k=0}^n V_k - \frac{\mu_Q(n/\mu_\sigma)}{\mu_\sigma}n}{n} \xrightarrow{d.} Z_{1,b,c}$, as $n \to \infty$.

(iii) *If* $\delta \in (2,4)$ *then there is some* $b > 0$ *such that* $\frac{\sum_{k=0}^{n} V_k - \frac{\mu_Q}{\mu_\sigma} n}{n^{2/\delta}} \xrightarrow{d.} Z_{\delta/2,b}$*, as* $n \to \infty$*.*

(iv) *If* $\delta = 4$ *then there is some* $b > 0$ *such that* $\frac{\sum_{k=0}^{n} V_k - \frac{\mu_Q}{\mu_\sigma} n}{[n\log(n)]^{1/2}} \xrightarrow{d.} Z_{2,b}$*, as* $n \to \infty$*.*

(v) *If* $\delta > 4$ *then there is some* $b > 0$ *such that* $\frac{\sum_{k=0}^{n} V_k - \frac{\mu_Q}{\mu_\sigma} n}{n^{1/2}} \xrightarrow{d.} Z_{2,b}$*, as* $n \to \infty$*.*

To prove Lemma 4.6 we will need two general results about the limiting distributions of sums of i.i.d. random variables. The first result is a particular case of convergence to stable distributions for sums of i.i.d. random variables with regularly varying tails. The second result concerns sums of a random number of i.i.d. random variables with finite variance.

Theorem 4.7 (Special case of Theorem 1, page 172 (for $\alpha \geq 2$) and Theorem 2, page 175 (for $\alpha < 2$) in [5])**.** *Let* Z *be a random variable with distribution such that:*

$$\mathbb{P}(Z > x) \sim Cx^{-\alpha}, \text{ as } x \to \infty, \text{ for some constants } C > 0 \text{ and } \alpha > 0.$$
$$\mathbb{P}(Z < x_0) = 0, \text{ for some } x_0 \in (-\infty, 0].$$

Also, let $Z_1, Z_2, \ldots$ *be i.i.d. random variables distributed as* Z*. Then the following hold:*

(i) *If* $\alpha \in (0,1)$, $\frac{\sum_{k=1}^{n} Z_k}{n^{1/\alpha}} \xrightarrow{d.} Z_{\alpha,b}$*, for some* $b > 0$*.*

(ii) *If* $\alpha = 1$, $\frac{\sum_{k=1}^{n} Z_k - n \cdot \mathbb{E}[Z \cdot \mathbb{1}_{\{Z \leq n\}}]}{n} \xrightarrow{d.} Z_{1,b,c}$*, for some* $b > 0$ *and* $c \in \mathbb{R}$*.*

(iii) *If* $\alpha \in (1,2)$, $\frac{\sum_{k=1}^{n} Z_k - n \cdot \mathbb{E}(Z)}{n^{1/\alpha}} \xrightarrow{d.} Z_{\alpha,b}$*, for some* $b > 0$*.*

(iv) *If* $\alpha = 2$, $\frac{\sum_{k=1}^{n} Z_k - n \cdot \mathbb{E}(Z)}{[n\log(n)]^{1/2}} \xrightarrow{d.} Z_{2,b}$*, for some* $b > 0$*.*

(v) *If* $\alpha > 2$, $\frac{\sum_{k=1}^{n} Z_k - n \cdot \mathbb{E}(Z)}{n^{1/2}} \xrightarrow{d.} Z_{2,b}$*, for some* $b > 0$*.*

Theorem 4.8 (Theorem 3.1, page 17 in [6])**.** *Let* $(Z_k)_{k\geq 1}$ *be i.i.d. random variables with* $\mathbb{E}(Z_k) = 0$ *and* $Var(Z_k) = \sigma_Z^2 \in (0,\infty)$*. Also, let* $(N_n)_{n\geq 1}$ *be a sequence of random variables such that* $N_n/n \to \theta$*, in probability, for some* $\theta \in (0,\infty)$*. Then* $\left(\sum_{k=1}^{N_n} Z_k\right)/\left(\sigma_Z\sqrt{n\theta}\right) \xrightarrow{d.} N(0,1)$*, as* $n \to \infty$*.*

In addition to these two theorems we will also need the following lemma for the proof of part (v) of Lemma 4.6.

Lemma 4.9. *For any* $\delta > 1$ *the following hold:*

(i) $i_n/n \to 1/\mu_\sigma$ *a.s. under* $P_{0,\pi}^V$*.*

(ii) *There exists some constant* $c_{14} > 0$ *such that:*

$$\lim_{n\to\infty} P_{0,\pi}^V(n - \sigma_{i_n} > k) \leq c_{14} \cdot k^{1-\delta}, \quad \text{for each } k \in \mathbb{N}.$$
$$\lim_{n\to\infty} P_{0,\pi}^V(\sigma_{i_n+1} - n > k) \leq c_{14} \cdot k^{1-\delta}, \quad \text{for each } k \in \mathbb{N}.$$

Proof of Lemma 4.9 (i). Since $\Delta_{\sigma,i}$, $i \geq 1$, are i.i.d., the times $(\sigma_i)_{i\geq 0}$ are the renewal times for a (delayed) renewal process. By definition i_n is the number of renewals up to time n (excluding σ_0). So, by [19, Proposition 3.5.1] $i_n/n \to 1/\mu_\sigma$ a.s. □

Proof of Lemma 4.9 (ii). Since the Markov chain (R_k) is uniformly ergodic it is aperiodic, which implies the Markov chain on pairs (V_k, R_k) is aperiodic (due to ellipticity of the cookie stacks), which implies the discrete renewal process with renewal times $(\sigma_i)_{i\geq 0}$ is itself aperiodic. Let $A_{m,k}$ be the event that there are no renewal times σ_i in $\{m, m{+}1, \dots, m{+}k{-}1\}$ and let Δ_σ be a random variable (on some probability space) with the common distribution of the random variables $\Delta_{\sigma,i}$, $i \geq 1$. It follows from aperiodicity and [19, Example 4.3 (C)] that $P_{0,\pi}^V(A_{m,k}) \to \mathbb{E}[(\Delta_\sigma - k)^+]/\mathbb{E}[\Delta_\sigma]$ as $m \to \infty$, for each fixed $k \in \mathbb{N}$. Moreover, by Proposition 4.3, $\mathbb{E}[(\Delta_\sigma - k)^+]/\mathbb{E}[\Delta_\sigma] \leq c_{14} \cdot k^{1-\delta}$, for some $c_{14} > 0$ and all $k \in \mathbb{N}$. Thus,

$$\begin{aligned}\lim_{n\to\infty} P_{0,\pi}^V(n - \sigma_{i_n} > k) &= \lim_{n\to\infty} P_{0,\pi}^V(\text{no renewals times in } \{n-k, n-k+1, \dots, n\}) \\ &= \lim_{n\to\infty} P_{0,\pi}^V(A_{n-k,k+1}) \leq \lim_{n\to\infty} P_{0,\pi}^V(A_{n-k,k}) \leq c_{14} \cdot k^{1-\delta}.\end{aligned}$$

and

$$\begin{aligned}\lim_{n\to\infty} P_{0,\pi}^V(\sigma_{i_n+1} - n > k) &= \lim_{n\to\infty} P_{0,\pi}^V(\text{no renewals times in } \{n+1, \dots, n+k\}) \\ &= \lim_{n\to\infty} P_{0,\pi}^V(A_{n+1,k}) \leq c_{14} \cdot k^{1-\delta}.\end{aligned}$$ □

Proof of Lemma 4.6. The proofs of the various parts of the lemma will be given separately, but we begin with one important general observation that is necessary for the proof of several parts. If $1 \leq \ell \leq m_n$ and $|\sigma_{m_n} - n| \leq \ell$, then

$$\left|\sum_{k=0}^{n} V_k - \sum_{i=1}^{m_n} Q_i\right| \;\leq\; Q_0 + \sum_{k=\sigma_{m_n}-\ell}^{\sigma_{m_n}+\ell} V_k \;\leq\; Q_0 + \sum_{k=\sigma_{(m_n-\ell)}}^{\sigma_{(m_n+\ell)}} V_k \;=\; Q_0 + \sum_{i=m_n-\ell+1}^{m_n+\ell} Q_i$$

since the random variables V_k are nonnegative and $\sigma_{m_n-\ell} \leq \sigma_{m_n} - \ell < \sigma_{m_n} + \ell \leq \sigma_{m_n+\ell}$. Thus, if (a_n) and (b_n) are any sequences of positive real numbers such that $b_n \to \infty$ and $a_n \to \infty$ with $a_n = o(n)$, then

$$\begin{aligned}&\limsup_{n\to\infty}\; P_{0,\pi}^V\Big(\Big|\sum_{k=0}^{n} V_k - \sum_{i=1}^{m_n} Q_i\Big| > 2b_n\Big) \\ &\quad\leq \limsup_{n\to\infty}\left[P_{0,\pi}^V(|\sigma_{m_n} - n| > a_n) + P_{0,\pi}^V(Q_0 > b_n) + P_{0,\pi}^V\Big(\sum_{i=m_n-\lfloor a_n\rfloor+1}^{m_n+\lfloor a_n\rfloor} Q_i > b_n\Big)\right] \\ &\quad\leq \limsup_{n\to\infty} P_{0,\pi}^V(|\sigma_{m_n} - n| > a_n) + \limsup_{n\to\infty} P_{0,\pi}^V\Big(\sum_{i=1}^{2\lfloor a_n\rfloor} Q_i > b_n\Big). \end{aligned} \tag{4.8}$$

In the last line we have used the fact that the random variables $(Q_i)_{i\geq 1}$ are i.i.d., and also the fact that Q_0 is $P_{0,\pi}^V$ a.s. finite (due to Corollary 4.5), which implies $P_{0,\pi}^V(Q_0 > b_n) \to 0$, if $b_n \to \infty$.

Proof of (i):
With $\delta \in (1,2)$ it follows from Proposition 4.3 and Theorem 4.7-(iii) that $\frac{\sum_{i=1}^m \Delta_{\sigma,i} - m\mu_\sigma}{m^{1/\delta}} \xrightarrow{d.} Z_{\delta,B_1}$, as $m \to \infty$, for some $B_1 > 0$. Also, since σ_0 is a.s. finite $\frac{\sigma_0 + (m_n\mu_\sigma - n)}{m_n^{1/\delta}} \to 0$ a.s, as $n \to \infty$. Hence,

$$\frac{\sigma_{m_n} - n}{m_n^{1/\delta}} = \left[\frac{\sigma_0 + (m_n\mu_\sigma - n)}{m_n^{1/\delta}} + \frac{\sum_{i=1}^{m_n} \Delta_{\sigma,i} - m_n\mu_\sigma}{m_n^{1/\delta}}\right] \xrightarrow{d.} Z_{\delta,B_1},\ \text{as } n \to \infty. \tag{4.9}$$

Further, by Proposition 4.4 and Theorem 4.7-(i), there is some $B_2 > 0$ such that

$$\frac{\sum_{i=1}^m Q_i}{m^{2/\delta}} \xrightarrow{d.} Z_{\delta/2,B_2},\ \text{as } m \to \infty. \tag{4.10}$$

Now,

$$\frac{\sum_{k=0}^n V_k}{n^{2/\delta}} = \left(\frac{m_n^{2/\delta}}{n^{2/\delta}} \cdot \frac{\sum_{i=1}^{m_n} Q_i}{m_n^{2/\delta}}\right) + \frac{\sum_{k=0}^n V_k - \sum_{i=1}^{m_n} Q_i}{n^{2/\delta}} \equiv (I) + (II).$$

Since $\frac{m_n^{2/\delta}}{n^{2/\delta}} \to (1/\mu_\sigma)^{2/\delta}$ it follows from (4.10) and (4.6) that $(I) \xrightarrow{d.} Z_{\delta/2,b}$, for some $b > 0$. So it will suffice to show that $(II) \xrightarrow{p.} 0$. To do this, fix an arbitrary $\epsilon > 0$. Applying (4.8) with $a_n = n^p$, $p \in (1/\delta, 1)$, and $b_n = \epsilon n^{2/\delta}$ gives

$$\begin{aligned}
&\limsup_{n\to\infty} P_{0,\pi}^V\Big(\Big|\sum_{k=0}^n V_k - \sum_{i=1}^{m_n} Q_i\Big| > 2\epsilon n^{2/\delta}\Big) \\
&\qquad \le \limsup_{n\to\infty} P_{0,\pi}^V(|\sigma_{m_n} - n| > n^p) + \limsup_{n\to\infty} P_{0,\pi}^V\Big(\sum_{i=1}^{2\lfloor n^p\rfloor} Q_i > \epsilon n^{2/\delta}\Big).
\end{aligned}$$

The first term on the right hand side is 0 by (4.9), and the second term on the right hand side is 0 by (4.10). Since $\epsilon > 0$ is arbitrary, it follows that $(II) \xrightarrow{p.} 0$.

Proof of (ii):
With $\delta = 2$ it follows from Proposition 4.4 that $P_{0,\pi}^V(Q_i > x) \sim c_{13} x^{-1}$, for $i \ge 1$, which implies

$$\mu_Q(t) \sim c_{13}\log(t), \text{ as } t \to \infty \quad \text{and} \quad \lim_{n\to\infty} [\mu_Q(n/\mu_\sigma) - \mu_Q(m_n)] = 0. \tag{4.11}$$

Also, similar arguments as in the proof of part (i) of the lemma using Propositions 4.3 and 4.4 along with parts (ii) and (iv) of Theorem 4.7 show that there are constants $B_1, B_2 > 0$ and $C \in \mathbb{R}$ such that:

$$\frac{\sigma_{m_n} - n}{\sqrt{m_n \log(m_n)}} \xrightarrow{d.} Z_{2,B_1}, \text{ as } n \to \infty. \tag{4.12}$$

$$\frac{\sum_{i=1}^m Q_i - m\mu_Q(m)}{m} \xrightarrow{d.} Z_{1,B_2,C}, \text{ as } m \to \infty. \tag{4.13}$$

Now,

$$\begin{aligned}
&\frac{\sum_{k=0}^n V_k - \frac{\mu_Q(n/\mu_\sigma)}{\mu_\sigma} n}{n} \\
&= \left(\frac{m_n}{n} \cdot \frac{\sum_{i=1}^{m_n} Q_i - m_n\mu_Q(m_n)}{m_n}\right) + \frac{m_n\mu_Q(m_n) - \frac{n}{\mu_\sigma}\mu_Q(n/\mu_\sigma)}{n} + \frac{\sum_{k=0}^n V_k - \sum_{i=1}^{m_n} Q_i}{n} \\
&\equiv (I) + (II) + (III).
\end{aligned}$$

Since $m_n/n \to 1/\mu_\sigma$ it follows from (4.13) and (4.7) that $(I) \xrightarrow{d.} Z_{1,b,c}$, for some $b > 0$ and $c \in \mathbb{R}$. Also, by (4.11), $(II) \to 0$ (deterministically). So, it will suffice to show that $(III) \xrightarrow{p.} 0$. To do this, fix an arbitrary $\epsilon > 0$. Applying (4.8) with $a_n = n^p$, $p \in (1/2, 1)$, and $b_n = \epsilon n$ gives

$$\begin{aligned}
&\limsup_{n\to\infty} P_{0,\pi}^V\Big(\Big|\sum_{k=0}^n V_k - \sum_{i=1}^{m_n} Q_i\Big| > 2\epsilon n\Big) \\
&\qquad \le \limsup_{n\to\infty} P_{0,\pi}^V(|\sigma_{m_n} - n| > n^p) + \limsup_{n\to\infty} P_{0,\pi}^V\Big(\sum_{i=1}^{2\lfloor n^p\rfloor} Q_i > \epsilon n\Big).
\end{aligned}$$

The first term on the right hand side is 0 by (4.12), and the second term on the right hand side is 0 by (4.11) and (4.13). Since $\epsilon > 0$ is arbitrary, it follows that $(III) \xrightarrow{p.} 0$.

Proof of (iii) and (iv):
Define $d(n) = n^{2/\delta}$ if $\delta \in (2,4)$ and $d(n) = [n\log(n)]^{1/2}$ if $\delta = 4$. We wish to show that

$$\frac{\sum_{k=0}^{n} V_k - \frac{\mu_Q}{\mu_\sigma} n}{d(n)} \xrightarrow{d.} Z_{\delta/2,b}\ , \quad \text{for some } b > 0.$$

Similar arguments as in the proof of part (i) of the lemma using Propositions 4.3 and 4.4 along with parts (iii), (iv), and (v) of Theorem 4.7 show that there are constants $B_1, B_2 > 0$ such that:

$$\frac{\sigma_{m_n} - n}{m_n^{1/2}} \xrightarrow{d.} Z_{2,B_1},\ \text{as } n \to \infty. \tag{4.14}$$

$$\frac{\sum_{i=1}^{m} Q_i - m\mu_Q}{d(m)} \xrightarrow{d.} Z_{\delta/2,B_2},\ \text{as } m \to \infty. \tag{4.15}$$

Now,

$$\begin{aligned}\frac{\sum_{k=0}^{n} V_k - \frac{\mu_Q}{\mu_\sigma} n}{d(n)} &= \left(\frac{d(m_n)}{d(n)} \cdot \frac{\sum_{i=1}^{m_n} Q_i - m_n\mu_Q}{d(m_n)}\right) + \frac{\mu_Q(m_n - n/\mu_\sigma)}{d(n)} + \frac{\sum_{k=0}^{n} V_k - \sum_{i=1}^{m_n} Q_i}{d(n)} \\ &\equiv (I) + (II) + (III).\end{aligned}$$

Since $d(m_n)/d(n) \to (1/\mu_\sigma)^{2/\delta}$ it follows from (4.15) and (4.6) that $(I) \xrightarrow{d.} Z_{\delta/2,b}$, for some $b > 0$. Also, $(II) \to 0$ (deterministically). So, it will suffice to show that $(III) \xrightarrow{p.} 0$. To do this, fix an arbitrary $\epsilon > 0$. Applying (4.8) with $a_n = n^{1/2}(\log n)^{1/4}$ and $b_n = \epsilon d(n)$ gives

$$\begin{aligned}&\limsup_{n\to\infty} P_{0,\pi}^{V}\Big(\Big|\sum_{k=0}^{n} V_k - \sum_{i=1}^{m_n} Q_i\Big| > 2\epsilon d(n)\Big) \\ &\leq \limsup_{n\to\infty} P_{0,\pi}^{V}\Big(|\sigma_{m_n} - n| > n^{1/2}(\log n)^{1/4}\Big) + \limsup_{n\to\infty} P_{0,\pi}^{V}\Big(\sum_{i=1}^{2\lfloor n^{1/2}(\log n)^{1/4}\rfloor} Q_i > \epsilon d(n)\Big).\end{aligned}$$

The first term on the right hand side is 0 by (4.14), and the second term on the right hand side is 0 by (4.15). Since $\epsilon > 0$ is arbitrary, it follows that $(III) \xrightarrow{p.} 0$.

Proof of (v):
Let $Z_i = Q_i - \frac{\mu_Q}{\mu_\sigma}\Delta_{\sigma,i}$, $i \geq 1$. Note that with $\delta > 4$ the random variables $(Z_i)_{i\geq 1}$ are i.i.d. under $P_{0,\pi}^{V}$ with mean 0 and finite variance, due to Propositions 4.3 and 4.4. By (4.3),

$$\frac{\sum_{k=0}^{n} V_k - \frac{\mu_Q}{\mu_\sigma} n}{n^{1/2}} \geq \frac{\sum_{i=1}^{i_n} Z_i}{n^{1/2}} + \frac{\frac{\mu_Q}{\mu_\sigma}(\sigma_{i_n} - \sigma_0 - n)}{n^{1/2}} \equiv (I) + (II)$$

and

$$\frac{\sum_{k=0}^{n} V_k - \frac{\mu_Q}{\mu_\sigma} n}{n^{1/2}} \leq \frac{\sum_{i=1}^{i_n+1} Z_i}{n^{1/2}} + \frac{\frac{\mu_Q}{\mu_\sigma}(\sigma_{i_n+1} - \sigma_0 - n) + Q_0}{n^{1/2}} \equiv (I') + (II').$$

By Lemma 4.9-(i), $i_n/n \xrightarrow{p.} 1/\mu_\sigma$. So, it follows from Theorem 4.8 that $(I) \xrightarrow{d.} Z_{2,b}$ and $(I') \xrightarrow{d.} Z_{2,b}$ where $b \equiv \mathrm{Var}(Z_i)/(2\mu_\sigma)$. Furthermore, by Corollary 4.5 and Lemma 4.9-(ii), $(II) \xrightarrow{p.} 0$ and $(II') \xrightarrow{p.} 0$. This completes the proof. □

The next lemma gives the limiting distribution of the hitting times τ_n^X.

Lemma 4.10. *Let $v = 1/(1+2\frac{\mu_Q}{\mu_\sigma})$ be the velocity of the ERW $(X_n)_{n\geq 0}$ from (4.4). Then under the probability measure $P_{0,\pi}$ for the ERW (X_n) the following hold:*

(i) If $\delta \in (1,2)$ then there is some $b>0$ such that $\frac{\tau_n^X}{n^{2/\delta}} \xrightarrow{d.} Z_{\delta/2,b}$, as $n\to\infty$.

(ii) If $\delta = 2$ then there are constants $b>0$ and $c\in\mathbb{R}$ such that $\frac{\tau_n^X - \left[1+\frac{2\mu_Q(n/\mu_\sigma)}{\mu_\sigma}\right]n}{n} \xrightarrow{d.} Z_{1,b,c}$, as $n\to\infty$.

(iii) If $\delta\in(2,4)$ then there is some $b>0$ such that $\frac{\tau_n^X-\frac{1}{v}n}{n^{2/\delta}} \xrightarrow{d.} Z_{\delta/2,b}$, as $n\to\infty$.

(iv) If $\delta=4$ then there is some $b>0$ such that $\frac{\tau_n^X-\frac{1}{v}n}{[n\log(n)]^{1/2}} \xrightarrow{d.} Z_{2,b}$, as $n\to\infty$.

(v) If $\delta>4$ then there is some $b>0$ such that $\frac{\tau_n^X-\frac{1}{v}n}{n^{1/2}} \xrightarrow{d.} Z_{2,b}$, as $n\to\infty$.

Proof. Since (X_n) is $P_{0,\pi}$ a.s. right transient with any $\delta>1$, $\limsup_{n\to\infty}\sum_{k<0} D_{n,k}$ is $P_{0,\pi}$ a.s. finite. So, $\left(\sum_{k<0} D_{n,k}\right)/n^\alpha \xrightarrow{p.} 0$, for any $\alpha>0$. Thus, parts (i)-(v) of this lemma follow directly from (i)-(v) of Lemma 4.6 using (2.8) and Lemma 2.4, along with (4.7) for part (ii) and (4.6) for the other parts. (Note that the constants b, c are modified from Lemma 4.6.) □

The proof of Theorem 1.13 below will be based on Lemma 4.10, but first we will need one final lemma about backtracking probabilities. The same statement has been given in [16, Lemma 6.1] for the case of (IID) environments, and our proof will be very similar, but must be adjusted slightly to use the renewal times $(\sigma_i)_{i\geq 0}$, instead of the successive times at which $V_k=0$.

Lemma 4.11. *Assume that $\delta>1$ and let c_{14} be as in Lemma 4.9. Then*

$$P_{0,\pi}\Big(\inf_{m\geq \tau_{n+k}^X} X_m \leq n\Big) \leq c_{14}\cdot k^{1-\delta}\ , \quad \text{for all } k,n\in\mathbb{N}.$$

In particular, $P_{0,\pi}\Big(\inf_{m\geq\tau_{n+k}^X} X_m\leq n\Big)\to 0$, as $k\to\infty$, uniformly in n.

Proof. As in the proof of Lemma 4.9 we consider the renewal times σ_i and note that, for each fixed k,

$$\lim_{N\to\infty} P_{0,\pi}^V\Big(\text{no renewals in times } \{N,N+1,\dots,N+k-1\}\Big) \leq c_{14}\cdot k^{1-\delta}.$$

Using this along with Lemma 2.4 gives

$$\begin{aligned}
&P_{0,\pi}\Big(\inf_{m\geq\tau_{n+k}^X} X_m\leq n\Big) = \lim_{N\to\infty} P_{0,\pi}\Big(\inf_{\tau_{n+k}^X\leq m<\tau_N^X} X_m\leq n\Big)\\
&\leq \lim_{N\to\infty} P_{0,\pi}\Big(D_{N,j}\geq 1, \forall j\in\{n+1,n+2,\dots,n+k\}\Big)\\
&= \lim_{N\to\infty} P_{0,\pi}^V\Big(V_j\geq 1,\forall j\in\{N-n-k,\dots,N-n-1\}\Big)\\
&\leq \lim_{N\to\infty} P_{0,\pi}^V\Big(\text{no renewals in times } \{N-n-k,\dots,N-n-1\}\Big)\\
&\leq c_{14}\cdot k^{1-\delta}.
\end{aligned}$$

□

Proof of Theorem 1.13. Recall that for the probability measure $\mathbf{P}$ on environments, the marginal distribution of S_0 is ϕ. We will first prove the theorem in the case that $\phi = \pi$ is the stationary distribution. We will then extend to the case of general ϕ.

Case 1: $\phi = \pi$.
Let $X_n^+ = \sup_{i \le n} X_i$ and $X_n^- = \inf_{i \ge n} X_i$. Since $X_n^- \le X_n \le X_n^+$, for all n, it will suffice to show that

(a) Parts (i)-(v) of the theorem all hold with X_n replaced with X_n^+, and

(b) Parts (i)-(v) of the theorem all hold with X_n replaced with X_n^-.

Now observe that, for any $n, m, k \in \mathbb{N}$,

$$\{X_n^+ < m\} = \{\tau_m^X > n\} \tag{4.16}$$

and

$$\{X_n^+ < m\} \subset \{X_n^- < m\} \subset \Big(\{X_n^+ < m + k\} \cup \big\{ \inf_{i \ge \tau_{m+k}^X} X_i < m \big\}\Big). \tag{4.17}$$

Standard computations using (4.16) and Lemma 4.10 give (a) for parts (i), (iii), (iv), and (v) of the theorem (with a modified value of the constant b in parts (iii)-(v)). The proof of (a) for part (ii) of the theorem is a bit more subtle and will be given below, following closely the method in [9, Appendix B]. The second statement (b) follows from (a) using (4.17) and Lemma 4.11.

We proceed now to the proof of (a) for part (ii) of the theorem. Assume $\delta = 2$ and define $D(t) = c + 1 + \frac{2\mu_Q(t/\mu_\sigma)}{\mu_\sigma}$ and $a = \mu_\sigma/(2c_{13})$, where c is the same constant from Lemma 4.10-(ii). Then, by Lemma 4.10-(ii) and (4.7),

$$\frac{\tau_n^X - D(n)n}{n} \xrightarrow{d.} Z_{1,b}, \text{ for some } b > 0. \tag{4.18}$$

Also, since $P_{0,\pi}^V(Q_i > x) \sim c_{13} x^{-1}$, $i \ge 1$, it follows from the definition of $\mu_Q(t)$ that

$$D(t) \sim \frac{1}{a}\log(t), \text{ as } t \to \infty \quad \text{and} \quad \lim_{n\to\infty} D(k_n) - D(n) = 0, \text{ if } k_n \sim n. \tag{4.19}$$

For $t > 0$, let $\Gamma(t) = \inf\{s > 0 : sD(s) \ge t\}$. Note that $D(t) \sim \frac{1}{a}\log(t) \Longrightarrow \Gamma(t) \sim at/\log(t)$. Further we claim that

$$\Gamma(t) D(\Gamma(t)) = t + o(\Gamma(t)), \text{ as } t \to \infty. \tag{4.20}$$

To see this note that the function $g(s) = sD(s)$ is right continuous and strictly increasing for all sufficiently large s. Moreover, jump discontinuities in this function $g(s)$ can occur only at $s = k\mu_\sigma$ for integer k, and at such an s the size of the jump discontinuity is $s[\frac{2}{\mu_\sigma} \cdot \frac{s}{\mu_\sigma} P_{0,\pi}^V(Q_i = \frac{s}{\mu_\sigma})] = 2(\frac{s}{\mu_\sigma})^2 P_{0,\pi}^V(Q_i = \frac{s}{\mu_\sigma})$. It follows from these observations and the definition of $\Gamma(t)$ that

$$|\Gamma(t)D(\Gamma(t)) - t| \le 2(\Gamma(t)/\mu_\sigma)^2 P_{0,\pi}^V(Q_i = \Gamma(t)/\mu_\sigma), \tag{4.21}$$

for all sufficiently large t. Now, since $P_{0,\pi}^V(Q_i > x) \sim c_{13}x^{-1}$, we have $xP_{0,\pi}^V(Q_i = x) \to 0$, as $x \to \infty$. So, the right hand side of (4.21) is $o(\Gamma(t))$, which proves (4.20).

Now, for $x \in \mathbb{R}$ and $n \in \mathbb{N}$, let $k_{n,x} = \max\{\lceil \Gamma(n) + \frac{xn}{(\log n)^2} \rceil, 0\}$. Note that $k_{n,x} \sim \Gamma(n)$ as $n \to \infty$, since $\Gamma(n) \sim an/\log(n)$. Using this fact along with (4.19) and (4.20) and, again, the tail asymptotics for $\Gamma(n)$ shows that, for any fixed x,

$$\lim_{n\to\infty} \frac{n - k_{n,x} D(k_{n,x})}{k_{n,x}} = \lim_{n\to\infty} \frac{n - \left[\Gamma(n) + \frac{xn}{(\log n)^2}\right] D(\Gamma(n))}{\Gamma(n) + \frac{xn}{(\log n)^2}} = \frac{-x}{a^2}. \tag{4.22}$$

Further, since X_n^+ takes only integer values it follows from (4.16) that, for all sufficiently large n,

$$P_{0,\pi}^V\left(\frac{X_n^+ - \Gamma(n)}{n/(\log n)^2} < x\right) = P_{0,\pi}^V\left(\frac{\tau_{k_{n,x}}^X - k_{n,x}D(k_{n,x})}{k_{n,x}} > \frac{n - k_{n,x}D(k_{n,x})}{k_{n,x}}\right).$$

Taking the limit of both sides as $n \to \infty$ and using (4.18) and (4.22) shows that $\lim_{n\to\infty} P_{0,\pi}^V\left(\frac{X_n^+ - \Gamma(n)}{n/(\log n)^2} < x\right) = \mathbb{P}(Z_{1,b} > -x/a^2)$, which implies $\frac{X_n^+ - \Gamma(n)}{a^2 n/(\log n)^2} \xrightarrow{d.} -Z_{1,b}$.

Case 2: General ϕ.
We will extend from Case 1 using a coupling argument. Let $(S_k^\phi)_{k\in\mathbb{Z}}$ and $(S_k^\pi)_{k\in\mathbb{Z}}$ denote the stack sequences when S_0 has marginal distribution ϕ and π respectively. Couple these processes as follows. First sample $(S_k^\phi)_{k\le 0}$ and $(S_k^\pi)_{k\le 0}$ independently. Then run the Markov chains (S_k^ϕ) and (S_k^π) forward in time independently, starting from the given values of S_0^ϕ and S_0^π, until the first time $N > 0$ such that $S_N^\phi = S_N^\pi$ (due to the uniform ergodicity hypothesis N is a.s. finite). After the chains first meet at time N, run them forward together, so that $S_k^\phi = S_k^\pi$, for all $k \ge N$.

Now, let ω^ϕ and ω^π be the corresponding environments given by $\omega^\phi(k,i) = S_k^\phi(i)$ and $\omega^\pi(k,i) = S_k^\pi(i)$, as in (1.11), and couple the random walks (X_n^ϕ) and (X_n^π) in these environments as follows:

- Let $(\theta_{k,i})_{k\in\mathbb{Z},i\in\mathbb{N}}$ be i.i.d. Uniform([0,1]) random variables.
- Set $X_0^\phi = X_0^\pi = 0$. Then define inductively:

$$X_{n+1}^\phi = \begin{cases} X_n^\phi + 1\ , & \text{if } \theta_{X_n^\phi, I_n^\phi} < \omega^\phi(X_n^\phi, I_n^\phi) \\ X_n^\phi - 1\ , & \text{else} \end{cases} \qquad \text{and}$$

$$X_{n+1}^\pi = \begin{cases} X_n^\pi + 1\ , & \text{if } \theta_{X_n^\pi, I_n^\pi} < \omega^\pi(X_n^\pi, I_n^\pi) \\ X_n^\pi - 1\ , & \text{else} \end{cases}$$

where $I_n^\phi = |\{0 \le m \le n : X_m^\phi = X_n^\phi\}|$ and $I_n^\pi = |\{0 \le m \le n : X_m^\pi = X_n^\pi\}|$.

In words, the walk (X_n^ϕ) jumps right on its i-th visit to site k if $\theta_{k,i} < \omega^\phi(k,i)$, and left otherwise. Similarly, the walk (X_n^π) jumps right on its i-th visit to site k if $\theta_{k,i} < \omega^\pi(k,i)$, and left otherwise.

With this construction both walks (X_n^π) and (X_n^ϕ) have the correct averaged laws. Moreover, due to the coupling between environments $\omega^\pi(k,i) = \omega^\phi(k,i)$, for all $k \ge N$ and $i \in \mathbb{N}$. So, both walks (X_n^π) and (X_n^ϕ) have the same theoretical "jump sequence" at each site $k \ge N$. That is, both walks will jump the same direction from any site $k \ge N$ on their i-th visit to that site, if such an i-th visit occurs.

Now, since it is assumed that $\delta > 1$ in all cases of the theorem, we know both walks are right transient. So, with probability 1, each walk eventually reaches site N and also returns to this site after any leftward excursion from it. Combining this with the previous observation about matching jump sequences at all sites $k \ge N$ shows that

$$\limsup_{n\to\infty} |X_n^\phi - X_n^\pi| < \infty \text{ a.s.}$$

Since (1.4)-(1.8) hold with $X_n = X_n^\pi$, by Case 1, it follows that (1.4)-(1.8) also hold with $X_n = X_n^\phi$. □

A Proof of Lemma 2.2

Proof of Lemma 2.2. Fix $x \in \mathbb{N}$. Clearly, $P_{x,s}(A^+) = 0$ if $P_{x,s}(X_n \to +\infty) = 0$, so we need only show that $P_{x,s}(A^+) > 0$ if $P_{x,s}(X_n \to +\infty) > 0$. By definition, $P_{x,s}(A^+) = \mathbf{E}_s[P_x^\omega(A^+)]$

and $P_{x,s}(X_n \to +\infty) = \mathbf{E}_s[P_x^\omega(X_n \to +\infty)]$, so it will suffice to show the following claim.

Claim: Let $\omega \in \Omega$ be any environment satisfying $\omega(k,i) \in (0,1)$, for all $k \in \mathbb{Z}$ and $i \in \mathbb{N}$. Then $P_x^\omega(A^+) > 0$ if $P_x^\omega(X_n \to +\infty) > 0$.

Proof of Claim: For $m \in \mathbb{N}$ and any nearest neighbor path $\zeta = (x_0, \ldots, x_m) \in \mathbb{Z}^{m+1}$ with $x_0 = x$, define the events A_m and A_ζ by

$$A_m = \{X_n > x, \forall n \geq m \text{ and } X_n \to +\infty\} \quad \text{and} \quad A_\zeta = A_m \cap \{X_0 = x_0, \ldots, X_m = x_m\}.$$

If $P_x^\omega(X_n \to +\infty) > 0$, then there must be some finite path $\zeta = (x_0, \ldots, x_m)$ such that $x_0 = x$, $x_m = x + 1$, and $P_x^\omega(A_\zeta) > 0$. This, of course, implies $P_x^\omega(A_m | X_0 = x_0, \ldots, X_m = x_m) > 0$.

We construct from $\zeta = (x_0, \ldots, x_m)$ the reduced path $\widetilde{\zeta} = (\widetilde{x}_0, \ldots, \widetilde{x}_{\widetilde{m}})$ by setting $\widetilde{x}_0 = x_0 = x$, and then removing from the tail $(x_1, \ldots, x_m)$ all steps in any leftward excursions from site x. For example,

$$\text{if } \zeta = (x, \mathbf{x-1}, \mathbf{x-2}, \mathbf{x-1}, \mathbf{x}, x+1, x, \mathbf{x-1}, \mathbf{x}, \mathbf{x-1}, \mathbf{x}, x+1, x+2, x+3, x+2, x+1)$$
$$\text{then } \widetilde{\zeta} = (x, x+1, x, x+1, x+2, x+3, x+2, x+1)$$

(where we denote the removed steps in bold for visual clarity). By construction, $\widetilde{x}_{\widetilde{m}} = x_m = x+1$ and the number of visits to each site $k \geq x+1$ up to time m if $(X_0, \ldots, X_m) = \zeta$ is exactly the same as the number of visits to each site $k \geq x+1$ up to time $\widetilde{m}$ if $(X_0, \ldots, X_{\widetilde{m}}) = \widetilde{\zeta}$. Thus,

$$P_x^\omega(A_{\widetilde{m}} | X_0 = \widetilde{x}_0, \ldots, X_{\widetilde{m}} = \widetilde{x}_{\widetilde{m}}) = P_x^\omega(A_m | X_0 = x_0, \ldots, X_m = x_m) > 0,$$

which implies

$$\begin{aligned} P_x^\omega(A^+) &\geq P_x^\omega(X_0 = \widetilde{x}_0, \ldots, X_{\widetilde{m}} = \widetilde{x}_{\widetilde{m}}) \cdot P_x^\omega(A^+ | X_0 = \widetilde{x}_0, \ldots, X_{\widetilde{m}} = \widetilde{x}_{\widetilde{m}}) \\ &\geq P_x^\omega(X_0 = \widetilde{x}_0, \ldots, X_{\widetilde{m}} = \widetilde{x}_{\widetilde{m}}) \cdot P_x^\omega(A_{\widetilde{m}} | X_0 = \widetilde{x}_0, \ldots, X_{\widetilde{m}} = \widetilde{x}_{\widetilde{m}}) > 0. \end{aligned}$$

(Note that $P_x^\omega(X_0 = \widetilde{x}_0, \ldots, X_{\widetilde{m}} = \widetilde{x}_{\widetilde{m}}) > 0$, since we assume $\omega(k,i) \in (0,1)$, $\forall k \in \mathbb{Z}$, $i \in \mathbb{N}$.) □

B Proof of Lemma 3.1

The proof of Lemma 3.1 is based on the following much more general, but less explicit, condition for transience vs. recurrence of Markov chains on the nonnegative integers given in [13].

Lemma B.1 (Theorem A.1 of [13]). *Let $(Z_k)_{k \geq 0}$ be an irreducible, time-homogenous Markov chain on state space $\mathbb{N}_0$. Let $\mathbb{P}_x(\cdot)$ be the probability measure for the Markov chain started from $Z_0 = x$, and let $\mathbb{E}_x(\cdot)$ be the corresponding expectation operator.*

(i) If there exists a function $F : \mathbb{N}_0 \to (0, \infty)$ such that $\lim_{x \to \infty} F(x) = \infty$ and $\mathbb{E}_x[F(Z_1)] \leq F(x)$ for all sufficiently large x, then the Markov chain (Z_k) is recurrent.

(ii) If there exists a function $F : \mathbb{N}_0 \to (0, \infty)$ such that $\lim_{x \to \infty} F(x) = 0$ and $\mathbb{E}_x[F(Z_1)] \leq F(x)$ for all sufficiently large x, then the Markov chain (Z_k) is transient.

The function $F(x)$ is called a Lyapunov function. Our proof of Lemma 3.1 using Lemma B.1 follows closely the proof of Theorem 1.3 in [13]. In particular, we will use the same choice of Lyapunov functions $F(x)$ and Taylor expand in the same fashion. However, controlling the error term in the Taylor expansion becomes somewhat more involved, because of the weaker concentration condition we assume for the transition probabilities.

Proof of Lemma 3.1 (i). Let $F(x) : [0,\infty) \to (0,\infty)$ be a smooth function such that $F(x) = \ln\ln(x)$, for $x > 3$. Then for all $x > 3$

$$\begin{aligned} F'(x) &= \frac{1}{x\ln(x)}, \\ F''(x) &= -\frac{1}{x^2\ln(x)} - \frac{1}{x^2\ln^2(x)}, \\ F'''(x) &= \frac{2}{x^3\ln(x)} + \frac{3}{x^3\ln^2(x)} + \frac{2}{x^3\ln^3(x)}. \end{aligned} \tag{B.1}$$

By Taylor's Theorem with remainder

$$F(Z_1) = F(x) + F'(x)(Z_1 - x) + \frac{1}{2}F''(x)(Z_1-x)^2 + \frac{1}{6}F'''(\xi)(Z_1-x)^3, \text{ for each } x \in \mathbb{N},$$

where ξ is some (random, depending on Z_1) number between Z_1 and x. Thus, for all positive integer $x > 3$,

$$\begin{aligned} &\mathbb{E}_x[F(Z_1)] \\ &= F(x) + \frac{\mathbb{E}_x[Z_1 - x]}{x\ln(x)} + \frac{1}{2}\left[-\frac{1}{x^2\ln(x)} - \frac{1}{x^2\ln^2(x)}\right]\mathbb{E}_x[(Z_1-x)^2] \ + \ \frac{1}{6}\mathbb{E}_x[F'''(\xi)(Z_1-x)^3] \\ &= F(x) + \frac{1}{x\ln(x)}\left[\rho(x) - \frac{1}{2}\left(1 + \frac{1}{\ln(x)}\right)\nu(x)\right] \ + \ \frac{1}{6}\mathbb{E}_x[F'''(\xi)(Z_1-x)^3]. \end{aligned}$$

So,

$$\begin{aligned} \mathbb{E}_x[F(Z_1)] \le F(x) \iff &\frac{1}{x\ln(x)}\left[\rho(x) - \frac{1}{2}\left(1+\frac{1}{\ln(x)}\right)\nu(x)\right] + \frac{1}{6}\mathbb{E}_x[F'''(\xi)(Z_1-x)^3] \le 0 \\ \iff &\theta(x) \le 1 + \frac{1}{\ln(x)} - \frac{1}{3}\frac{x\ln(x)}{\nu(x)}\mathbb{E}_x[F'''(\xi)(Z_1-x)^3]. \end{aligned}$$

Therefore, in light of Lemma B.1 and the assumption that $\liminf_{x\to\infty}\nu(x) > 0$, it will suffice to show

$$\mathbb{E}_x[F'''(\xi)(Z_1-x)^3] = o\Big(\frac{1}{x\ln^2(x)}\Big). \tag{B.2}$$

Now,

$$\begin{aligned} &\left|\mathbb{E}_x[F'''(\xi)(Z_1-x)^3]\right| \le \mathbb{E}_x\left|F'''(\xi)(Z_1-x)^3\right| \\ &= \begin{cases} \mathbb{E}_x\left|F'''(\xi)(Z_1-x)^3\mathbb{1}\{Z_1 \le x/2\}\right| + \mathbb{E}_x\left|F'''(\xi)(Z_1-x)^3\mathbb{1}\{x/2 < Z_1 \le 2x\}\right| \\ + \ \mathbb{E}_x\left|F'''(\xi)(Z_1-x)^3\mathbb{1}\{Z_1 > 2x\}\right| \end{cases} \\ &\equiv (I) + (II) + (III). \end{aligned} \tag{B.3}$$

Below we will show that $(I) = O\left(e^{-x^{1/4}}\right)$, $(II) = O\left(\dfrac{1}{\ln(x)\cdot x^{4/3}}\right)$, and $(III) = O\left(e^{-x^{1/4}}\right)$, which establishes (B.2).

Bound on (I): Define $C = \max_{x\in[0,\infty)} |F'''(x)| < \infty$. Using (3.1) gives,

$$\begin{aligned}(I) &\equiv \mathbb{E}_x \left|F'''(\xi)(Z_1-x)^3 \mathbb{1}\{Z_1 \le x/2\}\right| \le Cx^3 \cdot \mathbb{P}_x(Z_1 \le x/2) \le Cx^3 \cdot \mathbb{P}_x(|Z_1 - x| \ge x/2) \\ &\le Cx^3 \cdot C_1\left[1 + (1/2)^{2/3}x^{1/3}\right]e^{-C_2(1/2)^{2/3}x^{1/3}} = O\left(e^{-x^{1/4}}\right).\end{aligned}$$

Bound on (II): By (B.1), for all $x > 6$,

$$\begin{aligned}(II) &\equiv \mathbb{E}_x \left|F'''(\xi)(Z_1-x)^3 \mathbb{1}\{x/2 < Z_1 \le 2x\}\right| \\ &\le \frac{7}{(x/2)^3 \ln(x/2)} \mathbb{E}_x\Big(|Z_1 - x|^3 \mathbb{1}\{x/2 < Z_1 \le 2x\}\Big).\end{aligned} \tag{B.4}$$

We define $\widetilde{Z}_1 = Z_1 \cdot \mathbb{1}_{\{x/2<Z_1\le 2x\}} + x \cdot \mathbb{1}_{\{Z_1 \notin (x/2,2x]\}}$, and bound the expectation on the right hand side as follows:

$$\begin{aligned}\mathbb{E}_x\Big(|Z_1 - x|^3 \mathbb{1}\{x/2 < Z_1 \le 2x\}\Big) &= \mathbb{E}_x\left(|\widetilde{Z}_1 - x|^3\right) \\ &= \int_0^x \mathbb{P}_x(|\widetilde{Z}_1 - x| > y) \cdot 3y^2 dy \\ &\le 3\int_0^x \mathbb{P}_x(|Z_1 - x| > y)\cdot y^2 dy \\ &= 3\int_0^1 \mathbb{P}_x(|Z_1 - x| > \epsilon x)\cdot x^3\epsilon^2 d\epsilon \qquad \text{(substitute } y = \epsilon x\text{)} \\ &\le 3x^3 \int_0^1 \epsilon^2 \cdot C_1(1+\epsilon^{2/3}x^{1/3})e^{-C_2\epsilon^{2/3}x^{1/3}} d\epsilon \qquad \text{(by (3.1))} \\ &= \frac{9}{2}x^3 \int_0^1 C_1(t^{7/2} + t^{9/2}x^{1/3})e^{-C_2 x^{1/3} t} dt \qquad \text{(substitute } t = \epsilon^{2/3}\text{)} \\ &\le \frac{9}{2}x^3 \int_0^1 C_1(t^3 + t^4 x^{1/3})e^{-C_2 x^{1/3} t} dt \\ &= O\left(x^{5/3}\right) \qquad \text{(repeated integration by parts).}\end{aligned} \tag{B.5}$$

Combining (B.4) and (B.5) shows that $(II) = O\left(\dfrac{1}{\ln(x)\cdot x^{4/3}}\right)$.

Bound on (III): By (B.1), for all $x > 3$,

$$(III) \equiv \mathbb{E}_x \left|F'''(\xi)(Z_1-x)^3 \mathbb{1}\{Z_1 > 2x\}\right| \le \frac{7}{x^3\ln(x)}\mathbb{E}_x\Big(|Z_1 - x|^3 \mathbb{1}\{Z_1 > 2x\}\Big). \tag{B.6}$$

We define $\widetilde{Z}_1 = Z_1 \cdot \mathbb{1}_{\{Z_1 > 2x\}} + x \cdot \mathbb{1}_{\{Z_1 \le 2x\}}$ and write the expectation on the right hand side as

$$\begin{aligned}&\mathbb{E}_x\Big(|Z_1 - x|^3 \mathbb{1}\{Z_1 > 2x\}\Big) = \mathbb{E}_x\Big(|\widetilde{Z}_1 - x|^3\Big) = \int_0^\infty \mathbb{P}_x(|\widetilde{Z}_1 - x| > y)\cdot 3y^2 dy \\ &= 3\left[\int_0^x \mathbb{P}_x(Z_1 - x > x) y^2 dy \;+\; \int_x^\infty \mathbb{P}_x(Z_1 - x > y) y^2 dy\right].\end{aligned} \tag{B.7}$$

The first integral in the brackets on the right hand side of (B.7) is easily bounded using (3.1):

$$\int_0^x \mathbb{P}_x(Z_1 - x > x)y^2 dy \le \int_0^x \left[C_1\left(1 + x^{1/3}\right)e^{-C_2 x^{1/3}}\right] y^2 dy = O\left(e^{-x^{1/4}}\right). \tag{B.8}$$

The second integral in the brackets on the right hand side of (B.7) is bounded as follows:

$$\begin{aligned}\int_x^\infty \mathbb{P}_x(Z_1 - x > y)y^2 dy &= \int_x^\infty \mathbb{P}_x(|Z_1 - x| > y)y^2 dy \quad (Z_1 \text{ is nonnegative})\\ &= \int_1^\infty \mathbb{P}_x(|Z_1 - x| > \epsilon x)x^3\epsilon^2 d\epsilon \quad (\text{substitute } y = \epsilon x)\\ &\le x^3 \int_1^\infty \epsilon^2 \cdot C_1 \left(1 + x^{1/3}\epsilon^{1/3}\right) e^{-C_2 x^{1/3}\epsilon^{1/3}} d\epsilon \quad (\text{by (3.1)})\\ &= 3x^3 \int_1^\infty C_1(t^8 + t^9 x^{1/3})e^{-C_2 x^{1/3} t} dt \quad (\text{substitute } t = \epsilon^{1/3})\\ &= O\left(e^{-x^{1/4}}\right) \quad (\text{repeated integration by parts}). \end{aligned} \tag{B.9}$$

Combining (B.7), (B.8), and (B.9) shows that

$$\mathbb{E}_x\left(|Z_1 - x|^3 \mathbb{1}\{Z_1 > 2x\}\right) = O\left(e^{-x^{1/4}}\right). \tag{B.10}$$

Hence, by (B.6), $(III) = O\left(e^{-x^{1/4}}\right)$. □

Proof of Lemma 3.1 (ii). Let $F(x) : [0, \infty) \to (0, \infty)$ be a smooth function such that $F(x) = \frac{1}{\ln(x)}$, for $x > 2$. Then for all $x > 2$

$$\begin{aligned} F'(x) &= -\frac{1}{x \ln^2(x)},\\ F''(x) &= \frac{1}{x^2 \ln^2(x)} + \frac{2}{x^2 \ln^3(x)},\\ F'''(x) &= -\frac{2}{x^3 \ln^2(x)} - \frac{6}{x^3 \ln^3(x)} - \frac{6}{x^3 \ln^4(x)}. \end{aligned} \tag{B.11}$$

Thus, by Taylor's Theorem with remainder, for all positive integer $x > 2$

$$\begin{aligned} &\mathbb{E}_x[F(Z_1)]\\ &= F(x) - \frac{\mathbb{E}_x[Z_1 - x]}{x \ln^2(x)} + \frac{1}{2}\left[\frac{1}{x^2 \ln^2(x)} + \frac{2}{x^2 \ln^3(x)}\right]\mathbb{E}_x[(Z_1 - x)^2] + \frac{1}{6}\mathbb{E}_x[F'''(\xi)(Z_1 - x)^3]\\ &= F(x) + \frac{1}{x \ln^2(x)}\left[-\rho(x) + \frac{1}{2}\left(1 + \frac{2}{\ln(x)}\right)\nu(x)\right] + \frac{1}{6}\mathbb{E}_x[F'''(\xi)(Z_1 - x)^3] \end{aligned}$$

where ξ is some (random) number between Z_1 and x. So,

$$\begin{aligned} &\mathbb{E}_x[F(Z_1)] \le F(x)\\ &\iff \frac{1}{x \ln^2(x)}\left[-\rho(x) + \frac{1}{2}\left(1 + \frac{2}{\ln(x)}\right)\nu(x)\right] + \frac{1}{6}\mathbb{E}_x[F'''(\xi)(Z_1 - x)^3] \le 0\\ &\iff \theta(x) \ge 1 + \frac{2}{\ln(x)} + \frac{1}{3}\frac{x \ln^2(x)}{\nu(x)}\mathbb{E}_x[F'''(\xi)(Z_1 - x)^3]. \end{aligned}$$

Therefore, in light of Lemma B.1 and the assumption that $\liminf_{x\to\infty} \nu(x) > 0$, it will suffice to show

$$\mathbb{E}_x[F'''(\xi)(Z_1 - x)^3] = o\Big(\frac{1}{x \ln^3(x)}\Big). \tag{B.12}$$

Now, $|\mathbb{E}_x[F'''(\xi)(Z_1 - x)^3]| \le (I) + (II) + (III)$, where terms (I), (II), and (III) are defined just as in (B.3), but with our new Lyapunov function $F(x) = 1/\ln(x)$. The exact same proof as above for part (i) of the lemma shows that $(I) = O\left(e^{-x^{1/4}}\right)$. The bounds on (II) and (III) are also quite similar:

$$(II) \overset{\text{(B.11)}}{\leq} \left[\frac{14}{(x/2)^3 \ln^2(x/2)}\right] \cdot \mathbb{E}_x\Big(|Z_1 - x|^3 \mathbb{1}\{x/2 < Z_1 \leq 2x\}\Big) \overset{\text{(B.5)}}{=} O\left(\frac{1}{\ln^2(x) \cdot x^{4/3}}\right).$$

$$(III) \overset{\text{(B.11)}}{\leq} \left[\frac{14}{x^3 \ln^2(x)}\right] \cdot \mathbb{E}_x\Big(|Z_1 - x|^3 \mathbb{1}\{Z_1 > 2x\}\Big) \overset{\text{(B.10)}}{=} O\left(e^{-x^{1/4}}\right).$$

Combining these estimates on (I), (II), and (III) shows that $\mathbb{E}_x[F'''(\xi)(Z_1 - x)^3] = O\left(\frac{1}{\ln^2(x) \cdot x^{4/3}}\right)$, which implies (B.12). □

C Proof of Propositions 4.1 and 4.2

Our proof of Propositions 4.1 and 4.2 is based on the following slightly more general proposition, which we isolate in this form because it seems it may be applicable for analyzing other critical branching (or branching-type) processes.

Proposition C.1. *Let $(Z_k)_{k\geq 0}$ be an irreducible, time-homogeneous Markov chain on state space $\mathbb{N}_0$, and denote by P_x^Z the probability measure for the Markov chain (Z_k) started from $Z_0 = x$. Also, let E_x^Z and Var_x^Z denote, respectively, expectation and variance with respect to the measure P_x^Z. Assume that the following four conditions are satisfied.*

(A) Monotonicity: *For any $x, y, z \in \mathbb{N}_0$ with $x > y$* [8]*,*

$$P_x^Z(Z_1 \geq z) \geq P_y^Z(Z_1 \geq z). \tag{C.1}$$

(B) Expectation and Variance: *There exist constants $\alpha > 0$ and $\beta < 0$ such that*

$$E_x^Z(Z_1) = x + \alpha\beta + O(x^{-1/3}) \quad \text{and} \quad \text{Var}_x^Z(Z_1) = 2\alpha x + O(x^{2/3}). \tag{C.2}$$

(C) Single Step Concentration Estimate: *There exist constants $C_1, C_2 > 0$ such that:*

$$\begin{aligned} P_x^Z(|Z_1 - x| \geq \epsilon x) &\leq C_1(1 + \epsilon^{2/3}x^{1/3})e^{-C_2\epsilon^{2/3}x^{1/3}}, \quad \text{for all } x \in \mathbb{N} \text{ and } 0 < \epsilon \leq 1. \\ P_x^Z(|Z_1 - x| \geq \epsilon x) &\leq C_1(1 + \epsilon^{1/3}x^{1/3})e^{-C_2\epsilon^{1/3}x^{1/3}}, \quad \text{for all } x \in \mathbb{N} \text{ and } \epsilon \geq 1. \end{aligned} \tag{C.3}$$

(D) Exit Probability Concentration Estimates: *There exist constants $C_3, C_4 > 0$ and $N \in \mathbb{N}$ such that for all $x \geq N$ the following hold:*

$$\sup_{0 \leq z < x} P_z^Z(Z_{\tau_{x^+}^Z} > x + y | \tau_{x^+}^Z < \tau_0^Z) \leq \begin{cases} C_3(1 + y^{2/3}x^{-1/3})e^{-C_4 y^{2/3}x^{-1/3}}, & \text{for } 0 \leq y \leq x \\ C_3(1 + y^{1/3})e^{-C_4 y^{1/3}}, & \text{for } y \geq x \end{cases} \tag{C.4}$$

$$\sup_{x < z < 4x} P_z^Z(Z_{\tau_{x^-}^Z \wedge \tau_{(4x)^+}^Z} < x - y) \leq C_3(1 + y^{2/3}x^{-1/3})e^{-C_4 y^{2/3}x^{-1/3}}, \quad \text{for } 0 \leq y \leq x. \tag{C.5}$$

Here, as usual, $\tau_x^Z = \inf\{k > 0 : Z_k = x\}$, and $\tau_{x^+}^Z$, $\tau_{x^-}^Z$ are defined by $\tau_{x^+}^Z = \inf\{k > 0 : Z_k \geq x\}$, $\tau_{x^-}^Z = \inf\{k > 0 : Z_k \leq x\}$.

[8] Note that this condition (along with the Markov property) implies one can couple together versions of the process (Z_k) starting from two different initial conditions $x > y$ in such a way that $Z_k^{(x)} \geq Z_k^{(y)}$, for all $k \geq 0$.

Then there exist constants $C_5, C_6 > 0$ such that, as $t \to \infty$,

$$P_0^Z(\tau_0^Z > t) \sim C_5 t^{(\beta-1)} \quad \textit{and} \quad P_0^Z\Big(\sum_{k=0}^{\tau_0^Z} Z_k > t\Big) \sim C_6 t^{(\beta-1)/2}. \tag{C.6}$$

Proof of Propositions 4.1 and 4.2. We simply apply Proposition C.1 to the (irreducible, time-homogeneous) Markov chain $(\widehat{V}_k)_{k\geq 0}$ with transition probabilities give by (2.14). We consider this Markov chain under the family of measure $P_{x,s}^V$, $x \in \mathbb{N}_0$, so that $\widehat{V}_0 = V_0 = x$ deterministically. Thus, the measure $P_{x,s}^V$ for $(\widehat{V}_k)$ is the equivalent of the measure P_x^Z for the Markov chain (Z_k) in Proposition C.1.

By construction the Markov chain $(\widehat{V}_k)$ satisfies the monotonicity condition (A) of Proposition C.1. Also, by Lemmas 2.12 and 2.13, condition (B) is satisfied with $\alpha = \mu_s > 0$ and $\beta = (1-\delta) < 0$. Finally, condition (C) is satisfied due to Lemma 2.11 and condition (D) is satisfied due to Lemma 2.14. Thus, the proposition is applicable and we have $P_{0,s}^V(\tau_0^{\widehat{V}} > t) \sim c_{10} t^{-\delta}$ and $P_{0,s}^V\Big(\sum_{k=0}^{\tau_0^{\widehat{V}}} \widehat{V}_k > t\Big) \sim c_{11} t^{-\delta/2}$ for some constants $c_{10}, c_{11} > 0$. □

Proof of Proposition C.1 (Sketch). Our proof of Proposition C.1 follows very closely the approach used in [8] to prove *Theorems 2.1* and *2.2*, so we will provide only a rough sketch. Throughout we will use *italics* when referring to all theorems, lemmas, sections... etc. from [8], to distinguish from the corresponding items in our paper. For the sake of comparison we restate *Theorems 2.1* and *2.2* below explicitly.

Theorem C.2 (*Theorems 2.1 and 2.2* of [8])**.** *Let $\widetilde{\boldsymbol{P}}$ be a probability measure on cookie environments satisfying (IID), (BD), and (ELL), and let $(\widetilde{V}_k)_{k\geq 0}$ be the associated backward branching process for ERW in this environment. Assume $\delta > 0$ (where δ is given by* (1.2)*). Then there exists constants $\widetilde{C}_1, \widetilde{C}_2 > 0$ such that*

$$P_0^{\widetilde{V}}(\tau_0^{\widetilde{V}} > x) \sim \widetilde{C}_1 x^{-\delta} \quad \textit{and} \quad P_0^{\widetilde{V}}\Big(\sum_{k=0}^{\tau_0^{\widetilde{V}}} \widetilde{V}_k > x\Big) \sim \widetilde{C}_2 x^{-\delta/2}. \tag{C.7}$$

The Markov chain $(Z_k)_{k\geq 0}$ of our proposition is the equivalent of the process $(\widetilde{V}_k)_{k\geq 0}$, under the correspondence $\beta = 1-\delta$. More precisely, if $\delta > 1$ then $(\widetilde{V}_k)$ satisfies conditions (A)-(D) of the proposition with $\beta = 1-\delta$ and $\alpha = 1$, and the decay rates in (C.6) and (C.7) are the same with $\beta = 1-\delta$. The value of α does not effect the decay rates in (C.6), only the constants C_5 and C_6.

The main elements of the proof of *Theorems 2.1* and *2.2* in [8] are *Lemmas 3.1-3.3* and *3.5* in *Section 3*, and *Lemmas 5.1-5.3* and *Corollary 5.5* in *Section 5*. *Lemma 3.1* establishes convergence of the discrete process $(\widetilde{V}_k)$ to a limiting diffusion $(\widetilde{Y}_t)$, and the other lemmas in *Section 3* give properties of the limiting diffusion. The lemmas of *Section 5* then give tight estimates on exit probabilities of the process $(\widetilde{V}_k)$ from certain intervals. The entire series of additional lemmas and propositions used to establish *Theorems 2.1* and *2.2* in *Sections 4, 6, 7* and *8* use only the results of *Sections 3* and *5*, along with the fact that the process $(\widetilde{V}_k)$ is an irreducible Markov chain on state space $\mathbb{N}_0$ that is monotonic in the sense of (C.1). Essentially the goal of these other sections is to show that the discrete process $(\widetilde{V}_k)$ has the same type of scaling properties as the limiting diffusion $(\widetilde{Y}_t)$, and to do this requires some technical work, using the estimates of *Section 5*.

Now, let us compare to our situation for the process (Z_k). Lemma C.3, given below in Section C.1, is the analog of *Lemma 3.1*, and Lemma C.4 in Section C.1 is the analog of *Lemmas 3.2, 3.3,* and *3.5*. Also, Condition (D), which we assume to hold for the process

(Z_k), is the analog of *Lemma 5.1* in [8], where the corresponding property of the process $(\widetilde{V}_k)$ is proven. Finally, Lemmas C.7, C.8, and C.9 in Section C.2 are, respectively, the analogs of *Lemma 5.2*, *Lemma 5.3*, and *Corollary 5.5* in [8].

With the analogous results to the lemmas in *Section 3* and *Section 5* of [8] established, the proof or our Proposition C.1 for the process (Z_k) proceeds almost the same way, line by line, as the proof of *Theorems 2.1* and *2.2* for the process $(\widetilde{V}_k)$. So, we will not repeat it. However, for the sake of completeness, let us point out the few small differences in our analog lemmas from the originals.

1. The bound of $\exp(-a^{n/10})$ in our Lemma C.7-(i) is instead $\exp(-a^{n/4})$ in *Lemma 5.2*-(i). This is irrelevant for how the lemma is applied; a bound of $\exp(-a^{cn})$, for any $c > 0$, would be sufficient.

2. The concentration bounds (C.4) and (C.5) in our condition (D) are instead the following in *Lemma 5.1*:

$$\sup_{0 \le z < x} P_z^{\widetilde{V}}(\widetilde{V}_{\tau^{\widetilde{V}}_{x+}} > x + y | \tau^{\widetilde{V}}_{x+} < \tau^{\widetilde{V}}_0) \le C(e^{-cy^2/x} + e^{-cy}) \ , \ y \ge 0. \tag{C.8}$$

$$\sup_{x < z < 4x} P_z^{\widetilde{V}}(\widetilde{V}_{\tau^{\widetilde{V}}_{x-} \wedge \tau^{\widetilde{V}}_{(4x)+}} < x - y) \le Ce^{-cy^2/x} \ , \ 0 \le y \le x. \tag{C.9}$$

The latter bounds are slightly stronger than ours. However, when x is large, with either (C.8) and (C.9) or with (C.4) and (C.5), the right hand sides become small only when $y \gg x^{1/2}$. In [8] the inequalities (C.8) and (C.9) are applied either when y is of order x or larger, or in other instances when y is of order $x^{2/3}$, giving respectively bounds on the right hand side which are $O(e^{-cy})$ or $O(e^{-cx^{1/3}})$. If instead (C.4) and (C.5) are used these bounds reduce to, respectively, $O(e^{-y^{1/4}})$ and $O(e^{-x^{1/10}})$. But, again, this is irrelevant for how the estimates are applied; any sort of stretched exponential decay would be sufficient. With our weaker estimates slightly larger error terms arise in the proofs, but they always remain negligible in comparison with all other terms. □

C.1 Diffusion approximation lemmas

In the statement of the following lemma $(Z_k)_{k \ge 0}$ is an irreducible, time-homogeneous Markov chain on state space $\mathbb{N}_0$ satisfying (A)-(D), as in Proposition C.1.

Lemma C.3 (Diffusion Approximation, Analog of Lemma 3.1 from [8])**.** *Fix any $0 < \epsilon < y < \infty$, and let $(Y(t))_{t \ge 0}$ be the solution of*

$$dY(t) = \alpha\beta dt + \sqrt{2\alpha Y(t)^+} dB(t) \ , \ Y(0) = y \tag{C.10}$$

where $(B(t))_{t \ge 0}$ is a standard 1-dimensional Brownian motion[9]*. Also, let $Y_\epsilon(t) = Y(t \wedge \tau^Y_\epsilon)$. For each $n \in \mathbb{N}$, let $(Z_{n,k})_{k \ge 0}$ be a process with the distribution of $(Z_k)_{k \ge 0}$ when $Z_0 = \lfloor yn \rfloor$, and define $Y_{\epsilon,n}(t) = \dfrac{Z_{n, \lfloor nt \rfloor \wedge \kappa_{\epsilon,n}}}{n}$, where $\kappa_{\epsilon,n} = \inf\{k \ge 0 : Z_{n,k} \le \epsilon n\}$. In addition, let $\tau_{\epsilon,n} = \kappa_{\epsilon,n}/n$. Then:*

- *(i)* $(Y_{\epsilon,n}(t))_{t \ge 0} \xrightarrow{J_1} (Y_\epsilon(t))_{t \ge 0}$*, as $n \to \infty$, where $\xrightarrow{J_1}$ denotes convergence in distribution with respect to the Skorokhod J_1 topology.*
- *(ii)* $\tau_{\epsilon,n} \xrightarrow{d.} \tau^Y_\epsilon$*, as $n \to \infty$.*
- *(iii)* $\int_0^{\tau_{\epsilon,n}} Y_{\epsilon,n}(t)dt \xrightarrow{d.} \int_0^{\tau^Y_\epsilon} Y_\epsilon(t)dt$*, as $n \to \infty$.*

[9] Note that the process $(\widehat{Y}(t))_{t \ge 0}$ defined by $\widehat{Y}(t) = 2Y(t)/\alpha$ is a squared Bessel process of generalized dimension 2β, since it satisfies $d\widehat{Y}(t) = 2\beta dt + 2\sqrt{\widehat{Y}(t)^+} dB(t)$.

The following properties of the diffusion $Y(t)$ are established in [8] for the case $\alpha = 1$, with $\beta \equiv 1 - \delta$, but generalize to any $\alpha > 0$. Indeed note that if Y^α is the solution to (C.10) with a given value α and Y^1 is the solution to (C.10) with $\alpha = 1$ (both with the same value of β and initial value y) then $(Y^\alpha(t))_{t\geq 0} \overset{d.}{=} (Y^1(\alpha t))_{t\geq 0}$. All properties of Y^α follow from the corresponding properties for Y^1 and this relation.

Lemma C.4 (Properties of Limiting Diffusion, Analog of Lemmas 3.2, 3.3, and 3.5 from [8])**.** *Let $P_y^Y(\cdot)$ be the probability measure for the process $(Y(t))_{t\geq 0}$ in* (C.10) *with given initial value $Y(0) = y > 0$.*

(i) $\exists\, K_1, K_2 > 0$ *such that* $P_1^Y(\tau_0^Y > x) \sim K_1 x^{\beta-1}$ *and* $P_1^Y(\int_0^{\tau_0^Y} Y(t)dt > x^2) \sim K_2 x^{\beta-1}$*, as $x \to \infty$.*

(ii) For $0 \leq a < y < b$, $P_y^Y(\tau_a^Y < \tau_b^Y) = (b^{1-\beta} - y^{1-\beta})/(b^{1-\beta} - a^{1-\beta})$.

(iii) The process $(Y(t))_{t\geq 0}$ when $Y(0) = 1$ has the same law as the process $(\frac{Y(ty)}{y})_{t\geq 0}$ when $Y(0) = y$.

In the remainder of Section C.1 we will use the generic probability measure $\mathbb{P}$ and corresponding expectation operator $\mathbb{E}$ for all random variables, including the Markov chains $(Z_{n,k})_{k\geq 0}$ of Lemma C.3. The proof of Lemma C.3 is based on the following result from [9].

Lemma C.5. *([9, Lemma 7.1]) Let $b \in \mathbb{R}$ and $D > 0$, and let $(\mathcal{Y}(t))_{t\geq 0}$ be the solution of*

$$d\mathcal{Y}(t) = b\,dt + \sqrt{D\mathcal{Y}(t)^+}dB(t)\ ,\ \mathcal{Y}(0) = y > 0 \tag{C.11}$$

where $(B(t))_{t\geq 0}$ is a standard 1-dimensional Brownian motion. Let $(\mathcal{Z}_{n,k})_{k\geq 0}$, $n \in \mathbb{N}$, be integer-valued (time-homogenous) Markov chains such that $\mathcal{Z}_{n,0} = \lfloor ny_n \rfloor$, where $y_n \to y$ as $n \to \infty$, and such that conditions (i) and (ii) below are satisfied.

(i) There is a sequence of positive integers $(N_n)_{n\geq 1}$ such that $N_n \to \infty$ with $N_n = o(n)$, a function $f : \mathbb{N} \to [0, \infty)$ with $f(x) \to 0$ as $x \to \infty$, and a function $g : \mathbb{N} \to [0, \infty)$ with $g(x) \searrow 0$ as $x \to \infty$, such that:

$$(E)\quad |\mathbb{E}(\mathcal{Z}_{n,k+1} - \mathcal{Z}_{n,k}|\mathcal{Z}_{n,k} = m) - b| \leq f(m \vee N_n).$$

$$(V)\quad \left|\frac{\mathrm{Var}(\mathcal{Z}_{n,k+1}|\mathcal{Z}_{n,k} = m)}{m \vee N_n} - D\right| \leq g(m \vee N_n).$$

(ii) For each $T, a > 0$

$$\mathbb{E}\left[\max_{1\leq k\leq (Tn)\wedge t_n} (\mathcal{Z}_{n,k} - \mathcal{Z}_{n,k-1})^2\right] = o(n^2)\ ,\quad \text{as } n \to \infty$$

where $t_n \equiv \inf\{k \geq 0 : \mathcal{Z}_{n,k} \geq an\}$.

Then $(\mathcal{Y}_n(t))_{t\geq 0} \xrightarrow{J_1} (\mathcal{Y}(t))_{t\geq 0}$, as $n \to \infty$, where $\mathcal{Y}_n(t) \equiv \mathcal{Z}_{n,\lfloor nt \rfloor}/n$.

We will also need the following lemma, which is a minor extension of Lemma 3.3 in [12] where the same result is stated in the case $D = 2$.

Lemma C.6. *For a function f in the Skorokhod space $D[0, \infty)$, define $\tau_\epsilon^f = \inf\{t \geq 0 : f(t) \leq \epsilon\}$, and define $\varphi_\epsilon(f) \in D[0, \infty)$ by $(\varphi_\epsilon(f))(t) = f(t \wedge \tau_\epsilon^f)$. Let ψ be any of the following three mappings defined on $D[0, \infty)$:*

$$f \mapsto \tau_\epsilon^f \in [0, \infty], \quad f \mapsto \varphi_\epsilon(f) \in D[0, \infty), \quad f \mapsto \int_0^{\tau_\epsilon^f} f^+(t)dt \in [0, \infty].$$

Denote by Cont$(\psi) = \{f \in D[0, \infty) : \psi$ is continuous at $f\}$ the set of continuity points of ψ. Then the solution $\mathcal{Y} = (\mathcal{Y}(t))_{t\geq 0}$ of (C.11) *satisfies $\mathbb{P}(\mathcal{Y} \in \text{Cont}(\psi)) = 1$, for any $b \in \mathbb{R}$, $D > 0$, and $0 < \epsilon < y < \infty$.*

Proof of Lemma C.3. For concreteness take $N_n = \lfloor n^{1/2} \rfloor$ and define, for $n \in \mathbb{N}$, integer-valued (time homogeneous) Markov chains $\mathcal{Z}_n \equiv (\mathcal{Z}_{n,k})_{k\geq 0}$ by

$$\mathcal{Z}_{n,0} = \lfloor yn \rfloor \quad \text{and}$$
$$\mathbb{P}(\mathcal{Z}_{n,k+1} = x+z|\mathcal{Z}_{n,k} = x) = \begin{cases} \mathbb{P}(Z_{n,k+1} = x+z|Z_{n,k} = x), \ z \in \mathbb{Z} \text{ and } x \geq N_n \\ \mathbb{P}(Z_{n,k+1} = N_n+z|Z_{n,k} = N_n), \ z \in \mathbb{Z} \text{ and } x < N_n. \end{cases} \tag{C.12}$$

Thus, $\mathcal{Z}_{n,0} = Z_{n,0} \equiv \lfloor yn \rfloor$, and if the Markov chain $(\mathcal{Z}_{n,k})_{k\geq 0}$ is currently at level $x \geq N_n$, then it has the same transition probabilities as the Markov chain $(Z_{n,k})_{k\geq 0}$. On the other hand, if the Markov chain $(\mathcal{Z}_{n,k})_{k\geq 0}$ is currently at some level $x < N_n$, then the difference $(\mathcal{Z}_{n,k+1} - \mathcal{Z}_{n,k})$ between the current value and next value has the same law as the difference $(Z_{n,k+1} - Z_{n,k})$ when $Z_{n,k} = N_n$.

With this construction the chains $(Z_{n,k})_{k\geq 0}$ and $(\mathcal{Z}_{n,k})_{k\geq 0}$ can be naturally coupled until the first time k that they fall below level ϵn (for all n large enough that $N_n = \lfloor n^{1/2} \rfloor < \epsilon n$). Thus, by Lemma C.6 and the continuous mapping theorem it will suffice to show the following claim to establish the proposition.

Claim: Define $(\mathcal{Y}_n(t))_{t\geq 0}$ by $\mathcal{Y}_n(t) = \mathcal{Z}_{n,\lfloor nt \rfloor}/n$. Then $(\mathcal{Y}_n(t))_{t\geq 0} \xrightarrow{J_1} (Y(t))_{t\geq 0}$, where $(Y(t))_{t\geq 0}$ is the solution of (C.10).

Proof of Claim: We apply Lemma C.5 with $b = \alpha\beta$ and $D = 2\alpha$. By definition $(Z_{n,k})_{k\geq 0}$ has the distribution of $(Z_k)_{k\geq 0}$ when $Z_0 = \lfloor yn \rfloor$, where the Markov chain $(Z_k)_{k\geq 0}$ is originally defined in Proposition C.1. It follows immediately from condition (B) in the statement of this proposition and the definition (C.12) that the Markov chains $(\mathcal{Z}_{n,k})_{k\geq 0}$ satisfy conditions (E) and (V) in (i) of Lemma C.5. Indeed, $f(x)$ and $g(x)$ are both $O(x^{-1/3})$. To show (ii) of Lemma C.5 we define $M_n = \max_{1\leq k\leq (Tn)\wedge t_n} |\mathcal{Z}_{n,k} - \mathcal{Z}_{n,k-1}|$ and write the expectation as

$$\mathbb{E}\left[\max_{1\leq k\leq (Tn)\wedge t_n} (\mathcal{Z}_{n,k} - \mathcal{Z}_{n,k-1})^2\right] = \mathbb{E}(M_n^2) = \int_0^\infty \mathbb{P}(M_n^2 > x)dx = \int_0^\infty \mathbb{P}(M_n > x^{1/2})dx.$$

The last integral may be decomposed as

$$\int_0^\infty \mathbb{P}(M_n > x^{1/2})dx = \int_0^{n^{3/2}} \mathbb{P}(M_n > x^{1/2})dx + \int_{n^{3/2}}^\infty \mathbb{P}(M_n > x^{1/2})dx. \tag{C.13}$$

The first integral on the right hand side of (C.13) is at most $n^{3/2}$. Using the definition (C.12) along with condition (C) in the statement of Proposition C.1 and the union bound estimate

$$\mathbb{P}(M_n > x^{1/2}) \leq Tn\left[\max_{m\leq an} \mathbb{P}\Big(|\mathcal{Z}_{n,k} - \mathcal{Z}_{n,k-1}| > x^{1/2}\Big|\mathcal{Z}_{n,k-1} = m\Big)\right]$$

one finds the second integral on the right hand side of (C.13) is o(1), as $n \to \infty$. This establishes (ii) of Lemma C.5, and, hence, the claim. □

C.2 Exit probability lemmas

In the statement of the following lemmas $(Z_k)_{k\geq 0}$ is an irreducible, time-homogeneous Markov chain on state space $\mathbb{N}_0$ satisfying (A)-(D), as in Proposition C.1.

Lemma C.7 (Analog of Lemma 5.2 from [8]). *Fix any $a \in (1,2]$, and define $\mathcal{I}_n = [a^n - a^{\frac{2}{3}n}, a^n + a^{\frac{2}{3}n}]$ and $\gamma_n = \inf\{k \geq 0 : Z_k \notin (a^{n-1}, a^{n+1})\}$. Then, for all sufficiently large n the following hold:*

(i) $P_x^Z\Big(\mathit{dist}\big(Z_{\gamma_n},(a^{n-1},a^{n+1})\big)\geq a^{\frac{2}{3}(n-1)}\Big)\leq \exp(-a^{n/10})$, *for each* $x\in\mathcal{I}_n$.

(ii) $\Big|P_x^Z(Z_{\gamma_n}\leq a^{n-1})-a^{1-\beta}/(1+a^{1-\beta})\Big|\leq a^{-n/4}$, *for each* $x\in\mathcal{I}_n$.

Lemma C.8 (Analog of Lemma 5.3 from [8]). *For each* $a\in(1,2]$ *there is some* $\ell_a\in\mathbb{N}$ *such that if* $\ell,m,u,x\in\mathbb{N}$ *satisfy* $\ell_a\leq\ell<m<u$ *and* $x\in\mathcal{I}_m$ *(where* $\mathcal{I}_m$ *is as in Lemma C.7) then*

$$\frac{h^-_{a,\ell}(m)-1}{h^-_{a,\ell}(u)-1}\leq P_x^Z\Big(\tau^Z_{(a^\ell)^-}>\tau^Z_{(a^u)^+}\Big)\leq\frac{h^+_{a,\ell}(m)-1}{h^+_{a,\ell}(u)-1}$$

where $h^{\pm}_{a,\ell}(i)=\prod_{j=\ell+1}^{i}(a^{1-\beta}\mp a^{-\lambda j})$, $i>\ell$, *and* λ *is some small positive number not depending on* ℓ.

Lemma C.9 (Analog of Corollary 5.5 from [8]). *For each* $x\in\mathbb{N}_0$ *there exists* $C=C(x)>0$ *such that*

$$P_x^Z(\tau^Z_{n^+}<\tau^Z_0)\leq C/n^{1-\beta}\ ,\quad \text{for all } n\in\mathbb{N}.$$

Moreover, for each $\epsilon>0$ *there exists* $c=c(\epsilon)>0$ *such that*

$$P_n^Z(\tau^Z_0>\tau^Z_{(cn)^+})<\epsilon\ ,\quad \text{for all } n\in\mathbb{N}.$$

We will prove Lemma C.7 below. The proof is similar to the proof of Lemma 5.2 in [8], but the remainder term r_k^n in (C.17) must be bounded differently, because we do not have the same explicit form for the transition probabilities of the Markov chain (Z_k). The proofs of Lemmas C.8 and C.9 are essentially the same as the proofs of their counterparts in [8], and are therefore omitted.

Proof of Lemma C.7. Part (i) follows directly from condition (D) in the statement of Proposition C.1 where the Markov chain (Z_k) is defined. To prove (ii), fix $a\in(1,2]$ and let $g\in C_c^\infty([0,\infty))$ be any non-negative function such that $g(t)=t^{1-\beta}$ for $t\in(\frac{2}{3a},\frac{3a}{2})$. Then, for each $n\in\mathbb{N}$, define a process $W^n\equiv(W_k^n)_{k\geq0}$ by

$$W_k^n=g\Big(\frac{Z_{k\wedge\gamma_n}}{a^n}\Big).$$

Let $\mathcal{F}_k=\sigma(Z_0,\dots,Z_k)\supseteq\sigma(W_0^n,\dots,W_k^n)$. At the end of the main proof we will establish the following two claims.

Claim 1: There exists some $B_1=B_1(a)>0$ such that

$$E_x^Z(\gamma_n)\leq B_1a^n,\ \text{for each } n\in\mathbb{N} \text{ and } x\in\mathcal{I}_n.$$

Claim 2: The process $(W_k^n)_{k\geq0}$ is "close" to being a martingale, when n is large, in the following sense: There exists some $B_2=B_2(a)>0$ such that

$$|E_x^Z(W_{k+1}^n|\mathcal{F}_k)-W_k^n|\leq B_2a^{-\frac{4}{3}n}\ a.s.,\quad \text{for each } k\in\mathbb{N}_0,\ n\in\mathbb{N}, \text{ and } x\in\mathcal{I}_n.$$

We now show how these two claims can be used to prove the lemma. Assume $Z_0=x\in\mathcal{I}_n$ and define a process $(\mathcal{R}_k^n)_{k\geq0}$ by

$$\mathcal{R}_0^n=0\quad\text{and}\quad\mathcal{R}_k^n=\sum_{j=1}^{k\wedge\gamma_n}\Big[E_x^Z(W_j^n|\mathcal{F}_{j-1})-W_{j-1}^n\Big]\ ,\ k\geq1.$$

Observe that $(W_k^n - \mathcal{R}_k^n)_{k\ge 0}$ is a martingale with initial value W_0^n. Moreover, by Claims 1 and 2,

$$\left|E_x^Z(\mathcal{R}_{\gamma_n}^n)\right| \le E_x^Z\left(\sum_{j=1}^{\gamma_n}\left|E_x^Z(W_j^n|\mathcal{F}_{j-1}) - W_{j-1}^n)\right|\right) \le E_x^Z(\gamma_n)\cdot B_2 a^{-\frac{4}{3}n} \le B_3 a^{-\frac{1}{3}n}. \quad \text{(C.14)}$$

Since $|\mathcal{R}_k^n| \le \sum_{j=1}^{\gamma_n}\left|E_x^Z(W_j^n|\mathcal{F}_{j-1}) - W_{j-1}^n)\right|$, for all k, this shows that $(\mathcal{R}_k^n)_{k\ge 0}$ is uniformly integrable, and $(W_k^n)_{k\ge 0}$ is also uniformly integrable, since $|W_k^n| \le \|g\|_\infty$ with probability 1. Thus, the martingale $(W_k^n - \mathcal{R}_k^n)_{k\ge 0}$ is itself uniformly integrable, so we may apply the optional stopping theorem to conclude

$$W_0^n = E_x^Z(W_{\gamma_n}^n) - E_x^Z(\mathcal{R}_{\gamma_n}^n). \quad \text{(C.15)}$$

Combining (C.14) and (C.15) and using the fact that $g(t)$ is equal to $t^{1-\beta}$ on $(\frac{2}{3a}, \frac{3a}{2})$ shows that

$$W_0^n - B_3 a^{-\frac{1}{3}n} \le E_x^Z(W_{\gamma_n}^n) \le \begin{cases} P_x^Z(Z_{\gamma_n} \in [a^{n+1}, a^{n+1} + a^{\frac{2}{3}(n-1)})) \cdot (a + a^{-\frac{1}{3}(n+2)})^{1-\beta} \\ +P_x^Z(Z_{\gamma_n} \in (a^{n-1} - a^{\frac{2}{3}(n-1)}, a^{n-1}]) \cdot a^{-(1-\beta)} \\ +E_x^Z[W_{\gamma_n}^n \cdot \mathbb{1}\{Z_{\gamma_n} \notin (a^{n-1} - a^{\frac{2}{3}(n-1)}, a^{n+1} + a^{\frac{2}{3}(n-1)})\}] \end{cases}$$

and

$$W_0^n + B_3 a^{-\frac{1}{3}n} \ge E_x^Z(W_{\gamma_n}^n) \ge \begin{cases} P_x^Z(Z_{\gamma_n} \in [a^{n+1}, a^{n+1} + a^{\frac{2}{3}(n-1)})) \cdot a^{1-\beta} \\ +P_x^Z(Z_{\gamma_n} \in (a^{n-1} - a^{\frac{2}{3}(n-1)}, a^{n-1}]) \cdot (a^{-1} - a^{-\frac{1}{3}(n+2)})^{1-\beta} \\ +E_x^Z[W_{\gamma_n}^n \cdot \mathbb{1}\{Z_{\gamma_n} \notin (a^{n-1} - a^{\frac{2}{3}(n-1)}, a^{n+1} + a^{\frac{2}{3}(n-1)})\}]. \end{cases}$$

Using part (i) and the fact that $W_{\gamma_n}^n$ is bounded by $\|g\|_\infty$ gives

$$W_0^n = (1-p)\cdot a^{1-\beta} \ + \ p \cdot a^{-(1-\beta)} \ + \ O(a^{-\frac{1}{3}n})$$

uniformly in the initial value $Z_0 = x \in \mathcal{I}_n$, where $p \equiv P_x^Z(Z_{\gamma_n} \le a^{n-1})$. Now, using the definition $W_0^n = g(Z_0/a^n) = g(x/a^n)$ shows also that $W_0^n = 1 + O(a^{-\frac{1}{3}n})$, uniformly in $x \in \mathcal{I}_n$. So, we have

$$1 = (1-p)\cdot a^{1-\beta} \ + \ p \cdot a^{-(1-\beta)} \ + \ O(a^{-\frac{1}{3}n})$$

uniformly in $x \in \mathcal{I}_n$. Solving for p gives $p = a^{1-\beta}/(1 + a^{1-\beta}) + O(a^{-\frac{1}{3}n})$, which implies (ii).

Proof of Claim 1: It will suffice to prove the claim for all sufficiently large n. Assume n is large enough that $\mathcal{I}_n \subset (a^{n-1}, a^{n+1})$, and define $\gamma_n^- = \inf\{k \ge 0 : Z_k \le a^{n-1}\}$. By monotonicity of the process $(Z_k)_{k\ge 0}$ with respect to its initial condition

$$P_z^Z(\gamma_n < a^n) \ge P_z^Z(\gamma_n^- < a^n) \ge P_{\lfloor a^{n+1}\rfloor}^Z(\gamma_n^- < a^n) \text{ , for all } z \in (a^{n-1}, a^{n+1}).$$

Further, by Lemma C.3-(ii), $\liminf_{n\to\infty} P_{\lfloor a^{n+1}\rfloor}^Z(\gamma_n^- < a^n) > 0$. So, there exist some $c > 0$ and $n_0 \in \mathbb{N}$ such that

$$P_z^Z(\gamma_n < a^n) \ge c \text{ , } \forall n \ge n_0 \text{ and } z \in (a^{n-1}, a^{n+1}). \quad \text{(C.16)}$$

Let $t_0 = 0$ and $t_{i+1} = t_i + \lceil a^n \rceil$, $i \ge 0$. By (C.16) $P_x^Z(\gamma_n > t_{i+1}|\gamma_n > t_i) \le 1 - c$, for all $n \ge n_0$, $x \in \mathcal{I}_n$, and $i \ge 0$. Thus, for each $n \ge n_0$, $x \in \mathcal{I}_n$, and $m \ge 0$

$$P_x^Z(\gamma_n > m \cdot \lceil a^n \rceil) = P_x^Z(\gamma_n > t_m) \le (1-c)^m.$$

This implies the claim.

Proof of Claim 2: Since $|W_k^n|$ is bounded by $\|g\|_\infty$, for all n,k, it will suffice to show the claim for sufficiently large n. Throughout we assume n is sufficiently large that $(\frac{a^{n-1}-a^{\frac{2}{3}n}}{a^n}, \frac{a^{n+1}+a^{\frac{2}{3}n}}{a^n}) \subset (\frac{2}{3a}, \frac{3a}{2})$ and that $Z_0 = x \in \mathcal{I}_n$. All $O(\cdot)$ estimates stated will be uniform in $x \in \mathcal{I}_n$ and $k \in \mathbb{N}_0$. By Taylor's Theorem,

$$g\Big(\frac{Z_{k+1}}{a^n}\Big) = g\Big(\frac{Z_k}{a^n}\Big) + g'\Big(\frac{Z_k}{a^n}\Big)\frac{Z_{k+1}-Z_k}{a^n} + \frac{1}{2}g''\Big(\frac{Z_k}{a^n}\Big)\frac{(Z_{k+1}-Z_k)^2}{a^{2n}} + \frac{1}{6}g'''(t)\frac{(Z_{k+1}-Z_k)^3}{a^{3n}}$$

where t is some random point between Z_k/a_n and Z_{k+1}/a^n. Thus, on the event $\{\gamma_n > k\}$, we have

$$\begin{aligned} E_x^Z(W_{k+1}^n|\mathcal{F}_k) - W_k^n &= E_x^Z\Big[g\Big(\frac{Z_{k+1}}{a^n}\Big)\Big|\mathcal{F}_k\Big] - g\Big(\frac{Z_k}{a^n}\Big) \\ &= \frac{1}{a^n}g'\Big(\frac{Z_k}{a^n}\Big)E_x^Z[Z_{k+1}-Z_k|\mathcal{F}_k] + \frac{1}{2a^{2n}}g''\Big(\frac{Z_k}{a^n}\Big)E_x^Z[(Z_{k+1}-Z_k)^2|\mathcal{F}_k] + r_k^n \\ &= \frac{1}{a^n}g'\Big(\frac{Z_k}{a^n}\Big)\Big[\alpha\beta + O\Big(a^{-\frac{1}{3}n}\Big)\Big] + \frac{1}{2a^{2n}}g''\Big(\frac{Z_k}{a^n}\Big)\Big[2\alpha Z_k + O\Big(a^{\frac{2}{3}n}\Big)\Big] + r_k^n \end{aligned} \quad \text{(C.17)}$$

by (C.2), where the remainder r_k^n satisfies

$$|r_k^n| \le \frac{1}{6}\|g'''\|_\infty E_x^Z\Big[\frac{|Z_{k+1}-Z_k|^3}{a^{3n}}\Big|\mathcal{F}_k\Big].$$

Now, for $t \in (\frac{2}{3a}, \frac{3a}{2})$, $tg''(t) = -\beta g'(t)$. So, for $k < \gamma_n$, $\frac{Z_k}{a^n}g''(\frac{Z_k}{a^n}) = -\beta g'(\frac{Z_k}{a^n})$. Plugging this relation back into (C.17) and simplifying we find that, on the event $\{\gamma_n > k\}$,

$$\begin{aligned} E_x^Z(W_{k+1}^n|\mathcal{F}_k) - W_k^n &= \frac{1}{a^n}g'\Big(\frac{Z_k}{a^n}\Big)\cdot O\Big(a^{-\frac{1}{3}n}\Big) \;+\; \frac{1}{2a^{2n}}g''\Big(\frac{Z_k}{a^n}\Big)\cdot O\Big(a^{\frac{2}{3}n}\Big) \;+\; r_k^n \\ &= O\Big(a^{-\frac{4}{3}n}\Big) + r_k^n. \end{aligned} \quad \text{(C.18)}$$

The remainder term r_k^n can be bounded as

$$\begin{aligned} |r_k^n| &\le \frac{\|g'''\|_\infty}{6a^{3n}} \max_{z\in(a^{n-1},a^{n+1})} E_x^Z\Big(|Z_{k+1}-Z_k|^3\Big|Z_k = z\Big) \\ &= \frac{\|g'''\|_\infty}{6a^{3n}} \max_{z\in(a^{n-1},a^{n+1})} E_z^Z(|Z_1 - z|^3). \end{aligned} \quad \text{(C.19)}$$

We split this last expectation into three pieces:

$$\begin{aligned} E_z^Z(|Z_1-z|^3) = E_z^Z\Big(\mathbb{1}\{Z_1 \le z/2\}\cdot|Z_1-z|^3\Big) + E_z^Z\Big(\mathbb{1}\{z/2 < Z_1 \le 2z\}\cdot|Z_1-z|^3\Big) \\ + E_z^Z\Big(\mathbb{1}\{Z_1 > 2z\}\cdot|Z_1-z|^3\Big). \end{aligned}$$

By (C.3) the first term on the right hand side is $O(e^{-z^{1/4}})$, and using (C.3) and calculations exactly as in the derivation of (B.5) and (B.10) shows that the second and third terms are, respectively, $O(z^{5/3})$ and $O(e^{-z^{1/4}})$. Plugging these estimates back into (C.19) gives $|r_k^n| = O(a^{-\frac{4}{3}n})$, and combining that with (C.18) shows also that

$$E_x^Z(W_{k+1}^n|\mathcal{F}_k) - W_k^n = O(a^{-\frac{4}{3}n})\ , \ \text{ on the event } \{\gamma_n > k\}.$$

Of course, on the event $\{\gamma_n \le k\}$, $W_{k+1}^n = W_k^n$ (deterministically), so this proves the claim. □

D Proof of Propositions 4.3 and 4.4

In this section we use Propositions 4.1 and 4.2 to prove Propositions 4.3 and 4.4. It is assumed throughout, if not otherwise specified, that $R_0 = s$ and $V_0 = 0$. Thus, $\tau_0 = 0$, $\widehat{V}_0 = 0$, and $\sigma_0^V = \tau_{\tau_0^{\widehat{V}}}$ (where $(\tau_k)_{k\geq 0}$ is as in (2.10)). The details of the proofs are somewhat technical, but the general ideas are fairly simple, so we will present these first before proceeding to the formal proofs. First, for Proposition 4.3, observe that if $\tau_0^{\widehat{V}} = m$, for some large m, then

$$\sigma_0^V = \tau_m = \sum_{k=1}^{m} (\tau_k - \tau_{k-1}) \approx m \cdot \mu_s.$$

So, by Proposition 4.1, for large x,

$$P_{0,s}^V(\sigma_0^V > x) \approx P_{0,s}^V(\tau_0^{\widehat{V}} > x/\mu_s) \approx c_{10} \cdot (x/\mu_s)^{-\delta} = c_{12} \cdot x^{-\delta}.$$

Next, for Proposition 4.4, note that if $\sum_{k=0}^{\sigma_0^V} V_k$ is large, generally it will be the case that σ_0^V is large as well, and also that V_k will be relatively large for most times k between 0 and σ_0^V (because the V_k process is unlikely to remain close to 0 very long without hitting 0). Thus, by Lemma 2.11, the $\widehat{V}_j$ process, and also the V_k process, will not fluctuate too much too rapidly, relative to their current values. So, very roughly speaking, we should expect that

$$\sum_{k=0}^{\sigma_0^V} V_k = \sum_{j=0}^{\tau_0^{\widehat{V}}-1} \sum_{k=\tau_j}^{\tau_{j+1}-1} V_k \approx \sum_{j=0}^{\tau_0^{\widehat{V}}-1} \sum_{k=\tau_j}^{\tau_{j+1}-1} \widehat{V}_j \approx \sum_{j=0}^{\tau_0^{\widehat{V}}-1} \widehat{V}_j \cdot \mu_s = \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j \cdot \mu_s,$$

when either the sum on the right hand side or left hand side (equivalently both sums) are large. Therefore, by Proposition 4.2, we should expect that, for large x,

$$P_{0,s}^V\left(\sum_{k=0}^{\sigma_0^V} V_k > x\right) \approx P_{0,s}^V\left(\sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j > x/\mu_s\right) \approx c_{11} \cdot (x/\mu_s)^{-\delta/2} = c_{13} \cdot x^{-\delta/2}.$$

Proof of Proposition 4.3. Fix any $\epsilon > 0$. It will suffice to show that

$$\limsup_{n\to\infty} n^\delta \cdot P_{0,s}^V(\sigma_0^V > n) \leq c_{12} + \epsilon \tag{D.1}$$

and

$$\liminf_{n\to\infty} n^\delta \cdot P_{0,s}^V(\sigma_0^V > n) \geq c_{12} - \epsilon. \tag{D.2}$$

Proof of (D.1): Choose $\rho > 0$ sufficiently small that $(1+\rho)^\delta \mu_s^\delta (c_{10} + \rho) + \rho \leq c_{10}\mu_s^\delta + \epsilon = c_{12} + \epsilon$. For $n \in \mathbb{N}$, let $m = m(n) = (1+\rho)\mu_s n$ (it is not assumed that m is an integer). Then $\{\sigma_0^V > m\} \subseteq \{\tau_0^{\widehat{V}} > n\} \cup \{\tau_n > m\}$. So,

$$m^\delta P_{0,s}^V(\sigma_0^V > m) \leq m^\delta \left[P_{0,s}^V(\tau_0^{\widehat{V}} > n) + P_{0,s}^V(\tau_n > m)\right]. \tag{D.3}$$

By Proposition 4.1, for all sufficiently large n,

$$m^\delta P_{0,s}^V(\tau_0^{\widehat{V}} > n) = \frac{m^\delta}{n^\delta} \cdot \left[n^\delta P_{0,s}^V(\tau_0^{\widehat{V}} > n)\right] \leq (1+\rho)^\delta \mu_s^\delta \cdot (c_{10} + \rho). \tag{D.4}$$

Also, by Lemma 2.5, $P_{0,s}^V(\tau_n > m) = P_{0,s}^V\big(\sum_{i=1}^n(\tau_i - \tau_{i-1}) > (1+\rho)\mu_s n\big)$ decays exponentially in n, since the random variables $(\tau_i - \tau_{i-1})_{i\geq 1}$ are i.i.d. with exponential tails and mean μ_s. Thus, for all sufficiently large n,

$$m^\delta P_{0,s}^V(\tau_n > m) \leq \rho. \tag{D.5}$$

Combining the estimates (D.3), (D.4), and (D.5) shows that, for all sufficiently large n,

$$m^\delta P_{0,s}^V(\sigma_0^V > m) \leq (1+\rho)^\delta \mu_s^\delta(c_{10}+\rho) + \rho \leq c_{12} + \epsilon,$$

which proves (D.1).

Proof of (D.2): Choose $\rho \in (0,1)$ sufficiently small that $(1-\rho)^\delta \mu_s^\delta(c_{10}-\rho) - \rho \geq c_{10}\mu_s^\delta - \epsilon = c_{12} - \epsilon$. For $n \in \mathbb{N}$, let $m = m(n) = (1-\rho)\mu_s n$ (again, it is not assumed that m is an integer). Then $\{\sigma_0^V > m\} \supseteq \{\tau_0^{\widehat{V}} > n\} \cap \{\tau_n \geq m\}$. So,

$$P_{0,s}^V(\sigma_0^V > m) \geq P_{0,s}^V(\tau_0^{\widehat{V}} > n, \tau_n \geq m) \geq P_{0,s}^V(\tau_0^{\widehat{V}} > n) - P_{0,s}^V(\tau_n < m). \tag{D.6}$$

Now, by Proposition 4.1, $P_{0,s}^V(\tau_0^{\widehat{V}} > n) \geq n^{-\delta}(c_{10} - \rho)$, for all sufficiently large n. Also, by Lemma 2.5, $P_{0,s}^V(\tau_n < m) = P_{0,s}^V\big(\sum_{i=1}^n(\tau_i - \tau_{i-1}) < (1-\rho)\mu_s n\big)$ decays exponentially in n. So, $P_{0,s}^V(\tau_n < m) \leq \rho \cdot m^{-\delta}$, for all sufficiently large n. Plugging these estimates back into (D.6) shows that,

$$m^\delta P_{0,s}^V(\sigma_0^V > m) \geq \left(\frac{m}{n}\right)^\delta (c_{10}-\rho) - \rho = (1-\rho)^\delta \mu_s^\delta (c_{10}-\rho) - \rho \geq c_{12} - \epsilon$$

for all sufficiently large n, which proves (D.2). □

Proof of Proposition 4.4. Fix any $\epsilon_1 \in (0, \frac{1}{22})$ and $\epsilon_2 \in (0, \frac{1}{4}\epsilon_1)$. Then, for $n \in \mathbb{N}$, define the following random variables:

- $T_0 = 0$ and $T_{i+1} = \inf\{k \geq T_i + \lfloor n^{\epsilon_1} \rfloor : R_k = s\}$, $i \geq 0$.
- $i_{\max} = \max\{i \geq 0 : T_i \leq \sigma_0^V\}$, $k_{\max} = T_{i_{\max}}$, and $j_{\max}$ is the unique j such that $\tau_{j_{\max}} = k_{\max}$.
- $K_i = \{T_{i-1}, T_{i-1}+1, \ldots, T_i\}$ and $J_i = \{j \in \mathbb{N}_0 : \tau_j \in K_i\}$, $i \geq 1$.
- $j_i^{\max} = \max\{j : j \in J_i\}$ and $j_i^{\min} = \min\{j : j \in J_i\}$.
- $J_i^0 = J_i \backslash \{j_i^{\max}\}$ and $\widetilde{J}_i = \{j_i^{\min}, j_i^{\min}+1, \ldots, j_i^{\min} + \lfloor 2n^{\epsilon_1} \rfloor\}$.
- $K_i^0 = K_i \backslash \{T_i\}$ and $\widetilde{K}_i = \{\tau_{j_i^{\min}}, \tau_{j_i^{\min}}+1, \ldots, \tau_{j_i^{\min}+\lfloor 2n^{\epsilon_1}\rfloor}\} = \{T_{i-1}, T_{i-1}+1, \ldots, \tau_{j_i^{\min}+\lfloor 2n^{\epsilon_1}\rfloor}\}$.

Also, denote $V_{\max} = \max\{V_k : 0 \leq k \leq \sigma_0^V\}$ and $\Delta_{\tau,j} = \tau_j - \tau_{j-1}$, for $j \geq 1$, and define the following events:

- $E_i = \big\{\max_{k \in K_i} |V_{T_{i-1}} - V_k| > 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)}\big\}$, $i \geq 1$.
- $F_i = \left\{\left|\sum_{j \in J_i^0}(\mu_s - \Delta_{\tau,j+1})\right| > n^{\frac{1}{2}\epsilon_1 + \epsilon_2}\right\}$, $i \geq 1$.
- $A_1 = \{\sigma_0^V > n^{\frac{1}{2}(1+\epsilon_2)}\}$.
- $A_2 = \{V_{\max} > 2n^{\frac{1}{2}(1+\epsilon_2)}\}$.
- $A_3 = \{\exists 1 \leq i \leq i_{\max}$ such that E_i occurs$\}$.
- $A_4 = \{\exists 1 \leq i \leq i_{\max}$ such that F_i occurs$\}$.
- $A_5 = \{\exists 1 \leq i \leq i_{\max}+1$ such that $T_i - T_{i-1} > 2n^{\epsilon_1}\}$.
- $G = A_1^c \cap A_2^c \cap A_3^c \cap A_4^c \cap A_5^c$ (the "good event").

At the end of the main proof we will establish the following claim.

Claim 1: For all sufficiently large n, $P_{0,s}^V(G^c) \le 5n^{-(\frac{\delta}{2}+\frac{\delta\epsilon_2}{4})}$.

By the triangle inequality,

$$\left|\sum_{k=0}^{\sigma_0^V} V_k - \mu_s \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j\right| \le \left|\sum_{k=0}^{\sigma_0^V} V_k - \sum_{i=1}^{i_{\max}} V_{T_{i-1}}(T_i - T_{i-1})\right| + \left|\sum_{i=1}^{i_{\max}} V_{T_{i-1}}(T_i - T_{i-1}) - \mu_s \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j\right|$$

$$\le \begin{cases} \left|\sum_{k=0}^{k_{\max}-1} V_k - \sum_{i=1}^{i_{\max}} V_{T_{i-1}}(T_i - T_{i-1})\right| + \sum_{k=k_{\max}}^{\sigma_0^V} V_k \\ + \left|\sum_{j=0}^{j_{\max}-1} \mu_s \widehat{V}_j - \sum_{i=1}^{i_{\max}} V_{T_{i-1}}(T_i - T_{i-1})\right| + \mu_s \sum_{j=j_{\max}}^{\tau_0^{\widehat{V}}} \widehat{V}_j \end{cases}$$

$$\equiv (I) + (II) + (III) + (IV).$$

We now show how each of the terms (I), (II), (III), and (IV) can be bounded on the event G.

Bound on (I): On the event G,

$$(I) \equiv \left|\sum_{k=0}^{k_{\max}-1} V_k - \sum_{i=1}^{i_{\max}} V_{T_{i-1}}(T_i - T_{i-1})\right| = \left|\sum_{i=1}^{i_{\max}} \sum_{k \in K_i^0} V_k - \sum_{i=1}^{i_{\max}} \sum_{k \in K_i^0} V_{T_{i-1}}\right|$$

$$\overset{(a)}{\le} \sum_{i=1}^{i_{\max}} \sum_{k \in K_i^0} 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)} = k_{\max} \cdot 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)} \le \sigma_0^V \cdot 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)}$$

$$\overset{(b)}{\le} n^{\frac{1}{2}(1+\epsilon_2)} 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)} \overset{(c)}{\le} 2n^{\frac{5}{6}+\frac{11}{6}\epsilon_1}.$$

Step (a) follows from the fact that $G \subset A_3^c$, step (b) follows from the fact that $G \subset A_1^c$, and step (c) follows from the fact that $\epsilon_2 < \epsilon_1$.

Bound on (II) and (IV): First note that $(IV) \equiv \mu_s \cdot \sum_{j=j_{\max}}^{\tau_0^{\widehat{V}}} \widehat{V}_j \le \mu_s \cdot \sum_{k=k_{\max}}^{\sigma_0^V} V_k \equiv \mu_s \cdot (II)$, so it will suffice to bound (II). Now, on the event G,

$$(II) \le V_{\max} |\sigma_0^V - k_{\max}| \overset{(a)}{\le} 2n^{\frac{1}{2}(1+\epsilon_2)} \cdot 2n^{\epsilon_1} \overset{(b)}{\le} 4n^{\frac{1}{2}(1+3\epsilon_1)}.$$

Step (a) follows from the fact that $G \subset A_2^c$ and $G \subset A_5^c$, and step (b) follows from the fact that $\epsilon_2 < \epsilon_1$.

Bound on (III): On the event G,

$$(III) \equiv \left|\sum_{j=0}^{j_{\max}-1} \mu_s \widehat{V}_j \;-\; \sum_{i=1}^{i_{\max}} V_{T_{i-1}}(T_i - T_{i-1})\right|$$

$$= \left|\sum_{i=1}^{i_{\max}} \sum_{j \in J_i^0} \mu_s \widehat{V}_j \;-\; \sum_{i=1}^{i_{\max}} \sum_{j \in J_i^0} V_{T_{i-1}} \cdot \Delta_{\tau,j+1}\right|$$

$$= \left|\sum_{i=1}^{i_{\max}} \sum_{j \in J_i^0} \mu_s (\widehat{V}_j - V_{T_{i-1}}) \;+\; \sum_{i=1}^{i_{\max}} \sum_{j \in J_i^0} V_{T_{i-1}}(\mu_s - \Delta_{\tau,j+1})\right|$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{i_{\max}} \sum_{j \in J_i^0} \mu_s |\widehat{V}_j - V_{T_{i-1}}| \;+\; \sum_{i=1}^{i_{\max}} V_{T_{i-1}} \left| \sum_{j \in J_i^0} (\mu_s - \Delta_{\tau,j+1}) \right| \\
&\overset{(a)}{\leq} \sum_{i=1}^{i_{\max}} \sum_{j \in J_i^0} \mu_s \cdot 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)} \;+\; \sum_{i=1}^{i_{\max}} 2n^{\frac{1}{2}(1+\epsilon_2)} \cdot n^{\epsilon_1/2+\epsilon_2} \\
&\overset{(b)}{\leq} \tau_0^{\widehat{V}} \cdot \mu_s \cdot 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)} \;+\; \frac{\sigma_0^V}{\lfloor n^{\epsilon_1} \rfloor} 2n^{\frac{1}{2}(1+\epsilon_2)} n^{\epsilon_1/2+\epsilon_2} \\
&\overset{(c)}{\leq} n^{\frac{1}{2}(1+\epsilon_2)} \cdot \mu_s \cdot 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)} \;+\; \left(n^{\frac{1}{2}(1+\epsilon_2)} 2n^{-\epsilon_1} \right) \cdot 2n^{\frac{1}{2}(1+\epsilon_2)} n^{\epsilon_1/2+\epsilon_2} \\
&\overset{(d)}{\leq} 2\mu_s n^{\frac{5}{6}+\frac{11}{6}\epsilon_1} \;+\; 4n^{1+2\epsilon_2-\epsilon_1/2}.
\end{aligned}
$$

Step (a) follows from the fact that $G \subset A_2^c$, $G \subset A_3^c$, and $G \subset A_4^c$. Step (b) follows from the relations $j_{\max} \leq \tau_0^{\widehat{V}}$ and $i_{\max} \leq k_{\max}/\lfloor n^{\epsilon_1} \rfloor \leq \sigma_0^V / \lfloor n^{\epsilon_1} \rfloor$. Step (c) follows from the inequality $\tau_0^{\widehat{V}} \leq \sigma_0^V$ and the fact that $G \subset A_1^c$. Finally, Step (d) follows from the fact that $\epsilon_2 < \epsilon_1$.

Now, let $\alpha = \max\left\{ (\frac{5}{6} + \frac{11}{6}\epsilon_1), (1 + 2\epsilon_2 - \frac{\epsilon_1}{2}) \right\}$. By the choice of ϵ_1 and ϵ_2, we have $\frac{1}{2}(1+3\epsilon_1) < \alpha$ and $\alpha < 1$. Combining the estimates on the terms (I)-(IV) we find that, on the event G,

$$
\begin{aligned}
\Big| \sum_{k=0}^{\sigma_0^V} V_k - \mu_s \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j \Big| &\leq (I) + (II) + (III) + (IV) \\
&\leq 2n^\alpha + 4n^\alpha + [2\mu_s n^\alpha + 4n^\alpha] + 4\mu_s n^\alpha \leq 16\mu_s n^\alpha.
\end{aligned}
$$

Using this estimate along with Claim 1 and Proposition 4.2 we can now establish the proposition.

Given any $\epsilon \in (0,1)$:

- Choose $N_1 \in \mathbb{N}$ such that $P_{0,s}^V(G^c) \leq 5n^{-(\frac{\delta}{2}+\frac{\delta\epsilon_2}{4})}$, for $n \geq N_1$ (possible by Claim 1).
- Choose $N_2 \in \mathbb{N}$ such that $16\mu_s n^\alpha \leq \epsilon n$, for $n \geq N_2$ (possible since $\alpha < 1$).
- Finally, choose $N_3 \in \mathbb{N}$ such that, for all $n \geq N_3$,

$$
\begin{aligned}
P_{0,s}^V\Big(\sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j > \frac{1-\epsilon}{\mu_s} \cdot n \Big) &\leq (c_{11}+\epsilon) \left(\frac{1-\epsilon}{\mu_s} \cdot n \right)^{-\delta/2} \quad \text{and} \\
P_{0,s}^V\Big(\sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j > \frac{1+\epsilon}{\mu_s} \cdot n \Big) &\geq (c_{11}-\epsilon) \left(\frac{1+\epsilon}{\mu_s} \cdot n \right)^{-\delta/2}.
\end{aligned}
$$

  This is possible by Proposition 4.2.

Then for each $n \geq N_0 \equiv \max\{N_1, N_2, N_3\}$ we have

$$
\begin{aligned}
&P_{0,s}^V\Big(\sum_{k=0}^{\sigma_0^V} V_k > n \Big) \leq P_{0,s}^V\Big(\mu_s \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j > (1-\epsilon)n \Big) + P_{0,s}^V\Big(\Big| \sum_{k=0}^{\sigma_0^V} V_k - \mu_s \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j \Big| > \epsilon n \Big) \\
&\leq (c_{11}+\epsilon) \left(\frac{1-\epsilon}{\mu_s} \cdot n \right)^{-\delta/2} + P_{0,s}^V(G^c) \leq (c_{11}+\epsilon) \left(\frac{1-\epsilon}{\mu_s} \cdot n \right)^{-\delta/2} + 5n^{-(\frac{\delta}{2}+\frac{\delta\epsilon_2}{4})}
\end{aligned}
$$

and

$$P^V_{0,s}\Big(\sum_{k=0}^{\sigma_0^V} V_k > n\Big) \geq P^V_{0,s}\Big(\mu_s \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j > (1+\epsilon)n\Big) - P^V_{0,s}\Big(\Big|\sum_{k=0}^{\sigma_0^V} V_k - \mu_s \sum_{j=0}^{\tau_0^{\widehat{V}}} \widehat{V}_j\Big| > \epsilon n\Big)$$
$$\geq (c_{11}-\epsilon)\left(\frac{1+\epsilon}{\mu_s}\cdot n\right)^{-\delta/2} - P^V_{0,s}(G^c) \geq (c_{11}-\epsilon)\left(\frac{1+\epsilon}{\mu_s}\cdot n\right)^{-\delta/2} - 5n^{-(\frac{\delta}{2}+\frac{\delta\epsilon_2}{4})}.$$

Since $\epsilon \in (0,1)$ was arbitrary it follows that

$$\lim_{n\to\infty} n^{\delta/2}\cdot P^V_{0,s}\Big(\sum_{k=0}^{\sigma_0^V} V_k > n\Big) = \mu_s^{\delta/2}\cdot c_{11} = c_{13}.$$

This concludes the main proof of the proposition, and it remains only now to show Claim 1. To do this, though, we will first need to establish some auxiliary claims that will be used in its proof.

Claim 2: For all sufficiently large n,

$$P^V_{0,s}(E_i|V_{T_{i-1}} = x) \leq e^{-n^{\epsilon_1/2}}, \text{ for each } 0 \leq x \leq n^{\frac{1}{2}(1+\epsilon_1)} \text{ and } i \geq 1.$$

Proof. By Lemma 2.11, we have that for all sufficiently large n

$$P^V_{0,s}\Big(\max_{\tau_j\leq k\leq\tau_{j+1}} |V_{\tau_j} - V_k| > n^{\frac{1}{3}(1+\epsilon_1)}\Big| V_{\tau_j} = x\Big) \leq e^{-n^{\frac{1}{19}(1+\epsilon_1)}},\ 0 \leq x \leq 2n^{\frac{1}{2}(1+\epsilon_1)}. \quad \text{(D.7)}$$

For $i \geq 1$, define $T_0^{(i)} = T_{i-1}$ and $T_j^{(i)} = \inf\{k > T_{j-1}^{(i)} : R_k = s\}$, $j \geq 1$. Then, let

$$A_j^{(i)} = \Big\{\max_{T_{j-1}^{(i)}\leq k\leq T_j^{(i)}} |V_{T_{j-1}^{(i)}} - V_k| \leq n^{\frac{1}{3}(1+\epsilon_1)}\Big\}.$$

Observe that, for all sufficiently large n and $1 \leq j \leq 2n^{\epsilon_1}$, if $V_{T_{i-1}} \leq n^{\frac{1}{2}(1+\epsilon_1)}$ and $A_1^{(i)},\ldots,A_j^{(i)}$ all occur then $V_{T_j^{(i)}} \leq n^{\frac{1}{2}(1+\epsilon_1)} + j\cdot n^{\frac{1}{3}(1+\epsilon_1)} \leq 2n^{\frac{1}{2}(1+\epsilon_1)}$. Thus, by (D.7), for all sufficiently large n and $0 \leq x \leq n^{\frac{1}{2}(1+\epsilon_1)}$ we have

$$P^V_{0,s}\Big(\max_{k\in\widetilde{K}_i} |V_{T_{i-1}} - V_k| > 2n^{\epsilon_1} n^{\frac{1}{3}(1+\epsilon_1)}\Big| V_{T_{i-1}} = x\Big)$$
$$\leq P^V_{0,s}\left(\exists 1 \leq j \leq 2n^{\epsilon_1} : \big(A_j^{(i)}\big)^c \text{ occurs} \Big| V_{T_{i-1}} = x\right)$$
$$\leq \sum_{j=1}^{\lfloor 2n^{\epsilon_1}\rfloor} P^V_{0,s}\left(\big(A_j^{(i)}\big)^c \Big| A_1^{(i)},\ldots,A_{j-1}^{(i)}, V_{T_{i-1}} = x\right) \leq 2n^{\epsilon_1}\cdot e^{-n^{\frac{1}{19}(1+\epsilon_1)}}. \quad \text{(D.8)}$$

Also, since $|\widetilde{K}_i| \geq \lfloor 2n^{\epsilon_1}\rfloor$ (deterministically) and the event $\{K_i \not\subset \widetilde{K}_i\}$ is independent of the value of $V_{T_{i-1}}$ it follows from (2.26) that, for any x,

$$P^V_{0,s}(K_i \not\subset \widetilde{K}_i|V_{T_{i-1}} = x) = P^V_{0,s}(K_i \not\subset \widetilde{K}_i) \leq P^V_{0,s}(|K_i| > \lfloor 2n^{\epsilon_1}\rfloor) \leq c_3 e^{-c_4\lfloor n^{\epsilon_1}\rfloor}. \quad \text{(D.9)}$$

Combing the estimates (D.8) and (D.9) shows that, for all sufficiently large n and $0 \leq x \leq n^{\frac{1}{2}(1+\epsilon_1)}$,

$$P^V_{0,s}(E_i|V_{T_{i-1}} = x) \leq 2n^{\epsilon_1} e^{-n^{\frac{1}{19}(1+\epsilon_1)}} + c_3 e^{-c_4\lfloor n^{\epsilon_1}\rfloor} \leq e^{-n^{\epsilon_1/2}}.$$

Claim 3: For all sufficiently large n,

$$P_{0,s}^V(F_i) \le e^{-n^{\epsilon_1\epsilon_2}}, \text{ for each } i \ge 1.$$

Proof. Define $j_0 \equiv j_1^{\max} = \inf\{j : \tau_j \ge \lfloor n^{\epsilon_1} \rfloor\} = \inf\{j : \sum_{\ell=1}^{j} \Delta_{\tau,\ell} \ge \lfloor n^{\epsilon_1} \rfloor\}$. Since the (R_k) process is Markovian,

$$P_{0,s}^V(F_i) = P_{0,s}^V(F_1) = P_{0,s}^V\Big(\Big|\sum_{j=1}^{j_0}(\mu_s - \Delta_{\tau,j})\Big| > n^{\frac{1}{2}\epsilon_1+\epsilon_2}\Big) \ , \ \text{ for all } i \ge 1.$$

Let $N_n^+ = \lfloor (n^{\epsilon_1} + n^{\frac{1}{2}(\epsilon_1+\epsilon_2)})/\mu_s \rfloor$ and $N_n^- = \lfloor (n^{\epsilon_1} - n^{\frac{1}{2}(\epsilon_1+\epsilon_2)})/\mu_s \rfloor$. Define events B_1 and B_2 by

$$B_1 = \Big\{\Big|\sum_{j=1}^{N_n^+} \Delta_{\tau,j} - \mu_s N_n^+\Big| \le (N_n^+)^{\frac{1}{2}+\epsilon_2}\Big\} \quad \text{and} \quad B_2 = \Big\{\Big|\sum_{j=1}^{N_n^-} \Delta_{\tau,j} - \mu_s N_n^-\Big| \le (N_n^-)^{\frac{1}{2}+\epsilon_2}\Big\}.$$

Since the random variables $(\Delta_{\tau,j})_{j\ge1}$ are i.i.d. with mean μ_s and exponentials tails (due to (2.26)), it follows from Lemma 2.5 that there exist some constants $C_1, C_2 > 0$ such that

$$P_{0,s}^V\Big(\Big|\sum_{j=1}^{m}(\Delta_{\tau,j} - \mu_s)\Big| > \epsilon m\Big) \le C_1 e^{-C_2\epsilon^2 m} \ , \ \text{ for all } 0 < \epsilon < 1 \text{ and } m \in \mathbb{N}.$$

Using this with $m = N_n^+, N_n^-$ and $\epsilon = (N_n^+)^{-1/2+\epsilon_2}, (N_n^-)^{-1/2+\epsilon_2}$, respectively, shows that

$$P_{0,s}^V(B_1^c) \le C_1 e^{-C_2 (N_n^+)^{2\epsilon_2}} \quad \text{and} \quad P_{0,s}^V(B_2^c) \le C_1 e^{-C_2 (N_n^-)^{2\epsilon_2}}.$$

Hence, for all sufficiently large n,

$$P_{0,s}^V\big((B_1 \cap B_2)^c\big) \le C_1 e^{-C_2 (N_n^+)^{2\epsilon_2}} + C_1 e^{-C_2 (N_n^-)^{2\epsilon_2}} \le e^{-n^{\epsilon_1\epsilon_2}}.$$

So, it will suffice to show that, for all sufficiently large n,

$$\Big|\sum_{j=1}^{j_0}(\Delta_{\tau,j} - \mu_s)\Big| \le n^{\frac{1}{2}\epsilon_1+\epsilon_2} \quad \text{on the event } B_1 \cap B_2. \tag{D.10}$$

Now, since $\epsilon_1 < 1/2$, $n^{\epsilon_1\epsilon_2} < n^{\frac{1}{2}\epsilon_2}$. Using this fact and a little bit of algebra it follows from the definitions of B_1 and B_2 that, for all sufficiently large n, on the event $B_1 \cap B_2$

$$n^{\epsilon_1} < \sum_{j=1}^{N_n^+} \Delta_{\tau,j} < n^{\epsilon_1} + 2n^{\frac{1}{2}(\epsilon_1+\epsilon_2)} \quad \text{and} \quad n^{\epsilon_1} - 2n^{\frac{1}{2}(\epsilon_1+\epsilon_2)} < \sum_{j=1}^{N_n^-} \Delta_{\tau,j} < n^{\epsilon_1}.$$

Together these inequalities imply $N_n^- \le j_0 \le N_n^+$. So, by the definitions of N_n^+ and N_n^-,

$$\sum_{j=1}^{j_0}(\Delta_{\tau,j} - \mu_s) < [n^{\epsilon_1} + 2n^{\frac{1}{2}(\epsilon_1+\epsilon_2)}] - N_n^-\mu_s \le 4n^{\frac{1}{2}(\epsilon_1+\epsilon_2)} \text{ and}$$
$$\sum_{j=1}^{j_0}(\Delta_{\tau,j} - \mu_s) > [n^{\epsilon_1} - 2n^{\frac{1}{2}(\epsilon_1+\epsilon_2)}] - N_n^+\mu_s \ge -4n^{\frac{1}{2}(\epsilon_1+\epsilon_2)}$$

on the event $B_1 \cap B_2$, for all sufficiently large n. Since $4n^{\frac{1}{2}(\epsilon_1+\epsilon_2)} \le n^{\frac{1}{2}\epsilon_1+\epsilon_2}$ for all sufficiently large n, this shows that (D.10) holds for all sufficiently large n.

Claim 4: Denote $\widehat{V}_{\max} = \max\{\widehat{V}_j : 0 \le j \le \tau_0^{\widehat{V}}\}$. There exists some $c_{15} > 0$ such that

$$P_{0,s}^V(\widehat{V}_{\max} > t) \le c_{15} t^{-\delta}, \quad \text{for all } t \in [0, \infty). \tag{D.11}$$

Proof. As noted above in the proof of Propositions 4.1 and 4.2 in Appendix C, the Markov chain $(\widehat{V}_k)_{k \ge 0}$ satisfies all conditions (A)-(D) of Proposition C.1 with $\alpha = \mu_s$ and $\beta = 1 - \delta$. Thus, all the lemmas in Appendix C which hold for a Markov chain (Z_k) satisfying these properties apply to $(\widehat{V}_k)$. We will use Lemma C.9. By this lemma there exists some constant $C = C(0)$ such that

$$P_{0,s}^V(\widehat{V}_{\max} > n) = P_{0,s}^V(\tau_{(n+1)^+}^{\widehat{V}} < \tau_0^{\widehat{V}}) \le C(n+1)^{-\delta}, \quad \text{for all } n \in \mathbb{N}.$$

This implies (D.11).

Proof of Claim 1. We will show that $P_{0,s}^V(A_i) \le n^{-(\frac{\delta}{2} + \frac{\delta \epsilon_2}{4})}$, for each $i = 1, \ldots, 5$. The estimate for A_1 follows directly from Proposition 4.3. The bounds for the other events are given below.

Bound for A_2:
We decompose $P_{0,s}^V(V_{\max} > 2n^{\frac{1}{2}(1+\epsilon_2)})$ as

$$P_{0,s}^V\left(V_{\max} > 2n^{\frac{1}{2}(1+\epsilon_2)}\right) = \begin{cases} P_{0,s}^V(\widehat{V}_{\max} > n^{\frac{1}{2}(1+\epsilon_2)}, V_{\max} > 2n^{\frac{1}{2}(1+\epsilon_2)}) \\ +P_{0,s}^V(\widehat{V}_{\max} \le n^{\frac{1}{2}(1+\epsilon_2)}, V_{\max} > 2n^{\frac{1}{2}(1+\epsilon_2)}, \tau_0^{\widehat{V}} > n) \\ +P_{0,s}^V(\widehat{V}_{\max} \le n^{\frac{1}{2}(1+\epsilon_2)}, V_{\max} > 2n^{\frac{1}{2}(1+\epsilon_2)}, \tau_0^{\widehat{V}} \le n) \end{cases}$$
$$\equiv (I) + (II) + (III).$$

By Claim 4, $(I) \le P_{0,s}^V(\widehat{V}_{\max} > n^{\frac{1}{2}(1+\epsilon_2)}) \le c_{15}(n^{\frac{1}{2}(1+\epsilon_2)})^{-\delta}$. Also, by Proposition 4.1, $(II) \le P_{0,s}^V(\tau_0^{\widehat{V}} > n) \le 2c_{10} n^{-\delta}$, for all sufficiently large n. Term (III) is estimated as follows:

$$\begin{aligned} (III) &\le P_{0,s}^V\left(\exists\, 0 \le j \le n-1 : \widehat{V}_j \equiv V_{\tau_j} \le n^{\frac{1}{2}(1+\epsilon_2)} \text{ and } \max_{\tau_j \le k \le \tau_{j+1}} |V_{\tau_j} - V_k| > n^{\frac{1}{2}(1+\epsilon_2)}\right) \\ &\le \sum_{j=0}^{n-1} P_{0,s}^V\left(V_{\tau_j} \le n^{\frac{1}{2}(1+\epsilon_2)} \text{ and } \max_{\tau_j \le k \le \tau_{j+1}} |V_{\tau_j} - V_k| > n^{\frac{1}{2}(1+\epsilon_2)}\right) \\ &\le \sum_{j=0}^{n-1} P_{0,s}^V\left(\max_{\tau_j \le k \le \tau_{j+1}} |V_{\tau_j} - V_k| > n^{\frac{1}{2}(1+\epsilon_2)} \Big| V_{\tau_j} \le n^{\frac{1}{2}(1+\epsilon_2)}\right) \\ &\le n \cdot \left[\max_{0 \le x \le n^{\frac{1}{2}(1+\epsilon_2)}} P_{x,s}^V\left(\max_{0 \le k \le \tau_s^R} |V_k - x| > n^{\frac{1}{2}(1+\epsilon_2)}\right)\right]. \end{aligned} \tag{D.12}$$

By (2.49) the right hand side of (D.12) is at most $ne^{-n^{1/6}}$, for all sufficiently large n (note that although (2.49) is not directly applicable when $x = 0$, $P_{0,s}^V(\max_{0 \le k \le \tau_s^R} V_k > t) \le P_{1,s}^V(\max_{0 \le k \le \tau_s^R} V_k > t)$, for any $t > 0$, so the bound still holds in this case as well). Combining these estimates on terms (I), (II), and (III) we find that, for all sufficiently large n,

$$P_{0,s}^V(A_2) \equiv P_{0,s}^V\left(V_{\max} > 2n^{\frac{1}{2}(1+\epsilon_2)}\right) \le c_{15}(n^{\frac{1}{2}(1+\epsilon_2)})^{-\delta} + 2c_{10} n^{-\delta} + ne^{-n^{1/6}} \le n^{-(\delta/2 + \delta\epsilon_2/4)}.$$

Bound for A_3:
By construction, $i_{\max} \le \tau_0^{\widehat{V}}$. So, by Proposition 4.1, for all sufficiently large n

$$P_{0,s}^V(i_{\max} > n) \le P_{0,s}^V(\tau_0^{\widehat{V}} > n) \le 2c_{10}n^{-\delta}. \tag{D.13}$$

Combining this with Claim 2 and Claim 4 shows that, for all sufficiently large n,

$$\begin{aligned} P_{0,s}^V(A_3) &= \begin{cases} P_{0,s}^V(A_3, \widehat{V}_{\max} > n^{\frac{1}{2}(1+\epsilon_2)}) + P_{0,s}^V(A_3, \widehat{V}_{\max} \le n^{\frac{1}{2}(1+\epsilon_2)}, i_{\max} > n) \\ + P_{0,s}^V(A_3, \widehat{V}_{\max} \le n^{\frac{1}{2}(1+\epsilon_2)}, i_{\max} \le n) \end{cases} \\ &\le \begin{cases} P_{0,s}^V(\widehat{V}_{\max} > n^{\frac{1}{2}(1+\epsilon_2)}) + P_{0,s}^V(i_{\max} > n) \\ + P_{0,s}^V(\exists 1 \le i \le n : V_{T_{i-1}} \le n^{\frac{1}{2}(1+\epsilon_2)} \text{ and } E_i \text{ occurs}) \end{cases} \\ &\le c_{15}(n^{\frac{1}{2}(1+\epsilon_2)})^{-\delta} + 2c_{10}n^{-\delta} + ne^{-n^{\epsilon_1/2}} \\ &\le n^{-(\delta/2+\delta\epsilon_2/4)}. \end{aligned}$$

Bound for A_4:
By (D.13) and Claim 3 we have, for all sufficiently large n,

$$\begin{aligned} P_{0,s}^V(A_4) &= P_{0,s}^V(A_4, i_{\max} > n) + P_{0,s}^V(A_4, i_{\max} \le n) \\ &\le P_{0,s}^V(i_{\max} > n) + P_{0,s}^V(\exists 1 \le i \le n : F_i \text{ occurs}) \\ &\le 2c_{10}n^{-\delta} + ne^{-n^{\epsilon_1\epsilon_2}} \\ &\le n^{-(\delta/2+\delta\epsilon_2/4)}. \end{aligned}$$

Bound for A_5:
By (D.9), $P_{0,s}^V(|K_i| > 2n^{\epsilon_1}) \le c_3 e^{-c_4\lfloor n^{\epsilon_1}\rfloor}$, for each i. Using this along with (D.13) shows that, for all sufficiently large n,

$$\begin{aligned} P_{0,s}^V(A_5) &= P_{0,s}^V(A_5, i_{\max} > n) + P_{0,s}^V(A_5, i_{\max} \le n) \\ &\le P_{0,s}^V(i_{\max} > n) + P_{0,s}^V(\exists 1 \le i \le n+1 : |K_i| > 2n^{\epsilon_1}) \\ &\le 2c_{10}n^{-\delta} + (n+1) \cdot c_3 e^{-c_4\lfloor n^{\epsilon_1}\rfloor} \\ &\le n^{-(\delta/2+\delta\epsilon_2/4)}. \end{aligned}$$

□

References

[1] G. Amir, N. Berger, and T. Orenshtein. Zero-one law for directional transience of one dimensional excited random walks. *Annales de l'Institut Henri Poincaré - Probabilités Statistiques*, 52(1):47–57, 2016.

[2] A. Basdevant and A. Singh. On the speed of a cookie random walk. *Probab. Theory Related Fields*, 141:625–645, 2008.

[3] A. Basdevant and A. Singh. Rate of growth of a transient cookie random walk. *Elec. J. Probab.*, 13:811–851, 2008.

[4] I. Benjamini and D. B. Wilson. Excited random walk. *Elec. Comm. Probab.*, 8:86–92, 2003.

[5] B. V. Gnedenko and A. N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, 1968. Translated from the Russian by K. L. Chung.

[6] A. Gut. *Stopped Random Walks*. Springer, second edition, 2009.

[7] H. Kesten, M. V. Kozlov, and F. Spitzer. A limit law for random walk in a random environment. *Composito Math.*, 30:145–168, 1975.

[8] E. Kosygina and T. Mountford. Limit laws of transient excited random walks on integers. *Annales de l'Institut Henri Poincaré - Probabilités Statistiques*, 47(2):575–600, 2011.

[9] E. Kosygina and J. Peterson. Excited random walks with Markovian cookie stacks. *Annales de l'Institut Henri Poincaré - Probabilités Statistiques*, 53(3):1458–1497, 2017.

[10] E. Kosygina and M. Zerner. Positively and negatively excited random walks on integers, with branching processes. *Elec. J. Probab.*, 13:1952–1979, 2008.

[11] E. Kosygina and M. Zerner. Excited random walks: results, methods, open problems. *Bull. Inst. Math. Acad. Sin. (N.S.)*, 8(1):105–157, 2013.

[12] E. Kosygina and M. Zerner. Excursions of excited random walks on integers. *Elec. J. Probab.*, 19(25):1–25, 2014.

[13] G. Kozma, T. Orenshtein, and I. Shinkar. Excited random walk with periodic cookies. *Annales de l'Institut Henri Poincaré - Probabilités Statistiques*, 52(3):1023–1049, 2016.

[14] D. Levin, Y. Peres, and E. Wilmer. *Markov Chains and Mixing Times*. Providence, R.I. American Mathematical Society, 2009.

[15] T. Mountford, L. P. R. Pimentel, and G. Valle. On the speed of the one-dimensional excited random walk in the transient regime. *Alea*, 2:279–296, 2006.

[16] J. Peterson. Large deviations and slowdown asymptotics for one-dimensional excited random walks. *Elec. J. Probab.*, 17(48):1–24, 2012.

[17] V. V. Petrov. *Sums of Independent Random Variables*. Spring-Verlag, 1975. Translated from the Russian by A. A. Brown.

[18] R. G. Pinsky. Transience/recurrence and the speed of a one-dimensional random walk in a "have your cookie and eat it" environment. *Annales de l'Institut Henri Poincaré - Probabilités Statistiques*, 46(4):949–964, 2010.

[19] S. Ross. *Stochastic Processes*. Wiley, second edition, 1996.

[20] B. Tóth. The "true" self-avoiding walk with bond repulsion on **Z**: limit theorems. *Ann. Probab.*, 23(4):1523–1556, 1995.

[21] B. Tóth. Generalized Ray-Knight theory and limit theorems for self-interacting random walks on $\mathbf{Z}^1$. *Ann. Probab.*, 24(3):1324–1367, 1996.

[22] M. Zerner. Multi-excited random walks on integers. *Probab. Theory Related Fields*, 133:98–122, 2005.

Acknowledgments. I thank Jonathon Peterson for helpful discussion on excited random walks, and in particular for making me aware of the diffusion approximation method for analyzing the backward branching process from previous works. I also thank an anonymous referee for comments that led to significant improvements in the paper.