ALEA, Lat. Am. J. Probab. Math. Stat. 12 (1), 427–450 (2015)

Excursions and occupation times of critical excited random walks

Dmitry Dolgopyat and Elena Kosygina

Department of Mathematics
University of Maryland,
4417 Mathematics Building
College Park, MD 20742, USA
E-mail address: dmitry@math.umd.edu
URL: http://www.math.umd.edu/~dolgop/

Department of Mathematics
Baruch College, Box B6-230
One Bernard Baruch Way
New York, NY 10010, USA
E-mail address: elena.kosygina@baruch.cuny.edu
URL: http://www.baruch.cuny.edu/math/facultystaff/ekosygina/kosygina.html

Abstract. We consider excited random walks (ERWs) on integers in i.i.d. environments with a bounded number of excitations per site. The emphasis is primarily on the critical case for the transition between recurrence and transience which occurs when the total expected drift δ at each site of the environment is equal to 1 in absolute value. Several crucial estimates for ERWs fail in the critical case and require a separate treatment. The main results discuss the depth and duration of excursions from the origin for $|\delta| = 1$ as well as occupation times of negative and positive semi-axes and scaling limits of ERW indexed by these occupation times. We also point out that the limiting proportions of the time spent by a non-critical recurrent ERW (i.e. when $|\delta| < 1$) above or below zero converge to beta random variables with explicit parameters given in terms of δ. The last observation can be interpreted as an ERW analog of the arcsine law for the simple symmetric random walk.

1. Introduction and main results

1.1. *Model description.* We consider an exited random walk (ERW) on $\mathbb{Z}$ with nearest neighbor jumps which evolves in a random "cookie environment". Each site of the lattice contains a stack of "cookies" $\omega_x := (\omega_x(1), \omega_x(2), \dots.)$. A cookie

Received by the editors October 26, 2014; accepted April 17, 2015.
2010 *Mathematics Subject Classification.* Primary: 60K37, 60F17, 60J80. Secondary: 60J60.
Key words and phrases. Random cookie environment, branching process, diffusion approximation, perturbed Brownian motion, beta distribution.

$\omega_x(i) \in [0,1]$, $x \in \mathbb{Z}$, $i \in \mathbb{N}$, encodes the probability that the walk jumps to the right upon the i-th visit to x. We assume that the cookie stacks ω_x, $x \in \mathbb{Z}$, are spatially i.i.d. and that there is a non-random $M \geq 0$, the number of excitations per site, such that $\omega_x(i) = 1/2$ for all $i > M$ and $x \in \mathbb{Z}$, i.e. starting from the $(M+1)$-th visit to a site the walk makes only unbiased jumps from this site.

More formally, we suppose that an environment $\omega \in \Omega = [0,1]^{\mathbb{Z}\times\mathbb{N}}$ is chosen according to a probability measure $\mathbb{P}$ which satisfies the following three assumptions.

(IID) (Independence) The cookie stacks $\omega_x(\cdot)$, $x \in \mathbb{Z}$, are i.i.d. under $\mathbb{P}$.

(WEL) (Weak ellipticity) For all $x \in \mathbb{Z}$

$$\mathbb{P}(\omega_x(i) > 0 \ \forall i \in \mathbb{N}) > 0 \quad \text{and} \quad \mathbb{P}(\omega_x(i) < 1 \ \forall i \in \mathbb{N}) > 0.$$

($\mathrm{BD_M}$) (At most M excitations per site) $\mathbb{P}(\omega_x(i) = 1/2 \ \forall x \in \mathbb{Z},\ i > M) = 1$.

Given an environment $\omega \in \Omega$, we shall use the usual coin-toss construction of a random walk, albeit we should keep a record of the number of visits of the walk to each site and use appropriately biased coins for the first M visits to each site. Namely, let $(\eta_x(i))_{x\in\mathbb{Z},i\in\mathbb{N}}$ be independent (under some probability measure P_ω) Bernoulli random variables such that $P_\omega(\eta_x(i) = 1) = 1 - P_\omega(\eta_x(i) = 0) = \omega_x(i)$ for all $x \in \mathbb{Z}$, $i \in \mathbb{N}$. Set $X_0 = x$, $x \in \mathbb{Z}$, and define recursively

$$X_{n+1} = X_n + 2\eta_{X_n}(\#\{k \in \{0,1,\dots,n\} : X_k = X_n\}) - 1, \quad n \in \{0\} \cup \mathbb{N}.$$

The probability measure $P_{\omega,x}$ induced on the space of random walk paths which start from x is called the *quenched* measure. The probability measure on the product space of environments and random walk paths originating at x defined by

$$P_x(\cdot) = \mathbb{E}\left[P_{\omega,x}(\cdot)\right] = \int_\Omega P_{\omega,x}(\cdot)\, d\mathbb{P}(\omega)$$

is called the *averaged* measure. Observe that ERW is not a Markov process with respect to either of these measures.

Below we shall only quote the facts needed to put our results into the context of previous work. For an overview of various ERW models, methods, and results the reader is referred to Kosygina and Zerner (2013).

1.2. *Excursions from the origin.* Let $T_k := \inf\{n \geq 0 : X_n = k\}$, $k \in \mathbb{Z}$, be the time of the first visit to k and $T_0^r := \inf\{n \geq 1 : \ X_n = 0\}$ be the first strictly positive time at which the random walk visits the origin.

Under our assumptions, several phase transitions are known to be characterized by the expected total drift stored in a single cookie stack

$$\delta := \mathbb{E}\left[\sum_{i=1}^{M}(2\omega_0(i) - 1)\right]. \tag{1.1}$$

The excited random walk $(X_n)_{n\geq 0}$

(i) is transient, i.e. $|X_n| \to \infty$ P_0-a.s., iff $|\delta| > 1$ (see Kosygina and Zerner (2013, Theorem 3.10) and the references therein or a combination of Kosygina and Zerner (2014, Corollary 7.10) and Remark A.6 below)[1];

(ii) is ballistic, i.e. there is a constant $v \neq 0$ such that P_0-a.s. $\lim_{n\to\infty} X_n/n = v$, iff $|\delta| > 2$ (see Kosygina and Zerner (2013, Theorem 5.2) and the references therein);

[1] for $|\delta| \leq 1$ X is recurrent, i.e. returns to the origin infinitely often P_0-a.s..

(iii) is strongly transient, i.e. $E^0\left[T_0^r \,|\, T_0^r < \infty\right] < \infty$, iff $|\delta| > 3$ (see Kosygina and Zerner (2014, Corollary 1.2));

(iv) after diffusive scaling converges under P_0 to a Brownian motion iff $|\delta| > 4$ or $\delta = 0$ (see Kosygina and Zerner (2013, Theorems 6.1, 6.3, 6.5, 6.7) and the references therein).

Remark 1.1. The velocity v in (ii) as well as all constants b, c, c_i, $i \geq 1$, which appear below depend on the distribution of a single cookie stack ω_0 under $\mathbb{P}$. They are not, in general, functions of δ (see Kosygina and Zerner (2013, Remark 5.8) for a discussion about v).

The phase transition in (iii) emerged in the study of the depth and duration of excursions of ERW. Since our first result is about excursions in the critical case $|\delta| = 1$ we shall first quote the original relevant theorem.

Theorem 1.2 (Kosygina and Zerner (2014), Theorem 1.1)**.** *Let $\delta \in \mathbb{R} \setminus \{1\}$. Then there are constants $c_1, c_2 \in (0, \infty)$ such that*

$$\lim_{n\to\infty} n^{|\delta-1|} P_1(T_n < T_0 < \infty) = c_1, \tag{1.2}$$

$$\lim_{n\to\infty} n^{|\delta-1|/2} P_1(n < T_0 < \infty) = c_2. \tag{1.3}$$

Moreover, if $\delta = 1$ then every $\varepsilon > 0$,

$$\lim_{n\to\infty} n^{\varepsilon} P_1(T_n < T_0) = \lim_{n\to\infty} n^{\varepsilon} P_1(T_0 > n) = \infty. \tag{1.4}$$

If $|\delta| \neq 1$[2] *then there is a constant $c_3 \in (0, \infty)$ such that*

$$\lim_{n\to\infty} n^{||\delta|-1|/2} P_0(n < T_0^r < \infty) = c_3. \tag{1.5}$$

Moreover, if $|\delta| = 1$ then for every $\varepsilon > 0$,

$$\lim_{n\to\infty} n^{\varepsilon} P_0(T_0^r > n) = \infty. \tag{1.6}$$

This theorem immediately implies (iii) but provides very little information about the tail of the return time in the critical case. Our first result fills in this gap.

Theorem 1.3. *If $\delta = 1$ then there is a constant $c_4 \in (0, \infty)$ such that*

$$\lim_{n\to\infty} (\ln n)\, P_1(T_n < T_0) = c_4; \tag{1.7}$$

$$\lim_{n\to\infty} (\ln n)\, P_1(T_0 > n) = 2c_4. \tag{1.8}$$

Moreover, if $|\delta| = 1$ then

$$\lim_{n\to\infty} (\ln n)\, P_0(T_0^r > n) = c_5 := \begin{cases} 2c_4\mathbb{E}[\omega_0(1)], & \text{if } \delta = 1; \\ 2c_4\mathbb{E}[1-\omega_0(1)], & \text{if } \delta = -1. \end{cases} \tag{1.9}$$

The key statements of Theorem 1.3 are (1.7) and (1.8). The last conclusion follows easily from (1.8), (1.3) with $\delta = -1$, and the following remark by conditioning on the first step (see Kosygina and Zerner (2014, (6.2))).

[2] In (1.5) of Kosygina and Zerner (2014) both $\delta = 1$ and $\delta = -1$ should have been excluded.

Remark 1.4. There is a useful symmetry in our model. If the environment $(\omega_x)_{x\in\mathbb{Z}}$ is replaced with $(\widetilde{\omega}_x)_{x\in\mathbb{Z}}$ where $\widetilde{\omega}_x(i) = 1-\omega_x(i)$, for all $i\in\mathbb{N},\ x\in\mathbb{Z}$, then $\widetilde{X}$, the ERW corresponding to the new environment, satisfies

$$\widetilde{X} \overset{\mathrm{d}}{=} -X, \tag{1.10}$$

where $\overset{\mathrm{d}}{=}$ denotes the equality in distribution. Thus, it is sufficient to consider only excursions to the right (for all δ). The corresponding results for excursions to the left will follow by symmetry. Thus from now on we shall assume without loss of generality that $\delta \geq 0$.

1.3. *Occupation times and scaling limits.* Unless stated otherwise we shall assume that all processes start at the origin at time 0. Let $B = (B(t)),\ t\geq 0$, denote a standard Brownian motion and $W_{\alpha,\beta} = (W_{\alpha,\beta}(t)),\ t\geq 0$, be an (α,β)-perturbed Brownian motion, i.e. the solution of the equation

$$W_{\alpha,\beta}(t) = B(t) + \alpha \sup_{s\leq t} W_{\alpha,\beta}(s) + \beta \inf_{s\leq t} W_{\alpha,\beta}(s). \tag{1.11}$$

Reflected α-perturbed Brownian motion, $W_\alpha = (W_\alpha(t)),\ t\geq 0$, is the solution of

$$W_\alpha(t) = B(t) + \alpha \sup_{s\leq t} W_\alpha(s) + \frac{1}{2} L^{W_\alpha}(t), \tag{1.12}$$

where $L^{W_\alpha}(t)$ is the local time of W_α at zero. Equation (1.11) has a path-wise unique solution if $(\alpha,\beta)\in(-\infty,1)\times(-\infty,1)$, and (1.12) has a path-wise unique solution when $\alpha<1$ (Davis (1996); Perman and Werner (1997); Chaumont and Doney (1999)). In both cases the solution is adapted to the filtration of B. If $\beta=0$ then the solution of (1.11) can be written explicitly:

$$W_{\alpha,0}(t) = B(t) + \frac{\alpha}{1-\alpha} \sup_{s\leq t} B(s). \tag{1.13}$$

Throughout the paper we use $\Rightarrow$ to denote the weak convergence of random variables and $\overset{J_1}{\Rightarrow}$ for the weak convergence of stochastic processes with respect to the standard Skorokhod topology J_1 on $D([0,\infty))$, the space of càdlàg functions on $[0,\infty)$.[3]

The following two theorems describe scaling limits of recurrent ERWs.

Theorem 1.5 (Dolgopyat and Kosygina (2012), Theorem 1.1)**.** *Let* $\delta\in[0,1)$. *Then under* P_0

$$\frac{X_{[n\cdot]}}{\sqrt{n}} \overset{J_1}{\Rightarrow} W_{\delta,-\delta}(\cdot) \ \text{as } n\to\infty.$$

Theorem 1.6 (Dolgopyat and Kosygina (2012), Theorem 1.2)**.** *Let* $\delta=1$ *and* $B^*(t) := \max_{s\leq t} B(s)$. *Then there exists a constant* $b\in(0,\infty)$ *such that under* P_0

$$\frac{X_{[n\cdot]}}{b\sqrt{n}\log n} \overset{J_1}{\Rightarrow} B^*(\cdot) \ \text{as } n\to\infty.$$

---

[3]Since all limiting processes below have continuous paths, we can also claim the convergence with respect to the uniform topology on $D([0,T])$ for each $T>0$ (see Billingsley (1999, Section 15)).

At the first glance it appears counter-intuitive that for $\delta = 1$ the limiting process is transient while the original process is recurrent. However, the running maximum of Brownian motion is a natural limit of $W_{\alpha,\beta}((1-\alpha)^2\,\cdot)$ as $\alpha \uparrow 1$ (see the discussion right after Theorem 1.7).

Theorem 1.5 suggests that the rescaled occupation times of positive and negative semi-axes of non-critical recurrent ERW should converge to those of the reflected perturbed Brownian motion. The latter was studied in detail, and the next theorem quotes results from the literature. Let

$$A^+(t) := \int_0^t \mathbb{1}_{\{W_{\alpha,\beta}(u)\ge 0\}}\,du, \quad A^-(t) := \int_0^t \mathbb{1}_{\{W_{\alpha,\beta}(u)< 0\}}\,du, \ t \ge 0,$$

and $T_\pm(t) := \inf\{s:\ A^\pm(s) > t\}$, $t \ge 0$, be the right continuous inverses of $A^\pm(\cdot)$. Denote by $Z(a,b)$ a Beta-distributed random variable with parameters a and b.

Theorem 1.7. *For all $\alpha, \beta < 1$ the following holds:*

(a) *Carmona et al. (1998, equation (8))*

$$\frac{A^+(t)}{t} \overset{\mathrm{d}}{=} Z\left(\frac{1-\beta}{2}, \frac{1-\alpha}{2}\right) \quad \textit{and} \quad \frac{A^-(t)}{t} \overset{\mathrm{d}}{=} Z\left(\frac{1-\alpha}{2}, \frac{1-\beta}{2}\right).$$

(b) *Chaumont and Doney (2000, Theorem 1)*

$$W_{\alpha,\beta}(T^+(\cdot)) \overset{\mathrm{d}}{=} W_\alpha(\cdot) \quad \textit{and} \quad -W_{\alpha,\beta}(T^-(\cdot)) \overset{\mathrm{d}}{=} W_\beta(\cdot).$$

Theorem 1.7 implies that $W_{\alpha,\beta}((1-\alpha)^2\,\cdot) \Rightarrow B^*(\cdot)$ as $\alpha \uparrow 1$. Indeed, the Brownian scaling of $W_{\alpha,\beta}$ (Carmona et al. (1998, Proposition 2.3)) allows to rewrite the above convergence as

$$(1-\alpha)W_{\alpha,\beta}(\cdot) \Rightarrow B^*(\cdot) \text{ as } \alpha \uparrow 1. \tag{1.14}$$

By Theorem 1.7(a) $\alpha \uparrow 1$ the process $W_{\alpha,\beta}$ stays most of the time in $[0,\infty)$ (recall that $\mathbb{E}\,[A^+(t)/t] = (1-\beta)/(2-\alpha-\beta)$). By Theorem 1.7(b) we conclude that the limit in (1.14) should be independent of β. On the other hand, if $\beta = 0$ then (1.13) tells us that $(1-\alpha)W_{\alpha,0}(\cdot)$ has the same law as $(1-\alpha)B(\cdot) + \alpha \sup_{s\le \cdot} B(s)$. This implies (1.14).

The next corollary follows from Theorems 1.5 and 1.7 by the continuous mapping theorem (see Section 4 for details).

Corollary 1.8. *Suppose that $\delta \in [0,1)$. Let*

$$A_n^+ := \sum_{i=0}^n \mathbb{1}_{\{X_i \ge 0\}} \quad \textit{and} \quad A_n^- := \sum_{i=0}^n \mathbb{1}_{\{X_i < 0\}}, \ n \ge 0,$$

and $T_m^\pm := \inf\{n \ge 0:\ A_n^\pm > m\}$, $m \ge 0$. Then

(a) $$\frac{A_n^+}{n} \Rightarrow Z\left(\frac{1+\delta}{2}, \frac{1-\delta}{2}\right) \quad \textit{and} \quad \frac{A_n^-}{n} \Rightarrow Z\left(\frac{1-\delta}{2}, \frac{1+\delta}{2}\right) \quad \textit{as } n \to \infty;$$

(b) $$\frac{X_{T^+_{\lfloor m\cdot\rfloor}}}{\sqrt{m}} \overset{J_1}{\Rightarrow} W_\delta(\cdot) \qquad \textit{and} \qquad -\frac{X_{T^-_{\lfloor m\cdot\rfloor}}}{\sqrt{m}} \overset{J_1}{\Rightarrow} W_{-\delta}(\cdot) \qquad \textit{as } m \to \infty.$$

Consider now the critical case $\delta = 1$. It is clear from Theorem 1.6 that the proportion of time spent in $(-\infty, 0)$ by an ERW with $\delta = 1$ should converge to 0 (see Lemma 4.1 below). Since the critical ERW is recurrent and satisfies ($\mathrm{BD_M}$), $A_n^- \to \infty$ as $n \to \infty$. But how fast does A_n^- increase? Our last theorem answers this question on a logarithmic scale and also provides scaling limits of $X_{T^\pm_{\lfloor m\cdot\rfloor}}$ when $\delta = 1$.

Theorem 1.9. *Let $\delta = 1$ and $A_n^\pm$, $T_n^\pm$, $n \geq 0$, be as in Corollary 1.8. Then under P_0*

(a) $\dfrac{\log A_n^-}{\log n} \Rightarrow U$ *as $n \to \infty$, where U is uniform on $[0,1]$ random variable;*

(b) $-\dfrac{X_{T^-_{\lfloor m\cdot\rfloor}}}{\sqrt{m}} \overset{J_1}{\Rightarrow} W_{-1}(\cdot)$ *as $m \to \infty$;*

(c) *there is a constant $b \in (0,\infty)$ such that* $\dfrac{X_{T^+_{\lfloor m\cdot\rfloor}}}{b\sqrt{m}\log m} \overset{J_1}{\Rightarrow} B^*(\cdot)$ *as $m \to \infty$.*

Part (a) of the above theorem informally says that $A_n^- \asymp n^U$ where U is a standard uniform random variable. See Section 4.2 for a heuristic derivation of this asymptotics. Part (b) is just a simple extension of the last claim of Corollary 1.8 to $\delta = 1$. This reflects the fact that if we consider an ERW with $\delta = 1$ only at the times when it visits the negative half-line then such process is not critical and can be treated essentially in the same way as the case $\delta \in [0,1)$. The situation is different if we look at an ERW with $\delta = 1$ only when it visits the positive half-line, since neither $W_{\alpha,\beta}$ nor W_α exists for $\alpha = 1$. But in view of Theorem 1.6 the statement of part (c) is not surprising.

1.4. *Organization of the paper.* In Section 2 we explain the connection between ERWs and some branching processes. The main theorem of Section 2, Theorem 2.1, is an important tool for the proofs of our main results. We illustrate this by deriving Theorem 1.3 as a simple corollary of Theorem 2.1. The proof of Theorem 2.1 is given in Section 3. In Section 4 we prove Corollary 1.8 and Theorem 1.9. Proofs of some technical results are collected in Appendices A and B.

2. Connection with branching processes

In this section we construct the relevant branching process (BP) and restate (1.7) and (1.8) in terms of the tails of the extinction time and the total progeny of these BPs.

We shall use the same environment $\omega \in \Omega$ and Bernoulli random variables $(\eta_x(i))_{x\in\mathbb{Z},i\in\mathbb{N}}$ as in the construction of the ERW. This will provide us with a natural coupling between the ERW and the BP. We define here only the BP V which corresponds to right excursions of the walk.[4] For $x \in \{0\} \cup \mathbb{N}$ let

$$S_x(0) = 0, \quad S_x(m) := \inf\left\{k \geq 1 : \sum_{i=1}^k (1 - \eta_x(i)) = m\right\} - m, \quad m \in \mathbb{N}.$$

Thus, $S_x(m)$ is the number of "successes" before the m-th "failure" in the sequence $\eta_x(i)$, $i \in \mathbb{N}$. Define the process $V = (V_n)_{n\geq 0}$ which starts with y particles in generation 0 by

$$V_0 = y, \quad V_n = S_n(V_{n-1}), \quad n \in \mathbb{N}. \tag{2.1}$$

[4]The BP corresponding to left excursions, V^-, is constructed in a symmetric way and will be introduced in Section 4.2.

If there were no biased coins, V would be a Galton-Watson process with mean 1 geometric offspring distribution. Our process uses up to M possibly biased coins in each generation, therefore, strictly speaking, it is not a "true" branching process. We could recast it as a branching process with migration (see Kosygina and Zerner (2008, Section 3)) but, since we do not use any results from branching processes literature, we shall not need this step.

For $y \in [0,\infty)$ we shall denote by P_y^V the (averaged) probability measure corresponding to the process V which starts with $\lfloor y \rfloor$ particles in generation 0. For $x \in [0,\infty)$ define $\tau_x^V := \inf\{n \in \mathbb{N} : V_n \geq x\}$ and $\sigma_x^V := \inf\{n \in \mathbb{N} : V_n \leq x\}$. When there is no danger of confusion we shall drop the superscript V.

Theorem 2.1. *Let $\delta = 1$ and $V = (V_n)_{n\geq 0}$ be defined by (2.1). Then for each $y \in \mathbb{N}$ there is a constant $c_6(y) \in (0,\infty)$ such that*

$$\lim_{n\to\infty} (\ln n)\, P_y^V(\sigma_0^V > n) = c_6(y), \tag{2.2}$$

$$\lim_{n\to\infty} (\ln n)\, P_y^V\left(\sum_{i=0}^{\sigma_0^V - 1} V_i > n\right) = 2c_6(y). \tag{2.3}$$

Assume for the moment Theorem 2.1 and derive Theorem 1.3.

Proof of Theorem 1.3: The proof is essentially the same as that of Theorem 1.1 in Kosygina and Zerner (2014). Let the ERW start with $x = 1$ and the corresponding BP start with $y = 1$. Observe that, since ERW and BP are constructed from the same $(\eta_x(i))_{x\in\mathbb{Z}, i\in\mathbb{N}}$, we have

$$\sigma_0^V = \max\{X_n : n < T_0\} \quad \text{and} \quad T_0 \mathbb{1}_{\{T_0<\infty\}} = \left(2\sum_{n=0}^{\sigma_0^V-1} V_n - 1\right)\mathbb{1}_{\{\sigma_0^V<\infty\}}.$$

Therefore, (1.7) and (1.8) with $c_4 = c_6(1)$ follow from (2.2) and (2.3). To show (1.9) we start ERW with $x = 0$ and condition on the first step. Since $P_{\omega,\pm 1}(T_0 \geq n)$ do not depend on $\omega_0(\cdot)$,

$$\begin{aligned} P_0(T_0^r > n) &= \mathbb{E}[P_{\omega,0}(T_0^r > n)] \\ &= \mathbb{E}[\omega_0(1)P_{\omega,1}(T_0 \geq n)] + \mathbb{E}[(1-\omega_0(1))P_{\omega,-1}(T_0 \geq n)] \\ &= \mathbb{E}[\omega_0(1)]P_1(T_0 \geq n) + \mathbb{E}[(1-\omega_0(1))]P_{-1}(T_0 \geq n). \end{aligned}$$

By (WEL), $\mathbb{E}[\omega_0(1)] > 0$ and $\mathbb{E}[(1-\omega_0(1))] > 0$. If $\delta = 1$ then $(\ln n)P_1(T_0 \geq n) \to 2c_4$ as $n \to \infty$ by (1.8). By Remark 1.4 and (1.3) with $\delta = -1$, $nP_{-1}(T_0 \geq n)$ converges to a constant. We conclude that

$$\lim_{n\to\infty} (\ln n) P_0(T_0^r > n) = 2c_4\mathbb{E}[\omega_0(1)]. \tag{2.4}$$

The result for $\delta = -1$ follows by symmetry. □

3. Proof of Theorem 2.1

The proof of Theorem 2.1 depends on a number of additional facts which we state below and prove in Appendix A.

Lemma 3.1. *Let $\delta = 1$ and $y \in \mathbb{N}$. Then there is a constant $c_6(y) \in (0,\infty)$ such that*

$$\lim_{n\to\infty} (\ln n)\, P_y(\tau_n < \sigma_0) = c_6(y).$$

Lemma 3.2. *Let $\delta = 1$. For every $y \in \mathbb{N}$ and $\alpha > 1$*

$$\lim_{n\to\infty} (\ln n) P_y \left(\sum_{i=0}^{\sigma_0 - 1} \mathbb{1}_{\{V_i \le n\}} > n^\alpha \right) = 0.$$

Lemma 3.3. *Let $\delta = 1$. For every $h > 0$*

$$\lim_{n\to\infty} P_n(\sigma_0 > hn) = 1; \tag{3.1}$$

$$\lim_{n\to\infty} P_n \left(\sum_{i=0}^{\sigma_0 - 1} V_i > hn^2 \right) = 1. \tag{3.2}$$

The following results, which will be referred to as (DA), Diffusion Approximation, and (OS), "Overshoot", respectively, are borrowed from previous works.

Lemma 3.4 (Diffusion approximation)**.** *Let $\delta = 1$. Fix an arbitrary $\varepsilon > 0$ and $y > \varepsilon$. Let $Y^{\varepsilon,n}(0) = [ny]$ and $Y^{\varepsilon,n}(t) = \dfrac{V_{[nt]\wedge\sigma_{\varepsilon n}}}{n}$, $t \ge 0$. Then, under the averaged measure, $Y^{\varepsilon,n} \overset{J_1}{\Rightarrow} Y$, where Y is the solution of*

$$dY(t) = dt + \sqrt{2Y(t)}\, dB(t), \quad Y(0) = y, \tag{3.3}$$

stopped when Y reaches level ε.

Lemma 3.4 is an immediate consequence of Proposition 3.2 and Lemma 3.3 in Kosygina and Zerner (2014).

Lemma 3.5 ("Overshoot", Lemma 5.1 in Kosygina and Mountford (2011))**.** *There are constants $c_7, c_8 > 0$ and $N \in \mathbb{N}$ such that for all $x \ge N$ and $y \ge 0$*

$$\max_{0\le z<x} P_z(V_{\tau_x} > x + y \,|\, \tau_x < \sigma_0) \le c_7 \left(e^{-c_8 y^2/x} + e^{-c_8 y} \right)$$

and

$$\max_{x<z<4x} P_z(V_{\sigma_x \wedge \tau_{4x}} < x - y) \le c_7 e^{-c_8 y^2/x}.$$

Proof of Theorem 2.1: We start with the proof of (2.2).

Lower bound for (2.2). For every $y \in \mathbb{N}$ we have by the strong Markov property and monotonicity in the starting point that

$$\begin{aligned} P_y(\sigma_0 > n) &\ge P_y(\sigma_0 > n, \tau_n < \sigma_0) \\ &= P_y(\sigma_0 > n \,|\, \tau_n < \sigma_0) P_y(\tau_n < \sigma_0) \ge P_n(\sigma_0 > n) P_y(\tau_n < \sigma_0). \end{aligned}$$

Using Lemma 3.1 and (3.1) we get

$$\liminf_{n\to\infty} (\ln n) P_y(\sigma_0 > n) \ge c_6(y).$$

Upper bound for (2.2). Fix an arbitrary $\alpha > 1$ and notice that for all $m > y$

$$\begin{aligned} (\ln m)\, P_y(\sigma_0 > m^\alpha) &\le (\ln m)\, P_y(\sigma_0 > m^\alpha, \tau_m \le m^\alpha) + (\ln m)\, P_y(\tau_m \wedge \sigma_0 > m^\alpha) \\ &\le (\ln m)\, P_y(\sigma_0 > \tau_m) + (\ln m)\, P_y \left(\sum_{i=1}^{\sigma_0 - 1} \mathbb{1}_{\{V_i \le m\}} > m^\alpha \right). \end{aligned}$$

As $m \to \infty$, the first term in the right hand side converges to $c_6(y)$ by Lemma 3.1 and the second term vanishes due to Lemma 3.2.

Define $m = m(\alpha, n)$ by the condition $m^\alpha \le n < (m+1)^\alpha$. Then we get

$$\limsup_{n\to\infty} (\ln n)\, P_y(\sigma_0 > n) \le \lim_{\alpha\downarrow 1} \lim_{m\to\infty} \alpha(\ln(m+1)) P_y(\sigma_0 > m^\alpha) = c_6(y),$$

which matches the lower bound.

We turn now to the proof of (2.3). It is enough to show that

$$\lim_{n\to\infty} (\ln n) P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > n^2\right] = c_6(y). \tag{3.4}$$

Lower bound for (3.4). By Lemma 3.1 and (3.2) ,

$$\begin{aligned}\liminf_{n\to\infty} (\ln n) P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > n^2\right] &\ge \liminf_{n\to\infty} (\ln n) P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > n^2, \tau_n < \sigma_0\right] \\ &\ge \lim_{n\to\infty} (\ln n) P_y[\tau_n < \sigma_0] \lim_{n\to\infty} P_n\left[\sum_{i=0}^{\sigma_0-1} V_i > n^2\right] = c_6(y).\end{aligned}$$

Upper bound for (3.4). The reasoning is very similar to the one we gave for (2.2). Fix $\alpha > 1$. Using the sequence $m = m(\alpha, n)$ such that $m^\alpha \le n < (m+1)^\alpha$ we get

$$\limsup_{n\to\infty} (\ln n) P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > n^2\right] \le \alpha \limsup_{m\to\infty} \ln(m+1) P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > m^{2\alpha}\right].$$

Therefore, if we show that for every $\alpha > 1$

$$\limsup_{m\to\infty} (\ln m) P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > m^{2\alpha}\right] \le c_6(y), \tag{3.5}$$

then letting $\alpha \to 1$ and using the lower bound we shall obtain (3.4). Notice that

$$(\ln m)\, P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > m^{2\alpha}, \tau_m < \sigma_0\right] \le (\ln m)\, P_y(\tau_m < \sigma_0),$$

which by Lemma 3.1 converges to $c_6(y)$ as $m \to \infty$. Finally,

$$\begin{aligned}P_y\left[\sum_{i=0}^{\sigma_0-1} V_i > m^{2\alpha}, \tau_m > \sigma_0\right] &\le P_y(\sigma_0 > m^{2\alpha-1}, \tau_m > \sigma_0) \\ &\le P_y\left[\sum_{i=0}^{\sigma_0-1} \mathbb{1}_{\{V_i \le m\}} > m^{2\alpha-1}\right].\end{aligned}$$

By Lemma 3.2 the last expression is $o(1/\ln m)$ as $m \to \infty$, and we get (3.5). □

4. Proofs of Corollary 1.8 and Theorem 1.9

4.1. *Proof of Corollary 1.8.* Part (a) of Corollary 1.8 follows from the following lemma. Observe that this lemma also covers the case $\delta = 1$. This will be needed later in the section.

Lemma 4.1. *Let $\delta \in [0,1]$. Then as $n \to \infty$*

$$\frac{A_n^+}{n} \Rightarrow Z\left(\frac{1+\delta}{2}, \frac{1-\delta}{2}\right),$$

where we set $Z(1,0) \equiv 1$.

Proof: This lemma is an easy consequence of Theorems 1.5 and 1.7 (for $\delta \in [0,1)$), Theorem 1.6 (for $\delta = 1$), and the continuous mapping theorem. To unify the notation let

$$X_{\delta,n}(\cdot) := \begin{cases} \dfrac{X_{[n\cdot]}}{\sqrt{n}}, & \text{if } \delta \in [0,1); \\ \dfrac{X_{[n\cdot]}}{b\sqrt{n}\log n}, & \text{if } \delta = 1; \end{cases} \qquad W_{1,-1} := B^*.$$

Define $\varphi : D([0,1]) \to \mathbb{R}$ by

$$\varphi(\omega) = \int_0^1 \mathbb{1}_{[0,\infty)}(\omega(t))\, dt. \tag{4.1}$$

Note that the Lebesgue measure of the set $\mathcal{Z} := \{t \in [0,1] : \ W_{\delta,-\delta}(t) = 0\}$ is 0 P-a.s.. Indeed,

$$E\int_0^1 \mathbb{1}_{\mathcal{Z}}(t)\, dt = \int_0^1 P(W_{\delta,-\delta}(t) = 0)\, dt = 0.$$

where the last equality follows from the fact that $W_{-\delta,\delta}$ has a density (see Carmona et al. (1998, Proposition 2.3 and Section 3.3)). Then, if P is the measure corresponding to $W_{\delta,-\delta}$ then by Proposition B.1 the map φ is continuous P-a.s. (as P is supported on continuous functions) and

$$\varphi(X_{\delta,n}) = \frac{1}{n}\sum_{k=0}^{n} \mathbb{1}_{[0,\infty)}(X_k) = \frac{A_n^+}{n} \Rightarrow \varphi(W_{\delta,-\delta}) \overset{\mathrm{d}}{=} Z\left(\frac{1+\delta}{2}, \frac{1-\delta}{2}\right).$$

The last equality follows from Theorem 1.7(a) for $\delta \in [0,1)$ and is trivial for $\delta = 1$. □

It is enough to show the second part of Corollary 1.8(b). The proof of the first part is similar. For every $R > 0$ consider the map $\psi : D([0,\infty)) \to D([0,R])$ defined by

$$\psi(\omega(s),\, 0 \le s < \infty) = (-\omega(T^-(s)),\, 0 \le T^-(s) \le R), \tag{4.2}$$

where $T^-(s) := \inf\left\{t \ge 0 : \int_0^t \mathbb{1}_{(-\infty,0)}(\omega(r))\, dr > s\right\}$. By Proposition B.2, ψ is continuous P-a.s. (P is the measure which corresponds to $W_{\delta,-\delta}$). The desired statement now follows from Theorem 1.5 and Theorem 1.7(b) by the continuous mapping theorem.

4.2. *Heuristics and the proof of Theorem 1.9(a).* We start by introducing some additional notation which will be used throughout the rest of Section 4. Denote by d_n the number of down-crossings of X from 0 to -1 up to time n inclusively and by u_n the number of up-crossings of X from 0 to 1 up to time n inclusively. Rename V to V^+ (for this section only) and introduce the process V^- which corresponds to left excursions of the walk. Namely, for $x \le 0$ let

$$F_x(0) = 0, \quad F_x(m) := \inf\left\{k \ge 1 : \ \sum_{i=1}^{k} \eta_x(i) = m\right\} - m, \quad m \in \mathbb{N}.$$

Thus, $F_x(m)$ is the number of "failures" before the m-th "success" in the sequence $\eta_x(i)$, $i \in \mathbb{N}$. Define the process $V^- = (V_n^-)_{n\ge 0}$ which starts with y particles in generation 0 by

$$V_0^- = y, \quad V_n^- = F_{-n}(V_{n-1}^-), \quad n \in \mathbb{N}. \tag{4.3}$$

If $V_0^\pm = k$ then denote by $\Sigma_k^\pm := \sum_{j=0}^{\sigma_0-1} V_j^\pm$ the total progeny of the BP $V^\pm$ over its lifetime and observe that

$$2\Sigma^+_{u_n-1} \le A_n^+ \le 2\Sigma^+_{u_n} + d_n + 1 \quad \text{and} \quad 2\Sigma^-_{d_n-1} - d_n \le A_n^- \le 2\Sigma^-_{d_n}. \tag{4.4}$$

To see why the first set of the above inequalities holds, note that A_n^+ falls in between the total duration (including visits to 0) of the first u_n-1 and the first u_n excursions to the right. Since the number of up-crossings from one level to the next in each excursion is equal to the number of down-crossings, by coupling with the BP we obtain the estimates in terms of the total progeny of the BP which starts with u_n-1 and u_n particles respectively. Since A_n^+ includes the number of visits to zero, we have to add to the upper bound the number of visits to 0 after which the walker stepped to the left, i.e. d_n. An additional 1 in the upper bound for A_n^+ accounts for the possibility that $X_n = 0$, in which case we have to count the up- or down-crossing in the next step from that point. The second set of inequalities is obtained similarly. The only difference is that by our definition A_n^- does not include the time spent at 0.

Informal discussion. We shall explain where the uniform distribution in Theorem 1.9(a) comes from. Recall that Y is a diffusion process satisfying (3.3). It is a half of a squared Bessel process of dimension 2. Let $\tau_x = \inf\{t \ge 0 : Y(t) = x\}$, $x > 0$. The uniform distribution appears naturally in the following lemma.

Lemma 4.2. *Let $Y^*(t) = \max_{s\le t} Y(s)$ and $Y(0) = y > 1$. Then*

$$\frac{\ln y}{\ln Y^*(\tau_1)} \overset{\text{d}}{=} U.$$

Proof: It is easy to check that $\ln Y(t)$, $t \ge 0$, is a local martingale and so for all $R > y$

$$P_y(\tau_R < \tau_1) = \frac{\ln y}{\ln R}. \tag{4.5}$$

For $x \in (0, 1)$ we have

$$P_y\left(\frac{\ln y}{\ln Y^*(\tau_1)} \le x\right) = P_z(Y^*(\tau_1) \ge y^{1/x}) = P_y(\tau_{y^{1/x}} \le \tau_1) \overset{(4.5)}{=} x. \qquad \square$$

The next step is to observe that for a large starting point y the area under the path of Y up to τ_1 is roughly the square of $Y^*(\tau_1)$.

Lemma 4.3. *Let $Y(0) = y > 1$. Then*

$$\frac{\ln \int_0^{\tau_1} Y(s)\,ds}{\ln Y^*(\tau_1)} \Rightarrow 2 \quad \text{as } y \to \infty. \tag{4.6}$$

The proof of Lemma 4.3 is omitted as we use it only for this informal discussion. It can be proven in the same way as Lemma 4.8. The next statement immediately follows from Lemmas 4.2 and 4.3.

Corollary 4.4. *Let $Y(0) = y > 1$. Then*

$$\frac{2\ln y}{\ln \int_0^{\tau_1} Y(s)\,ds} \Rightarrow U \quad \text{as } y \to \infty. \tag{4.7}$$

The key part of the proof of Theorem 1.9(a) is the following analog of (4.7): let $V_0^+ = n$, then

$$\frac{2\ln n}{\ln \Sigma_n^+} \Rightarrow U \quad \text{as } n\to\infty. \tag{4.8}$$

Notice that (4.8) could not be obtained from (4.7) simply by the diffusion approximation, since we consider V^+ all the way down to the extinction time and Y does not hit zero with probability 1. In the next subsection we prove BP versions of Lemmas 4.2 and 4.3 (see Lemmas 4.7 and 4.8) and obtain (4.8).

Once we know (4.8), it is relatively simple to arrive at the conclusion of Theorem 1.9(a). We want to show that $\ln A_n^-/\ln n \Rightarrow U$. Consider the following chain of substitutions as $n\to\infty$:

$$\frac{\ln A_n^-}{\ln n} \overset{(4.4)}{\longleftrightarrow} \frac{\ln \Sigma_{d_n}^-}{\ln n} \overset{\text{L. }4.5}{\longleftrightarrow} \frac{2\ln d_n}{\ln n} \overset{(4.10)}{\longleftrightarrow} \frac{2\ln u_n}{\ln n} \overset{\text{L. }4.1}{\longleftrightarrow} \frac{2\ln u_n}{\ln A_n^+} \overset{(4.4)}{\longleftrightarrow} \frac{2\ln u_n}{\ln \Sigma_{u_n}^+} \overset{(4.10)}{\longleftrightarrow} \frac{2\ln n}{\ln \Sigma_n^+},$$

where the last ratio converges to U by (4.8). The actual proof combines the last three steps into a single argument. Below we state Lemmas 4.5 and 4.6 mentioned above, and use them together with (4.8) to derive Theorem 1.9(a). The proofs of Lemmas 4.5 and 4.6 are postponed until Section 4.4.

Lemma 4.5. *For every $\nu>0$, $x\in[0,1]$, and all sufficiently large n*

$$P_0\left(\frac{2\ln d_n}{\ln n}\le x-\nu\right)-\nu \le P_0\left(\frac{\ln\Sigma_{d_n}^-}{\ln n}\le x\right)\le P_0\left(\frac{2\ln d_n}{\ln n}\le x+\nu\right)+\nu.$$

Lemma 4.6. *The following statements hold with probability 1 as $n\to\infty$:*

$$u_n\to\infty;\quad d_n\to\infty; \tag{4.9}$$

$$\frac{d_n}{u_n}\to 1. \tag{4.10}$$

Proof of Theorem 1.9(a): By (4.4), Lemma 4.5, and Lemma 4.6 it is enough to show that $2(\ln u_n)/(\ln n)\Rightarrow U$ as $n\to\infty$.

Let $x\in(0,1)$. Fix an arbitrary $\nu>0$ and $\varepsilon\in(0,1/2)$. Then

$$\begin{aligned}
&P\left(\frac{2\ln u_n}{\ln n}\le x\right)\\
&\quad\le P\left(2\Sigma^+_{\lfloor n^{x/2}\rfloor}\ge(1-2\varepsilon)n\right)+P\left(\frac{2\ln u_n}{\ln n}\le x,\,2\Sigma^+_{\lfloor n^{x/2}\rfloor}<(1-2\varepsilon)n\right)\\
&\quad= P\left(2\Sigma^+_{\lfloor n^{x/2}\rfloor}\ge(1-2\varepsilon)n\right)+P\left(u_n\le\lfloor n^{x/2}\rfloor,\,2\Sigma^+_{\lfloor n^{x/2}\rfloor}<(1-2\varepsilon)n\right).
\end{aligned}$$

Note that by Lemma 4.1 with probability at least $1-\nu/2$ for all large n we have that

$$(1-\varepsilon)n+1\le A_n^+\overset{(4.4)}{\le} 2\Sigma^+_{u_n}+d_n+1. \tag{4.11}$$

Moreover, on the set $\{u_n\le\lfloor n^{x/2}\rfloor\}$ we have by coupling that

$$\Sigma^+_{u_n}\le\Sigma^+_{\lfloor n^{x/2}\rfloor}. \tag{4.12}$$

Inequalities (4.11) and (4.12) imply that

$$2\Sigma^+_{\lfloor n^{x/2}\rfloor} \ge (1-\varepsilon)n - d_n \ge n\left(1-\varepsilon-\frac{d_n}{n}\right).$$

By Lemma 4.1, $d_n/n \le A_n^-/n \to 0$ in probability. Therefore, for all large n

$$P\left(u_n \le \lfloor n^{x/2}\rfloor, 2\Sigma^+_{\lfloor n^{x/2}\rfloor} < (1-2\varepsilon)n\right) \le \nu.$$

We conclude that for all sufficiently large n

$$\begin{aligned} P\left(\frac{2\ln u_n}{\ln n} \le x\right) &\le P\left(2\Sigma^+_{\lfloor n^{x/2}\rfloor} \ge (1-2\varepsilon)n\right) + \nu \\ &\le P\left(\frac{2\ln\lfloor n^{x/2}\rfloor}{\ln \Sigma^+_{\lfloor n^{x/2}\rfloor}} \le x+\nu\right) + \nu \overset{(4.8)}{\le} x + 3\nu. \end{aligned}$$

Towards a lower bound, observe that by coupling $\{2\Sigma^+_z > n\} \subset \{u_n \le z\}$ for all $z \in \mathbb{N}$. Using this fact, Lemma 4.6, and (4.8) we get for all sufficiently large n that

$$\begin{aligned} P\left(\frac{2\ln u_n}{\ln n} \le x\right) &\ge P\left(u_n \le \lfloor n^{x/2}\rfloor,\ 2\Sigma^+_{\lfloor n^{x/2}\rfloor} > n\right) \\ &= P\left(2\Sigma^+_{\lfloor n^{x/2}\rfloor} > n\right) \ge P\left(\frac{2\ln\lfloor n^{x/2}\rfloor}{\ln \Sigma^+_{\lfloor n^{x/2}\rfloor}} \le x-\nu\right) \ge x - 3\nu. \ \square \end{aligned}$$

4.3. *The lifetime maximum and progeny of a critical BP.* In this subsection we prove (4.8). It is an immediate consequence of the following two lemmas.

Lemma 4.7. *Let $V_0^+ = n > 1$. Then*

$$\frac{\ln n}{\ln \max_{j<\sigma_0} V_j^+} \Rightarrow U \quad \text{as } n \to \infty.$$

Lemma 4.8. *Let $V_0^+ = n > 1$. Then*

$$\frac{\ln \sum_{j=0}^{\sigma_0-1} V_j^+}{\ln \max_{j<\sigma_0} V_j^+} \Rightarrow 2 \quad \text{as } n \to \infty.$$

Proof of Lemma 4.7: For every $x \in (0,1)$

$$P_n(\max_{j<\sigma_0} V_j^+ \ge n^{1/x}) = P_n(\tau_{n^{1/x}} < \sigma_0).$$

The proof will be complete if we can show that the last probability converges to x as $n \to \infty$. Fix a large enough $y \in \mathbb{N}$ to satisfy the conditions of Lemma A.1. Then

$$\begin{aligned} P_n(\tau_{n^{1/x}} < \sigma_0) &= P_n(\tau_{n^{1/x}} < \sigma_y) + P_n(\tau_{n^{1/x}} < \sigma_0 \,|\, \tau_{n^{1/x}} > \sigma_y)P_n(\tau_{n^{1/x}} > \sigma_y) \\ &\le P_n(\tau_{n^{1/x}} < \sigma_y) + P_y(\tau_{n^{1/x}} < \sigma_0). \end{aligned}$$

By Lemma A.1 the first term in the right-hand side of the above inequality is bounded above by $\lceil \log_2 n\rceil / \lfloor x^{-1}\log_2 n\rfloor$ which converges to x as $n \to \infty$. By Corollary A.4

$$P_y(\tau_{n^{1/x}} < \sigma_0) \le \frac{c_{10}(y)}{\lfloor x^{-1}\ln n\rfloor} \to 0 \ \text{ as } n \to \infty.$$

The lower bound is even easier. By Lemma A.1 and Remark A.2

$$\liminf_{n\to\infty} P_n(\tau_{n^{1/x}} < \sigma_0) \ge \liminf_{n\to\infty} P_n(\tau_{n^{1/x}} < \sigma_y) \ge x. \qquad \square$$

Proof of Lemma 4.8: Fix $\varepsilon \in (0,1)$, let $k_0 = \lfloor \log_2 n \rfloor$.

Lower "tail". To get a bound on the probability that the ratio in Lemma 4.8 is not less than $2-\varepsilon$, we split the path space of the process V^+ according to its lifetime maximum. On each event $\tau_{2^k} < \sigma_0 < \tau_{2^{k+1}}$, $k \ge k_0$, we shall take into account only the values of V^+ from the time τ_{2^k} up until the time $\sigma'_{2^{k-1}} := \inf\{i > \tau_{2^k} : V_i^+ \le 2^{k-1}\}$. On the time interval $\{i \in \mathbb{N} : \tau_{2^k} \le i < \sigma'_{2^{k-1}}\}$ the process V^+ stays above 2^{k-1} and below 2^{k+1}. Thus,

$$\begin{aligned}
&P_n\left(\sum_{i=0}^{\sigma_0-1} V_i^+ \le \left(\max_{i<\sigma_0} V_i^+\right)^{2-\varepsilon}\right)\\
&\quad \le \sum_{k=k_0}^{\infty} P_n\left(\sum_{i=\tau_{2^k}}^{\sigma'_{2^{k-1}}-1} V_i^+ \le 2^{(k+1)(2-\varepsilon)}, \tau_{2^k} < \sigma_0 < \tau_{2^{k+1}}\right)\\
&\quad \le \sum_{k=k_0}^{\infty} P_n\left(2^{k-1}(\sigma'_{2^{k-1}} - \tau_{2^k}) \le 2^{(k+1)(2-\varepsilon)}, \tau_{2^k} < \sigma_0 < \tau_{2^{k+1}}\right)\\
&\quad \le \sum_{k=k_0}^{\infty} E_n\left(\mathbb{1}_{\{\tau_{2^k}<\sigma_0\}} P_n\left(\sigma'_{2^{k-1}} - \tau_{2^k} \le 2^{k(1-\varepsilon)+3}, \sigma_0 < \tau_{2^{k+1}} \,\middle|\, \mathcal{F}_{\tau_{2^k}}\right)\right)\\
&\quad \le \sum_{k=k_0}^{\infty} P_n(\tau_{2^k} < \sigma_0) P_{2^k}\left(\sigma'_{2^{k-1}} \le 2^{k(1-\varepsilon)+3}, \sigma_0 < \tau_{2^{k+1}}\right) = \sum_{k=k_0}^{\infty} A_{n,k} B_k,
\end{aligned}$$

where $A_{n,k} = P_n(\tau_{2^k} < \sigma_0)$ and $B_k = P_{2^k}\left(\sigma'_{2^{k-1}} \le 2^{k(1-\varepsilon)+3}, \sigma_0 < \tau_{2^{k+1}}\right)$, which we estimate separately.

Let $\ell_0 < k_0$ be fixed as in Lemma A.1, k_0 be sufficiently large, and $k \ge k_0 + 2$ (for $k = k_0, k_0+1$ we shall use the trivial bound $A_{n,k} \le 1$). Then

$$\begin{aligned}
A_{n,k} &= P_n(\tau_{2^k} < \sigma_{2^{\ell_0}}) + P_n(\tau_{2^k} < \sigma_0 \,|\, \sigma_{2^{\ell_0}} < \tau_{2^k}) P_n(\sigma_{2^{\ell_0}} < \tau_{2^k})\\
&\le P_{2^{k_0+1}}(\tau_{2^k} < \sigma_{2^{\ell_0}}) + P_{2^{\ell_0}}(\tau_{2^k} < \sigma_0) \overset{\text{L. A.1, L. 3.1}}{\le} \frac{k_0+1}{k} + \frac{C(\ell_0)}{k}.
\end{aligned}$$

Fix an arbitrary $\nu > 0$. If k_0 is large enough then for all $k \ge k_0$

$$\begin{aligned}
B_k &\le P_{2^k}(\sigma_{2^{k-1}} \le 2^{k(1-\varepsilon)+3}, \sigma_0 < \tau_{2^{k+1}}, V^+_{\sigma_{2^{k-1}}} \ge 2^{k-2})\\
&\qquad + P_{2^k}(\sigma_{2^{k-1}} < \tau_{2^{k+1}}, V^+_{\sigma_{2^{k-1}}} < 2^{k-2})\\
&\overset{\text{(OS)}}{\le} E_{2^k}\left[\mathbb{1}_{\{\sigma_{2^{k-1}} \le 2^{k(1-\varepsilon)+3}\}} P_{2^k}(\sigma_0 < \tau_{2^{k+1}}, V^+_{\sigma_{2^{k-1}}} \ge 2^{k-2} \,|\, \mathcal{F}_{2^{k-1}})\right]\\
&\qquad + c_7 \exp(-c_9 2^k)\\
&\le P_{2^k}(\sigma_{2^{k-1}} \le 2^{k(1-\varepsilon)+3}) P_{2^{k-2}}(\sigma_{2^{\ell_0}} < \tau_{2^{k+1}}) + c_7 \exp(-c_9 2^k)\\
&\overset{\text{(DA), L. A.1}}{\le} \frac{\nu}{k-\ell_0} + c_7 \exp(-c_9 2^k).
\end{aligned}$$

Substituting the estimates for $A_{n,k}$ and B_k we get that for all sufficiently large n

$$P_n\left(\sum_{i=0}^{\sigma_0-1} V_i^+ \le \left(\max_{i<\sigma_0} V_i^+\right)^{2-\varepsilon}\right) \le 3\nu + \nu(k_0+1+C(\ell_0)) \sum_{k=k_0}^{\infty} \frac{1}{k(k-\ell_0)} < C_1(\ell_0)\nu.$$

Upper "tail". To get a bound on the probability that the ratio in Lemma 4.8 is at least $2+\varepsilon$ we let

$$O_j := \frac{1}{2^j}\sum_{i=0}^{\sigma_0-1} \mathbb{1}_{\{2^j\le V_i^+<2^{j+1}\}},\quad k^* := \lfloor \log_2 \max_{i<\sigma_0} V_i^+\rfloor,\quad m_k = \lfloor 2^{\varepsilon k-2}\rfloor,$$

and use a crude "union bound":

$$P_n\left(\sum_{i=0}^{\sigma_0-1} V_i^+ \ge \left(\max_{i<\sigma_0} V_i^+\right)^{2+\varepsilon}\right) \le P_n\left(\sum_{j=0}^{k^*} 2^{2j+1} O_j \ge 2^{k^*(2+\varepsilon)}\right)$$
$$\le P_n\left(\max_{0\le j\le k^*} O_j \ge m_{k^*}\right) \le \sum_{k=k_0}^{\infty}\sum_{j=0}^{k} P_n(O_j\ge m_k). \quad (4.13)$$

To estimate the (rescaled) time O_j which the process V^+ spends in the interval $[2^j,2^{j+1})$, $j\ge 0$, we define

$$\rho_0^{(j)} := \inf\{i\ge 0:\ V_i^+\in[2^j,2^{j+1})\},\quad \rho_m^{(j)} := \inf\{i\ge \rho_{m-1}^{(j)}+2^j:\ V_i^+\in[2^j,2^{j+1})\}$$

for $m\in\mathbb{N}$. Then by the strong Markov property

$$P_n(O_j\ge m_k)\le P_n(\rho_{m_k}^{(j)}<\sigma_0)\le P_n(\rho_{m_k}^{(j)}<\sigma_0\,|\,\rho_{m_k-1}^{(j)}<\sigma_0)P_n(\rho_{m_k-1}^{(j)}<\sigma_0)$$
$$\le \left(\max_{2^j\le x<2^{j+1}} P_x(\rho_1^{(j)}<\sigma_0)\right)^{m_k} P_n(\rho_0^{(j)}<\sigma_0).$$

We notice that by (DA) there is a $c>0$ such that $P_{2^{j+1}}(\sigma_{2^{j-1}}<2^{j-1})>c$ for all $j\ge 2$, and choosing ℓ_0 as in Lemma A.1 we get that if $(\ell_0+1)\wedge c_{10}(2^{\ell_0})<j\le k$ where c_{10} is from Corollary A.4 then

$$\max_{2^j\le x<2^{j+1}} P_x(\rho_1^{(j)}<\sigma_0)\le 1-\min_{2^j\le x<2^{j+1}} P_x(\rho_1^{(j)}>\sigma_0,\sigma_{2^{j-1}}<2^{j-1})$$
$$\le 1-\min_{2^j\le x<2^{j+1}} P_x(\rho_1^{(j)}>\sigma_0\,|\,\sigma_{2^{j-1}}<2^{j-1})P_{2^{j+1}}(\sigma_{2^{j-1}}<2^{j-1})$$
$$\le 1-cP_{2^{j-1}}(\sigma_0<\tau_{2^j})\le 1-cP_{2^{j-1}}(\sigma_0<\tau_{2^j},\sigma_{2^{\ell_0}}<\tau_{2^j})$$
$$\le 1-cP_{2^{\ell_0}}(\sigma_0<\tau_{2^j})P_{2^{j-1}}(\sigma_{2^{\ell_0}}<\tau_{2^j}) \underset{\text{Rem.A.2}}{\overset{\text{Cor.A.4}}{\le}} 1-c\left(1-\frac{c_{10}(\ell_0)}{j}\right)\frac{1}{j}\le 1-\frac{c'}{k}.$$

Choosing k_0 large enough we can also ensure that for all $k\ge k_0$

$$\max_{0\le j\le(\ell_0+1)\wedge c_{10}(\ell_0)}\ \max_{2^j\le x<2^{j+1}} P_x(\rho_1^{(j)}<\sigma_0)\le 1-c'/k.$$

Substituting these estimates in (4.13) we conclude that

$$P_n\left(\sum_{i=0}^{\sigma_0-1} V_i^+\ge\left(\max_{i<\sigma_0}V_i^+\right)^{2+\varepsilon}\right)\le\sum_{k=k_0}^{\infty}(k+1)\left(1-\frac{c'}{k}\right)^{m_k}\to 0\ \text{ as } n\to\infty. \quad \square$$

4.4. *Proofs of Lemmas 4.5 and 4.6.* We shall need the following result.

Lemma 4.9 ((4.4) from Theorem 4.1 of Kosygina and Zerner (2014))**.** *Let $(Y^-(t))$, $t\ge 0$, be the solution of*

$$dY^-(t) = -dt+\sqrt{2Y^-(t)}\,dB(t),\quad Y^-(0)=1,\quad t\in[0,\tau_0].$$

Then for every $h > 0$

$$\lim_{n\to\infty} P^{V^-}\left(\Sigma_n^- > hn^2\right) = P_1^{Y^-}\left(\int_0^{\tau_0} Y^-(s)\,ds > h\right). \tag{4.14}$$

Proof of Lemma 4.5: Upper bound:

$$\begin{aligned} P_0(\Sigma_{d_n}^- \le n^x) &\le P_0(d_n \le n^{x/2}\ln n) + P_0(\Sigma_{d_n}^- \le n^x, d_n > n^{x/2}\ln n) \\ &\le P_0(d_n \le n^{x/2}\ln n) + P_{\lfloor n^{x/2}\ln n\rfloor}(\Sigma^-_{\lfloor n^{x/2}\ln n\rfloor} \le n^x) \\ &\overset{(4.14)}{\le} P_0(d_n \le n^{x/2}\ln n) + \nu. \end{aligned}$$

Lower bound:

$$\begin{aligned} P_0(\Sigma_{d_n}^- \le n^x) &\ge P_0(d_n \le n^{x/2}/\ln n) - P_0(\Sigma_{d_n}^- > n^x, d_n \le n^{x/2}/\ln n) \\ &\ge P_0(d_n \le n^{x/2}/\ln n) - P_{\lfloor n^{x/2}/\ln n\rfloor}(2\Sigma^-_{\lceil n^{x/2}\ln n\rceil} > n^x) \\ &\ge P_0(d_n \le n^{x/2}/\ln n) - \nu. \end{aligned}$$

□

Proof of Lemma 4.6: Let L_n be the number of visits of X to 0 up to time n inclusively. Since $0 \le L_n - (u_n + d_n) \le 1$ and the ERW with $\delta = 1$ is recurrent, we have that $L_n - u_n - 1 \le d_n \le L_n - u_n$, $L_n \to \infty$ a.s., and both (4.9) and (4.10) would follow if we show that

$$\frac{u_n}{L_n} \to \frac{1}{2} \quad \text{as } n \to \infty \text{ a.s.}. \tag{4.15}$$

Notice that

$$\frac{\sum_{i=M+1}^{L_n} \eta_0(i)}{L_n} \le \frac{u_n}{L_n} \le \frac{M + \sum_{i=M+1}^{L_n} \eta_0(i)}{L_n}. \tag{4.16}$$

As $L_n \to \infty$ a.s. as $n \to \infty$, the rightmost and leftmost ratios in (4.16) a.s. converge to $1/2$ by the strong law of large numbers for Bernoulli trials. □

4.5. *Proof of Theorem 1.9(b),(c).*

Proof of Theorem 1.9(b): Let X_n^0 denote the excited random walk in the cookie environment obtained by removing all cookies from the positive semi-axis. The same proof as for Dolgopyat and Kosygina (2012, Theorem 1.1) shows that

$$\frac{X^0_{\lfloor n\cdot\rfloor}}{\sqrt{n}} \overset{J_1}{\Rightarrow} W_{0,-1}. \tag{4.17}$$

Namely, we write $X_n^0 = B_n^0 + C_n^0$, where $B_0^0 = C_0^0 = 0$ and

$$B^0_{n+1} - B^0_n = X^0_{n+1} - X^0_n, \quad C^0_{n+1} - C^0_n = 0$$

if X^0 visited X_n^0 at least M times before time n and

$$B^0_{n+1} - B^0_n = 0, \quad C^0_{n+1} - C^0_n = X^0_{n+1} - X^0_n$$

otherwise. Then we can show that

$$\left(\frac{B^0_{\lfloor n\cdot\rfloor}}{\sqrt{n}}, \frac{C^0_{\lfloor n\cdot\rfloor}}{\sqrt{n}}\right) \overset{J_1}{\Rightarrow} \left(B(\cdot), -\min_{s\le\cdot} B(s)\right),$$

and obtain (4.17). We refer to Dolgopyat and Kosygina (2012) for full details. Since there is an obvious coupling such that $X^0_{T_k^-} = X_{T_k^-}$, $k \ge 0$, the result follows from Theorem 1.7(b) and the continuity of the map ψ defined in (4.2). □

Proof of Theorem 1.9(c): This result admits the same proof as the one for Corollary 1.8(b) but, since $A_n^+/n \to 1$ for $\delta = 1$, we can give a simpler derivation.

Without loss of generality we show the convergence on $D([0,1])$. Write

$$X_{T_m^+} = X_m + (X_{T_m^+} - X_m).$$

By Lemma 4.1 for each $\varepsilon, \nu > 0$ and all large n

$$P\left(\max_{m\le n}(T_m^+ - m) \ge \varepsilon n\right) \le \nu.$$

On the other hand, given arbitrary positive ε and ν we can choose $\lambda > 0$ so that

$$P\left(\sup_{0\le s\le t\le s+\lambda\le 1+\lambda}(B^*(t) - B^*(s)) > \varepsilon\right) \le \nu.$$

The above inequalities and Theorem 1.6 imply that for any fixed $\varepsilon, \nu > 0$ and all sufficiently large n

$$P\left(\max_{m\le n}|X_{T_m^+} - X_m| > \varepsilon\sqrt{n}\ln n\right) < \nu.$$

Theorem 1.6 and the "convergence together" theorem Billingsley (1999, Theorem 3.1) imply the desired result. □

Appendix A. Proofs of Lemmas 3.1 - 3.3

Throughout this section we assume that $\delta = 1$ unless stated otherwise. The following lemma plays an important role in proofs of Lemmas 3.1 and 3.2.

Lemma A.1 (Main lemma)**.** *Let*

$$h^\pm(n) := n \pm \frac{1}{n} \qquad \textit{for all } n \in \mathbb{N}.$$

Then there is $\ell_0 \in \mathbb{N}$ such that if $\ell, m, u, x \in \mathbb{N}$ satisfy $\ell_0 \le \ell < m < u$ and $|x - 2^m| \le 2^{2m/3}$ then

$$\frac{h^-(u) - h^-(m)}{h^-(u) - h^-(\ell)} \le P_x[\sigma_{2^\ell} < \tau_{2^u}] \le \frac{h^+(u) - h^+(m)}{h^+(u) - h^+(\ell)}. \tag{A.1}$$

Remark A.2. A little algebra shows that the lower bound is at least $1 - m/u$.

The proof of Lemma A.1 is the same as that of Lemma 5.3 in Kosygina and Mountford (2011) where we take $a = 2$, $h_a^\pm(n) = n \pm 1/n$, and use the following result instead of Kosygina and Mountford (2011, Lemma 5.2).

Lemma A.3. *Consider the process V with $|V_0 - 2^n| \le 2^{2n/3}$ and let $T := \inf\{k \ge 0 : V_k \notin (2^{n-1}, 2^{n+1})\}$. Then for all sufficiently large n*

$$P(\mathrm{dist}(V_T, (2^{n-1}, 2^{n+1})) \ge 2^{2(n-1)/3}) \le \exp(-2^{n/4}); \tag{A.2}$$

$$\left|P(V_T \le 2^{n-1}) - \frac{1}{2}\right| \le 2^{-n/4}. \tag{A.3}$$

The proof of the above lemma repeats the one of Kosygina and Mountford (2011, Lemma 5.2) where we use our process V, set $a = 2$, and $s(x) = \ln x$ on $(3^{-1}, 3)$.

Corollary A.4. *For every $y \in \mathbb{N}$ there is a constant $c_{10}(y)$ such that for every $n \in \mathbb{N}$*

$$(\ln n)P_y(\sigma_0 > \tau_n) \le c_{10}(y).$$

The proof of this corollary is the same as that of (5.4) in Kosygina and Mountford (2011) and uses Lemma A.1 instead of Lemma 5.3 of Kosygina and Mountford (2011).

Corollary A.5. $P_y^V(\sigma_0^V<\infty)=1$ *for every* $y\in\mathbb{N}$.

Proof: By Corollary A.4 and the fact that $P_y(\sigma_0=\infty,\tau_n=\infty)=0$ for $n>y$,

$$P_y(\sigma_0=\infty)=P_y(\sigma_0=\infty,\tau_n<\infty)\le P_y(\sigma_0>\tau_n)\le\frac{c_{10}(y)}{\ln n}\to 0\quad\text{as } n\to\infty.\ \square$$

Remark A.6. Corollary A.5, the first statement of Kosygina and Zerner (2014, Corollary 7.9), and symmetry imply that ERW with $|\delta|=1$ is recurrent without using any results from the literature on branching processes. A direct proof of recurrence and transience results for $|\delta|\neq 1$ was obtained in Kosygina and Zerner (2014, Corollary 7.10).

Proof of Lemma 3.1: For every $n>2$ there is an $m\in\mathbb{N}$ such that $2^m\le n<2^{m+1}$ and for this m

$$(\ln 2^m)P_y(\sigma_0>\tau_{2^{m+1}})\le(\ln n)P_y(\sigma_0>\tau_n)\le(\ln 2^{m+1})P_y(\sigma_0>\tau_{2^m}).$$

If we can show the existence of

$$g(y):=\lim_{m\to\infty}mP_y(\sigma_0>\tau_{2^m})\in(0,\infty),\tag{A.4}$$

then we get

$$\begin{aligned}&\limsup_{n\to\infty}(\ln n)P_y(\sigma_0>\tau_n)\le\ln 2\lim_{m\to\infty}(m+1)P_y(\sigma_0>\tau_{2^m})=(\ln 2)g(y)\\&=\ln 2\lim_{m\to\infty}mP_y(\sigma_0>\tau_{2^{m+1}})\le\liminf_{n\to\infty}(\ln n)P_y(\sigma_0>\tau_n),\end{aligned}$$

and the desired statement follows. Therefore, we shall show (A.4). Let $\ell=(\lfloor\log_2 y\rfloor+1)\vee\ell_0$, where ℓ_0 is the same as in Lemma A.1. Then

$$\begin{aligned}mP_y(\sigma_0>\tau_{2^m})&=m\left[\prod_{j=\ell+1}^m P_y(\sigma_0>\tau_{2^j}\,|\,\sigma_0>\tau_{2^{j-1}})\right]P_y(\sigma_0>\tau_{2^\ell})\\&=\ell P_y(\sigma_0>\tau_{2^\ell})\left[\prod_{j=\ell+1}^m\frac{j}{j-1}P_y(\sigma_0>\tau_{2^j}\,|\,\sigma_0>\tau_{2^{j-1}})\right].\end{aligned}$$

We need to prove that the last product converges. For this it is sufficient to show that

$$\sum_{j=\ell+1}^\infty\left|\frac{j}{j-1}P_y(\sigma_0>\tau_{2^j}\,|\,\sigma_0>\tau_{2^{j-1}})-1\right|<\infty.$$

Lemma A.1 and Corollary A.4 allow us to obtain the necessary estimates.

$$\begin{aligned}&\frac{j}{j-1}P_y(\sigma_0>\tau_{2^j}\,|\,\sigma_0>\tau_{2^{j-1}})-1\ge\frac{j}{j-1}P_{2^{j-1}}(\sigma_0>\tau_{2^j})-1\\&\quad\ge\frac{j}{j-1}P_{2^{j-1}}(\sigma_{2^\ell}>\tau_{2^j})-1\overset{(A.1)}{\ge}\frac{j}{j-1}\,\frac{j-1+\frac{1}{j-1}-\ell-\frac{1}{\ell}}{j+\frac{1}{j}-\ell-\frac{1}{\ell}}-1\\&\quad=\frac{\frac{2j-1}{j^2(j-1)^2}-\frac{\ell}{j(j-1)}-\frac{1}{\ell j(j-1)}}{1+\frac{1}{j^2}-\frac{\ell}{j}-\frac{1}{\ell j}}\ge\frac{2}{j^2(j-1)}-\frac{\ell}{j(j-1)}-\frac{1}{\ell j(j-1)}.\end{aligned}$$

The right hand side of the above expression is a term of an absolutely convergent series.

Set $x := 2^{j-1} + 2^{2(j-1)/3}$. Then

$$\frac{j}{j-1} P_y(\sigma_0 > \tau_{2^j} \,|\, \sigma_0 > \tau_{2^{j-1}}) \le \frac{j}{j-1}\big(P_x(\sigma_0 > \tau_{2^j}) + P_y(V_{\tau_{2^{j-1}}} > x \,|\, \sigma_0 > \tau_{2^{j-1}})\big).$$

By (OS) the last term decays exponentially fast in j, and we shall concentrate on the first term in the right hand side of the above inequality. For all sufficiently large j

$$\begin{aligned}
&\frac{j}{j-1} P_x(\sigma_0 > \tau_{2^j}) - 1\\
&\quad \le \frac{j}{j-1} P_x(\sigma_{2^\ell} > \tau_{2^j}) - 1 + \frac{j}{j-1} P_x(\sigma_0 > \tau_{2^j} \,|\, \sigma_{2^\ell} < \tau_{2^j}) P_x(\sigma_{2^\ell} < \tau_{2^j})\\
&\quad \overset{(A.1)}{\le} \frac{j}{j-1}\,\frac{j-1}{j} - 1 + \frac{j}{j-1} P_{2^\ell}(\sigma_0 > \tau_{2^j}) \frac{j + \frac{1}{j} - (j-1) - \frac{1}{j-1}}{j + \frac{1}{j} - \ell - \frac{1}{\ell}}\\
&\quad \le \frac{j}{(j-1)(j-\ell-1)} P_{2^\ell}(\sigma_0 > \tau_{2^j}) \overset{\text{Cor. A.4}}{\le} \frac{C(\ell)}{(j-1)(j-\ell-1)}.
\end{aligned}$$

Again the last expression is a term of a convergent series, and we are done. □

The proof of Lemma 3.2 depends on an estimate of the time the branching process V spends in an interval $[x, 2x)$ before extinction.

Lemma A.7. *For every $\alpha > 1$ there is a constant $c_{11}(\alpha) \in (0,1)$ such that for all $k, x, y \in \mathbb{N}$*

$$P_y\left(\sum_{j=0}^{\sigma_0 - 1} \mathbb{1}_{[x,2x)}(V_j) > 2kx^\alpha\right) \le P_y(\rho_0 < \sigma_0)(1 - c_{11}(\alpha))^k,$$

where $\rho_0 := \inf\{j \ge 0 : V_j \in [x, 2x)\}$;

Proof: The proof is very similar to the one of Proposition 6.1 in Kosygina and Mountford (2011). There are two differences. First, everywhere in the proof of Proposition 6.1 the statement (ii) should be replaced with the following: there is a constant $c = c(\alpha) > 0$ such that for all $x \in \mathbb{N}$

$$P_{x/2}(\sigma_0 < \tau_{x^\alpha}) > c. \tag{A.5}$$

Second, the stopping times ρ_j, $j \in \mathbb{N}$, should be defined as follows: ρ_0 was defined above,

$$\rho_j = \inf\{r \ge \rho_{j-1} + 2x^\alpha : V_r \in [x, 2x)\}, \quad j \ge 1.$$

Below we show (A.5). The rest of the proof is the same as that of Kosygina and Mountford (2011, Proposition 6.1).

To prove (A.5) we fix a large $y \in \mathbb{N}$ and observe that by Corollary A.4 and Remark A.2 for all $x > 2y + 1$

$$\begin{aligned}
P_{x/2}(\sigma_0 < \tau_{x^\alpha}) &= P_{x/2}(\sigma_0 < \tau_{x^\alpha} \,|\, \sigma_y < \tau_{x^\alpha}) P_{x/2}(\sigma_y < \tau_{x^\alpha})\\
&\ge P_y(\sigma_0 < \tau_{x^\alpha}) P_{x/2}(\sigma_y < \tau_{x^\alpha}) = (1 - P_y(\sigma_0 > \tau_{x^\alpha})) P_{x/2}(\sigma_y < \tau_{x^\alpha})\\
&\ge \left(1 - \frac{c_{10}(y)}{\alpha \ln x}\right)\left(1 - \frac{\ln(x/2)}{\alpha \ln x}\right) \ge \left(1 - \frac{c_{10}(y)}{\alpha \ln x}\right)\frac{\alpha - 1}{\alpha} > c > 0.
\end{aligned}$$

Adjusting the constant c if necessary we obtain (A.5) for all $x \in \mathbb{N}$. $\square$

Proof of Lemma 3.2: For every $n \in \mathbb{N}$ let $k \in \mathbb{N}$ be such that $2^{k-1} \le n < 2^k$. We can always write α as $\alpha' + \lambda$ where $\alpha' > 1$ and $\lambda > 0$. Then

$$\begin{aligned} P_y\left(\sum_{j=0}^{\sigma_0-1} \mathbb{1}_{\{V_j \le n\}} > n^\alpha\right) &\le P_y\left(\sum_{j=0}^{\sigma_0-1} \mathbb{1}_{\{V_j < 2^k\}} > 2^{\alpha(k-1)}\right) \\ &\le P_y\left(\sum_{j=0}^{\sigma_0-1}\sum_{i=1}^{k} \mathbb{1}_{[2^{i-1},2^i)}(V_j) > 2^{\lambda(k-1)}(1-2^{-\alpha'})\sum_{i=1}^{k} 2^{\alpha'(i-1)}\right) \\ &\le \sum_{i=1}^{k} P_y\left(\sum_{j=0}^{\sigma_0-1} \mathbb{1}_{[2^{i-1},2^i)}(V_j) > 2^{\lambda(k-1)}(1-2^{-\alpha'})2^{\alpha'(i-1)}\right) \\ &\overset{\text{Lem. A.7}}{\le} k(1-c_{11}(\alpha'))^{\lfloor 2^{\lambda(k-1)-1}(1-2^{-\alpha'})\rfloor}. \end{aligned}$$

Multiplying by $\ln n$ which is less than $k \ln 2$ we get that as $n \to \infty$

$$(\ln n)\, P_y\left(\sum_{j=0}^{\sigma_0-1} \mathbb{1}_{\{V_j \le n\}} > n^\alpha\right) \le (\ln 2)k^2(1-c_{11}(\alpha'))^{\lfloor 2^{\lambda(k-1)-1}(1-2^{-\alpha'})\rfloor} \to 0. \quad \square$$

Before we turn to the proof of Lemma 3.3 we present its continuous space-time version.

Lemma A.8. *Let Y be the diffusion defined by (3.3) which starts at* 1 *and* $\tau_\varepsilon := \inf\{t \ge 0 : Y(t) = \varepsilon\}$. *Then for every $h > 0$*

$$\lim_{\varepsilon\to 0} P_1^Y(\tau_\varepsilon > h) = 1; \tag{A.6}$$

$$\lim_{\varepsilon\to 0} P_1^Y\left(\int_0^{\tau_\varepsilon} Y(t)\,dt > h\right) = 1. \tag{A.7}$$

Lemma A.8 follows from the fact that 0 is an inaccessible point for the two-dimensional squared Bessel process. The details are left to the reader.

Proof of Lemma 3.3: We prove only (3.2), since the proof of (3.1) is the same (it uses (A.6) instead of (A.7)). Notice that

$$\lim_{n\to\infty} P_n\left(\sum_{i=0}^{\sigma_0-1} V_i > hn^2\right) \ge \lim_{\varepsilon\to 0}\lim_{n\to\infty} P_n\left(\sum_{i=0}^{\sigma_{\varepsilon n}-1} V_i > hn^2\right).$$

By the diffusion approximation, for every $\varepsilon \in (0,1)$

$$\lim_{n\to\infty} P_n\left(\sum_{i=0}^{\sigma_{\varepsilon n}-1} V_i > hn^2\right) = P_1^Y\left(\int_0^{\tau_\varepsilon} Y(t)\,dt > h\right),$$

and by (A.7),

$$\lim_{\varepsilon\to 0} P_1^Y\left(\int_0^{\tau_\varepsilon} Y(t)\,dt > h\right) = 1. \quad \square$$

Appendix B. Continuity of maps φ and ψ

Denote by meas A the Lebesgue measure of set A.

Proposition B.1. *Let P be a probability measure supported on $C([0,1])$ such that P-a.s.*

$$\operatorname{meas}\{t \in [0,1] : \omega(t) = 0\} = 0. \tag{B.1}$$

Then the map φ defined by (4.1) is P-a.s. continuous.

Proof: It is sufficient to show continuity at every $\omega \in C([0,1])$ which satisfies (B.1). Let $\varpi \in D([0,1])$ and $\sup_{t\in[0,1]} |\varpi(t) - \omega(t)| \le \nu$.[5] Then

$$\int_0^1 \mathbb{1}_{[\nu,\infty)}(\omega(t))\,dt \le \varphi(\varpi) \le \int_0^1 \mathbb{1}_{[-\nu,\infty)}(\omega(t))\,dt \quad \text{and}$$

$$|\varphi(\omega) - \varphi(\varpi)| \le \int_0^1 \mathbb{1}_{[-\nu,\nu]}(\omega(t))\,dt = \operatorname{meas}\{t \in [0,1] :\, -\nu \le \omega(t) \le \nu\}. \tag{B.2}$$

Since $\{t \in [0,1] :\, -\nu \le \omega(t) \le \nu\} \searrow \{t \in [0,1] :\, \omega(t) = 0\}$ and $\operatorname{meas}\{t \in [0,1] : \omega(t) = 0\} = 0$, given $\varepsilon > 0$ we can choose $\nu > 0$ such that the right-hand side of (B.2) is less than ε. □

Proposition B.2. *Let P be a probability measure supported on $C([0,\infty))$ such that P-a.s.*

$$\mathit{meas}\{t \ge 0 :\, \omega(t) = 0\} = 0 \quad \mathit{and} \quad \mathit{meas}\{t \ge 0 : \omega(t) < 0\} = \infty. \tag{B.3}$$

Then the map ψ defined by (4.2) is P-a.s. continuous.

Proof: It is sufficient to show continuity at every $\omega \in C([0,\infty))$ which satisfies (B.3). Fix such an ω and let $\varepsilon > 0$. Recall that $T^-_\omega(s) := \inf\{t \ge 0 :\, \operatorname{meas}\{r \in [0,t] :\, \omega(r) < 0\} > s\}$. Given $R > 0$ let M be chosen so that $T^-_\omega(M) = R + 1$. We need to find ν such that if $\varpi \in D([0,\infty))$ satisfies

$$\sup_{t\in[0,M]} |\varpi(t) - \omega(t)| < \nu \tag{B.4}$$

then

$$\sup_{t\in[0,R]} |\omega(T^-_\omega(t)) - \varpi(T^-_\varpi(t))| < \varepsilon. \tag{B.5}$$

We denote $\lim_{s\uparrow t} \varpi(s)$ by $\varpi(t-)$. Note that due to (B.4) for $t \in (0,M]$ we have

$$|\varpi(t-0) - \omega(t)| < \nu. \tag{B.6}$$

Choose h such that

$$\sup_{t',t''\in[0,M]:|t'-t''|<3h} |\omega(t') - \omega(t'')| < \varepsilon/8. \tag{B.7}$$

Next choose $\nu < \varepsilon/8$ such that

$$\operatorname{meas}\{t \in [0,M] :\, |\omega(t)| \le \nu\} < h. \tag{B.8}$$

[5] Recall that for $\omega \in C([0,1])$ the Skorokhod convergence to ω implies the uniform convergence (see Billingsley (1999, the last paragraph on p. 128)). Thus, it is sufficient to work with the sup norm.

Let ϖ satisfy (B.4). Then for $t \in [0, R]$ we have

$$|\omega(T^-_\omega(t)) - \varpi(T^-_\varpi(t))| \le |\omega(T^-_\omega(t)) - \omega(T^-_\varpi(t))| + |\omega(T^-_\varpi(t)) - \varpi(T^-_\varpi(t))| \\ \le |\omega(T^-_\omega(t)) - \omega(T^-_\varpi(t))| + \nu. \tag{B.9}$$

For $f \in D([0,\infty))$ let $A^-_f(t) := \text{meas}\{s \in [0,t] : f(s) < 0\} = \int_0^t \mathbb{1}_{(-\infty,0)}(f(s))\, ds$. The definition implies that $A^-_f \in C([0,\infty))$ and $A^-_f(T^-_f(t)) \equiv t$. Note that due to (B.4) we have

$$A^-_{\omega+\nu}(s) \le A^-_\varpi(s) \le A^-_{\omega-\nu}(s)$$

and due to (B.8) we have

$$A^-_{\omega-\nu}(s) - h \le A^-_\omega(s) \le A^-_{\omega+\nu}(s) + h.$$

Therefore,

$$t - h = A^-_\omega(T^-_\omega(t)) - h \le A^-_{\omega+\nu}(T^-_\omega(t)) \le A^-_\varpi(T^-_\omega(t)) \le A^-_{\omega-\nu}(T^-_\omega(t)) \\ \le A^-_\omega(T^-_\omega(t)) + h = t + h.$$

We now consider 4 cases.

(I) $t - h \le A^-_\varpi(T^-_\omega(t)) \le t$ (which implies that $T^-_\omega(t) \le T^-_\varpi(t)$) and $\omega(u) < 0$ for $u \in [T^-_\omega(t), T^-_\varpi(t)]$.

Then, since $A^-_\varpi(s) - A^-_\varpi(r) \ge s - r - \text{meas}\{u \in [r,s] : \varpi(u) \ge 0\}$ for $s \ge r$ and $\varpi(u) \ge 0 \overset{(B.4)}{\Rightarrow} \omega(u) \ge -\nu$ for all $u \in [T^-_\omega(t), T^-_\varpi(t)]$, we have by (B.8) that

$$h \ge A^-_\varpi(T^-_\varpi(t)) - A^-_\varpi(T^-_\omega(t)) \ge T^-_\varpi(t) - T^-_\omega(t) - h.$$

Hence, $T^-_\varpi(t) - T^-_\omega(t) \le 2h$ and so by (B.7) $|\omega(T^-_\varpi(t)) - \omega(T^-_\omega(t))| \le \varepsilon/8$.

(II) $t - h \le A^-_\varpi(T^-_\omega(t)) \le t$ and $\omega(\cdot)$ has zeroes on $[T^-_\omega(t), T^-_\varpi(t)]$.

Let a be the first zero and b be the last zero of $\omega(\cdot)$ on $[T^-_\omega(t), T^-_\varpi(t)]$. Notice that $\omega(T^-_\omega(t)) \le 0$. Thus, $\omega(s) \le 0$ for $s \in [T^-_\omega(t), a]$ and the same argument as in case (I) shows that

$$|\omega(T^-_\omega(t))| = |\omega(T^-_\omega(t)) - \omega(a)| \le \varepsilon/8.$$

Moreover if $\omega(T^-_\varpi(t)) \le 0$ then by the same argument we also have

$$|\omega(T^-_\varpi(t))| = |\omega(T^-_\varpi(t)) - \omega(b)| \le \varepsilon/8.$$

On the other hand, if $\omega(T^-_\varpi(t)) > 0$ then, since ω is continuous and $\varpi(T^-_\varpi(t)-)) \le 0$, we get

$$|\omega(T^-_\varpi(t))| = \omega(T^-_\varpi(t)) \le \omega(T^-_\varpi(t)) - \varpi(T^-_\varpi(t)-) < \nu < \varepsilon/8.$$

In either case we obtain

$$|\omega(T^-_\varpi(t)) - \omega(T^-_\omega(t))| \le \varepsilon/4.$$

(III) $t < A^-_\varpi(T^-_\omega(t)) \le t + h$ and $\varpi(u) < 0$ for $u \in [T^-_\varpi(t), T^-_\omega(t)]$.

Then $h \ge A^-_\varpi(T^-_\omega(t)) - A^-_\varpi(T^-_\varpi(t)) = T^-_\omega(t) - T^-_\varpi(t)$, and so by (B.7)

$$|\omega(T^-_\omega(t)) - \omega(T^-_\varpi(t))| \le \varepsilon/8.$$

(IV) $t < A^-_\varpi(T^-_\omega(t)) \le t + h$ and $\varpi(\cdot)$ takes non-negative values somewhere on $[T^-_\varpi(t), T^-_\omega(t)]$.

Let

$$a = \inf\{u \in [T_{\varpi}^{-}(t), T_{\omega}^{-}(t)] : \varpi(u) \geq 0\} \quad \text{and}$$
$$b = \inf\{u \in [T_{\varpi}^{-}(t), T_{\omega}^{-}(t)] : \varpi(u)\varpi(s) > 0 \ \forall s \in [u, T_{\omega}^{-}(t))\}.$$

Observe that by (B.4) and continuity of $\omega(\cdot)$ it holds that $|\omega(a)| < \nu$ and $|\omega(b)| < \nu$. Next, the same argument as in case (III) shows that $|\omega(a) - \omega(T_{\varpi}^{-}(t))| \leq \varepsilon/8$. Moreover, if $\varpi(T_{\omega}^{-}(t)-) < 0$ then we also have that $|\omega(T_{\omega}^{-}(t)) - \omega(b)| \leq \varepsilon/8$, whereas if $\varpi(T_{\omega}^{-}(t)-) \geq 0$ then, since $\omega(T_{\omega}^{-}(t)) \leq 0$, we conclude that

$$|\omega(T_{\omega}^{-}(t))| \leq |\omega(T_{\omega}^{-}(t)) - \varpi(T_{\omega}^{-}(t)-)| < \nu < \varepsilon/8.$$

Putting everything together we see that in case (IV)

$$|\omega(T_{\omega}^{-}(t)) - \omega(T_{\varpi}^{-}(t))| \leq |\omega(T_{\omega}^{-}(t)) - \omega(b)| + |\omega(b) - \omega(a)| + |\omega(a) - \omega(T_{\varpi}^{-}(t))| < \varepsilon/2.$$

Combining (B.9) with the above estimates for cases (I)–(IV) we obtain that for all $t \in [0, R]$

$$|\omega(T_{\omega}^{-}(t)) - \varpi(T_{\varpi}^{-}(t))| < 5\varepsilon/8.$$

This implies (B.5) and concludes the proof of the proposition. □

Acknowledgments

D. Dolgopyat was partially supported by the NSF grant DMS 1362064. E. Kosygina was partially supported by the Simons Foundation through a Collaboration Grant for Mathematicians #209493 and Simons Fellowship in Mathematics, 2014–2015. Parts of this work were done during authors' visit to the Fields Institute in Spring of 2011. The paper was completed during the second author's stay at the Institut Mittag-Leffler in the Fall of 2014. The authors thank both Institutes for support and excellent working conditions.

References

P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition (1999). ISBN 0-471-19745-9. A Wiley-Interscience Publication. MR1700749.

P. Carmona, F. Petit and M. Yor. Beta variables as times spent in $[0, \infty[$ by certain perturbed Brownian motions. *J. London Math. Soc. (2)* **58** (1), 239–256 (1998). MR1670130.

L. Chaumont and R.A. Doney. Pathwise uniqueness for perturbed versions of Brownian motion and reflected Brownian motion. *Probab. Theory Related Fields* **113** (4), 519–534 (1999). MR1717529.

L. Chaumont and R.A. Doney. Some calculations for doubly perturbed Brownian motion. *Stochastic Process. Appl.* **85** (1), 61–74 (2000). MR1730618.

B. Davis. Weak limits of perturbed random walks and the equation $Y_t = B_t + \alpha \sup\{Y_s\colon\ s \leq t\} + \beta \inf\{Y_s\colon\ s \leq t\}$. *Ann. Probab.* **24** (4), 2007–2023 (1996). MR1415238.

D. Dolgopyat and E. Kosygina. Scaling limits of recurrent excited random walks on integers. *Electron. Commun. Probab.* **17**, no. 35, 14 (2012). MR2965748.

E. Kosygina and T. Mountford. Limit laws of transient excited random walks on integers. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** (2), 575–600 (2011). MR2814424.

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

E. Kosygina and M.P.W. Zerner. Excited random walks: results, methods, open problems. *Bull. Inst. Math. Acad. Sin. (N.S.)* **8** (1), 105–157 (2013). MR3097419.

E. Kosygina and M.P.W. Zerner. Excursions of excited random walks on integers. *Electron. J. Probab.* **19**, no. 25, 25 (2014). MR3174837.

M. Perman and W. Werner. Perturbed Brownian motions. *Probab. Theory Related Fields* **108** (3), 357–383 (1997). MR1465164.