SCALING LIMITS OF RECURRENT EXCITED RANDOM WALKS ON INTEGERS

DMITRY DOLGOPYAT AND ELENA KOSYGINA

Abstract. We describe scaling limits of recurrent excited random walks (ERWs) on $\mathbb{Z}$ in i.i.d. cookie environments with a bounded number of cookies per site. We allow both positive and negative excitations. It is known that ERW is recurrent if and only if the expected total drift per site, δ, belongs to the interval $[-1,1]$. We show that if $|\delta|<1$ then the diffusively scaled ERW under the averaged measure converges to a $(\delta,-\delta)$-perturbed Brownian motion. In the boundary case, $|\delta|=1$, the space scaling has to be adjusted by an extra logarithmic term, and the weak limit of ERW happens to be a constant multiple of the running maximum of the standard Brownian motion, a transient process.

1. Introduction and main results

Given an arbitrary positive integer M let

$$\Omega_M := \big\{((\omega_z(i))_{i\in\mathbb{N}})_{z\in\mathbb{Z}} \mid \omega_z(i)\in[0,1], \text{ for } i\in\{1,2,\dots,M\}$$
$$\text{and } \omega_z(i)=1/2, \text{ for } i>M,\ z\in\mathbb{Z}\big\}.$$

An element of Ω_M is called a cookie environment. For each $z\in\mathbb{Z}$, the sequence $\{\omega_z(i)\}_{i\in\mathbb{N}}$ can be thought of as a stack of cookies at site z. The number $\omega_z(i)$ represents the transition probability from z to $z+1$ of a nearest-neighbor random walk upon the i-th visit to z. If $\omega_z(i)\ge 1/2$ (resp. $\omega_z(i)<1/2$) the corresponding cookie is called non-negative (resp. negative).

Let $\mathbb{P}$ be a probability measure on Ω_M, which satisfies the following two conditions:

(A1) Independence: the sequence $(\omega_z(\cdot))_{z\in\mathbb{Z}}$ is i.i.d. under $\mathbb{P}$;

(A2) Non-degeneracy: $\mathbb{E}\left[\prod_{i=1}^M \omega_0(i)\right]>0$ and $\mathbb{E}\left[\prod_{i=1}^M (1-\omega_0(i))\right]>0$.

For $x\in\mathbb{Z}$ and $\omega\in\Omega_M$ consider an integer valued process $X:=(X_j),\ j\ge 0$, on some probability space $(\mathcal{X},\mathcal{F},P_{x,\omega})$, which $P_{x,\omega}$-a.s. satisfies $P_{x,\omega}(X_0=x)=1$ and

$$P_{x,\omega}(X_{n+1}=X_n+1\,|\,\mathcal{F}_n)=1-P_{x,\omega}(X_{n+1}=X_n-1\,|\,\mathcal{F}_n)=\omega_{X_n}(L_{X_n}(n)),$$

where $\mathcal{F}_n\subset\mathcal{F}$, $n\ge 0$, is the natural filtration of X and $L_m(n):=\sum_{j=0}^n \mathbb{1}_{\{X_j=m\}}$ is the number of visits to site m by X up to time n. Informally speaking, upon each visit to a site the walker eats the topmost cookie from the stack at that site and makes one step to the right or to the left with probabilities prescribed by this cookie. The consumption

---

2000 Mathematics Subject Classification. Primary: 60K37, 60F17, 60G50.

Key words: excited random walk, cookie walk, branching process, random environment, perturbed Brownian motion.

of a cookie $\omega_z(i)$ induces a drift of size $2\omega_z(i)-1$. Since $\omega_z(i)=1/2$ for all $i>M$, the walker will make unbiased steps from z starting from the $(M+1)$-th visit to z. Let δ be *the expected total drift per site*, i.e.

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

The parameter δ plays a key role in the classification of the asymptotic behavior of the walk. For a fixed $\omega\in\Omega$ the measure $P_{\omega,x}$ is called *quenched*. The *averaged* measure P_x is obtained by averaging over environments, i.e. $P_x(\,\cdot\,):=\mathbb{E}\left(P_{x,\omega}(\,\cdot\,)\right)$.

There is an obvious symmetry between positive and negative cookies: if the environment $(\omega_z)_{z\in\mathbb{Z}}$ is replaced by $(\omega'_z)_{z\in\mathbb{Z}}$ where $\omega'_z(i)=1-\omega_z(i)$, for all $i\in\mathbb{N},\ z\in\mathbb{Z}$, then X', the ERW corresponding to the new environment, satisfies $X'\overset{\mathrm{d}}{=}-X$, where $\overset{\mathrm{d}}{=}$ denotes the equality in distribution. Thus, it is sufficient to consider only non-negative δ (this, of course, allows both negative and positive cookies), and we shall always assume this to be the case.

ERW on $\mathbb{Z}$ in a non-negative cookie environment and its natural extension to $\mathbb{Z}^d$ (when there is a direction in $\mathbb{R}^d$ such that the projection of a drift induced by every cookie on that direction is non-negative) were considered previously by many authors (see, for example, [4], [22], [23], [2], [3], [17] [5], [9], [16], and references therein).

Our model allows both positive and negative cookies but restricts their number per site to M. This model was studied in [14], [15], [20], [19]. It is known that the process is recurrent (i.e. for $\mathbb{P}$-a.e. ω it returns to the starting point infinitely often) if and only if $\delta\le 1$ ([14]). For transient (i.e. not recurrent) ERW, there is a rich variety of limit laws under P_0 ([15]).

In this paper we study scaling limits of recurrent ERW under P_0. The functional limit theorem for recurrent ERW in stationary ergodic non-negative cookie environments on strips $\mathbb{Z}\times(\mathbb{Z}/L\mathbb{Z})$, $L\in\mathbb{N}$, under the quenched measure was proven in [9]. Our results deal only with i.i.d. environments on $\mathbb{Z}$ with bounded number of cookies per site but remove the non-negativity assumption on the cookies. We are also able to treat the boundary case $\delta=1$. Extensions of these results and results of [15] to strips, or $\mathbb{Z}^d$ for $d>1$, or the "boundary" case for the model treated in [9] are still open problems.

To state our results we need to define the candidates for limiting processes. Let $D([0,\infty))$ be the Skorokhod space of càdlàg functions on $[0,\infty)$ and denote by $\overset{J_1}{\Rightarrow}$ the weak convergence in the standard (J_1) Skorokhod topology on $D([0,\infty))$. Unless stated otherwise, all processes start at the origin at time 0. Let $B=(B(t)),\ t\ge 0$, denote a standard Brownian motion and $X_{\alpha,\beta}=(X_{\alpha,\beta}(t)),\ t\ge 0$, be an (α,β)-perturbed Brownian motion, i.e. the solution of the equation

$$X_{\alpha,\beta}(t)=B(t)+\alpha\sup_{s\le t}X_{\alpha,\beta}(s)+\beta\inf_{s\le t}X_{\alpha,\beta}(s), \tag{2}$$

For $(\alpha,\beta)\in(-\infty,1)\times(-\infty,1)$ the equation (2) has a pathwise unique solution that is adapted to the filtration of B and is a.s. continuous ([18], [7]). Now we can state the results of our paper.

Theorem 1 (Non-boundary case). *If $\delta \in [0,1)$ then*

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

We note that there are other known random walk models which after rescaling converge to a perturbed Brownian motion (see, e.g., [8, 21]).

Theorem 2 (Boundary case). *Let $\delta = 1$ and $B^*(t) := \max_{s \le t} B(s)$, $t \ge 0$. Then there exists a constant $D > 0$ such that*

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

Observe that for $\delta = 1$ the limiting process is transient while the original process is recurrent. To prove Theorem 2 we consider the process $S_j := \max_{0 \le i \le j} X_i$, $j \ge 0$, and show that after rescaling it converges to the running maximum of Brownian motion. The stated result then comes from the fact that with an overwhelming probability the maximum amount of "backtracking" of X_j from S_j for $j \le [Tn]$ is of order $\sqrt{n}$, which is negligible on the scale $\sqrt{n}\log n$ (see Lemma 10).

2. Notation and preliminaries

Assume that $\delta \ge 0$ and $X_0 = 0$. Let $T_x = \inf\{j \ge 0 : X_j = x\}$ be the first hitting time of $x \in \mathbb{Z}$. Set

$$S_n = \max_{k \le n} X_k, \quad I_n = \min_{k \le n} X_k, \quad R_n = S_n - I_n + 1, \quad n \ge 0.$$

At first, we recall the connection with branching processes exploited in [2], [3], [14], and [15].

For $n \in \mathbb{N}$ and $0 \le k \le n$ define

$$D_{n,k} = \sum_{j=0}^{T_n - 1} \mathbb{1}_{\{X_j = k,\ X_{j+1} = k-1\}},$$

the number of jumps from k to $k-1$ before time T_n. Then

$$T_n = n + 2\sum_{k \le n} D_{n,k} = n + 2\sum_{0 \le k \le n} D_{n,k} + 2\sum_{k<0} D_{n,k}. \tag{3}$$

Consider the "backward" process $(D_{n,n}, D_{n,n-1} \dots, D_{n,0})$. Obviously, $D_{n,n} = 0$ for every $n \in \mathbb{N}$. Moreover, given $D_{n,n}, D_{n,n-1}, \dots, D_{n,k+1}$, we can write

$$D_{n,k} = \sum_{j=1}^{D_{n,k+1}+1} (\#\text{ of jumps from } k \text{ to } k-1 \text{ between the } (j-1)\text{-th and } j\text{-th jump from } k \text{ to } k+1 \text{ before time } T_n),\ k = 0, 1, \dots, n-1.$$

Here we used the observation that the number of jumps from k to $k+1$ before time T_n is equal to $D_{n,k+1} + 1$ for all $0 \le k \le n-1$. It follows from the definition that $(D_{n,n}, D_{n,n-1} \dots, D_{n,0})$ is a Markov process. Moreover, it can be recast as a branching process with migration (see [14], Section 3, as well as [15], Section 2). Let $V := (V_k),\ k \geq 0$, be the process such that $V_0 = 0$ and

$$(V_0, V_1, \dots, V_n) \overset{\mathrm{d}}{=} (D_{n,n}, D_{n,n-1} \dots, D_{n,0}) \qquad \text{for all } n \in \mathbb{N}.$$

Denote by $\sigma \in [1, \infty]$ and $\Sigma \in [0, \infty]$ respectively the lifetime and the total progeny over the lifetime of V, i.e. $\sigma = \inf\{k > 0 : V_k = 0\}$, $\Sigma = \sum_{k=0}^{\sigma-1} V_k$. The probability measure that corresponds to V will be denoted by P_0^V. The following result will be used several times throughout the paper.

Theorem 3 ([15], Theorems 2.1 and 2.2)**.** *Let $\delta > 0$. Then*

$$\lim_{n\to\infty} n^\delta P_0^V(\sigma > n) = C_1 \in (0, \infty); \tag{4}$$

$$\lim_{n\to\infty} n^\delta P_0^V\left(\Sigma > n^2\right) = C_2 \in (0, \infty). \tag{5}$$

We shall need to consider V over many lifetimes. Let $\sigma_0 = 0$, $\Sigma_0 = 0$,

$$\sigma_i = \inf\{k > \sigma_{i-1} : V_k = 0\}, \quad \Sigma_i = \sum_{k=\sigma_{i-1}}^{\sigma_i - 1} V_k, \quad i \in \mathbb{N}. \tag{6}$$

Then $(\sigma_i - \sigma_{i-1}, \Sigma_i)_{i\in\mathbb{N}}$ are i.i.d. under P_0^V, $(\sigma_i - \sigma_{i-1}, \Sigma_i) \overset{\mathrm{d}}{=} (\sigma, \Sigma)$, $i \in \mathbb{N}$.

3. Non-boundary case: two useful lemmas

Let $\delta \in [0, 1)$. First of all, we show that by time n the walker consumes almost all the drift between I_n and S_n.

Lemma 4. *Assume that $\delta \in [0, 1)$. Given $\gamma_1 > \delta$, there exist $\gamma_2 > 0$ and $\theta \in (0, 1)$ such that for all $1 \leq \ell \leq n$*

$$P_0\left(\sum_{m=n-\ell}^{n-1} \mathbb{1}_{\{L_m(T_n) < M\}} > \ell^{\gamma_1}\right) \leq \theta^{\ell^{\gamma_2}} \quad \textit{and} \tag{7}$$

$$P_0\left(\sum_{m=-(n-1)}^{-(n-\ell)} \mathbb{1}_{\{L_m(T_{-n}) < M\}} > \ell^{\gamma_1}\right) \leq \theta^{\ell^{\gamma_2}}. \tag{8}$$

Proof. We shall start with (7) and use the connection with branching processes. Since the event we are interested in depends only on the environment and the behavior of the walk on $\{n - \ell, n - \ell + 1, \dots\}$, we may assume without loss of generality that the process starts at $n - \ell$ and, thus, by translation invariance consider only the case $\ell = n$.

Let $L^V_k(n)=\sum_{j=0}^n \mathbb{1}_{\{V_j=k\}}$. We have

(9)
$$P_0\left(\sum_{m=0}^{n-1}\mathbb{1}_{\{L_m(T_n)<M\}}>n^{\gamma_1}\right)\le P_0\left(\sum_{m=0}^{n}\mathbb{1}_{\{D_{n,m}<M\}}>n^{\gamma_1}\right)=P_0^V\left(\sum_{m=0}^{n}\mathbb{1}_{\{V_m<M\}}>n^{\gamma_1}\right)$$
$$\le M\max_{0\le k<M}P_0^V\left(\sum_{m=0}^{n}\mathbb{1}_{\{V_m=k\}}>\frac{n^{\gamma_1}}{M}\right)=M\max_{0\le k<M}P_0^V\left(L^V_k(n)>\frac{n^{\gamma_1}}{M}\right).$$

At first, consider the case $\delta\in(0,1)$. Let $k=0$. Then (see (4) and (6)) for all sufficiently large n we get

$$P_0^V\left(L^V_0(n)>\frac{n^{\gamma_1}}{M}\right)\le\prod_{i=1}^{[n^{\gamma_1}/M]}P_0^V\left(\sigma_i-\sigma_{i-1}\le n\right)\le\left(1-\frac{C_1}{2n^\delta}\right)^{[n^{\gamma_1}/M]}.$$

Since $\gamma_1>\delta$, this implies the desired estimate for $k=0$.

Let $k\in\{1,2,\dots,M-1\}$. Then for any $\varepsilon>0$

$$P_0^V\left(L^V_k(n)>\frac{n^{\gamma_1}}{M}\right)=$$
$$P_0^V\left(L^V_k(n)>\frac{n^{\gamma_1}}{M},L^V_0(n)>\frac{\varepsilon n^{\gamma_1}}{2M}\right)+P_0^V\left(L^V_k(n)>\frac{n^{\gamma_1}}{M},L^V_0(n)\le\frac{\varepsilon n^{\gamma_1}}{2M}\right)$$
$$\le P_0^V\left(L^V_0(n)>\frac{\varepsilon n^{\gamma_1}}{2M}\right)+P_0^V\left(L^V_0(n)\le\frac{\varepsilon n^{\gamma_1}}{2M}\,\Big|\,L^V_k(n)>\frac{n^{\gamma_1}}{M}\right).$$

We only need to estimate the last term. Notice that by (A2) there is $\varepsilon>0$ such that $P_0^V(V_{j+1}=0\,|\,V_j=k)\ge\varepsilon$ for all $k\in\{1,2,\dots,M-1\}$ and $j\in\mathbb{N}$. Therefore, the last term is bounded above by the probability that in at least $[n^{\gamma_1}/M]$ independent Bernoulli trials with probability of success in each trial of at least ε there are at most $[\varepsilon n^{\gamma_1}/(2M)]$ successes. This probability is bounded above by $\exp(-cn^{\gamma_1}/M)$ for some positive $c=c(\varepsilon)$. This completes the proof of (7) for $\delta>0$.

If $\delta=0$ we modify the environment by increasing slightly the drift (to the right) in the first cookie at each site. Let $\widetilde{V}$ be the branching process corresponding to the modified environment. There is a natural coupling between V and $\widetilde{V}$ such that $\widetilde{V}_j\le V_j$, $j\in\{0,1,\dots,n\}$. Accordingly,

$$\sum_{j=0}^n\mathbb{1}_{\{V_j<M\}}\le\sum_{j=0}^n\mathbb{1}_{\{\widetilde{V}_j<M\}},$$

and (7) for $\delta=0$ follows from the result for $\delta>0$ and the second line of (9).

Next after replacing X by $-X$ proving (8) reduces to proving (7) for $\delta\le0$ and $\gamma_1>0$. As above, the result for $\delta\le0$ can be deduced from the result for $\delta\in(0,\gamma_1)$ by coupling of the corresponding branching processes. □

Next we show that $\sqrt{n}$ is a correct scaling in Theorem 1.

Lemma 5. *Assume that $\delta \in [0,1)$. There exists $\theta \in (0,1)$ such that for all $L > 0$, $\ell \in \mathbb{N} \cup \{0\}$, and $n \in \mathbb{N}$*

$$P_0\left(T_{\ell+n} - T_\ell \le \frac{n^2}{L}\right) \le \theta^{\sqrt{L}} \quad \text{and} \quad P_0\left(T_{-\ell-n} - T_{-\ell} \le \frac{n^2}{L}\right) \le \theta^{\sqrt{L}}.$$

Proof. We shall prove the first inequality for $\delta \in (0,1)$. The case $\delta = 0$ and the second inequality are handled in exactly the same way as in the proof of Lemma 4.

Since $T_{n+\ell} - T_\ell \ge \sum_{k=\ell}^{n+\ell} D_{n+\ell,k} \overset{\mathrm{d}}{=} \sum_{j=0}^n V_j$, it is enough to show that

$$P_0^V\left(\sum_{j=0}^n V_j \le \frac{n^2}{L}\right) \le \theta^{\sqrt{L}}.$$

Notice that by the Markov property and the stochastic monotonicity of V in the initial number of particles

$$(10)\quad P_0^V\left(\sum_{j=0}^{m+k} V_j \le n\right) \le P_0^V\left(\sum_{j=m+1}^{m+k} V_j \le n \,\Big|\, \sum_{j=0}^{m} V_j \le n\right) P_0^V\left(\sum_{j=0}^{m} V_j \le n\right) \le P_0^V\left(\sum_{j=0}^{k} V_j \le n\right) P_0^V\left(\sum_{j=0}^{m} V_j \le n\right).$$

Suppose that we can show that there exist $K, n_0 \in \mathbb{N}$ such that for all $n \ge n_0$

$$(11)\quad P_0^V\left(\sum_{j=0}^{Kn} V_j \le n^2\right) \le \frac{1}{2}.$$

Then using (10) and (11) we get that for all $L > 4K^2$ and $n \ge \sqrt{L} n_0$

$$P_0^V\left(\sum_{j=0}^{n} V_j \le \frac{n^2}{L}\right) \le \left(P_0^V\left(\sum_{j=0}^{[2Kn/\sqrt{L}]} V_j \le \frac{n^2}{L}\right)\right)^{[\sqrt{L}/(2K)]} \le \left(P_0^V\left(\sum_{j=0}^{2K[n/\sqrt{L}]} V_j \le 4\left[\frac{n}{\sqrt{L}}\right]^2\right)\right)^{[\sqrt{L}/(2K)]} \le \left(\left(\frac{1}{2}\right)^{1/(4K)}\right)^{\sqrt{L}},$$

and we are done.

To prove (11), we observe that due to (4) the sequence $\sigma_m/m^{1/\delta}$, $m \in \mathbb{N}$, has a limiting distribution ([10], Theorem 3.7.2) and, thus, if K is large then $P_0(\sigma_{[(\sqrt{K}n)^\delta]} > Kn) \le 1/4$

for all large enough n. We conclude that there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

$$P_0^V\left(\sum_{j=0}^{Kn} V_j \leq n^2\right) \leq \frac{1}{4} + P_0^V\left(\sum_{j=0}^{\sigma_{[(\sqrt{K}n)^\delta]}} V_j \leq n^2, \sigma_{[(\sqrt{K}n)^\delta]} \leq Kn\right)$$

$$\leq \frac{1}{4} + P_0^V\left(\sum_{i=1}^{[(\sqrt{K}n)^\delta]} \Sigma_i \leq n^2\right) \leq \frac{1}{4} + \prod_{i=1}^{[(\sqrt{K}n)^\delta]} P_0^V\left(\Sigma_i \leq n^2\right) \overset{(5)}{\leq} \frac{1}{4} + \left(1 - \frac{C_2}{2n^\delta}\right)^{[(\sqrt{K}n)^\delta]}.$$

This immediately gives (11) if K is chosen sufficiently large. □

4. Non-boundary case: Proof of Theorem 1

Let $\Delta_n = X_{n+1} - X_n$ and

$$B_n = \sum_{k=0}^{n-1} \left(\Delta_k - E_{0,\omega}(\Delta_k|\mathcal{F}_k)\right), \quad C_n = \sum_{k=0}^{n-1} E_{0,\omega}(\Delta_k|\mathcal{F}_k). \tag{12}$$

Then $X_n = B_n + C_n$, where $(B_n),\ n \geq 0$ is a martingale. Define

$$X^{(n)}(t) := \frac{X_{[nt]}}{\sqrt{n}}, \quad B^{(n)}(t) := \frac{B_{[nt]}}{\sqrt{n}}, \quad C^{(n)}(t) := \frac{C_{[nt]}}{\sqrt{n}}, \quad t \geq 0,\ n \in \mathbb{N}.$$

Theorem 1 is an easy consequence of the following three lemmas, the first of which holds for the quenched and the last two for the averaged measures.

Lemma 6. *Let B be a standard Brownian motion. Then $B^{(n)} \overset{J_1}{\Rightarrow} B$ as $n \to \infty$ for $\mathbb{P}$-a.e. ω.*

Lemma 7. *For each $t \geq 0$ and $\varepsilon > 0$*

$$P_0\left(\sup_{k \leq nt} \frac{|C_k - \delta R_k|}{\sqrt{n}} > \varepsilon\right) \to 0.$$

Lemma 8. *The sequence $X^{(n)}$, $n \geq 1$, is tight in $D([0,\infty))$. Moreover, if X is a limit point of this sequence and P is the corresponding measure on $D([0,\infty))$ then $P(X \in C([0,\infty))) = 1$.*

Proof of Theorem 1 assuming Lemmas 6–8. Since $X^{(n)}$, $n \geq 1$, is tight and $B^{(n)} \overset{J_1}{\Rightarrow} B$ as $n \to \infty$, the sequence $C^{(n)}$, $n \geq 1$, as the difference of two tight sequences is also tight. We can assume by choosing a subsequence that $X^{(n)} \overset{J_1}{\Rightarrow} X$, where X is continuous by Lemma 8. The mapping $x(\cdot) \mapsto r^x(\cdot) := \sup_{s \leq \cdot} x(s) - \inf_{s \leq \cdot} x(s)$ is continuous on $C([0,t])$. Therefore, by the continuous mapping theorem

$$r^{X^{(n)}}(\cdot) = \frac{R_{[n\cdot]}}{\sqrt{n}} \overset{J_1}{\Rightarrow} r^X(\cdot). \tag{13}$$

The tightness of $C^{(n)}$, $n \geq 1$, (13), Lemma 7, and the "convergence together" result ([6], Theorem 3.1) imply that $C^{(n)} \overset{J_1}{\Rightarrow} \delta r^X$ as $n \to \infty$.

Now we have a vector-valued sequence of processes $(X^{(n)}, B^{(n)}, C^{(n)})$, $n \geq 1$, that is tight. Therefore, along a subsequence, this 3-dimensional process converges to $(X, B, \delta r^X)$. Since $X^{(n)} = B^{(n)} + C^{(n)}$, we get that $X = B + \delta r^X$. □

We shall conclude this section with proofs of Lemmas 6–8.

Proof of Lemma 6. We shall use the functional limit theorem for martingale differences ([6], Theorem 18.2). Let $\xi_{nk} = n^{-1/2}(\Delta_{k-1} - E_{0,\omega}(\Delta_{k-1}|\mathcal{F}_{k-1}))$, $k, n \in \mathbb{N}$. Due to rescaling and the fact that ERW moves in unit steps, it is obvious that the Lindeberg condition,

$$\sum_{k\leq nt} E_{0,\omega}[\xi_{nk}^2 \mathbb{1}_{\{|\xi_{nk}|\geq\varepsilon\}}] \to 0 \quad \text{as } n\to\infty \text{ for every } t\geq 0 \text{ and } \varepsilon>0,$$

is satisfied. Thus, we just have to show the convergence of the quadratic variation process, i.e. for $\mathbb{P}$-a.e. ω for each $t \geq 0$

$$\sum_{k\leq nt} E_{0,\omega}(\xi_{nk}^2|\mathcal{F}_{k-1}) = \frac{[nt]}{n} - \frac{1}{n}\sum_{k\leq nt}\left(E_{0,\omega}(\Delta_{k-1}|\mathcal{F}_{k-1})\right)^2 \Rightarrow t \tag{14}$$

as $n \to \infty$. Since

$$0 \leq \frac{1}{n}\sum_{k\leq nt}\left(E_{0,\omega}(\Delta_{k-1}|\mathcal{F}_{k-1})\right)^2 \leq \frac{M}{n} R_{[nt]},$$

it is enough to prove that $P_{0,\omega}(R_{[nt]} > \varepsilon n) \to 0$ a.s. for each $\varepsilon > 0$. We have

$$P_{0,\omega}(R_{[nt]} > \varepsilon n) \leq P_{0,\omega}(T_{[\varepsilon n/3]} \leq nt) + P_{0,\omega}(T_{-[\varepsilon n/3]} \leq nt) =: f_{n,\varepsilon}(\omega, t).$$

By Fubini's theorem and Lemma 5,

$$\mathbb{E}\left(\sum_{n=1}^{\infty} f_{n,\varepsilon}(\omega,t)\right) = \sum_{n=1}^{\infty}\mathbb{E} f_{n,\varepsilon}(\omega,t) = \sum_{n=1}^{\infty}\left(P_0\left(T_{[\varepsilon n/3]}\leq nt\right) + P_0\left(T_{-[\varepsilon n/3]}\leq nt\right)\right) < \infty.$$

This implies that $f_{n,\varepsilon}(\omega,t) \to 0$ a.s. as $n \to \infty$ and completes the proof. □

Proof of Lemma 7. Let $d_m = \sum_{i=1}^{M}(2\omega_m(i) - 1)$ be the total drift stored at site m, $m \in \mathbb{Z}$. Then

$$C_k - \delta R_k = \sum_{m=I_k}^{S_k}(d_m - \delta) - \sum_{m=I_k}^{S_k}\mathbb{1}_{\{L_m(k)<M\}}\sum_{j=L_m(k)+1}^{M}(2\omega_m(j)-1).$$

By Lemma 5, given $\nu > 0$, we can choose K sufficiently large so that $P_0(R_{[nt]} > K\sqrt{n}) < \nu/2$ for all $n \in \mathbb{N}$. We have

$$P_0\left(\sup_{k\leq nt}\frac{|C_k - \delta R_k|}{\sqrt{n}} > \varepsilon\right) \leq P_0\left(\max_{k\leq nt}\frac{\left|\sum_{m=I_k}^{S_k}(d_m-\delta)\right|}{R_k}\frac{R_k}{\sqrt{n}} > \frac{\varepsilon}{2}, \frac{R_{[nt]}}{\sqrt{n}} \leq K\right) \tag{15}$$
$$+ P_0\left(\frac{M}{\sqrt{n}}\sum_{m=I_{[nt]}}^{S_{[nt]}}\mathbb{1}_{\{L_m([nt])<M\}} > \frac{\varepsilon}{2}, \frac{R_{[nt]}}{\sqrt{n}} \leq K\right) + \frac{\nu}{2}.$$

By the strong law of large numbers $\lim_{(a+b)\to\infty}(a+b)^{-1}\sum_{m=-a}^{b}(d_m-\delta) = 0$ ($\mathbb{P}$-a.s.). Therefore, for $\mathbb{P}$-a.e. ω there is an $r(\omega)\in\mathbb{N}$ such that $R_k^{-1}\left|\sum_{m=I_k}^{S_k}(d_m-\delta)\right| \le \varepsilon/(2K)$ whenever $R_k \ge r(\omega)$, and the first term in the right-hand side of (15) does not exceed

$$P_0\left(\frac{2(M+1)r(\omega)}{\sqrt{n}} > \frac{\varepsilon}{2}, \frac{R_{[nt]}}{\sqrt{n}} \le K\right) \le \mathbb{E}\left(P_{0,\omega}\left(r(\omega) > \frac{\varepsilon\sqrt{n}}{4(M+1)}\right)\right) \to 0 \text{ as } n\to\infty.$$

Thus, we only need to estimate the second term in the right-hand side of (15).

Divide the interval $[I_{[nt]}, S_{[nt]}]$ into subintervals of length $n^{1/4}$. By Lemma 4, given $\gamma_1 \in (\delta, 1)$, with probability at least $1-\theta^{n^{\gamma_2/4}}Kn^{1/4}$ all subintervals except the two extreme ones have at most $n^{\gamma_1/4}$ points which are visited less than M times. Hence, for n sufficiently large

$$P_0\left(\frac{M}{\sqrt{n}}\sum_{m=I_{[nt]}}^{S_{[nt]}}\mathbb{1}_{\{L_m([nt])<M\}} > \frac{\varepsilon}{2}, \frac{R_{[nt]}}{\sqrt{n}} \le K\right) \le$$

$$P_0\left(\sum_{m=I_{[nt]}}^{S_{[nt]}}\mathbb{1}_{\{L_m([nt])<M\}} > n^{(1+\gamma_1)/4} + 2n^{1/4}, \frac{R_{[nt]}}{\sqrt{n}} \le K\right) \le \theta^{n^{\gamma_2/4}}Kn^{1/4},$$

and the proof is complete. □

Proof of Lemma 8. The idea of the proof is the following. If $X^{(n)}$ has large fluctuations then either $B^{(n)}$ has large fluctuations or $C^{(n)}$ has large fluctuations. $B^{(n)}$ is unlikely to have large fluctuations, since it converges to the Brownian motion. By Lemma 4, C_n can have large fluctuations only if S_n increases or I_n decreases. However by Lemma 5 neither I_n nor S_n can change too quickly. Let us give the details.

To prove both statements of Lemma 8 it is enough to show that there exists $C_3, \alpha > 0$ such that for all $\ell\in\mathbb{N}$ and sufficiently large n, $n > 2^\ell$,

$$P_0(\cup_{k<2^\ell}\Omega_{n,k,\ell}) \le C_3 2^{-\alpha\ell}, \tag{16}$$

where

$$\Omega_{n,k,\ell} = \left\{\left|X^{(n)}\left(\frac{k+1}{2^\ell}\right) - X^{(n)}\left(\frac{k}{2^\ell}\right)\right| > 2^{-\ell/8}\right\}$$

(see e.g. the last paragraph in the proof of Lemma 1 in [12], Chapter III, Section 5).

Let

$$m_1 := \left[\frac{kn}{2^\ell}\right],\ m_2 := \left[\frac{(k+1)n}{2^\ell}\right],\ J := \frac{1}{4}n^{1/2}2^{-\ell/8}. \tag{17}$$

Then

$$\Omega_{n,k,\ell} = \{|X_{m_2} - X_{m_1}| > 4J\} \subset \Omega^B_{n,k,\ell}\cup\Omega^C_{n,k,\ell},$$

where

$$\Omega^B_{n,k,\ell} = \{|B_\tau - B_{m_1}| > J, \tau\le m_2\},\quad \Omega^C_{n,k,\ell} = \{|C_\tau - C_{m_1}| > 3J, \tau \le m_2\},$$

$\tau := \inf\{m > m_1 : |X_m - X_{m_1}| > 4J\}$ and B_n and C_n are defined in (12).

Since $(B_{j+m_1}-B_{m_1})$, $j\ge 0$, is a martingale, whose quadratic variation grows at most linearly, the maximal inequality and Burkholder-Davis-Gundy inequality ([13], Theorem 2.11 with $p=4$) imply that

$$P_{0,\omega}(\Omega^B_{n,k,\ell})\le P_{0,\omega}\left(\max_{m_1\le j\le m_2}|B_j-B_{m_1}|>J\right)\le \frac{C(m_2-m_1)^2}{J^4}\le C'2^{-3\ell/2}.$$

Hence, $P_0\left(\cup_{k<2^\ell}\Omega^B_{n,k,\ell}\right)\le C'2^{-\ell/2}$.

To control $P_0(\Omega^C_{n,k,\ell})$ consider the following intervals:

$$A_1=(-\infty,I_{m_1})\cap\Gamma,\quad A_2=[I_{m_1},S_{m_1}]\cap\Gamma,\quad A_3=(S_{m_1},\infty)\cap\Gamma,$$

where $\Gamma=[X_{m_1}-4J,X_{m_1}+4J]$. Then

$$\Omega^C_{n,k,\ell}\subset\bigcup_{s=1}^{3}\left\{\sum_{j=m_1}^{\tau-1}|E_{0,\omega}(\Delta_j\,|\,\mathcal{F}_j)|\mathbb{1}_{\{X_j\in A_s\}}>J,\tau\le m_2\right\}$$
$$\subset\bigcup_{s=1}^{3}\left\{\sum_{j=m_1}^{m_2-1}|E_{0,\omega}(\Delta_j\,|\,\mathcal{F}_j)|\mathbb{1}_{\{X_j\in A_s\}}>J\right\}=:\bigcup_{s=1}^{3}\Omega^C_{n,k,\ell,s}.$$

To estimate $P_0(\Omega^C_{n,k,\ell,3})$ note that to accumulate a drift larger than J the walk should visit at least $[J/M]$ distinct sites, i.e.

$$\Omega^C_{n,k,\ell,3}\subset\{T_{S_{m_1}+[J/M]}-T_{S_{m_1}+1}\le m_2-m_1\}.$$

Let $\bar J=[J/(2M)]$ and $\bar\ell=\ell/8$. There exists an $m\in\mathbb{N}$ such that $S_{m_1}+1\le m\bar J\le(m+1)\bar J\le S_{m_1}+[J/M]$. Using Lemma 5, we can find $K>1$ such that $P_0(S_n>K\sqrt{n})<2^{-\bar\ell}$ for all sufficiently large n. Therefore,

$$P_0(\Omega^C_{n,k,\ell,3})\le 2^{-\bar\ell}+P_0\left(\cup_{m<2^{\bar\ell+3}MK}\Omega^\dagger_{n,m,\ell},S_n\le K\sqrt{n}\right),$$

where $\Omega^\dagger_{n,m,\ell}=\left\{T_{(m+1)\bar J}-T_{m\bar J}\le m_2-m_1\right\}$. Since $m_2-m_1\le C\bar J^2/2^{6\bar\ell}$ for some constant $C>0$, Lemma 5 implies that there is $\hat\theta<1$ such that and all sufficiently large n

$$P_0\left(\cup_{m<2^{\bar\ell+3}KM}\Omega^\dagger_{n,m,\ell}\right)\le\sum_{m<2^{\bar\ell+3}KM}P_0\left(\Omega^\dagger_{n,m,\ell}\right)\le 2^{\bar\ell+3}KM\hat\theta^{2^{3\ell}}<C''2^{-\ell}.$$

$P_0(\cup_{k<2^\ell}\Omega^C_{n,k,\ell,1})$ is estimated in the same way.

We consider now A_2, which is a random subinterval of $[-m_1,m_1]$ and, on $\Omega^C_{n,k,\ell,2}$, has length between J/M and $8J$. To estimate $P_0(\Omega^C_{n,k,\ell,2})$ we notice that by Lemma 4, outside of an event of exponentially small (in J^{γ_2}) probability, the number of cookies that are left in A_2 at time m_1 does not exceed CJ^{γ_1}, where $\gamma_1<1$. Even if the walker consumes all cookies in that interval, it can not build up a drift of size $J\gg CJ^{\gamma_1}$ (for J large). With this idea in mind, we turn now to a formal proof.

As we noted above, on $\Omega^C_{n,k,\ell,2}$, we have $A_2\in\mathcal{I}$, where $\mathcal{I}$ denotes the set of all intervals of the form

$$[a,b]\text{ such that }a,b\in\mathbb{Z},\ -m_1\le a<b\le m_1,\ J/M\le b-a\le 8J.$$

The cardinality of $\mathcal{I}$ does not exceed $16m_1J \le Cn^{3/2}$. Therefore,

$$P_0(\Omega^C_{n,k,\ell,2}) \le Cn^{3/2} \max_{A\in\mathcal{I}} P_0\left(\sum_{j=m_1}^{m_2-1} |E_{0,\omega}(\Delta_j \,|\, \mathcal{F}_j)| \mathbb{1}_{\{X_j\in I\}} > J, A_2 = A\right). \tag{18}$$

By the definition of A_2, the walk necessarily crosses the interval A_2 by the time m_1. The leftover drift in A_2 is at most M times the number of sites in A_2, which still have at least one cookie. Writing A as $[a,b]$, $a,b\in\mathbb{Z}$, $a<b$, we can estimate the last probability by

$$P_0\left(M\sum_{m=a}^{b} \mathbb{1}_{\{L_m(T_a\vee T_b)<M\}} > J\right) = P_0\left(\sum_{m=a}^{b} \mathbb{1}_{\{L_m(T_a\vee T_b)<M\}} > J/M\right).$$

If $a\ge 0$ we can apply Lemma 4 and get that for all sufficiently large n (such that $(8J)^{\gamma_1} \le J/M$)

$$\begin{aligned} P_0\left(\sum_{m=a}^{b} \mathbb{1}_{\{L_m(T_a\vee T_b)<M\}} > J/M\right) &\le P_0\left(\sum_{m=a}^{b} \mathbb{1}_{\{L_m(T_b)<M\}} > (b-a)^{\gamma_1}\right) \\ &\le \theta^{(b-a)^{\gamma_2}} \le \theta^{(J/M)^{\gamma_2}}. \end{aligned} \tag{19}$$

The case $b\le 0$ is similar. Finally, consider the case $a<0<b$. Then

$$\begin{aligned} P_0\left(\sum_{m=a}^{b} \mathbb{1}_{\{L_m(T_a\vee T_b)<M\}} > J/M\right) &\le P_0\left(\sum_{m=a}^{0} \mathbb{1}_{\{L_m(T_a)<M\}} > J/(2M)\right) \\ &+ P_0\left(\sum_{m=0}^{b} \mathbb{1}_{\{L_m(T_b)<M\}} > J/(2M)\right). \end{aligned} \tag{20}$$

If $b\le J/(2M)$ then the last term in (20) is 0. But for $J/(2M) < b \le 8J$ we have that $b^{\gamma_1} \le J/(2M)$ for all sufficiently large J. Lemma 4 implies that

$$P_0\left(\sum_{m=0}^{b} \mathbb{1}_{\{L_m(T_b)<M\}} > J/(2M)\right) \le P_0\left(\sum_{m=0}^{b} \mathbb{1}_{\{L_m(T_b)<M\}} > b^{\gamma_1}\right) \le \theta^{b^{\gamma_2}} \le \theta^{(J/(2M))^{\gamma_2}}.$$

The first term in the right-hand side of (20) is estimated in the same way. We conclude that for some constant C and all sufficiently large n

$$P_0(\cup_{k<2^\ell}\Omega^C_{n,k,\ell,2}) \le Cn^{3/2} 2^\ell \theta^{(J/(2M))^{\gamma_2}} < 2^{-\ell}.$$

This completes the proof of (16) establishing Lemma 8. □

5. Boundary case: Proof of Theorem 2.

Let $\delta = 1$. For $t\ge 0$ and $n\ge 2$ set

$$T^{(n)}(x) := \frac{T_{[nx]}}{n^2/\log^2 n}, \quad X^{(n)}(t) := \frac{X_{[nt]}}{\sqrt{n}\log n}, \quad S^{(n)}(t) := \frac{S_{[nt]}}{\sqrt{n}\log n}.$$

Let $\Sigma_j,\ j\ge 0$ be i.i.d. positive integer-valued random variables defined in (6). They satisfy (5) with $\delta=1$ and by [11, Chapter 9, Section 6] for some constant $a>0$

$$\frac{\sum_{j=0}^{[n\cdot]}\Sigma_j}{n^2}\overset{J_1}{\Rightarrow} aH(\cdot)\quad \text{as } n\to\infty, \tag{21}$$

where $H:=(H(x)),\ x\ge 0$, is a stable subordinator with index 1/2. More precisely,

$$H(x)=\inf\{t\ge 0:\ B(t)=x\}. \tag{22}$$

We shall need the following two lemmas.

Lemma 9. *The finite dimensional distributions of $T^{(n)}$ converge to those of cH, where $c>0$ is a constant and H is given by (22).*

Lemma 10. *For every $\varepsilon>0$, $T>0$*

$$\lim_{n\to\infty} P_0\left(\sup_{0\le t\le T}(S^{(n)}(t)-X^{(n)}(t))>\varepsilon\right)=0.$$

Theorem 2 is an easy consequence of these lemmas.

Proof of Theorem 2. Lemma 9 implies that the finite dimensional distributions of the process $S^{(n)}$ converge to those of DB^*, where $D>0$ is a constant. Since the trajectories of $S^{(n)}$ are monotone and the limiting process B^* is continuous, we conclude that $S^{(n)}$ converges weakly to DB^* in the (locally) uniform topology (see [1], Corollary 1.3 and Remark (e) on p. 588). Finally, by Lemma 10 for each $T>0$

$$\sup_{0\le t\le T}(S^{(n)}(t)-X^{(n)}(t))\to 0$$

in P_0 probability. By the "converging together" theorem ([6, Theorem 3.1]) we conclude that $X^{(n)}$ converges weakly to DB^* in the (locally) uniform topology, and, thus, in J_1. □

Proof of Lemma 9. Let $k\in\mathbb{N}$ and $0=x_0<x_1<\dots<x_k$. We have to show that for any $0=t_0<t_1<t_2<\dots<t_k$

$$\begin{aligned}P_0(T^{(n)}(x_k)-T^{(n)}(x_i)\le t_{k-i},\ \forall i=0,1,2,\dots,k-1)\\ \to P(T(x_k)-T(x_i)\le t_{k-i},\ \forall i=0,1,\dots,k-1),\ \text{as } n\to\infty,\end{aligned}$$

where $T(\cdot)=cH(\cdot)$ for some $c>0$.

At time $T_{[nx_k]}$ consider the structure of the corresponding branching process as we look back from $[nx_k]$. Notice that $D_{[nx_i],j}\le D_{[nx_k],j}$ for $i\le k$ and all j. This simple observation will allow us to get bounds on $T_{[nx_i]}$, $i=1,2,\dots,k-1$, in terms of the structure of downcrossings at time $T_{[nx_k]}$. This means that we can use the same copy of the branching process V to draw conclusions about all hitting times $T_{[nx_i]}$, $i=1,2,\dots,k$.

We shall use notation (6) and let $N^{(0)}=0$,

$$N^{(k-i)}=\min\{m\in\mathbb{N}:\ \sigma_m\ge[nx_k]-[nx_i]\},\quad i=0,1,2,\dots,k-1.$$

Since

$$2\sum_{j=1}^{N^{(k-i)}-1}\Sigma_j \le T_{[nx_k]}-T_{[nx_i]} \le nx_k - nx_i + 2\sum_{j=1}^{N^{(k-i)}}\Sigma_j,$$

we have

$$\begin{aligned}(23)\quad & P_0(T^{(n)}(x_k)-T^{(n)}(x_i)\le t_{k-i},\ \forall i=0,1,2,\dots,k-1)\\ & \le P\left(2\sum_{j=1}^{N^{(k-i)}-1}\Sigma_j \le n^2 t_{k-i}/\log^2 n,\ \forall i=0,1,2,\dots,k-1\right)\end{aligned}$$

and

$$\begin{aligned}(24)\quad & P_0(T^{(n)}(x_k)-T^{(n)}(x_i)\le t_{k-i},\ \forall i=0,1,2,\dots,k-1)\\ & \ge P\left([nx_k]-[nx_i]+2\sum_{j=1}^{N^{(k-i)}}\Sigma_j \le n^2 t_{k-i}/\log^2 n,\ \forall i=0,1,2,\dots,k-1\right).\end{aligned}$$

Next we provide some control on $N^{(k-i)}$, $i=0,1,\dots,k-1$, and on the maximal lifetime over $[nx_k]$ generations. Theorem 3 and [10, Theorem 3.7.2] imply that $\sigma_n/(n\log n)\Rightarrow b^{-1}$ for some positive constant b. From this it is easily seen that

$$\frac{\min\{m\in\mathbb{N}:\ \sigma_m>n\}}{nb/\log n}\Rightarrow 1 \quad \text{as } n\to\infty. \tag{25}$$

Recalling our definition of $N^{(k-i)}$ we get that for every $\varepsilon,\nu>0$ there is n_0 such that for all $n\ge n_0$

$$P\left(1-\nu\le\frac{N^{(k-i)}}{\bar N^{(k-i)}}\le 1+\nu,\ i=0,\dots,k-1\right)>1-\varepsilon,$$

where $\bar N^{(k-i)}=b(x_k-x_i)n/\log n$. In particular, for $C=(1+\nu)bx_k$ we have that

$$P\left(N^{(k)}\le\frac{Cn}{\log n}\right)>1-\varepsilon.$$

Define $\lambda_n=(\log n)^{-1/2}$ (any sequence λ_n, $n\in\mathbb{N}$, such that $\lambda_n\to 0$ and $\lambda_n\log n\to\infty$ will work) and notice that by Theorem 3 there is n_1 such that for all $n\ge n_1$

$$P\left(\max_{1\le i\le Cn/\log n}(\sigma_i-\sigma_{i-1})\le n\lambda_n\right)\ge\left(1-\frac{2C_1}{n\lambda_n}\right)^{Cn/\log n}>1-\varepsilon.$$

Thus, on a set Ω_ε of measure at least $1-2\varepsilon$ for all $n\ge n_0\vee n_1$ the number of lifetimes of the branching process V covering $[nx_k]-[nx_i]$ generations, $i=0,1,2,\dots,k-1$, is well controlled and the maximal lifetime over $[nx_k]$ generations does not exceed $n\lambda_n$. In particular, on Ω_ε, the number of lifetimes in any interval $([nx_i],[nx_{i+1}])$, $i=0,1,\dots,k-1$, goes to infinity as $n\to\infty$.

Finally, on Ω_ε we get from (23) and (21) that

$$\begin{aligned}
&P_0(T^{(n)}(x_k) - T^{(n)}(x_i) \le t_{k-i},\ \forall i = 0,1,2,\dots,k-1) \\
&\qquad \le P\left(2 \sum_{j=1}^{(1-\nu)\bar N^{(k-i)}-1} \Sigma_j \le n^2 t_{k-i}/\log^2 n,\ \forall i = 0,1,2,\dots,k-1\right) \\
&\qquad = P\left(\frac{\sum_{j=1}^{(1-\nu)\bar N^{(k-i)}-1}\Sigma_j}{((1-\nu)n/\log n)^2} \le \frac{t_{k-i}}{2(1-\nu)^2},\ \forall i = 0,1,2,\dots,k-1\right) \\
&\qquad \to P(aH(b(x_k - x_i)) \le (1-\nu)^{-2} t_{k-i}/2\ \forall i = 0,1,2,\dots,k-1) \\
&\qquad\quad = P(2ab^2(H(x_k) - H(x_i)) \le t_{k-i}(1-\nu)^{-2})\ \forall i = 0,1,2,\dots,k-1).
\end{aligned}$$

The lower bound is shown starting from (24) in exactly the same way. Letting $\nu \to 0$ and then $\varepsilon \to 0$ we obtain the statement of the lemma with $T(\cdot) = 2ab^2 H(\cdot) =: cH(\cdot)$. □

Proof of Lemma 10. Without loss of generality we can consider $t \in [0,1]$. Fix some $\nu > 0$. We have

$$(26)\quad P_0\left(\sup_{0\le t\le 1}(S^{(n)}(t) - X^{(n)}(t)) > \varepsilon\right) \le P_0(S_n \ge K\sqrt{n}\ln n) + P_0\left(\max_{0\le m\le n}(S_m - X_m) > \varepsilon\sqrt{n}\ln n, S_n < K\sqrt{n}\ln n\right).$$

By Lemma 9 we can find $K > 0$ such that for all large n

$$P_0(S_n \ge K\sqrt{n}\ln n) \le P_0(T_{[K\sqrt{n}\ln n]} \le n) < \nu.$$

To estimate the last term in (26) we shall use properties of the branching process V. Let $N = \min\{m \in \mathbb{N} : \sigma_m > K\sqrt{n}\ln n\}$. Then the last term in (26) is bounded by

$$\begin{aligned}
&P_0^V\left(\max_{i\le N}(\sigma_i - \sigma_{i-1}) \ge \varepsilon\sqrt{n}\ln n\right) \le \\
&\qquad P_0^V(N > C\sqrt{n}) + P_0^V\left(\max_{i\le C\sqrt{n}}(\sigma_i - \sigma_{i-1}) \ge \varepsilon\sqrt{n}\ln n, N \le C\sqrt{n}\right) \overset{(25)}{\le} \\
&\qquad\qquad \nu + P_0^V\left(\max_{i\le C\sqrt{n}}(\sigma_i - \sigma_{i-1}) \ge \varepsilon\sqrt{n}\ln n\right),
\end{aligned}$$

for some large C and all sufficiently large n. Finally, from (4) we conclude that for all large enough n the last probability does not exceed

$$1 - \left(1 - \frac{2C_1}{\varepsilon\sqrt{n}\ln n}\right)^{[C\sqrt{n}]} < \nu.$$

This completes the proof. □

Acknowledgments. The authors are grateful to the Fields Institute for Research in Mathematical Sciences for support and hospitality. D. Dolgopyat was partially supported by the NSF grant DMS 0854982. E. Kosygina was partially supported by a Collaboration Grant for Mathematicians (Simons Foundation) and the PSC CUNY Award # 64603-00 42. The authors also thank the anonymous referee for careful reading of the paper and remarks which helped to improve the exposition.

References

[1] D. Aldous (1989) Stopping times and tightness, II. *Ann. Probab.* **17**, no. 2. 586–595.
[2] A.-L. Basdevant, A. Singh (2008). On the speed of a cookie random walk. *Probab. Theory Related Fields* **141**, no. 3-4, 625–645.
[3] A.-L. Basdevant, A. Singh (2008). Rate of growth of a transient cookie random walk. *Electron. J. Probab.* **13**, no. 26, 811–851.
[4] I. Benjamini, D.B. Wilson (2003). Excited random walk. *Electron. Comm. Probab.* **8**, 86–92
[5] J. Bérard, A. Ramírez (2007). Central limit theorem for the excited random walk in dimension $d \geq 2$. *Elect. Comm. in Probab.* **12**, no. 30, 303–314
[6] P. Billingsley (1999). Convergence of probability measures. Second edition. John Wiley & Sons, Inc., New York, x+277 pp.
[7] L. Chaumont, R. A. Doney (1999). Pathwise uniqueness for perturbed versions of Brownian motion and reflected Brownian motion, *Probab. Theory Related Fields* **113**, no. 4, 519–534.
[8] B. Davis (1996). 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** 2007–2023.
[9] D. Dolgopyat (2011). Central limit theorem for excited random walk in the recurrent regime. *ALEA, Lat. Am. J. Prob. Mat. Stat.* **8**, 259–268.
[10] R. Durrett (2010). Probability: theory and examples. Fourth edition. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, x+428 pp.
[11] I. I. Gikhman, A. V. Skorokhod (1965). Introduction to the theory of stochastic processes. "Nauka", Moscow, 656 pp.. English translation: (1969) W. B. Saunders Co., Philadelphia, Pa.-London-Toronto, Ont. xiii+516 pp.
[12] I. I. Gikhman, A. V. Skorokhod (1974). The theory of stochastic processes. I. Translated from the Russian by S. Kotz. Die Grundlehren der mathematischen Wissenschaften, Band 210. Springer-Verlag, New York-Heidelberg. viii+570 pp.
[13] P. Hall, C. C. Heyde (1980). Martingale limit theory and its application. Academic Press, Inc. New York-London, xii+308 pp.
[14] E. Kosygina, M. Zerner (2008). Positively and negatively excited random walks on integers. with branching processes, *Electron. J. Probab.* **13**, no. 64, 1952–1979.
[15] E. Kosygina, T. Mountford (2011). Limit laws of transient excited random walks on integers. *Ann. Inst. H. Poincaré Probab. Statist.* **47**, no. 2, 575–600.
[16] M. Menshikov, S. Popov, A. Ramirez, M. Vachkovskaya (2011). On a general many-dimensional excited random walk. *Ann. Probab.* to appear.
[17] T. Mountford, L.P.R. Pimentel and G. Valle (2006). On the speed of the one-dimensional excited random walk in the transient regime. *Alea* **2**, 279–296.
[18] M. Perman, W. Werner (1997). Perturbed Brownian motions, *Prob. Theory Related Fields*, **108**, no. 3, 357-383.
[19] J. Peterson (2012). Large deviations and slowdown asymptotics for one-dimensional excited random walks. arXiv:1201.0318.
[20] R. Rastegar, A. Roiterstein (2011). Maximum occupation time of a transient excited random walk on $\mathbb{Z}$. arXiv:1111.1254.

[21] B. Toth (1996). Generalized Ray-Knight theory and limit theorems for self-interacting random walks on $\mathbb{Z}^1$. *Ann. Probab.* **24** 1324–1367.

[22] M.P.W. Zerner (2005). Multi-excited random walks on integers. *Probab. Theory Related Fields* **133**, 98 – 122

[23] M.P.W. Zerner (2006). Recurrence and transience of excited random walks on $\mathbb{Z}^d$ and strips. *Electron. Comm. Probab.* **11**, no. 12, 118 – 128

Department of Mathematics
University of Maryland
4417 Mathematics Building
College Park, MD 20742, USA
dmitry@math.umd.edu

Department of Mathematics
Baruch College, Box B6-230
One Bernard Baruch Way
New York, NY 10010, USA
elena.kosygina@baruch.cuny.edu