QUENCHED LIMIT THEOREMS FOR NEAREST NEIGHBOUR RANDOM WALKS IN 1D RANDOM ENVIRONMENT

D. DOLGOPYAT[1] AND I. GOLDSHEID[2]

Abstract. It is well known that random walks in one dimensional random environment can exhibit subdiffusive behavior due to presence of traps. In this paper we show that the passage times of different traps are asymptotically independent exponential random variables with parameters forming, asymptotically, a Poisson process. This allows us to prove weak quenched limit theorems in the subdiffusive regime where the contribution of traps plays the dominating role.

1. Introduction

Let $\omega = \{p_i\}, i \in \mathbb{Z}$ be an i.i.d. sequence of random variables, $0 < p_i < 1$. The sequence ω is called *environment* (or *random environment*). Let $(\Omega, \mathbf{P})$ be the corresponding probability space with Ω being the set of all environments and $\mathbf{P}$ the probability measure on Ω. The expectation with respect to this measure will be denoted by $\mathbf{E}$. Given an ω we define a random walk $X = \{X_n,\ n \geq 0\}$ on $\mathbb{Z}$ in the environment ω by setting $X_0 = 0$ and

$$\mathbb{P}_\omega(X_{n+1} = X_n{+}1 | X_0 \dots X_n) = p_{X_n} \quad \mathbb{P}_\omega(X_{n+1} = X_n{-}1 | X_0 \dots X_n) = q_{X_n}$$

where $q_n = 1 - p_n$. Denote by $\mathfrak{X} = \{X\}$ the space of all trajectories of the walk starting from zero. A *quenched* (fixed) environment ω thus provides us with a conditional probability measure $\mathbb{P}_\omega$ on $\mathfrak{X}$. The expectation with respect to $\mathbb{P}_\omega$ will be denoted by $\mathbb{E}_\omega$. In turn, these two measures naturally generate the so called *annealed* measure on the direct product $\Omega \times \mathfrak{X}$ which is a semi-direct product $\mathrm{P} := \mathbf{P} \ltimes \mathbb{P}_\omega$. However, with a very slight abuse of notation, $\mathbf{P}$ and $\mathbf{E}$ will also denote the latter measure and the corresponding expectation; the exact meaning of the corresponding probabilities and expectations will always be

[1]Department of Mathematics and Institute of Physical Science and Technology, University of Maryland, College Park, MD, 20742, USA

[2]School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, Great Britain

clear from the context. The term *annealed walk* will be used to discuss properties of the above random walk with respect to the annealed probability.

From now on we assume that

(A) $\mathbf{E}(\ln(p/q)) > 0$.

(B) $\mathbf{E}\left(\frac{q}{p}\right)^s = 1$ for some $s > 0$.

(C) There is a constant ε_0 such that $\varepsilon_0 \leq p_n \leq 1 - \varepsilon_0$ with probability 1.

(D) The support of $\ln(q/p)$ is non-arithmetic.

Assumption (A) implies (see [25]) that $X_n \to \infty$ with probability 1. Assumption (B) means that even though the walker goes to $+\infty$ there are some sites where the drift points in the opposite direction. We note that (A) and (B) are essentially equivalent to each other. Indeed, since $\mathbf{E}\left(\frac{q}{p}\right)^h$ is a convex function of h, (B) implies (A). On the other hand, the existence of finite s in (B) follows from (A) if and only if $\mathbf{P}(q > p) > 0$. It is convenient to have both these conditions on the list for reference purposes.

(C) is a standard ellipticity assumption which prevents the walker from getting stuck at finitely many vertices for a long time. Most of our results can be proved under a weaker assumption, namely $\mathbf{E}\left((\frac{q}{p})^s \ln \frac{q}{p}\right) < \infty$ as in [15]. However, this would lead to more technical, longer, less transparent proofs; also the estimates of some remainders (see e. g. Theorem 2) would become weaker.

(D) is a technical assumption which we don't use in our proofs but which is used in the proof of Lemma 3.6 borrowed from [13]. It is satisfied by a generic distribution of p_n.

We will be mostly interested in the case $s \in (0, 2]$ which implies that the annealed distribution of X_n does not satisfy the standard Central Limit Theorem ([15]). Since X_n is transient it looks monotonically increasing on a large scale and hence it makes sense to study the hitting time $\tilde{T}_N := \min(n : X_n = N)$ which can roughly be viewed as the inverse function of X_n. This approach was used already in the pioneering papers [25] and [15]. In particular, in [15] the annealed behavior of X_n was derived from that of $\tilde{T}_N$. The latter is described by the following

Theorem 1. *([15]) The annealed random walk X has the following properties:*

(a) If $s < 1$ then the distribution of $\frac{\tilde{T}_N}{N^{1/s}}$ converges to a stable law with index s.

(b) If $1 < s < 2$ *then there is a constant* u *such that the distribution of* $\frac{\tilde{T}_N - Nu}{N^{1/s}}$ *converges to a stable law with index* s*.*

(c) If $s > 2$ *then there is a constant* u *such that the distribution of* $\frac{\tilde{T}_N - Nu}{N^{1/2}}$ *converges to a normal distribution.*

(d) If $s = 1$ *then there is a sequence* $u_N \sim cN \ln N$ *such that the distribution of* $\frac{\tilde{T}_N - u_N}{N}$ *converges to a stable law with index* 1*.*

(e) If $s = 2$ *then there is a constant* u *such that the distribution of* $\frac{\tilde{T}_N - Nu}{\sqrt{N \ln N}}$ *converges to a normal distribution.*

The proof of this theorem given in [15] makes use of the connection between random walks in random environment and branching processes. Another proof of Theorem 1 was given in [7, 4]. These papers make use of the notion of potential introduced by Ya. G. Sinai in [26] for the study of the recurrent case (when $\mathbf{E}(\ln(p/q)) = 0$).

The results for quenched limits (that is when a typical environment is fixed) are relatively recent. To prove an almost sure quenched limit theorem for $\tilde{T}_N$ one can make use of the representation

$$\tilde{T}_N = \sum_{i=1}^{N} \tau_i, \tag{1.1}$$

where τ_i is the time the walk starting from $i-1$ needs in order to reaches i for the first time. The advantage of this approach is due to the fact that if the environment ω is fixed then τ_i are independent random variables and this was used by many authors starting from the pioneering paper [25].

If $s > 2$ then one can prove the almost sure Central Limit Theorem (CLT) for $\tilde{T}_N$ checking that the sequence $\{\tau_i\}$ in (1.1) satisfies the Lindeberg condition for almost all ω (and for that one only needs the environment $\{p_i\}$ to be stationary, see e.g. [9]). Proving the CLT for X_n in this regime is a more delicate matter and this was done in [9] for several classes of environments (including the i.i.d. case) and independently in [16] for the i.i.d. environments. It has to be mentioned that, in the case of i.i.d. environments, it is easy to derive the annealed CLT from the related quenched CLT but this may not be easy for other classes of environments and in fact may not always be true.

In this paper, unlike in [15], we thus don't have to analyze the case $s > 2$. However, we explain at the end of Section 6 that it is not difficult to adapt the argument of that section to handle also the diffusive regime.

For $s < 2$, an important step was made in [17] and [19] where it was proved that it is impossible to have almost sure quenched limit theorems in this regime. Namely, for almost all ω no non-trivial distributional limit of $\frac{\tilde{T}_N - u_N}{v_N}$ exists for any choice of sequences $u_N(\omega)$ and $v_N(\omega)$.

In [6] it was proved that in the sub-ballistic regime $0 < s < 1$ the coordinate of the random walk becomes, as time $t \to \infty$, localized at the bottom of one of the finitely many valleys which are defined in terms of Sinai's potential.

The main goals of this paper is to present a complete description of the limiting behaviour of the random walk which turns out to be much more interesting than expected before. As will be seen below, the particularity of the sub-diffusive regime is that it is the asymptotic behaviour of the random environment that implies the limiting behaviour of the random walk. We show that $\tilde{T}_N$ viewed as a function of two(!) random parameters, $X(\cdot)$ and ω (the trajectory of the walk and the environment), does exhibit a limiting behaviour as $N \to \infty$ which for $0 < s < 2$ can be described explicitly in terms of a point Poisson process (Theorem 2). Namely, it turns out that for large fixed N and $\omega \in \Omega_N$ (where $\mathbf{P}(\Omega_N) \to 1$ as $N \to \infty$) the properly normalized $\tilde{T}_N$ is a linear combination of independent exponential random variables with coefficients of this combination depending only on ω and forming a point Poisson process. As a corollary, one obtains the results from [17] and [19] as well as a new proof of Theorem 1. In the case $s = 2$ we show that the CLT holds (Theorem 3); however, we provide a heuristic argument which shows that, in contrast with the case $s > 2$, the CLT does not hold for almost all ω but rather just for $\omega \in \Omega_N$.

The backbone of our approach is formed by the study of occupation times; such studies were initiated in [21, 23, 8]. In view of this technique it is more natural to consider the occupation time T_N of the interval $[0, N)$ rather than $\tilde{T}_N$. These two random variables have the same asymptotic behaviour (see Lemma 2.1) and therefore the results for $\tilde{T}_N$ follow easily from those for occupation times.

The main difference between our and other existing approaches is that:

− We introduce a Poisson process describing the "trapping properties" of the environment.

− This process allows us to separate explicitly the contribution to the occupation time (or, equivalently, hitting time) coming from the environment and the walk (and thus prove Theorem 2).

– It also allows us to answer some other interesting questions about the limiting behaviour of the walk (e.g., about the limiting behaviour of the distribution of the maximal occupation times, Theorem 4).

Similar results are valid in a more general setting of random walks in random environment on a strip and in particular for walks with bounded jumps. This will be a subject of a separate paper.

The layout of the paper is the following. In Section 2 we state our main results. In Section 3 we collect background information and prove some auxiliary results. In Section 4 we deduce Theorem 2 dealing with the case $s < 2$ from the fact that the set of sites with high expected number of visits has asymptotically Poisson distribution (Lemma 4.4). The proof of Lemma 4.4 itself is given in Section 5. The case when $s = 2$ (Theorem 3) requires a different approach (namely, we use big block-small block method of Bernstein) which is presented in Section 6. In Section 7 we explain how to modify the proof of Theorem 2 to obtain Theorem 4. In Appendix A we derive some previously known theorems from our results. Appendix B contains the derivation of the quenched limit theorem from our main result (Theorem 3).

We shall use the following convention about the constants appearing in the paper. The values of the constants can change from entry to entry unless it is explicitly stated otherwise.

After completing the paper we learned that Corollary 1 was proved independently by J. Peterson and G. Samorodnitsky [18] and by N. Enriquez, C. Sabot, L. Tournier, O. Zindy [5] using a different approach.

2. Main results

Throughout the paper the following definitions and notations will be used.

Definition. The *occupation time* T_N of the interval $[0, N)$ is the total time the walk X_n starting from 0 spends on this (semi-open) interval during its life time. In other words, $T_N = \#\{n : \ 0 \le n < \infty, \ 0 \le X_n \le N-1\}$

Remark. We thus use the following convention: starting from a site j counts as one visit of the walk to j.

The occupation time of a site j is defined similarly and is denoted by ξ_j. Observe that T_N (and ξ_j) is equal to the number of visits by the walk to $[0, N)$ (respectively, to site j). Since our random walk is transient to the right, both T_N and ξ_j are, $\mathbf{P}$-almost surely, finite random variables. It is clear from these definitions that

$$T_N = \sum_{j=0}^{N-1} \xi_j.$$

The following lemma shows that T_N and the hitting time $\tilde{T}_N$ have the same asymptotic behaviour.

Lemma 2.1. *For any* $\varepsilon > 0$

$$\mathbf{P}\left(\frac{|T_N - \tilde{T}_N|}{N^{1/s}} > \varepsilon\right) \to 0 \text{ as } N \to \infty.$$

Proof. It is easy to see that

$$\tilde{T}_N = \#\{n :\ 0 < n \le \tilde{T}_N,\ X_n \in [0, N-1]\} + \#\{n :\ 0 < n \le \tilde{T}_N,\ X_n < 0\}$$

and

$$T_N = \#\{n :\ 0 \le n \le \tilde{T}_N,\ X_n \in [0, N-1]\} + \#\{n :\ n > \tilde{T}_N,\ X_n \in [0, N-1]\}.$$

Since the first terms in these formulae are equal, $|T_N - \tilde{T}_N|$ can be estimated above by a sum of two random variables: the number of visits to the left of 0 and the number of visits to the left of N after $\tilde{T}_N$:

$$|T_N - \tilde{T}_N| \le \#\{n :\ n \ge 0,\ X_n < 0\} + \#\{n :\ n > \tilde{T}_N,\ X_n < N\}$$

The first term in this estimate is bounded for $\mathbf{P}$-almost all ω. Since $\tilde{T}_N$ is a hitting time, the second term has, for a given ω, the same distribution as $\#\{n :\ n > 0,\ X_n < N \,|\, X_0 = N\}$ (due to the strong Markov property). Finally, the latter is a stationary sequence with respect to the annealed measure and therefore is stochastically bounded. Hence the lemma. □

Remark. The difference between T_N and $\tilde{T}_N$ is thus negligible and yet there is a sharp contrast between their presentations by sums introduced above. Namely, unlike the τ_i's, the ξ_j's are not independent. Moreover, as we shall see below, there are whole random regions on $[0, N]$ where the knowledge of just one ξ_j essentially determines the values all the others. In fact, namely this strong interdependence of ξ_j's implies some of the main results of this paper.

From now on we shall deal mainly with t_N which is a normalized version of T_N, namely we set

$$\mathsf{t}_N = \begin{cases} \frac{T_N}{N^{1/s}} & \text{if } 0 < s < 1, \\ \frac{T_N - \mathbb{E}_\omega(T_N)}{N^{1/s}} & \text{if } 1 \le s < 2, \\ \frac{T_N - \mathbb{E}_\omega(T_N)}{\sqrt{N \ln N}} & \text{if } s = 2. \end{cases}$$

It is also important and natural to have control over the $\mathbb{E}_\omega(T_N)$. The corresponding normalized quantity is defined as follows:

$$\mathfrak{u}_N = \begin{cases} \frac{\mathbb{E}_\omega(T_N)}{N^{1/s}} & \text{if } 0 < s < 1, \\ \frac{\mathbb{E}_\omega(T_N)-u_N}{N} & \text{if } s = 1, \\ \frac{\mathbb{E}_\omega(T_N)-\mathbf{E}(T_N)}{N^{1/s}} & \text{if } 1 < s < 2, \\ \frac{\mathbb{E}_\omega(T_N)-\mathbf{E}(T_N)}{\sqrt{N \ln N}} & \text{if } s = 2, \end{cases}$$

where u_N is the same as in Theorem 1. Set

$$F_N^\omega(x) = \mathbb{P}_\omega\left(\mathfrak{t}_N \le x\right). \tag{2.1}$$

We can consider F_N^ω as a random variable with values in the space $\mathcal{X}$ of distributions on the line. Endowed with topology of weak convergence $\mathcal{X}$ is a topological space with topology given by the metric

$$d(F_1, F_2) = \inf\{\varepsilon : F_2(x-\varepsilon) - \varepsilon < F_1(x) < F_2(x+\varepsilon) + \varepsilon\}. \tag{2.2}$$

The result from [17, 19] cited above states that these processes are not concentrated near one point (at least for $0 < s < 2$). We shall show that nevertheless the limiting behaviour of the sequence $\mathfrak{t}_N$ can be described in terms of a marked point Poisson process which we shall now introduce.

We start with a point Poisson process. Given a $\mathbf{c} > 0$, let $\Theta = \{\Theta_j\}$ be a point Poisson process [1] on $(0, \infty)$ with intensity $\frac{\mathbf{c}}{\theta^{1+s}}$. For a given collection of points $\{\Theta_j\}$ let $\{\Gamma_{\Theta_j}\}$ be a collection of i.i.d. random variables with mean 1 exponential distribution which are labeled by the points $\{\Theta_j\}$. In the sequel we shall use a concise notation $\{\Gamma_j\}$ for $\{\Gamma_{\Theta_j}\}$. We can now consider a new process $(\Theta, \Gamma) = \{(\Theta_j, \Gamma_j)\}$ which is often called the marked point Poisson process. We note that (Θ, Γ) is in fact a point Poisson process on $(0, \infty) \times (0, \infty)$ with intensity $\frac{\mathbf{c}}{\theta^{1+s}} \times e^{-x}$. We shall denote by $E(\cdot)$, $\mathrm{Var}(\cdot)$, etc. the expectations, variances, etc. with respect to the distribution of (Θ, Γ) and by $P_\Theta(\cdot)$ the conditional probability distribution of Γ conditioned on Θ.

Set

$$Y = \begin{cases} \sum_j \Theta_j \Gamma_j & \text{if } 0 < s < 1 \\ \sum_j \Theta_j (\Gamma_j - 1) & \text{if } 1 \le s < 2 \end{cases}. \tag{2.3}$$

Observe that Y is finite almost surely. Indeed, there are only finitely many points with $\Theta_j \ge 1$. Next, if $0 < s < 1$ let

$$\tilde{Y} = \sum_{\Theta_j < 1} \Theta_j \Gamma_j.$$

[1] For reader's convenience we collect some facts about the Poisson processes in section 3.1.

Then

$$E(\tilde{Y}) = \int_0^1 \frac{\mathbf{c}\theta d\theta}{\theta^{1+s}} = \frac{\mathbf{c}}{1-s} < \infty.$$

In case $1 \leq s < 2$ let

$$\tilde{Y}_\delta = \sum_{\delta < \Theta_j < 1} \Theta_j(\Gamma_j - 1).$$

Then $E(\tilde{Y}_\delta) = 0$ and

$$\mathrm{Var}(\tilde{Y}_\delta) = \int_\delta^1 \frac{\mathbf{c}\theta^2 d\theta}{\theta^{1+s}} = \frac{\mathbf{c}}{2-s}\left(1 - \delta^{2-s}\right).$$

Denote by $\Theta^{(\delta)}$ a point Poisson process on $\mathbb{R}_\delta := [\delta, \infty)$ with intensity $\frac{\mathbf{c}}{\theta^{1+s}}$ and let $(\Theta^{(\delta)}, \Gamma)$ be a point process with Γ being as above. $(\Theta^{(\delta)}, \Gamma)$ can be viewed as a restriction of (Θ, Γ) a smaller phase space. It is important that $\Theta^{(\delta)}$ and $(\Theta^{(\delta)}, \Gamma)$ converge weakly, as $\delta \to 0$, to Θ and (Θ, Γ) respectively. Namely, for a given Θ define the conditional distribution function of Y by $F^\Theta(x) = P_\Theta(Y \leq x) \equiv P(Y \leq x \,|\, \Theta)$. Since Θ is a random parameter, F^Θ is a random variable taking values in $\mathcal{X}$. Next, for $1 \leq s < 2$ set

$$F_\delta^\Theta(x) = P(\sum_{\delta < \Theta_j} \Theta_j(\Gamma_j - 1) \leq x \,|\, \Theta)$$

and $F_\delta^\Theta(x)$ is defined similarly for $0 < s < 1$. Then for P-almost all Θ $d(F_\delta^\Theta, F^\Theta) \to 0$ as $\delta \to 0$.

Let $\mathfrak{F}_\delta$ be the set of all finite subsets of $\mathbb{R}_\delta$ and $\Theta^{(N,\delta)} : \Omega \mapsto \mathfrak{F}_\delta$ be a sequence of point processes defined on the space of environments Ω and taking values in $\mathfrak{F}_\delta$. The standard definitions of the relevant sigma-algebra and measurability can be found e.g. in [22]. In the constructions below such sequences will be arising in a natural way and it will always be clear that the relevant mappings are measurable. Set $|\Theta^{(N,\delta)}| \equiv \mathrm{Card}(\Theta^{(N,\delta)})$. We need the following

Definition. A sequence of random point processes $\Theta^{(N,\delta)} = \{\Theta_j^{(N,\delta)}\}$ defined on Ω converges weakly to a Poisson process $\Theta^{(\delta)}$ if for any $k \geq 1$ and any bounded continuous symmetric function $H_k : \mathbb{R}_\delta^k \mapsto \mathbb{R}$ of k variables

$$\lim_{N\to\infty} \mathbf{E}\left(H_k(\Theta^{(N,\delta)})\, I_{|\Theta^{(N,\delta)}|=k}\right) = E\left(H_k(\Theta^{(\delta)})\, I_{|\Theta|=k}\right).$$

Suppose next that $\Gamma^{(N,\delta)}$ is a collection of random variables defined on Ω and labeled by the points of $\Theta^{(N,\delta)} = \{\Theta_j^{(N,\delta)}\}$:

$$\left(\Theta^{(N,\delta)}, \Gamma^{(N,\delta)}\right) = \{(\Theta_j^{(N,\delta)}(\omega), \Gamma_{\Theta_j}^{(N,\delta)}(\omega))\}$$

As in the definition of (Θ, Γ), we shall write $\{\Gamma_j^{(N,\delta)}\}$ for $\{\Gamma_{\Theta_j}^{(N,\delta)}\}$. Finally, the process $(\Theta^{(N,\delta)}, \Gamma^{(N,\delta)}) = \{(\Theta_j^{(N,\delta)}, \Gamma_j^{(N,\delta)})\}$ can be viewed as a mapping

$$(\Theta^{(N,\delta)}, \Gamma^{(N,\delta)}) : \Omega \mapsto \mathfrak{F}_\delta \times \tilde{\mathfrak{F}},$$

where $\tilde{\mathfrak{F}}$ is the set of all finite subsets of $[0, \infty)$. The weak convergence of this sequence of processes to $(\Theta^{(\delta)}, \Gamma)$ is defined as above with the only difference that now we have to deal with symmetric continuous functions $H_k : (\mathbb{R}_\delta \times [0, \infty))^k \mapsto \mathbb{R}$.

Definition. $\{\Gamma_j^{(N,\delta)}(\omega)\}$ is said to be asymptotically i.i.d. with distribution $\boldsymbol{\nu}$ and asymptotically independent of the environment if for any $k \geq 1$ and any bounded continuous symmetric function H_k : $(\mathbb{R}_\delta \times [0, \infty))^k \mapsto \mathbb{R}$ of k pairs of variables $(\Theta, \Gamma) = ((\Theta_1, \Gamma_1), ..., (\Theta_k, \Gamma_k))$

$$\lim_{N \to \infty} \mathbf{E}\left[I_{|\Theta^{(N,\delta)}|=k} \left|\mathbb{E}_\omega\left(H_k((\Theta^{(N,\delta)}, \Gamma^{(N,\delta)}))\right) - \bar{H}_k(\Theta^{(N,\delta)})\right|\right] = 0,$$

where

$$\bar{H}_k(\Theta_1, \ldots, \Theta_k) = \int \ldots \int H_k((\Theta_1, \Gamma_1), \ldots, (\Theta_k, \Gamma_k)) d\boldsymbol{\nu}(\Gamma_1) \ldots \boldsymbol{\nu}(\Gamma_1).$$

Note that here $H_k((\Theta^{(N,\delta)}, \Gamma^{(N,\delta)}))$ is well defined because $|\Gamma^{(N,\delta)}| = |\Theta^{(N,\delta)}| = k$. We can now state our main result.

Theorem 2. *For $0 < s < 2$ and a $\delta > 0$ there is a sequence $\Omega_{N,\delta} \subset \Omega$ such that $\lim_{N\to\infty} \mathbf{P}(\Omega_{N,\delta}) = 1$ and a sequence of random point processes*

$$(\Theta^{(N,\delta)}, \Gamma^{(N,\delta)}) : \Omega \times \mathfrak{X} \mapsto \mathfrak{F}_\delta \times \tilde{\mathfrak{F}},$$

such that
(i) The component $\Theta^{(N,\delta)}$ depends only on ω and converges weakly to a point Poisson process $\Theta^{(\delta)}$ on $[\delta, \infty)$ with intensity $\frac{\bar{c}}{\theta^{1+s}}$ (with some constant $\bar{c} > 0$).
(ii) The component $\Gamma^{(N,\delta)} = \{\Gamma_j^{(N,\delta)}\}$ is a collection of asymptotically i.i.d. random variables with mean 1 exponential distribution which also are asymptotically independent of the environment.
(iii) The $\mathfrak{t}_N$ and $\mathfrak{u}_N$ can be presented in the following form:
(a) If $0 < s < 1$ then for $\omega \in \Omega_{N,\delta}$
(2.4)

$$\mathfrak{t}_N = \sum_j \Theta_j^{(N,\delta)} \Gamma_j^{(N,\delta)} + R_N, \quad \textit{where } R_N \geq 0 \textit{ and } \mathbf{E}(1_{\Omega_{N,\delta}} R_N) = \mathcal{O}(\delta^{1-s})$$

$$\mathfrak{u}_N = \sum_j \Theta_j^{(N,\delta)} + \hat{R}_N, \quad \textit{where } \hat{R}_N \geq 0,\ \mathbf{E}(\hat{R}_N) = \mathcal{O}(\delta^{1-s})$$

(b) If $s=1$ then for $\omega \in \Omega_{N,\delta}$ and a given $1/2<\kappa<1$

$$\mathfrak{t}_N=\sum_j \Theta_j^{(N,\delta)}(\Gamma_j^{(N,\delta)}-1)+R_N, \quad \textit{where } \mathbf{E}\left[1_{\Omega_{N,\delta}}\mathbb{E}_\omega(R_N^2)\right]^\kappa=\mathcal{O}(\delta^{2\kappa-1})$$

$$\mathfrak{u}_N=\sum_j \Theta_j^{(N,\delta)}-\bar{c}|\ln\delta|+\hat{R}_N, \quad \textit{where } \mathbf{E}(|\hat{R}_N|^2)=\mathcal{O}(\delta)$$

(c) If $1<s<2$ then for $\omega \in \Omega_{N,\delta}$

$$\mathfrak{t}_N=\sum_j \Theta_j^{(N,\delta)}(\Gamma_j^{(N,\delta)}-1)+R_N, \quad \textit{where } \mathbf{E}\left[1_{\Omega_{N,\delta}}\mathbb{E}_\omega(R_N^2)\right]=\mathcal{O}(\delta^{2-s})$$

$$\mathfrak{u}_N=\sum_j \Theta_j^{(N,\delta)}-\frac{\bar{c}}{(s-1)\delta^{s-1}}+\hat{R}_N, \quad \textit{where } \mathbf{E}(\hat{R}_N^2)=\mathcal{O}(\delta^{2-s})$$

Remark. The estimates of the remainders in the statements of Theorem 2 are not uniform in N but are uniform in δ. More precisely, e. g. the relation $\mathbf{E}(|\hat{R}_N|^2)=\mathcal{O}(\delta)$ means that for any $\delta>0$ there is N_δ and a constant C (which does not depend on δ) such that $\mathbf{E}(|\hat{R}_N|^2)\le C\delta$ if $N>N_\delta$.

Remark. Note that the dependence of $\Theta^{(N,\delta)}$ on ω persists as $N\to\infty$ whereas $\Gamma^{(N,\delta)}$ becomes "almost" independent of ω. More precisely, for $K\gg 1$ and sufficiently large N the events $B_k:=\{|\Theta^{(N,\delta)}|=k\}$, $0\le k\le K$, form, up to a set of a small probability, a partition of Ω. Obviously

$$\lim_{N\to\infty}\mathbf{P}\{|\Theta^{(N,\delta)}|=k\}=\frac{e^{-\tilde{c}\delta^{-s}}(\tilde{c}\delta^{-s})^k}{k!},$$

where $\tilde{c}=\bar{c}/s$. In contrast, if $\omega\in B_k$ then $\Gamma^{(N,\delta)}(\omega,X)$ is a collection of k random variables which converge weakly as $N\to\infty$ to a collection of k i.i.d. standard exponential random variables. Thus the only dependence of $\Gamma^{(N,\delta)}(\omega,X)$ on ω and δ which persists as $N\to\infty$ is reflected by the fact that $|\Theta^{(N,\delta)}|=|\Gamma^{(N,\delta)}|$.

It is natural to expect that Theorem 2 implies weak convergence of the relevant distributions. Namely, both $F_N^\omega(x)$ defined by (2.1) and $F^\Theta(x)=P(Y\le x\,|\,\Theta)$ can be viewed as monotone random processes with x playing the role of the time of the process and with random parameters ω and Θ respectively or, eqivalently, as random variable taking values in $\mathcal{X}$. We say that $F_N^\omega\Rightarrow F^\Theta$ as $N\to\infty$ if for any continuous function $\varphi:\mathcal{X}\mapsto\mathbb{R}$

$$\lim_{N\to\infty}\mathbf{E}(\varphi(F_N^\omega))=E(\varphi(F^\Theta)).$$

The following corollary follows from Theorem 2.

Corollary 1. *(a) If $0 < s < 2$, $s \neq 1$ then F^ω_N converges weakly to F^Θ.*

(b) If $1 < s < 2$ then $\left(F^\omega_N, \frac{\mathbb{E}_\omega(T_N) - \mathbf{E}(T_N)}{N^{1/s}}\right)$ converges weakly to

$$\left(F^\Theta, \lim_{\delta\to 0}\left(\sum_{\Theta_j > \delta} \Theta_j - \frac{\bar{c}}{(s-1)\delta^{s-1}}\right)\right).$$

(c) If $s = 1$ then there exists $u_N \sim \bar{c} N \ln N$ such that $\left(F^\omega_N, \frac{\mathbb{E}_\omega(T_N) - u_N}{N}\right)$ converges weakly to

$$\left(F^\Theta, \lim_{\delta\to 0}\left(\sum_{\Theta_j > \delta} \Theta_j + \bar{c}\ln\delta\right)\right)$$

where $E\left(\sum_{\delta<\Theta_j<1} \Theta_j\right) = -\bar{c}\ln\delta$.

The proof of this corollary follows from a general statement about weak convergence of monotone random processes. A brief explanation of relevant ideas is given in Appendix B.

Remark. Similar limiting distributions were obtained in [24] for a simpler model of 'random climbing' where the particle moves forward with unit speed and with intensity 1 it slides back to a nearest point of intensity λ Poisson process.

We also recover the result of [19].

Corollary 2. *For $0 < s < 2$ and $\mathbf{P}$-almost every environment ω the sequence $\mathfrak{t}_N(\omega, X)$ has no limiting distribution as $N \to \infty$. Moreover, fix a finite sequence $a_j > 0$. Let $\mathbf{F}$ be the distribution function of*

$$\begin{cases} \sum_j a_j \Gamma_j, & 0 < s < 1, \\ \sum_j a_j(\Gamma_j - 1), & 1 \le s < 2. \end{cases} \tag{2.5}$$

Then with probability one there exists a sequence $N_k(\omega)$ such that $d(F^\omega_{N_k(\omega)}, \mathbf{F}) \to 0$ as $k \to \infty$.

Consequently, for $\mathbf{P}$-almost every environment ω any distribution that can be obtained as a limit of distributions of type (2.5) can also be obtained as a weak limit of $\mathfrak{t}_{N_k}(\omega, X)$ as $k \to \infty$, where N_k depends on ω and $\{a_j\}$.

The proof of this statement will be given in the Appendix A.

We complete the picture by stating the result for the case $s = 2$.

Theorem 3. *If* $s = 2$ *then there are constants* D_1, D_2 *such that* $(\mathfrak{t}_N, \mathfrak{u}_N)$ *converge weakly to* $(\mathcal{N}_1, \mathcal{N}_2)$ *where* $\mathcal{N}_1$ *and* $\mathcal{N}_2$ *are independent Gaussian random variables with zero means and variances* D_1 *and* D_2 *respectively. Moreover,* $\mathfrak{t}_N$ *is asymptotically independent of the environment.*

Remark. For $s = 2$ the fact that $\mathfrak{u}_N$ is asymptotically normal was proved in [12] and so to prove Theorem 3 it is enough to show that for any $\varepsilon > 0$

$$\mathbf{P}\left(\sup_x |F_N^\omega(x) - F_{\mathcal{N}_1}(x)| > \varepsilon\right) \to 0 \quad \text{as} \quad N \to \infty. \tag{2.6}$$

Indeed F_N^ω and $\mathfrak{u}_N = \frac{\mathbb{E}_\omega(T_N) - \mathbf{E}(T_N)}{\sqrt{N \ln N}}$ are evidently asymptotically independent since the distribution of the latter depends only on the environment and the distribution of the former is asymptotically the same for the set of ωs of asymptotically full measure.

It is well known that the reason why the hitting times do not always satisfy the Central Limit Theorem is the presence of traps which slow down the particle. It will be seen in the proofs that Theorems 2 and 3 state that if traps are ordered according to the expected time the walker spends inside the trap then the asymptotic distribution of traps is Poissonian with intensity $\frac{\mathbf{c}}{\theta^{1+s}}$. This result holds regardless of the value of s. However, if $s \geq 2$ then the time spent inside the traps is smaller than the time spent outside of the traps.

Let as before ξ_n be the number of visits to n and $\xi_N^* = \max_{[0,N]} \xi_n$.

Theorem 4. *If* $s > 0$ *then* $\frac{\xi_N^*}{N^{1/s}}$ *converges to* $\max_j \hat{\Theta}_j$*, where* $\hat{\Theta}$ *is a Poisson process on* $(0, \infty)$ *with intensity* $\frac{\bar{\mathbf{c}}}{\theta^{1+s}}$ *for some constant* $\bar{\mathbf{c}}$. *Accordingly*

$$\mathbf{P}\left(\xi_N^* < x N^{1/s}\right) \to \exp\left[-\frac{\bar{\mathbf{c}}}{s} x^{-s}\right].$$

Theorem 4 shows that the fact that traps are Poisson distributed is useful even for $s > 2$.

Corollary 3. *If* $0 < s < 1$ *then as* $N \to \infty$

$$\limsup \frac{\xi_N^*}{T_N} > 0$$

almost surely.

Remark. Corollary 3 is a minor modification of the result of [8]. Namely, in [8] the authors consider not all visits to site n but only visits before $\tilde{T}_N$. By Lemma 2.1 this difference is not essential since most visits occur before $\tilde{T}_N$.

3. Preliminaries.

3.1. **Poisson process.** The proofs of the facts listed below can be found in monographs [20, 22].

Let (X, μ) be a measure space. Recall that a Poisson process is a point process on X such that
(a) if $A \subset X$, $\mu(A)$ is finite, and $N(A)$ is the number of points in A then $N(A)$ has a Poisson distribution with parameter $\mu(A)$;
(b) if $A_1, A_2 \dots A_k$ are disjoint subsets of X then $N(A_1), N(A_2) \dots N(A_k)$ are mutually independent.

If $X \subset \mathbb{R}^d$ and μ has a density f with respect to the Lebesgue measure we say that f is the intensity of the Poisson process.

Lemma 3.1. *(a) If $\{\Theta_j\}$ is a Poisson process on X with measure μ and $\psi : X \to \tilde{X}$ is a measurable map then $\tilde{\Theta}_j = \psi(\Theta_j)$ is a Poisson process with measure $\tilde{\mu}$ where $\tilde{\mu}(\tilde{A}) = \mu(\psi^{-1}\tilde{A})$.*

In particular if $X = \tilde{X} = \mathbb{R}$ and ψ is invertible then the intensity of $\tilde{\Theta}$ is

$$\tilde{f}(\theta) = f(\psi^{-1}(\theta)) \left| \frac{d\psi}{d\theta} \right|^{-1}. \tag{3.1}$$

(b) Let (Θ_j, Γ_j) be a point process on $X \times Z$. Then (Θ_j, Γ_j) is a Poisson process on $X \times Z$ with measure $\mu \times \boldsymbol{\nu}$ where $\boldsymbol{\nu}$ is a probability measure if and only if $\{\Theta_j\}$ is a Poisson process on X with measure μ and $\{\Gamma_j\}$ are Z-valued random variables which are i.i.d. with distribution $\boldsymbol{\nu}$ and independent of $\{\Theta_k\}$.

(c) If in (b) $X = Z = \mathbb{R}$ then $\tilde{\Theta} = \{\Gamma_j \Theta_j\}$ is a Poisson process. Its intensity is

$$\tilde{f}(\theta) = \int f\left(\theta\Gamma^{-1}\right)\Gamma^{-1}\boldsymbol{\nu}(d\Gamma).$$

Lemma 3.2. *Let Θ be Poisson process on X, $\psi : X \to \mathbb{R}$ a measurable function with $\int |\psi(\theta)| d\mu(\theta) < \infty$ then*

$$V = \sum_j \psi(\theta_j)$$

is finite with probability 1, the characteristic function of V is given by

$$E(\exp(ivV)) = \exp\left[\int \left(e^{iv\psi(\theta)} - 1\right) d\mu(\theta)\right], \tag{3.2}$$

and

$$E(V) = \int \psi(\theta) d\mu(\theta). \tag{3.3}$$

If in addition to the above conditions $\int \psi^2(\theta)d\mu(\theta) < \infty$ *then*

$$\mathrm{Var}(V) = \int \psi^2(\theta)d\mu(\theta) \tag{3.4}$$

Remark. Proofs of the statements listed in Lemmas 3.1 and 3.2 can be found in [20].

Lemma 3.3. *(a) If* $0 < s < 1$ *and* Θ_j *is a Poisson process with intensity* $\theta^{-(1+s)}$ *then* $\sum_j \Theta_j$ *has a stable distribution of index* s.

(b) If $1 < s < 2$ *and* Θ_j *is a Poisson process with intensity* $\theta^{-(1+s)}$ *then*

$$\lim_{\delta\to 0}\left[\left(\sum_{\delta<\Theta_j}\Theta_j\right)-\frac{1}{(s-1)\delta^{s-1}}\right]$$

has a stable distribution of index s.

(c) If $s = 1$ *and* Θ_j *is a Poisson process with intensity* θ^{-2} *then*

$$\lim_{\delta\to 0}\left[\left(\sum_{\delta<\Theta_j}\Theta_j\right)-|\ln\delta|\right]$$

has a stable distribution of index 1.

Remark. The proof of Lemma 3.3 follows from a direct computation of the characteristic function of the relevant sums in (a), (b), (c) using formula (3.2). We also note that the expressions under the limit sign in (b) and (c) are equal to $\sum_{\delta<\Theta_j}\Theta_j - E_\Theta\left(\sum_{\delta<\Theta_j}\Theta_j\right)$. One thus could say that the existence of the limit means that the series $\sum_j(\Theta_j - E_\Theta(\Theta_j))$ converges. However, for this interpretation of one has to introduce an ordering relation on the random sets $\{\Theta_j\}$ (see [22]).

3.2. **Backtracking.** As was mentioned before the analysis of our paper relies on the fact that the random walk is unlikely to backtrack. The precise statement we shall use is the following.

Lemma 3.4. *([8], Lemma 3.3) There exist* $C > 0, \beta < 1$ *such that*

$$\mathbf{P}(X \text{ visits } n \text{ after } n+m) \leq C\beta^m.$$

Remark. Here and below letters $\beta,\ \bar\beta,\ \beta_1,\ \beta_2$, etc. always denote a positive constant which is strictly smaller than 1. The precise meaning of these is always clear from the context.

3.3. **Occupation times. Recurrence relation.** As before, let ξ_n be the number of visits to the site n and set $\rho_n = \mathbb{E}_\omega \xi_n$. Observe that ξ_n has geometric distribution with parameter $1/\rho_n$.

Lemma 3.5. *If $X_0 = 0$ then for $n \geq 0$*

$$\rho_n = p_n^{-1} q_{n+1}\rho_{n+1} + p_n^{-1} = p_n^{-1}(1 + \alpha_{n+1} + \alpha_{n+1}\alpha_{n+2} + ...), \tag{3.5}$$

where $\alpha_j = \frac{q_j}{p_j}$.

Proof. Let η_n^+ and η_n^- be the number of passages of the edge $[n, n+1]$ in the forward, respectively, backward direction. Denote $\sigma_n^\pm = \mathbb{E}_\omega \eta_n^\pm$. We have

$$\rho_n = \sum_j \mathbb{P}_\omega(X_j = n) \text{ and } \sigma_n^+ = \sum_j \mathbb{P}_\omega(X_j = n, X_{j+1} = n+1).$$

Thus $\sigma_n^+ = \rho_n p_n$. Likewise $\sigma_n^- = \rho_{n+1} q_{n+1}$. Since $X_n \to +\infty$ we have that $\eta_n^+ - \eta_n^- = 1$ for $n \geq 0$. Hence $\rho_n p_n - \rho_{n+1} q_{n+1} = 1$ which implies the first relation in (3.5). Iterating this relation k times gives

$$\rho_n = p_n^{-1}\alpha_{n+1}\dots\alpha_{n+k-1} q_n \rho_n + (1 + \alpha_{n+1} + \dots + \alpha_{n+1}\dots\alpha_{n+k-1}) p_n^{-1}. \tag{3.6}$$

Since $\mathbf{E}(\ln \alpha) < 0$ we see that the first term in (3.6) tends to 0 as $k \to \infty$ almost surely and this proves the second relation in (3.5). $\square$

For future references, we record a useful bound for ρ_{n-k} in terms of ρ_n. For $k \geq 1$ set $A_{n,k} := \prod_{j=1}^{k-1} \alpha_{n-k+j}$ (with $A_{n,1} = 1$),$B_{n,k} := 1 + \alpha_{n-k+1} + \dots + \alpha_{n-k+1}\dots\alpha_{n-1}$. Then (3.6), with n replaced by $n-k$, can be rewritten as

$$\rho_{n-k} = p_{n-k}^{-1} q_n A_{n,k} \rho_n + p_{n-k}^{-1} B_{n,k}. \tag{3.7}$$

Set $\bar{c} := \varepsilon_0^{-1}$, where ε_0 is from condition (C). It follows from (3.7) that

$$\rho_{n-k} \leq \bar{c} A_{n,k} \rho_n + \bar{c} B_{n,k} \tag{3.8}$$

and (3.8) implies that

$$\rho_{n-k} \leq \bar{c} A_{n,k} \rho_n + \bar{c} k \varepsilon_0^{-k}. \tag{3.9}$$

Note that $A_{n,k}$ and ρ_n are independent random variables.

Next, we introduce

$$\begin{aligned} z_n &:= 1 + \alpha_{n+1} + \alpha_{n+1}\alpha_{n+2} + ... + \alpha_{n+1}...\alpha_{n+m} + ... \\ &= 1 + \alpha_{n+1} + \alpha_{n+1}\alpha_{n+2} + ... + \alpha_{n+1}...\alpha_{n+m} z_{n+m} \end{aligned} \tag{3.10}$$

In particular, we have that $z_n = 1 + \alpha_{n+1} z_{n+1}$, where α_{n+1} and z_{n+1} are independent random variables and the sequence $\{z_n\}_{-\infty<n<\infty}$ considered backward in time forms a Markov chain. Note that $\rho_n = p_n^{-1} z_n$

is a function on the phase space of the Markov chain $\{p_n, z_n\}$ (where p_n and z_n are independent).

We also need a slightly more general Markov process u_n defined by

$$u_n = \sigma_{n+1} + \alpha_{n+1} u_{n+1} \tag{3.11}$$

where the pair $(\sigma_{n+1}, \alpha_{n+1})$ is independent of $\{u_{n+m}\}_{m=1}^{\infty}$.

The stationary distributions of z_n, ρ_n, and u_n have a very important heavy tail property which plays a crucial role in our proofs and is described by the following

Lemma 3.6. *Suppose that $0 < c_0 \le \sigma_n \le c_1$ for some $c_0,\ c_1$ with probability 1 and that conditions (A) – (D) are satisfied. Then*

(a) Let ν denote the stationary measure for the process u_n. Then there is a constant C such that $\lim_{x\to+\infty} x^s \nu(u > x) = C$, where $s > 0$ satisfies $\mathbf{E}(\alpha^s) = 1$ (as in condition (B)).

(b) There exists a $c > 0$ such that $\lim_{x\to+\infty} x^s \mathbf{P}(z_n > x) = c$.

(c) There exists $c^* > 0$ such that $\lim_{x\to+\infty} x^s \mathbf{P}(\rho_n > x) = c^*$.

Proof. (a) is proven in [13] (under more general conditions). (b) is a particular case of (a). (c) follows from (b) since $\rho_n = p_n^{-1} z_n$ and

$$\mathbf{P}(p_n^{-1} z_n > x) = \mathbf{E}[\mathbf{P}(z_n > x p_n | p_n)] \sim \mathbf{E}(c x^{-s} p^{-s}) = c x^{-s} \mathbf{E}(p^{-s}),$$

where the first equality is due to the total probability formula and the second to the independence of p_n and z_n. We also see that $c^* = c\mathbf{E}(p^{-s})$. □

Lemma 3.7. *There exist $\varepsilon_1 > 0, \varepsilon_2 > 0, 0 < \beta < 1$ such that for any $\delta > 0$ there are N_δ and $C = C_\delta > 0$ such that for $N > N_\delta$ one has:*
(a) If $k \le \varepsilon_1 \ln N$ then

$$\mathbf{P}(\rho_n \ge \delta N^{1/s}, \rho_{n-k} \ge \delta N^{1/s}) \le \frac{C\beta^k}{N};$$

(b) If $k \ge \varepsilon_1 \ln N$ then

$$\mathbf{P}(\rho_n \ge \delta N^{1/s}, \rho_{n-k} \ge \delta N^{1/s}) \le C N^{-(\varepsilon_2+1)}.$$

Proof. (a) It follows from (3.9) that if ε_1 is chosen so that $-\varepsilon_1 \ln \varepsilon_0 \le \frac{1}{3s}$ and N is sufficiently large then

$$\rho_{n-k} \le \bar{c}\rho_n A_{n,k} + \bar{c}\varepsilon_1 (\ln N) \varepsilon_0^{-\varepsilon_1 \ln N} \le \bar{c}\rho_n A_{n,k} + \bar{c} N^{\frac{1}{2s}}.$$

Next, there exist $\beta_1, \beta_2 < 1$ such that

$$\mathbf{P}(\alpha_{n-1} \ldots \alpha_{n-k} \ge \beta_1^k) \le \beta_2^k. \tag{3.12}$$

Indeed, if $0 < h < \min(1, s)$ and β_1 is such that $\mathbf{E}(\alpha^h) < \beta_1^h < 1$ then it follows from the Markov's inequality that

$$\mathbf{P}(\alpha_{n-1} \dots \alpha_{n-k} \geq \beta_1^k) \leq \frac{(\mathbf{E}(\alpha^h))^k}{\beta_1^{hk}} \equiv \beta_2^k. \tag{3.13}$$

We can now choose N_δ so that for $N > N_\delta$ we shall have

$$\begin{aligned}\mathbf{P}(\rho_n \geq \delta N^{1/s}, \rho_{n-k} \geq \delta N^{1/s}) &\leq \mathbf{P}(\rho_n \geq \delta N^{1/s}, \bar{c}\rho_n A_{n,k} + \bar{c} N^{\frac{1}{2s}} \geq \delta N^{1/s}) \\ &\leq \mathbf{P}(\rho_n \geq \delta N^{1/s}, \bar{c}\rho_n A_{n,k} \geq \frac{\delta}{2} N^{1/s})\end{aligned}$$

Finally, the right hand side in the above inequality is estimated as follows:

$$\begin{aligned}&\mathbf{P}(\rho_n \geq \delta N^{1/s},\, \bar{c}\rho_n A_{n,k} \geq \frac{\delta}{2} N^{1/s}) \\ &= \mathbf{P}(\rho_n \geq \delta N^{1/s},\, \bar{c}\rho_n A_{n,k} \geq \frac{\delta}{2} N^{1/s},\, A_{n,k} \leq \beta_1^k) \\ &+ \mathbf{P}(\rho_n \geq \delta N^{1/s},\, \bar{c}\rho_n A_{n,k} \geq \frac{\delta}{2} N^{1/s},\, A_{n,k} > \beta_1^k) \\ &\leq \mathbf{P}\left(\rho_n \geq \frac{\beta_1^{-k}\delta N^{1/s}}{2\bar{c}}\right) + \mathbf{P}(\rho_n > \delta N^{1/s} \text{ and } A_{n,k} > \beta_1^k) \leq \text{Const}\frac{\beta_1^{ks} + \beta_2^k}{N},\end{aligned}$$

where the last step makes use of Lemma 3.6 (hence the dependence of the Const on δ) and of independence of ρ_n and $A_{n,k}$.

(b) For any $\varepsilon_3 > 0$ we can write

$$\begin{aligned}&\mathbf{P}(\rho_n \geq \delta N^{1/s},\, \rho_{n-k} \geq \delta N^{1/s}) \\ &\leq \mathbf{P}(\delta N^{1/s} \leq \rho_n \leq \delta N^{\frac{1+\varepsilon_3}{s}},\, \rho_{n-k} \geq \delta N^{1/s}) + \mathbf{P}(\rho_n > \delta N^{\frac{1+\varepsilon_3}{s}}) \\ &\leq \frac{\bar{\bar{c}}}{N^{1+\varepsilon_3}} + \mathbf{P}(\delta N^{1/s} \leq \rho_n \leq \delta N^{\frac{1+\varepsilon_3}{s}},\, \rho_{n-k} \geq \delta N^{1/s}),\end{aligned} \tag{3.14}$$

where the last step follows from Lemma 3.6. It is clear from (3.8) that the last term in (3.14) can be estimated above by

$$\begin{aligned}&\leq \mathbf{P}(\delta N^{1/s} \leq \rho_n \leq \delta N^{\frac{1+\varepsilon_3}{s}},\, \bar{c}A_{n,k}\rho_n + \bar{c}B_{n,k} \geq \delta N^{1/s}) \\ &\leq \mathbf{P}(\delta N^{1/s} \leq \rho_n \leq \delta N^{\frac{1+\varepsilon_3}{s}},\, \bar{c}A_{n,k}\delta N^{\frac{1+\varepsilon_3}{s}} + \bar{c}B_{n,k} \geq \delta N^{1/s}) \\ &= \mathbf{P}(\delta N^{1/s} \leq \rho_n \leq \delta N^{\frac{1+\varepsilon_3}{s}})\, \mathbf{P}(\bar{c}A_{n,k}\delta N^{\frac{1+\varepsilon_3}{s}} + \bar{c}B_{n,k} \geq \delta N^{1/s}),\end{aligned} \tag{3.15}$$

where the last step is due to the independence of ρ_n and $(A_{n,k}, B_{n,k})$. Next, let $1 > h > 0$ be such that $\bar{\beta} = \mathbf{E}(\alpha^h) < 1$, then $\mathbf{E}(B_{n,k}^h) \leq$

$(1-\bar\beta)^{-1}$. By Markov's inequality

(3.16)

$$\mathbf{P}(\bar c A_{n,k}\delta N^{\frac{1+\varepsilon_3}{s}}+\bar c B_{n,k}\geq \delta N^{1/s})\leq \bar c^h\frac{\mathbf{E}(\delta^h N^{\frac{1+\varepsilon_3}{s}h}A^h_{n,k}+B^h_{n,k})}{\delta^h N^{h/s}}$$
$$\leq \bar{\bar c} N^{\frac{\varepsilon_3 h}{s}}\bar\beta^k+\bar{\bar c}N^{\frac{-h}{s}}.$$

Since $k\geq \varepsilon_1\ln N$, we have that $N^{\frac{\varepsilon_3 h}{s}}\bar\beta^k\leq N^{\frac{\varepsilon_3 h}{s}+\varepsilon_1\ln\bar\beta}=N^{-\bar\varepsilon}$ (with ε_3 sufficiently small so that to make $\bar\varepsilon$ strictly positive). Finally, it follows from Lemma 3.6, (3.15) and (3.16) that

(3.17)

$$\mathbf{P}(\delta N^{1/s}\leq\rho_n\leq\delta N^{\frac{1+\varepsilon_3}{s}},\ \rho_{n-k}\geq\delta N^{1/s})\leq \mathrm{Const}N^{-1-\min(\bar\varepsilon,\,h/s)}.$$

The proof of (b) now follows from (3.17) and (3.14). □

Next, we need the fact that ρ_n is exponentially mixing by which we mean that for a typical realization of α the dependence of ρ_{n-k} on ρ_n decays exponentially. To prove this we use (3.7). We formulate this statement as follows. Given a $\hat\rho_n$ define for $k>0$

$$\hat\rho_{n-k}=p^{-1}_{n-k}\hat\rho_n q_n A_{n,k}+p^{-1}_{n-k}B_{n,k}. \tag{3.18}$$

We are mainly interested in the case when the difference between $\hat\rho_n$ and ρ_n is large. More specifically we assume that $\hat\rho_n^h\geq\mathbf{E}(\rho_n^h)+2$, where $0<h<\min(1,s)$ is as in (3.13). Then the following holds.

Lemma 3.8. *Let $\hat\rho_{n-k}$ be defined by* (3.18) *and ρ_n be the stationary sequence satisfying* (3.7). *Then there exist $K>0$ and β_1, $\beta_3<1$ such that for $k>K\ln\hat\rho_n$*

$$\mathbf{P}\left(|\rho_{n-k}-\hat\rho_{n-k}|\geq\beta_1^k\right)\leq\beta_3^k.$$

Proof. It follows from (3.7) and (3.18) that

$$|\rho_{n-k}-\hat\rho_{n-k}|\leq\bar c A_{n,k}\,|\rho_n-\hat\rho_n|\,.$$

Consider the same $0<h<1$, β_1, and β_2 as in (3.12), (3.13) and set $\beta_3=(1+\beta_2)/2$. Then

$$\mathbf{P}\left(|\rho_{n-k}-\hat\rho_{n-k}|\geq\beta_1^k\right)\leq\mathbf{P}\left(\bar c A_{n,k}\,|\rho_n-\hat\rho_n|\geq\beta_1^k\right)\leq\beta_2^k[\mathbf{E}(\rho_n^h)+\hat\rho_n^h]\leq\beta_3^k.$$

Here the first inequality is obvious. The second one is due to the Markov inequality, to (3.12), and to the independence of ρ_n and $A_{n,k}$. Finally, one easily checks that the third one holds for $k>K\ln\hat\rho_n$, where $K:=2h/\ln(0.5+0.5\beta_2^{-1})+1$ (this is where the condition $\hat\rho_n\geq\mathbf{E}(\rho_n^h)+2$ is used). □

3.4. **Occupation times. Correlations.** The proofs of Lemmas 3.10 and 3.11 will make use of several elementary equalities and inequalities concerned with a Markov chain $Y = \{Y_t,\ t \geq 0\}$ with a phase space $\{1,2,3\}$ and a transition matrix

$$\begin{pmatrix} \bar p & \bar q & 0 \\ \bar{\bar q} & \bar{\bar p} & \varepsilon \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.19}$$

Namely, let $\bar\eta$ and $\bar{\bar\eta}$ be the total numbers of visit by Y to sites 1 and 2 respectively. Set $U_1 = E(\bar\eta|Y_0=1)$, $U_2 = E(\bar\eta|Y_0=2)$, $V_1 = E(\bar{\bar\eta}|Y_0 = 1)$, $V_2 = E(\bar{\bar\eta}|Y_0 = 2)$. It follows easily from the standard first step analysis that

$$U_1 = \frac{\varepsilon + \bar{\bar q}}{\varepsilon \bar q}, \quad U_2 = \frac{\bar{\bar q}}{\varepsilon \bar q}, \quad V_1 = V_2 = \frac{1}{\varepsilon}. \tag{3.20}$$

Next, set $W_i = E(\bar\eta\bar{\bar\eta}|Y_0 = i)$, where $i = 1,\ 2$. Once again, by the first step analysis, one easily obtains that

$$W_1 = \bar p W_1 + \bar q W_2 + V_1, \quad W_2 = \bar{\bar q} W_1 + \bar{\bar p} W_2 + U_2. \tag{3.21}$$

Solving (3.21) gives

$$W_1 = V_1(U_1+U_2), \quad W_2 = U_2(V_1+V_2) \tag{3.22}$$

and hence

$$\mathrm{Cov}(\bar\eta, \bar{\bar\eta}|Y_0=1) = \mathrm{Cov}(\bar\eta, \bar{\bar\eta}|Y_0=2) = V_1U_2. \tag{3.23}$$

It is a standard fact that $\bar\eta$ conditioned on $Y_0 = 1$ has geometric distribution whose parameter is thus U_1^{-1}. If our Markov chain starts from 1 it must visit 2 before being absorbed by 3. Hence the distribution of $\bar{\bar\eta}$ conditioned on $Y_0 = 1$ is the same as the distribution of $\bar{\bar\eta}$ conditioned on $Y_0 = 2$ and is geometric with parameter $V_2^{-1} = \varepsilon$. We therefore have that $\mathrm{Var}(\bar\eta|Y_0 = 1) = U_1^2 - U_1$ and $\mathrm{Var}(\bar{\bar\eta}|Y_0 = 1) = V_2^2 - V_2$. We can now compute the correlation coefficient of $\bar\eta$ and $\bar{\bar\eta}$ which, taking into account (3.20), can be presented as follows:

(3.24)

$$\mathrm{Corr}(\bar\eta, \bar{\bar\eta}|Y_0=1) = \frac{V_1U_2}{\sqrt{(U_1^2-U_1)(V_2^2-V_2)}} = \frac{\bar{\bar q}}{\bar{\bar q}+\varepsilon}(1-U_1^{-1})^{-\frac12}(1-V_2^{-1})^{-\frac12}.$$

This formula implies lower and upper bounds for correlations in two different regimes: (a) when $\bar{\bar q}/\varepsilon \to 0$ and (b) when $\varepsilon \to 0$ while $\bar q$, $\bar{\bar q}$ remain separated from 0. Here is the precise statement we need.

Lemma 3.9. *(a) Suppose that $U_1 \geq 1+c$, $V_2 \geq 1+c$, where $c>0$. Then*

$$\mathrm{Corr}(\bar\eta, \bar{\bar\eta}|Y_0=1) \leq \mathrm{Const}\, \frac{\bar{\bar q}}{\varepsilon} \equiv \mathrm{Const}\, \bar{\bar q} V_1. \tag{3.25}$$

where the constant in this formula depends only on c.

(b) If $\bar{\bar q} \geq c$ and $\bar q \geq c$ for some $c>0$ then for ε small enough, or, equivalently, U_1 large enough

$$\mathrm{Corr}(\bar\eta, \bar{\bar\eta}|Y_0=1) \geq 1-\frac{\varepsilon}{c}, \quad \mathrm{Corr}(\bar\eta, \bar{\bar\eta}|Y_0=1) \geq 1-\frac{1}{cU_1}. \tag{3.26}$$

Proof. (a) Inequality (3.25) is an immediate corollary of (3.24).

(b) (3.24) can be written as

$$\mathrm{Corr}(\bar\eta, \bar{\bar\eta}|Y_0=1) = \frac{\bar{\bar q}}{\bar{\bar q}+\varepsilon}\left(1-\frac{\varepsilon \bar q}{\bar{\bar q}+\varepsilon}\right)^{-\frac12}(1-\varepsilon)^{-\frac12}.$$

If $\frac{\varepsilon}{\bar{\bar q}} < 1$ then it follows from here that

$$\mathrm{Corr}(\bar\eta, \bar{\bar\eta}|Y_0=1) = 1-\left(1-\frac{\bar q+\bar{\bar q}}{2}\right)\frac{\varepsilon}{\bar{\bar q}} + \mathcal{O}\left(\left(\frac{\varepsilon}{\bar{\bar q}}\right)^2\right). \tag{3.27}$$

Due to (3.20) and conditions of the Lemma we have $\varepsilon = \frac{\bar{\bar q}}{\bar q}\left(U_1^{-1}+\mathcal{O}(U_1^{-2})\right)$ and hence

$$\mathrm{Corr}(\bar\eta, \bar{\bar\eta}|Y_0=1) = 1-\left(1-\frac{\bar q+\bar{\bar q}}{2}\right)\frac{1}{\bar q U_1} + \mathcal{O}(U_1^{-2}). \tag{3.28}$$

(3.26) is now a simple corollary of (3.27) and (3.28). □

The next two lemmas follow from Lemma 3.9 and reflect the fact that the correlations between the number of visits to nearby sites are strong but decay rapidly as the spatial distance increases.

Lemma 3.10. *There is a $C>0$ such that for* **P***-almost all ω and $n \geq 0$*

$$\mathrm{Corr}_\omega(\xi_n, \xi_{n+1}) \geq 1-\frac{C}{\rho_n}. \tag{3.29}$$

Proof. Let ω be such that the random walk X runs away to $+\infty$ with $\mathbb{P}_\omega$ probability 1 (which is the case for **P**-almost all ω). For a given $n \geq 0$ consider a Markov chain $Y=\{Y_t,\ t\geq 0\}$, with the state space $\{n,\ n+1,\ as\}$, where $n,\ n+1$ are sites on $\mathbb{Z}$ and as is an absorbing state. Let $k_0<k_1<\ldots<k_\tau$ be the sequence of all moments such that $X_{k_j}\in\{n,n+1\}$; we set $Y_t=X_{k_t}$ if $t\leq\tau$ and $Y_t=as$ if $t>\tau$.

It is easy to see that the transition matrix of Y is as in (3.19) with transition probabilities given by

$$\bar{p} = q_n, \quad \bar{q} = p_n, \quad \bar{\bar{q}} = q_{n+1},$$

$$\bar{\bar{p}} = \mathbb{P}_\omega\{X_k \text{ starting from } n+1 \text{ returns to } n+1 \text{ before visiting } n\},$$

$$\varepsilon = \mathbb{P}_\omega\{X_t \text{ starting from } n+1 \text{ never returns to } n+1\}.$$

Also, in this context, $\bar{\eta} = \xi_n$, $\bar{\bar{\eta}} = \xi_{n+1}$ and hence $V_1 = \rho_n$. Next, $\bar{q}$, $\bar{\bar{q}}$ are separated from 0 because of condition (C) from Section 1. All conditions of Lemma 3.9 are thus satisfied and hence, for ρ_ns which are sufficiently large, (3.29) follows from (3.26). □

Lemma 3.11. *(a) There exist sets Ω_N, $K > 0$ such that $\mathbf{P}(\Omega_N^c) \leq N^{-100}$ and if $\omega \in \Omega_N$ then for all $0 \leq n_1,\ n_2 \leq N$ with $n_2 > n_1 + K \ln N$ we have*

$$\mathrm{Corr}_\omega(\xi_{n_1}, \xi_{n_2}) \leq N^{-100}.$$

(b) If K is sufficiently large then for each N there exist random variables $\{\bar{\xi}_n\}_{n=0}^N$ such that for each $\omega \in \Omega_N$ and any sequence $0 \leq n_1 < n_2 \cdots < n_k \leq N$ with $n_{j+1} > n_j + K \ln N$, the variables $\{\bar{\xi}_{n_j}\}_{j=0}^k$ are mutually independent and

$$\mathbf{P}(\bar{\xi}_n = \xi_n \textit{ for } n = 0, \ldots, N) \geq 1 - \frac{C}{N^{100}}. \tag{3.30}$$

Proof. (a) Consider a Markov chain Y which is defined as in the proof of Lemma 3.10 with the difference that its state space is $\{n_1, n_2, as\}$ and that $\bar{\eta} = \xi_{n_1}$, $\bar{\bar{\eta}} = \xi_{n_2}$. Then by (3.25)

$$\mathrm{Corr}_\omega(\xi_{n_1}, \xi_{n_2}) \leq \mathrm{Const}\, \bar{\bar{q}} \rho_{n_2}.$$

But, by Lemma 3.6, $\rho_n \leq N^{\frac{103}{s}}$ except for the set of measure $\mathcal{O}(N^{-103})$. Now Lemma 3.4 guarantees that we can choose K so that if the sites are separated by $K \ln N$ then $\bar{\bar{q}} < N^{-(101+103/s)}$ except for the set of measure $\mathcal{O}(N^{-103})$. This proves (a) for fixed n_1, n_2 on a set of measure $\geq 1 - \mathcal{O}(N^{-103})$ which in turn implies the wanted result.

(b) Let $\bar{\xi}_n$ be the number of visits to the site n before the first visit to $n + \frac{K \ln N}{2}$. It follows from this definition that $\{\bar{\xi}_{n_j}\}_{j=0}^k$ are mutually independent. Next,

$$\mathbf{P}(\bar{\xi}_n = \xi_n) \leq \mathbf{P}(X \text{ visits } n \text{ after } n + 0.5K \ln N)$$

Now (3.30) follows from Lemma 3.4. □

4. Proof of Theorem 2.

Our goal is to show that the main contribution to T_N comes from the terms where ρ_n is large. However, the set where ρ_n is large has an additional structure. Namely, if ρ_n is large the same is true for $\rho_{n\pm 1}$ and more generally for ρ_{n_1} and ρ_{n_2} when n_1 and n_2 are close to n; this implies that the corresponding ξ_{n_1} and ξ_{n_2} are strongly correlated. But if n_1 and n_2 are far apart then ρ_{n_1} and ρ_{n_2}, and also ξ_{n_1} and ξ_{n_2}, are almost independent. In the arguments below we need to take care about this additional structure.

But first we show that terms where $\rho_n < \delta N^{1/s}$ can be neglected.

The following convention will be used throughout this section: the summations are over suitable n, n_1, n_2 in $[0, N-1]$.

Lemma 4.1. *Let $\delta > 0$. Then there is N_δ such that for $N > N_\delta$ the following holds:*

(a) If $0 < s < 1$ then

$$\mathbf{E}\left(\sum_{\rho_n < \delta N^{1/s}} \xi_n\right) \leq \mathrm{Const} N^{1/s} \delta^{1-s}.$$

(b) If $1 < s < 2$ then there is a set $\tilde{\Omega}_{N,\delta}$ such that $\mathbf{P}(\tilde{\Omega}^c_{N,\delta}) \leq N^{-100}$ and

$$\mathbf{E}\left(1_{\tilde{\Omega}_{N,\delta}} \mathbb{E}_\omega\left[\left(\sum_{\rho_n < \delta N^{1/s}} (\xi_n - \rho_n)\right)^2\right]\right) \leq \mathrm{Const} N^{2/s} \delta^{2-s}.$$

(c) If $0 < s < 1$ then

$$\mathbf{E}\left(\sum_{\rho_n < \delta N^{1/s}} \rho_n\right) \leq \mathrm{Const} N^{1/s} \delta^{1-s}.$$

(d) If $1 < s < 2$ then

$$\mathrm{Var}\left(\sum_{\rho_n < \delta N^{1/s}} \rho_n\right) \leq \mathrm{Const} N^{2/s} \delta^{2-s}.$$

(e) If $s = 1$ then given $\frac{1}{2} < \kappa < 1$ there is a set $\tilde{\Omega}_{N,\delta}$ such that $\mathbf{P}(\tilde{\Omega}^c_{N,\delta}) \leq N^{-100}$ and

$$\mathbf{E}\left(1_{\tilde{\Omega}_{N,\delta}}\left(\mathrm{Var}_\omega\left(\sum_{\rho_n < \delta N} (\xi_n - \rho_n)\right)\right)^\kappa\right) \leq \mathrm{Const} N^{2\kappa} \delta^{2\kappa - 1}, \tag{4.1}$$

$$\mathbf{E}\left(\left(\sum_{\rho_n<\delta N}(\rho_n-\mathbf{E}\left(\rho I_{\rho<\delta N}\right))\right)^2\right)\leq \text{Const}N^2\delta. \tag{4.2}$$

Proof. Set $\chi_n = I_{\rho_n<\delta N^{1/s}}$; this concise notation will be used only within the proof of Lemma 4.1.

(a) Denote $Y_\delta = \sum_{\rho_n<\delta N^{1/s}} \xi_n$. Then

$$\mathbf{E}(Y_\delta) = N\mathbf{E}(\rho I_{\rho<\delta N^{1/s}}).$$

By Lemma 3.6 this expectation is bounded by $\text{Const}N^{1/s}\delta^{1-s}$ proving our claim.

(b) Denote $\tilde{Y}_\delta = \sum_{\rho_n<\delta N^{1/s}}(\xi_n - \rho_n)$. Then $\mathbb{E}_\omega(\tilde{Y}_n) = 0$ and so it suffices to show that $\text{Var}_\omega(\tilde{Y}_\delta) = o(\delta^{2-s}N^{2/s})$ except for ω from a set of small probability. Due to Lemma 3.11 for most ωs we have

(4.3)

$$\begin{aligned}\text{Var}_\omega(\tilde{Y}_\delta) &= \left|o(1) + \sum_{n_2-K\ln N<n_1<n_2} 2\chi_{n_1}\chi_{n_2}\text{Cov}_\omega\left(\xi_{n_1}\xi_{n_2}\right) + \sum_n \chi_n \text{Var}_\omega\left(\xi_n\right)\right| \\ &\leq 1 + \text{Const}\sum_{n_2-K\ln N<n_1\leq n_2} \rho_{n_1}\rho_{n_2}\chi_{n_1}\chi_{n_2}\end{aligned}$$

where the summation is over pairs with $\rho_{n_i} < \delta N^{1/s}$. The last step uses Cauchy-Schwartz inequality and the fact that ξ_n has geometric distribution, namely $|\text{Cov}_\omega\left(\xi_{n_1}\xi_{n_2}\right)| \leq \sqrt{\text{Var}_\omega\left(\xi_{n_1}\right)\text{Var}_\omega\left(\xi_{n_2}\right)} \leq \rho_{n_1}\rho_{n_2}$.

Next, we estimate the expectation of the last sum in (4.3). Using (3.7) we can write

$$\rho_{n-k}\rho_n = p_{n-k}^{-1}\rho_n^2 q_n \alpha_{n-1}\dots\alpha_{n-k+1} + (\alpha_{n-1}\dots\alpha_{n-k+1}+\dots+1)p_{n-k}^{-1}\rho_n.$$

Since ρ_n and $\{\alpha_j,\ j<n\}$ are independent we obtain

$$\mathbf{E}\left(\rho_{n-k}\rho_n\chi_n\right) \leq \text{Const}\left[\beta^{k-1}\mathbf{E}\left(\rho_n^2\chi_n\right) + \mathbf{E}\left(\rho_n\chi_n\right)\sum_{j=0}^{k-2}\beta^j\right],$$

where $\beta = \mathbf{E}(\alpha) < 1$. Thus

$$\mathbf{E}\left(\sum_{k=0}^{K\ln N}\left(\rho_{n-k}\rho_n\chi_n\right)\right) \leq \text{Const}\left[\mathbf{E}\left(\rho_n^2\chi_n\right) + \ln N\mathbf{E}\left(\rho_n\chi_n\right)\right]. \tag{4.4}$$

Hence

$$\begin{aligned}\mathbf{E}\left(\sum_{n_2-K\ln N<n_1\leq n_2}\rho_{n_1}\rho_{n_2}\chi_{n_1}\chi_{n_2}\right) &\leq \mathbf{E}\left(\sum_{n_2-K\ln N<n_1\leq n_2}\rho_{n_1}\rho_{n_2}\chi_{n_2}\right) \\ \leq \text{Const}\sum_{n_2}\left[\mathbf{E}\left(\left(\rho_{n_2}\right)^2\chi_{n_2}\right) + \ln N\mathbf{E}(\rho_{n_2})\right] &\leq \text{Const}\delta^{2-s}N^{2/s}.\end{aligned}$$

(c) The proof of (c) is the same as proof of (a).

(d) We assume first that

$$\nu([\delta N^{1/s}-N^{-100},\delta N^{1/s}+N^{-100}])\leq N^{-50} \tag{4.5}$$

where ν is a stationary distribution of ρ_n.

We claim that if (4.5) holds and $n_2>n_1+\bar{K}\ln N$ then

$$\mathrm{Cov}\,(\rho_{n_1}\chi_{n_1},\rho_{n_2}\chi_{n_2})<\frac{1}{N^3} \tag{4.6}$$

provided that $\bar{K}$ is large enough.

Indeed let $\{\hat{\rho}_n\}_{n\in[n_1,n_2-1]}$ be a sequence such that $\hat{\rho}_{n_2}$ has distribution ν and is independent of $\{\rho_n\}$,

$$\hat{\rho}_n=p_n^{-1}q_{n+1}\hat{\rho}_{n+1}+p_n^{-1}$$

for $n\leq n_2-1$ and

$$\mathbf{P}(|\rho_{n_1}-\hat{\rho}_{n_1}|>\beta_1^{\bar{k}})<\beta_2^{\bar{k}} \tag{4.7}$$

with $\bar{k}=\bar{K}\ln N-1$ (existence of such a sequence follows from Lemma 3.8). Denote

$$\eta_n=\rho_n\chi_n,\quad \hat{\eta}_n=\hat{\rho}_n I_{\hat{\rho}_n<\delta N^{1/s}},\quad A=\{|\eta_{n_1}-\hat{\eta}_{n_1}|>N^{-100}\}.$$

Note that by (4.5) $A\subset A'\bigcup A''$, where $A'=\{|\rho_{n_1}-\hat{\rho}_{n_1}|>N^{-100}\}$ so that by (4.7)

$$\mathbf{P}(A')\leq N^{-100},$$

provided that $\bar{K}$ is large enough, and $A''=\{|\hat{\rho}_{n_1}-\delta N^{1/s}|<N^{-100}\}$ so that $\mathbf{P}(A'')\leq N^{-50}$ due to (4.5).

Since $\hat{\eta}_{n_1}$ is independent of η_{n_2} we have

$$|\mathrm{Cov}(\eta_{n_2},\eta_{n_1})|=|\mathrm{Cov}(\eta_{n_2},\eta_{n_1}-\hat{\eta}_{n_1})|=|\mathbf{E}(\eta_{n_2}(\eta_{n_1}-\hat{\eta}_{n_1}))|$$

$$\leq\mathbf{E}(\eta_{n_2}|\eta_{n_1}-\hat{\eta}_{n_1}|I_{A^c})+\mathbf{E}(\eta_{n_2}|\eta_{n_1}-\hat{\eta}_{n_1}|I_{A'})+\mathbf{E}(\eta_{n_2}|\eta_{n_1}-\hat{\eta}_{n_1}|I_{A''}).$$

Since $\eta_n<\delta N^{1/s},\hat{\eta}_n<\delta N^{1/s}$ the first term is at most $\delta N^{1/s}N^{-100}$, the second term is at most $\delta^2N^{2/s}\mathbf{P}(A')\leq\delta^2N^{2/s}N^{-100}$ and the third term is at most $\delta^2N^{1/s}\mathbf{P}(A'')\leq\delta^2N^{2/s}N^{-50}$ proving (4.6).

(4.6) implies that if $\bar{K}$ is sufficiently large then

$$\mathrm{Var}\left(\sum_n\rho_n\chi_n\right)\leq 1+\left|\sum_{|n_1-n_2|<\bar{K}\ln N}\mathrm{Cov}(\rho_{n_1}\chi_{n_1},\rho_{n_2}\chi_{n_2})\right|.$$

The estimate of the last sum is exactly the same as in part (b). This completes the proof of part (d) under the assumption that (4.5) holds.

To prove (4.6) without this assumption take $\bar{\delta}=\bar{\delta}_N$ satisfing (4.5) and such that

$$\delta-2N^{-50}<\bar{\delta}<\delta.$$

(The existence of $\bar{\delta}$ with the required properties follows from the fact that we can partition $[\delta - 2N^{-50}, \delta]$ into disjoint segments of length $2N^{-100}$ and note that by the pigeonhole principle all segments can not have large measure).

Split $\eta_n = \bar{\eta}_n + \bar{\bar{\eta}}_n$ where

$$\bar{\eta}_n = \rho_n I_{\rho_n < \bar{\delta} N^{1/s}}, \quad \bar{\bar{\eta}}_n = \rho_n I_{\bar{\delta} N^{1/s} \leq \rho_n < \delta N^{1/s}}.$$

Then since the cut off at $\bar{\delta} N^{1/s}$ satisfies (4.5) we have

$$\mathrm{Var}(\sum_n \bar{\eta}_n) \leq C \delta^{2/s} N^{2/s}$$

so it remains to show that

$$\mathrm{Var}(\sum_n \bar{\bar{\eta}}_n) \leq C \delta^{2/s} N^{2/s}. \tag{4.8}$$

Introducing

$$\tilde{\eta} = \hat{\rho}_n I_{\bar{\delta} N^{1/s} \leq \rho_n < \delta N^{1/s}}, \quad \tilde{A}'' = \{|\tilde{\eta} - \bar{\delta} N^{1/s}| < N^{100} \text{ or } |\tilde{\eta} - \delta N^{1/s}| < N^{100}\}$$

and proceeding as in the proof of (4.6) we see that if $|n_2 - n_1| > \bar{K} \ln N$ then

$$|\mathrm{Cov}(\bar{\bar{\eta}}_{n_2}, \bar{\bar{\eta}}_{n_1})| \leq 2\delta N^{2/s} N^{-100} + \mathbf{E}(\bar{\bar{\eta}}_{n_2} |\bar{\bar{\eta}}_{n_1} - \tilde{\eta}_{n_1}| I_{\tilde{A}''}).$$

The last term can be bounded by

$$\delta N^{1/s} \mathbf{E}(\bar{\bar{\eta}}_{n_2} I_{\tilde{A}''}) = \delta N^{1/s} \mathbf{E}(\bar{\bar{\eta}}_{n_2}) \mathbf{P}(\tilde{A}'')$$

where we have used independence of $\bar{\bar{\eta}}_{n_2}$ and $\tilde{\eta}_{n_1}$. Next by Lemma 3.6

$$\mathbf{E}(\bar{\bar{\eta}}_{n_2}) \leq \delta N^{1/s} \mathbf{P}(\bar{\bar{\eta}}_{n_2} \neq 0) \leq C \delta N^{(1/s)-1}$$

and $\mathbf{P}(\tilde{A}'') \leq \frac{C}{\delta N}$. Accordingly

$$|\mathrm{Cov}(\bar{\bar{\eta}}_{n_2}, \bar{\bar{\eta}}_{n_1})| \leq C \delta N^{(2/s)-2}$$

so that

$$\sum_{|n_2 - n_1| > \bar{K} \ln N} |\mathrm{Cov}(\bar{\bar{\eta}}_{n_2}, \bar{\bar{\eta}}_{n_1})| \leq C \delta N^{(2/s)}.$$

On the other hand arguing as before we obtain

$$\sum_{|n_2 - n_1| \leq \bar{K} \ln N} |\mathrm{Cov}(\bar{\bar{\eta}}_{n_2}, \bar{\bar{\eta}}_{n_1})| \leq C \delta^{2-s} N^{(2/s)}.$$

This proves (4.8) completing the proof of part (d) of Lemma 4.1.

(e) In view of (4.3) in order to prove (4.1) it suffices to estimate

$$\mathbf{E}\left(\sum_{n_1 \leq n_2 < n_1 + K \ln N} \chi_{n_1} \chi_{n_2} \rho_{n_1} \rho_{n_2}\right)^{\kappa}$$

We have

$$\mathbf{E}\left(\sum_{n_1\le n_2<n_1+K\ln N}\chi_{n_1}\chi_{n_2}\rho_{n_1}\rho_{n_2}\right)^\kappa \le \sum_{n_1\le n_2<n_1+K\ln N}\mathbf{E}\left(\chi_{n_1}\chi_{n_2}\left(\rho_{n_1}\rho_{n_2}\right)^\kappa\right).$$

Using that $\mathbf{E}(\alpha^\kappa)<1$ we can proceed as in part (b) to estimate the last sum by

$$C\sum_n\left[\mathbf{E}\left((\rho_n)^{2\kappa}\chi_n\right)+(\ln N)\mathbf{E}\left((\rho_n)^\kappa\chi_n\right)\right]\le \tilde{C}N^{2\kappa}\delta^{2\kappa-1}.$$

This completes the proof of (4.1).

Next we prove (4.2). We assume that (4.5) holds, this assumption can be removed in the same way as it was done in part (d). Since the estimate of $\mathrm{Cov}(\eta_{n_1},\eta_{n_2})$ in case $|n_2-n_1|>\bar{K}\ln N$ did not use the fact that $s>1$ we need to bound

$$\sum_{|n_1-n_2|\le\bar{K}\ln N}\mathrm{Cov}(\eta_{n_1},\eta_{n_2})=\sum_{|n_1-n_2|\le\bar{K}\ln N}\mathbf{E}(\eta_{n_1},\eta_{n_2})+o(N\ln N).$$

Without loss of generality we can assume that $n_1\le n_2$. To simplify notation we put $n_2=n, n_1=n-k$. Using the same notation as in (3.8) it is enough to bound

$$\mathbf{E}((\rho_n^2A_{n,k}+\rho_nB_{n,k})\chi_n\chi_{n-k}). \tag{4.9}$$

Since $\mathbf{E}(\alpha)=1$ we have $\mathbf{E}(B_{n,k})\le Ck$. Since $B_{n,k}$ is independent of ρ_n the second term in (4.9) can be bounded by

$$\mathbf{E}(\rho_n\chi_n)\mathbf{E}(B_{n,k})\le C\delta\ln Nk\le\tilde{C}\delta\ln^2 N$$

so the main contribution come from the first term. We shall use the fact that there are constants $\varepsilon>0$ and $\bar{\beta}<1$ such that for some constant C

$$\mathbf{E}(A_{n,k}I_{A_{n,k}<e^{\varepsilon k}})<C\bar{\beta}^k. \tag{4.10}$$

We split

$$\mathbf{E}(\rho_n^2A_{n,k}\chi_n\chi_{n-k})=\mathbf{E}(\rho_n^2A_{n,k}I_{\rho_n<\delta Ne^{-\varepsilon k}}\chi_{n-k})+\mathbf{E}(\rho_n^2A_{n,k}I_{\delta Ne^{-\varepsilon k}\le\rho_n<\delta N}\chi_{n-k}).$$

The first term is bounded by

$$\mathbf{E}(\rho_n^2I_{\rho_n<\delta Ne^{-\varepsilon k}})\mathbf{E}(A_{n,k})=\mathbf{E}(\rho_n^2I_{\rho_n<\delta Ne^{-\varepsilon k}})\le C\delta Ne^{-\varepsilon k}.$$

Since $\rho_{n-k}\ge A_{n,k}\rho_n$ the second term is bounded by

$$\mathbf{E}(\rho_n^2\chi_nA_kI_{A_{N,k}<e^{\varepsilon k}})=\mathbf{E}(\rho_n^2\chi_n)]\mathbf{E}(A_kI_{A_{N,k}<e^{\varepsilon k}})\le C\delta N\bar{\beta}^k.$$

Summing over n and k we get

$$\sum_{n=1}^N\sum_{k\le\bar{K}\ln N}\mathbf{E}(\rho_n^2A_{n,k}+\rho_nB_{n,k}\chi_n\chi_{n-k})\le C\delta N^2$$

as claimed. It remains to establish (4.10). The estimate (3.12) implies that there exists $\hat{\varepsilon} > 0$ and $\hat{\beta} < 1$ such that

$$\mathbf{P}(A_{n,k} > e^{-\hat{\varepsilon}k}) \leq \hat{\beta}^k.$$

Hence

$$\mathbf{E}(A_{n,k} I_{A_{n,k}<e^{\varepsilon k}}) = \mathbf{E}(A_{n,k} I_{A_{n,k}<e^{-\hat{\varepsilon}k}}) + \mathbf{E}(A_{n,k} I_{e^{-\hat{\varepsilon}k} \leq A_{n,k}<e^{\varepsilon k}}) \leq e^{-\hat{\varepsilon}k} + e^{\varepsilon k}\hat{\beta}^k$$

which decays exponentially provided that ε is small enough. This proves (4.10). □

Lemma 4.1 allows us to concentrate on sites where $\rho_n \geq \delta N^{1/s}$. In view of Lemma 3.6 for each fixed δ we expect to have finitely many such points on $[0, N]$ (namely the expected number of points is $\mathcal{O}(\delta^{-s})$).

Definition. Let $M = M_N := \ln \ln N$. We shall say that n is a *massive* site if $\rho_n \geq \delta N^{1/s}$. A site $n \in [0, N-1]$ is *marked* if it is massive and $\rho_{n+j} < \delta N^{1/s}$ for $1 \leq j \leq M$. For n marked the interval $[n-M, n]$ is called the *cluster* associated to n.

Note that this definition implies that the distinct clusters are disjoint.

It may happen that not all massive sites belong to one of the clusters. This situation is controlled by the following

Lemma 4.2. *There is $\beta < 1$ such that for $n \in [0, N-1]$*

$$\mathbf{P}\left(\rho_n \geq \delta N^{1/s} \text{ and } n \text{ is not in a cluster}\right) \leq \text{Const}\frac{\beta^M}{N}. \tag{4.11}$$

Proof. Suppose that n is a massive point which is not in a cluster. Then consider all massive points n_i such that $n < n_1 < ... < n_k < n + M$. Note that such points exist because otherwise n would have been a marked point. Let now $n^* > n_k$ be the nearest to n_k massive point. Then by construction $n^* \geq n+M$. Also $n^* \leq n+2M$ because otherwise n_k would have been a marked point and n would belong to the n_k-cluster. Hence the event

$$\{n \text{ is massive and not in a cluster}\} \subset \bigcup_{n' \in [n+M, n+2M]} \{\rho_n \geq \delta N^{1/s},\ \rho_{n'} \geq \delta N^{1/s}\}.$$

By Lemma 3.7(b) we obtain

$$\mathbf{P}\left(n \text{ is massive and not in a cluster}\right) \leq \sum_{n'=n+M}^{n+2M} \mathbf{P}\left(\rho_n \geq \delta N^{1/s},\ \rho_{n'} \geq \delta N^{1/s}\right) \leq \text{Const}\frac{\beta^M}{N}$$

which proves our statement. □

It is clear from the just presented proof that the event

{there is n which is massive and not in a cluster}

belongs to the set of environments

(4.12)
$$\bar{\Omega}_N^\delta := \{\exists n_1, n_2 : M < n_2 - n_1 \le 2M \text{ and } \rho_{n_1} \ge \delta N^{1/s}, \rho_{n_2} \ge \delta N^{1/s}\}.$$

Then again by Lemma 3.7(b) we have that

(4.13)
$$\mathbf{P}\left(\bar{\Omega}_N^\delta\right) \le \sum_{n=0}^{N-1} \sum_{n_1=n-2M}^{n-M} \mathbf{P}\{\rho_{n_1} \ge \delta N^{1/s},\, \rho_n \ge \delta N^{1/s}\} \le \text{Const}\beta^M.$$

Obviously $\mathbf{P}(\bar{\Omega}_N^\delta) \to 0$ as $N \to \infty$.

It is clear from the definitions that $\mathbf{P}$(n is massive and in a cluster) $\ge$ $\mathbf{P}$(n is marked). The following lemma shows that in fact these quantities are of the same order of smallness.

Lemma 4.3.

(4.14)
$$\mathbf{P}\left(\rho_n \ge \delta N^{1/s} \textit{ and } n \textit{ is in a cluster}\right) \le \text{Const}\mathbf{P}\left(n \textit{ is marked}\right).$$

Proof. The event

$$\{n \text{ is massive and in a cluster}\} \subset \bigcup_{k=0}^{M} \{\rho_n \ge \delta N^{1/s},\, n+k \text{ is marked}\}.$$

Since ρ_n is a stationary sequence we have

$$\mathbf{P}\{\rho_n \ge \delta N^{1/s},\, n+k \text{ is marked}\} =$$
$$\mathbf{P}\{\rho_{n-k} \ge \delta N^{1/s} \,|\, n \text{ is marked}\} \times \mathbf{P}\{n \text{ is marked}\}.$$

Fix $h \in (0, s)$ and let $\beta = \mathbf{E}(\alpha^h)$. Note that $\beta < 1$. We shall now prove that

$$\mathbf{P}\{\rho_{n-k} \ge \delta N^{1/s} \,|\, n \text{ is marked}\} \le \text{Const}\beta^k.$$

Since M is growing very slowly we have for $N \ge N_\delta$ that $M\varepsilon_0^{-M} \ll 0.5\delta N^{1/s}$. Then (3.9) implies that $\rho_{n-k} \le \bar{c}A_{n,k}\rho_n + 0.5\delta N^{1/s}$ and therefore

$$\mathbf{P}\{\rho_{n-k} \ge \delta N^{1/s} \,|\, n \text{ is marked}\} \le \mathbf{P}\{\bar{c}A_{n,k}\rho_n > 0.5\delta N^{1/s} | \, n \text{ is marked}\}.$$

For n marked $\rho_{n+1} < \delta N^{1/s}$ and hence $\rho_n < 2\varepsilon_0^{-1}\delta N^{1/s}$. Since $A_{n,k}$ and $\{\rho_j\}_{j \ge n}$ are independent we have, with $C = 2\bar{c}\varepsilon_0^{-1}$:

$$\mathbf{P}\{\bar{c}A_{n,k}\rho_n > 0.5\delta N^{1/s} | \, n \text{ is marked}\}$$
$$\le \mathbf{P}\{CA_{n,k}\delta N^{1/s} > 0.5\delta N^{1/s} | \, n \text{ is marked}\} = \mathbf{P}\{A_{n,k} > 0.5C^{-1}\} \le \text{Const}\beta^k.$$

(Once again, last step is due to the Markov inequality.) Finally we obtain

$$\mathbf{P}\,(n \text{ is massive and is in a cluster}) \leq$$

$$\leq \mathrm{Const}(\sum_{k=0}^{M} \beta^k) \times \mathbf{P}\{n \text{ is marked}\} \leq \mathrm{Const}\,\mathbf{P}\{n \text{ is marked}\}.$$

□

We shall now turn to the analysis of the properties of clusters. The next lemma is the main technical result of the paper. It will be proved in Section 5. We need one more

Definition. For each marked point n, we set

(4.15)

$$a_n = \rho_n/\delta N^{1/s}, \quad b_n = \frac{\Sigma_{j=0}^{M}\rho_{n-j}}{\rho_n} \text{ and } m_n = \Sigma_{j=0}^{M}\rho_{n-j} = \delta N^{1/s} a_n b_n.$$

We call m_n the mass of the cluster.

Lemma 4.4. *For a given $\delta > 0$ the following holds:*

(a) The point process $\{(\frac{n}{N}, a_n, b_n) \;:\; n \text{ is marked }\}$ converges as $N \to \infty$ to a point process $\{(t_j, \tilde{a}_j, \tilde{b}_j)\}$ where t_j form a Poisson process with a constant intensity $\tilde{c}\delta^{-s}$.

(b) For a given (finite) collection $\{t_j\}$ the corresponding collection $\{(\tilde{a}_j, \tilde{b}_j)\}$ consists of i.i.d. random variables which are independent of $\{t_j\}$ (except that both collections have the same cardinality). The distributions of the pair $(\tilde{a}, \tilde{b})$ does not depend on δ.

(c) Consequently[2] *$\{(\frac{n}{N}, \frac{m_n}{N^{1/s}})\}$ converges to a Poisson process $\Lambda^\delta = \{(t_j, \Theta_j)\}$ on $[0, 1] \times [\delta, \infty)$. Moreover if $\bar{\Theta}_j = \frac{\Theta_j}{\delta}$ then the distribution of $\bar{\Theta}_j$ is independent of δ. Let ζ denote this distribution. Then there is a constant $\bar{c}$ such that*

$$\lim_{x \to +\infty} x^s \zeta(\bar{\Theta}_j > x) = \bar{c}. \tag{4.16}$$

Lemma 4.5. *As $\delta \to 0$ Λ^δ converges to a Poisson process Λ on $[0, 1] \times (0, \infty)$. Let $\{\Theta_j\}$ be the projection of Λ onto the second coordinate. Then there exists* $\mathbf{c}$ *such that $\{\Theta_j\}$ is a Poisson process with intensity $\frac{\mathbf{c}}{\theta^{1+s}}$.*

Proof. Consider the measure $\zeta_\delta(A) = \zeta(G_\delta(A))$, where $G_\delta(\theta) = (\frac{\theta}{\delta})$ and $G_\delta(A)$ denotes the image of A under G_δ. By Lemmas 4.4(c) and 3.1(a) Λ^δ is a Poisson process with measure λ_δ where $d\lambda_\delta = \tilde{c}\delta^{-s} dt d\zeta_\delta$.

[2] the first statement of part (c) of Lemma 4.4 follows from parts (a) - (b) and Lemma 3.1.

By (4.16), as $\delta \to 0$, $\lambda_\delta \Rightarrow \lambda$ where $d\lambda = \mathbf{c} dt \frac{d\theta}{\theta^{1+s}}$, where $\mathbf{c} = s\tilde{c}\bar{c}$, $\tilde{c}$ is the constant from Lemma 4.4(a) and $\bar{c}$ is the constant from Lemma 4.4(c). Hence, as $\delta \to 0$, $\Lambda^\delta \Rightarrow \Lambda$, where Λ is a Poisson process with measure λ. Now the result follows from Lemma 3.1(a). □

Remark. Once we know that Λ^δ has a limit as $\delta \to 0$ the intensity of the limiting process can also be found by a scaling argument. Namely, for each κ, we have $\Lambda = \lim_{\delta\to 0} \Lambda^{\kappa\delta}$. Λ^δ depends on δ in two ways. First its intensity is proportional to δ^{-s}. Second, $\Theta_j/\delta = \tilde{a}_j\tilde{b}_j$. Recall that the distribution of $\tilde{a}_j\tilde{b}_j$ is independent of δ. Therefore replacing δ by $\kappa\delta$ replaces $\Theta \to \kappa\Theta$ and multiplies the intensity by κ^{-s}. In other words rescaling $\{\Theta_j\}$ by κ amounts to multiplying its intensity by κ^{-s}. Now the result follows from (3.1).

We are now in a position to finish the proof of Theorem 2. We shall do that in the case $0 < s < 1$. In all other cases the proof is similar.

Present the time spent by the walk in $[0, N)$ as

$$T_N = \sum_{n=0}^{N-1} \xi_n = S_1 + S_2 + S_3, \tag{4.17}$$

where

$$S_1 = \sum_{n:\, \rho_n < \delta N^{1/s},\, n \notin \text{any cluster}} \xi_n$$

$$S_2 = \sum_{n:\, \rho_n \geq \delta N^{1/s},\, n \text{ is not in a cluster}} \xi_n$$

$$S_3 = \sum_{n:\, n \text{ is in a cluster}} \xi_n.$$

By Lemma 4.1, (a) we have that $\mathbf{E}(S_1) \leq \text{Const} N^{1/s}\delta^{1-s}$. Next by (4.12), (4.13) we have that

$$\mathbf{P}(S_2 > 0) \leq \mathbf{P}\left(\bar{\Omega}_N^\delta\right) \to 0 \text{ as } N \to \infty.$$

We readily have that for $\omega \notin \bar{\Omega}_N^\delta$

$$\mathbf{t}_N = N^{-\frac{1}{s}} S_3 + N^{-\frac{1}{s}} S_1 = N^{-\frac{1}{s}} S_3 + R_N,$$

where $R_N := N^{-\frac{1}{s}} S_1$ and satisfies the requirements of (a), Theorem 2.

Next, consider S_3 which comes from the sum over the clusters and is the main contribution to T_N. Let us present it as follows:

$$N^{-\frac{1}{s}} S_3 = \sum_{n:\, n \text{ is marked}} N^{-\frac{1}{s}} \sum_{j=0}^{M} \xi_{n-j}.$$

In turn

$$\sum_{j=0}^{M}\xi_{n-j}=\sum_{j=1}^{M}\left(\frac{\xi_{n-j}}{\rho_{n-j}}-\frac{\xi_n}{\rho_n}\right)\rho_{n-j}+\frac{\xi_n}{\rho_n}\sum_{j=0}^{M}\rho_{n-j}$$

Next, using Lemma 3.10 and the fact that ξ_n is a geometric random variable and therefore $\mathrm{Var}_\omega(\xi_n)=\rho_n^2-\rho_n$ one obtains

$$\left\|\frac{\xi_{n-j}}{\rho_{n-j}}-\frac{\xi_n}{\rho_n}\right\|\leq\sum_{k=n-j}^{n-1}\left\|\frac{\xi_k}{\rho_k}-\frac{\xi_{k+1}}{\rho_{k+1}}\right\|\leq\text{Const}\sum_{k=n-j}^{n-1}\frac{1}{\sqrt{\rho_k}}.$$

Here and below $\|f\|:=\sqrt{\mathbb{E}_\omega(|f|^2)}$ with f being a function on the space of trajectories of the walk.

For $n-j$ belonging to a cluster, that is $(n-j)\in[n-M,n]$, we have that $\rho_{n-j}\geq c\varepsilon_0^M\rho_n\geq cN^{-\bar\varepsilon}\rho_n$. (Remember that if ε_1 in Lemma 3.7 is small enough then $\bar\varepsilon$ can be made very small which is what we shall use in this proof.) Thus

$$\left\|\sum_{j=1}^{M}\left(\frac{\xi_{n-j}}{\rho_{n-j}}-\frac{\xi_n}{\rho_n}\right)\rho_{n-j}\right\|\leq\text{Const}\frac{N^{\bar\varepsilon/2}}{\sqrt{\rho_n}}\sum_{j=1}^{M}M\rho_{n-j}\leq\text{Const}\frac{N^{\bar\varepsilon}}{\sqrt{\rho_n}}\sum_{j=1}^{M}\rho_{n-j}$$

If for n marked we set

$$\zeta_n=m_n^{-1}\sum_{j=1}^{M}\left(\frac{\xi_{n-j}}{\rho_{n-j}}-\frac{\xi_n}{\rho_n}\right)\rho_{n-j}$$

then $\|\zeta_n\|\leq\text{Const}\frac{N^{\bar\varepsilon}}{\sqrt{\rho_n}}\to 0$ as $N\to\infty$ and we have

$$\frac{\sum_{j=0}^{M}\xi_{n-j}}{N^{1/s}}=\left(\frac{\xi_n}{\rho_n}+\zeta_n\right)\frac{m_n}{N^{1/s}},$$

Next ξ_n/ρ_n is asymptotically exponential with mean 1 since ξ_n is geometric with parameter $1/\rho_n$. Also by Lemma 3.11(b) $\{\frac{\xi_n}{\rho_n}\}_{n \text{ is marked}}$ are asymptotically independent. On the other hand $\{\frac{m_n}{N^{1/s}}\}_{n\text{ is marked}}$ are asymptotically Poisson by Lemma 4.4. In other words,

$$(\hat\Theta^{N,\delta},\hat\Gamma^{N,\delta})=\left(\left\{\frac{m_n}{N^{1/s}},\frac{\xi_n}{\rho_n}+\zeta_n\right\}_{n\text{ is }\delta-\text{marked}}\right).$$

satisfy the condition of Theorem 2 except that (i) is replaced by the condition

$(\widetilde{i})$ $\Gamma_j^{N,\delta}$ converges weakly to a Poisson process on $[\delta,\infty)$ with measure μ^δ. Moreover for each $[s,t]\in(0,\infty)$ $\mu^\delta([s,u])\to\mu([s,u])$ as $\delta\to 0$ where $\mu([s,u])=\int_s^u\frac{\mathbf{c}}{x^{1+s}}dx$ (the last statement follows from Lemma 4.5).

This statement appears slightly weaker than we need but a general argument from real analysis allows to upgrade $(\widetilde{i})$ to (i).

Namely $(\widetilde{i})$ shows that for each ε and $\bar{\delta}$ there is a number $N(\varepsilon, \bar{\delta})$ such that for any collection of disjoint intervals $[s_1, u_1], \dots, [s_l, u_l]$ in $[\bar{\delta}, \infty)$ with $\mu([s_j, u_j]) > \varepsilon$ for each $k_1, \dots, k_l$ we have

$$\left|\mathbf{P}\left(\Lambda_{N,\bar{\delta}}(s_j, u_j) = k_j \quad \text{for all } j \in [1, l]\right) - \prod_{j=1}^{l} \frac{\lambda_{j,\bar{\delta}}^{k_j}}{k_j!} e^{-\lambda_{j,\bar{\delta}}}\right| < \varepsilon. \tag{4.18}$$

where $\Lambda_{N,\bar{\delta}}(s_j, u_j)$ is the number of $\bar{\delta}$-clusters with mass between s_j and u_j and $\lambda_{j,\bar{\delta}} = \mu^{\bar{\delta}}([s_j, u_j])$. Choose a sequence ε_k converging to 0 (for example, $\varepsilon_k = \frac{1}{k}$ will do) and let δ_k be such that for each $\bar{\delta} < \delta_k$ and each $[s, u] \in [\bar{\delta}, \infty)$ we have

$$|\mu^{\bar{\delta}}([s, u]) - \mu([s, u])| < \varepsilon_k. \tag{4.19}$$

Then for $N > N(\varepsilon_k, \delta_k)$ both (4.18) and (4.19) are valid. Let $k(N)$ be the largest number such that $N > N(\varepsilon_k, \delta_k)$. Then (4.18) and (4.19) imply that

$$(\Theta^{N,\delta}, \Gamma^{N,\delta}) = \left(\left\{\frac{m_n}{N^{1/s}}, \frac{\xi_n}{\rho_n} + \zeta_n\right\}_{n \text{ is } \delta_{k(N)}-\text{ marked}, \, m_n > \delta N^{1/s}}\right)$$

satisfies (i).

5. Poisson Limit for expected occupation times.

To understand the asymptotic properties of the distribution of a_n defined in (4.15) we need the following

Lemma 5.1. *(a) For each $m > 0$ and $y \geq 1$*

$$\begin{aligned} \mu^m(y) &:= \lim_{N\to\infty} N\mathbf{P}\left(\frac{\rho_n}{N^{1/s}} \geq \delta y, \rho_{n+1} < \delta N^{1/s}, \dots, \rho_{n+m} < \delta N^{1/s}\right) \\ &= \delta^{-s} c\mathbf{E}[(D_0^s y^{-s} - \max_{1\leq j\leq m} D_j^s) I_{\max_{1\leq j\leq m} D_j < D_0 y^{-1}}], \end{aligned} \tag{5.1}$$

where $D_j := p_{n+j}^{-1} \alpha_{n+j+1} \dots \alpha_{n+m}$ (and by convention $D_m := p_{n+m}^{-1}$).
(b) There exists $\mu^\infty(y) = \lim_{m\to\infty} \mu^m(y)$.
(c) $\mu^\infty(1) > 0$.

Proof. (a) We shall make use of (3.10) and the relation $\rho_n = p_n^{-1} z_n$. For a fixed m and $0 \leq j \leq m$ we can write

$$\rho_{n+j} = p_{n+j}^{-1} \alpha_{n+j+1} \dots \alpha_{n+m} z_{n+m} + \mathcal{O}(1).$$

The inequalities $\rho_n/N^{1/s} \geq \delta y$ and $\rho_{n+j} < \delta N^{1/s}$ in (5.1) are equivalent to

$z_{n+m} \geq N^{1/s}\delta(D_0^{-1}y + \mathcal{O}(N^{-1/s}))$ and $z_{n+m} < N^{1/s}\delta(D_j^{-1} + \mathcal{O}(N^{-1/s}))$

respectively. Thus

$$\mathbf{P}\left(\frac{\rho_n}{N^{1/s}} \geq \delta y, \rho_{n+1} < \delta N^{1/s}, \dots, \rho_{n+m} < \delta N^{1/s}\right) =$$
$$\mathbf{P}\left(\delta(D_0^{-1}y + \mathcal{O}(N^{-1/s}))N^{1/s} \leq z_{n+m} < N^{1/s}\delta \min_{1\leq j\leq m}(D_j^{-1} + \mathcal{O}(N^{-1/s}))\right).$$

Since z_{n+m} and $\{p_{n+j}\}_{j\leq m}$ are independent, we can compute the following limit by conditioning on $\{p_{n+j}\}_{j\leq m}$ and using Lemma 3.6:

$$\lim_{N\to\infty} N\mathbf{P}\left(\rho_n N^{-1/s} \geq y, \rho_{n+1} < \delta N^{1/s}, \dots, \rho_{n+m} < \delta N^{1/s} | \{p_{n+j}\}_{j\leq m}\right) =$$
$$\delta^{-s}c(D_0^s y^{-s} - \max_{1\leq j\leq m} D_j^s) I_{\max_{1\leq j\leq m} D_j < D_0 y^{-1}}$$

and the convergence is uniform in $\{p_{n+j}\}_{j\leq m}$.

To compute the limit (5.1), it remains to take the expectation with respect to $\{p_{n+j}\}_{j\leq m}$:

$$\lim_{N\to\infty} N\mathbf{P}\left(\rho_n N^{-1/s} \geq y, \rho_{n+1} < \delta N^{1/s}, \dots, \rho_{n+m} < \delta N^{1/s}\right) =$$
$$\delta^{-s}c\mathbf{E}[(D_0^s y^{-s} - \max_{1\leq j\leq m} D_j^s) I_{\max_{1\leq j\leq m} D_j < D_0 y^{-1}}].$$

This completes the proof of part (a).
(b) The probability $\mathbf{P}\left(\frac{\rho_n}{N^{1/s}} \in [c,d], \rho_{n+1} \leq \delta N^{1/s}, \dots, \rho_{n+m} \leq \delta N^{1/s}\right)$ is a monotonically decaying function of m. Hence the proof.
(c) If $\mu^\infty(1) = 0$ then $N\mathbf{P}(n \text{ is marked}) \to 0$ as $N \to \infty$. Then

$$\begin{aligned}\mathbf{P}(\rho_n \geq \delta N^{1/s}) &\leq \mathbf{P}(\rho_n \geq \delta N^{1/s} \text{ and } n \text{ is not in a cluster}) \\ &\quad + \mathbf{P}(\rho_n \geq \delta N^{1/s} \text{ and } n \text{ is in a cluster}) \\ &\leq \text{Const}\frac{\beta^M}{N} + \text{Const}\mathbf{P}(n \text{ is marked}),\end{aligned}$$

where the estimates for the first and second term are provided by Lemmas 4.2 and 4.3 respectively. But then $N\mathbf{P}(\rho_n \geq \delta N^{1/s}) \to 0$ as $N \to \infty$ contradicting Lemma 3.6. This proves (c). □

Lemma 5.1 gives the limiting distribution of $\tilde{a}$ in Lemma 4.4. Namely, $\mathbf{P}(\tilde{a} > y) = 1$ if $y \leq 1$ and for $y > 1$ we have
(5.2)
$$\mathbf{P}(\tilde{a} > y) = \lim_{N\to\infty} \mathbf{P}(\rho_n > N^{1/s}\delta y \,|\, n \text{ is marked}) = \mu^\infty(y)/\mu^\infty(1).$$

Next we address the distribution of $\tilde{b}$.

Lemma 5.2. *(a) The distribution of* $\frac{\sum_{j=0}^{M}\rho_{n-j}}{\rho_n}$ *conditioned on* $\rho_n \geq \delta N^{1/s}$ *converges as* $N\to\infty$ *to the distribution of*

$$1+p_{-1}^{-1}q_0+p_{-2}^{-1}q_0\alpha_{-1}+\cdots+p_{-k}^{-1}q_0\alpha_{-1}\dots\alpha_{-k+1}+...$$

(b) Let

$$\bar{b}_j=p_{j-1}^{-1}+p_{j-2}\alpha_{j-1}+p_{j-3}\alpha_{j-1}\alpha_{j-2}+\dots \tag{5.3}$$

so that the distribution of $\tilde{b}$ *is the same as the distribution of* $1+q_0\bar{b}_0$. *Then there is a constant* $\hat{c}>0$ *such that* $\lim_{x\to+\infty}x^s\mathbf{P}(\bar{b}_0>x)=\hat{c}$.

Proof. According to (3.7)

$$\rho_{n-j}=p_{n-j}^{-1}q_n\alpha_{n-1}\dots\alpha_{n-j+1}\rho_n+\mathcal{O}(K^M).$$

Since $K^M\ll N^{1/s}$ we see that

$$\frac{\sum_{j=0}^{M}\rho_{n-j}}{\rho_n}=1+p_{n-1}^{-1}q_n+p_{n-2}^{-1}q_n\alpha_{n-1}+\cdots+p_{n-M}^{-1}q_n\alpha_{n-1}\dots\alpha_{n-M+1}+o(1).$$

As $N\to\infty$, also $M=M_N\to\infty$ and so the limiting distribution of the above expression is the same as the distribution of

$$1+p_{-1}^{-1}q_0+p_{-2}^{-1}q_0\alpha_{-1}+\cdots+p_{-k}^{-1}q_0\alpha_{-1}\dots\alpha_{-k+1}+... \tag{5.4}$$

This completes the proof of (a). Next note that $\bar{b}_{n+1}=\alpha_n\bar{b}_n+p_n^{-1}$. Hence (b) follows from part (a) of Lemma 3.6. □

Next take $\varepsilon_5<\varepsilon_4<\varepsilon_2$ where ε_2 is from Lemma 3.7(b). Divide $[1,\infty)$ (the set of all possible values of a_n) into finitely many disjoint intervals $I_1,I_2,\dots,I_{d_1}$. Divide $[0,N]$ into a union of long intervals L_j of length N^{ε_4} separated by short intervals of length N^{ε_5}. (Intervals are numbered in decreasing order). By Lemma 3.6 the total number of clusters originated in short intervals tends to 0 in probability since the total number of sites in the union of short intervals is $o(N)$. Observe that by Lemmas 3.7 and 5.1

$$\mathbf{P}\left(n \text{ is marked, } a_n\in I_m \text{ and } \rho_{n-k}\leq\delta N^{1/s}, k=M\dots N^{\varepsilon_4}\right)$$
$$\sim\frac{\mu_\infty(I_m)}{N}\left(1-\mathcal{O}\left(\beta^M\right)\right).$$

Lemma 5.2 now implies that if we divide $[1,\infty)\times[1,\infty)$ into finitely many disjoint rectangles $J_1,J_2\dots J_{d_1}$ then

$$\mathbf{P}\left(n \text{ is marked, } (a_n,b_n)\in J_m \text{ and } \rho_{n-k}\leq\delta N^{1/s}, k=M\dots N^{\varepsilon_4}\right)$$
$$\sim\frac{\tilde{\mu}_\infty(J_m)}{N}\left(1-\mathcal{O}\left(\beta^M\right)\right)$$

where $\tilde{\mu}_\infty$ is a measure on $[1,\infty)\times[1,\infty)$.

Let V_j be the vector whose m-th component is

$$\mathrm{Card}(n \in L_j : n - marked, (a_n, b_n) \in J_m).$$

Then

$$\mathbf{P}(V_j = e_m) \sim \tilde{\mu}(J_m) N^{\varepsilon_4 - 1}, \quad \mathbf{P}(|V_j| > 1) = o(N^{\varepsilon_4 - 1})$$

so

$$\mathbf{E}(\exp(i\langle v, V_j\rangle) = 1 + N^{\varepsilon_4 - 1} \sum_m \tilde{\mu}(J_m)(e^{iv_m} - 1) + o(N^{\varepsilon_4 - 1}). \tag{5.5}$$

where v_m denotes the m-th component of vector v.

Next, divide $[0, 1]$ into intervals $K_1, K_2 \dots K_{d_2}$. Pick d_2 vectors $u^{(1)}, u^{(2)} \dots u^{(d_2)}$ and let $v^{(j)} = u^{(l)}$ if $jN^{\varepsilon_4} \in K_l$. We claim that

$$\ln \mathbf{E}(\exp(i \sum_{k=1}^{j} \langle v^{(k)}, V_k\rangle) = \sum_{k=1}^{j} \sum_m \tilde{\mu}(J_m)(e^{iv_m^{(j)}} - 1) + o(jN^{\varepsilon_4 - 1}). \tag{5.6}$$

This holds because V_j is almost independent of $V_1, V_2 \dots V_{j-1}$. Namely, by Lemma 3.8 the value of ρ_n at the left endpoint of L_j could influence V_j only if ρ_{n-k} is $\beta_1^{N^{\varepsilon_5}}$-close to the boundary of I_m. However if N is large then the probability that there is $n - k \in L_j$ such ρ_{n-k} is close to the boundary of I_m is $o(N^{\varepsilon_4 - 1})$ and hence arguing as in the proof of (5.5) we obtain (5.6). Taking $j \sim N^{1-\varepsilon_4}$ we obtain

$$\ln \mathbf{E}(\exp(i \sum_{k=1}^{j} \langle v^{(k)}, V_k\rangle) = \sum_{l=1}^{d_2} \sum_{m=1}^{d_1} |K_l| \tilde{\mu}(J_m)(e^{iu_m^{(l)}} - 1) + o(1)$$

which implies parts (a) and (b) of Lemma 4.4.

It remains to prove (4.16). To do this note that $\bar{\Theta}$ has the same distribution as $\tilde{a} + \tilde{a}q\bar{b}$ where the distribution of $\tilde{a}$ and $\bar{b}$ is given by (5.2) and (5.3) respectively. Now the proof of (4.16) is the same as the proof of part (c) of Lemma 3.6.

6. Case $s = 2$: proof of Theorem 3

To prove Theorem 3 we follow the approach used in [3].

We split

$$\sum_{n=1}^{N} (\xi_n - \rho_n) = S_L + S_M + S_H$$

where S_H corresponds to the high values of ρ_n, namely, $\rho_n > \sqrt{N} \ln^{100} N$, S_M corresponds to the moderate values of ρ_n, namely, $\frac{\sqrt{N}}{\ln^{100} N} \le \rho_n \le \sqrt{N} \ln^{100} N$ and S_L corresponds to the low values of ρ_n, namely, $\rho_n <$

$\frac{\sqrt{N}}{\ln^{100} N}$. We begin by showing that high and moderate values of ρ_n can be ignored. First, by Lemma 3.6

$$\mathbf{P}(S_H \neq 0) \leq N\mathbf{P}(\rho_n > \sqrt{N}\ln^{100} N) \leq \frac{C}{\ln^{200} N}.$$

Second, arguing as in the proof of Lemma 4.1(b) we see that

$$\mathbf{E}(S_M^2) \leq \text{Const} \sum \mathbf{E}\left((\rho_n)^2 I_{\sqrt{N}/\ln^{100} N < \rho_n < \sqrt{N}\ln^{100} N}\right) \leq \text{Const} N \ln\ln N$$

and hence $S_M/\sqrt{N \ln N}$ converges to 0 in probability.

Therefore the main contribution comes from S_L. To handle it use Bernstein's method. Divide the interval $[0, N]$ into blocks of length $L_N = \ln^{10} N$ and $l_N = \ln^2 N$ following each other. More precisely the j-th big block is

$$I_j = [j(L_N + l_N), (j+1)L_N + jl_N - 1]$$

and j-th small block is

$$J_j = [(j+1)L_N + jl_N, (j+1)(L_N + l_N) - 1].$$

Accordingly, we split $S_L = S_L^{big} + S_L^{small}$, where S_L^{big} (S_L^{small}) is the contribution to S_L coming from big (small) blocks. Arguing as in the proof of Lemma 4.1(b) we see that

$$\mathbf{E}(\text{Var}_\omega(S_L^{small})) \leq C \sum_{n \in \text{small blocks}} \left[\mathbf{E}((\rho_n)^2) + \mathbf{E}(\rho_n l_N)\right] I_{\rho_n < \sqrt{N}/\ln^{100} N}$$
$$\leq C\left(N \ln N \frac{l_N}{L_N} + N\frac{l_N^2}{L_N}\right)$$

and hence the main contribution comes from the big blocks.

Next we modify ξ_n as follows. If $n \in I_j$ let $\tilde{\xi}_n$ be the number of visits to the site n before our walk reaches I_{j+1}. Let $\tilde{\rho}_n = \mathbb{E}_\omega(\tilde{\xi}_n)$. Observe that $\tilde{\xi}_n$ corresponds to imposing absorbing boundary conditions at the beginning of I_{j+1} so $\tilde{\rho}_n = p_n^{-1} q_{n+1} \tilde{\rho}_n + p_n^{-1}$ with absorbing boundary condition at $\bar{n} := (L_N + l_N)(j+1)$. Hence

$$\rho_n - \tilde{\rho}_n = \frac{q_{\bar{n}}}{q_n} \alpha_n \dots \alpha_{\bar{n}-1} \rho_{\bar{n}}$$

and so by (3.12)

$$\left| \sum_{n \in \text{big blocks}} [\tilde{\rho}_n - \rho_n] \right| < 1 \tag{6.1}$$

except for the set of probability tending to 0 as $N \to \infty$. Also by Lemma 3.4

$$\mathbf{P}(\tilde{\xi}_n = \xi_n \text{ for } n = 0 \dots N) \to 1 \text{ as } N \to \infty.$$

Let

$$\tilde{S} = \sum_{n \in \text{big blocks}} (\tilde{\xi}_n - \tilde{\rho}_n) I_{\tilde{\rho}_n < \sqrt{N}/\ln^{100} N}.$$

By the foregoing discussion it is enough to show that

(6.2) with $\mathbf{P}$ probability close to 1 the quenched distribution of $\tilde{S}$ is close to normal.

We claim that the following limit exists (in probability)

$$\lim_{N \to \infty} \frac{\mathrm{Var}_\omega(\tilde{S})}{N \ln N} = D_2. \tag{6.3}$$

Before proving (6.3) let us show how to complete the proof of (6.2). Let

$$\tilde{S}_j = \sum_{n \in I_j} (\tilde{\xi}_n - \tilde{\rho}_n) I_{\tilde{\rho}_n < \sqrt{N}/\ln^{100} N}$$

be the contribution of the j-th block. Since summation is taken over n with $\rho_n < \sqrt{N}/\ln^{100} N$ and $\tilde{\xi}_n$ has geometric distribution we have for $k \in \mathbb{N}$

$$\mathbb{P}_\omega\left(\tilde{S}_j > \frac{\sqrt{N} L_N k}{\ln^{100} N}\right) \leq C e^{-k} L_N. \tag{6.4}$$

Indeed $\tilde{S}_j > \frac{\sqrt{N} L_N k}{\ln^{100} N}$ implies that $\tilde{\xi}_n > \frac{\sqrt{N} k}{\ln^{100} N}$ for some n in the block. (6.3) and (6.4) show that $\sum_j \tilde{S}_j$ satisfies the Lindeberg condition. It remains to establish (6.3). To this end we prove two facts.

$$(A) \quad \forall \varepsilon > 0 \exists M : \mathbf{P}\left(\frac{\sum_{n_1 < n_2 - M} \mathrm{Cov}_\omega(\tilde{\xi}_{n_1}, \tilde{\xi}_{n_1})}{N \ln N} > \varepsilon\right) < \varepsilon \quad \text{and}$$

$$(B) \quad \forall k \quad \frac{\sum_n \mathrm{Cov}_\omega(\tilde{\xi}_n, \tilde{\xi}_{n-k})}{N \ln N} \Rightarrow \frac{\mathbf{E}(\alpha)^k c^*}{2} \text{ in probability}$$

where c^* is the constant from Lemma 3.6.

The remaining part of Section 6 is devoted to the proofs of statements (A) and (B). We will drop tildes in $\tilde{\xi}$ and $\tilde{\rho}$ in order to simplify notation. (Note that the results of Sections 3.3 and 3.4 are valid for $\tilde{\rho}, \tilde{\xi}$ since the arguments in those sections did not depend on the boundary conditions. cf. also (6.1)).

To obtain (A) we show that there is $\theta < 1$ such that

$$\mathbf{E}(|\mathrm{Cov}_\omega(\xi_{n-k}, \xi_n)| \, |\mathcal{F}_n) \leq C\theta^k (\rho_n)^2 \tag{6.5}$$

where $\mathcal{F}_n$ denotes the σ-algebra generated by $\{p_m\}_{m \geq n}$.

Pick a small $\epsilon > 0$ and consider two cases

(I) $\rho_n > (1+\epsilon)^k$. Then we use that

$$|\text{Cov}_\omega(\xi_{n-k}, \xi_n)| \leq \sqrt{\text{Var}_\omega(\xi_{n-k})\text{Var}_\omega(\xi_n)} \leq C\rho_{n-k}\rho_n$$

and that

$$\mathbf{E}(\rho_{n-k}|\mathcal{F}_n) \leq \mathbf{E}(\alpha)^k \rho_n + C.$$

(II) $\rho_n \leq (1+\epsilon)^k$. Then by (3.25)

$$|\text{Cov}_\omega(\xi_{n-k}, \xi_n)| \leq C\text{Var}_\omega(\xi_n)\rho_{n-k}q^*$$

where q^* is the probability to visit $n-k$ before n starting from $n-1$. Hence

$$\mathbf{E}(|\text{Cov}_\omega(\xi_{n-k}, \xi_n)||\mathcal{F}_n) \leq C\rho_n\sqrt{\mathbf{E}((\rho_{n-k})^2|\mathcal{F}_n)\mathbf{E}(q^*|\mathcal{F}_n)}$$

We have

$$\mathbf{E}((\rho_{n-k})^2|\mathcal{F}_n) \leq \rho_n^2 + Ck$$

since $s = 2$ whereas $\mathbf{E}(q^*|\mathcal{F}_n) \leq C\theta^k$ by Lemma 3.4.

Summing (6.5) over k we obtain (A).

To prove (B) observe that by Lemma 3.10 for fixed k we have

$$\text{Cov}_\omega(\xi_{n-k}, \xi_n) = \rho_{n-k}\rho_n + \mathcal{O}(\rho_n)$$

where the implicit constant depends on k. Let $Z_n = \text{Cov}_\omega(\xi_n, \xi_{n-k})$. Since $\mathbf{E}(\rho_{n-k}|\mathcal{F}_n) = \rho_n\mathbf{E}(\alpha)^k + C$ we get

$$\mathbf{E}(Z_n I_{\rho_n < \sqrt{N}/\ln^{100} N}) = \mathbf{E}((\rho_n)^2 I_{\rho_n < \sqrt{N}/\ln^{100} N})\mathbf{E}(\alpha)^k + \mathcal{O}\left(\mathbf{E}\left(\rho_n\right)\right).$$

Next

$$\text{Var}\left(\sum_n Z_n\right) = \sum \text{Var}(Z_n) + 2\sum_{n_1<n_2} \text{Cov}(Z_{n_1}, Z_{n_2}).$$

Observe that Z_{n_1} and Z_{n_2} are independent if n_1, n_2 belong to different blocks and so we can limit summation over n_1, n_2 in the same block. Since

$$Z_n = \rho_n^2\alpha_{n-k}\dots\alpha_{n-1} + \mathcal{O}(\rho_n)$$

Lemma 3.6 gives

$$\text{Var}(Z_n) \leq \text{Const}\frac{N}{\ln^{200} N}.$$

By Cauchy-Schwartz inequality

$$\text{Cov}(Z_{n_1}, Z_{n_2}) \leq \text{Const}\frac{N}{\ln^{200} N}.$$

Therefore

$$\text{Var}(\sum_n Z_n) \leq \text{Const}\frac{NL_N^2}{\ln^{100} N}.$$

This completes the proof of (B).

Remark. The same argument allows one to handle the case $s > 2$. Actually this case is simpler since there is no need to introduce cut-offs. We do not provide the details here since the case $s > 2$ had been studied in detail in [10, 16][3], where stronger almost sure quenched limit theorems were obtained (as has already been mentioned in the Introduction). However, such almost sure statement can not be extended to $s = 2$ for two (related) reasons. First the quenched variance of ρ_n is not integrable and so (6.3) can not be upgraded to almost sure convergence [1]. Secondly even though the contribution of the site with largest ρ_n is much smaller than the contribution of the remaining sites with probability close to 1, still $\mathbf{P}(\max_n \rho_n > \sqrt{N}\ln^{100} N)$ decays quite slowly (as $\ln^{-200} N$) and so from time to time we will see the situation where the site with largest ρ_n can not be ignored: the distribution of the sequence T_N would alternate between that of the normal and exponential variables. So our method does not recover the results of [10, 16] but it allows us to reprove Theorem 1 for all values of s (as in [15]).

7. Maximum occupation time.

Here we prove Theorem 4. Consider the following process $\hat{\Lambda}_N^\delta = \{(\frac{n_j}{N}, \frac{\hat{m}_j}{N^{1/s}})\}$ where n_j are marked points and $\hat{m}_j$ is the maximum of ρ_n inside the j-th cluster. The proof of Theorem 4 relies on the following fact.

Lemma 7.1. *(a) As $N \to \infty$ $\hat{\Lambda}_N^\delta$ converges to a Poisson process $\hat{\Lambda}^\delta = \{t_j, \boldsymbol{\theta}_j)\}$ on $[0, 1] \times [\delta, \infty)$.*

(b) As $\delta \to 0$ $\hat{\Lambda}^\delta$ converges to the Poisson process $\hat{\Lambda}$.

(c) There exists a constant $\hat{c}$ such that $\hat{\Lambda}$ has the intensity $\frac{\hat{c}}{\boldsymbol{\theta}^{1+s}}$.

The proof of this lemma is similar to the proofs of Lemmas 4.4 and 4.5 and therefore will be omitted.

Next we show that the low values of ρ are unlikely to contribute to the maximal occupation times. Fix $\theta > 0$. Denote

$$\Omega_{N,k} = \{\exists n \leq N : N^{1/s} 2^{-(k+1)} < \rho_n \leq N^{1/s} 2^{-k} \text{ and } \xi_n > \theta N^{1/s}\}$$

and set

$$\Phi_{N,k,n} = \{N^{1/s} 2^{-(k+1)} < \rho_n \leq N^{1/s} 2^{-k}\}.$$

Then by Lemma 3.6

$$\mathbf{P}(\Omega_{N,k}) \leq N\mathbf{P}\left(\Phi_{N,k,n}\right) \mathbf{P}\left(\xi_n > \theta N^{1/s} | \Phi_{N,k,n}\right) \leq \text{Const} 2^{ks} \mathbf{P}\left(\xi_n > \theta N^{1/s} | \Phi_{N,k,n}\right)$$

[3] [10] considers a class of environments which is much larger than the one treated in our paper.

Since ξ_n has a geometric distribution with parameter ρ_n^{-1} we have that

$$\mathbf{P}\left(\xi_n > \theta N^{1/s} | \Phi_{N,k,n}\right) \leq (1-\rho_n^{-1})^{\theta N^{1/s}} \leq \text{Const} e^{-c2^k}.$$

Summing these bounds over $k \geq \log_2(1/\delta)$ we see that the points from outside of the clusters can be ignored. The rest of the proof of Theorem 4 is similar to the proof of Theorem 2. Namely Lemma 3.10 implies that with high probability the maximum occupation time inside the j-th cluster occurs at the site $\hat{n}_j$ such that $\rho_{\hat{n}_j} = \hat{m}_j$. This shows that if δ is sufficiently small then with probability close to 1 $\xi^*_N = \max_j \hat{m}_j \frac{\xi_{\hat{n}_j}}{\hat{m}_j}$ where the maximum is taken over the δ-clusters. For large N the $\frac{\xi_{\hat{n}_j}}{\hat{m}_j}$ is asymptotically exponential with mean 1. Therefore letting $N \to \infty$ and $\delta_N \to 0$ we obtain that the distribution of $\frac{\xi^*_N}{N^{1/s}}$ is asymptotically the same as that of

$$\max_j \hat{\boldsymbol{\theta}}_j \Gamma_j$$

where $\hat{\Lambda} = \{(t_j, \hat{\boldsymbol{\theta}}_j)\}$ and Γ_j are i.i.d random variables independent of $\hat{\Lambda}$ and having mean 1 exponential distribution. It remains to notice that by Lemma 3.1 $\{\hat{\boldsymbol{\theta}}_j \Gamma_j\}$ also form a Poisson process with intensity $\frac{\bar{\mathbf{c}}}{\theta^{1+s}}$.

Appendix A. Annealed distribution.

Here we show how our results allow us to recover the known facts about the annealed distribution.

A.1. **Proof of Theorem 1.** By Lemma 2.1 we can consider T_N instead of $\tilde{T}_N$.

If $0 < s < 1$ then our result follows from Theorem 2(a), Lemma 3.1(c) and Lemma 3.3(a).

If $1 < s < 2$ we have

$$\frac{T_N - \mathbf{E}(T_N)}{N^{1/s}} = \mathfrak{t}_n - \mathfrak{u}_n = \sum \Theta_j^{(N,\delta)} (\Gamma_j^{(N,\delta)} - 1) + R_N + \hat{R}_N. \tag{A.1}$$

By Theorem 2(b) given $\varepsilon > 0$ we can find δ_0 such that for $\delta < \delta_0$

$$\mathbf{P}(|R_N + \hat{R}_N| > \varepsilon) < \varepsilon.$$

Also by Theorem 2(b) as $N \to \infty$ the sum in (A.1) converges to

$$\sum_{\Theta_j > \delta} \Theta_j (\Gamma_j - 1) \tag{A.2}$$

where Θ_j is a Poisson process on $(0, \infty)$ with intensity $\frac{\bar{c}}{\theta^{1+s}}$ and $\{\Gamma_j\}$ are iid random variables independent of $\{\Theta_j\}$ and having standard exponential distribution. On the other hand by Lemmas 3.1(c) and

3.3(b) if we drop the restriction that $\Theta_j > \delta$ in (A.2) then the sum will have a stable distribution of index s. Therefore to complete the proof it suffices to show that given $\varepsilon > 0$ we can find δ_0 such that for $\delta < \delta_0$

$$\mathbf{P}(|\bar{R}_\delta| > \varepsilon) < \varepsilon \tag{A.3}$$

where

$$\bar{R}_\delta = \sum_{\Theta_j < \delta} \Theta_j(\Gamma_j - 1).$$

Note that $\mathbf{E}(\bar{R}_\delta) = 0$ due to independence of Γs and Θs while by (3.4) $\mathrm{Var}(\bar{R}_\delta) = C\delta^{2-s}$. This proves (A.3) completing the proof of Theorem 1 in case $1 < s < 2$.

The proof in case $s = 1$ is similar to $1 < s < 2$ case. In case $s = 2$ the result follows from Theorem 3. Finally the case $s > 2$ was discussed in the introduction.

A.2. **Proof of Corollary 2.** To prove Corollary 2 it suffices to show that given $\varepsilon > 0$ the event

$$d(F_N^\omega, \mathbf{F}) < \varepsilon$$

occurs infinitely often.

Given an interval $I = [n_1, n_2]$ let T_I be the total time the walker spends inside I before $\tilde{T}_{n_2}$, the hitting time of n_2. Note that the quenched distribution functions $F^\omega_{T_{I_1}} \dots F^\omega_{T_{I_k}}$ of T_{I_j} are independent if the intervals $I_1, I_2 \dots I_k$ have disjoint interiors. On the other hand we can choose a sequence N_m growing so fast that

$$\mathbf{P}\left(d(F^\omega_{N_m}, F^\omega_{(m)}) > \frac{\varepsilon}{4}\right) < \frac{1}{m^{10}}. \tag{A.4}$$

where $F^\omega_{(m)} = F_{T_{[N_{m-1}, N_m]}/N_m^{1/s}}$. Accordingly in view of the Borel-Cantelli Lemma it suffices to show that the event

$$d(F^\omega_{(m)}, \mathbf{F}) < \frac{\varepsilon}{2} \quad \text{occurs infinitely often.} \tag{A.5}$$

By Corollary 1 there exists $c = c(\varepsilon) > 0$ such that

$$\mathbf{P}\left(d(F^\omega_{N_m}, \mathbf{F}) < \frac{\varepsilon}{4}\right) > c.$$

Now (A.4) implies that for large m we have

$$\mathbf{P}\left(d(F^\omega_{(m)}, \mathbf{F}) < \frac{\varepsilon}{2}\right) > \frac{c}{2}$$

and hence (A.5) follows from Borel-Cantelli Lemma.

A.3. **Proof of Corollary 3.** Let N_k be a rapidly growing sequence (see condition (A.6) below for the precise requirement on the growth of N_k.).

For $n \in [N_k, N_{k+1})$ let $\tilde{\xi}_n$ be the number of visits to the site n before the first visit to N_{k+1}. Note that $\tilde{\xi}_n$ only depends on the environment between in $[N_k, N_{k+1})$ since it counts visits to site n by the walk with absorbing boundary at N_{k+1} and reflecting boundary at $N_k - 1$. Let $\xi^{(k)} = \max_{[N_k,N_{k+1})} \tilde{\xi}_n$. Then $\xi^{(k)} \leq \xi^*_{N_{k+1}}$. Assume that N_k grows so fast that

$$\text{(A.6)} \quad \sum_k \mathbf{P}\left(T_{N_{k+1}} - \bar{T}^{(k)} > \frac{\bar{T}^{(k)}}{2}\right) < \infty, \text{ where } \bar{T}_k = \sum_{n \in [N_k, N_{k+1})} \tilde{\xi}_n$$

(such a sequence exists because $\bar{T}^{(k)} \geq N_{k+1} - N_k$). In view of the (A.6) it suffices to show that

$$\liminf \frac{\xi^{(k)}}{\bar{T}^{(k)}} > 0$$

almost surely. By Theorem 4 and Lemma 2.1

$$\mathbf{P}\left(\frac{\xi^{(k)}}{\bar{T}_{N_k}} > c_1\right) > c_2.$$

Corollary 3 now follows from the Borel-Cantelli Lemma.

Appendix B. Proof of Corollary 1.

We consider only the first statement of Corollary 1 and, moreover, we suppose that $0 < s < 1$. The proofs for all other cases are similar. We observe the following.

(1) Denote by $\mathfrak{t}_{N,\delta} = \sum_j \Theta_j^{(N,\delta)} \Gamma_j^{(N,\delta)}$ and define $F^\omega_{N,\delta}$ by $F^\omega_{N,\delta}(x) = \mathbb{P}_\omega\left(\mathfrak{t}_{N,\delta} \leq x\right)$. Obviously $\mathfrak{t}_N - \mathfrak{t}_{N,\delta} = R_N$ when $\omega \in \Omega_{N,\delta}$ (see (2.4)). Due to the estimate of R_N and the fact that $\mathbf{P}(\Omega_{N,\delta}) \to 0$ as $N \to \infty$ we have that for a given $\varepsilon > 0$ there are $N(\varepsilon)$ and $\delta(\varepsilon)$ such that $\mathbf{P}\left(d(F^\omega_{N,\delta}, F^\omega_N)\right) < \varepsilon$ for all $N > N(\varepsilon)$ and $\delta < \delta(\varepsilon)$.

(2) Consider $Y_\delta = \sum_{\Theta_j > \delta} \Theta_j \Gamma_j$ and let $F^\Theta_\delta(x) = P(Y_\delta \leq x | \Theta)$. If $\delta(\varepsilon)$ is small enough then $P(d(F^\Theta, F^\Theta_\delta) > \varepsilon) < \varepsilon$ if $\delta < \delta(\varepsilon)$.

(3) Consider $\mathcal{X}^{[a,b]} = \{F : F(b) - F(a) \geq 1 - \varepsilon\} \subset \mathcal{X}$ such that $P\left(\mathcal{X}^{[a,b]}\right) \geq 1 - \varepsilon$. Obviously, such a, b can be found for any $\varepsilon > 0$.

(4) The above properties imply that in order to establish that F^ω_N converges weakly to F^Θ it suffices to show that the sequence of ransom processes $F^\omega_{N,\delta}(\cdot)$ restricted to $[a, b]$ converges weakly to the restriction of the processes $F^\Theta_\delta(\cdot)$ to $[a, b]$. In turn, due to the monotonicity of

these random processes, this convergence follows from the convergence of the finite-dimensional distributions of $F^{\omega}_{N,\delta}(\cdot)$ to those of $F^{\Theta}_{\delta}(\cdot)$.

(5) It remains to prove that for any $x_j \in [a,b],\ j = 1, ..., n$, and any $u_j \in [0,1]\ j = 1, ..., n$,

$$\lim_{N\to\infty} \mathbf{P}\left(F^{\omega}_{N,\delta}(x_j) \le u_j,\ j = 1, ..., n\right) = P\left(F^{\Theta}_{\delta}(x_j) \le u_j,\ j = 1, ..., n\right).$$

The latter follows from statements (i) and (ii) of Theorem 2.

References

[1] Aaronson J. *An introduction to infinite ergodic theory,* Math. Surv. & Monographs **50** (1997) AMS Providence, RI.

[2] Bolthausen E., Goldsheid I. *Recurrence and transience of random walks in random environments on a strip,* Comm. Math. Phys. **214** (2000) 429–447.

[3] Chernov N., Dolgopyat D. *Anomalous current in periodic Lorentz gases with infinite horizon,* Russian Math. Surveys **64** (2009) 651–699.

[4] Enriquez N., Sabot C., Tournier L., Zindy O. *Stable fluctuations for ballistic random walks in random environment on* $\mathbb{Z}$, ArXiv preprint 10041333.

[5] N. Enriquez, C. Sabot, L. Tournier, O. Zindy *Annealed and quenched fluctuations for ballistic random walks in random environment on Z*, preprint.

[6] Enriquez N., Sabot C., Zindy O. *Aging and quenched localization for one-dimensional random walks in random environment in the sub-ballistic regime,* Bull. Soc. Math. France **137** (2009) 423-452.

[7] Enriquez N., Sabot C., Zindy O. *Limit laws for transient random walks in random environment on* $\mathbb{Z}$, Annales de l'institut Fourier **59**, no. 6 (2009) 2469-2508.

[8] Gantert N., Shi Z. *Many visits to a single site by a transient random walk in random environment,* Stochastic Process. Appl. **99** (2002) 159–176.

[9] Goldsheid I. *Simple transient random walks in one-dimensional random environment: the central limit theorem,* Prob. Th., Related Fields **139** (2007) 41–64.

[10] Goldsheid I. *Linear and sub-linear growth and the CLT for hitting times of a random walk in random environment on a strip,* Prob. Th., Related Fields **141** (2008) 471–511.

[11] Guivarch Y. *Heavy tail properties of stationary solutions of multidimensional stochastic recursions,* In *Dynamics & stochastics,* 85–99, IMS Lecture Notes Monogr. Ser., **48** (2006) Inst. Math. Stat., Beachwood, OH.

[12] Guivarch Y., Le Page E. *On spectral properties of a family of transfer operators and convergence to stable laws for affine random walks,* Erg. Th. Dynam. Systems **28** (2008) 423–446.

[13] Kesten H. *Random difference equations and renewal theory for products of random matrices,* Acta Math. **131** (1973), 207–248.

[14] Kesten H. *Renewal theory for functionals of a Markov chain with general state space,* Ann. Prob. **2** (1974) 355–386.

[15] Kesten H., Kozlov M., Spitzer F. *A limit law for random walk in a random environment,* Compositio Math. **30** (1975) 145–168.

[16] Peterson J. *Limiting distributions and large deviations for random walks in random environments,* PhD Thesis - University of Minnesota, 2008.

[17] Peterson J. *Quenched limits for transient, ballistic, sub-gaussian one-dimensional random walk in random environment,* Ann. Inst. H. Poincaré, Probab. Statist. **45** (2009) 685–709.

[18] Peterson J., Samorodnitsky G. *Weak quenched limiting distributions for transient one-dimensional random walk in a random environment,* preprint.

[19] Peterson J., Zeitouni O. *Quenched limits for transient, zero-speed one-dimensional random walk in random environment,* Ann. Prob. **37** (2009) 143-188.

[20] Resnick S. *Extreme Values, Regular Variation And Point Processes*, Springer, 2007, 320 pp.

[21] Revesz P. *Random walk in random and non-random environments,* 2d edition. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005. xvi+380 pp.

[22] Samorodnitsky G. and Taqqu M. *Stable non-Gaussian random processes,* Chapman & Hall/CRC, 1994, 640 pp.

[23] Shi Z. *A local time curiosity in random environment,* Stochastic Process. Appl. **76** (1998) 231–250.

[24] Shiga T., Tanaka H. *Infinitely divisible random probability distributions with an application to a random motion in a random environment,* Electron. J. Probab. **11** (2006), paper 44, 1144–1183.

[25] Solomon F.*Random walks in a random environment,* Ann. Prob. **3**, (1975) 1–31 .

[26] Sinai Ya. G. *The limiting behavior of a one-dimensional random walk in a random medium,* Theory Prob. Appl. **27** (1982), 256–268 .