RANDOMLY TRAPPED RANDOM WALKS ON $\mathbb{Z}^d$

JIŘÍ ČERNÝ AND TOBIAS WASSMER

Abstract. We give a complete classification of scaling limits of randomly trapped random walks and associated clock processes on $\mathbb{Z}^d$, $d \geq 2$. Namely, under the hypothesis that the discrete skeleton of the randomly trapped random walk has a slowly varying return probability, we show that the scaling limit of its clock process is either deterministic linearly growing or a stable subordinator. In the case when the discrete skeleton is a simple random walk on $\mathbb{Z}^d$, this implies that the scaling limit of the randomly trapped random walk is either Brownian motion or the Fractional Kinetics process, as conjectured in [BCČR14].

1. Introduction

Randomly trapped random walks (RTRWs) were introduced in [BCČR14] for two main reasons. On one hand they generalize several classical models of trapped random walks such as the continuous-time random walk or the symmetric Bouchaud trap model. On the other hand they provide a tool to describe random walks on some classical random structures such as the incipient critical Galton Watson tree or the invasion percolation cluster on a regular tree.

In [BCČR14] the authors define the RTRW on general graphs and study in depth the model on $\mathbb{Z}$. They give a complete classification of scaling limits, showing that the limit of a RTRW on $\mathbb{Z}$ is one of the following four processes: (i) Brownian motion, (ii) Fractional Kinetics process, (iii) FIN singular diffusion, or (iv) a new class of processes called spatially subordinated Brownian motion. They further give sufficient conditions for convergence to the respective limits and study in detail how the different limits arise. For RTRW on $\mathbb{Z}^d$, $d \geq 2$, they conjectured that only the first two of the above scaling limits are possible, that is RTRW on $\mathbb{Z}^d$ converges after rescaling either to the Brownian motion or to the Fractional Kinetics process. We prove this conjecture here.

Let us briefly introduce the model, its formal definition is given in Section 2 below. The RTRW on $\mathbb{Z}^d$ is a particular class of random walk in random environment. Its law is determined by two inputs: (i) its step distribution, that is a probability measure ν on $\mathbb{Z}^d$, and (ii) a probability distribution μ on the space of all probability measures on $(0, \infty)$ characterising its waiting times. The random environment of the RTRW is given by an i.i.d. collection $\pi = (\pi_x)_{x \in \mathbb{Z}^d}$ of μ-distributed probability measures. For fixed π, the RTRW $X = (X(t))_{t \geq 0}$ is a continuous-time process such that, whenever at vertex x, it stays there a random duration sampled from the distribution π_x and then moves on according to the transition kernel $\nu(\,\cdot\, - x)$. If the process X visits x again at a later time, the duration of this next visit at x is sampled again and independently from the distribution π_x. We always assume that X starts at $0 \in \mathbb{Z}^d$ and use $\mathbb{P}$ for the annealed distribution of the process X.

From the description above it is apparent that the RTRW is a time change of the discrete-time random walk $(Y(n))_{n \geq 0}$ on $\mathbb{Z}^d$ with one-step distribution ν. Formally, X can be

Date: May 30, 2014.

written as

$$X(t) = Y(S^{-1}(t)), \tag{1.1}$$

where the time-change process $S : \mathbb{N} \to [0, \infty)$, the *clock process*, measures the time needed for a given number of steps of the RTRW and S^{-1} is its right-continuous inverse. In view of (1.1) it should not be surprising that the scaling behaviour of X is (essentially) determined by the scaling behaviour of the clock process.

While we are primarily interested in Y being a simple random walk on $\mathbb{Z}^d$, $d \geq 2$, it does not complicate the proofs to make the following far less restrictive assumption on the random walk Y, that is on the one-step distribution ν: Let $r_n : \mathbb{N} \to [0, 1]$ be the probability that Y does not return to its starting point in n steps,

$$r_n = \mathbb{P}[Y(k) \neq Y(0) \text{ for } k = 1, \dots, n].$$

Assumption A. *The function r_n can be written as $r_n = \frac{1}{\ell^*(n)}$ for a slowly varying function $\ell^* : \mathbb{N} \to [1, \infty)$.*

Assumption A is obviously fulfilled for all transient random walks, where $1/\ell^*(n) \to \gamma$ for some $\gamma \in (0, 1)$, but there are also recurrent walks satisfying it with $\ell^*(n) \to \infty$ as $n \to \infty$. In particular, the classical result of Kesten and Spitzer [KS63, Theorem 3] implies that this assumption holds for all aperiodic[1] genuinely d-dimensional random walks in $d \geq 2$.

We can state our first main theorem giving the complete classification of the scaling limits of the clock process.

Theorem 1.1. *Let $S : \mathbb{N} \to [0, \infty)$ be the clock process of the RTRW. Suppose that Assumption A holds and there is a sequence $a_N \nearrow \infty$ such that for all but countably many $t \in [0, \infty)$*

$$S_N(t) := \frac{1}{a_N} S(\lfloor Nt \rfloor) \xrightarrow{N \to \infty} \mathcal{S}(t) \qquad \text{in } \mathbb{P}\text{-distribution}, \tag{1.2}$$

where $\mathcal{S} : [0, \infty) \to [0, \infty)$ is a cadlag process satisfying the non-triviality assumption

$$\limsup_{t \to \infty} \mathcal{S}(t) = \infty \qquad \mathbb{P}\text{-a.s.} \tag{1.3}$$

Then one of the following two cases occurs:

(i) The limit clock process is linear, $\mathcal{S}(t) = Mt$ for some constant $M > 0$, and the normalizing sequence satisfies $a_N = N\ell(N)$ for some slowly varying function ℓ.

(ii) The limit clock process is an α-stable subordinator, $\mathcal{S} = V_\alpha$, $\alpha \in (0, 1)$, and the normalizing sequence satisfies $a_N = N^{1/\alpha}\ell(N)$ for some slowly varying function ℓ.

In order to study the scaling limits of the RTRW itself, we need a more restrictive assumption:

Assumption B. *The one-step distribution ν of the random walk Y is centred, has finite range, and it is genuinely d-dimensional, that is $\mathbb{E}[Y(1)] = 0$, $\mathbb{P}[|Y(1)| > C] = 0$ for some $C < \infty$, and the linear span of the set $\{x : \nu(x) > 0\}$ is $\mathbb{R}^d$.*

This assumption ensures that the scaling limit of Y is a d-dimensional Brownian motion: There exists a $d \times d$ matrix $\mathcal{A}$ such that

$$Y_N(t) := \frac{1}{\sqrt{N}} \mathcal{A} Y(\lfloor Nt \rfloor) \tag{1.4}$$

[1] The aperiodicity assumption can easily be removed by considering the lazy version of Y first, applying [KS63] and transferring the result back to the original random walk.

converges to a standard d-dimensional Brownian motion. Note that by the remark after Assumption A, for $d \geq 2$ Assumption A is implied by Assumption B.

Our second main result classifies the possible scaling limits of RTRW and confirms the conjecture of [BCČR14].

Theorem 1.2. *Let $d \geq 2$ and $X : [0, \infty) \to \mathbb{Z}^d$ be the RTRW. Suppose that Assumption B holds and there is a sequence $a_N \nearrow \infty$ such that the processes*

$$X_N(t) := \frac{1}{\sqrt{N}} \mathcal{A} X(a_N t) = Y_N(S_N^{-1}(t)) \tag{1.5}$$

converge in $\mathbb{P}$-distribution on the space D^d of cadlag $\mathbb{R}^d$-valued functions equipped with the Skorohod J_1-topology to some process $\mathcal{X} : [0, \infty) \to \mathbb{R}^d$ satisfying the non-triviality assumption

$$\limsup_{t \to \infty} |\mathcal{X}(t)| = \infty \qquad \mathbb{P}\text{-a.s.} \tag{1.6}$$

Then one of the following two cases occurs:

(i) $a_N = N\ell(N)$ and $\mathcal{X}(t) = B(M^{-1}t)$ for some constant $M > 0$, some slowly varying function ℓ, and a standard d-dimensional Brownian motion B.

(ii) $a_N = N^{1/\alpha}\ell(N)$ for some slowly varying function ℓ and a parameter $\alpha \in (0, 1)$, and $\mathcal{X}(t) = B(V_\alpha^{-1}(t))$, where B is a standard d-dimensional Brownian motion and $V_\alpha^{-1}(t) = \inf\{s \geq 0 : V_\alpha(s) > t\}$ is the right-continuous inverse of an α-stable subordinator V_α which is independent of B (i.e. $\mathcal{X}$ is the Fractional Kinetics process).

Let us make a few remarks about our setting and results. The definition of the RTRW we give here is slightly more general than the one in [BCČR14] since we allow the discrete skeleton to be more general than the simple random walk only. Assumption A on the discrete skeleton is taken from [FM13]. This assumption can be used to show weak laws of large numbers for the range of the random walk and for some related quantities. We would like to point out that the only place in the proof of Theorem 1.1 where we use $\mathbb{Z}^d$-specific properties of the random walk is in the derivation of these laws of large numbers. In particular, Theorem 1.1 classifying the possible scaling limits of the clock process can be shown to hold for the RTRW on any countable state space where the discrete-time skeleton is a Markov chain satisfying such laws of large numbers for the range and the related quantities.

Our setting generalizes several previous results, let us mention some of them. Mostly, the models studied in the literature involve trapped random walks with some kind of heavy-tailed waiting times, with the aim to show convergence of rescaled clock processes to an α-stable subordinator.

In the so-called continuous-time random walk (CTRW), introduced in [MW65], all π_x are deterministically identical heavy-tailed probability distributions, that is for some $\alpha \in (0, 1)$ and $c > 0$,

$$\pi_x[u, \infty) = cu^{-\alpha}(1 + o(1)) \text{ as } u \to \infty. \tag{1.7}$$

Independently of the nature of the discrete skeleton Y, the clock process is then a sum of i.i.d. heavy-tailed random variables, and it is well known that it converges after normalization to a stable subordinator. The scaling limits of the CTRW were studied in more detail in [MS04].

In the symmetric Bouchaud trap model (BTM) the discrete skeleton Y is simple random walk and the π_x are exponential random variables with means m_x that are i.i.d. heavy-tailed random variables satisfying e.g.

$$\mathbb{P}[m_x > u] = cu^{-\alpha}(1 + o(1)) \text{ as } u \to \infty, \tag{1.8}$$

The BTM on $\mathbb{Z}^2$ was for the first time studied in [BČM06] where the authors show convergence of the clock process to a stable subordinator and use this to derive aging properties of the model. In [BČ07] it is then shown, in the case of the BTM on $\mathbb{Z}^d$, $d \geq 2$, that the rescaled random walks and clock processes converge jointly to a Brownian motion and a stable subordinator, and therefore the scaling limit of the BTM is the Fractional Kinetics process.

A general model of trapped random walk where the waiting times are exponential with heavy-tailed means as in (1.8) is studied in [FM13]. As mentioned above, they consider discrete skeleton to be an arbitrary random walk on $\mathbb{Z}^d$ satisfying Assumption A. Instead of scaling limits, which require additional restrictions as in our Assumption B, [FM13] focus on the so-called age process, which is related to the clock process and describes the 'depth of the trap in which the process stays at a given time'.

Our setting is restricted to the fact that the discrete skeleton Y is independent of the random environment π. There are however interesting models where this is not the case, for example the asymmetric Bouchaud trap model (ABTM). In [BČ11] for $d \geq 3$ and in [Mou11] with different methods for $d \geq 5$ it is shown that the scaling limit for ABTM is also Fractional Kinetics. Yet another approach to prove convergence of rescaled clock processes to a stable subordinator is given in [GS13], their setting includes the ABTM as a special case.

The most of the above mentioned previous results are quenched, that is the convergence holds for almost every realisation of the environment. On the contrary, our results are annealed, that is averaged over the environment, but this is not an issue for the classification theorem.

We also believe that when the annealed convergence takes place as in Theorem 1.1, then the quenched convergence holds true as well. In high dimensions ($d \geq 5$) this could be proved similarly as in [Mou11], using techniques from [BS02], see also the additional condition in [FM13] under which their annealed result holds quenched. In low dimensions these methods fail due to many self-intersections of the discrete skeleton. An adaptation of more complicated methods which give the quenched convergence in low dimensions (like the coarse-graining procedure of [BČM06, BČ07] or the techniques of [GS13]) to the RTRW seems to be non-trivial and is out of the scope of this paper.

We conclude the introduction by giving sufficient conditions for convergence in both cases of our main theorems. Given the collection of probability measures $\pi = (\pi_x)_{x\in\mathbb{Z}^d}$, let $m_x = \int u\pi_x(du) \in (0, \infty]$ be the mean and $\hat{\pi}_x(\lambda) = \int e^{-\lambda u}\pi_x(du)$ the Laplace transform of π_x. Note that in the next theorem Assumption A is not needed, we only need Y to be non-degenerate.

Theorem 1.3. *Let X be RTRW in $d \geq 1$. If $\nu \neq \delta_0$ and the annealed expected waiting time is finite, $\mathbb{E}[m_0] = M < \infty$, then the rescaled clock processes S_N with normalization $a_N = N$ converge in $\mathbb{P}$-distribution on D^1 equipped with the Skorohod J_1-topology to the linear process $\mathcal{S}(t) = Mt$. If in addition Assumption B holds, then the rescaled processes X_N with $a_N = N$ converge in $\mathbb{P}$-distribution on D^d equipped with the Skorohod J_1-topology, and the limit is $\mathcal{X}(t) = B(M^{-1}t)$ as in (i) of Theorem 1.2.*

For convergence to Fractional Kinetics we have the following sufficient criterium. In Section 5 we will sketch some examples of RTRWs that satisfy this criterium with different functions f.

Theorem 1.4. *Let X be RTRW with discrete skeleton Y satisfying Assumption A with slowly varying function ℓ^*. Assume that there is a normalizing sequence $a_N \nearrow \infty$ such that for any positive real number $r > 0$,*

$$-\log \mathbb{E}\left[\hat{\pi}_0(\lambda/a_N)^{r\ell^*(N)}\right] = f(r)\lambda^\alpha \frac{\ell^*(N)}{N}(1+o(1)) \text{ as } N \to \infty. \tag{1.9}$$

Then the rescaled clock processes S_N with normalization a_N converge in $\mathbb{P}$-distribution on D^1 equipped with the Skorohod M_1-topology to an α-stable subordinator V_α. If in addition Assumption B holds, then the rescaled processes X_N converge in $\mathbb{P}$-distribution on D^1 equipped with the Skorohod J_1-topology, and the limit is the FK process as in (ii) of Theorem 1.2.

The rest of this paper is structured as follows. In Section 2 we give precise definitions of the model and introduce some notation used through the paper. Theorem 1.1 and Theorem 1.2 are proved in Sections 3 and 4 respectively, and Section 5 deals with Theorems 1.3 and 1.4. Finally, in Section 6 we prove one technical lemma which is used in the proof of Theorem 1.1. In Appendix A we explain how Assumption A on the escape probability implies the laws of large numbers that we mentioned above.

2. Setting and notations

We start by giving a formal definition of the RTRW. Recall that ν is a probability measure on $\mathbb{Z}^d$ and μ a probability measure on the space of probability measures on $(0,\infty)$. To avoid trivial situations, we assume that $\nu \neq \delta_0$.

Given μ and ν, let $\pi = (\pi_x)_{x\in\mathbb{Z}^d}$ be an i.i.d. sequence of probability measures with marginal μ, and $\xi = (\xi_i)_{i\geq 1}$ an i.i.d. sequence with marginal ν independent of π defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Define

$$Y(n) = \xi_1 + \cdots + \xi_n$$

to be a random walk with step distribution ν and denote by $L(x,n) = \sum_{k=0}^n \mathbf{1}_{\{Y(k)=x\}}$ its local time.

Given a realisation of π, let further $(\tau_x^i)_{x\in\mathbb{Z}^d, i\geq 1}$ be a collection of independent random variables, independent of ξ, such that every τ_x^i has distribution π_x, defined on the same probability space. The clock process of the RTRW, $S : \mathbb{N} \to [0,\infty)$ is then defined by $S(0) = 0$ and

$$S(n) = \sum_{x\in\mathbb{Z}^d} \sum_{i=1}^{L(x,n-1)} \tau_x^i = \sum_{k=0}^{n-1} \tau_{Y(k)}^{L(Y(k),k)} \qquad \text{for } n \geq 1.$$

Finally, we define the RTRW $X = (X(t))_{t\geq 0}$ by

$$X(t) = Y(k) \qquad \text{for } S(k) \leq t < S(k+1),$$

or equivalently

$$X(t) = Y(S^{-1}(t)),$$

where $S^{-1}(t) = \inf\{k \geq 0 : \ S(k) > t\}$ is the right-continuous inverse of S.

Under $\mathbb{P}$, the process X has exactly the law described in the introduction. The random variable τ_x^i denotes the duration of the i-th visit of the vertex x. We refer to $\mathbb{P}$ as *annealed* distribution of X.

We write D^d for the space of the $\mathbb{R}^d$-valued cadlag functions on $[0,\infty)$, and when needed $D^d(J_1)$, $D^d(M_1)$, etc. to point out which of Skorohod topologies we use on this space. We refer to [Whi02, Chapter 3.3] for an introduction and [Whi02, Chapters 12–13] for details on these topologies.

It will be useful to introduce the sequence of successive waiting times

$$\tilde{\tau}_k = \tau_{Y(k)}^{L(Y(k),k)}, \qquad k \geq 0.$$

With this notation,

$$S(n) = \sum_{k=0}^{n-1} \tilde{\tau}_k. \tag{2.1}$$

We now show that $\tilde{\tau}_k$ is ergodic, which will be used in the proof of Theorem 1.3. To this end let $\mathcal{P}'$ be the law on $\Omega' := [0,\infty)^{\mathbb{N}}$ of the sequence $(\tilde{\tau}_k)_{k\geq 0}$ and let θ be the left shift on Ω', $\theta(\tilde{\tau}_1, \tilde{\tau}_2, \dots) = (\tilde{\tau}_2, \tilde{\tau}_3, \dots)$.

Lemma 2.1. *The left-shift θ acts ergodically on $(\Omega', \mathcal{P}')$.*

Proof. To show that θ is measure-preserving we follow the environment as 'viewed from the particle'. Namely, let $\Theta : \Omega \to \Omega$ be such that if $\omega' = \Theta(\omega)$, then

$$\begin{aligned}
\xi_i(\omega') &= \xi_{i+1}(\omega), \qquad i \geq 1,\\
\pi_x(\omega') &= \pi_{x+\xi_1(\omega)}(\omega), \qquad x \in \mathbb{Z}^d,\\
\tau_x^i(\omega') &= \begin{cases} \tau^i_{x+\xi_1(\omega)}, & \text{if } x \neq -\xi_1(\omega),\\ \tau^{i-1}_{x+\xi_1(\omega)}, & \text{if } x = -\xi_1(\omega). \end{cases}
\end{aligned}$$

From the independence of ξ from π and τ, and from the i.i.d. properties of π and $\tau_x^{\cdot}$ for every x, it is easy to see that the law of $X \circ \Theta$ agrees with the law of X, that is Θ is $\mathbb{P}$-preserving. Since, in addition, $\tilde{\tau}(\Theta(\omega)) = \theta(\tilde{\tau}(\omega))$ and $\mathcal{P}' = \tilde{\tau} \circ \mathbb{P}$, this implies that θ is $\mathcal{P}'$-preserving.

To prove the ergodicity, we show that θ is strongly mixing. To this end it is sufficient to verify that

$$\big|\mathcal{P}[\theta^{-n} A \cap B] - \mathcal{P}[A]\mathcal{P}[B]\big| \xrightarrow{n\to\infty} 0 \tag{2.2}$$

for all cylinder sets $A = \{\tilde{\tau}_i \in A_i,\ i \in I\}$, $B = \{\tilde{\tau}_j \in B_j,\ j \in J\}$, where $I, J \subset \mathbb{Z}$ are finite sets and $A_i, B_j \subset \mathbb{R}$ are Borel sets, see e.g. [Pet83, Prop 2.5.3]. Since $\tilde{\tau}_k$ and $\tilde{\tau}_l$ are independent with respect to the annealed measure whenever $Y(k) \neq Y(l)$, we have

$$\begin{aligned}
\big|\mathbb{P}[\theta^{-n} A \cap B] - \mathbb{P}[A]\mathbb{P}[B]\big| &\leq \mathbb{P}[Y(i+n) = Y(j) \text{ for some } i \in I, j \in J]\\
&\leq \sum_{i\in I, j\in J} \mathbb{P}[Y(i+n) = Y(j)].
\end{aligned}$$

Now for n large enough the Markov property for Y implies that $\mathbb{P}[Y(i+n) = Y(j)] = \mathbb{P}[Y(i+n-j) = 0]$, which tends to 0 as $n \to \infty$ for every (non-degenerate) random walk, see e.g. [Spi76, P7.6]. Since I and J are finite, (2.2) follows. ☐

3. Proof of Theorem 1.1

In this section we prove Theorem 1.1. To this end we assume that the rescaled clock processes S_N converge to some process $\mathcal{S}$ in the way as stated in Theorem 1.1. In the next two lemmas we study the properties of $\mathcal{S}$.

Lemma 3.1. *The limit clock process $\mathcal{S}$ has stationary increments and is self-similar with index $\rho > 0$, i.e. $\mathcal{S}(t) \stackrel{d}{=} \lambda^\rho \mathcal{S}(t/\lambda)$. Moreover, the normalizing sequence is of the form $a_N = N^\rho \ell(N)$, for the same $\rho > 0$ and some slowly varying function ℓ.*

Proof. Stationarity of the increments follows immediately from (2.1) and the stationarity of the sequence $\tilde{\tau}$ of successive waiting times which was proved in Lemma 2.1. To see the self-similarity, fix $\lambda > 0$ and t such that condition (1.2) holds for t and t/λ, and $\mathcal{S}(t)$, $\mathcal{S}(\lambda t)$ are not identically zero, which is possible thanks to (1.3). Then,

$$\mathcal{S}(t) = \lim_{N\to\infty} \frac{1}{a_N} S(Nt) = \lim_{N\to\infty} \frac{a_{\lambda N}}{a_N} \frac{1}{a_{\lambda N}} S\left(\lambda N \frac{t}{\lambda}\right) \stackrel{d}{=} \mathcal{S}\left(\frac{t}{\lambda}\right) \lim_{N\to\infty} \frac{a_{\lambda N}}{a_N}.$$

Since $\mathcal{S}(t)$ and $\mathcal{S}(t/\lambda)$ are not identically zero, it follows that $\frac{a_{\lambda N}}{a_N}$ must converge to some constant $c(\lambda)$, yielding the scale invariance. Moreover, elementary results of the theory of regularly varying functions imply that $c(\lambda) = \lambda^\rho$ for some $\rho \in \mathbb{R}$, and that a_N is regularly varying of index ρ, that is $a_N = N^\rho \ell(N)$ for some slowly varying function $\ell(N)$. Since $a_N \to \infty$, we have of course that $\rho > 0$. □

Lemma 3.2. *The limit clock process $\mathcal{S}$ has independent increments.*

Let us postpone the proof of this lemma and show Theorem 1.1 first.

Proof of Theorem 1.1. By Lemmas 3.1 and 3.2, $\mathcal{S}$ has stationary and independent increments and is self-similar with index ρ. From this and the fact that $\mathcal{S} \geq 0$ it follows that either $\rho = 1$ and $\mathcal{S}(t) = Mt$ for some $M \in (0, \infty)$, or $\rho > 1$ and $\mathcal{S}$ is an increasing α-stable Lévy process with $\alpha = \rho^{-1} \in (0,1)$, that is an α-stable subordinator. Lemma 3.1 gives the normalizing sequence a_N as claimed. □

In order to show Lemma 3.2 we need three technical lemmas which are consequences of laws of large numbers for the range-like objects related to the random walk Y, as mentioned in the introduction.

The first lemma states that for any given times $0 = t_0 < t_1 < \cdots < t_n = t$, the number of vertices visited by the random walk Y in more than one of the time intervals $[\lfloor t_{i-1} N\rfloor, \lfloor t_i N\rfloor - 1]$ is small. To this end, let

$$R(k) = \{Y(0), \dots, Y(k-1)\}$$

be the range of the random walk Y at time $k-1$, R_N^i be the 'range between $t_{i-1}N$ and $t_i N$',

$$R_N^i = \{Y(k) : \ k = \lfloor N t_{i-1}\rfloor, \dots, \lfloor N t_i\rfloor - 1\},$$

O_N^i be the set of the points visited only in this time interval,

$$O_N^i = \left\{x \in R_N^i : \ x \notin R_N^j \text{ for all } j \neq i\right\},$$

and M_N^i be the set of points visited in more than one of them, $M_N^i = R_N^i \setminus O_N^i$.

Lemma 3.3. *If Y verifies Assumption A, then for any choice of time points $0 = t_0 < t_1 < \cdots < t_n = t$,*

$$\lim_{N\to\infty} |M_N^i| \frac{\ell^*(N)}{N} = 0 \qquad \textit{in } \mathbb{P}\textit{-probability for all } i = 1, \dots, n.$$

Proof. The size of the sets O_N^i can be bounded by

$$|R(\lfloor Nt\rfloor)| - \left|\bigcup_{j\neq i} R_N^j\right| \leq |O_N^i| \leq |R(\lfloor Nt_i\rfloor)| - |R(\lfloor Nt_{i-1}\rfloor)|. \tag{3.1}$$

Applying the laws of large numbers from Lemma A.1 and the Markov property at times $\lfloor Nt_i \rfloor$, it follows that for every $i = 1, \dots, n$,

$$|R(\lfloor Nt_i \rfloor)| \frac{\ell^*(N)}{Nt_i} \xrightarrow{N\to\infty} 1, \quad \text{and} \quad \Big| \bigcup_{j\neq i} R_N^j \Big| \frac{\ell^*(N)}{N(t_n - t_i + t_{i-1})} \xrightarrow{N\to\infty} 1$$

in probability. Inserting this into (3.1) yields a law of large numbers for $|O_N^i|$,

$$|O_N^i| \frac{\ell^*(N)}{N(t_i - t_{i-1})} \xrightarrow{N\to\infty} 1$$

in probability. By Lemma A.1 and the Markov property again, $|R_N^i|$ satisfies the same law of large numbers as $|O_N^i|$. Using $|M_N^i| = |R_N^i| - |O_N^i|$ the claim follows. $\square$

The second lemma will help to control the contribution of frequently visited vertices to the clock process. Fix $t > 0$, and for $K > 0$ define the set of 'frequently visited vertices'

$$\mathcal{F}_{N,K} = \big\{ x : L(x, \lfloor Nt \rfloor - 1) \geq K\ell^*(N) \big\}. \tag{3.2}$$

Let $F_{N,K}$ be the 'number of visits to $\mathcal{F}_{N,K}$'

$$F_{N,K} = \sum_{x\in\mathcal{F}_{N,K}} L(x, \lfloor Nt \rfloor - 1). \tag{3.3}$$

Lemma 3.4. *If Y verifies Assumption A, then there is a constant $c > 0$ such that for every $\varepsilon > 0$*

$$\mathbb{P}[F_{N,K} \geq \epsilon Nt] \leq \epsilon \qquad \textit{for all } N \textit{ large enough,}$$

with

$$K = K(\epsilon) = -c \log\left(\epsilon^2\right). \tag{3.4}$$

Proof. We claim that for ϵ small enough and N large enough,

$$\mathbb{E}\left[F_{N,K}\right] \leq \epsilon^2 Nt. \tag{3.5}$$

Applying the Markov inequality then yields the desired result.

To show (3.5), let $\psi_k = \mathbf{1}_{\{Y(l)\neq Y(k)\,\forall l<k\}}$ be the indicator of the event that a 'new' vertex is found at time k. Then

$$F_{N,K} = \sum_{k=0}^{\lfloor Nt \rfloor -1} \psi_k L(Y_k, \lfloor Nt \rfloor - 1)\mathbf{1}_{\{L(Y_k,\lfloor Nt\rfloor -1)\geq K\ell^*(N)\}}.$$

Using the Markov property and the fact that $L(Y_k, \lfloor Nt \rfloor - 1)$ is stochastically dominated by $L(0, \lfloor Nt \rfloor - 1)$,

$$\mathbb{E}[F_{N,K}] \leq \mathbb{E}\big[L(0, \lfloor Nt \rfloor - 1)\mathbf{1}_{\{L(0,\lfloor Nt\rfloor -1)\geq K\ell^*(N)\}}\big] \sum_{k=0}^{\lfloor Nt \rfloor -1} \mathbb{E}[\psi_k]. \tag{3.6}$$

By (A.2), $\sum_{k=0}^{\lfloor Nt\rfloor -1} \mathbb{E}[\psi_k] = \mathbb{E}[|R(\lfloor Nt \rfloor)|] = Nt/\ell^*(N)$. On the other hand, denoting by $\tilde{H}_0$ the first return time of Y to 0, for every $k \geq 1$

$$\mathbb{P}[L(0, \lfloor Nt \rfloor - 1) \geq k] \leq \big(\mathbb{P}[\tilde{H}_0 \leq \lfloor Nt \rfloor]\big)^{k-1} = (1 - \ell^*(\lfloor Nt \rfloor)^{-1})^{k-1},$$

and thus $L(0, \lfloor Nt \rfloor - 1)$ is stochastically dominated by a geometric random variable with parameter $1/\ell^*(\lfloor Nt \rfloor)$. If G is a geometric variable with parameter p, then for every $M \in \mathbb{N}$,

$$\mathbb{E}[G\mathbf{1}_{\{G\geq M\}}] = (1-p)^{M-1}\Big(M - 1 + \frac{1}{p}\Big).$$

Hence,

$$\begin{aligned}\mathbb{E}\big[L(0,\lfloor Nt\rfloor-1)\mathbf{1}_{\{L(0,\lfloor Nt\rfloor-1)\ge K\ell^*(N)\}}\big]\\ \le\Big(1-\frac{1}{\ell^*(\lfloor Nt\rfloor)}\Big)^{K\ell^*(\lfloor Nt\rfloor)-1}\big((K+1)\ell^*(\lfloor Nt\rfloor)-1\big),\end{aligned}\tag{3.7}$$

and the claim (3.5) follows by inserting K as in (3.4) and combining (3.6), (3.7). □

The last of the technical lemmas allows to control the influence of an arbitrary subset of waiting times to the sum of all waiting times if the subset is small.

Lemma 3.5. *Let $\mathcal{B}_N\subset\{0,1,\dots,\lfloor Nt\rfloor-1\}$ be a random set, depending on the trajectory of the random walk Y up to time $\lfloor Nt\rfloor-1$ only. If Assumption A holds, then for every $t>0$ and $\delta>0$,*

$$\lim_{\epsilon\to0}\lim_{N\to\infty}\mathbb{P}\Bigg[\sum_{k\in\mathcal{B}}\tilde\tau_k\ge\delta S(\lfloor Nt\rfloor),\ |\mathcal{B}|\le\epsilon N\Bigg]=0.$$

The proof of this lemma is surprisingly lengthy and is therefore postponed to Section 6. The main source of complications comes from the fact that we cannot make any assumptions on the moments of the waiting times τ_x^i. It is also essential to use some properties of the random walk Y, as it is easy to construct counterexamples to the lemma when τ_x^i are not summed along the trajectory of Y.

With the above three lemmas we can now show Lemma 3.2.

Proof of Lemma 3.2. Fix times $0=t_0<t_1<\dots<t_n=t$. Consider first the following alternative construction of the clock process S. On the same space $(\Omega,\mathcal{F},\mathbb{P})$, let for every $x\in\mathbb{Z}^d$ independently $(\pi_{x,j})_{j=1,\dots,n}$ be i.i.d. μ-distributed probability measures, and given a realisation of these measures, let $(\tau_{x,j}^i)_{x\in\mathbb{Z}^d,j,i\ge1}$ be independent random variables such that every $\tau_{x,j}^i$ has distribution $\pi_{x,j}$. For every vertex $x\in\mathbb{Z}^d$, let $j(x)$ be be such that the first visit to x occurs in the time interval $[\lfloor Nt_{j(x)-1}\rfloor,\lfloor Nt_{j(x)}\rfloor-1]$. Define a new process $S':\mathbb{N}\to[0,\infty)$ by

$$S'(k)=\sum_{x\in\mathbb{Z}^d}\sum_{i=1}^{L(x,k-1)}\tau_{x,j(x)}^i.$$

One can think of choosing the distributions π_x at the time of the first visit in x according to the time interval in which this first visit occurs. Constructed in this way, S' has clearly the same distribution as the original clock process S.

We now define an approximation $\tilde S$ of S' which collects time $\tau_{x,j}^{L(x,k)}$ whenever at a vertex x at time $k\in[\lfloor Nt_{j-1}\rfloor,\lfloor Nt_j\rfloor-1]$,

$$\tilde S(m)=\sum_{j=1}^{n}\sum_{k=\lfloor Nt_{j-1}\rfloor}^{(m\wedge\lfloor Nt_j\rfloor)-1}\tau_{Y(k),j}^{L(Y(k),k)}.\tag{3.8}$$

$\tilde S$ can be viewed as the clock for which the whole environment π is being refreshed at all times $\lfloor Nt_j\rfloor$. Therefore, by the independence structure of the $\tau_{x,j}^i$'s, the increments $(\tilde S(\lfloor Nt_j\rfloor)-\tilde S(\lfloor Nt_{j-1}\rfloor))_{j=1,\dots,n}$ are mutually independent. In addition, for every j, the increment $\tilde S(\lfloor Nt_j\rfloor)-\tilde S(\lfloor Nt_{j-1}\rfloor)$ is independent of the increments $\{\xi_k:\ k\notin[\lfloor Nt_{j-1}\rfloor,\lfloor Nt_j\rfloor-1]\}$ of the random walk Y.

To conclude the proof it is now sufficient to show that for all $j = 1, \dots, n$ and every $\delta > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left[\left|\tilde{S}(\lfloor Nt_j \rfloor) - S'(\lfloor Nt_j \rfloor)\right| > \delta S'(\lfloor Nt_j \rfloor)\right] = 0. \tag{3.9}$$

This implies that the limit process $\mathcal{S}$ has independent increments. Indeed, note that (3.9) readily implies $\frac{\tilde{S}(\lfloor Nt_j \rfloor)}{S'(\lfloor Nt_j \rfloor)} \to 1$ in $\mathbb{P}$-probability for all j. This means that whenever $\frac{1}{a_N} S'(\lfloor Nt_j \rfloor) \overset{d}{\to} \mathcal{S}(t_j)$, then also $\frac{1}{a_N} \tilde{S}(\lfloor Nt_j \rfloor) \overset{d}{\to} \mathcal{S}'(t_j)$, and therefore the increments $(\mathcal{S}(t_j) - \mathcal{S}(t_{j-1}))_{j=1,\dots,n}$ are independent, whenever (1.2) is satisfied for the times t_j. By easy approximation arguments this also holds for the at most countably many t_j's that do not satisfy (1.2). Since the times t_j are chosen arbitrarily, it follows that the process $\mathcal{S}$ has independent increments.

In order to show (3.9), note that the difference of $\tilde{S}(\lfloor Nt_j \rfloor)$ and $S'(\lfloor Nt_j \rfloor)$ originates in the waiting times in vertices visited in multiple time intervals. Recalling the sets M_N^j from Lemma 3.2,

$$\left|\tilde{S}(\lfloor Nt_j \rfloor) - S'(\lfloor Nt_j \rfloor)\right| \le \sum_{l=1}^{j} \sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_{x,l}^i.$$

It is therefore sufficient to show that for each $j = 1, \dots, n$ and $1 \le l \le j$, and every $\delta > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left[\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_{x,l}^i \ge \delta S'(\lfloor Nt_j \rfloor)\right] = 0,$$

The probability above is bounded by

$$\mathbb{P}\left[(1+\delta)\left(\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_{x,l}^i\right) \ge \delta\left(\sum_{x\in R(\lfloor Nt_j \rfloor)\setminus M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_{x,j(x)}^i + \sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_{x,l}^i\right)\right].$$

Note that, by definition of the random variables $\tau_{x,j}^i$, requiring the above probability to tend to 0 as $N \to \infty$ is the same as requiring

$$\lim_{N\to\infty} \mathbb{P}\left[\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_x^i \ge \delta S(\lfloor Nt_j \rfloor)\right] = 0, \tag{3.10}$$

for each $j = 1, \dots, n$ and $1 \le l \le j$, and every $\delta > 0$, where here S is the original clock process, i.e. the sum of the τ_x^i's which have distributions π_x.

Fix $\epsilon > 0$ small, set K as in (3.4), recall the definition of $\mathcal{F}_{N,K}$ from (3.2) (with t_j instead of t), and write

$$\begin{aligned}
&\mathbb{P}\left[\sum_{x\in M_N^l} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_x^i \ge \delta S(\lfloor Nt_j \rfloor)\right] \\
&\le \mathbb{P}\left[\sum_{x\in M_N^l\setminus\mathcal{F}_{N,K}} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_x^i \ge \frac{\delta}{2} S(\lfloor Nt_j \rfloor)\right] + \mathbb{P}\left[\sum_{x\in M_N^l\cap\mathcal{F}_{N,K}} \sum_{i=1}^{L(x,\lfloor Nt_j \rfloor-1)} \tau_x^i \ge \frac{\delta}{2} S(\lfloor Nt_j \rfloor)\right].
\end{aligned} \tag{3.11}$$

By Lemma 3.3 we can choose N large enough such that $\mathbb{P}[|M_N^l| > \epsilon N/\ell^*(N)] \le \epsilon$. Then the first term on the right-hand side of (3.11) is bounded by

$$\mathbb{P}\left[\sum_{x\in M_N^l\setminus\mathcal{F}_{N,K}} \sum_{i=1}^{L(x,\lfloor Nt_j\rfloor-1)} \tau_x^i \ge \frac{\delta}{2}S(\lfloor Nt_j\rfloor),\ |M_N^l| \le \epsilon N/\ell^*(N)\right] + \epsilon$$

$$= \mathbb{P}\left[\sum_{k\in\mathcal{B}_1} \tilde{\tau}_k \ge \frac{\delta}{2}S(\lfloor Nt_j\rfloor),\ |M_N^l| \le \epsilon N/\ell^*(N)\right] + \epsilon.$$

Here $\mathcal{B}_1$ is the set of all times where a vertex in $M_N^l \setminus \mathcal{F}_{N,K}$, i.e. with $L(x, \lfloor Nt_j\rfloor - 1) \le K\ell^*(N)$ is visited. But if $|M_N^l| \le \epsilon N/\ell^*(N)$, then $|\mathcal{B}_1| \le \epsilon KN$. Since $\epsilon K \to 0$ as $\epsilon \to 0$ by the definition of K, we can apply Lemma 3.5 to get that the first term on the right-hand side of (3.11) converges to 0 when $N \to \infty$ and then $\varepsilon \to 0$.

The second term on the right-hand side of (3.11) can be bounded similarly. Recalling $F_{N,K}$ (for t_j) from (3.3), it is bounded from above by

$$\mathbb{P}[F_{N,K} \ge \epsilon N] + \mathbb{P}\left[\sum_{k\in\mathcal{B}_F} \tilde{\tau}_k \ge \frac{\delta}{2}S(\lfloor Nt_j\rfloor), F_{N,K} \le \epsilon N\right].$$

Here $\mathcal{B}_F$ is the set of times where a frequently visited vertex is visited, i.e. $|\mathcal{B}_F| = F_{N,K}$. Applying Lemma 3.4 to the first term and Lemma 3.5 to the second, this converges to zero as $N \to \infty$ and $\varepsilon \to 0$, and (3.10) follows. This finishes the proof of the lemma. □

4. Proof of Theorem 1.2

The goal of this section is to prove the classification theorem for the RTRW, Theorem 1.2. This will be done using Theorem 1.1. At first we should however show that the assumptions of Theorem 1.2 allow to verify the hypotheses of Theorem 1.1.

Proposition 4.1 ($d \ge 1$)**.** *Let X_N be as in* (1.5)*. Suppose that Assumption B holds and that X_N converge in the sense of Theorem 1.2. Then the clock processes S_N, defined as in* (1.2)*, converge in $\mathbb{P}$-distribution on $D^d(M_1')$ to some process $\mathcal{S}$. If $\mathcal{S}(0) = 0$, then the convergence holds with respect to the Skorohod M_1-topology.*

We first use this proposition to show Theorem 1.2.

Proof of Theorem 1.2. By Proposition 4.1, S_N converge to some process $\mathcal{S}$ in distribution on $D^d(M_1')$. This convergence implies the convergence of $S_N(t)$ to $\mathcal{S}(t)$ for all but countably many $t \in [0, \infty)$, cf. [Whi02, Theorem 11.6.6 and Corollary 12.2.1]. The non-triviality assumption (1.6) implies (1.3). We can thus apply Theorem 1.1. By this theorem there are only two possibilities, either $\mathcal{S}(t) = Mt$ or $\mathcal{S}(t) = V_\alpha(t)$. Since in both cases $\mathcal{S}(0) = 0$, the convergence of S_N actually holds in the M_1-topology.

The possible limits $\mathcal{S}$ are in the subspace $D^1_{u,\uparrow\uparrow}$ of unbounded strictly increasing functions from $[0, \infty)$ to $\mathbb{R}$, and their inverses are continuous. By [Whi02, Corollary 13.6.4], the inverse map from the space $D^1_{u,\uparrow}(M_1)$ of unbounded non-decreasing functions to $D^1(J_1)$ is continuous at $D^1_{u,\uparrow\uparrow}$, therefore S_N^{-1} converge to $\mathcal{S}^{-1}$ in $\mathbb{P}$-distribution on $D^1(J_1)$. Moreover, the rescaled random walks Y_N converge in $\mathbb{P}$-distribution on $D^d(J_1)$ to a standard d-dimensional Brownian motion B.

To proceed, we need to show that B and the limit clock process $\mathcal{S}$ are independent. This is trivial for the case $\mathcal{S}(t) = Mt$, so we may assume that $\mathcal{S} = V_\alpha$. We will use [Kal02, Lemma 15.6] which applied to our situation states that if $B, \mathcal{S}$ are such that

$B(0) = \mathcal{S}(0) = 0$ and the process $(B, \mathcal{S}) \in D^{d+1}$ has independent increments and no fixed jumps, $\mathcal{S}$ is a.s. a step process and $\Delta B \cdot \Delta \mathcal{S} = 0$ a.s., then B and $\mathcal{S}$ are independent. The only assumption that remains to be verified is that $(B, \mathcal{S})$ has jointly independent increments.

For fixed times $0 = t_0 < t_1 < \cdots < t_n = t$, consider the version $\tilde{S}(\lfloor Nt \rfloor)$ from (3.8) in the proof of Lemma 3.2. We have seen that every increment $\tilde{S}(\lfloor Nt_i \rfloor) - \tilde{S}(\lfloor Nt_{i-1} \rfloor)$ is independent of the increments $\{\xi_k : k \notin [\lfloor Nt_{j-1} \rfloor, \lfloor Nt_j \rfloor - 1]\}$ of the random walk Y. Since there is such version $\tilde{S}(\lfloor Nt \rfloor)$ for every choice of times t_j, and every such $\tilde{S}(\lfloor Nt \rfloor)$ converges to $\mathcal{S}$ after normalization, we obtain that for the limit $\mathcal{S}$ every increment $\mathcal{S}(t) - \mathcal{S}(s)$ is independent of $\{B(u) : u \notin [t, s]\}$. Since both B and $\mathcal{S}$ have independent increments, this implies that $(B, \mathcal{S})$ has jointly independent increments. Applying [Kal02, Lemma 15.6] it follows that the two limit processes B and $\mathcal{S}$, and thus also B and $\mathcal{S}^{-1}$ are independent.

It follows that (Y_N, S_N^{-1}) converge in distribution on $D^d(J_1) \times D^1_{u,\uparrow}(J_1)$ to $(B, \mathcal{S}^{-1})$. By [Whi02, Theroem 13.2.2], the composition map from $D^d(J_1) \times D^1_{u,\uparrow}(J_1)$ to $D^d(J_1)$ taking $(y(t), s(t))$ to $y(s(t))$ is continuous at (y, s) if y is continuous and s non-decreasing. From this we conclude that the compositions $X_N(t) = Y_N(S_N^{-1}(t))$ converge in distribution on $D^d(J_1)$ to $B(\mathcal{S}^{-1}(t))$ as required. □

For the proof of Proposition 4.1 we will relate the clock process S to the quadratic variation process of the RTRW X and then apply [JS03, Corollary VI.6.29] which states that under some conditions, whenever a sequence of processes converges in distribution, then so does the sequence of their quadratic variations.

We need some definitions first. For a d-dimensional pure-jump process Z, let $Z^{(i)}$ denote the i-th coordinate of Z, and let $\Delta Z^{(i)}(t) = Z^{(i)}(t) - Z^{(i)}(t-)$ be the jump size of $Z^{(i)}$ at time t. The quadratic variation process $[Z, Z]_t$ is a $d \times d$ matrix-valued process, where the (i, j)-th entry is the quadratic covariation of the i-th and j-th coordinate of Z, which is

$$[Z^{(i)}, Z^{(j)}]_t = \sum_{0<s\leq t} \Delta Z^{(i)}(s) \Delta Z^{(j)}(s).$$

We proceed by relating the inverse S_N^{-1} of the clock process to the quadratic variation process of X_N.

Lemma 4.2. *Let $[X_N, X_N]_t$ be the quadratic variation process of X_N, and define $\sigma^2 = \mathbb{E}[|\mathcal{A}\xi_j|^2]$ (recall (1.4) and (1.5) for the notation). Then for every $t > 0$,*

$$\frac{\operatorname{trace}[X_N, X_N]_t}{\sigma^2 S_N^{-1}(t)} \xrightarrow{N\to\infty} 1 \text{ in } \mathbb{P}\text{-probability.}$$

Proof. Easy computation yields

$$\operatorname{trace}[X_N, X_N]_t = \sum_{i=1}^{d} \sum_{0<s\leq t} (\Delta X_N^{(i)}(t))^2 = \frac{1}{N} \sum_{j \leq S^{-1}(a_N t)} |\mathcal{A}\xi_j|^2.$$

The process S^{-1} has increments of size 1, and since the times between increments are a.s. finite, $S^{-1}(a_N t) \nearrow \infty$ a.s. as $N \to \infty$. Therefore, since $\sigma^2 = \mathbb{E}[|\mathcal{A}\xi_j|^2] < \infty$ by Assumption B, the law of large numbers implies

$$\mathbb{P}\left[\left|\frac{N}{S^{-1}(a_N t)} \operatorname{trace}[X_N, X_N]_t - \sigma^2\right| > \epsilon\right] \longrightarrow 0 \text{ as } N \to \infty \text{ for every } \epsilon > 0.$$

Noting that $\frac{1}{N} S^{-1}(a_N t) = S_N^{-1}(t)$ finishes the proof. □

We now check that the assumptions for [JS03, Corollary VI.6.29] are fullfilled.

Lemma 4.3. *If Assumption B holds, then the rescaled processes X_N are local martingales with bounded increments.*

Proof. The increments are bounded by Assumption B. The local martingale property is unaffected by linear scaling, it is hence sufficient to prove it for the process X.

We show that the sequence of stopping times $\sigma_l = \inf\{t > 0 : S^{-1}(t) \geq l\}$, $l \geq 1$, is a localizing sequence for X, i.e. we show that $(X(t \wedge \sigma_l))_{t \geq 0}$ is a martingale for every $l \geq 1$.

We introduce the filtration $\mathcal{F}_t = \sigma(Y(k), S(k) : k \leq t)$. Obviously, Y is an $\mathcal{F}$-martingale, and $S^{-1}(t)$ is an $\mathcal{F}$-stopping time for every $t \geq 0$, with $S^{-1}(t) \geq S^{-1}(s)$ for $t \geq s$. The natural filtration for X, $\mathcal{G}_t = \mathcal{F}_{S^{-1}(t)}$, is right-continuous (see [Kal02, Proposition 7.9]), and X is $\mathcal{G}$-adapted. The sequence of random variables σ_l is indeed an increasing sequence of $\mathcal{G}$-stopping times (σ_l is the time at which the process X jumps for the l-th time). Moreover, by definition $S^{-1}(t \wedge \sigma_l) = S^{-1}(t) \wedge S^{-1}(\sigma_l) \leq S^{-1}(\sigma_l) = l$. Applying Doob's optional sampling theorem (see e.g. [Kal02, Theorem 7.12]) to the discrete-time martingale Y and the bounded stopping time $S^{-1}(t \wedge \sigma_l)$, we obtain

$$\mathbb{E}\left[X(t \wedge \sigma_l) \mid \mathcal{G}_s\right] = \mathbb{E}\left[Y\left(S^{-1}(t \wedge \sigma_l)\right) \mid \mathcal{F}_{S^{-1}(s)}\right] = Y\left(S^{-1}(t \wedge \sigma_l) \wedge S^{-1}(s)\right) = X(s \wedge \sigma_l).$$

This completes the proof. □

We can now prove Proposition 4.1.

Proof of Proposition 4.1. By Lemma 4.3, X_N are local martingales with bounded increments. [JS03, Corollary VI.6.29] then implies that the quadratic variation processes $[X_N, X_N]_t$ converge component-wise on $D^1(J_1)$ to the quadratic variation process $[\mathcal{X}, \mathcal{X}]_t$ of $\mathcal{X}$. Since all jumps of the processes $[X_N^{(i)}, X_N^{(i)}]_t$, $i = 1, \ldots, d$, are positive, [Whi02, Theorem 12.7.3 (continuity of addition at limits with jumps of common sign)] yields that $\operatorname{trace}[X_N, X_N]_t$ converges to some non-decreasing process in $D^1(M_1)$. From Lemma 4.2 it then follows that the inverses S_N^{-1} of the rescaled clock processes converge to some non-decreasing process $\mathcal{S}^{-1}(t)$ in $D^1(M_1)$.

For non-decreasing functions $x \in D^1$ the right-continuous inverse satisfies $(x^{-1})^{-1} = x$, and thus $S_N = (S_N^{-1})^{-1}$. Hence, by [Whi02, Theorem 13.6.1], which ensures the continuity of the inverse operation, S_N converges to $\mathcal{S}$ in $D^1(M_1)$ provided that $\mathcal{S}(0) = (\mathcal{S}^{-1})^{-1}(0) = 0$.

If we do not know whether $\mathcal{S}(0) = 0$, this theorem does not apply. This issue can be solved by weakening the topology from M_1 to M_1' (see [Whi02, Section 13.6.2] for details). In particular, [Whi02, Theorem 13.6.2] yields that S_N converge to $\mathcal{S}$ in distribution in $D^1(M_1')$. □

5. Proofs of sufficiency criteria

Theorem 1.3, giving a sufficient criterium for convergence to Brownian motion, is an immediate consequence of the ergodicity of the sequence of successive waiting times.

Proof of Theorem 1.3. Consider $\tilde{\tau} = (\tilde{\tau}_k)_{k \geq 0}$ and let $\theta : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ be the left-shift along the sequence, which by Lemma 2.1 acts ergodically along $\tilde{\tau}$.

If $\mathbb{E}[\tilde{\tau}_0] = M$ is finite, the function $f(\tilde{\tau}) = \tilde{\tau}_0$ is integrable, and we can apply the ergodic theorem to f to get

$$\lim_{N\to\infty} \frac{1}{N} S(\lfloor Nt \rfloor) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{\lfloor Nt \rfloor -1} \tilde{\tau}_k = \lim_{N\to\infty} t \frac{1}{Nt} \sum_{k=0}^{\lfloor Nt \rfloor -1} f(\theta^k(\tilde{\tau}))$$
$$= t\mathbb{E}\left[f(\tilde{\tau})\right] = Mt \text{ almost surely.}$$

Thus we have that the rescaled clock processes S_N converge in distribution on $D^1(J_1)$ to Mt, where the normalization is $a_N = N$. If additionally Assumption B holds, using the same arguments as in the proof of Theorem 1.2 we conclude that the X_N converge and the limit $\mathcal{X}$ is as in case (i) of Theorem 1.2. □

Before starting the proof of Theorem 1.4, which deals with the convergence to the Fractional Kinetics, we briefly sketch some examples that illustrate how different functions f in condition (1.9) arise.

First, consider the CTRW defined in (1.7). Then the waiting times τ_x^i lie in the domain of attraction of an α-stable law, thus the quenched Laplace transform (which is deterministic here) satisfies

$$\hat{\pi}_0(\lambda/a_N) = \exp\{-c\,(\lambda/a_N)^\alpha\,(1+o(1))\} \qquad \text{as } N\to\infty$$

for some $c > 0$. Taking this to the power $r\ell^*(N)$ it follows that the CTRW satisfies condition (1.9) with $a_N = c^{1/\alpha}N^{1/\alpha}$ and $f(r) = r$.

Secondly, consider the following simplified Bouchaud trap model (cf. (1.8)). Let $\pi_x = \delta_{\tau_x}$ where the τ_x, $x \in \mathbb{Z}^d$, are heavy-tailed i.i.d. random variables, that is

$$\mathbb{P}[\tau_x > u] = u^{-\alpha}(1+o(1)) \text{ as } u\to\infty.$$

Then the quenched Laplace transform satisfies

$$\hat{\pi}_0(\lambda/a_N) = \exp\{-\lambda a_N^{-1}\tau_x\}.$$

Taking this to the power $r\ell^*(N)$ and taking the expectation over τ_x, this is the Laplace transform of a random variable in the domain of attraction of an α-stable law, evaluated at $\lambda r\frac{\ell^*(N)}{a_N}$, which is

$$\mathbb{E}\left[\hat{\pi}_0(\lambda/a_N)^{r\ell^*(N)}\right] = \exp\left\{-c\lambda^\alpha r^\alpha \left(\frac{\ell^*(N)}{a_N}\right)^\alpha (1+o(1))\right\} \text{ as } N\to\infty$$

for some $c > 0$. Choosing $a_N = c^{1/\alpha}N^{1/\alpha}\ell^*(N)^{1-1/\alpha}$, condition (1.9) is then satisfied for $f(r) = r^\alpha$.

To see that $f(r)$ can be more than just a power of r, consider the following mixture of the above two models. For some $p \in (0,1)$, let each π_x with probability p be a heavy-tailed distribution with

$$\mathbb{P}[\tau_x^i > u] = u^{-\alpha}\ell^*(u)^{\alpha-1}(1+o(1)) \text{ as } u\to\infty,$$

and with probability $1-p$ be δ_{τ_x} where the τ_x are heavy-tailed random variables with

$$\mathbb{P}[\tau_x > u] = u^{-\alpha}(1+o(1)) \text{ as } u\to\infty.$$

Then, by combining the arguments above, condition (1.9) is satisfied with the normalization $a_N = cN^{1/\alpha}\ell^*(N)^{1-1/\alpha}$ and $f(r) = pr + (1-p)r^\alpha$.

Proof of Theorem 1.4. By Theorem 1.1 it is sufficient to show that

$$\lim_{N\to\infty} \mathbb{E}[\exp\{-\lambda S_N(t)\}] = e^{-ct\lambda^\alpha} \tag{5.1}$$

for some $c \in (0,\infty)$, this is equivalent to convergence of S_N to an α-stable subordinator. Using the independence of the π_x's, recalling that $\hat{\pi}_x$ denotes the Laplace transform of π_x, we have

$$\begin{aligned}\mathbb{E}\Big[\exp\Big\{-\frac{\lambda}{a_N}S(\lfloor Nt\rfloor)\Big\}\Big|Y\Big] &= \mathbb{E}\Big[\exp\Big\{-\frac{\lambda}{a_N}\sum_{x\in\mathbb{Z}^d}\sum_{i=1}^{L(x,\lfloor Nt\rfloor-1)}\tau_x^i\Big\}\Big|Y\Big] \\ &= \prod_{x\in\mathbb{Z}^d}\mathbb{E}\Big[\hat{\pi}_x(\lambda/a_N)^{L(x,\lfloor Nt\rfloor-1)}\Big\}\Big|Y\Big].\end{aligned} \tag{5.2}$$

Treating the case when Y is transient first, let $R^k(Nt) = \{x\in\mathbb{Z}^d : L(x,\lfloor Nt\rfloor-1) = k\}$. By Lemma A.1, $R^k(Nt)/(Nt) \xrightarrow{N\to\infty} \gamma^2(1-\gamma)^{k-1}$ in probability. Using the translation invariance, the right-hand side of (5.2) can be written as

$$\exp\Big\{\sum_{k=1}^{\infty}|R_k(Nt)|\log\mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^k\big]\Big\}.$$

For arbitrary $M\in\mathbb{N}$, using the law of large numbers for $R^k(Nt)$ and assumption (1.9),

$$\sum_{k=1}^{M}|R^k(Nt)|\log\mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^k\big] \xrightarrow{N\to\infty} -t\lambda^\alpha\sum_{k=1}^{M} f(k/\gamma)\gamma^2(1-\gamma)^{k-1}, \tag{5.3}$$

in probability. Applying Jensen's inequality, it is easy to see that $f(k)$ grows at most linearly with k, so the right-hand side of the above expression converges as $M\to\infty$ to a finite value, by assumptions of the theorem. On the other hand, by Jensen's inequality again, for every $\delta > 0$

$$\begin{aligned}&\mathbb{P}\Big[-\sum_{k=M}^{\infty}|R^k(Nt)|\log\mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^k\big]\geq\delta\Big] \\ &\leq \mathbb{P}\Big[-\log\mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)\big]\sum_{k=M}^{\infty}|R^k(Nt)|k\geq\delta\Big].\end{aligned} \tag{5.4}$$

By the Markov inequality, for $0 < c_1 < -\log(1-\gamma)$, $\mathbb{P}[R^k(Nt)/(Nt) \geq e^{-c_1k}] \leq e^{-c'k}$ uniformly for all $k \geq M$ and N large enough, and thus by a union bound

$$\mathbb{P}\big[\exists k\geq M \text{ such that } R^k(Nt)/(Nt)\geq e^{-c_1k}\big] \leq Ce^{-c'M} \tag{5.5}$$

uniformly in N. Using (5.5) and the fact that $\log\mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)\big]$ is finite by assumption, it follows that the left-hand side of (5.4) converges to 0 in probability when $N\to\infty$ and then $M\to\infty$, and therefore (5.3) also holds with $M=\infty$. Using the bounded convergence theorem, it then follows that

$$\begin{aligned}\mathbb{E}\Big[\exp\Big\{-\frac{\lambda}{a_N}S(\lfloor Nt\rfloor)\Big\}\Big] &= \mathbb{E}\Big[\exp\Big\{\sum_{k=1}^{\infty}|R^k(Nt)|\log\mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^k\big]\Big\}\Big] \\ &\xrightarrow{N\to\infty}\exp\Big\{-t\lambda^\alpha\sum_{k=1}^{\infty}f(k/\gamma)\gamma^2(1-\gamma)^{k-1}\Big\},\end{aligned}$$

which proves (5.1) in the transient case.

To treat the recurrent case, we fix $\beta > 0$ small and define for $k \geq 1$

$$R_\beta^k(Nt) = \{x \in \mathbb{Z}^d : (k-1)\beta\ell^*(N) < L(x, \lfloor Nt \rfloor - 1) \leq k\beta\ell^*(N)\}.$$

By Lemma A.1, $R_\beta^k(Nt)\ell^*(N)/(Nt) \xrightarrow{N\to\infty} e^{-(k-1)\beta} - e^{-k\beta}$ in probability. The right-hand side of (5.2) can be bounded from above by

$$\exp\Big\{\sum_{k=1}^\infty |R_\beta^k(Nt)| \log \mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^{(k-1)\ell^*(N)}\big]\Big\},$$

and from below by

$$\exp\Big\{\sum_{k=1}^\infty |R_\beta^k(Nt)| \log \mathbb{E}\big[\hat{\pi}_0(\lambda/a_N)^{k\ell^*(N)}\big]\Big\}.$$

Following the same steps as in the transient case, it can be easily shown that

$$\begin{aligned}\exp\Big\{-t\lambda^\alpha \sum_{k=1}^\infty f(\beta(k-1))\big(e^{-(k-1)\beta} - e^{-k\beta}\big)\Big\} &\leq \lim_{N\to\infty} \mathbb{E}\big[e^{-\lambda S_N(t)}\big] \\ \leq \exp\Big\{-t\lambda^\alpha \sum_{k=1}^\infty f(\beta k)\big(e^{-(k-1)\beta} - e^{-k\beta}\big)\Big\}&\end{aligned}$$

Since f is a monotonous function, the sums in the above expression can be viewed as lower and upper Riemann sums for the integral $\int_0^\infty f(x)e^{-x}\,dx$ to which they tend when $\beta \to 0$. This integral is finite by assumption, and (5.1) is proved in the recurrent case. □

6. Ignoring small sets

In this section we prove Lemma 3.5 which allows us to ignore small sets when dealing with the clock process.

We first assume that the random walk Y is transient, that is $1/\ell^*(n) \to \gamma \in (0,1)$ as $n \to \infty$. We start by noting that for every $x \in R(\lfloor Nt \rfloor)$ and $i \in \{1, \dots, L(x, \lfloor Nt \rfloor - 1)\}$, since $(\tau_x^i)_{i\geq 1}$ are i.i.d.,

$$\mathbb{E}\left[\frac{\tau_x^i}{S(\lfloor Nt \rfloor)} \;\Big|\; Y\right] = \mathbb{E}\left[\frac{\tau_x^1}{S(\lfloor Nt \rfloor)} \;\Big|\; Y\right]. \tag{6.1}$$

For fixed $0 \leq l < \lfloor Nt \rfloor$, let $x = Y(l)$ and $i = L(Y(l), l)$, that is $\tilde{\tau}_l = \tau_x^i$. Using (6.1) and the fact that $(\tau_x^1)_{x\in\mathbb{Z}^d}$ are i.i.d. under $\mathbb{P}$,

$$\begin{aligned}\mathbb{E}\left[\frac{\tilde{\tau}_l}{S(\lfloor Nt \rfloor)} \;\Big|\; Y\right] &= \mathbb{E}\left[\frac{\tau_x^i}{\sum_{y\in R(\lfloor Nt \rfloor)} \sum_{j=1}^{L(x,\lfloor Nt \rfloor - 1)} \tau_y^j} \;\Big|\; Y\right] \\ &\leq \mathbb{E}\left[\frac{\tau_x^1}{\sum_{y\in R(\lfloor Nt \rfloor)} \tau_y^1} \;\Big|\; Y\right] \\ &= \frac{1}{|R(\lfloor Nt \rfloor)|} \sum_{x\in R(\lfloor Nt \rfloor)} \mathbb{E}\left[\frac{\tau_x^1}{\sum_{y\in R(\lfloor Nt \rfloor)} \tau_y^1} \;\Big|\; Y\right] \\ &= \frac{1}{|R(\lfloor Nt \rfloor)|}.\end{aligned} \tag{6.2}$$

By the law of large numbers for $R(n)$ (Lemma A.1) in the transient case, there is a constant $C < \infty$ such that for all N large enough

$$\mathbb{P}\left[|R(\lfloor Nt \rfloor)| < CN\right] < \epsilon.$$

Hence, for N large enough,

$$\mathbb{P}\bigg[\sum_{l\in\mathcal{B}} \tilde{\tau}_l \geq \delta S(\lfloor Nt\rfloor),\ |\mathcal{B}| \leq \epsilon N\bigg] \leq \mathbb{P}\bigg[\sum_{l\in\mathcal{B}} \tilde{\tau}_l \geq \delta S(\lfloor Nt\rfloor),\ |\mathcal{B}| \leq \epsilon N,\ |R(\lfloor Nt\rfloor)| \geq CN\bigg] + \epsilon.$$

Using the Markov inequality and (6.2), this is bounded from above by

$$\begin{aligned}
&\leq \frac{1}{\delta}\mathbb{E}\left[\sum_{l\in\mathcal{B}} \mathbb{E}\left[\frac{\tilde{\tau}_l}{S(\lfloor Nt\rfloor)} \;\Big|\; Y\right] \mathbf{1}_{\{|\mathcal{B}|\leq \epsilon N\}} \mathbf{1}_{\{|R(\lfloor Nt\rfloor)|\geq CN\}}\right] + \epsilon. \\
&\leq \frac{1}{\delta}\mathbb{E}\left[\frac{|\mathcal{B}|}{|R(\lfloor Nt\rfloor)|} \mathbf{1}_{\{|\mathcal{B}|\leq \epsilon N\}} \mathbf{1}_{\{|R(\lfloor Nt\rfloor)|\geq CN\}}\right] + \epsilon \\
&\leq \frac{\epsilon}{C\delta} + \epsilon.
\end{aligned}$$

Letting $N \to \infty$ and then $\epsilon \to 0$ completes the proof of the lemma in the transient case.

We now consider the recurrent case. Let $R_{\mathcal{B}} = \{Y(l) : l \in \mathcal{B}\}$, and for $x \in R_{\mathcal{B}}$ let $L_{\mathcal{B}}(x) = |\{l \in \mathcal{B} : Y(l) = x\}|$. Fix some small $\beta > 0$ and let

$$\begin{aligned}
R_{>\beta} &= \{x \in R(\lfloor Nt\rfloor) : L(x, \lfloor Nt\rfloor - 1) > \beta \ell^*(N)\}, \\
R_{\leq\beta} &= \{x \in R(\lfloor Nt\rfloor) : L(x, \lfloor Nt\rfloor - 1) \leq \beta \ell^*(N)\}.
\end{aligned}$$

By Lemma A.1, the sizes of $R_{>\beta}$ and $R_{\leq\beta}$ satisfy weak laws of large numbers with respective averages $Nte^{-\beta}/\ell^*(N)(1 + o(1))$ and $Nt(1 - e^{-\beta})/\ell^*(N)(1 + o(1))$. In particular for $C_\beta = (1 - \epsilon)e^{-\beta}t$ and $c_\beta = (1 + \epsilon)(1 - e^{-\beta})t$, for all N large enough,

$$\mathbb{P}\left[|R_{>\beta}| < C_\beta \frac{N}{\ell^*(N)}\right] + \mathbb{P}\left[|R_{\leq\beta}| > c_\beta \frac{N}{\ell^*(N)}\right] \leq \epsilon.$$

Therefore, for N large enough,

$$\mathbb{P}\bigg[\sum_{l\in\mathcal{B}} \tilde{\tau}_l \geq \delta S(\lfloor Nt\rfloor),\ |\mathcal{B}| \leq \epsilon N\bigg]$$

$$\leq \mathbb{P}\bigg[\sum_{x\in R_{\mathcal{B}}\cap R_{>\beta}} \sum_{i=1}^{L_{\mathcal{B}}(x)} \tau_x^i \geq \frac{\delta}{2} S(\lfloor Nt\rfloor),\ |\mathcal{B}| \leq \epsilon N,\ |R_{>\beta}| \geq C_\beta \frac{N}{\log N}\bigg] \tag{6.3}$$

$$+ \mathbb{P}\bigg[\sum_{x\in R_{\mathcal{B}}\cap R_{\leq\beta}} \sum_{i=1}^{L_{\mathcal{B}}(x)} \tau_x^i \geq \frac{\delta}{2} S(\lfloor Nt\rfloor),\ |R_{>\beta}| \geq C_\beta \frac{N}{\log N},\ |R_{\leq\beta}| \leq c_\beta \frac{N}{\log N}\bigg] + \epsilon. \tag{6.4}$$

Using (6.1) and the similar reasoning as in the transient case, since $(\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i)_{x\in\mathbb{Z}^d}$ are i.i.d. with respect to the annealed measure, we have for $x\in R_{\mathcal{B}}\cap R_{>\beta}$,

$$\begin{aligned}\mathbb{E}\left[\frac{\sum_{i=1}^{L_{\mathcal{B}}(x)}\tau_x^i}{S(\lfloor Nt\rfloor)}\;\Big|\;Y\right] &= \frac{L_{\mathcal{B}}(x)}{\beta\ell^*(N)}\mathbb{E}\left[\frac{\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i}{S(\lfloor Nt\rfloor)}\;\Big|\;Y\right]\\ &\le \frac{L_{\mathcal{B}}(x)}{\beta\ell^*(N)}\mathbb{E}\left[\frac{\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i}{\sum_{y\in R_{>\beta}}\sum_{i=1}^{\beta\ell^*(N)}\tau_y^i}\;\Big|\;Y\right]\\ &= \frac{L_{\mathcal{B}}(x)}{|R_{>\beta}|\beta\ell^*(N)}.\end{aligned}$$

Therefore, using the Markov inequality,

$$\begin{aligned}(6.3) &\le \frac{2}{\delta}\mathbb{E}\left[\sum_{x\in R_{\mathcal{B}}\cap R_{>\beta}}\mathbb{E}\left[\frac{\sum_{i=1}^{L_{\mathcal{B}}(x)}\tau_x^i}{S(\lfloor Nt\rfloor)}\;\Big|\;Y\right]\mathbf{1}_{\{|\mathcal{B}|\le\epsilon N\}}\mathbf{1}_{\left\{|R_{>\beta}|\ge C_\beta\frac{N}{\ell^*(N)}\right\}}\right]\\ &\le \frac{2}{\delta}\mathbb{E}\left[\sum_{x\in R_{\mathcal{B}}\cap R_{>\beta}}\frac{L_{\mathcal{B}}(x)}{|R_{>\beta}|\beta\ell^*(N)}\mathbf{1}_{\{|\mathcal{B}|\le\epsilon N\}}\mathbf{1}_{\left\{|R_{>\beta}|\ge C_\beta\frac{N}{\ell^*(N)}\right\}}\right]\\ &\le \frac{2\epsilon}{\delta\beta C_\beta}.\end{aligned}\tag{6.5}$$

where for the last inequality we used the fact that $\sum_x L_{\mathcal{B}}(x)\le|\mathcal{B}|\le\varepsilon N$.

It remains to bound (6.4). Using again the fact that $(\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i)_{x\in\mathbb{Z}^d}$ are i.i.d. with respect to the annealed measure and independent of Y,

$$\begin{aligned}&\mathbb{P}\Big[\sum_{x\in R_{\mathcal{B}}\cap R_{\le\beta}}\sum_{i=1}^{L_{\mathcal{B}}(x)}\tau_x^i\ge\frac{\delta}{2}S(\lfloor Nt\rfloor)\;\Big|\;Y\Big]\\ &\le \mathbb{P}\Big[\left(1+\frac{\delta}{2}\right)\sum_{x\in R_{\mathcal{B}}\cap R_{\le\beta}}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i\ge\frac{\delta}{2}\sum_{x\in R_{>\beta}\cup(R_{\mathcal{B}}\cap R_{\le\beta})}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i\;\Big|\;Y\Big]\\ &\le \frac{2+\delta}{\delta}\mathbb{E}\Big[\frac{\sum_{x\in R_{\mathcal{B}}\cap R_{\le\beta}}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i}{\sum_{x\in R_{>\beta}\cup(R_{\mathcal{B}}\cap R_{\le\beta})}\sum_{i=1}^{\beta\ell^*(N)}\tau_x^i}\;\Big|\;Y\Big]\\ &= \frac{2+\delta}{\delta}\frac{|R_{\mathcal{B}}\cap R_{\le\beta}|}{|R_{>\beta}\cup(R_{\mathcal{B}}\cap R_{\le\beta})|}.\end{aligned}$$

Therefore,

$$(6.4)\le\frac{2+\delta}{\delta}\frac{c_\beta}{C_\beta}=\frac{2+\delta}{\delta}\frac{1+\epsilon}{1-\epsilon}\left(e^\beta-1\right).\tag{6.6}$$

Combining (6.3)–(6.6) and letting $N\to\infty$, then $\epsilon\to0$ and finally $\beta\to0$ finishes the proof of the lemma in the recurrent case. □

Appendix A. Laws of large numbers for range-like objects

We prove here that Assumption A implies weak laws of large numbers for several range-related quantities. The proofs are based on the classical paper [DE51], see also [Rév13, Chapter 21].

Recall that

$$R(n) = \{x \in \mathbb{Z}^d : \ L(x, n-1) > 0\}$$

is the range of the random walk Y up to time $n-1$. In the recurrent case, i.e. if $\ell^*(n) \to \infty$, define for $k \geq 1$ and $\beta > 0$

$$R^k_\beta(n) = \{x \in \mathbb{Z}^d : \ L(x, n-1) \in ((k-1), k]\beta\ell^*(n)\}$$

the set of vertices visited $(k-1)\beta\ell^*(n)$ to $k\beta\ell^*$ times up to time $n-1$. In the transient case, if $1/\ell^*(n) \to \gamma \in (0,1)$, let for $k \geq 1$

$$R^k(n) = \{x \in \mathbb{Z}^d : \ L(x, n-1) = k\}$$

the vertices visited exactly k times up to time $n-1$.

We say that a sequence of random variables Z_n satisfies the weak law of large numbers if $Z_n/EZ_n \xrightarrow{n\to\infty} 1$ in probability.

Lemma A.1. *(i) If Assumption A holds, then $|R(n)|$ satisfies the weak law of large numbers with*

$$\mathbb{E}[|R(n)|] = \frac{n}{\ell^*(n)}(1 + o(1)) \ \textit{as } n \to \infty.$$

(ii) If in addition $\ell^*(n) \to \infty$ as $n \to \infty$, then $|R^k_\beta(n)|$ satisfies the weak law of large numbers for every $k \geq 1$ and $\beta > 0$, and

$$\mathbb{E}[|R^k_\beta(n)|] = (e^{-(k-1)\beta} - e^{-k\beta})\frac{n}{\ell^*(n)}(1 + o(1)) \ \textit{as } n \to \infty.$$

(iii) If, on the other hand, $1/\ell^*(n) \to \gamma \in (0,1)$, then $|R^k(n)|$ satisfies the weak law of large numbers for every $k \geq 1$, and

$$\mathbb{E}[|R^k(n)|] = \gamma^2(1-\gamma)^{k-1}n(1 + o(1)) \ \textit{as } n \to \infty.$$

Proof. Note that for the simple random walk in $d \geq 3$ and $d = 2$ respectively, part (i) is a classical result from [DE51], part (iii) was hinted at in [ET60, Theorem 12] and proved in [Pit74], whereas part (ii) is a direct consequence of [DE51, Theorem 4] and [Čer07, Theorem 2]. Part (i) above is proved exactly as in [DE51]. We include its proof, since proofs of (ii) and (iii) are its extensions. Let ψ_k be the indicator of the event that a new vertex is found at time k,

$$\psi_k = \mathbf{1}_{\{Y(l) \neq Y(k) \text{ for all } 0 \leq l < k\}},$$

with $\psi_0 = 1$. Recall that ξ_i denote the i.i.d. increments of the random walk Y. Then,

$$\begin{aligned}
\mathbb{E}[\psi_k] &= \mathbb{P}\left[Y(k) \neq Y(k-1), \ Y(k) \neq Y(k-2), \ldots, \ Y(k) \neq Y(0)\right] \\
&= \mathbb{P}\left[\xi_k \neq 0, \ \xi_k + \xi_{k-1} \neq 0, \ldots, \ \xi_k + \cdots + \xi_1 \neq 0\right] \\
&= \mathbb{P}\left[\xi_1 \neq 0, \ \xi_1 + \xi_2 \neq 0, \ldots, \ \xi_1 + \cdots + \xi_k \neq 0\right] \\
&= \mathbb{P}[Y(l) \neq 0 \text{ for } l = 1, \ldots, k] = r_k.
\end{aligned} \tag{A.1}$$

For a slowly varying function ℓ, $\sum_{k=1}^n \ell(k) = n\ell(n)(1 + o(1))$ as $n \to \infty$ (see e.g. [Sen76, p. 55]). Therefore, by Assumption A,

$$\mathbb{E}[|R(n)|] = \sum_{k=0}^{n-1} \mathbb{E}[\psi_k] = \frac{n}{\ell^*(n)}(1 + o(1)) \text{ as } n \to \infty. \tag{A.2}$$

To prove the weak law of large numbers, we compute the variance. First note that for $i \le j$, by the Markov property,

$$\begin{aligned}\mathbb{E}[\psi_i \psi_j] &= \mathbb{E}\left[\mathbf{1}_{\{Y(l)\neq Y(i),\ 0\le l<i\}}\mathbf{1}_{\{Y(l)\neq Y(j),\ 0\le l<j\}}\right] \\ &\le \mathbb{E}\left[\mathbf{1}_{\{Y(l)\neq Y(i),\ 0\le l<i\}}\mathbf{1}_{\{Y(l)\neq Y(j),\ i\le l<j\}}\right] = \mathbb{E}[\psi_i]\mathbb{E}[\psi_{j-i}].\end{aligned} \tag{A.3}$$

Then,

$$\begin{aligned}\operatorname{Var}|R(n)| &= \sum_{0\le i,j\le n-1} \mathbb{E}[\psi_i\psi_j] - \mathbb{E}[\psi_i]\mathbb{E}[\psi_j] \\ &\le 2\sum_{i=0}^{n-1}\sum_{j=i}^{n-1} \mathbb{E}[\psi_i]\left(\mathbb{E}[\psi_{j-i}] - \mathbb{E}[\psi_j]\right) \\ &\le 2\sum_{i=0}^{n-1} \mathbb{E}[\psi_i]\left(\max_{k=0,\dots,n-1}\sum_{j=k}^{n-1}\mathbb{E}[\psi_{j-k}] - \mathbb{E}[\psi_j]\right).\end{aligned} \tag{A.4}$$

By (A.1), $\mathbb{E}[\psi_k]$ is non-increasing, therefore the maximum in (A.4) is attained in $k = \frac{n}{2}$. The parenthesis in (A.4) can then be estimated using elementary properties of slowly varying functions,

$$\begin{aligned}\sum_{j=\frac{n}{2}}^{n-1}\mathbb{E}[\psi_{j-\frac{n}{2}}] - \mathbb{E}[\psi_j] &= \sum_{j=0}^{\frac{n}{2}-1}\frac{1}{\ell^*(j)} - \sum_{j=\frac{n}{2}}^{n-1}\frac{1}{\ell^*(j)} \\ &= \sum_{j=0}^{\frac{n}{2}-1}\frac{1}{\ell^*(j)} - \left(\sum_{j=0}^{n-1}\frac{1}{\ell^*(j)} - \sum_{j=0}^{\frac{n}{2}-1}\frac{1}{\ell^*(j)}\right) \\ &= 2\frac{\frac{n}{2}}{\ell^*(\frac{n}{2})}(1+o(1)) - \frac{n}{\ell^*(n)}(1+o(1)) = \frac{n}{\ell^*(n)}o(1) \text{ as } n\to\infty.\end{aligned}$$

Inserting this into (A.4), we obtain

$$\operatorname{Var}|R(n)| \le 2\sum_{i=0}^{n-1}\mathbb{E}[\psi_i]\frac{n}{\ell^*(n)}o(1) = o\left(\left(\frac{n}{\ell^*(n)}\right)^2\right) \text{ as } n\to\infty,$$

and the weak law of large numbers for $|R(n)|$ follows by usual arguments.

Before turning to part (ii), we note the following fact on return times. Let as before $H_0^1 = \inf\{i > 0 :\ Y(i) = 0\}$ denote the time of the first return to 0, and $H_0^k = \inf\{i > H_0^{k-1} :\ Y(i) = 0\}$ the time of the k-th return to 0. Let $T_i = H_0^i - H_0^{i-1}$ (with $H_0^0 = 0$) be the successive return times. By the Markov property the $(T_i)_{i\ge1}$ are i.i.d., and $\mathbb{P}[T_i > n] = r_n = \frac{1}{\ell^*(n)}$ by Assumption A. If $\ell^*(k)\to\infty$, the T_i are a.s. finite and have slowly varying tail. It is well known (e.g. [Dar52, Theorem 3.2]) that for such i.i.d. random variables T_i,

$$\frac{\sum_{i=1}^n T_i}{\max_{i=1}^n T_i} \to 1 \text{ in probability as } n\to\infty. \tag{A.5}$$

Since $\ell^*(cn)\sim\ell^*(n)$ as $n\to\infty$,

$$\mathbb{P}\big[\max\{T_i : 1\le i\le \beta\ell^*(n)\}\le cn\big] = \left(1-\frac{1}{\ell^*(cn)}\right)^{\beta\ell^*(n)} = e^{-\beta}(1+o(1)) \text{ as } n\to\infty. \tag{A.6}$$

From (A.5) and (A.6) we obtain for every $c > 0$ and $\beta > 0$

$$\mathbb{P}\left[L(0, cn) \geq \beta\ell^*(n)\right] = \mathbb{P}\left[\sum_{i=1}^{\beta\ell^*(n)} T_i \leq cn\right] = e^{-\beta}(1 + o(1)). \tag{A.7}$$

For part (ii) we only prove the statement for $R_\beta(n) = R^1_\beta(n)$, the statement for $k > 1$ follows easily by subtracting the claims with β replaced by βk and $\beta(k-1)$. Consider ψ_k as above, and additionally define functions $\varphi_k = \mathbf{1}_{\{L(Y(k), n-1) \leq \beta\ell^*(n)\}}$. Using the Markov property and translation invariance,

$$\begin{aligned}\mathbb{E}[|R_\beta(n)|] &= \sum_{k=0}^{n-1} \mathbb{E}[\psi_k\varphi_k] = \sum_{k=0}^{n-1} \mathbb{E}[\psi_k]\mathbb{P}\left[L(0, n-1-k) \leq \beta\ell^*(n)\right] \\ &= \sum_{k=0}^{n-1} \frac{1}{\ell^*(k)}\mathbb{P}\left[\sum_{i=1}^{\beta\ell^*(n)} T_i \geq n-1-k\right].\end{aligned} \tag{A.8}$$

If $k \leq (1-\delta)n$ for some $\delta > 0$, then we can apply (A.7). Bounding the probability by one in the remaining cases, we see that (A.8) is bounded from above by

$$\begin{aligned}\mathbb{E}[|R_\beta(n)|] &\leq \sum_{k=0}^{(1-\delta)n} \frac{1}{\ell^*(k)}(1 - e^{-\beta})(1 + o_\delta(1)) + \sum_{(1-\delta)n < k < n} \frac{1}{\ell^*(k)} \\ &= \frac{(1-\delta)n}{\ell^*(n)}(1 - e^{-\beta})(1 + o_\delta(1)) + \frac{\delta n}{\ell^*(n)},\end{aligned}$$

and from below by

$$\mathbb{E}[|R_\beta(n)|] \geq \sum_{k=0}^{(1-\delta)n} \frac{1}{\ell^*(k)}(1 - e^{-\beta})(1 + o_\delta(1)) = \frac{(1-\delta)n}{\ell^*(n)}(1 - e^{-\beta})(1 + o_\delta(1)).$$

Sending $\delta \to 0$ proves the statement for $\mathbb{E}[|R_\beta(n)|]$.

To bound the variance, we first note that for $i < i + \delta n \leq j \leq (1-\delta)n$, by the Markov property and using Assumption A and (A.7),

$$\begin{aligned}\mathbb{E}\left[\psi_i\varphi_i\psi_j\varphi_j\right] &\leq \mathbb{E}\left[\psi_i \mathbf{1}_{\{L(Y(i), i+\delta n) \leq \beta\ell^*(n)\}} \mathbf{1}_{\{Y(k) \neq Y(j),\ i+\delta n \leq k < j\}} \mathbf{1}_{\{L(Y(j), n) \leq \beta\ell^*(n)\}}\right] \\ &= \mathbb{E}[\psi_i]\mathbb{P}\big[L(0, \delta n) \leq \beta\ell^*(n)\big]\mathbb{E}\big[\psi_{j-i-\delta n}\big]\mathbb{P}\big[L(0, \delta n) \leq \beta\ell^*(n)\big] \\ &= \frac{1}{\ell^*(i)}(1 - e^{-\beta})\frac{1}{\ell^*(j-i-\delta n)}(1 - e^{-\beta})(1 + o_\delta(1)).\end{aligned} \tag{A.9}$$

The variance of $|R_\beta(n)|$ is

$$\operatorname{Var}|R_\beta(n)| = 2\sum_{0 \leq i \leq j \leq n-1} \mathbb{E}[\psi_i\varphi_i\psi_j\varphi_j] - \mathbb{E}[\psi_i\varphi_i]\mathbb{E}[\psi_j\varphi_j]. \tag{A.10}$$

For $i < i + \delta n \le j \le (1-\delta)n$ we can use (A.9) and (A.8) to get

$$\begin{aligned}
&\sum_{i<i+\delta n\le j\le(1-\delta)n} \mathbb{E}[\psi_i\xi_i\psi_j\xi_j] - \mathbb{E}[\psi_i\xi_i]\mathbb{E}[\psi_j\xi_j] \\
&\le (1-e^{-\beta})^2 \sum_{i<i+\delta n\le j\le(1-\delta)n} \frac{1}{\ell^*(i)}\left(\frac{1}{\ell^*(j-i-\delta n)}(1+o_\delta(1)) - \frac{1}{\ell^*(j)}(1+o_\delta(1))\right) \\
&= (1-e^{-\beta})^2 \sum_{i=0}^{(1-2\delta)n} \frac{1}{\ell^*(i)}\left(\sum_{j=0}^{(1-2\delta)n-i}\frac{1}{\ell^*(j)}(1+o_\delta(1)) - \sum_{j=i+\delta n}^{(1-\delta)n}\frac{1}{\ell^*(j)}(1+o_\delta(1))\right) \qquad \text{(A.11)} \\
&= o_\delta\left(\left(\frac{n}{\ell^*(n)}\right)^2\right).
\end{aligned}$$

For the remaining i, j, using (A.3) we have

$$\begin{aligned}
\sum_{i=0}^{n-1} \sum_{\substack{i\le j<i+\delta n\\(1-\delta)n<j<n}} &\mathbb{E}[\psi_i\xi_i\psi_j\xi_j] - \mathbb{E}[\psi_i\xi_i]\mathbb{E}[\psi_j\xi_j] \\
&\le \sum_{i=0}^{n-1} \sum_{\substack{i\le j<i+\delta n\\(1-\delta)n<j<n}} \mathbb{E}[\psi_i\psi_j] \\
&\le \sum_{i=0}^{n-1} \sum_{\substack{i\le j<i+\delta n\\(1-\delta)n<j<n}} \mathbb{E}[\psi_i]\mathbb{E}[\psi_{j-i}] \qquad \text{(A.12)} \\
&\le \sum_{i=0}^{n-1} \frac{1}{\ell^*(i)}\left(\sum_{j=i}^{i+\delta n-1}\frac{1}{\ell^*(j)} + \sum_{j=(1-\delta)n+1}^{n-1}\frac{1}{\ell^*(j)}\right) \\
&\le 2\frac{\delta n^2}{(\ell^*(n))^2}(1+o_\delta(1)).
\end{aligned}$$

Inserting (A.11), (A.12) into (A.10) and taking $\delta \to 0$ yields $\operatorname{Var}|R_\beta(n)| = o((\mathbb{E}|R_\beta(n)|)^2)$ and the weak law of large numbers follows.

Finally, part (iii) is proved in the same way as part (ii). The only difference is that instead of using (A.7) we note that $L(0,\infty)$ is a geometric random variable with parameter γ, therefore for every $c > 0$,

$$\mathbb{P}\left[L(0, cn) = k\right] = \gamma(1-\gamma)^{k-1}(1+o(1)) \text{ as } n \to \infty.$$

This completes the proof. □

References

[BČ07] G. Ben Arous and J. Černý, *Scaling limit for trap models on $\mathbb{Z}^d$*, Ann. Probab. **35** (2007), no. 6, 2356–2384. MR 2353391

[BČ11] M. T. Barlow and J. Černý, *Convergence to fractional kinetics for random walks associated with unbounded conductances*, Probab. Theory Related Fields **149** (2011), no. 3-4, 639–673. MR 2776627

[BCČR14] G. Ben Arous, M. Cabezas, J. Černý, and R. Royfman, *Randomly Trapped Random Walks*, To appear in Ann. Probab. (2014).

[BČM06] G. Ben Arous, J. Černý, and T. Mountford, *Aging in two-dimensional Bouchaud's model*, Probab. Theory Related Fields **134** (2006), no. 1, 1–43. MR 2221784

[BS02] E. Bolthausen and A.-S. Sznitman, *On the static and dynamic points of view for certain random walks in random environment*, Methods Appl. Anal. **9** (2002), no. 3, 345–375, Special issue dedicated to Daniel W. Stroock and Srinivasa S. R. Varadhan on the occasion of their 60th birthday. MR 2023130

[Čer07] J. Černý, *Moments and distribution of the local time of a two-dimensional random walk*, Stochastic Process. Appl. **117** (2007), no. 2, 262–270. MR 2290196

[Dar52] D. A. Darling, *The influence of the maximum term in the addition of independent random variables*, Trans. Amer. Math. Soc. **73** (1952), 95–107. MR 0048726

[DE51] A. Dvoretzky and P. Erdős, *Some problems on random walk in space*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, University of California Press, Berkeley and Los Angeles, 1951, pp. 353–367. MR 0047272

[ET60] P. Erdős and S. J. Taylor, *Some problems concerning the structure of random walk paths*, Acta Math. Acad. Sci. Hungar. **11** (1960), 137–162. (unbound insert). MR 0121870

[FM13] L. R. G. Fontes and P. Mathieu, *On the dynamics of trap models in $\mathbb{Z}^d$*, Proceedings of the London Mathematical Society (2013).

[GS13] V. Gayrard and A. Svejda, *Convergence of clock processes on infinite graphs and aging in Bouchaud's asymmetric trap model on $\mathbb{Z}^d$*, arXiv:math/1309.3066 (2013).

[JS03] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR 1943877

[Kal02] O. Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169

[KS63] H. Kesten and F. Spitzer, *Ratio theorems for random walks. I*, J. Analyse Math. **11** (1963), 285–322. MR 0162279

[Mou11] J.-C. Mourrat, *Scaling limit of the random walk among random traps on $\mathbb{Z}^d$*, Ann. Inst. Henri Poincaré Probab. Stat. **47** (2011), no. 3, 813–849. MR 2841076

[MS04] M. M. Meerschaert and H.-P. Scheffler, *Limit theorems for continuous-time random walks with infinite mean waiting times*, J. Appl. Probab. **41** (2004), no. 3, 623–638. MR 2074812

[MW65] E. W. Montroll and G. H. Weiss, *Random walks on lattices. II*, J. Mathematical Phys. **6** (1965), 167–181. MR 0172344

[Pet83] K. Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics, vol. 2, Cambridge University Press, Cambridge, 1983. MR 833286

[Pit74] J. H. Pitt, *Multiple points of transient random walks*, Proc. Amer. Math. Soc. **43** (1974), 195–199. MR 0386021

[Rév13] P. Révész, *Random walk in random and non-random environments*, third ed., World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2013. MR 3060348

[Sen76] E. Seneta, *Regularly varying functions*, Lecture Notes in Mathematics, no. Nr. 508, Springer-Verlag, 1976.

[Spi76] Frank Spitzer, *Principles of random walk*, second ed., Springer-Verlag, New York-Heidelberg, 1976, Graduate Texts in Mathematics, Vol. 34. MR 0388547

[Whi02] W. Whitt, *Stochastic-process limits*, Springer Series in Operations Research, Springer-Verlag, New York, 2002, An introduction to stochastic-process limits and their application to queues. MR 1876437

University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Wien, Austria
E-mail address: `tobias.wassmer@univie.ac.at`