On the dynamics of trap models in $\mathbb{Z}^d$

L. R. G. Fontes and P. Mathieu

Abstract
We consider trap models in $\mathbb{Z}^d$. These are stochastic processes in a random environment as follows. The environment is given by a family $\tau = (\tau_x, x \in \mathbb{Z}^d)$ of positive independent and identically distributed random variables in the basin of attraction of an $\alpha$-stable law, $0 < \alpha < 1$. Given $\tau$, our process is a continuous time Markov pure jump process, whose jump chain is a, in principle, generic random walk in $\mathbb{Z}^d$, $d \geq 1$, independent of $\tau$, and $\tau$ represents the holding time averages of the continuous time process. We may think of the sites of $\mathbb{Z}^d$ as traps, and of $\tau_x$ as the depth of trap $x$. We are interested in the trap process, namely the process that associates to time $t$ the depth of the currently visited trap. Our first result is the convergence of the law of that process under suitable scaling. The limit process is given by the jumps of a certain $\alpha$-stable subordinator at the inverse of another $\alpha$-stable subordinator, correlated with the first subordinator. For that result, the requirements on the underlying random walk are (a) the validity of a law of large numbers for its range and (b) the slow variation at infinity of the tail of the distribution of its time of return to the origin: they include all transient random walks as well as all random walks in $d \geq 2$, and also many one-dimensional random walks, but not the simple symmetric case. We then derive ageing results for our process, namely scaling limits for some two-time correlation functions of the process; a strong form of those results requires an assumption of transience, stronger than $a, b$ above. The scaling limit result mentioned above is an averaged result with respect to the environment. Under an additional condition on the size of the intersection of the ranges of two independent copies of the underlying random walk, roughly saying that it is small compared with size of the range, we derive a stronger scaling limit result, roughly stating that it holds in probability with respect to the environment. With that additional condition, we also strengthen the ageing results, from the averaged version mentioned above, to convergence in probability with respect to the environment.

1. Introduction
We begin with a precise definition of random walks among random traps. These are constructed through the following two-step procedure. We first choose a probability measure on $\mathbb{Z}^d \setminus \{0\}$, say $\pi$, and let $\tau = (\tau_x, x \in \mathbb{Z}^d)$ be a collection of positive real numbers attached to the points of $\mathbb{Z}^d$. The random walk in the trap environment $\tau$ is then the continuous time Markov process with values in $\mathbb{Z}^d$ that starts at the origin and has generator

$$
\mathcal{L}^\tau f(x) = \frac{1}{\tau_x} \sum_y (f(y) - f(x))\pi(y - x),
$$

(1.1)

so that when sitting at point $x \in \mathbb{Z}^d$, the process waits for an exponentially distributed time of mean $\tau_x$ and then jumps to point $x + y$, where $y$ is sampled from the distribution $\pi$. This procedure is then iterated with independent hopping times and jumps. We will take $\tau$ random so that the $(\tau_x, x \in \mathbb{Z}^d)$ is an independent identically distributed family of random variables whose common law belongs to the domain of attraction of a stable law of index $\alpha \in (0, 1)$. The model thus defined is therefore an example of a random walk in a random environment. We
denote by $\mathbb{P}$ the probability thus defined. More precisely, $\mathbb{P}$ is a probability measure on the product space $\Omega \times \mathcal{D}([0, \infty), \mathbb{Z}^d)$, where $\Omega = (0, \infty)^{\mathbb{Z}^d}$ is the space of trap environments and $\mathcal{D}([0, \infty), \mathbb{Z}^d)$ is the space of càdlàg trajectories from $[0, \infty)$ to $\mathbb{Z}^d$. The first marginal of $\mathbb{P}$ is of the form $Q^{\mathbb{Z}^d}$, where $Q$ belongs to the domain of attraction of a stable law of index $\alpha \in (0, 1)$, and the conditional law of the second marginal given $\tau$ is the law of the random walk in the trap environment $\tau$. We use the notation $(X_t, t \geq 0)$ for the canonical projections defined on $\mathcal{D}([0, \infty), \mathbb{Z}^d)$ that give the position of the random walker at times $(t \geq 0)$.

**Remark 1.** It is not difficult to see that if we choose $Q$ with compact support in $(0, +\infty)$, then the behaviour of the random walk with traps is very similar to the random walk without traps. For instance, if $\pi$ is symmetric with finite support, then one finds that the scaling limit of $\mathcal{X}$ under diffusive scaling is a Brownian motion. Fluctuations of the environment only affect the value of the effective diffusivity. In order to observe stronger slowing down effects, in particular ageing, one has to choose heavy tailed $\tau$ as we do here.

This process is an example of a trap model in the spirit of J.-P. Bouchaud. One important aspect of it is the lack of dependence of $\pi$ on $\tau$. (A class of models where there is such a dependence, known as asymmetric trap models, have also been considered in the physics and mathematics literature. See below. Unless explicitly mentioned, we do not discuss these models here.) Such processes were initially introduced in the context of statistical mechanics as toy models for spin glasses and in order to illustrate the phenomenon of ageing; see [12, 14] or [13] for instance. In usual models of spin glasses, the Hamiltonian is a random Gaussian field of large variance. At low temperature, it is natural to guess that the main contributions to the dynamics come from states of low energy. As the statistical properties of extremes of log Gaussian fields, the Gibbs factors in this context, are described by random variables with polynomial tail, the choice of a law in the basin of attraction of a stable law for $\tau$, which plays a similar role in the simplified model, is also natural. Note that the parameter $\alpha$ can then be interpreted as the temperature; see [2, 19, Section 3.2].

The ageing property refers to the following phenomenon: as time increases, the process visits a larger and larger part of its state space and therefore increases its probability to find a location $x$ where $\tau_x$ is large. Since the time the process stays at location $x$ before jumping off is of order $\tau_x$, some slow down effect might take place. One way to measure how much the process is slow is to compute quantities of the form

$$\Pi(s, t) = \mathbb{P}(X_r = X_t, r \in [t, t+s]),$$

which are generally called ageing functions in this context. The Markov property implies that

$$\Pi(s, t) = \mathbb{E}(e^{-s/\tau_{X_t}}),$$

and thus we observe that a nontrivial limit for $\Pi(s, t)$ as $s, t \to \infty$, with $s$ and $t$ related in a given way, implies that, at large time $t$, $\tau_{X_t}$ should be of order $s$, so that the (order of the) ‘age’ of the process can actually be approximately read from its position at large times. Thus, in order to describe ageing, we are led to considering the asymptotics of the age (or trap) process $A = (A_t = \tau_{X_t}, t \geq 0)$.

The first computations of Bouchaud and Dean in [12, 14] consisted in describing the asymptotics of trapped random walks on a large complete graph and in some appropriate scaling. Since then, the subject has developed into a rich mathematical theory. Mathematical papers treating the model in the complete graph include [15, 20]. Although one motivation is certainly to understand the physicists’ claims and prove ageing for as realistic as possible models of spin glasses, see [4, 5] and more recently [3], it also turns out that trapping and ageing effects also play a role in models without any connection to spin glass theory such as
random walks with random conductances or random walks on Galton–Watson trees; see \([1, 10]\).

The main strategy used in these papers has a strong potential theoretic flavour: for a given realization of the trap environment \(\tau\), one tries to identify, among the different points, \(x\) with large \(\tau_x\) which will be hit by the random walk. We refer the reader to [8] for a presentation of this point of view in an abstract setting. One advantage of this approach is that it does not seem to require the state space to have many symmetries. It provides strong forms of ageing properties that are valid for a given realization of the traps. On the other hand, this machinery is often quite heavy to use.

As far as trap models on \(\mathbb{Z}^d\) are concerned, excluding the asymmetric case (where \(\pi\) depends on \(\tau\) in a specific way, as mentioned above; see [1, 6, 17, 25]), only the case of the simple symmetric random walk was investigated so far. It corresponds to \(\pi\) being the uniform law on the nearest neighbours of the origin. Then the paths of the process \(\mathcal{X}\), that is, the sequence of the different points visited by \(\mathcal{X}\) moves, that is, the different hopping times at the successive locations are given by the environment \(\tau\). The one-dimensional case happens to be special: due to the strong recurrence properties of the simple symmetric random walk on \(\mathbb{Z}\), the process localizes. This localization effect, ageing properties and scaling limits are precisely described in [18]. The scaling limit is a singular diffusion now known under the name of FIN.

In higher dimension \(d \geq 2\), the scaling limit is known to be the so-called Fractional Kinetics process: \(d\)-dimensional Brownian motion time changed by the inverse of an independent stable subordinator as proved in [7, 9, 16], the strategy being similar to [8]. Besides a number of estimates on the Green kernel of simple symmetric random walk in \(\mathbb{Z}^d\), the proof in the \(d = 2\) case involves rather sophisticated renormalization technics. The same result also follows in \(d \geq 5\) as a particular case of results in [25], where a different approach is developed.

What do we do here? A first motivation of this paper is to derive ageing properties for a more general class of random walks than the nearest neighbours case, in the form of an appropriate scaling limit of the age process, as suggested by our discussion above; see Theorem 5. In doing so we hope to clarify which properties of the random walk are truly relevant for ageing. Observe, in particular, that the usual recurrence versus transience dichotomy does not apply here (as we can already conclude from the results of [7] for the simple symmetric case). As an outcome, we obtain a new proof of the scaling limit that applies to any genuinely \(d\)-dimensional random walk for \(d \geq 2\). This proof is more conceptual than the approach previously used by other authors. Indeed we need to know very little about specific estimates for the transition probabilities or Green kernels. We also completely avoid the renormalization step, even in the \(d = 2\) case. It should also be mentioned that Theorem 5 is an annealed result with respect to the environment, as opposed to the analogue quenched result of [7]. In particular, it applies in situations where \(\mathcal{X}\) does not have a quenched scaling limit, see Remark 36. It can, however, be strengthened with little more effort, say, half way towards a quenched result, under a natural additional condition on the intersections of the ranges of independent copies of our process (see Theorem 35).

Observe that in our general setting it does not make sense to look for scaling limits of the process \(\mathcal{X}\) itself. Indeed the underlying random walk (with increments distributed according to \(\pi\)) may not have a non trivial scaling limit. We choose then to focus on the age process \(\mathcal{A}\), which is a natural object in the ageing context, as we discuss next. The expression of the ageing function \(\Pi(s, t)\), suggests that we should look at the limiting law of \(\mathcal{A}_t/s\) to derive an ageing property. We actually provide a more complete answer by describing the scaling limit of the full process \(\mathcal{A}\), thus establishing a fuller ageing picture. This scaling limit is expressed as the value of the jump of some subordinator computed at the inverse of another subordinator. Interestingly, this scaling limit is universal, even if the scale on which the process \(\mathcal{A}\) lives depends on the random walk (and in particular is linear if and only if the random walk is transient).
The topology under which we are able to establish Theorem 5 is quite weak, though, due to the nature of the age process. Obtaining a scaling limit result for ageing functions like (1.2) requires more work, done in Section 5 (for (1.2) and two other examples) under the stronger assumption of transience, and in Subsection 6.1 with the additional condition of Section 6; see Theorems 25 and 39. Integrated forms of those results follow from Theorem 5, under the original assumptions, as discussed separately at the end of those sections.

The remainder of this paper is organized as follows. In Section 2, we have a detailed presentation of the model, assumptions and one result (annealed scaling limit of $A$), with some more discussion. In Section 3, we discuss some preliminary results on random walks that are used subsequently. In Section 4, we prove the annealed result just mentioned, and in Sections 5 and 6, we state and prove our further scaling limit results for some ageing functions, and in a stronger than annealed sense, as discussed above.

2. Model and results

As in the introduction, let $\pi$ be a probability measure on $\mathbb{Z}^d \setminus \{0\}$, and let $\tau = (\tau_x, \ x \in \mathbb{Z}^d)$ be a collection of positive numbers attached to the points of $\mathbb{Z}^d$ chosen as follows. Let $Q$ be a probability measure on $(0, \infty)$ that belongs to the domain of attraction of a stable law of index $\alpha \in (0, 1)$. In other words, we assume that

$$Q(u, \infty) = \ell(u)u^{-\alpha}, \quad u > 0,$$

(2.1)

where $\ell$ is a slowly varying function at infinity. We choose for $\tau = \{\tau_x, \ x \in \mathbb{Z}^d\}$

(2.2)

a family of independent random variables with law $Q$. More precisely, we endow the product space $\Omega = (0, \infty)^{\mathbb{Z}^d}$ with the law $Q = Q^{\mathbb{Z}^d}$.

We consider the Markov generator

$$\mathcal{L}^\tau f(x) = \frac{1}{\tau_x} \sum_y (f(y) - f(x))\pi(y - x).$$

(2.3)

Let $P^\tau_x$ be the law of the Markov process $X$ generated by $\mathcal{L}^\tau$ and started at $x$ on path space $\mathcal{D}([0, \infty), \mathbb{Z}^d)$. We recall that $(X_t, \ t \geq 0)$ denotes the canonical projections on $\mathcal{D}([0, \infty), \mathbb{Z}^d)$.

We define the age process (as in the previous section):

$$A = (A_t = \tau X_t, \ t \geq 0).$$

(2.4)

The so-called annealed law of the process $X$ is the semidirect product measure on $\Omega \times \mathcal{D}([0, \infty), \mathbb{Z}^d)$ defined by

$$\mathbb{P}(A \times B) = \int_A dQ(\tau)P^\tau_0(B),$$

(2.5)

where $A$ and $B$ are measurable subsets of $\Omega$ and $\mathcal{D}([0, \infty), \mathbb{Z}^d)$, respectively.

In order to state our assumptions we introduce an auxiliary random walk: let $\xi_1, \xi_2, \ldots$ be independent and identically distributed (i.i.d.) $\mathbb{Z}^d$-valued random vectors with distribution $\pi$ and define

$$X_0 = 0, \quad X_n = \sum_{i=1}^n \xi_i, \quad n \geq 1.$$

(2.6)

$X = (X_n, \ n \geq 0)$ is a version of the jump chain of $X$.

Let us also define the range of $X$ (up to time $n$),

$$\mathcal{R}_n = \mathcal{R}_n(X) = \{z \in \mathbb{Z}^d: X_i = z \text{ for some } i \leq n\}, \quad n \geq 0,$$

(2.7)
and make
\[ R_n = |R_n| \quad \text{and} \quad \rho_n = \mathbb{E}(R_n). \] (2.8)

Also define for \( n \geq 1 \)
\[ r_n = \mathbb{P}(X_1 \neq 0, \ldots, X_n \neq 0). \] (2.9)

We will at times think of \((r_n)\) as a function.

Our first result requires the following assumptions.

**Assumption A** (Law of large numbers for the range).
\[ \lim_{n \to \infty} \frac{R_n}{\rho_n} = 1 \quad \text{in probability}. \] (2.10)

**Assumption B** (Slow variation of \( r \)).
\[ r : \mathbb{N} \to [0, 1] \text{ given in } (2.9) \text{ above is slowly varying at infinity}. \]

**Remark 2.** All transient random walks in \( \mathbb{Z}^d, d \geq 1 \), including all random walks in \( d \geq 3 \), obviously satisfy Assumptions A and B. But all planar random walks \([21, 22]\), and one-dimensional \( \beta \)-stable random walks with \( \beta \leq 1 \) \([24]\) also satisfy Assumptions A and B.

Before stating our first result, we describe the form of the scaling limit of \( \mathcal{A} \): we introduce an \( \alpha \)-stable subordinator \( \Upsilon = (\Upsilon_r)_{r \geq 0} \) and a family of independent mean 1 exponential random variables \( \{T_r; r \geq 0\} \), and let
\[ V_s = \int_0^s T_r d\Upsilon_r, \quad s \geq 0, \] (2.11)
and \( W = V^{-1} \) be the inverse of \( V \) (see Remark 11). Let finally
\[ Z_t = \Upsilon_{W_t} - \Upsilon_{W_{t-}}, \quad t \geq 0. \] (2.12)

**Remark 3.** Note that \( V \) is itself an \( \alpha \)-stable subordinator. (This will be relevant in our discussion on ageing in Section 5; see Remark 31) We may regard \( V \) and \( Z \) as processes in the random environment \( \Upsilon \). In fact, given \( \Upsilon \), they are both Markovian. We may think of \( \Upsilon \) as the scaling limit of the (relevant) environment of the trap model. The overall distribution of \( Z \) (integrated over the joint distribution of \( \Upsilon \) and \( \{T_r; r \geq 0\} \)) makes it a self-similar process of index 1.

**Remark 4.** Let \( c > 0 \) be the constant such that \( \mathbb{E}(\exp\{-\lambda \Upsilon_1\}) = \exp\{-c\lambda^\alpha\} \). One readily checks that the distribution of \( Z \) does not depend on that constant. Here and below we will denote also by \( \mathbb{P} \) and \( \mathbb{E} \) the probability and expectation underlying the distributions of \( \Upsilon \) and \( T_r \).

We are now ready to state our convergence result, but first let us introduce a few notations. For \( \varepsilon > 0, \ t \geq 0, \) let
\[ \mathcal{A}_t^{(\varepsilon)} = \varepsilon q_\varepsilon \mathcal{A}_{e^{-\varepsilon}t}, \] (2.13)
where \( q_\varepsilon \) is a slowly varying function at 0 to be further specified below, and denote \( \mathcal{A}^{(\varepsilon)} = (\mathcal{A}_t^{(\varepsilon)}) \) and \( Z = (Z_t) \). Let \( D, D_T \) denote the spaces of càdlàg real functions defined on \([0, \infty), [0, T]\), respectively. Let \( d_T \) denote the \( L_1 \) distance in \( D_T \), and \( d = \sum_{n=1}^\infty 2^{-n}(d_n \wedge 1) \).
ON THE DYNAMICS OF TRAP MODELS IN $\mathbb{Z}^d$  1567

Theorem 5. Suppose Assumptions A and B are in force. Then as $\varepsilon \searrow 0$

$$A^{(\varepsilon)} \rightarrow Z$$

(2.14)

in distribution on $(D,d)$, where $(q_{\varepsilon})_{\varepsilon>0}$ is nonincreasing, slowly varying at 0 and satisfies $1 \leq q := \lim_{\varepsilon \to 0} q_{\varepsilon} \leq \infty$, with $q < \infty$ if and only if $X$ is transient.

Note that in the transient case we have linear scaling, and sublinear scaling in the recurrent case. A more precise description of $q_{\varepsilon}$ is given in (2.17).

Here is a rough evaluation on the form of the spatial scaling in (2.13) leading to (2.14). After $n$ steps, the random walk $X$ has explored $R_n \sim \rho_n$ new sites, each one visited of the order of $1/r_n$ times, so the walk should be visiting a trap of depth of order $s_{\rho_n} =: \nu_n$, where $s_n$ is of the order of the largest of $n$ independent copies of $\tau_0$. Under our assumptions, the main contributions to the total time spent by the process to give $n$ steps, namely the times spent at the deepest traps encountered up to that number of steps, are roughly independent, so that the total time should be of order $v_n/r_n =: \phi_n$. Then to times of order $\varepsilon^{-1}$ correspond energies of order $\nu_{\nu_n}$, where

$$\nu_{\varepsilon} = \phi^{-1}(\varepsilon^{-1}),$$

(2.15)

and $\phi^{-1}$ is the inverse of $\phi$. We thus should take

$$\frac{1}{v_{\nu_n}} = \frac{1}{r_{\nu_n} \phi_{\nu_n}} \sim \frac{\varepsilon}{r_{\nu_n}}$$

(2.16)

as spatial scaling, where the latter approximate equality follows from the inverse relation between $\phi$ and $\nu$. From the assumptions on $r$ and the distribution of $\tau_0$, it follows that $\nu$ is nonincreasing, unbounded and regularly varying at infinity. Since $r \in [0,1)$ is nonincreasing and slowly varying at infinity, we conclude that

$$q_{\varepsilon} = 1/r_{\nu_n}$$

(2.17)

has the stated properties.

The rough discussion of the last paragraph also gives rise to the form of the limiting process, once one realizes that (under our assumptions) the total times spent on the few contributing traps, when divided by the respective means, and suitably scaled to account for the number of distinct visits, are approximately i.i.d. mean 1 exponentials. A striking aspect is the universality of the limiting process even for cases where the scalings are distinct (linear and sublinear as pointed out above).

A remark about the topology: small traps do not contribute to the limit and were they to be completely disregarded the convergence would take place in the usual Skorohod topology, but they are there and mix up with large traps in a way which the $J_1$ topology (and other more usual ones, like the $M_1$ topology) is too fine to handle, and so we resort to a rougher topology.

One important technical aspect to consider in order to prove the scaling limit of an ageing function like (1.3) is to show that for an arbitrary fixed time $t$, $A_t^{(\varepsilon)}$ converges to $Z_t$ as $\varepsilon \to 0$ in a strong enough form. This does not follow from Theorem 5, mainly due to the topological issue just discussed. We state and prove such a convergence result, Lemma 29 involving suitable versions of the relevant processes, before establishing the main results of that section, namely the ageing results stated in Theorem 25 (see also Remark 32). Our approach requires the strengthening of Assumptions A and B to transience of $X$. Scaling limit results for integrated versions of the ageing functions herein considered, and potentially others, follow from Theorem 5, or rather the arguments in its proof, directly, under the original assumptions; see Subsection 5.1.
We close with a brief discussion on the stronger than annealed version for Theorem 5. Roughly, the annealed aspect of Theorem 5 follows from our approach of considering a version of the $\tau$ variables placed over the range of the underlying discrete random walk, thus fixing it, in such a way that they in a certain sense converge almost surely (to the increments of a stable subordinator, $\Upsilon$); see details in the proof of Theorem 5. This turns out to be convenient for the analysis leading to Theorem 5, but when going back to the original $\tau$, we have only an annealed result. But of course, when we consider the distribution of the process given $\tau$, we integrate with respect to the underlying discrete random walk, and the averaging involved in this could lead to a stronger result. A condition is required, however, annealed convergence is all we have in, say, the asymmetric simple one-dimensional case. One is introduced in Section 6, saying roughly that independent realizations of the trajectory of $X$ intersect little. With that additional condition, we state and prove stronger convergence results.

3. Preliminaries on random walks

In this section, we establish a few facts concerning discrete time random walks that follow from Assumptions A and B made in Section 2, as well as from other assumptions we will consider below. These results will be used later in the sections ahead.

Let $X = (X_n, n \geq 0)$ be the random walk introduced in Section 2, and define

$$ u_n = \mathbb{P}(X_n = 0), \quad U_n = \sum_{i=0}^{n} u_i, \quad L_n = \sum_{i=0}^{n} 1\{X_i = 0\}, \quad n \geq 0. $$

$L_n$ is the occupation time of the origin up to step $n$. We will also write $L_{x}$ for a positive real $x$, and it means $L_{\lfloor x \rfloor}$, similarly for $U_x$ and $r_x$.

Our first remark is that

$$ \rho_n = \sum_{k=0}^{n} r_k, \quad (3.2) $$

with $\rho$ and $r$ defined in (2.8, 2.9). The formula is proved as follows (see [27, p. 36]). First, we have that

$$ R_n = \sum_{k=0}^{n} 1\{R_k = R_{k-1} + 1\}, $$

where $R$ was defined in (2.7). Upon noting that

$$ R_k = R_{k-1} + 1 \Leftrightarrow X_k - X_{k-1} \neq 0; \quad X_k - X_{k-2} \neq 0, \ldots, X_k \neq 0, $$

we conclude that

$$ \mathbb{P}(R_k = R_{k-1} + 1) = \mathbb{P}(\xi_k \neq 0; \xi_k + \xi_{k-1} \neq 0, \ldots, X_k \neq 0) $$

$$ = \mathbb{P}(X_1 \neq 0; X_2 \neq 0, \ldots, X_k \neq 0) = r_k. $$

It then follows from (3.2) and Assumption B that

$$ \lim_{n \to \infty} \rho_n = 1. \quad (3.3) $$

(This readily follows from the first displayed equation on [26, p. 55].)

Our second remark is the following result.

**Lemma 6.** Under Assumption B, and provided $\lim_{n \to \infty} r_n = 0$ (that is, if $X$ is recurrent), the law of $r_n L_n$ approximates a mean 1 exponential distribution as $n \to \infty$. 
ON THE DYNAMICS OF TRAP MODELS IN \( \mathbb{Z}^d \)

**Proof.** Given \( u \geq 0 \), we have that

\[
\{ r_n L_n > u \} = \{ \eta_1 + \cdots + \eta_k \leq n \},
\]

where the \( \eta_j, j \geq 1 \), are the successive increments of return times to the origin by \( X \), and \( k = \lceil u/r_n \rceil + 1 \).

Therefore,

\[
\mathbb{P}(r_n L_n > u) = \mathbb{P}(\eta_1 + \cdots + \eta_k \leq n) = \mathbb{P}(\bar{\eta}_n \leq 1),
\]

(3.4)

where \( \bar{\eta}_n = (\eta_1 + \cdots + \eta_k)/n \).

A straightforward computation of the Laplace transform of \( \bar{\eta}_n \) yields

\[
\mathbb{E}(e^{-\lambda \bar{\eta}_n}) = \{ e^{-(\lambda/n)\eta_1} \}^k = \left\{ 1 - \frac{\lambda}{n} \int_0^\infty r(x) e^{-\lambda x} dx \right\}^k = \left\{ 1 - \int_0^\infty r(yn/\lambda) e^{-y} dy \right\}^k,
\]

(3.5)

where \( r(x) = \mathbb{P}(\eta_1 > x) = r_{[x]} \).

Assumption B and Seneta [26, Theorems 2.6 and 2.7] imply that the integral on the right-hand side of (3.5) is asymptotic to \( r_n \) as \( n \to \infty \). This and the form of \( k \) imply that

\[
\lim_{n \to \infty} \mathbb{E}(e^{-\lambda \bar{\eta}_n}) = e^{-u}
\]

(3.6)

for all \( \lambda > 0 \), and this implies that the law of \( \bar{\eta}_n \) converges to that of a (n extended) random variable which takes the value 0 with probability \( e^{-u} \), and the value \( \infty \) with the complementary probability. It follows that

\[
\lim_{n \to \infty} \mathbb{P}(r_n L_n > u) = e^{-u}
\]

(3.7)

for every \( u > 0 \). \( \square \)

**Corollary 7.** Under Assumption B, we have that

\[
\lim_{n \to \infty} r_n U_n = 1.
\]

(3.8)

**Proof.** In the transient case, this follows from \( r_n \to r_\infty > 0, L_n \to L_\infty, \) a Geometric random variable with mean \( r_\infty^{-1} \), and monotone convergence.

In the recurrent case, from the first equality in (3.4), we find that

\[
\mathbb{P}(r_n L_n > u) \leq \mathbb{P} \left( \max_{1 \leq j \leq k} \eta_j \leq n \right) = (1 - r_n)^k.
\]

(3.9)

From the form of \( k \), and the fact that \( \lim_{n \to \infty} (1 - r_n)^{1/r_n} = e^{-1} \), we find that for every \( c > e^{-1} \) and large enough \( n \), the right-hand side of (3.8) is dominated by \( c^n \) for all large enough \( n \).

Dominated convergence now yields

\[
r_n U_n = \int_0^\infty \mathbb{P}(r_n L_n > u) du \to \int_0^\infty e^{-u} du = 1.
\]

(3.10)

as \( n \to \infty \), since, from (3.1), \( r_n U_n = r_n \mathbb{E}(L_n) = \mathbb{E}(r_n L_n) \). \( \square \)

Another corollary of Lemma 6 is as follows. For \( x \in \mathbb{Z}^d \), let

\[
\ell(x, n) = \sum_{j=0}^n 1\{X_j = x\} T_j,
\]

(3.11)

where \( T_1, T_2, \ldots \) are i.i.d. mean 1 exponential random variables independent of \( X \).
Corollary 8. Under Assumption B, we have that $r_n \ell(0,n)$ converges weakly to a mean 1 exponential distribution.

Proof. In the recurrent case, the result follows immediately from Lemma 6 and the law of large numbers, once we observe that

$$\ell(0,n) = \sum_{i=1}^{L_n} T'_i,$$

where $T'_1, T'_2, \ldots$ are i.i.d. mean 1 exponential random variables independent of $X$.

In the transient case, $L_n$ converges as $n \to \infty$ to a geometrically distributed random variable, say $L_\infty$. Also $\lim_{n \to \infty} r_n = r_\infty > 0$, and one readily checks that $r_\infty \sum_{i=1}^{L_\infty} T'_i$ is a mean 1 exponential random variable.

Lemma 9. Under Assumption B, we have that for every $0 < a < b < \infty$

$$L_{bn} - L_{an} \to 0$$

in probability as $n \to \infty$.

Remark 10. Since $L_{bn} - L_{an}$ is an integer, we have that the probability of no return to 0 of $X$ during $[an, bn]$ goes to 1 as $n \to \infty$ for every fixed $0 < a < b < \infty$.

Proof. Let $N = N_n(a, b)$ denote the random variable on the left of (3.13). Using the Markov property, one readily checks that the conditional distribution of $L_{(b+1)n} - L_{an}$ given $N \geq 1$ dominates the unconditional one of $L_n$. We conclude that

$$U_{(b+1)n} - U_{an} = \mathbb{E}[L_{(b+1)n} - L_{an}] \geq \mathbb{E}[L_{(b+1)n} - L_{an}; N \geq 1]$$

$$\geq \mathbb{E}[L_n] \mathbb{P}(N \geq 1) = U_n \mathbb{P}(N \geq 1).$$

We then have that $\mathbb{P}(N \geq 1) \leq (U_{(b+1)n} - U_{an})/U_n$ and the result follows from Assumption B and (3.8).

4. Convergence

This section is devoted to the proof of Theorem 5. The main idea of the proof is to work with a version of the process in which the environment is dynamically constructed along the trajectory of the walk and coupled to a stable process. Once this version is defined, it is rather elementary to get the convergence in a quenched form.

Remark 11. We often in this and other sections consider the inverse $g$ of a (possibly random) monotonic unbounded function $f : D \to [0, \infty)$, where $D$ may be either $\mathbb{N}$ or $[0, \infty)$, defined (as usually)

$$g(x) = \inf\{y \in D : f(y) > x\}$$

for $x \in [0, \infty)$. (Some times, we will write $f_n$ instead of $f(n)$.)

Let $X$ and $\tau$ be as in the previous sections. Assumptions A and B are in force. In the build up to the proof of Theorem 5, we start by considering a particular construction of the law of $X'$ and $A$ under $\mathbb{P}$, as follows. Let $T_0, T_1, T_2, \ldots$ be a family of independent mean 1 exponential
random variables. Consider the following random function $C : \mathbb{N} \to [0, \infty)$:

$$C_n = \sum_{i=0}^{n} \tau X_i T_i, \quad n \geq 0,$$

(4.2)

and let $I$ denote its inverse. We may call $C = (C_n)$ the clock process associated to (this particular construction of) the trap model. Now define for $t \geq 0$

$$Y_t = \tau X_t,$$

(4.3)

Note that $(X_t, Y_t)$ has the same law as $\mathcal{X}$ under $\mathbb{P}$, and $Y = (Y_t)$ has the same law as $\mathcal{A}$ under $\mathbb{P}$. Thus, making

$$Y^{(\varepsilon)}_t = \varepsilon q \varepsilon Y^{(-1)}_t, \quad t \geq 0,$$

(4.4)

we have that $Y^{(\varepsilon)} = (Y^{(\varepsilon)}_t)$ has the same law as $\mathcal{A}^{(\varepsilon)}$ under $\mathbb{P}$.

We will work with a particular version of $Y^{(\varepsilon)}$ where a specific version of the (scaled) $\tau$ random variables coupled to $Y$ are effectively placed over the range of $X$. We define this specific version in (4.15). The proof of Theorem 5 is then divided into first showing that is indeed a version, and next establishing convergence of the version. The ingredients of the version are the above defined $X$ and $Y$, and a family of i.i.d. mean 1 exponential random variables

$$\{\hat{T}^{(j)}_i, j \geq 0, i \geq 1\}.$$  

(4.5)

Let us start by enumerating the full range of $X$

$$\mathcal{R} := \cup_{n \geq 1} \mathcal{R}_n =: \{\hat{X}_0, \hat{X}_1, \ldots\},$$

(4.6)

in chronological order (in the natural way, that is, given $x, y \in \mathcal{R}$, we have $x < y$ if and only if $X$ hits $x$ before it does $y$); let $\psi : \mathbb{N} \to \mathbb{N}$ be the map

$$\psi(n) = m \quad \text{if and only if} \quad X_n = \hat{X}_m.$$  

(4.7)

We now consider properly scaled $\tau_0$-distributed random variables, to be eventually placed over $\mathcal{R}$, following the order therein. For that let us (re)introduce

$$v_n = s_n \rho_n \quad \text{and} \quad \phi_n = \frac{v_n}{r_n},$$  

(4.8)

where

$$s_n = \inf\{t \geq 0 : Q(t, \infty) \leq n^{-1}\}.$$  

(4.9)

Note that all of this functions (of $n$), namely $v, \phi, s$ are nondecreasing and unbounded. They were discussed above (in the paragraph of (2.16)); we have now a more precise definition of $s_n$. Due to the regularly varying characters of $Q$ (see (2.1)) and $r$ (Assumption B), $v, \phi, s$ are all regularly varying with a common index $\alpha^{-1}$.

It follows from elementary properties of monotonicity and regular variation of the above functions that

$$\nu \to \infty$$

(4.10)

as $\varepsilon \to 0$, where $\nu$ was defined in (2.15).

Now let

$$\bar{\varepsilon} = 1/\rho_{\nu},$$

(4.11)

and make, for $x \in \mathbb{N},$

$$\bar{\tau}^{(\varepsilon)}_x = G^{-1}(\bar{\varepsilon}^{-1/\alpha} (Y_{\bar{\varepsilon}^x + \bar{\varepsilon} - Y_{\bar{\varepsilon}x})), $$

(4.12)

where $G$ is defined by

$$\mathbb{P}(\mathcal{Y}_1 > G(y)) = \mathbb{P}(\tau_0 > y), \quad y \geq 0.$$  

(4.13)
By elementary properties of the subordinator \( \Upsilon \) and the definition of \( G \), it readily follows that 
\[ \{ \tilde{\tau}_x^{(\varepsilon)} \mid x \in \mathbb{N} \} \text{ is an i.i.d. family with } \tilde{\tau}_0^{(\varepsilon)} \overset{d}{=} \tau_0, \]
where \( \overset{d}{=} \) denotes equality in distribution.

We are now ready to define the version of the age process which will be used in the proof of Theorem 5. We start by defining the version of the clock process. For \( n \geq 0 \), let 
\[ \tilde{C}_n^{(\varepsilon)} = \sum_{y \in \mathbb{N}} \tilde{\tau}_y^{(\varepsilon)} L(\hat{X}_y, n) \sum_{i=1}^{\hat{T}_i^{(y)}}, \]
and let \( \tilde{I}^{(\varepsilon)} \) denote the inverse of \( \tilde{C}^{(\varepsilon)} \). Then make 
\[ \tilde{\tau}_t^{(\varepsilon)} = \tilde{\tau}_{\psi(\tilde{I}^{(\varepsilon)} t)}, \quad t \geq 0. \]

**Remark 12.** One may think of the latter version as one in which the environmental variables of the age process \( \tilde{\tau}_x^{(\varepsilon)} = \{ \tilde{\tau}_x^{(\varepsilon)} \mid x \in \mathbb{N} \} \) are placed over the range of \( \hat{X}_x, \{ \hat{X}_x \mid x \in \mathbb{N} \} \), respectively. The coupled construction of \( \tilde{\tau}^{(\varepsilon)} \) yields its strong convergence, when properly rescaled, to its limiting counterpart \( \Upsilon \).

**Proof of Theorem 5.** The result follows readily from Propositions 13 and 14 stated and proved below.

**Proposition 13.** For every \( \varepsilon > 0 \), \( \hat{Y}^{(\varepsilon)} \) and \( Y \) have the same distribution.

Let \( \hat{Y}_t^{(\varepsilon)} = \varepsilon q_{\hat{Y}^{(\varepsilon)} \psi^{-1} t}, \quad t \geq 0. \)

**Proposition 14.** For almost every \( \Upsilon \)
\[ \hat{Y}^{(\varepsilon)} \longrightarrow Z \quad (4.16) \]
as \( \varepsilon \searrow 0 \) in distribution on \((D,d)\).

**Remark 15.** The distribution referred to in the statement of Proposition 14 is the joint one of \( X \) and \( \{ \hat{T}_i^{(j)} \} \) (with \( \Upsilon \) fixed).

**Proof of Proposition 13.** As noted right below (4.13), we have that \( \{ \tau_{\hat{X} y} \mid y \in \mathbb{N} \} \) is independent of \( X \) and 
\[ \{ \tau_{\hat{X} y} \mid y \in \mathbb{N} \} \overset{d}{=} \{ \tilde{\tau}_y^{(\varepsilon)} \mid y \in \mathbb{N} \}. \] (4.17)

It thus follows that 
\[ C_n = \sum_{z \in \mathbb{Z}^d} \tau_z \sum_{i=1}^{L(z,n)} T_i^{(z)} = \sum_{y \in \mathbb{N}} \tau_y \sum_{i=1}^{L(\hat{X}_y,n)} T_i^{(\hat{X}_y)} = \sum_{y \in \mathbb{N}} \tilde{\tau}_y^{(\varepsilon)} \sum_{i=1}^{L(\hat{X}_y,n)} T_i^{(\hat{X}_y)} = \sum_{y \in \mathbb{N}} \tilde{\tau}_y^{(\varepsilon)} \sum_{i=1}^{L(\hat{X}_y,n)} \hat{T}_i^{(y)}, \]
(4.18)
as vectors indexed by \( n \), where \( \{ T_i^{(z)} \mid z \in \mathbb{Z}^d, i \geq 1 \} \) is an i.i.d. family of mean one exponential random variables.

**Proof of Proposition 14.** It will be implicit (and sometimes explicit) in the claims made below that they hold for almost every (a.e.)-\( \Upsilon \). We will use the symbols ‘\( P \)’ and ‘\( E \)’ to denote
the probability measure and expectation associated to the distribution of $X$ and $\{\hat{T}_i^{(j)}\}$, referred to sometimes below as dynamical random variables.

We start by defining the set of deep traps or $\delta$-traps. For $x \geq 0$, let $\mu(x) = Y(x) - Y(x-)$ and fix $\delta > 0$ arbitrarily. Consider

$$T_\delta = \{x \geq 0 : \mu(x) > \delta\} = \{x_1 < x_2 < \cdots\}. \quad (4.19)$$

Let now $y_i^{(\varepsilon)} = \lfloor \tilde{\varepsilon}^{-1} x_i \rfloor$, $i \geq 1$, and define

$$\mathcal{T}_\delta^{(\varepsilon)} = \{y_1^{(\varepsilon)}, y_2^{(\varepsilon)}, \ldots\}. \quad (4.20)$$

Our strategy will be to consider modified versions of $\hat{Y}(\varepsilon)$ and $Z$, respectively $\hat{Y}(\varepsilon, \delta)$ and $Z^{(\delta)}$, where only deep traps contribute (see (4.26)), and then show on the one hand that a version of (4.16) holds for the modified processes (see Lemma 16), and on the other hand that the errors due to the replacing $\hat{Y}(\varepsilon)$ with $\hat{Y}(\varepsilon, \delta)$ are negligible (see Lemma 22). One other aspect to be considered in adopting this modified version, besides the fact that the main contributions for $\hat{Y}(\varepsilon)$ come from its largest values, or $\delta$-traps, is that these contributions are interspersed with the lesser contributions of shallower traps. This is illustrated in Figure 1, where the $\delta$-traps are represented in solid horizontal lines and the heights of the lesser traps do not appear (but, perhaps inconsistently, the respective sojourn times do appear). This interspersion of small and large traps does not occur to $\hat{Y}(\varepsilon, \delta)$; see Figures 2 and 3.

Let us first define a modified version of $\tilde{Y}(\varepsilon)$ in terms of a restricted clock process as follows. Let

$$\tilde{Y}(\varepsilon, \delta)_{\tau}(\varepsilon, \delta)_{\psi}(\tilde{Y}(\varepsilon, \delta)_{\tilde{T}})_{\tilde{n}} = \sum_{j \geq 1} z_{y_j}^{(\varepsilon)}(y_j^{(\varepsilon)}), n), \quad (4.21)$$

where for $y \in \mathbb{N}$

$$\tilde{T}(y, n) = \sum_{i=1}^{L(X_n, n)} \tilde{T}^{(y)}_i, \quad (4.22)$$

and let $\tilde{I}(\varepsilon, \delta)$ denote the inverse of $\tilde{C}_n^{(\varepsilon, \delta)}$ as a function of $n$. Then make

$$\tilde{Y}(\varepsilon, \delta)_{\tau}(\varepsilon, \delta)_{\psi}(\tilde{T})_{\tilde{n}}(\tilde{y}_i^{(\varepsilon, \delta)})_{\tilde{T}}_{\tilde{t}}(\tau)_{\tilde{t}}, t > 0 \quad (4.23)$$

and

$$\tilde{Y}(\varepsilon, \delta)_{\tau}(\varepsilon, \delta)_{\psi}(\tilde{T})_{\tilde{n}}(\tilde{y}_i^{(\varepsilon, \delta)})_{\tilde{T}}_{\tilde{t}}(\tau)_{\tilde{t}}, t > 0. \quad (4.24)$$
Figure 2. Schematic realization of $\hat{Y}^{(\varepsilon, \delta)}$. Only values of $\delta$-traps appear, with sojourn times corresponding to actual times spent in those values by $\hat{Y}^{(\varepsilon)}$.

Figure 3. Schematic realization of $\hat{Y}^{(\varepsilon, \delta')}$, $\delta' < \delta$. See Figure 2.

Let us now define the modified versions of $Z$. Let

$$V^{(\delta)}_s = \sum_{i=1,2,\ldots; x_i \leq s} \mu(x_i)T_{x_i}, \quad s \geq 0, \quad (4.25)$$

and consider the inverse $V^{(\delta)}$ as a function of $s$, $W^{(\delta)}$. Then let

$$Z^{(\delta)}_t = \mu(W^{(\delta)}_t), \quad t \geq 0. \quad (4.26)$$

Theorem 5 readily follows from Lemmas 16, 21 and 22. $\square$

**Lemma 16.** Let $\delta > 0$ be fixed. Then as $\varepsilon \searrow 0$

$$\hat{Y}^{(\varepsilon, \delta)} \xrightarrow{\text{d}} Z^{(\delta)} \quad (4.27)$$

in distribution on $(D, J_1)$, where $J_1$ is the usual Skorohod metric on $D$.

**Proof.** We start by fixing a positive integer $K$, and analysing the behaviour of $\hat{Y}^{(\varepsilon, \delta)}$ till it hits the $K + 1$st trap of $\Sigma_\delta$; in other words, till $X$ hits $\hat{X}^{(\varepsilon)}_y$.
For $j = 1, 2, \ldots$ let us define $x_j^{(c)} = \hat{\varepsilon} y_j^{(c)}$ and
\[ \mu^{(c)}(x_j^{(c)}) = \varepsilon q \hat{\tau}_j^{(c)}, \] (4.28)

**Claim 1.** We claim that, outside an event of vanishing probability as $\varepsilon \to 0$, before $X$ hits $\hat{X}_{y_j^{(c)} + 1} \hat{Y}^{(\varepsilon, \delta)}$ does nothing except
(i) visiting $\mu^{(c)}(x_1^{(c)}), \mu^{(c)}(x_2^{(c)}), \ldots, \mu^{(c)}(x_K^{(c)})$, successively, without backtracking;
(ii) the sojourn times of those visits converge in distribution to an independent vector of $K$
 exponential random variables with means $\mu(x_1), \ldots, \mu(x_K)$, respectively.

Fontes, Isopi and Newman [18, Proposition 3.1] implies that for a.e.-$\Upsilon$
\[ \mu^{(c)}(x_j^{(c)}) \rightarrow \mu(x_j) \text{ as } \varepsilon \rightarrow 0, \] (4.29)
for every $j \geq 1$.

**Remark 17.** Some matching of the notation of [18] and the present one needs to be done in order to verify the above claim by resorting to that reference. $V$ there corresponds to $\tilde{V}$ here. Respectively, $\varepsilon$ corresponds to $\hat{\varepsilon}$. We have that $1/\nu \sim \varepsilon q$ above corresponds to $c \hat{\varepsilon}$, with $c$. introduced in (3.9) of [18], and our $\hat{\tau}_x^{(c)}$ corresponds to $\tau_x^{(\varepsilon)}$ of that reference. (4.29) follows then from the point process convergence statement in [18, Proposition 3.1].

Claim 1 and (4.29) are the main ingredients of the proof of Lemma 16.
In order to justify Claim 1, let us introduce $\zeta_i^{(c)}$, the hitting time of $\hat{X}_{y_i^{(c)}}$ by $X$, $i = 1, 2, \ldots,$ and consider
\[ L(\hat{X}_{y_i^{(c)}}, \zeta_i^{(c)}), \quad i = 1, 2, \ldots \] (4.30)

Let us state a result concerning these hitting times, proved at the end of this section.

**Lemma 18.** For all $s' > s > s'' \geq 0$ and $\delta > 0$, we have that outside an event of vanishing probability as $\varepsilon \to 0$
\[ L(\hat{X}_y, \nu \varepsilon) = 0 \text{ for all } y \geq \varepsilon^{-1}s', \text{ and } L(\hat{X}_y, \nu \varepsilon) > 0 \text{ for all } y \in \mathbb{N} \cap [0, \varepsilon^{-1}s'']. \] (4.31)

**Remark 19.** The above statement is equivalent to
\[ \frac{1}{\nu \varepsilon} \zeta_i^{(c)} \rightarrow x_i \] (4.32)
in probability as $\varepsilon \to 0$ for $i \geq 1$.

It follows from Lemma 18 that, outside an event of vanishing probability as $\varepsilon \to 0$, $\hat{\zeta}_i^{(\varepsilon, \delta)}$ vanishes if $s < x_1$, and is restricted to the first terms if $x_k < s < x_{k+1}$, $k = 1, \ldots, K$. Thus, given arbitrary $r_1 < s_1 < \cdots < r_{K+1} < s_{K+1}$ such that $0 < r_i < x_i < s_i$, $i = 1, \ldots, K + 1$, outside an event of vanishing probability as $\varepsilon \to 0$ we have that
\[ \nu \varepsilon (r_i - s_i) \leq \zeta_i^{(c)} \leq \nu \varepsilon s_i, \] (4.33)
and thus for $i = 1, \ldots, K$
\[ L(\hat{X}_{y_i^{(c)}}, \zeta_i^{(c)} + \nu \varepsilon (r_{i+1} - s_i)) \leq L(\hat{X}_{y_i^{(c)}}, \nu \varepsilon s_{i+1}) \leq L(\hat{X}_{y_i^{(c)}}, \zeta_i^{(c)} + \nu \varepsilon (s_{i+1} - r_i)). \] (4.34)
Furthermore, by Corollary 10, we obtain the following.

**Key facts** Outside an event of vanishing probability as \( \varepsilon \to 0 \), we have equalities in (4.34); namely
\[
L(\tilde{X}'_{y_j}, \zeta_{s_i} + \nu_e(r_i - s_i)) = L(\tilde{X}'_{y_j}, \zeta_{s_i+1}) = L(\tilde{X}'_{y_j}, \zeta_{s_i+1} + \nu_e(s_{i+1} - r_i)),
\]
\[i = 1, \ldots, K, \text{ and indeed also}
\]
\[
L(\tilde{X}'_{y_j}, \zeta_{s_i+1}) = L(\tilde{X}'_{y_j}, \zeta_{s_{i+1}}), \quad i = 1, \ldots, K.
\]

The first part of the Claim 1 at the beginning of this proof is thus established.

To argue the second part, we start by observing that the sojourn time of \( \tilde{Y}'_{x_j} \) on \( y_j \), \( j = 1, \ldots, K \), up until \( X \) hits \( \tilde{X}'_{y_{i+1}} \), is given by
\[
\tilde{T}_{x_j}^{(e)} = r_{\nu_e} \sum_{i=1}^{s_i} \tilde{T}_{y_j}^{(e)},
\]
so the respective time spent by \( \tilde{Y}'_{x_j} \) is \( \varepsilon \) times that, which can be rewritten as
\[
\mu^{(e)}(x_j) \tilde{T}_{j}^{(e)},
\]
where
\[
\tilde{T}_{j}^{(e)} = r_{\nu_e} \sum_{i=1}^{y_j} \tilde{T}_{y_j}^{(e)}.
\]

(Let us recall that \( q_{\nu_e} = r_{\nu_e} \).)

From the above we conclude that outside an event of vanishing probability as \( \varepsilon \to 0 \), up until \( X \) hits \( \tilde{X}'_{y_{i+1}} \), we have that
\[
\tilde{Y}'_{x_j}^{(e)} = \tilde{T}_{x_j}^{(e)} := \mu^{(e)}(\tilde{W}'_{x_j}^{(e)}),
\]
where \( \tilde{W}'_{x_j}^{(e)} \) is the inverse of
\[
\tilde{V}_{x_j}^{(e)} = \sum_{j=1, x_j \leq s_i} \mu^{(e)}(x_j) \tilde{T}_{j}^{(e)}.
\]

We now argue that \( \tilde{T}_{j}^{(e)} \), \( j = 1, 2, \ldots \), are asymptotically independent mean 1 exponential random variables as \( \varepsilon \to 0 \). By the key fact (4.35) and Remark 10, and the Markov property, we have that for \( a > 0 \), outside an event of vanishing probability as \( \varepsilon \to 0 \), each \( \tilde{T}_{j}^{(e)} \) coincides with
\[
L(\tilde{X}'_{y_j}, \zeta_j + \nu_e(r_i - s_i)) = L(\tilde{X}'_{y_j}, \zeta_{j+1}) = L(\tilde{X}'_{y_j}, \zeta_{j+1} + \nu_e(s_{j+1} - r_i)),
\]
\[j = 1, \ldots, K.
\]

Corollary 8 now implies that for any fixed \( a > 0 \) the distribution of (4.43) converges to a mean 1 exponential one as \( \varepsilon \to 0 \). This and (4.29) establish the exponentiality part of Claim 1(ii).
To establish the independence part, it is enough to argue that for $j \geq 2$ we have that $\tilde{T}_j^{(\varepsilon)}$ is asymptotically independent of the vector

$$\left(\tilde{T}_i^{(\varepsilon)}, \ i = 1, \ldots, j - 1\right). \tag{4.44}$$

By the key facts (4.35)–(4.36), we have that, outside an event of vanishing probability as $\varepsilon \to 0$,

$$L(X_{s_j^{(\varepsilon)}}) \zeta_j^{(\varepsilon)} + \nu_s (r_{j+1} - s_j)) \tilde{T}_j^{(\varepsilon)} = r_{\nu_s} \sum_{i=1}^{\tilde{T}_i^{(\varepsilon)}} \tilde{T}_i^{(\varepsilon)}. \tag{4.45}$$

Since $\zeta_j^{(\varepsilon)}$ is a stopping time, the strong Markov property yields the independence of the latter random variable and the vector in (4.44), and the claimed asymptotic independence is established. Claim 1 is thus established.

In order to complete the argument for this proof, we first observe that we may replace $D$ with $D_T$, $T > 0$ arbitrary.

Note that from the above arguments, it follows that for every $s \geq 0$ fixed, we have that

$$\tilde{V}^{(\varepsilon, \delta)}_s \to V^{(\delta)}_s \tag{4.46}$$

in distribution as $\varepsilon \to 0$. It is clear that

$$V^{(\delta)}_s \to \infty \quad \text{as } s \to \infty \tag{4.47}$$

almost surely, so given $\eta > 0$, we may choose $S \notin \mathbb{T}_\delta$ such that for all small enough $\varepsilon$

$$P(\tilde{V}^{(\varepsilon, \delta)}_S > T) > 1 - \eta. \tag{4.48}$$

Let $K = [\mathbb{T}_\delta \cap [0, S]]$. Then outside an event of probability at most $\eta$ the distance $D_T$ is bounded above by $D_{V^{(\varepsilon, \delta)}_S}$, and combining this with Claim 1 at the beginning of this proof, we get that outside an event of probability at most $2\eta$, uniformly in $\varepsilon$, within $[0, T]$, $\tilde{Y}^{(\varepsilon, \delta)}$ visits $\mu^{(\varepsilon)}(x^{(\varepsilon)}_1), \ldots, \mu^{(\varepsilon)}(x^{(\varepsilon)}_K)$ successively without backtracking (not necessarily all of them on $[0, T]$), with sojourn times converging to independent exponential random variables with means $\mu(x_1), \ldots, \mu(x_K)$, respectively. The result follows.

**Remark 20.**

1. We claimed above that the convergence in (4.46) holds for fixed times $s$. In the next section we will need a stronger version of that convergence, namely one that holds for the trajectories, under the $J_1$ metric. It is quite clear from the ingredients at end of the above proof that

$$\lim_{\varepsilon \to 0} (\tilde{V}^{(\varepsilon, \delta)}_t) \overset{d}{=} (V^{(\delta)}_t) \tag{4.49}$$

on $(D, J_1)$. Let us sketch a brief argument. Indeed, $(\tilde{V}^{(\varepsilon, \delta)}_t)$ is a jump process with jumps located at $x^{(\varepsilon)}_1, x^{(\varepsilon)}_2, \ldots$ with respective sizes $\mu^{(\varepsilon)}(x^{(\varepsilon)}_1), \mu^{(\varepsilon)}(x^{(\varepsilon)}_2), \ldots$. As established above, the jump locations converge to $x_1, x_2, \ldots$, and the jump sizes converge in distribution to independent exponentials of respective means $\mu(x_1), \mu(x_2), \ldots$, which is a description of $(V^{(\delta)}_t)$, and the convergences of jump locations and sizes imply $J_1$-convergence.

2. Another point to be used below: it immediately follows from (4.33) that if $s \notin \mathbb{T}_\delta$ then

$$\varepsilon \tilde{C}^{(\varepsilon, \delta)}_{V^{(\delta)}_s} = \tilde{V}^{(\varepsilon, \delta)}_s \tag{4.50}$$

outside an event of vanishing probability as $\varepsilon \to 0$. It follows from (4.46) that, for every fixed $s \geq 0$, $\varepsilon \tilde{C}^{(\varepsilon, \delta)}_{V^{(\delta)}_s} \to \tilde{V}^{(\delta)}_s$ in distribution as $\varepsilon \to 0$. 


In particular, we may conclude that given \( T, \eta, \delta > 0 \), there exist \( \varepsilon_0, S > 0 \) such that 
\[
\mathbb{P}(\varepsilon\bar{C}_{\nu, S} \leq T) \leq \mathbb{P}(\varepsilon\bar{C}_{\nu, S} \leq T) \leq \eta \quad \text{for all } \varepsilon < \varepsilon_0.
\]

**Lemma 21.** We have that 
\[
(V_t^{(\delta)}) \rightarrow (V_t) \quad \text{and} \quad (Z_t^{(\delta)}) \rightarrow (Z_t)
\]
almost surely on \((D, J_t)\) as \( \delta \to 0 \).

**Proof.** Using the fact that \( \sum_{x \in \mathbb{T}_\delta \cap [0, T]} \mu(x) \to 0 \) as \( \delta \to 0 \) for arbitrary \( T \), the proof follows readily. \( \square \)

We will claim several times below that for certain random variables depending on two parameters \( \varepsilon \) and \( \delta \), called generically now \( \Xi^{(\varepsilon, \delta)} \), it holds that \( \lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \mathbb{P}(\Xi^{(\varepsilon, \delta)}) = 0 \) in probability. This means that for every \( \eta > 0 \), we have that \( \lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \mathbb{P}(\Xi^{(\varepsilon, \delta)} > \eta) = 0 \).

**Lemma 22.**
\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} d((\bar{Y}_t^{(\varepsilon, \delta)}), (\hat{Y}_t^{(\varepsilon)})) = 0
\]
in probability.

**Proof.** To establish (4.52), we introduce a further auxiliary process. The reason for this is the following difficulty in the direct comparison between \( \bar{Y}^{(\varepsilon, \delta)} \) and \( \hat{Y}^{(\varepsilon)} \). The trajectories of those two processes are somewhat poorly aligned when the latter one is visiting \( \delta \)-traps, which is the only occasion they can agree, since the former process lives on \( \delta \)-traps. By poor alignment of these processes we mean that there are times when \( \bar{Y}^{(\varepsilon)} \) is visiting a given \( \delta \)-trap, while \( \hat{Y}^{(\varepsilon, \delta)} \) is visiting a different \( \delta \)-trap. This is perhaps apparent on a comparison of Figures 1 and 2. With the aim of minimizing this bad misalignment, we introduce a process, \( \bar{Y}^{(\varepsilon, \delta)} \), living on \( \delta \)-traps which is better aligned with \( \hat{Y}^{(\varepsilon)} \) in the sense that the poor alignment described above does not take place, and which is easily compared with \( \bar{Y}^{(\varepsilon, \delta)} \) as well. More on this discussion after the definition of \( \bar{Y}^{(\varepsilon, \delta)} \) next.

Let \( \bar{x}_0 = 0, \bar{x}_i = (x_i + x_{i+1})/2, i \geq 1 \), and define
\[
\bar{Y}_t^{(\varepsilon, \delta)} = \mu^{(\varepsilon)}(\bar{x}_i^{(\varepsilon)}) \quad \text{if} \quad \varepsilon\bar{C}_{\nu, \bar{x}_{i-1}} < t \leq \varepsilon\bar{C}_{\nu, \bar{x}_i}, \quad i \geq 1.
\]

Coming back to the alignment issue, we note that on any given finite time interval, outside an event of vanishing probability as \( \varepsilon \to 0 \), everytime \( \bar{Y}^{(\varepsilon)} \) is visiting a \( \delta \)-trap, \( \bar{Y}^{(\varepsilon, \delta)} \) is visiting the same trap. Since \( \bar{Y}^{(\varepsilon, \delta)} \) lives on \( \delta \)-traps, the \( d \) distance between \( \bar{Y}^{(\varepsilon, \delta)} \) and \( \bar{Y}^{(\varepsilon)} \) on a given time interval is, outside an event of vanishing probability as \( \varepsilon \to 0 \), bounded above by the size of the deepest \( \delta \)-trap in that interval multiplied by the total time spent by \( \bar{Y}^{(\varepsilon)} \) outside \( \delta \)-traps during that interval.

The comparison between \( \bar{Y}^{(\varepsilon, \delta)} \) and \( \bar{Y}^{(\varepsilon)} \) is also quite simple, perhaps simpler. Both processes live on \( \delta \)-traps, which they visit in the same order, with distinct sojourn times of order 1, which approach each other in the limit as \( \varepsilon \to 0 \). So, we indeed have the vanishing of the \( J_1 \) distance of \( \bar{Y}^{(\varepsilon, \delta)} \) and \( \bar{Y}^{(\varepsilon)} \) in the limit as \( \varepsilon \to 0 \).

We make these arguments more precisely now. We have that (4.52) follows from
\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} d((\bar{Y}_t^{(\varepsilon)}), (\bar{Y}_t^{(\varepsilon, \delta)})) = 0,
\]
\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} J_1((\bar{Y}_t^{(\varepsilon)}), (\bar{Y}_t^{(\varepsilon, \delta)})) = 0,
\]
in probability.
To establish (4.54), we claim that it is enough to consider $d_T$ instead of $d$, with

$$T = T(\varepsilon, S) = \varepsilon \sum_{y \in \mathbb{N}} \hat{\tau}_y^{(\varepsilon)} \hat{T}(y, \nu_{\varepsilon} S),$$  \hspace{1cm} (4.56)$$

$S$ fixed arbitrarily. To justify the claim, it suffices to argue that $T \to \infty$ in probability as $S \to \infty$ uniformly in $\varepsilon$ at a neighbourhood of the origin. This can be done as follows. $T$ clearly dominates $\varepsilon \mathcal{C}_{\nu_{\varepsilon} S}$ (see (4.21)); it then follows from (4.50) and (4.47) that the latter quantity diverges in probability as $S \to \infty$ uniformly in $\varepsilon$ around the origin. This closes the argument for the claim.

Reasoning now again as in the proof of Lemma 16, we find that outside an event of vanishing probability as $\varepsilon \to 0$, $\hat{Y}^{(\varepsilon)}$ and $\hat{Y}^{(\varepsilon, \delta)}$ coincide within $[0, T]$ whenever the first process is in a $\delta$-trap. Since the set of $\delta$-traps visited during $[0, T]$ is contained in $[0, S + 1]$, we conclude that the $d_T$ distance between the two processes is bounded above by

$$\max\{\mu^{(\varepsilon)}(x_i^{(\varepsilon)}), \ i \geq 1, \ x_i^{(\varepsilon)} \leq S + 1\} \varepsilon \sum_{y \in \mathbb{N} \setminus \{x_i^{(\varepsilon)}\}} \hat{\tau}_y^{(\varepsilon)} \hat{T}(y, \nu_{\varepsilon} S).$$ \hspace{1cm} (4.57)
then follows from Lemma 24 and the fact that the max is over a bounded set uniformly in \( \varepsilon \) and (4.29).

To establish (4.55), we again replace \( D \) by \( DT \), this time \( T \) deterministic, but otherwise arbitrarily fixed. For an arbitrary \( \eta > 0 \), let \( S \) be as in the second point of Remark 20. Then, arguing as in the proof of Lemma 16 (see the Claim 1 at the beginning of that proof, and also the paragraph of (4.40)), on the event that \( \varepsilon \bar{C}_{v_s} > T \) and outside an event of vanishing probability as \( \varepsilon \to 0 \), \( \bar{Y}_t^{(\varepsilon, \delta)} \) successively visits the set of states \( \{ \mu^{(\varepsilon)}(x^{(\varepsilon)}), i \geq 1: x^{(\varepsilon)}_i \leq S + 1 \} \) (not necessarily all of them by time \( T \), but in any case in that order), with respective sojourn times \( \{ \mu^{(\varepsilon)}(x^{(\varepsilon)}_i) \bar{T}_i^{(\varepsilon)}, i \geq 1: x^{(\varepsilon)}_i \leq S + 1 \} \). The same is of course true of \( \tilde{Y}_t^{(\varepsilon, \delta)} \), except that the sojourn times are given by \( \{ \tilde{S}_i^{(\varepsilon)} := \varepsilon \bar{C}_{v_s \varepsilon} - \varepsilon \bar{C}_{v_{\varepsilon, i-1}}, i \geq 1: x^{(\varepsilon)}_i \leq S + 1 \} \). Furthermore, for \( i \geq 1 \) such that \( x^{(\varepsilon)}_i \leq S + 1 \), we have that outside an event whose probability vanishes as \( \varepsilon \to 0 \), \( \tilde{S}_i^{(\varepsilon)} \geq \bar{S}_i^{(\varepsilon)} := \mu^{(\varepsilon)}(x^{(\varepsilon)}_i) \tilde{T}_i^{(\varepsilon)} \) for such \( i \), and for the same \( i \) the difference between \( \tilde{S}_i^{(\varepsilon)} \) and \( \bar{S}_i^{(\varepsilon)} \) is bounded above by

\[
\varepsilon \sum_{x \in \mathbb{Z} \cap \bar{x}_{i-1}, \bar{x}_i} \tilde{z}_x^{(\varepsilon)} \tilde{T}_x^{(\varepsilon)} (\varepsilon^{-1} x, \nu_S) \leq \varepsilon \sum_{y \in \mathbb{N} \cap \bar{y}_i^{(\varepsilon)}} \tilde{z}_y^{(\varepsilon)} \bar{T}_y^{(\varepsilon)} (y, \nu_S),
\]

and the latter expression vanishes in probability as \( \varepsilon \to 0 \) by Lemma 24. The result follows since \( \eta \) is arbitrary.

**Remark 23.** It follows from arguments in the last paragraph of the above proof together with other arguments that for all \( t > 0 \) fixed

\[
\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \sup P(\tilde{Y}_t^{(\varepsilon, \delta)} \neq \bar{Y}_t^{(\varepsilon, \delta)}) = 0.
\]

Indeed, for \( t < T \), with \( T, \eta \) and \( S \) as in the proof of (4.55) (see last paragraph of the above proof), and setting \( K = \max\{i \geq 1: x_i < S + 2\} \), let

\[
\tilde{S}_j^{(\varepsilon)} = \sum_{i=1}^j \tilde{S}_i^{(\varepsilon)}, \quad S_j^{(\varepsilon)} = \sum_{i=1}^j S_i^{(\varepsilon)}, \quad j = 1, \ldots, K,
\]

where \( \tilde{S}_j^{(\varepsilon)} \) and \( S_j^{(\varepsilon)} \) were defined right above (4.58). Then, since every random variable involved is continuous and \( \tilde{S}_j^{(\varepsilon)} \geq S_j^{(\varepsilon)}, 1 \leq j \leq K \) (outside an event whose probability vanishes as \( \varepsilon \to 0 \)), we have that \( \{ \tilde{Y}_t^{(\varepsilon, \delta)} \neq \bar{Y}_t^{(\varepsilon, \delta)} \} \) is almost surely contained in \( \bigcup_{j=1}^K \{ \tilde{S}_j^{(\varepsilon)} < t < S_j^{(\varepsilon)} \} \). Now the fact argued at end of the above proof that the difference between \( \tilde{S}_j^{(\varepsilon)} \) and \( S_j^{(\varepsilon)} \) vanishes as \( \varepsilon \to 0 \) in probability for \( j = 1, \ldots, K \), and the fact that \( \tilde{S}_1^{(\varepsilon)}, \ldots, \tilde{S}_K^{(\varepsilon)} \) are asymptotically independent, lead readily to the completion of the argument.

**Lemma 24.** We have that

\[
\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \sup \varepsilon \sum_{y \in \mathbb{N} \setminus \bar{y}_i^{(\varepsilon)}} \tilde{z}_y^{(\varepsilon)} \tilde{T}_y^{(\varepsilon)} (y, \nu_s) = 0
\]

in probability for every \( s \geq 0 \).

**Proof.** We start by resorting to the Markov property and Corollary 7 to find that for every \( s > 0 \)

\[
r_{\nu_s} E(\tilde{T}_y^{(\varepsilon)} (y, \nu_s s)) \leq r_{\nu_s} E(L(0, \nu_s s)) = r_{\nu_s} U_{\nu_s} \to 1 \quad \text{as} \quad \varepsilon \to 0.
\]

(4.62)
We next resort to Lemma 18 to restrict the sum in (4.61) on \( x \leq \tilde{\varepsilon}^{-1}s' \). By (4.62), the expectation of the restricted sum is bounded above by constant times

\[
\sum_{y \in \mathbb{N} \cap (\Xi^u_\delta)^\ast \cap [0, \tilde{\varepsilon}^{-1}s'] \cap [0, \tilde{\varepsilon}^{-1}s']}
\]

(recall that \( q_\varepsilon = r_{\nu_\varepsilon}^{-1} \)). We now claim that the \( \lim_{\varepsilon \to 0} \limsup_{\varepsilon \to 0} \) of (4.63) vanishes almost surely. In order to use results of [18] for that, let us extend to \( \tilde{\varepsilon} \mathbb{N} \) the measure \( \mu^{(\varepsilon)} \) defined above for rescaled \( \delta \)-trap sites only (see (4.28)). For \( x \in \tilde{\varepsilon} \mathbb{N} \), let

\[
\mu^{(\varepsilon)}(x) = \varepsilon q_\varepsilon \tilde{z}_\varepsilon^{(\varepsilon)}.
\]

Here we abuse notation by writing \( \mu^{(\varepsilon)}(\cdot), \mu(\cdot) \) instead of \( \mu^{(\varepsilon)}(\{\cdot\}), \mu(\{\cdot\}) \).

We may then rewrite the sum (4.63) as

\[
\mu^{(\varepsilon)}([0, s']) - \mu^{(\varepsilon)}([0, s'] \setminus \tilde{\varepsilon} \Xi^u_\delta)
\]

Now by (4.29) we have that the second term of (4.65) converges to \( \mu([0, s'] \cap \Xi_\delta) \) almost surely as \( \varepsilon \to 0 \). It follows from [18, Proposition 3.1] (the vague convergence part) that \( \mu^{(\varepsilon)}([0, s']) \to \mu([0, s']) = \Upsilon(s') \) almost surely as \( \varepsilon \to 0 \); see Remark 17. Here, we may use the fact that deterministic points like \( s' \) are continuity points of \( \Upsilon \) almost surely.

We then have that (4.63) converges almost surely to

\[
\mu([0, s'] \setminus \Xi_\delta)
\]

as \( \varepsilon \to 0 \). It is again a standard fact that the expression in (4.66) vanishes almost surely as \( \delta \to 0 \), and the claim and lemma follow.

**Proof of Lemma 18.** We start by pointing out that

\[
L(\tilde{X}_y, \nu_\varepsilon s) = 0 \quad \text{if and only if} \quad y > R_{\nu_\varepsilon, s}.
\]

Given \( \eta > 0 \) to be chosen below, consider the event

\[
\left\{ \left| \frac{R_{\nu_\varepsilon, s}}{\rho_{\nu_\varepsilon}} - 1 \right| > \eta \right\}.
\]

(4.68)

Assumption A implies that the probability of this event vanishes as \( \varepsilon \to 0 \).

Now write \( \rho_{\nu_\varepsilon} \) as

\[
\frac{\rho_{\nu_\varepsilon}}{\nu_\varepsilon \nu_{\nu_\varepsilon} \nu_{\nu_\varepsilon}^{-1}}
\]

From (3.3) and (4.10), we get that the \( \lim_{\varepsilon \to 0} \) of the first and third quotients in (4.69) both equal 1. Assumption B implies that the \( \lim_{\varepsilon \to 0} \) of the second quotient equals 1.

Now let us choose \( \eta > 0 \) small enough so that \( (1 + 2\eta)s \leq s' \) and \( (1 - 2\eta)s \leq s'' \). From the conclusion of the previous paragraph, we get that, for all \( \varepsilon \) small enough, the events

\[
\{R_{\nu_\varepsilon, s} \geq \varepsilon^{-1} x \text{ for some } x > s'\}
\]

\[
\{R_{\nu_\varepsilon, s} \leq \varepsilon^{-1} x \text{ for some } x \leq s''\}
\]

are contained in event (4.68), which, as already pointed out, has vanishing probability as \( \varepsilon \to 0 \).

Since outside the event

\[
\{R_{\nu_\varepsilon, s} \geq \varepsilon^{-1} x \text{ for some } x > s'\} \cup \{R_{\nu_\varepsilon, s} \leq \varepsilon^{-1} x \text{ for some } x \leq s''\}
\]

we have that (4.67) yields (4.31), the result follows.
5. Ageing

We will consider the following two ageing functions of \( (Y_t) \).

\[
\bar{R}(s, t) = \mathbb{P}(Y_t = Y_{t+s}), \quad (5.1)
\]

\[
\bar{\Pi}(s, t) = \mathbb{P}(Y_t = Y_{t+r} \text{ for all } r \in [0, s]), \quad (5.2)
\]

with the following result.

**Theorem 25.** If \( X \) is transient, then there exist non-trivial functions \( R, \Pi : [0, \infty) \to (0, 1] \) such that

\[
\lim_{t \to \infty} \bar{R}(\theta t, t) = R(\theta), \quad (5.3)
\]

\[
\lim_{t \to \infty} \bar{\Pi}(\theta t/q_{1/t}, t) = \Pi(\theta). \quad (5.4)
\]

**Remark 26.** We do not have any strong reason to believe that the result does not hold generally under Assumptions A and B. The restriction to transient processes is technical. Our argument below requires, roughly speaking, that during time \( t \) a single visit of \( Y \) to a deep trap lasts a length of time of order \( t \), and this occurs only in the transient case. In this case, both assumptions A and B hold, with \( \rho_n \sim n/q, r_n \sim 1/q, \) and \( q_{1/t} \sim q \), where \( 1 < q < \infty \) is the same constant appearing in the statement of Theorem 5.

We state below a (weaker) version of this result for integrated versions of the ageing functions, or equivalently, for those functions looked at suitable random times.

**Remark 27.** Let \( \tilde{X}_t = X_{I_t} \). In the literature one has rather considered the ageing functions

\[
R(s, t) = \mathbb{P}(\tilde{X}_t = \tilde{X}_{t+s}), \quad (5.5)
\]

\[
\Pi(s, t) = \mathbb{P}(\tilde{X}_t = \tilde{X}_{t+r} \text{ for all } r \in [0, s]). \quad (5.6)
\]

In case \( \tau_0 \) is a continuous random variable, then we of course have the identities \( R(\cdot, \cdot) = \bar{R}(\cdot, \cdot) \) and \( \Pi(\cdot, \cdot) = \bar{\Pi}(\cdot, \cdot) \), but not otherwise. In any case, one can show that ageing results like (5.3), (5.4) hold for \( R(\cdot, \cdot) \) and \( \Pi(\cdot, \cdot) \) as well, with \( R(\cdot) \) and \( \Pi(\cdot) \) as limiting ageing functions, respectively.

**Remark 28.** \( R \) and \( \Pi \) turn out to be identical. We have

\[
R(\theta) = \Pi(\theta) = \frac{\sin(\pi \alpha)}{\pi} \int_{\theta/(1+\theta)}^{1} s^{-\alpha} (1-s)^{\alpha-1} ds. \quad (5.7)
\]

See (5.18), (5.23) and Remark 31.

In order to prove Theorem 25, we will naturally consider the rescaled version \( \bar{Y}^{(\varepsilon)} \) of \( Y \) with the special strongly converging rescaled environment (see (4.15) above). One ingredient of the proof of (5.3) is a comparison to \( Y^{(\varepsilon, \delta)} \) (see (4.53) above) as follows.

**Lemma 29.** For all \( t > 0 \) fixed, if \( X \) is transient, then we have that

\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \mathbb{P}(Y^{(\varepsilon)}_t \neq \bar{Y}^{(\varepsilon, \delta)}_t) = 0 \quad (5.8)
\]

for a.e. \( \Upsilon \).
Proof. As in previous arguments, we will leave implicit many times below that claims hold for a.e. \( \mathcal{Y} \). \( P \) and \( E \) remain as notation for the probability and expectation with respect to the distribution of the dynamical random variables (\( X \) and \( \hat{T}_t^{(j)} \)). We first consider for \( T > 0 \)

\[
\int_0^T P(\tilde{Y}_{s}^{(\varepsilon)} \neq \tilde{Y}_{s}^{(\varepsilon, \delta)}) \, ds = E \int_0^T 1\{\tilde{Y}_{s}^{(\varepsilon)} \neq \tilde{Y}_{s}^{(\varepsilon, \delta)}\} \, ds. \tag{5.9}
\]

We will argue as in the proofs of Lemmas 16 and 22. Let us first fix an arbitrary \( \eta > 0 \), and then choose \( S \) as in the second point of Remark 20. Then on the event that \( \varepsilon \tilde{C}_{\nu, S}^{(\varepsilon)} > T \) and outside an event of vanishing probability as \( \varepsilon \to 0 \), the integral on the right of (5.9) is bounded above by

\[
\sum_{x \in \mathfrak{H}(\mathfrak{T}_X^{(\varepsilon)}) \cap S+1} \mu(x) \tilde{T}(\varepsilon^{-1}, \nu, S). \tag{5.10}
\]

Using now (4.62) and the fact argued below (4.63), we have that the \( \lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \) of the expectation of the integral is bounded above by \( T \eta \). Since \( \eta \) is arbitrary, we have that

\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \int_0^T P(\tilde{Y}_{s}^{(\varepsilon)} \neq \tilde{Y}_{s}^{(\varepsilon, \delta)}) \, ds = 0. \tag{5.11}
\]

We now fix \( \delta' > \delta \) and define \( I = \min\{i \geq 1 : x_i \in \mathfrak{T}_{\mathfrak{Y}}\} \). Let \( \zeta^{(\varepsilon)} \) be as in the proof of Lemma 16 (see paragraph of (4.30)). Let \( t > 0 \) be fixed and condition on

\[
\bar{C}^{(\varepsilon)} := \varepsilon \tilde{C}_{\nu, t-1-1}^{(\varepsilon)} = \mu(x_I^{(\varepsilon)}) \tilde{T}(y_I^{(\varepsilon)}, \zeta_I^{(\varepsilon)}) + \Delta^{(\varepsilon)}, \tag{5.12}
\]

where the latter summand, \( \Delta^{(\varepsilon)} \), is defined by this equality. Note that the former summand is an exponential random variable of mean \( \mu(x_I^{(\varepsilon)}) r_{\nu(\varepsilon^{-1})} \), and that given \( X \) (and \( \mathfrak{Y} \)) the summands are absolutely continuous random variables independent of each other.

It thus follows that

\[
P(\tilde{Y}_t^{(\varepsilon)} \neq \tilde{Y}_t^{(\varepsilon, \delta)} | X) \leq \int_0^T \left( \int_0^s b_s^{-1} e^{-b_s(s-r) f^{(\varepsilon)}(r)} \, dr \right) P(\tilde{Y}_t^{(\varepsilon)} \neq \tilde{Y}_t^{(\varepsilon, \delta)} | \bar{C}^{(\varepsilon)} = s, X) \, ds
\]

\[
+ P(\bar{C}^{(\varepsilon)} > t | X), \tag{5.13}
\]

where \( b_s^{-1} = \mu(x_I^{(\varepsilon)}) r_{\nu(\varepsilon^{-1})} \) and \( f^{(\varepsilon)} \) is the density of \( \Delta^{(\varepsilon)} \) given \( X \). The above integral is thus upper bounded by

\[
b_s \int_0^t P(\tilde{Y}_t^{(\varepsilon)} \neq \tilde{Y}_t^{(\varepsilon, \delta)} | \bar{C}^{(\varepsilon)} = s, X) \, ds. \tag{5.14}
\]

Now, by the independence of increments of \( (\tilde{C}_n^{(\varepsilon)}) \) given \( X \), the probability inside the latter integral can be written as

\[
P(\tilde{Y}_t^{(\varepsilon)} \neq \tilde{Y}_{t-s}^{(\varepsilon, \delta)} | X), \tag{5.15}
\]

where \( (\tilde{Y}_t^{(\varepsilon)}) \) and \( (\tilde{Y}_{t-s}^{(\varepsilon, \delta)}) \) are defined as \( (\tilde{Y}_t^{(\varepsilon)}) \) and \( (\tilde{Y}_{t-s}^{(\varepsilon, \delta)}) \), respectively, with \( \tilde{C}_n^{(\varepsilon)} - \tilde{C}_{\nu, t-1-1}^{(\varepsilon)} \) replacing \( \tilde{C}_n^{(\varepsilon)} \). We thus obtain (after integrating in \( X \))

\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \int_0^t P(\tilde{Y}_{t-s}^{(\varepsilon)} \neq \tilde{Y}_{t-s}^{(\varepsilon, \delta)}) \, ds = \lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \int_0^t P(\tilde{Y}_{t-s}^{(\varepsilon)} \neq \tilde{Y}_{t-s}^{(\varepsilon, \delta)}) \, ds = 0 \tag{5.16}
\]

similarly as we did (5.11). And we have that \( \lim_{\delta \to 0} \liminf_{\varepsilon \to 0} b_s^{-1} \geq \delta'/q > 0 \), since \( x_I \in \mathfrak{T}_{\mathfrak{Y}} \) and \( \lim_{\varepsilon \to 0} r_{\nu(\varepsilon^{-1})} = 1/q > 0 \). We thus get that the \( \lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \) of the first term in (5.13) vanishes, and since \( \delta' \) is arbitrary and \( \lim_{\delta' \to 0} \limsup_{\varepsilon \to 0} \tilde{C}^{(\varepsilon)} = 0 \) in probability, as can be readily checked, the result follows.
Remark 30. In the recurrent case $b_\varepsilon$ is not bounded, and thus the above argument breaks down.

Proof of Theorem 25. We will first replace $\hat{Y}^{(\varepsilon)}$ by $\hat{Y}^{(\varepsilon,\delta)}$ and then resort to Lemma 29 and Remark 23. By Lemma 16, we have that
\[
\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \hat{Y}^{(\varepsilon,\delta)} = Z
\] (5.17)
in distribution on $(D,J_1)$. From that and Lemma 29 and Remark 23, we claim that
\[
\lim_{\varepsilon \to 0} \bar{R}(\theta^{-1},\varepsilon^{-1}) = \lim_{\delta \to 0} \lim_{\varepsilon \to 0} \mathbb{P}(\hat{Y}_1^{(\varepsilon,\delta)} = \hat{Y}_{1+\theta}) = \mathbb{P}(Z_1 = Z_{1+\theta}) := R(\theta).
\] (5.18)
(5.3) then follows. The only point of the claim that needs arguing is the second equality. We first point out that from the construction of $\hat{Y}^{(\varepsilon,\delta)}$ (see (4.26)), since almost surely $\mu(x_i) \neq \mu(x_j)$ whenever $i \neq j$, we have from (4.29) that, for all fixed $\delta$ and all small enough $\varepsilon$, the probability in the second term in (5.18) equals
\[
\mathbb{P}(\hat{Y}_1^{(\varepsilon,\delta)} = \hat{Y}_{1+r}) \text{ for all } r \in [0,\theta])
\] (5.19)
plus a small error, and the latter probability equals
\[
\mathbb{P}([1,1+\theta] \cap \text{ range of } \hat{Y}^{(\varepsilon,\delta)} = \emptyset).
\] (5.20)
It readily follows from (4.49) and Remark 20.1 that the second term in (5.18) equals
\[
\mathbb{P}([1,1+\theta] \cap \text{ range of } V = \emptyset) = \mathbb{P}(Z_1 = Z_{1+r}, \text{ for all } r \in [0,\theta]).
\] (5.21)
Now the right-hand side of (5.21) equals that of (5.18) (since almost surely $\mu(x_i) \neq \mu(x_j)$ whenever $i \neq j$). (5.3) is then settled.

In the above argument, we felt the need to go through (5.19) and (5.20), since the (indicators of the) events in the second probability in (5.18) and in (5.19) are not almost surely continuous on $(D,J_1)$, but so is the event in (5.20).

As regards (5.4), we have that
\[
\Pi(\theta e^{-1}/q_\varepsilon,\varepsilon^{-1}) = \mathbb{P}(Y_{\varepsilon-1} = Y_{\varepsilon-1+r} \text{ for all } r \in [0,\theta e^{-1}/q_\varepsilon]) = \mathbb{E}(e^{-\theta/(\varepsilon q_{\varepsilon-1})}) = \mathbb{E}(e^{-\theta/\hat{Y}^{(\varepsilon)}})
\] (5.22)
and Lemma 29 implies that
\[
\lim_{\varepsilon \to 0} \mathbb{E}(e^{-\theta/\hat{Y}^{(\varepsilon)}_1}) = \lim_{\delta \to 0} \lim_{\varepsilon \to 0} \mathbb{E}(e^{-\theta/\hat{Y}^{(\varepsilon,\delta)}_1}) = \mathbb{E}(e^{-\theta/Z_1}) =: \Pi(\theta),
\] (5.23)
where the second equality follows from (5.17), which implies marginal convergence in distribution (since each fixed deterministic time is almost surely a continuity point of $Z$).

Remark 31. One quickly checks that
\[
\mathbb{P}(Z_1 = Z_{1+\theta}) = \mathbb{P}(Z_1 = Z_{1+r}, \text{ for all } r \in [0,\theta]),
\] (5.24)
which equals $\mathbb{P}([1,1+\theta] \cap \text{ range of } V = \emptyset)$, as noted in (5.21). One can then obtain the right-hand side of (5.7) as an expression for the latter probability. (This may be readily seen to follow from [11, Proposition 3.1], since $V$ is an $\alpha$-stable subordinator; see Remark 3.) We further notice that we can write
\[
\mathbb{P}(Z_1 = Z_{1+\theta}) = \mathbb{E}(e^{-\theta/Z_1}).
\] (5.25)
See Remark 28. (5.24) and (5.25) give us the Laplace transform of $1/Z_1$. An expression for the density of that variable can be found in (5.97) of [20].
Remark 32. Another ageing function which is natural on one side and not as considered in the literature as the above ones on the other side, and also fits well in the above picture, is the following one.

\[ \Omega(s, t) = \mathbb{P} \left( \sup_{r \in [0, t]} Y_r < \sup_{r \in [0, t+s]} Y_r \right). \] (5.26)

It was suggested in [18] as a ‘measure of the prospects for novelty in the system’. Since

\[ \sup_{r \in [0, t]} \hat{Y}_r(\varepsilon) = \sup_{r \in [0, t]} Y_r(\varepsilon, \delta) \] if \( \hat{Y}_r(\varepsilon) = \hat{Y}_r(\varepsilon, \delta) \), from Lemma 29, (4.55), Lemma 16 and (4.51), we have that

\[ \lim_{t \to \infty} \Omega(\theta t, t) = \mathbb{P} \left( \sup_{r \in [0, 1]} Z_r < \sup_{r \in [0, 1+\theta]} Z_r \right) =: \Omega(\theta). \] (5.27)

This is an example where the limiting ageing function requires full use of the process \( Z \); in the previous cases, the limits could be expressed in terms of the (clock) process \( V \) alone. We could not find an explicit expression for the right-hand side of (5.27).

5.1. Integrated ageing results

By considering integrated ageing functions, or equivalently ageing functions looked at random times, we may circumvent the difficulties exhibited by recurrence, see Remarks 26 and 30. As an example, let us consider one such integrated ageing function, and state the corresponding result.

Let

\[ \mathcal{R}(\lambda, \mu) = \mathbb{E}[\hat{R}(\lambda T, \mu T)] = \int_0^\infty e^{-t} \hat{R}(\lambda t, \mu t) dt, \] (5.28)

with \( \hat{R} \) as in (5.1), and \( T \) a mean one exponential random variable independent of every other random variable in the problem.

**Theorem 33.** If \( X \) satisfies Assumptions A and B, then

\[ \lim_{\lambda, \mu \to \infty} \frac{\lambda}{\mu} = \theta \] \[ \lim_{\lambda, \mu \to \theta} \mathcal{R}(\lambda, \mu) = R(\theta), \] (5.29)

with \( R \) as in Theorem 25.

**Proof.** Let us for simplicity take \( \lambda = \theta \mu \). We may then write

\[ \mathcal{R}(\lambda, \mu) = \mathbb{E} \int_0^\infty e^{-t} 1 \{ Y_{\mu t} = Y_{(\mu+\lambda)t} \} dt = \mathbb{E} \int_0^\infty e^{-t} 1 \{ \hat{Y}_t(\varepsilon) = \hat{Y}_t(\varepsilon, \delta) \} dt, \] (5.30)

where \( \varepsilon = \mu^{-1} \).

Given \( 0 < \eta < 1 \), (4.54) allows us to bound the right-hand side of (5.30) above and below by \( (1 \pm \eta) \) times

\[ \lim_{\varepsilon \to 0} \mathbb{E} \int_0^\infty e^{-t} 1 \{ \hat{Y}_t(\varepsilon, \delta) = \hat{Y}_t(\varepsilon, \delta, \theta) \} dt = \mathbb{E} \int_0^\infty e^{-t} 1 \{ Z_t(\varepsilon) = Z_t(\varepsilon, \theta) \} dt, \] (5.31)

respectively, as soon as \( \delta \) is close enough to 0, where the latter equality follows from Lemma 16. Since \( \eta \) is arbitrary, by Lemma 21 the left-hand side of (5.30) equals

\[ \mathbb{E} \int_0^\infty e^{-t} 1 \{ Z_t = Z_t(1+\theta)t \} dt = \int_0^\infty e^{-t} \mathbb{P}(Z_t = Z_t(1+\theta)t) dt, \] (5.32)
since every fixed \( t \) is almost surely a continuity point of \( Z \). Now by the self-similarity of index 1 exhibited by \( Z \), the probability inside the integral does not depend on \( t > 0 \), and the result follows, since \( P(Z_1 = Z(1+\theta)) = R(\theta) \).

Similar results can be argued for integrated versions of the ageing functions \( \tilde{\Pi} \) and \( \Omega \) above.

6. Stronger convergence

In this section, we strengthen the convergence results of Section 4 under an additional condition, which we now explain.

Let \( X' = (X'_n)_{n \geq 0} \) a random walk independent from and equally distributed with \( X \) and define

\[
I_n = I_n(X, X') = \{ z \in \mathbb{Z}^d : X_i = X'_j = z \text{ for some } 0 \leq i, j \leq n \} = R_n(X) \cap R_n(X'), \quad n \geq 0,
\]

as the set of intersection points of the paths of \( X \) and \( X' \) up to (discrete) time \( n \) (it can be seen also as indicated as the intersection of the ranges of \( X \) and \( X' \) up to time \( n \)). Let now

\[
I_n = |I_n|
\]

be the number of such intersection points. The additional condition we impose, in order that the results of this section hold, is as follows.

**Assumption C.**

\[
\frac{I_n}{E(R_n)} \rightarrow 0 \quad \text{in probability as } n \rightarrow \infty.
\]

**Remark 34.** The expectation of the quotient in (6.3) can be reexpressed as

\[
\frac{\sum_{x \in \mathbb{Z}^d} [P(T_x \leq n)]^2}{\sum_{x \in \mathbb{Z}^d} P(T_x \leq n)},
\]

where \( T_x = \inf\{n \geq 0 : X_n = x\} \). We readily find that (6.3) holds in either the general \( d \geq 2 \) transient case, or the one-dimensional nonintegrable increment, transient case, since in both these cases \( \limsup_{\|x\| \rightarrow \infty} P(T_x < \infty) = 0 \) (see, for example, [27, Proposition 25.3]), and in general \( \lim_{n \rightarrow \infty} E(R_n) = \infty \).

It also holds for two-dimensional mean zero, finite second moment random walks from results in [23]. We are uncertain about other recurrent planar walks, as well as 1-stable one-dimensional recurrent walks. (See Remark 2.)

Let \( B_u \) be the class of bounded uniformly continuous real functions on \((D,d)\). Here is the main result of this section. Let \( Y^{(\epsilon)} \) and \( Z \) be as in Theorem 5.

**Theorem 35.** Under Assumptions A–C, for every \( F \in B_u \), we have

\[
\mathbb{E}[F(Y^{(\epsilon)})|\tau] \rightarrow \mathbb{E}[F(Z)],
\]

in probability as \( \epsilon \rightarrow 0 \).

**Remark 36.** As anticipated at the end of Section 2, a condition like (6.3) is needed for the validity of the above result. A case where Assumptions A and B are satisfied, but not...
Assumption C, and (6.5) does not hold, is when 

X is one-dimensional simple asymmetric. This

is particularly clear in the totally asymmetric case, when 

C_n (see (4.2)) is a partial sum of

i.i.d. random variables in the basin of attraction of an \( \alpha \)-stable law, \( \alpha \in (0, 1) \), in which case it

is well known to only converge, when properly rescaled, in distribution. This prevents a result

of the form of (6.5) in that case.

Proof of Theorem 35. Let \( F \in B_u \) be fixed. We may and will restrict to \( F \) with bounded

support, say \([0, T]\), where \( T > 0 \) is arbitrary.

It follows from Theorem 5 that

\[
\mathbb{E}[F(Y^{(\varepsilon)})] \longrightarrow \mathbb{E}[F(Z)].
\] (6.6)

We will use this and (6.3) to get (6.5).

The strategy is to construct two sets of versions of \( Y^{(\varepsilon)} \). In each set of versions, the different

versions have independent dynamical random variables. The distinction is on the environmental

variables. On one set of versions the respective environments are also independent among

distinct versions, so that the versions are fully independent of one another: for this set the

empirical mean of \( F \) over the different versions yields the left-hand side of (6.6) in the limit

as the number of versions grows. On the other set of versions, we have a single environment

for all the versions, so the empirical mean of \( F \) over the different versions yields the left-

hand side of (6.5). However, this single environment can be constructed in a coupled way to

the independent environments of the first set of versions, so that the difference between the

empirical mean over the different first set of versions and that over the second set of versions vanishes

as \( \varepsilon \to 0 \). This yields the result as soon as the limits as the number of versions grows and as

\( \varepsilon \to 0 \) are suitably taken. We define the two sets of versions now.

Let \( X^{(1)}, X^{(2)}, \ldots \) and \( \tau^{(0)}, \tau^{(1)}, \ldots \) be i.i.d. copies of \( X \) and \( \tau \), respectively. For \( k \geq 1 \), let

\( R^{(k)} \) be defined as in (2.7), with \( X^{(k)} \) replacing \( X \). Let now \( Z_n^{(1)} = R_n^{(1)} \) and for \( k > 1 \)

\[
Z_n^{(k)} = R_n^{(k)} \setminus \bigcup_{i=1}^{k-1} R_n^{(i)}.
\] (6.7)

We then define for each \( N \geq 1 \)

\[
\tilde{\tau}^{(N)} = \{ \tilde{\tau}^{(N)}_x, x \in \mathbb{Z}^d \},
\] (6.8)

where \( \tilde{\tau}^{(N)}_x = \tilde{\tau}^{(k)}_x \), if \( x \in \mathbb{Z}^{(k)} \) for some \( k \geq 1 \), and \( \tilde{\tau}^{(0)}_x = \tau^{(0)}_x \), otherwise.

Remark 37. \( \tilde{\tau}^{(N)} \) and \( \tau \) are equally distributed for every \( N \geq 1 \), whether or not

\( X^{(1)}, X^{(2)}, \ldots \) are given. In particular, \( \tilde{\tau}^{(N)} \) is independent of \( X^{(1)}, X^{(2)}, \ldots \).

Now let us consider two classes of clock processes \( C^{(k,N)}, C^{(k)} : \mathbb{N} \to [0, \infty) \), \( k,N \geq 1 \):}

\[
C^{(k,N)}_n = \sum_{i=0}^{n} \tilde{\tau}^{(N)}_{X^{(k)}_i} T^{(k)}_i, \quad C^{(k)}_n = \sum_{i=0}^{n} \tau^{(k)}_{X^{(k)}_i} T^{(k)}_i, \quad n \geq 0,
\] (6.9)

where \( \{T^{(k)}_i, \ i \geq 1\} =: T^{(k)}, \ k \geq 1 \), are independent families of i.i.d. mean 1 exponentials, and

their respective inverses \( I^{(k,N)} \) and \( I^{(k)} \). Let then

\[
Y^{(k,N)}_t = \tau^{(N)}_{X^{(k)}_t}, \quad Y^{(k)}_t = \tau^{(k)}_{X^{(k)}_t}, \quad t \geq 0.
\] (6.10)

Remark 38. We remark that \( C^{(k,N)} = C^{(k)} = C \) and \( Y^{(k,N)} = Y^{(k)} = Y \) in distribution

for all \( k,N \); and that \( Y^{(k)} \), \( k \geq 1 \), are i.i.d.; and \( Y^{(k,N)} \), \( k \geq 1 \), are i.i.d. given \( \tilde{\tau}^{(N)} \) for all \( N \).
The latter fact follows from the independence of $\tilde{\tau}^{(N)}$ from $X^{(1)}, X^{(2)}, \ldots$ as remarked above (see Remark 37).

The reason why the argument we outlined above, and then started to fill the details of, works is that, as already remarked and used in the previous section, the contributions to the processes come from only a few deep traps (in the sense specified at the proof of Theorem 5, see (4.19) and (4.20)), which are far off one another, and, given Assumption $C$, are unlikely to lie in any intersection of ranges. For this reason the difference between $Y^{(k)}$ and $Y^{(k,N)}$ (suitably rescaled) come from sites which are not deep traps, and thus are negligible. In order to make this argument, let us fix $\varepsilon, \delta > 0$ and define for each $k, N \geq 1$

\[
\tau^{(k, \varepsilon, \delta)} = \left\{ \tau^{(k)}_x := \tau^{(k)}_x 1 \{ \tau^{(k)}_x > \delta(\varepsilon q_x)^{-1} \}, \quad x \in \mathbb{Z}^d \right\},
\]

\[
\tilde{\tau}^{(N, \varepsilon, \delta)} = \left\{ \tilde{\tau}^{(N)}_x := \tilde{\tau}^{(N)}_x 1 \{ \tilde{\tau}^{(N)}_x > \delta(\varepsilon q_x)^{-1} \}, \quad x \in \mathbb{Z}^d \right\}
\]

and let

\[
C^{(k, N, \varepsilon, \delta)}_n = \sum_{i=0}^n \tau^{(k, \varepsilon, \delta)}_X, \quad C^{(k, \varepsilon, \delta)}_n = \sum_{i=0}^n \tau^{(k, \varepsilon, \delta)}_Y, \quad n \geq 0,
\]

with $I^{(k, N, \varepsilon, \delta)}$ and $I^{(k, \varepsilon, \delta)}$ their respective inverses, and

\[
Y^{(k, \varepsilon, \delta)} = \varepsilon q_{\tau^{(N)}} X^{(k)}_1, \quad Y^{(k, N, \varepsilon, \delta)} = \varepsilon q_{\tilde{\tau}^{(N)}} X^{(k)}_1,
\]

\[
Y^{(k, \varepsilon, \delta)} = \varepsilon q_{\tau^{(N)}} X^{(k)}_1, \quad Y^{(k, \varepsilon, \delta)} = \varepsilon q_{\tilde{\tau}^{(N)}} X^{(k)}_1.
\]

We then have that $Y^{(k, N, \varepsilon)} = Y^{(k, \varepsilon)}$ in distribution; $Y^{(k, \varepsilon)}$, $k \geq 1$, are i.i.d.; and $Y^{(k, \varepsilon)}$, $k \geq 1$, are i.i.d. given $\tilde{\tau}^{(N)}$ for all $N$.

Consider now

\[
\frac{1}{K} \sum_{k=1}^K F(Y^{(k, N, \varepsilon)}) = \frac{1}{K} \sum_{k=1}^K F(Y^{(k, N, \varepsilon)}) + \frac{1}{K} \sum_{k=1}^K \Delta^{(k, N, \varepsilon, \delta)},
\]

where $K$ is an arbitrary positive integer, and $\Delta^{(k, N, \varepsilon, \delta)} = F(Y^{(k, N, \varepsilon)}) - F(Y^{(k, N, \varepsilon)}).

With an argument similar to the one giving (4.52), one finds that

\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \Delta^{(k, N, \varepsilon, \delta)} = 0
\]

in probability for all $k$.

Since $C^{(k, N)} = C^{(k)} = C$ in distribution, we have that, given $\eta > 0$, there exists $S = S_K > 0$ such that

\[
\mathbb{P}(C^{(k, N)}_{\nu_{\varepsilon} S} \leq \varepsilon^{-1} T) = \mathbb{P}(C^{(k)}_{\nu_{\varepsilon} S} \leq \varepsilon^{-1} T) \leq \frac{\eta}{2K},
\]

for all $N \geq 1$ and $\varepsilon$ sufficiently small (see the second point of Remark 20).

From now on we take $N = N_{\varepsilon} = 2\rho_{\nu_{\varepsilon} S}$.

For $k, \ell = 1, \ldots, K$, let

\[
A_{k, \ell} = \{ \tau^{(k)}_X > \delta(\varepsilon q_x)^{-1} \text{ and } X^{(k)}_i \in R^{(\ell)}_{\nu_{\varepsilon} S} \text{ for some } i \leq \nu_{\varepsilon} S \}, \quad A_K = \cup_{k, \ell=1}^K A_{k, \ell}.
\]

Let also

\[
I_{k, \ell} := I_{\nu_{\varepsilon} S}(X^{(k)}, X^{(\ell)})
\]

(see (6.1)).
Then, given \( \lambda > 0 \)
\[
\mathbb{P}(A_{k,\ell}) = \mathbb{P} \left( \sum_{x \in I_{k,\ell}} 1 \left\{ \tau_x^{(k)} > \delta (q_x)^{-1} \right\} \geq 1 \right) 
\leq \lambda N \mathbb{P}(\tau_0 > \delta (q_x)^{-1}) + \mathbb{P}(|I_{k,\ell}| > \lambda N) + \mathbb{P}(R_{\nu,S}^{(k)} > N). 
\] (6.19)

By Assumption B, (2.1), (3.3), (4.8), the first term on the right-hand side of (6.19) is bounded above by

\[
\lambda 3\delta^{-\alpha} \lambda m \mathbb{P}(\tau_0 > s_m) 
\] (6.20)

for all \( \varepsilon \) small enough, where \( m = \rho_{\nu_\varepsilon} \). By the definition of \( s_m \) (see (4.9)), we may replace \( m \mathbb{P}(\tau_0 > s_m) \) by 1 in (6.20).

Putting this, Assumptions A and C together, we conclude that \( \mathbb{P}(A_{k,\ell}) \to 0 \) as \( \varepsilon \to 0 \) for all \( k, \ell \), and thus

\[
\mathbb{P}(A_K) \to 0 
\] (6.21)

as \( \varepsilon \to 0 \) for every \( K \geq 1 \).

We now go back to (6.14). By the above, the first term on its right can be written as

\[
\frac{1}{K} \sum_{k=1}^{K} F(Y^{(k,\varepsilon)}) + \frac{1}{K} \sum_{k=1}^{K} \Delta^{(k,\varepsilon,\delta)} + 1 \{ A_K \cup B_K \} \frac{1}{K} \sum_{k=1}^{K} (F(Y^{(k,N,\varepsilon,\delta)}) - F(Y^{(k,\varepsilon,\delta)})),
\] (6.22)

where \( \Delta^{(k,\varepsilon,\delta)} = F(Y^{(k,\varepsilon,\delta)}) - F(Y^{(k,\varepsilon)}) \) and \( B_K = \bigcup_{k=1}^{K} \{ \{ C_{\nu_\varepsilon}^{(k)} \leq \varepsilon^{-1}T \} \cup \{ C_{\nu_\varepsilon}^{(k)} \leq \varepsilon^{-1}T \} \} \).

This follows from the fact that outside \( A_K \cup B_K \), we have that \( Y_t^{(k,\varepsilon,\delta)} = Y_t^{(k,N,\varepsilon,\delta)} \) for \( t \in [0, T] \) and \( 1 \leq k \leq K \), and all fixed \( \varepsilon, \delta, N \), as can be readily checked.

As in (6.15), we have that

\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \Delta^{(k,\varepsilon,\delta)} = 0 
\] (6.23)

in probability for all \( k \).

Now from (6.14), (6.22), and since \( F \in B_u \), we obtain

\[
|\mathbb{E}[F(Y^{(1,N,\varepsilon)})|\tilde{\tau}(N)] - \mathbb{E}[F(Z)]|
\leq \left| \frac{1}{K} \sum_{k=1}^{K} (F(Y^{(k,N,\varepsilon)}) - \mathbb{E}[F(Y^{(k,N,\varepsilon)})|\tilde{\tau}(N)]) \right| + \left| \frac{1}{K} \sum_{k=1}^{K} (F(Y^{(k,\varepsilon)}) - \mathbb{E}[F(Y^{(k,\varepsilon)})]) \right|
\] (6.24)

plus a term whose expectation is bounded above by

\[
|\mathbb{E}[F(Y^{(1,N,\varepsilon)}) - \mathbb{E}(F(Z))]| + |\Delta^{(1,N,\varepsilon,\delta)}| + |\Delta^{(1,\varepsilon,\delta)}| + 2\|F\|_\infty (\mathbb{P}(A_K) + \mathbb{P}(B_K)),
\] (6.25)

where we have used the fact that both \( \mathbb{E}[F(Y^{(k,\varepsilon)})] \) and \( \mathbb{E}[F(Y^{(k,N,\varepsilon)})|\tilde{\tau}(N)] \) are independent of \( k \).

Recalling now Remark 38, and using Jensen, we find that the right-hand side of (6.24) is bounded above by constant times \( K^{-1/2} \). This and (6.6), (6.15), (6.16), (6.21), (6.23) yield

\[
\limsup_{\varepsilon \to 0} \mathbb{E}[|F(Y^{(1,N,\varepsilon)})|\tilde{\tau}(N)] - \mathbb{E}[F(Z)] \leq \text{const} \left( K^{-1/2} + \eta \right). 
\] (6.26)

Since \( K \) and \( \eta \) are arbitrary and the left-hand side of (6.26) does not depend on either, we conclude that

\[
\mathbb{E}[F(Y^{(1,N,\varepsilon)})|\tilde{\tau}(N)] \to \mathbb{E}[F(Z)] 
\] (6.27)

in probability as \( \varepsilon \to 0 \), and the result follows from the fact that \( \mathbb{E}[F(Y^{(1,N,\varepsilon)})|\tilde{\tau}(N)] \) and \( \mathbb{E}[(Y^{(\varepsilon)})|\tau] \) have the same distribution for every \( \varepsilon > 0, N \geq 1 \). \( \square \)
6.1. Stronger ageing results

The above can be extended to strengthen the ageing results of the previous section, under the same conditions of both this and that sections.

**Theorem 39.** If $X$ is transient and Assumption C holds, then we have that

\[
\lim_{t \to \infty} \mathbb{P}(Y_t = Y_{t+\theta t} | \tau) = R(\theta),
\]

\[
\lim_{t \to \infty} \mathbb{P}(Y_t = Y_{t+r} \text{ for all } r \in [0, \theta t/q_1/t]) = \Pi(\theta),
\]

in probability as $t \to \infty$, where $R$ and $\Pi$ are as in Theorem 25 (and indeed, both equal the right-hand side of (5.24)).

**Sketch of proof.** An argument like that of Lemma 16 can be used to get that

\[
(Y_t^{(1,\varepsilon,\delta)}) \longrightarrow (Z_t^{(\delta)})
\]

as $\varepsilon \to 0$ in distribution on $(D, J_1)$ (in here, differently from Lemma 16 case, $\tau$ is integrated; the argument might of course use a version of $\tau$ such that (6.30) holds for $\tau$ in a set of full probability). We may then extend Theorem 35 with a similar proof to get

\[
E[G(Y_t^{(1,\varepsilon)}) | \tau] \longrightarrow E[G(Z^{(\delta)})],
\]

in probability as $\varepsilon \to 0$, for every $\delta > 0$, with $G : D \to \{0, 1\}$ such that either

\[
G(U) = 1\{U_1 = U_{1+r} \text{ for all } r \in [0, \theta]\} \text{ or } G(U) = 1\{U_1 = U_{1+\theta}\}, \quad U \in D.
\]

We further need to extend Lemma 29 to get that for all $t > 0$

\[
\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \sup \mathbb{P}(Y_t^{(1,\varepsilon)} \neq Y_t^{(1,\varepsilon,\delta)}) = 0
\]

(6.33)

(here the proof can be made again by using special versions of $\tau$ like in the argument for Lemma 29).

From (6.31) and (6.33), and since $E[G(Z^{(\delta)})] \to E[G(Z)]$ as $\delta \to 0$, we get

\[
\lim_{\varepsilon \to 0} E[G(Y_t^{(1,\varepsilon)}) | \tau] = \lim_{\delta \to 0} \lim_{\varepsilon \to 0} E[G(Y_t^{(1,\varepsilon,\delta)}) | \tau] = \lim_{\delta \to 0} E[G(Z^{(\delta)}) | \tau] = E[G(Z)],
\]

in probability as $\varepsilon \to 0$, for $G$ as in (6.32).

**Remark 40.** Under the same conditions of Theorem 39, and by the same reasoning as above, we have that

\[
\lim_{t \to \infty} \mathbb{P}(\sup_{r \in [0,t]} Y_r < \sup_{r \in [0,t+\theta t]} Y_r | \tau) = \Omega(\theta)
\]

in probability. (See Remark 32.)

**Remark 41.** With the extra condition of this section, arguing as in Subsection 5.1, with the help of Theorem 35, we may establish stronger results for integrated ageing functions. We state the following result as an example, in the spirit of Subsection 5.1.

**Theorem 42.** If $X$ satisfies Assumptions A–C, then

\[
\lim_{\lambda, \mu \to \infty} E[R(\lambda T, \mu T) | \tau] = R(\theta)
\]

in probability, with $R$ as in Theorem 25.
Acknowledgements. L. R. G. Fontes would like to thank Marina Vachkovskaia for discussions at an early stage of this project. The authors thank an anonymous referee of earlier versions of this paper for comments which led to substantial corrections and other improvements. The work of the first author is part of USP project MaCLinC and FAPESP project NeuroMat. He would like to thank the CMI, Université de Provence, Aix-Marseille I for hospitality and support during several visits in the last few years where this and related projects were developed.

References

1. M. Barlow and J. Černý, ‘Convergence to fractional kinetics for random walks associated with unbounded conductances’, Probab. Theory Related Fields 149 (2011) 639–673.
2. G. Ben Arous, L. Bogatchev and S. Molchanov, ‘Limit theorems for sums of random exponentials’, Probab. Theory Related Fields 132 (2005) 579–612.
3. G. Ben Arous, A. Bovier and J. Černý, ‘Universality of the REM for dynamics of mean-field spin glasses’, Comm. Math. Phys. 282 (2008) 663–695.
4. G. Ben Arous, A. Bovier and V. Gayrard, ‘Glauber dynamics of the random energy model. I. Metastable motion on the extreme states’, Comm. Math. Phys. 235 (2003) 379–425.
5. G. Ben Arous, A. Bovier and V. Gayrard, ‘Glauber dynamics of the random energy model. II. Aging below the critical temperature’, Comm. Math. Phys. 236 (2003) 1–54.
6. G. Ben Arous and J. Černý, ‘Bouchaud’s model exhibits two different aging regimes in dimension one’, Ann. Appl. Probab. 15 (2005) 1161–1192.
7. G. Ben Arous and J. Černý, ‘Scaling limit for trap models on \( \mathbb{Z}^d \)’, Ann. Probab. 35 (2007) 2356–2384.
8. G. Ben Arous and J. Černý, ‘The arcsine law as a universal aging scheme for trap models’, Comm. Pure Appl. Math. 61 (2008) 289–329.
9. G. Ben Arous, J. Černý and T. Mountford, ‘Aging in two-dimensional Bouchaud’s model’, Probab. Theory Related Fields 134 (2006) 1–43.
10. G. Ben Arous, A. Fribergh, N. Gantert and H. Hammond, ‘Biased random walks on a Galton–Watson tree with leaves’, Ann. Probab. 40 (2012) 280–338.
11. J. Bertoin, Subordinators: examples and applications, Lecture Notes in Mathematics 1717 (Springer, Berlin, 1999) 1–91.
12. J.-P. Bouchaud, ‘Weak ergodicity breaking and aging in disordered systems’, J. Phys. I France 2 (1992) 1705–1713.
13. J.-P. Bouchaud, L. Cugliandolo, J. Kurchan and M. Mézard, ‘Out of equilibrium dynamics in spin-glasses and other glassy systems’, Spin-glasses and random fields (ed. A. P. Young; World Scientific, Singapore, 1998).
14. J.-P. Bouchaud and D. S. Dean, ‘Aging on Parisi’s tree’, J. Phys. I France 5 (1995) 265–286.
15. A. Bovier and A. Faggionato, ‘Spectral characterisation of ageing: the REM-like trap model in the complete graph’, Ann. Appl. Probab. 15 (2005) 1997–2037.
16. J. Černý, PhD Thesis, Ecole Polytechnique Fédérale de Lausanne, 2003.
17. J. Černý, ‘On two-dimensional random walk among heavy-tailed conductances’, Electron. J. Probab. 16 (2011) 293–313.
18. L. R. G. Fontes, M. Isopi and C. M. Newman, ‘Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension’, Ann. Probab. 30 (2002) 579–604.
19. L. R. G. Fontes and P. H. S. Lima, ‘Convergence of symmetric trap models in the hypercube’, XVth International Congress on Mathematical Physics, 2006, Rio de Janeiro. New Trends in Mathematical Physics (Springer, Berlin, 2009) 285–297.
20. L. R. G. Fontes and P. Mathieu, ‘K-processes, scaling limit and aging for the trap model in the complete graph’, Ann. Probab. 36 (2008) 1322–1358.
21. N. C. Jain and W. E. Pruitt, ‘The range of recurrent random walk in the plane’, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 16 (1970) 279–292.
22. N. C. Jain and W. E. Pruitt, ‘The range of random walk’, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability III: Probability theory (University of California Press, Berkeley, 1972) 31–50.
23. J.-F. Le Gall, ‘Propriétés d’intersection des marches aléatoires. I. Convergence vers le temps local d’intersection’, Comm. Math. Phys. 104 (1986) 471–507.
24. J.-F. Le Gall and J. Rosen, ‘The range of stable random walks’, Ann. Probab. 19 (1991) 650–705.
25. J.-C. Mourrat, ‘Scaling limit of the random walk among random traps on \( \mathbb{Z}^d \)’, Ann. Inst. H. Poincaré Probab. Statist. 47 (2011) 813–849.
26. E. Seneta, Regularly varying functions. Lecture Notes in Mathematics 508 (Springer, Berlin, 1976).
27. F. Spitzer, Principles of random walks, 2nd edn (Springer, Berlin, 1976).
L. R. G. Fontes
IME-USP, Rua do Matão 1010
05508-090 São Paulo, SP
Brazil
lrenato@ime.usp.br

P. Mathieu
Aix Marseille Université, CNRS, Centrale Marseille
LATP, UMR 7353
13453 Marseille
France
pierre.mathieu@cmi.univ-mrs.fr