Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension

L. R. G. Fontes M. Isopi C. M. Newman

September 6, 2000

Abstract

Let $\tau = (\tau_i : i \in \mathbf{Z})$ denote i.i.d. positive random variables with common distribution F and (conditional on τ) let $X = (X_t : t \geq 0, X_0 = 0)$, be a continuous-time simple symmetric random walk on $\mathbf{Z}$ with inhomogeneous rates $(\tau_i^{-1} : i \in \mathbf{Z})$. When F is in the domain of attraction of a stable law of exponent $\alpha < 1$ (so that $\mathbf{E}(\tau_i) = \infty$ and X is subdiffusive), we prove that (X, τ), suitably rescaled (in space and time), converges to a natural (singular) diffusion $Z = (Z_t : t \geq 0, Z_0 = 0)$ with a random (discrete) speed measure ρ. The convergence is such that the "amount of localization", $\mathbf{E}\sum_{i\in\mathbf{Z}}[\mathbf{P}(X_t = i|\tau)]^2$ converges as $t \to \infty$ to $\mathbf{E}\sum_{z\in\mathbf{R}}[\mathbf{P}(Z_s = z|\rho)]^2 > 0$, which is independent of $s > 0$ because of scaling/self-similarity properties of (Z, ρ). The scaling properties of (Z, ρ) are also closely related to the "aging" of (X, τ). Our main technical result is a general convergence criterion for localization and aging functionals of diffusions/walks $Y^{(\epsilon)}$ with (nonrandom) speed measures $\mu^{(\epsilon)} \to \mu$ (in a sufficiently strong sense).

Mathematics Subject Classification 2000: Primary—60K37, 82C44, 60G18. Secondary—60F17.
Key words and phrases: aging, localization, quasidiffusions, disordered systems, scaling limits, random walks in random environments, self-similarity.

1 Introduction

In this paper we continue the study of localization in the one-dimensional Random Walk with Random Rates (RWRR), begun in [1] (or equivalently of chaotic time dependence in the related Voter Model with Random Rates (VMRR)—see below and [1]). We also relate localization to "aging", a phenomenon of considerable interest in out-of-equilibrium physical systems, such as glasses (see, e.g., [2] for a review).

The RWRR is a continuous-time simple symmetric random walk on $\mathbf{Z}^d$, $X = (X_t : t \geq 0, X_0 = 0)$, where the time spent at site i before taking a step has an exponential distribution of mean τ_i, and where the τ_i's are i.i.d. positive random variables with common distribution F; thus it is a random walk in the random environment, $\tau = (\tau_i : i \in \mathbf{Z}^d)$.

When F has a finite mean, it can be shown (e.g., by the convergence results of [3], as discussed below) that (for a.e. τ) there is a central limit theorem for X_t, and more generally an invariance principle, i.e., that $\epsilon X_{t/\epsilon^2}$ converges to a Brownian motion as $\epsilon \to 0$. On the other hand, when F has infinite mean with a power law tail of exponent $\alpha < 1$, one expects power law subdiffusive behavior (with an exponent depending on both α and d); for reviews of the physics literature on subdiffusivity in random environments, see, e.g., [4, 5]. Logarithmic subdiffusivity [6], as occurs in other commonly studied random walks in random environments [7], would presumably occur in an RWRR if the tail of F were itself logarithmic, but the more natural context for an RWRR is a power law tail for F.

More striking than subdiffusivity, and the main result of [1], is that for $\alpha < 1$, $d = 1$, there is localization in the sense that (for a.e. τ) as $t \to \infty$,

$$\sup_{i \in \mathbf{Z}} \mathbf{P}(X_t = i|\tau) \not\to 0 \tag{1.1}$$

or equivalently

$$\sum_{i \in \mathbf{Z}} [\mathbf{P}(X_t = i|\tau)]^2 \not\to 0. \tag{1.2}$$

An essential purpose of this paper is to relate this localization to an appropriate scaling limit of X, in which it turns out that Brownian motion is replaced by a singular diffusion Z (in a random environment) — singular here meaning that the single time distributions of Z are discrete. We remark that there is also localization in the random walks of [7, 6], as shown by Golosov [8], but both the localization and scaling limits are of a somewhat different character there (as one would expect in cases of logarithmic subdiffusivity).

Kawazu and Kesten [9] treated the similar problem of finding the scaling limit of a random walk with i.i.d. random bond rates λ_i (for transitions from i to $i+1$ and from $i+1$ to i). Their random walk is in fact also related to the VMRR and hence to our RWRR, with $\lambda_i = 1/(2\tau_i)$. The scaling limit of [9] (see also [10, 11]) for $\alpha < 1$, obtained by a similar approach based on [3] as the one used here, is also a diffusion, but one that is nonsingular in the sense that the single time distributions are continuous. Our analysis of the type of localization exhibited in (1.1)–(1.2) (i.e., at *individual* points) requires a stronger type of convergence to the scaling limit than was needed in [9], as we explain later.

A convenient quantity, with which to express the relation between localization and the scaling limit, is the "amount of localization" at time t, as measured by

$$q_t = \mathbf{E} \sum_{i \in \mathbf{Z}} [\mathbf{P}(X_t = i|\tau)]^2, \tag{1.3}$$

where the expectation is with respect to τ. A main result of this paper (immediate from Theorem (3.2)) is that as $t \to \infty$, q_t converges to a (nonrandom) $q_\infty \in (0, 1)$ (depending on $\alpha \in (0, 1)$), which can itself be expressed by a formula (see (1.9) below) analogous to (1.3) with the singular diffusion Z replacing the random walk X.

Our analysis of the scaling limit of (X, τ) will also yield results about aging of the RWRR. As in the extensive physics literature on the subject (see, e.g., [2] and the references therein), we will consider a quantity $R(t_w + t, t_w)$ that measures the behavior of the system at a time

t_w+t, after it has been aged for time t_w. Normal aging corresponds to there being a well-defined nontrivial limit function when t and t_w are scaled proportionally:

$$\mathcal{R}(\theta) \;=\; \lim_{t_w\to\infty; t/t_w\to\theta} R(t_w+t,t_w). \tag{1.4}$$

One interesting example of an R for which such a limit follows from our results is $R(t_w+t,t) = q_t(t_w)$, where

$$q_t(t_w) \;=\; \mathbf{E}\sum_{i\in\mathbf{Z}}[\mathbf{P}(X_{t_w+t}=i|\tau,X_{t_w})]^2. \tag{1.5}$$

Of course, $q_t(0) = q_t$, corresponding to the amount of localization after time t, starting from a fresh ($t_w = 0$) system with $X_0 = 0$ that has not been aged. As with q_∞, the limit function $\mathcal{R}(\theta)$ will be given by a formula (see (1.10)) like (1.5), but with X replaced by the diffusion Z. It follows from (1.10) that $\mathcal{R}(\theta)$ tends to 1 as $\theta\to 0$ and to q_∞ as $\theta\to\infty$. Other examples of RWRR quantities that exhibit normal aging are the (unconditional) probabilities $\mathbf{P}(X_{t_w+t}=X_{t_w})$, which we discuss below, and

$$\mathbf{P}(\max_{t_w\le t'\le t_w+t}\tau_{X_{t'}} > \max_{0\le t'\le t_w}\tau_{X_{t'}}), \tag{1.6}$$

which measures the prospects for "novelty" in this aging system.

Before explaining more about Z and it's random environment, we make a short digression to point out that q_∞ is a natural object of study also for the related VMRR (as it is for other similar spin systems with stochastic dynamics).

The one-dimensional (linear) VMRR is the continuous-time Markov process σ_t with state space $\{\sigma(i) : i\in\mathbf{Z}\} = \{-1,+1\}^{\mathbf{Z}}$ in which, at rate $1/\tau_i$, site i chooses (with equal probability) one of it's two neighbors (say i') and replaces $\sigma(i)$ with $\sigma(i')$. The initial state σ_0 is taken to be $\xi = (\xi_i : i\in\mathbf{Z})$, with the ξ_i's i.i.d. and equally likely to be $+1$ or -1. Chaotic Time Dependence (CTD) is said to occur if (conditional on (ξ,τ)) the distribution of σ_t has multiple subsequence limits as $t\to\infty$. (For a discussion of the possible occurence of CTD in other more physical spin systems, see [1, 12].) Since the alternative to CTD for this VMRR would be for the distribution to converge to the symmetric mixture of the degenerate measures on the constant (identically $+1$ or identically -1) states, CTD is equivalent to the existence of some predictability about the state for some arbitrarily large times, based on complete knowledge of the inital state (and the environment of rates). In [1], CTD is proved to occur for a fat-tailed F (with $\alpha<1$) by showing that (for a.e. (ξ,τ) and every k) $\mathbf{E}[\sigma_t(k)|\xi,\tau]$ does not converge as $t\to\infty$, whereas the absence of CTD would require convergence to zero. A natural quantity measuring the amount of CTD/predictability (see, e.g., [13]) is thus

$$\lim_{t\to\infty}\lim_{L\to\infty}(2L+1)^{-1}\sum_{k=-L}^{k=L}\mathbf{E}^2[\sigma_t(k)|\xi,\tau] = \lim_{t\to\infty}\mathbf{E}\{\mathbf{E}^2[\sigma_t(0)|\xi,\tau]\}. \tag{1.7}$$

But by the standard fact that a time-reversed voter model corresponds to coalescing random walks, it easily follows, by doing the outermost expectation first over ξ and then over τ, that

$$\mathbf{E}\{\mathbf{E}^2[\sigma_t(0)|\xi,\tau]\} = \mathbf{E}(\mathbf{E}\left\{\mathbf{E}^2[\sigma_t(0)|\tau,\xi]\middle|\tau\right\}) = \mathbf{E}\sum_{i\in\mathbf{Z}}[\mathbf{P}(X_t=i|\tau)]^2 = q_t. \tag{1.8}$$

Thus, in the VMRR, the natural dynamical order parameter for CTD is just q_∞.

Of course, it should be noted, that the existence of the $t \to \infty$ limit in (1.7) is not at all obvious—especially in view of CTD. (The $L \to \infty$ limit is a consequence of the spatial ergodicity of (ξ, τ).) Indeed, we prove its existence by expressing the $t \to \infty$ limit of (1.8) in terms of a scaling limit of (X, τ), i.e., by showing that as $t \to \infty$,

$$q_t \to \mathbf{E} \sum_{z \in \mathbf{R}} [\mathbf{P}(Z_s = z|\rho)]^2 > 0, \tag{1.9}$$

where (Z, ρ) is a (singular) one-dimensional diffusion Z in a random environment ρ. Here $s > 0$ is arbitrary, and by the singularity of Z, we mean that (conditional on ρ) the distribution of Z_s is discrete, even though Z is a bona-fide diffusion with continuous sample paths. We shall see why the above expression for q_∞, which describes the amount of localization of (Z, ρ) at time s does not in fact depend on s (as long as $s \neq 0$), a fact that may at first seem surprising (since $Z_s \to 0$ as $s \to 0$, almost surely). Indeed this lack of dependence follows from the scaling/self-similarity properties of (Z, ρ) which imply that (conditioned on ρ) the distribution of $s^{\alpha/(\alpha+1)} Z_s$ is a *random* measure on $\mathbf{R}$ whose distribution (arising from its dependence on ρ) does not depend on $s > 0$.

Analogously to (1.9), we have $\mathcal{R}(\theta)$ of (1.4)–(1.5) given by

$$\lim_{t' \to \infty} q_{\theta t'}(t') \;=\; \mathbf{E} \sum_{z \in \mathbf{R}} [\mathbf{P}(Z_{s+\theta s} = z|\rho, Z_s)]^2. \tag{1.10}$$

The validity of this limit also follows from the results and techniques of Sections 2 and 3 of the paper — see Remark 2.1. Here, the self-similarity properties of (Z, ρ) imply that the RHS of (1.10) depends only on θ and *not* on s (for $0 < s < \infty$), explaining the basic signature of normal aging — that the asymptotics of $q_t(t_w)$ depend only on the asymptotic ratio of t/t_w.

Another example of an RWRR localization quantity with normal aging behavior is

$$q'_t(t_w) \;=\; \mathbf{P}(X_{t_w+t} = X_{t_w}) \;=\; \mathbf{E}\,\mathbf{P}(X_{t_w+t} = X_{t_w}|\tau, X_{t_w}). \tag{1.11}$$

In this case, the asymptotic aging function, $\mathcal{R}'(\theta)$, would have limits of 1 and 0 respectively as $\theta \to 0$ and ∞. Interestingly, a related quantity,

$$q^*_t(t_w) \;=\; \mathbf{P}(X_{t_w+t'} = X_{t_w} \,\forall\, t' \in [0, t]), \tag{1.12}$$

exhibits what is known as "subaging" (see, e.g., [14], where a one-parameter family of models extending the RWRR are studied nonrigorously, for general d). I.e. (assuming, for simplicity, that the tail of F satisfies $u^\alpha \mathbf{P}(\tau_i > u) \to K \in (0, \infty)$), there is a nontrivial limit when $t/(t_w)^\eta \to \theta$ as $t_w \to \infty$, for some $0 < \eta < 1$; here $\eta = 1/(1+\alpha)$ (for $0 < \alpha < 1$). The difference in behavior between q' and q^* is due to the fact that during the time interval $[t_w, t_w + \theta t_w]$, each visit of the random walk to X_{t_w} takes an amount of time of order $t_w^{1/(1+\alpha)}$, but there are of order $t_w^{\alpha/(1+\alpha)}$ visits. A related fact, in the scaling limit, is that for $s, s' > 0$, the diffusion process Z has (for a.e. ρ)

$$\mathbf{P}(Z_{s+s'} = Z_s \,|\, \rho) \;>\; 0 \quad \text{but} \quad \mathbf{P}(Z_{s+s''} = Z_s \,\forall\, s'' \in [0, s'] \,|\, \rho) \;=\; 0. \tag{1.13}$$

This existence of different scaling regimes for *different* quantities in a single model may be compared and contrasted to the search for multiple scaling regimes in the *same* quantity (see, e.g., [14]), where $R(t_w + \theta t_w^{\eta}, t_w)$ and $R(t_w + \theta' t_w^{\eta'}, t_w)$ with $\eta \neq \eta'$ would both have nontrivial limits. (In fact, something weaker than this is claimed in [14] for the $q_t^*(t_w)$ of (1.12).)

What is this diffusion in a random environment, (Z, ρ)? The random environment ρ, the spatial scaling limit of the original environment τ of rates on $\mathbf{Z}$, is a random discrete measure, $\sum_i W_i \delta_{Y_i}$, where the countable collection of (Y_i, W_i)'s yields an inhomogeneous Poisson point process on $\mathbf{R} \times (0, \infty)$ with density measure $dy\, \alpha w^{-1-\alpha} dw$. Note that although ρ is discrete, the set of Y_i's is a.s. dense in $\mathbf{R}$ because the density measure is non-integrable at $w = 0$. Conditional on ρ, Z_s is a diffusion process (with $Z_0 = 0$) that can be expressed as a time change of a standard Brownian motion $B(t)$ with speed measure ρ, as follows [15].

Letting $\ell(t, x)$ denote the local time at x of $B(t)$, define

$$\phi_t^\rho := \int \ell(t, y)\, d\rho(y) \tag{1.14}$$

and the stopping time ψ_s^ρ as the first time t when $\phi_t^\rho = s$ (so that ψ^ρ is the inverse function of ϕ^ρ); then $Z_s = B(\psi_s^\rho)$. For (a deterministic) $s > 0$, the distribution of Z_s is a discrete measure whose atoms are precisely those of ρ; this is essentially because the set of times when Z is anywhere else than these atoms has zero Lebesgue measure.

To see the lack of dependence of the RHS's of (1.9) and (1.10) on s, we may proceed as follows. For $\lambda > 0$, consider the rescaled Brownian motion and environment,

$$B^\lambda(t) = \lambda^{-1/2} B(\lambda t); \quad \rho^\lambda = \sum_i (\lambda^{-1/2})^{1/\alpha} W_i \delta_{\lambda^{-1/2} Y_i}. \tag{1.15}$$

Since B^λ and ρ^λ are equidistributed with B and ρ, it follows that if we define a diffusion Z^λ as the time-changed B^λ using speed measure ρ^λ, then $(Z^\lambda, \rho^\lambda)$ is equidistributed with the original diffusion in a random environment (Z, ρ). On the other hand, on the original probability space on which B and ρ are defined, one has $Z_s^\lambda = \lambda^{-1/2} Z_{\lambda^{(\alpha+1)/(2\alpha)} s}$, so that the RHS's of (1.9) and (1.10) remain the same when s is replaced by $\lambda^{(\alpha+1)/(2\alpha)} s$, and thus cannot depend on s.

To best understand how (Z, ρ) arises as the scaling limit of (X, τ), one should use the fact that not only diffusions, but also random walks (or more accurately, birth-death processes) can be expressed as time-changed Brownian motions [3, 15]. In particular, if for any $\epsilon > 0$, we take as speed measure

$$\rho^{(\epsilon)} := \sum_{i \in \mathbf{Z}} c_\epsilon \tau_i \delta_{\epsilon i}, \tag{1.16}$$

where the parameter $c_\epsilon > 0$ is yet to be determined, and then do the time-change on the rescaled Brownian motion B^{1/ϵ^2}, the resulting process is a rescaling of the original random walk X, namely $Z_s^{(\epsilon)} = \epsilon X_{s/(c_\epsilon \epsilon)}$. When the distribution F of the τ_i's has a finite mean, then by the Law of Large Numbers, taking $c_\epsilon = \epsilon$, $\rho^{(\epsilon)}$ converges to (the mean of F times) Lebesgue measure and $Z^{(\epsilon)}$ converges to a Brownian motion as $\epsilon \to 0$ [3, 16]. On the other hand, if $1 - F(u) = L(u)/u^\alpha$ with $\alpha < 1$ and $L(u)$ slowly varying as $u \to \infty$, then by choosing c_ϵ appropriately (as $\epsilon^{1/\alpha}$ times another slowly varying function as $\epsilon \to 0$ — see (3.8) below) one has (from the classical theories of domains of attraction and extreme value statistics) convergence (in various senses, to be discussed) of $\rho^{(\epsilon)}$ to the random measure ρ.

The idea that there should also follow some kind of convergence of $(Z^{(\epsilon)}, \rho^{(\epsilon)})$ to (Z, ρ) should by now be quite clear. And indeed the basic convergence results of [3] are enough to imply, for example, that a functional like

$$\mathbf{E}\{[P(a \leq Z^{(\epsilon)}(s) \leq b \,|\, \rho^{(\epsilon)})]^2\} \tag{1.17}$$

(for *deterministic* a, b) converges to the corresponding quantity for (Z, ρ). But they are not sufficient to get localization quantities like $q_{s/(c_\epsilon \epsilon)} = \mathbf{E} \sum_{z \in \mathbf{R}} [P(Z^{(\epsilon)}(s) = z \mid \rho^{(\epsilon)})]^2$ to converge. As mentioned earlier, the work of [9] was also based on the time-changed Brownian motion approach of [3, 15], but for their random walk and scaling limit, the convergence results of [3] are sufficient.

The problem in our case is not primarily with the randomness of $\rho^{(\epsilon)}$ (i.e., of τ) and ρ, but occurs already when considering the nature of convergence of a process $Y^{(\epsilon)}(t)$ that is a Brownian motion time-changed with a *deterministic* speed measure $\mu^{(\epsilon)}$. The convergence results of [3] imply that if $\mu^{(\epsilon)} \to \mu$ vaguely, then (for example) one has weak convergence of the distribution $\bar{\mu}^{(\epsilon)}$ of $Y^{(\epsilon)}(t_0)$ to the corresponding $\bar{\mu}$. But we need stronger convergence.

This stronger convergence is the subject of Section 2, which contains the main technical result of the paper, in which weak convergence is combined with "point process" convergence. By point process convergence for (say) a discrete measure $\sum_i w_i^{(\epsilon)} \delta_{y_i^{(\epsilon)}}$ to $\sum_i w_i \delta_{y_i}$ (where we have expressed each sum so that the atoms are not repeated), we mean that the subset of $\mathbf{R} \times (0, \infty)$ consisting of all the $(y_i^{(\epsilon)}, w_i^{(\epsilon)})$'s converges to the set of all (y_i, w_i)'s — in the sense that every open disk (whose closure is a compact subset of $\mathbf{R} \times (0, \infty)$) containing exactly m of the (y_i, w_i)'s ($m = 0, 1, \ldots$) with none on its boundary, contains also exactly m of the $(y_i^{(\epsilon)}, w_i^{(\epsilon)})$'s for all small ϵ. Our technical result is that vague plus point process convergence for the speed measures $\mu^{(\epsilon)} \to \mu$ implies the same for the distributions at a fixed time t_0; i.e., $\bar{\mu}^{(\epsilon)} \to \bar{\mu}$.

Going from this result for a sequence of deterministic speed measures to our context of random speed measures requires a bit more work, which is presented in Section 3 of the paper. The way we handle that, which may be of independent interest, is to replace the the random measures $\rho^{(\epsilon)}$ which only converge (in our two senses) in distribution, by a different (but also natural) coupling for the various ϵ's than that provided by the space of the original τ_i's so that convergence becomes almost sure. We note that almost sure convergence was also obtained in the scaling limit results of [9] by means of a coupling argument, but there the coupling was an abstract one. In our situation, because of the need to handle point process convergence, a concrete coupling seems more suitable, in addition to being more natural.

We close the introduction by noting that we have restricted attention to the scaling limit of a single RWRR. In the context of the the VMRR, which originally led to our interest in localization, one should consider the scaling limit of coalescing RWRR's. Furthermore, one should also study the scaling limit of the VMRR directly. These issues will be taken up in future papers.

2 The continuity theorem

Let μ, $\mu^{(\epsilon)}$, $\epsilon > 0$, be non-identically-zero, locally finite measures on $\mathbf{R}$ and let μ_d, $\mu_d^{(\epsilon)}$, $\epsilon > 0$, be their discrete parts. Let Y_t, $Y_t^{(\epsilon)}$, $t \geq 0$, $\epsilon > 0$, $Y_0^{(\epsilon)} = Y_0 \equiv x$, be the Markov processes in one dimension obtained by time changing a standard Brownian motion through $\mu^{(\epsilon)}$, i.e., let $B = B(s)$, $s \geq 0$, be a standard Brownian motion (with $B(0) = 0$) and let

$$\phi_s(x) := \int \ell(s, y - x)\, d\mu(y), \quad \psi(x) = \psi_t(x) := \phi_t^{-1}(x), \quad Y_t = B(\psi_t(x)) + x; \tag{2.1}$$

$$\phi_s^{(\epsilon)}(x) := \int \ell(s, y-x)\, d\mu^{(\epsilon)}(y), \quad \psi^{(\epsilon)}(x) = \psi_t^{(\epsilon)}(x) := (\phi_t^{(\epsilon)})^{-1}(x), \quad Y_t^{(\epsilon)} = B(\psi_t^{(\epsilon)}(x)) + x, \tag{2.2}$$

where ℓ is the Brownian local time of B [3, 15]. Notice that, since $\ell(s,y)$ is nondecreasing in s for all y, $\phi_s(x)$ and $\phi_s^{(\epsilon)}(x)$ are nondecreasing in s and so their (right-continuous) inverses, $\psi_t(x)$ and $\psi_t^{(\epsilon)}(x)$, respectively, are well-defined. Processes described in this way are known in the literature as *quasidiffusions*, *gap diffusions* or *generalized diffusions* ([18, 20] and references therein). They generalize the usual diffusions in that the *speed measures* μ can be zero in intervals, thus including birth and death and other processes.

One fact about those processes we will need below is the following formula from p. 641 of [3]. Let $Y_0 = x$; for any Borel set A of the reals,

$$\int_0^t 1\{Y_s \in A\}\, ds = \int_A \ell_Y(t, x, y)\, d\mu(y) \tag{2.3}$$

almost surely, where $\ell_Y(t,x,y) = \ell(\psi_t(x), y - x)$.

We will use the arrows $\stackrel{v}{\to}$ and $\stackrel{pp}{\to}$ to denote vague and point process convergences, respectively. I.e., given a family ν, $\nu^{(\epsilon)}$, $\epsilon > 0$, of locally finite measures on $\mathbf{R}$ with their discrete parts ν_d, $\nu_d^{(\epsilon)}$, $\epsilon > 0$, we will write $\nu^{(\epsilon)} \stackrel{v}{\to} \nu$ to indicate vague convergence, i.e., that for all continuous real-valued functions f on $\mathbf{R}$ with bounded support $\int f(y)\, d\nu^{(\epsilon)}(y) \to \int f(y)\, d\nu(y)$. For the same family, we write $\nu_d^{(\epsilon)} \stackrel{pp}{\to} \nu_d$ or $\nu^{(\epsilon)} \stackrel{pp}{\to} \nu$ to indicate point process convergence, i.e., that if the atoms of ν_d, $\nu_d^{(\epsilon)}$ are, respectively, at the *distinct* locations y_i, $y_i^{(\epsilon)}$ with weights w_i, $w_i^{(\epsilon)}$, then the subsets $V^{(\epsilon)} \equiv \{(y_i^{(\epsilon)}, w_i^{(\epsilon)})\}$ of $\mathbf{R} \times (0, \infty)$ converge to $V = \{(y_i, w_i)\}$ as $\epsilon \to 0$ in the sense that for any open U whose closure $\bar{U}$ is a compact subset of $\mathbf{R} \times (0,\infty)$ such that its boundary contains no points of V, the number of points in $V^{(\epsilon)} \cap U$ (necessarily finite because U is bounded and at a finite distance from $\mathbf{R} \times \{0\}$) equals the number of points in $V \cap U$ for all small enough ϵ.

We are ready to state the main result of this section; its proof will begin after two corollaries are presented.

Theorem 2.1 *Let $\mu^{(\epsilon)}, \mu, Y^{(\epsilon)}, Y$ be as above and fix any deterministic $t_0 > 0$ and $x \in \mathbf{R}$. Let $\bar{\mu}^{(\epsilon)}$ denote the distribution of $Y_{t_0}^{(\epsilon)}$ (with $Y_0^{(\epsilon)} = x$), let $\bar{\mu}_d^{(\epsilon)}$ denote its discrete part, and define $\bar{\mu}, \bar{\mu}_d$ similarly for Y_{t_0}. Suppose*

$$\mu^{(\epsilon)} \stackrel{v}{\to} \mu \quad \textit{and} \quad \mu_d^{(\epsilon)} \stackrel{pp}{\to} \mu_d \quad \textit{as } \epsilon \to 0. \tag{2.4}$$

Then, as $\epsilon \to 0$,

$$\bar{\mu}^{(\epsilon)} \stackrel{v}{\to} \bar{\mu} \quad \textit{and} \quad \bar{\mu}_d^{(\epsilon)} \stackrel{pp}{\to} \bar{\mu}_d. \tag{2.5}$$

Remark 2.1 *To study limits involving two (or more) times (see, e.g., (1.5), (1.10), (1.11)), some straightforward extensions of Theorem 2.1 are useful. One of these is that (2.5) remains valid if $Y_0^{(\epsilon)} = x^{(\epsilon)}$ with $x^{(\epsilon)} \to x$. Another is that the single-time distribution $\bar{\mu}^{(\epsilon)}$ of $Y_{t_0}^{(\epsilon)}$ can be replaced by the multi-time distribution of $(Y_{t_1}^{(\epsilon)}, \ldots, Y_{t_m}^{(\epsilon)})$, with point process convergence for measures on $\mathbf{R}^m$ defined in the obvious way.*

Corollary 2.1 *Under the same hypotheses, the weights of the atoms of $\bar{\mu}_d^{(\epsilon)}$ and $\bar{\mu}_d$ satisfy*

$$\sum_j [\bar{w}_j^{(\epsilon)}]^2 \to \sum_i [\bar{w}_i]^2 \quad \text{as } \epsilon \to 0. \tag{2.6}$$

Remark 2.2 *More explicitly, (2.6) takes the form*

$$\sum_{y \in \mathbf{R}} [\mathbf{P}(Y_{t_0}^{(\epsilon)} = y)]^2 \to \sum_{y \in \mathbf{R}} [\mathbf{P}(Y_{t_0} = y)]^2 \quad \text{as } \epsilon \to 0, \tag{2.7}$$

or, equivalently, if $Y_t^{(\epsilon)\prime}$ (resp. Y_t') is an independent copy of $Y_t^{(\epsilon)}$ (resp. Y_t), then

$$\mathbf{P}(Y_{t_0}^{(\epsilon)\prime} = Y_{t_0}^{(\epsilon)}) \to \mathbf{P}(Y_{t_0}' = Y_{t_0}) \quad \text{as } \epsilon \to 0. \tag{2.8}$$

Proof of Corollary 2.1.

The convergence $\bar{\mu}_d^{(\epsilon)} \overset{pp}{\to} \bar{\mu}_d$ implies that if we order the $(\bar{y}_i, \bar{w}_i)$'s (the locations and weights of the atoms of $\bar{\mu}_d$) so that $\bar{w}_{i_1} \geq \bar{w}_{i_2} \geq \ldots$, then the following condition holds.

Condition 1 *For each $l \geq 1$, there exists $j_l(\epsilon)$ such that*

$$(\bar{y}_{j_l(\epsilon)}^{(\epsilon)}, \bar{w}_{j_l(\epsilon)}^{(\epsilon)}) \to (\bar{y}_{i_l}, \bar{w}_{i_l}) \quad \text{as } \epsilon \to 0. \tag{2.9}$$

Condition 1 implies that

$$\liminf_{\epsilon \to 0} \sum_j [\bar{w}_j^{(\epsilon)}]^2 \geq \sup_k \sum_{l=1}^k [\bar{w}_{i_l}]^2 = \sum_i [\bar{w}_i]^2. \tag{2.10}$$

This condition, together with the distinctness of the $(\bar{y}_i, \bar{w}_i)$'s, also implies that for any k, the indices $j_1(\epsilon), \ldots, j_k(\epsilon)$ are distinct for small enough ϵ. Furthermore, it implies that if k and δ are such that $\bar{w}_{i_k} > \delta > \bar{w}_{i_{k+1}}$, then for small enough ϵ,

$$\sup\{\bar{w}_j^{(\epsilon)} : j \notin \{j_1(\epsilon), \ldots, j_k(\epsilon)\}\} < \delta. \tag{2.11}$$

To see this, note that otherwise, along some subsequence $\epsilon = \epsilon_l \to 0$, there would be an index $j^*(\epsilon) \notin \{j_1(\epsilon), \ldots, j_k(\epsilon)\}$ with $\liminf \bar{w}_{j^*(\epsilon)}^{(\epsilon)} \geq \delta$ and either (i) $\bar{y}_{j^*(\epsilon)}^{(\epsilon)} \to y^* \in (-\infty, +\infty)$ or else (ii) $|\bar{y}_{j^*(\epsilon)}^{(\epsilon)}| \to \infty$. Case (i) would contradict $\bar{\mu}_d^{(\epsilon)} \overset{pp}{\to} \bar{\mu}_d$, while case (ii) would imply that the family $\{\bar{\mu}^{(\epsilon)}\}$ is not tight, which would contradict $\bar{\mu}^{(\epsilon)} \overset{v}{\to} \bar{\mu}$ since $\bar{\mu}^{(\epsilon)}$ and $\bar{\mu}$ are all probability measures.

Using the above choice of k and δ, we thus have

$$\limsup_{\epsilon \to 0} \sum_j [\bar{w}_j^{(\epsilon)}]^2 \leq \sum_{l=1}^k [\bar{w}_{i_l}]^2 + \limsup_{\epsilon \to 0} \sum_j \delta \bar{w}_j^{(\epsilon)} = \sum_{l=1}^k [\bar{w}_{i_l}]^2 + \delta. \tag{2.12}$$

Letting $k \to \infty$ and $\delta \to 0$ completes the proof. □

Before continuing with the proof of Theorem 2.1, let us note that the above arguments yield another corollary. To state it, we denote by $\mathcal{D}(\nu)$ the mapping from $(0, \infty)$ to $\{0, 1, 2, \ldots, \infty\}$ with $\mathcal{D}(\nu)(w)$ the number of x's in $\mathbf{R}$ such that $\nu(\{x\}) = w$; i.e., $\mathcal{D}(\nu)$ is the set of weights w_i appearing in $V = \{y_i, w_i\}$ counting multiplicity. Of course, when ν is a totally finite measure, $\mathcal{D}(\nu)$ will not take the value ∞. The above arguments show that $\bar{\mu}^{(\epsilon)} \overset{v}{\to} \bar{\mu}$ and $\bar{\mu}_d \overset{pp}{\to} \bar{\mu}_d$ together imply that $\mathcal{D}(\bar{\mu}^{(\epsilon)}) \overset{pp}{\to} \mathcal{D}(\bar{\mu})$, where this latter point process convergence is defined in the obvious way, taking into account multiplicities. Thus we have

Corollary 2.2 *Under the same hypotheses,* $\mathcal{D}(\bar\mu^{(\epsilon)}) \stackrel{pp}{\to} \mathcal{D}(\bar\mu)$.

Proof of Theorem 2.1.

The first assertion in (2.5) is contained in Corollary 1 of [3]. Actually, the latter result is stronger. It states that $\{Y_t^{(\epsilon)},\, t\in[0,T]\}$ converges in distribution to $\{Y_t,\, t\in[0,T]\}$ as a process (in the Skorohod topology), $T>0$ arbitrary. We will indeed use the stronger result in the argument for point process convergence later on. The fixed t_0 case is a rather simple and straightforward consequence of the Brownian representation (2.1)-(2.2), so we next briefly indicate an argument.

Since $\ell(s,y)$ can be taken continuous in (s,y) and of bounded support in y for each s, the first assumption in (2.5) implies that $\phi_s^{(\epsilon)}(x)\to\phi_s(x)$ as $\epsilon\to0$ for all s. It follows that $\psi_t^{(\epsilon)}(x)\to\psi_t(x)$ as $\epsilon\to0$ for all t where $\psi(x)$ is continuous. It suffices now to argue that for any deterministic t, $\psi(x)$ is almost surely continuous at t. For that, notice that $\psi(x)$ is discontinuous at t (if and) only if $\phi(x)$ has a *plateau* at height t, i.e., only if $\phi_{T+s}(x)-\phi_T(x)=0$ for some $s>0$, where $T=\inf\{s'\ge0:\phi_{s'}(x)=t\}$. But, from the definition of $\phi(x)$ and monotonicity of ℓ, that means that

$$\ell(T+s,y-x)-\ell(T,y-x)=0 \quad \text{for } \mu\text{-almost every } y. \tag{2.13}$$

Now, the definition of T implies that $\phi_{T-s'}(x)<t$ for all $s'>0$. This implies that $B(T)=y_0-x$ for some y_0 in the support of μ. But given that, since T is a stopping time, $\ell(T+s,y_0-x)-\ell(T,y_0-x)$ is distributed like $\ell(s,0)$ and thus is strictly positive for all $s>0$. The continuity of ℓ now implies that there exists $\delta>0$ such that $\ell(T+s,y-x)-\ell(T,y-x)>0$ if $|y-y_0|<\delta$, which contradicts (2.13). This settles the first assertion in (2.5).

We claim that the second assertion in (2.5) follows from Condition 1 above (which we have yet to prove) and the first assertion in (2.5). Suppose it does not. According to the definitions above, that means that there exists an open set U_0 whose closure is in $\mathbf{R}\times(0,\infty)$ and a sequence (ϵ_n) tending to 0 as $n\to\infty$ such that $|V^{(\epsilon_n)}\cap U_0|\ne|V\cap U_0|$ for all n. By Condition 1 it must then be that $|V^{(\epsilon_n)}\cap U_0|>|V\cap U_0|$ for all large enough n. That means that either there exist $\hat\imath$, $w^*>0$ and sequences (ϵ'_j), (k_j) and (k'_j), with $\epsilon'_j\to0$ as $j\to\infty$ and $k_j\ne k'_j$ for all j, such that $\bar y_{k_j}^{(\epsilon'_j)},\bar y_{k'_j}^{(\epsilon'_j)}\to\bar y_{\hat\imath}$, $w_{k_j}^{(\epsilon'_j)}\to w_{\hat\imath}$ and $w_{k'_j}^{(\epsilon'_j)}\to w^*$ as $j\to\infty$ or there exist a point $(y^*,w^*)\in\mathbf{R}\times(0,\infty)\backslash V$ and sequences (ϵ'_j) and (k_j), with $\epsilon'_j\to0$ as $j\to\infty$, such that $(\bar y_{k_j}^{(\epsilon'_j)},w_{k_j}^{(\epsilon'_j)})\to(y^*,w^*)$ as $j\to\infty$. In either case we get a contradiction to vague convergence of $\bar\mu^{(\epsilon)}$ to $\bar\mu$ by taking a continuous function $\tilde f$ that approximates sufficiently well the indicator function of either $y_{\hat\imath}$ or y^*, depending on the case, and showing that $\int\tilde f\,d\bar\mu^{(\epsilon_j)}$ is bounded below away from $\int\tilde f\,d\bar\mu$.

It only remains to prove that Condition 1 holds. We will need the following three lemmas.

Lemma 2.1 *The set of locations of the atoms of* $\bar\mu$, $\{\bar y_i\}$, *is contained in that of* μ, $\{y_{i'}\}$.

Proof.

It is a result from the general theory of quasidiffusions [18, 19] that for a process Y' living on a finite interval I (with appropriate boundary conditions), there exists a symmetric continuous transition density $p'_I(t,x,y)$ which is strictly positive and such that

$$\mathbf{P}(Y'_t\in dy|Y'_0=x)=p'_I(t,x,y)\,\mu(dy) \quad \text{for } t>0,\ x,y\in I. \tag{2.14}$$

This would imply the result immediately if our process Y were such a finite interval process, but it is not. However, if we condition on its history being contained within a fixed interval, then we can use (2.14). The details are as follows.

Let $A_{t,l} = \{Y_s \in [-l,l] \text{ for } s \le t\}$, where $l > |x|$, $t \ge 0$. Then, on $A_{t,l}$, $\{Y_s, s \le t\}$ is distributed like $\{Y'_s, s \le t\}$ on the analogous $A'_{t,l}$, where Y' is the diffusion with speed measure $\mu' := \mu|_{(-l-1,l+1)}$ (and boundary conditions at $\pm(l+1)$ as in [18]). More precisely, there is a coupling between Y, Y' and a third process Y'' with speed measure μ' and *killing* boundary conditions at $\pm(l+1)$, such that $\{Y_s, s \le t\} = \{Y'_s, s \le t\}$ on $A''_{t,l}$. Thus

$$\mathbf{P}(Y_t = y_0 | Y_0 = x) - \epsilon_{t,l} = p'_I(t,x,y_0)\,\mu'(y_0) - \epsilon'_{t,l}, \tag{2.15}$$

where $I = [-l-1, l+1]$ and $0 \le \epsilon_{t,l}, \epsilon'_{t,l} \le \mathbf{P}((A''_{t,l})^c)$. If $\mu(y_0) = 0$, then $\mathbf{P}(Y_t = y_0|Y_0 = x) \le \mathbf{P}((A''_{t,l})^c)$ for all l. Then $\mathbf{P}(Y_t = y_0|Y_0 = x) \le \lim_{l\to\infty} \mathbf{P}((A''_{t,l})^c) = 0$. To obtain the vanishing of the last limit, we first notice that, for given $T > 0$, $\mathbf{P}(A''_{t,l}) \ge \mathbf{P}(B_s + x \in [-l,l],\, s \le T) - \mathbf{P}(\psi_t(x) > T)$, where B is the standard Brownian motion in (2.1). The latter probability is bounded above by $\mathbf{P}(\phi_T(x) \le t)$. Thus $\liminf_{l\to\infty} \mathbf{P}(A''_{t,l}) \ge 1 - \limsup_{T\to\infty} \mathbf{P}(\phi_T(x) \le t)$. From (2.1) and the known fact that almost surely $\lim_{T\to\infty} \ell(T,x') = \infty$ for all x', the latter lim sup is seen to vanish. The proof of the lemma is complete. $\square$

Lemma 2.2 *For all y_0, $\mathbf{P}(Y_t = y_0|Y_0 = x)$ is continuous in $t > 0$.*

Proof.

In view of Lemma 2.1, it suffices to consider the case where $\mu(y_0) > 0$. Let t', t be such that $|t'-t| \le 1$. Imitating the argument of the proof of that lemma,

$$\begin{aligned}&|\mathbf{P}(Y_{t'} = y_0|Y_0 = x) - \mathbf{P}(Y_t = y_0|Y_0 = x)| \\ &\le |p'_I(t,x,y_0) - p'_I(t',x,y_0)|\,\mu(y_0) \;+\; \mathbf{P}((A''_{t,l})^c) + \mathbf{P}((A''_{t',l})^c).\end{aligned} \tag{2.16}$$

Then $\lim_{t'\to t} |\mathbf{P}(Y_{t'} = y_0|Y_0 = x) - \mathbf{P}(Y_t = y_0|Y_0 = x)| \le \mathbf{P}((A''_{t,l})^c) + \mathbf{P}((A''_{t+1,l})^c)$ for all l and the result follows as in the proof of Lemma 2.1. $\square$

Lemma 2.3 *The set of locations of the atoms of $\bar\mu$, $\{\bar y_i\}$, contains that of μ, $\{y_{i'}\}$.*

Proof. This is a corollary to the continuity lemma just given and formula (2.3). From that formula, we have

$$\int_t^{t'} \mathbf{P}(Y_s = y_0|Y_0 = x)\,ds = \mathbf{E}(\ell_Y(t',x,y_0) - \ell_Y(t,x,y_0))\mu(y_0). \tag{2.17}$$

We claim that the above expectation is strictly positive for all x, y_0 and $t' > t$, if $\mu(y_0) > 0$. This is a consequence of the definition (see below (2.3)) of ℓ_Y and the fact that there is strictly positive probability that between the two stopping times, $\psi_t(x)$ and $\psi_{t'}(x)$, the Brownian motion B will pass through $y_0 - x$ and hence will strictly increase its local time there. Thus the integral on the left hand side of (2.17) is also strictly positive. This implies that for all x, in every open interval of the positive reals, there exists an s such that $\mathbf{P}(Y_s = y_0|Y_0 = x) > 0$. This and the continuity in time of these probabilities imply that, in every interval $(0,t)$ there exists an s such that $\mathbf{P}(Y_s = y_0|Y_0 = x)\mathbf{P}(Y_{t-s} = y_0|Y_0 = y_0) > 0$. By the Markov property and time homogeneity of Y, the latter product is a lower bound for $\mathbf{P}(Y_t = y_0|Y_0 = x)$. The lemma follows. $\square$

We return to the proof of Condition 1. Let i_l be such that $\bar{y}_{i_l} = y_{i'_l}$. Since $\mu_d^{(\epsilon)} \xrightarrow{pp} \mu_d$, there exists $j'_l(\epsilon)$ so that $(y^{(\epsilon)}_{j'_l(\epsilon)}, w^{(\epsilon)}_{j'_l(\epsilon)}) \to (y_{i'_l}, w_{i'_l})$ as $\epsilon \to 0$. Since, by Lemmas 2.1 and 2.3, $\{\bar{y}^{(\epsilon)}_j\} = \{y^{(\epsilon)}_{j'}\}$, we can define $j_l(\epsilon)$ so that $\bar{y}^{(\epsilon)}_{j_l(\epsilon)} = y^{(\epsilon)}_{j'_l(\epsilon)}$. We already then have $\bar{y}^{(\epsilon)}_{j_l(\epsilon)} \to \bar{y}_{i_l}$ $(= y_{i'_l})$. To obtain Condition 1, we must show that also $\bar{w}^{(\epsilon)}_{j_l(\epsilon)} \to \bar{w}_{i_l}$, i.e., that $\mathbf{P}(Y^{(\epsilon)}_{t_0} = \bar{y}^{(\epsilon)}_{j_l(\epsilon)}) \to \mathbf{P}(Y_{t_0} = \bar{y}_{i_l})$.

Let us simplify the notation a bit by setting $t_0 = 1$, $\bar{y}^{(\epsilon)}_{j_l(\epsilon)} = y^{(\epsilon)}_{j'_l(\epsilon)} = y_\epsilon$, $\bar{y}_{i_l} = y_{i'_l} = y_0$, $w^{(\epsilon)}_{j'_l(\epsilon)} = w_\epsilon$, $w_{i'_l} = w_0$, $\bar{w}^{(\epsilon)}_{j_l(\epsilon)} = \bar{w}_\epsilon$ and $\bar{w}_{i_l} = \bar{w}_0$. Thus, as $\epsilon \to 0$, we have $\mu^{(\epsilon)} \xrightarrow{v} \mu$, $y_\epsilon \to y_0$, $w_\epsilon \equiv \mu^{(\epsilon)}(y_\epsilon) \to \mu(y_0) = w_0$, and we must show that $\bar{w}_\epsilon \equiv \mathbf{P}(Y^{(\epsilon)}_1 = y_\epsilon) \to \mathbf{P}(Y_1 = y_0) = \bar{w}_0$.

We also already know that $Y^{(\epsilon)}_1 \to Y_1$ in distribution (i.e., $\bar{\mu}^{(\epsilon)} \xrightarrow{v} \bar{\mu}$). It follows that $\limsup_{\epsilon\to 0} \mathbf{P}(Y^{(\epsilon)}_1 = y_\epsilon) \le \mathbf{P}(Y_1 = y_0)$, since otherwise $\bar{\mu}^{(\epsilon)} \xrightarrow{v} \bar{\mu}$ would be violated. So we only need to prove

$$\liminf_{\epsilon\to 0} \mathbf{P}(Y^{(\epsilon)}_1 = y_\epsilon) \ge \mathbf{P}(Y_1 = y_0). \tag{2.18}$$

By convergence in distribution,

$$\begin{aligned}
\mathbf{P}(Y_1 = y_0) &= \lim_{\delta\to 0} \mathbf{P}(y_0 - \delta < Y_1 < y_0 + \delta) \\
&\le \lim_{\delta\to 0} \liminf_{\epsilon\to 0} \mathbf{P}(y_0 - \delta \le Y^{(\epsilon)}_1 \le y_0 + \delta) \\
&\le \lim_{\delta\to 0} \left[\liminf_{\epsilon\to 0} \mathbf{P}(Y^{(\epsilon)}_1 = y_\epsilon)\right] \left[\limsup_{\epsilon\to 0} \frac{\mathbf{P}(y_0 - \delta \le Y^{(\epsilon)}_1 \le y_0 + \delta)}{\mathbf{P}(Y^{(\epsilon)}_1 = y_\epsilon)}\right].
\end{aligned}$$

Hence to prove (2.18), it suffices to show that

$$\lim_{\delta\to 0} \limsup_{\epsilon\to 0} \frac{\mathbf{P}(y_0 - \delta \le Y^{(\epsilon)}_1 \le y_0 + \delta)}{\mathbf{P}(Y^{(\epsilon)}_1 = y_\epsilon)} \le 1 \tag{2.19}$$

or, equivalently, that

$$\lim_{\delta\to 0} \liminf_{\epsilon\to 0} \frac{\mathbf{P}(Y^{(\epsilon)}_1 = y_\epsilon)}{\mathbf{P}(y_0 - \delta \le Y^{(\epsilon)}_1 \le y_0 + \delta)} \ge 1. \tag{2.20}$$

Given any small $\delta > 0$, we want to find a small $\delta' = \delta'(\delta) > \delta$ with $\delta' \to 0$ as $\delta \to 0$ and small $\mathcal{T} = \mathcal{T}(\delta)$ with $0 < \mathcal{T} < 1$, such that the following will be valid.

(I$_\epsilon$) $\lim_{\delta\to 0} \liminf_{\epsilon\to 0} \mathbf{P}\left(y_0 - \delta' \le Y^{(\epsilon)}_{1-\mathcal{T}} \le y_0 + \delta' \,\middle|\, y_0 - \delta \le Y^{(\epsilon)}_1 \le y_0 + \delta\right) = 1$

(II$_\epsilon$) $\lim_{\delta\to 0} \liminf_{\epsilon\to 0} \mathbf{P}\left(y_0 - \delta \le Y^{(\epsilon)}_{1-\mathcal{T}} \le y_0 + \delta \,\middle|\, y_0 - \delta' \le Y^{(\epsilon)}_{1-\mathcal{T}} \le y_0 + \delta'\right) = 1$

(III$_\epsilon$) $\lim_{\delta\to 0} \liminf_{\epsilon\to 0} \mathbf{P}\left(\inf_{t\in[1-\mathcal{T},1]} Y^{(\epsilon)}_t \ge y_0 - \delta',\ \sup_{t\in[1-\mathcal{T},1]} Y^{(\epsilon)}_t \le y_0 + \delta' \,\middle|\, Y^{(\epsilon)}_{1-\mathcal{T}} \in [y_0 - \delta, y_0 + \delta]\right) = 1$

(IV$_\epsilon$) $\lim_{\delta\to 0} \liminf_{\epsilon\to 0} \mathbf{P}\left(Y^{(\epsilon)}_t = y_\epsilon \text{ for some } t \in [1 - \mathcal{T}, 1] \,\middle|\, Y^{(\epsilon)}_{1-\mathcal{T}} \in [y_0 - \delta, y_0 + \delta]\right) = 1$

If (I_ϵ)-(IV_ϵ) hold, then (2.20) would be a consequence of

$$\lim_{\delta'\to 0}\limsup_{\epsilon\to 0}\sup_{s>0}\mathbf{P}\left(Y_s^{(\epsilon)}\neq y_\epsilon \text{ and } Y_t^{(\epsilon)}\in[y_0-\delta',y_0+\delta'] \text{ for all } t\in[0,s]\,\middle|\,Y_0^{(\epsilon)}=y_\epsilon\right)=0. \tag{2.21}$$

But this follows from the assumptions $\mu^{(\epsilon)}\overset{v}{\to}\mu$ and $\mu^{(\epsilon)}\overset{pp}{\to}\mu$ (so that $\mu^{(\epsilon)}(y_\epsilon)\to\mu(y_0)$) as $\epsilon\to 0$ and the following lemma.

Lemma 2.4 *For any open interval I containing $y^{(\epsilon)}$,*

$$\mathbf{P}\left(Y_s^{(\epsilon)}\neq y_\epsilon \textit{ and } Y_t^{(\epsilon)}\in I \textit{ for all } t\in[0,s]\,\middle|\,Y_0^{(\epsilon)}=y_\epsilon\right)\leq 1-\frac{\mu^{(\epsilon)}(y_\epsilon)}{\mu^{(\epsilon)}(I)}. \tag{2.22}$$

Proof.

The first step of the proof is similar to part of the proof of Lemma 2.1, except here we use a coupling between the original process $Y_t^{(\epsilon)}$ and a different process on the finite interval, namely the process $\tilde{Y}_t$, whose speed measure is the finite measure $\mu^{(\epsilon)}\cdot 1_I$. (Basically, $\tilde{Y}_t$ has reflecting boundary conditions at both endpoints of I.) Let $A_{s,I}$ denote the event that $Y_t^{(\epsilon)}\in I$ for all $t\in[0,s]$; then we take a coupling in which $\{Y_t^{(\epsilon)},0\leq t\leq s\}$ equals $\{\tilde{Y}_t,0\leq t\leq s\}$ on $A_{s,I}$. Then the probability in (2.22) equals

$$\mathbf{P}(\tilde{Y}_s\neq y_\epsilon \text{ and } A_{s,I}|\tilde{Y}_0=y_\epsilon)\leq\mathbf{P}(\tilde{Y}_s\neq y_\epsilon|\tilde{Y}_0=y_\epsilon)=1-\mathbf{P}(\tilde{Y}_s=y_\epsilon|\tilde{Y}_0=y_\epsilon). \tag{2.23}$$

The proof is completed by applying the following lemma with Y replaced by $\tilde{Y}$, μ by $\mu^{(\epsilon)}\cdot 1_I$, and y by y_ϵ. □

Lemma 2.5 *Let $s\geq 0$ and $y\in\mathbf{R}$; then*

$$\mathbf{P}(Y_s=y|Y_0=y)\geq\frac{\mu(y)}{\mu(\mathbf{R})}. \tag{2.24}$$

Proof.

We may assume that $\mu(y)>0$ and $\mu(\mathbf{R})<\infty$, since otherwise the claim is trivially true. The Markov process Y has $[\mu(\mathbf{R})]^{-1}\mu$ as its unique invariant distribution. Let T_t denote the transition semigroup acting on $L^2(\mathbf{R},d\mu)$:

$$T_t\,:\,f(x)\longmapsto\mathbf{E}(f(Y_t)|Y_0=x). \tag{2.25}$$

Then $T_t=e^{t\mathcal{L}}$ is a continuous self-adjoint contraction semigroup, the generator $\mathcal{L}$ has a simple eigenvalue 0, with normalized eigenfunction the constant function $\Phi(x)=[\mu(\mathbf{R})]^{-1/2}$, and the rest of its spectrum strictly negative. Let $\Psi(x)$ be the (normalized in $L^2(\mathbf{R},d\mu)$) function $[\mu(y)]^{-1/2}1_y$. Then by the spectral theorem for the generator $\mathcal{L}$, and denoting by $(\cdot,\cdot)$ the inner product in $L^2(\mathbf{R},d\mu)$,

$$\mathbf{P}(Y_s=y|Y_0=y)=([\mu(y)]^{-1}1_y,T_s1_y)=(\Psi,e^{t\mathcal{L}}\Psi)=|(\Psi,\Phi)|^2+\int_{-\infty}^{0-}e^{sl}d\nu(l), \tag{2.26}$$

where ν (the spectral measure of Ψ restricted to $(-\infty,0)$) is a totally finite non-negative measure on $(-\infty,0)$. It follows that $\mathbf{P}(Y_s=y|Y_0=y)$ is non-increasing in s and converges, as $s\to\infty$, to $|(\Psi,\Phi)|^2=\mu(y)/\mu(\mathbf{R})$. The proof is complete. □

It remains to show that (I_ϵ)-(IV_ϵ) hold (for some $\delta'(\delta)$ and $\mathcal{T}(\delta)$). From the convergence in distribution (in the Skorohod topology) of the processes ([3]; see the first paragraph of this proof, following Corollary 2.2), we have, for example, that

$$\begin{aligned} \liminf_{\epsilon\to 0} \mathbf{P}\left(Y^{(\epsilon)}_{1-\mathcal{T}} \in [y_0-\delta', y_0+\delta']\right) &\geq \mathbf{P}\left(Y_{1-\mathcal{T}} \in (y_0-\delta', y_0+\delta')\right) \\ &= \lim_{\delta''\uparrow\delta'} \mathbf{P}\left(Y_{1-\mathcal{T}} \in [y_0-\delta'', y_0+\delta'']\right), \end{aligned} \tag{2.27}$$

$$\limsup_{\epsilon\to 0} \mathbf{P}\left(Y^{(\epsilon)}_{1} \in [y_0-\delta, y_0+\delta]\right) \leq \mathbf{P}\left(Y_{1} \in [y_0-\delta, y_0+\delta]\right), \tag{2.28}$$

$$\begin{aligned} &\liminf_{\epsilon\to 0}\ \mathbf{P}\left(Y^{(\epsilon)}_{t} \in [y_0-\delta', y_0+\delta'] \text{ for all } t\in[1-\mathcal{T},1]\right) \\ &\geq \lim_{\delta''\uparrow\delta'}\ \mathbf{P}\left(Y_{t} \in [y_0-\delta'', y_0+\delta''] \text{ for all } t\in[1-\mathcal{T},1]\right), \quad \text{etc.} \end{aligned} \tag{2.29}$$

Thus, (I_ϵ)-(III_ϵ) are seen to follow from the corresponding (I)-(III) with $Y^{(\epsilon)}$ replaced by Y (and $\liminf_{\epsilon\to 0}$ deleted). For (IV_ϵ), a different argument is required, because the usual notions of convergence in distribution of $Y^{(\epsilon)}$ to Y will not work directly for (IV_ϵ) and the analogous (IV). Instead, we replace (IV_ϵ) by a stronger condition (IV'_ϵ), stated in terms of the Brownian motion B of (2.1)–(2.2):

(IV'_ϵ) $\displaystyle\lim_{\delta\to 0}\liminf_{\epsilon\to 0} \mathbf{P}\left(Q_{\epsilon,x,[1-\mathcal{T},1]} \leq y_0-\delta,\ Q^{\epsilon,x,[1-\mathcal{T},1]} \geq y_0+\delta \,\middle|\, Y^{(\epsilon)}_{1-\mathcal{T}} \in [y_0-\delta, y_0+\delta]\right) = 1,$

where

$$Q_{\epsilon,x,[a,b]} = \inf_{s\in[\psi^{(\epsilon)}_a(x),\psi^{(\epsilon)}_b(x)]} (B(s)+x) \tag{2.30}$$

and $Q^{\epsilon,x,[a,b]}$ is defined analogously with inf replaced by sup. This condition is stronger because $\mu^{(\epsilon)}(y_\epsilon) > 0$ and so $Y^{(\epsilon)}$ cannot skip over y_ϵ. Now, as above, by the convergence in distribution of $Y^{(\epsilon)}$ to Y, it suffices to prove the corresponding condition (IV$'$) for Y.

It remains to show that (I)-(III) and (IV$'$) hold for some $\delta'(\delta)$ and $\mathcal{T}(\delta)$.
Since the distribution of Y_{t_0} has an atom at y_0, it follows that

$$\mathbf{P}(Y_{t_0} = y_0 | Y_{t_0} \in [y_0-\delta'', y_0+\delta'']) \to 1 \text{ as } \delta'' \to 0 \tag{2.31}$$

for each t_0. From this and Lemma 2.2 (and the vague continuity in t of the distribution of Y_t, from, e.g., [3]), (II) follows (provided $\delta' \to 0$ as $\delta \to 0$).
Similarly, assuming $\delta' \to 0$ as $\delta\to 0$, we can replace (I) by

(I$'$) $\displaystyle\lim_{\delta\to 0} \mathbf{P}\left(Y_{1-\mathcal{T}} = y_0 | Y_1 = y_0\right) = 1.$

But this follows, assuming $\mathcal{T}\to 0$ as $\delta\to 0$, again from the continuity of $\mathbf{P}(Y_t = y_0)$ in $t > 0$.
Let us take (IV$'$) now. The probability there is bounded from below by

$$\inf_{x\in\mathrm{supp}\mu\cap[y_0-\delta,y_0+\delta]} \mathbf{P}\left(\inf_{s\in[0,\psi_{\mathcal{T}}(x)]} B(s)+x \leq y_0-\delta,\ \sup_{s\in[0,\psi_{\mathcal{T}}(x)]} B(s)+x \geq y_0+\delta \,\middle|\, Y^{(\epsilon)}_0 = x\right). \tag{2.32}$$

Let S'_δ denote the time it takes for $B(s)+y_0-\delta$ to first reach $y_0+\delta$ and then come back to $y_0-\delta$. Then the expression in (2.32) is bounded from below by

$$\mathbf{P}(\phi_{S'_\delta}(y_0-\delta)<\mathcal{T}). \tag{2.33}$$

Now $S'_\delta\to 0$ as $\delta\to 0$ almost surely. From the almost sure continuity of ℓ, we have $\ell(S'_\delta, y-(y_0-\delta))\to\ell(0,y-y_0)\equiv 0$ as $\delta\to 0$ and from this and the monotonicity of ℓ in t and the fact that $\ell(t,\cdot)$ has compact support almost surely for every t, it follows straightforwardly that $\phi_{S'_\delta}(y_0-\delta)\to 0$ as $\delta\to 0$ almost surely. That means that the probability in (2.33) would tend to 1 as $\delta\to 0$ for any fixed $\mathcal{T}>0$ (i.e., not depending on δ). But then we can choose a sequence $\mathcal{T}=\mathcal{T}(\delta)$, with $\mathcal{T}\to 0$ as $\delta\to 0$, such that it still tends to 1 as $\delta\to 0$. This establishes (IV$'$).

Finally we need to choose δ' so that (III) is valid. The argument is analogous to the above one for (IV$'$). The probability in (III) is bounded from below by

$$\inf_{x\in\mathrm{supp}\mu\cap[y_0-\delta,y_0+\delta]}\mathbf{P}\left(\inf_{t\in[0,\mathcal{T}]}Y_t\geq y_0-\delta',\ \sup_{t\in[0,\mathcal{T}]}Y_t\leq y_0+\delta'\middle|\, Y_0=x\right). \tag{2.34}$$

Let Y'_t be a copy of Y_t, starting at $y_0-\delta$ at time 0 and let Y''_t be a copy of Y_t starting at $y_0+\delta$ at time 0. Let $T'_{\delta\delta'}$ and $T''_{\delta\delta'}$ denote the time it takes for Y'_t and Y''_t to first reach beyond $(y_0-\delta', y_0+\delta')$, respectively, and let $\mathcal{T}=\mathcal{T}(\delta)$ be as chosen in the previous paragraph with $\mathcal{T}\to 0$ as $\delta\to 0$. Then the expression in (2.34) is bounded from below by

$$\mathbf{P}(T'_{\delta\delta'}>\mathcal{T})+\mathbf{P}(T''_{\delta\delta'}>\mathcal{T})-1. \tag{2.35}$$

Let us take the first term of (2.35). Consider $S'_{\delta\delta'}$, the quantity corresponding to $T'_{\delta\delta'}$ for $B(s)+y_0-\delta$. Then, if δ' were fixed, we would have that $S'_{\delta\delta'}\to S'_{0\delta'}$ as $\delta\to 0$ almost surely. Now $T'_{\delta\delta'}>\mathcal{T}$ if $\phi_{S'_{\delta\delta'}}(y_0-\delta)>\mathcal{T}$. By the same reasoning as above, we have, for fixed δ', that $\phi_{S'_{\delta\delta'}}(y_0-\delta)\to\phi_{S'_{0\delta'}}(y_0)$ as $\delta\to 0$ almost surely. That means that, as $\delta\to 0$, the first term of (2.35) is bounded below by $\mathbf{P}(\phi_{S'_{0\delta'}}(y_0)>0)=1$ for any fixed $\delta'>0$, since $S'_{0\delta'}>0$ for any $\delta'>0$ and $\phi_s(y_0)>0$ for all $s>0$ almost surely. That means that $\mathbf{P}(T'_{\delta\delta'}>\mathcal{T}(\delta))\to 1$ as $\delta\to 0$ for any fixed $\delta'>0$ and the same can be argued analogously for the second term of (2.35). Thus we can choose a sequence $\delta'=\delta'(\delta)$, with $\delta'\to 0$ as $\delta\to 0$, such that (2.35) tends to 1 as $\delta\to 0$. This establishes (III) and the theorem. □

3 Scaling limit for a random walk with random rates

Let X_t, $t\geq 0$, $X_0=0$, be a continuous time random walk on $\mathbf{Z}$ with inhomogeneous rates given by $\lambda_i=\tau_i^{-1}$, $i\in\mathbf{Z}$, where τ_i, $i\in\mathbf{Z}$, are i.i.d. random variables such that $\mathbf{P}(\tau_0>0)=1$ and $\mathbf{P}(\tau_0>t)=L(t)/t^\alpha$, where L is a nonvanishing slowly varying function as $t\to\infty$ and $\alpha<1$.

Consider the Lévy process (see, e.g., [21, 22, 23]) V_x, $x\in\mathbf{R}$, $V_0=0$, with stationary and independent increments given by

$$\mathbf{E}\left[e^{ir(V_{x+x_0}-V_{x_0})}\right]=e^{\alpha x\int_0^\infty(e^{irw}-1)\,w^{-1-\alpha}\,dw} \tag{3.1}$$

for any $x_0\in\mathbf{R}$ and $x\geq 0$. It satisfies

$$\lim_{y\to\infty}y^\alpha\mathbf{P}(V_1>y)=1 \tag{3.2}$$

([17], Theorem XVII.5.3). Let ρ be the (random) Lebesgue-Stieltjes measure on the Borel sets of $\mathbf{R}$ associated to V, i.e.,

$$\rho((a,b]) = V_b - V_a,\ a,b \in \mathbf{R},\ a < b, \tag{3.3}$$

where we have chosen the process V to have sample paths that are right-continuous (with left-limits). Then

$$\frac{d\rho}{dx} = \frac{dV}{dx} = \sum_j w_j\, \delta(x - x_j), \tag{3.4}$$

where the (countable) sum is over the indices of an inhomogeneous Poisson point process $\{(x_j, w_j)\}$ on $\mathbf{R} \times (0,\infty)$ with density $dx\, \alpha w^{-1-\alpha}\, dw$.

For each $\epsilon > 0$, we want to define, in the fixed probability space on which V and ρ are defined, a sequence $\tau_i^{(\epsilon)}$, $i \in \mathbf{Z}$, of independent random variables such that

$$\tau_i^{(\epsilon)} \sim \tau_0 \quad \text{for every } i \in \mathbf{Z} \tag{3.5}$$

(where $\sim$ denotes equidistribution and τ_0 is as above) and with the following property: For a given family of constants c_ϵ, $\epsilon > 0$, let

$$\tilde{\rho}^{(\epsilon)} := \sum_{i=-\infty}^{\infty} c_\epsilon \tau_i^{(\epsilon)} \delta_{\epsilon i}; \tag{3.6}$$

we demand that

$$\tilde{\rho}^{(\epsilon)} \stackrel{v}{\to} \rho \quad \text{and} \quad \tilde{\rho}^{(\epsilon)} \stackrel{pp}{\to} \rho \quad \text{as } \epsilon \to 0 \text{ almost surely.} \tag{3.7}$$

The next proposition states that (3.5) and (3.7) hold for the following c_ϵ and $\tau_i^{(\epsilon)}$'s:

$$c_\epsilon = \left(\inf\{t \ge 0 : \mathbf{P}(\tau_0 > t) \le \epsilon\}\right)^{-1} \tag{3.8}$$

$$\tau_i^{(\epsilon)} = \frac{1}{c_\epsilon}\, g_\epsilon\Big(V_{\epsilon(i+1)} - V_{\epsilon i}\Big), \tag{3.9}$$

where g_ϵ is defined as follows. Let $G : [0,\infty) \to [0,\infty)$ satisfy

$$\mathbf{P}(V_1 > G(x)) = \mathbf{P}(\tau_0 > x) \quad \text{for all } x \ge 0. \tag{3.10}$$

G is well-defined since V_1 has a continuous distribution. Notice that that G is nondecreasing and right-continuous and thus has a nondecreasing and right-continuous generalized inverse G^{-1}. Let $g_\epsilon : [0,\infty) \to [0,\infty)$ be defined as

$$g_\epsilon(x) = c_\epsilon G^{-1}\Big(\epsilon^{-1/\alpha} x\Big) \quad \text{for all } x \ge 0. \tag{3.11}$$

Proposition 3.1 *(3.5) and (3.7) hold for c_ϵ and $\tau_i^{(\epsilon)}$ as in (3.8)-(3.11).*

Proof. We will prove (3.5) now and postpone (3.7) until later in this section (following Theorems 3.1 and 3.2).

To establish (3.5), by the stationarity of the increments of V, it suffices to take $i = 0$. Then, for $\epsilon > 0$

$$\mathbf{P}(\tau_0^{(\epsilon)} > t) = \mathbf{P}(g_\epsilon(V_\epsilon) > c_\epsilon t) = \mathbf{P}(V_\epsilon > g_\epsilon^{-1}(c_\epsilon t)) = \mathbf{P}(V_\epsilon > \epsilon^{1/\alpha} G(t)), \tag{3.12}$$

where g_ϵ^{-1} is the right-continuous inverse of g_ϵ, and we have used the easily checked fact that $g_\epsilon^{-1}(\cdot) = \epsilon^{1/\alpha} G(\cdot/c_\epsilon)$. The desired result (3.5) now follows by the scaling relation $V_\epsilon \sim \epsilon^{1/\alpha} V_1$ (see (3.1)) and (3.10). □

We consider next the scaling limit of the random walk X_t. Let

$$Z_t^{(\epsilon)} = \epsilon X_{t/(c_\epsilon \epsilon)}, \quad t \geq 0. \tag{3.13}$$

To study the limit of $Z^{(\epsilon)}$, in the presence of the random rates, which themselves converge vaguely and in the point process sense, but only in distribution, we will need a weak notion of vague and point process convergence, as follows. Let $\tilde{\mathcal{C}}_b$ be the class of bounded real functions f on the space $\mathcal{M}$ of locally finite measures (on $\mathbf{R}$) that are weakly continuous in the sense that $f(\mu_n) \to f(\mu)$ as $n \to \infty$ for all μ, μ_n, $n \geq 1$, in $\mathcal{M}$ such that *both* $\mu_n \stackrel{v}{\to} \mu$ *and* $\mu_n \stackrel{pp}{\to} \mu$ as $n \to \infty$.

Definition 3.1 *Let P, P_ϵ, $\epsilon > 0$ be probability measures on $\mathcal{M}$. We say that P_ϵ converges* doubly weakly *to P, denoted $P_\epsilon \stackrel{w}{\Rightarrow} P$, if, as $\epsilon \to 0$,*

$$\int f\, dP_\epsilon \to \int f\, dP \text{ for all } f \in \tilde{\mathcal{C}}_b. \tag{3.14}$$

We also use the notation $\pi^{(\epsilon)} \stackrel{w}{\Rightarrow} \pi$ for random measures $\pi^{(\epsilon)}$ and π, whose distributions are P_ϵ and P respectively.

Let Z_t be the (random) quasidiffusion Y_t as in (2.1)-(2.2) above, but with speed measure μ taken to be (random) discrete measure ρ of (3.3)–(3.4) associated with the Lévy process V. For $t_0 > 0$ fixed, let $\bar{\rho}$ and $\bar{\rho}^{(\epsilon)}$ be the (random) probability distributions of Z_{t_0} and $Z_{t_0}^{(\epsilon)}$, respectively. We can now state the following corollaries to Proposition 3.1, Theorem 2.1 and Corollary 2.1.

Theorem 3.1 *As $\epsilon \to 0$,*

$$\bar{\rho}^{(\epsilon)} \stackrel{w}{\Rightarrow} \bar{\rho}. \tag{3.15}$$

Theorem 3.2 *As $\epsilon \to 0$,*

$$\sum_{x \in \mathbf{R}} \mathbf{E} \left\{ \left[\bar{\rho}^{(\epsilon)}(\{x\}) \right]^2 \right\} \to \sum_{x \in \mathbf{R}} \mathbf{E} \left\{ [\bar{\rho}(\{x\})]^2 \right\}. \tag{3.16}$$

Proof of Theorem 3.1

$Z_t^{(\epsilon)}$ is distributed as a standard Brownian motion time changed through the speed measure $\rho^{(\epsilon)}$ (see (1.16) and the surrounding discussion in the Introduction). Let ρ and $\tilde{\rho}^{(\epsilon)}$ be as in (3.3) and (3.6) and let $\tilde{Z}^{(\epsilon)}$ be a standard Brownian motion time changed through $\tilde{\rho}^{(\epsilon)}$. By Proposition 3.1,

$$(Z^{(\epsilon)}, \rho^{(\epsilon)}) \sim (\tilde{Z}^{(\epsilon)}, \tilde{\rho}^{(\epsilon)}). \tag{3.17}$$

To obtain (3.15), it is thus enough, by (3.17) and dominated convergence to show that $\bar{\tilde{\rho}}^{(\epsilon)}$, the probability distribution of $\tilde{Z}_{t_0}^{(\epsilon)}$, (which is random because of its dependence on $\tilde{\rho}^{(\epsilon)}$ and hence on the Lévy process V) satisfies: $\bar{\tilde{\rho}}^{(\epsilon)} \xrightarrow{v} \bar{\rho}$ and $\bar{\tilde{\rho}}^{(\epsilon)} \xrightarrow{pp} \bar{\rho}$ almost surely. But that follows from Proposition 3.1 and Theorem 2.1. □

Proof of Theorem 3.2

Similarly, it suffices to prove that almost surely

$$\sum_{x\in\mathbf{R}} \left[\bar{\tilde{\rho}}^{(\epsilon)}(\{x\})\right]^2 \to \sum_{x\in\mathbf{R}} \left[\bar{\rho}(\{x\})\right]^2 \tag{3.18}$$

and that follows from Proposition 3.1 and Corollary 2.1. □

It remains to derive (3.7) and thus finish the proof of Proposition 3.1. For that we will need two main lemmas, as follows.

Lemma 3.1 *For any fixed $y > 0$, $g^{(\epsilon)}(y) \to y$ as $\epsilon \to 0$.*

Lemma 3.2 *For any $\delta' > 0$, there exist constants C' and C'' in $(0, \infty)$ such that*

$$g_\epsilon(x) \le C' x^{1-\delta'} \text{ for } \epsilon^{1/\alpha} \le x \le 1 \text{ and } \epsilon \le C''.$$

The proofs of these two main lemmas are based on the following four subsidiary lemmas, whose proofs are given later.

Lemma 3.3 $\dfrac{1}{\epsilon}\mathbf{P}\left(\tau_0 > \dfrac{1}{c_\epsilon}\right) \to 1$ *as $\epsilon \to 0$.*

Lemma 3.4 *For $y > 0$,* $\dfrac{1}{\epsilon}\mathbf{P}\left(\tau_0 > \dfrac{y}{c_\epsilon}\right) \to \dfrac{1}{y^\alpha}$ *as $\epsilon \to 0$.*

Lemma 3.5 *For any $\lambda > 0$, $\dfrac{c_\epsilon}{c_{\lambda\epsilon}} \to \lambda^{-1/\alpha}$ as $\epsilon \to 0$ and thus (by standard results, as in [17]) $c_\epsilon = \epsilon^{1/\alpha}\tilde{L}(\epsilon^{-1})$, where $\tilde{L}$ is a positive slowly varying function at infinity.*

Lemma 3.6 *There exists $\lambda > 0$ sufficiently small such that $G^{-1}(y) \le 1/c_{\lambda/y^\alpha}$ for $y \ge 1$ or, equivalently, $g_\epsilon(x) \le c_\epsilon / c_{\lambda\epsilon/x^\alpha}$ for $x \ge \epsilon^{1/\alpha}$.*

Proof of Lemma 3.1

Let g_ϵ^{-1} be the right-continuous generalized inverse of g_ϵ. To prove $g_\epsilon(y) \to y$, it suffices to prove that $g_\epsilon^{-1}(y) \to y$.

Now $G^{-1}(V_1) \sim \tau_0$, so $g_\epsilon(\epsilon^{1/\alpha}V_1) = c_\epsilon G^{-1}(\epsilon^{-1/\alpha}\epsilon^{1/\alpha}V_1) \sim c_\epsilon\tau_0$, and thus $\mathbf{P}(\tau_0 > y/c_\epsilon)$ equals

$$\mathbf{P}(c_\epsilon\tau_0 > y) = \mathbf{P}(g_\epsilon(\epsilon^{1/\alpha}V_1) > y) = \mathbf{P}(\epsilon^{1/\alpha}V_1 > g_\epsilon^{-1}(y)) = \mathbf{P}(V_1 > \epsilon^{-1/\alpha}g_\epsilon^{-1}(y)). \tag{3.19}$$

By (3.2),

$$\epsilon^{-1}\mathbf{P}(V_1 > \epsilon^{-1/\alpha}y) \to 1/y^\alpha \tag{3.20}$$

as $\epsilon \to 0$. By (3.19) and Lemma 3.4,

$$\epsilon^{-1}\mathbf{P}(V_1 > \epsilon^{-1/\alpha}g_\epsilon^{-1}(y)) = \epsilon^{-1}\mathbf{P}(\tau_0 > y/c_\epsilon) \to 1/y^\alpha \tag{3.21}$$

as $\epsilon \to 0$. This implies that $\mathbf{P}(V_1 > \epsilon^{-1/\alpha} g_\epsilon^{-1}(y))/\mathbf{P}(V_1 > \epsilon^{-1/\alpha} y) \to 1$ as $\epsilon \to 0$ and this plus (3.20) implies that $\limsup_{\epsilon\to 0} g_\epsilon^{-1}(y) \le y$ and $\liminf_{\epsilon\to 0} g_\epsilon^{-1}(y) \ge y$, completing the proof of Lemma 3.1. $\square$

Proof of Lemma 3.2

By Lemmas 3.5 and 3.6, for $x \ge \epsilon^{1/\alpha}$,

$$g_\epsilon(x) \le \lambda^{-1/\alpha} x \frac{\tilde{L}(\epsilon^{-1})}{\tilde{L}((x^\alpha/\lambda)\epsilon^{-1})} \tag{3.22}$$

for $\lambda > 0$ small enough; the value of λ will be chosen later. We now use a result from p. 274 of [17] about slowly varying functions, stating that $\tilde{L}(x) = a(x) e^{\int_1^x \frac{\Delta(y)}{y}\,dy}$, where $a(x) \to c \in (0,\infty)$ as $x \to \infty$ and $\Delta(y) \to 0$ as $y \to \infty$. The quotient in the right hand side of (3.22) then becomes

$$\frac{a(\epsilon^{-1})}{a((x^\alpha/\lambda)\epsilon^{-1})} \exp\left\{ \int_{(x^\alpha/\lambda)\epsilon^{-1}}^{\epsilon^{-1}} \frac{\Delta(y)}{y}\,dy \right\}. \tag{3.23}$$

If $\epsilon \le \lambda$ so that $(x^\alpha/\lambda)\epsilon^{-1} \ge 1/\lambda \ge \epsilon^{-1}$, then the absolute value of the latter integral is bounded above by

$$\delta \left| \int_{(x^\alpha/\lambda)\epsilon^{-1}}^{\epsilon^{-1}} \frac{1}{y}\,dy \right| \le \delta |\log(x^\alpha/\lambda)|, \tag{3.24}$$

where $\delta = \delta(\lambda) = \sup\{|\Delta(y)|, y > 1/\lambda\}$, and thus the exponential in (3.23) is bounded above (for $\lambda \le 1$, $x \le 1$) by

$$\lambda^{-\delta} x^{-\alpha\delta}. \tag{3.25}$$

Thus, given $\delta' > 0$, we choose $\lambda \in (0,1)$ such that $\alpha\delta(\lambda) \le \delta'$ and such that $a(y) \in [c/2, c]$ for $y \ge \lambda^{-1}$. The lemma now follows from (3.22)-(3.25) with $C' = 4\lambda^{-(1+\delta')/\alpha}$ and $C'' = \lambda$. $\square$

To complete the proof of our two main lemmas, it remains to prove the subsidiary Lemmas 3.3, 3.4, 3.5 and 3.6.

Proof of Lemma 3.3

By the definition (3.8) of c_ϵ, $\mathbf{P}(\tau_0 > c_\epsilon^{-1}) \le \epsilon$ and $\mathbf{P}(\tau_0 > x) > \epsilon$ for all $x < c_\epsilon^{-1}$. Thus, if the statement of the lemma is not true, then there must exist $\delta \in (0,1)$ and a sequence (ϵ_i) with $\epsilon_i > 0$ for all i and $\epsilon_i \to 0$ as $i \to \infty$ such that $\mathbf{P}(\tau_0 > c_{\epsilon_i}^{-1}) \le \delta\epsilon_i$ for all i. But then, given δ' such that $\delta^{1/\alpha} < \delta' < 1$, we have that $\mathbf{P}(\tau_0 > \delta' c_{\epsilon_i}^{-1}) > \epsilon_i$ and so

$$\frac{\mathbf{P}(\tau_0 > \delta' c_{\epsilon_i}^{-1})}{\mathbf{P}(\tau_0 > c_{\epsilon_i}^{-1})} \ge \delta^{-1} \tag{3.26}$$

for all i. Since $c_{\epsilon_i}^{-1} \to \infty$ and $\mathbf{P}(\tau_0 > \cdot)$ is regularly varying at infinity (with exponent $-\alpha$), it follows that for any $\lambda > 0$,

$$\lim_{t\to\infty} \frac{\mathbf{P}(\tau_0 > \lambda t)}{\mathbf{P}(\tau_0 > t)} = \lambda^{-\alpha}, \tag{3.27}$$

which contradicts (3.26) since $(\delta')^\alpha > \delta$. $\square$

Proof of Lemma 3.4

This is a consequence of Lemma 3.3, the fact that $c_\epsilon^{-1} \to \infty$ as $\epsilon \to 0$, and (3.27), from which it follows that

$$\frac{\mathbf{P}(\tau_0 > y/c_\epsilon)}{\mathbf{P}(\tau_0 > 1/c_\epsilon)} \to \frac{1}{y^\alpha}.$$

Proof of Lemma 3.5

By Lemma 3.3, $(\lambda\epsilon)^{-1}\mathbf{P}(\tau_0 > 1/c_{\lambda\epsilon}) \to 1$ or equivalently $\epsilon^{-1}\mathbf{P}(\tau_0 > 1/c_{\lambda\epsilon}) \to \lambda$ as $\epsilon \to 0$ while, by Lemma 3.4, $\epsilon^{-1}\mathbf{P}(\tau_0 > y/c_\epsilon) \to 1/y^\alpha$. This implies that taking $y^\alpha = \lambda^{-1}$ that $c_\epsilon \lambda^{1/\alpha}/c_{\lambda\epsilon} \to 1$ or $c_\epsilon/c_{\lambda\epsilon} \to \lambda^{-1/\alpha}$ as $\epsilon \to 0$.

Proof of Lemma 3.6

To show that $G^{-1}(y) \leq z$, it is enough to show that $G(z) > y$. Thus we want to prove that $G(1/c_{\lambda/y^\alpha}) > y$ for $y \geq 1$ and some $\lambda > 0$. By the definition (3.10) of G, $G(x) > y$ would be a consequence of $\mathbf{P}(V_1 > y) > \mathbf{P}(\tau_0 > x)$, where we take $x = 1/c_{\lambda/y^\alpha}$. Now there exists $K > 0$ such that $\mathbf{P}(V_1 > y) > K/y^\alpha$ for $y \geq 1$ (by (3.2)), so it suffices to show that $\mathbf{P}(\tau_0 > 1/c_{\lambda/y^\alpha}) \leq K/y^\alpha$ for $y \geq 1$ and some $\lambda > 0$; or, equivalently, taking $\epsilon = \lambda/y^\alpha$, it suffices to show that for some $\lambda > 0$ and all $\epsilon \leq \lambda$, $\mathbf{P}(\tau_0 > 1/c_\epsilon) \leq K\epsilon/\lambda$, or $\mathbf{P}(\tau_0 > 1/c_\epsilon)/\epsilon \leq K/\lambda$. By Lemma 3.3, we may choose λ small enough so that for $\epsilon \leq \lambda$, $\mathbf{P}(\tau_0 > 1/c_\epsilon)/\epsilon \leq 2$ and also small enough that $K/\lambda \geq 2$.

Completion of Proof of Proposition 3.1. We still have to prove (3.7). This will be done using our two main lemmas 3.1 and 3.2. The point process convergence of (3.7) would follow straightforwardly if we knew that $g_\epsilon(x_\epsilon) \to x_0$ as $\epsilon \to 0$ whenever $x_\epsilon \to x_0 > 0$. To obtain that, due to the monotonicity and right continuity of $g_\epsilon(\cdot)$, it suffices that $g_\epsilon(y) \to y$ as $\epsilon \to 0$ for any fixed $y > 0$, and that is given by Lemma 3.1.

We argue next why the vague convergence of (3.7) follows by using both Lemma 3.1 and Lemma 3.2. Let f be a continuous function with bounded support I. Then

$$\int f\, d\tilde{\rho}^{(\epsilon)} = \sum_{i \in \epsilon^{-1}I} f(\epsilon i) g_\epsilon(V_{\epsilon(i+1)} - V_{\epsilon i}). \tag{3.28}$$

For $y > 0$, let $J_y = \{i \in \epsilon^{-1}I : V_{\epsilon(i+1)} - V_{\epsilon i} \geq y\}$. To estimate (3.28), we treat separately the sums over J_δ, $J_{\epsilon^{1/a}} \backslash J_\delta$ and $\epsilon^{-1}I \backslash J_{\epsilon^{1/a}}$, with $\delta > \epsilon^{1/a}$. From Lemma 3.1, it follows that as $\epsilon \to 0$,

$$\sum_{i \in J_\delta} f(\epsilon i) g_\epsilon(V_{\epsilon(i+1)} - V_{\epsilon i}) \to \sum_{j : w_j \geq \delta} f(x_j) w_j, \tag{3.29}$$

where $\{(x_j, w_j)\}$ is the Poisson point process of (3.4).

By Lemma 3.2, we have that, given $\delta' > 0$ small enough (to be chosen shortly), for some finite constant C,

$$\sum_{i \in J_{\epsilon^{1/a}} \backslash J_\delta} f(\epsilon i) g_\epsilon(V_{\epsilon(i+1)} - V_{\epsilon i}) \leq C \sum_{i \in J_{\epsilon^{1/a}} \backslash J_\delta} (V_{\epsilon(i+1)} - V_{\epsilon i})^{1-\delta'}. \tag{3.30}$$

The latter sum is bounded above by

$$W_\delta := \sum_{j : x_j \in I, w_j \leq \delta} w_j^{1-\delta'}. \tag{3.31}$$

With $\delta' > 0$ chosen small enough so that $\delta' + \alpha < 1$, we claim that $W := \lim_{\delta \to 0} W_\delta = 0$ almost surely. Indeed, note that W is well defined in any case by monotonicity and is of course non-negative. We also have, by a standard Poisson process calculation, that

$$\mathbf{E}(W_\delta) = |I| \int_0^\delta w^{1-\delta'} w^{-1-\alpha}\, dw < \infty \tag{3.32}$$

for all $\delta > 0$ and $\mathbf{E}(W_\delta) \to 0$ as $\delta \to 0$. By dominated convergence, $\mathbf{E}(W) = 0$ and the claim follows.

Finally, by the definition (3.11) of g_ϵ and its monotonicity, we have that $g_\epsilon(x) \le g_\epsilon(\epsilon^{1/a}) = Cc_\epsilon$ for $x \le \epsilon^{1/a}$, where C is some finite constant. It then follows that

$$\sum_{i \in \epsilon^{-1} I \setminus J_{\epsilon^{1/a}}} f(\epsilon i) g_\epsilon(V_{\epsilon(i+1)} - V_{\epsilon i}) \le C' c_\epsilon \sum_{i \in \epsilon^{-1} I} \le C'' c_\epsilon \epsilon^{-1} \to 0 \tag{3.33}$$

as $\epsilon \to 0$, by Lemma 3.5, since $\alpha < 1$.

Combining the above estimates, we get that $\int f\, d\tilde{\rho}^{(\epsilon)}$ converges to

$$\lim_{\epsilon \to 0} \sum_{i \in \epsilon^{-1} I} f(\epsilon i) g_\epsilon(V_{\epsilon(i+1)} - V_{\epsilon i}) = \lim_{\delta \to 0} \sum_{j : w_j \ge \delta} f(x_j) w_j = \sum_j f(x_j) w_j = \int f\, d\rho. \square \tag{3.34}$$

Acknowledgements.

The authors have benefitted greatly from discussions with Anton Bovier about aging and with Raghu Varadhan about quasidiffusions. They also thank Maury Bramson, Amir Dembo, Alice Guionnet, Harry Kesten, Tom Kurtz, Enzo Marinari, Henry McKean and Alain Sznitman for useful comments and suggestions.

The research of L.F. is part of FAPESP theme project no. 99/11962-9 and CNPq PRONEX project no. 41/96/0923/00. It was also supported in part by CNPq research grant no. 300576/92-7, and travel grants from FAPESP no. 00/02773-7 and CNPq/NSF no. 910116/95-4.

The research of M.I. was supported in part by MURST under grants *Stime, ordinamenti stocastici e processi aleatori* and *Processi stocastici con struttura spaziale.*

The research of C.M.N. was supported in part by the N.S.F. under grant DMS-9803267.

Each of the authors thanks the departments and universities of their coauthors for hospitality and support provided during various visits.

References

[1] Fontes, L.R.; Isopi, M.; Newman, C.M., Chaotic time dependence in a disordered spin system. *Probab. Theory Relat. Fields* **115**, 417–443 (1999)

[2] Bouchaud, J.-P.; Cugliandolo, L.; Kurchan, J., Mézard, M., Out of equilibrium dynamics in spin-glasses and other glassy systems, in *Spin-glasses and Random Fields* (A.P. Young, Ed.). World Scientific, Singapore (1998)

[3] Stone, C., Limit theorems for random walks, birth and death processes, and diffusion processes. *Illinois J. Math.* **7**, 638–660 (1963)

[4] Havlin, S.; Ben-Avraham, D., Diffusion in disordered media. *Adv. Phys.* **36**, 695–798 (1987)

[5] Isichenko, M.B., Percolation, statistical topography and transport in random media. *Rev. Mod. Phys.* **64**, 961–1043 (1992)

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

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

[8] Golosov, A. O., Localization of random walks in one-dimensional random environments. *Comm. Math. Phys.* **92**, 491–506 (1984)

[9] Kawazu, K.; Kesten, H., On birth and death processes in symmetric random environment. *J. Statist. Phys.* **37**, 561–576 (1984)

[10] Schumacher, S., Diffusions with random coefficients. Ph.D. thesis, University of California, Los Angeles, 1984.

[11] Schumacher, S., Diffusions with random coefficients, in *Particle systems, random media and large deviations* (R. Durrett, Ed.). *Contemp. Math.* **41**, 351–356. American Mathematical Society, Providence (1985)

[12] Newman, C.M.; Stein, D.L., Equilibrium pure states and nonequilibrium chaos. *J. Stat. Phys.* **94**, 709–722 (1999)

[13] Nanda, S.; Newman, C.M.; Stein, D.L., Dynamics of Ising spin systems at zero temperature, in *On Dobrushin's Way. From Probability Theory to Statistical Physics* (R. Minlos, S. Shlosman and Y. Suhov, Eds.). *Amer. Math. Soc. Transl.* (2) **198**, 183–194 (2000)

[14] Rinn, B.; Maass, P.; Bouchaud, J.-P., Multiple scaling regimes in simple aging models. *Phys. Rev. Lett.* **84**, 5403–5406 (2000)

[15] Itô, K; McKean, H.P., *Diffusion processes and their sample paths*, Springer-Verlag, New York (1965)

[16] Kipnis, C.; Varadhan, S.R.S., Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104**, 1–19 (1986)

[17] Feller, W., *An introduction to probability theory and its applications*, Vol. II, Wiley (1966)

[18] Küchler, U., Some asymptotic properties of the transition densities of one-dimensional quasidiffusions, *Publ. Res. Inst. Math. Sci.* **16**, 245–268 (1980)

[19] Knight, F.B., Characterization of the Levy measures of inverse local times of gap diffusion, *Progr. Prob. Statist.* **1**, 53–78 (1981)

[20] Kotani, S., Watanabe, S., Kreĭn's spectral theory of strings and generalized diffusion processes, *Lecture Notes in Math.* **923**, 235–259 (1982)

[21] Resnick, S.I., *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag, New York (1987)

[22] Samorodnitsky, G., Taqqu, M.S., *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York (1994)

[23] Sato, K., *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge (1999)

Instituto de Matemática e Estatística — Universidade de São Paulo
Cx. Postal 66281 — 05315-970 São Paulo SP — Brasil
<lrenato@ime.usp.br>

Dipartimento di Matematica — Università di Bari
Via E. Orabona 4 — 70125 Bari — Italia
<isopi@dm.uniba.it>

Courant Institute of Mathematical Sciences — New York University
251 Mercer Street — New York, NY 10012 — USA
<newman@cims.nyu.edu>