The Annals of Probability
2004, Vol. 32, No. 1B, 996–1029

LARGE DEVIATIONS FOR RANDOM WALK IN RANDOM ENVIRONMENT WITH HOLDING TIMES

BY AMIR DEMBO,[1] NINA GANTERT[2] AND OFER ZEITOUNI[3]

Stanford University, Universität Karlsruhe, and Technion—Israel Institute of Technology and University of Minnesota

Suppose that the integers are assigned the random variables $\{\omega_x, \mu_x\}$ (taking values in the unit interval times the space of probability measures on $\mathbb{R}_+$), which serve as an environment. This environment defines a random walk $\{X_t\}$ (called a RWREH) which, when at x, waits a random time distributed according to μ_x and then, after one unit of time, moves one step to the right with probability ω_x, and one step to the left with probability $1-\omega_x$. We prove large deviation principles for X_t/t, both quenched (i.e., conditional upon the environment), with deterministic rate function, and annealed (i.e., averaged over the environment). As an application, we show that for random walks on Galton–Watson trees, quenched and annealed rate functions *along a ray* differ.

1. Introduction and statement of results. The study of random walks in random environments (RWRE) was initiated in the mid-1970s, and in the last decade there was a resurgence of interest and results for this model; see [16] and [18] for recent reviews. Much of the interest in the topic lies in *trapping* phenomena, a term coined to describe local "pockets" in the environment where the walker spends a relatively large time.

In this paper, we study large deviations for a generalization of the RWRE on $\mathbb{Z}$ that is obtained by allowing for random holding times. We begin by giving a formal definition of the random walk in random environment with holding times (RWREH). Fix $\varepsilon > 0$, and $S_\varepsilon := [\varepsilon, 1-\varepsilon] \times M_1^\varepsilon(\overline{\mathbb{R}}_+)$, where $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$ (with the usual one-point compactification at ∞) and $M_1^\varepsilon(\overline{\mathbb{R}}_+)$ denotes the space of Borel probability measures μ on $\overline{\mathbb{R}}_+$ such that $\mu(\mathbb{R}_+) \geq \varepsilon$. An *environment* $\bar{\omega} \in S_\varepsilon^{\mathbb{Z}} =: \overline{\Omega}_\varepsilon$ has coordinates $\bar{\omega}_x = (\omega_x, \mu_x) \in S_\varepsilon$. For each $\bar{\omega} \in \overline{\Omega}_\varepsilon$, we define the RWRE $\{Z_n\}$ on $\mathbb{Z}$ as the Markov process with $Z_0 = 0$ and transition probabilities

$$\widetilde{P}_{\bar{\omega}}(Z_{n+1} = z+1 | Z_n = z) = \omega_z,$$
$$\widetilde{P}_{\bar{\omega}}(Z_{n+1} = z-1 | Z_n = z) = 1-\omega_z.$$

Received June 2002; revised December 2002.

[1]Supported in part by NSF Grant DMS-00-72331, a US–Israel BSF grant and a Lady Davis fellowship at Technion, Israel.

[2]Supported in part by the DFG. Part of this work was done while the second author was visiting the Department of Electrical Engineering, Technion, Israel.

[3]Supported in part by a US–Israel BSF grant.

AMS 2000 subject classifications. 60J15, 60F10, 82C44, 60J80.

Key words and phrases. Random walk in random environment, large deviations, holding times.

Next define

$$\Theta_j = \sum_{i=0}^{j-1} \bigl(H_i(Z_i) + 1\bigr),$$

where $\{H_i(x)\}_{i\in\mathbb{N},x\in\mathbb{Z}}$ are independent random variables with μ_x being the law of $H_i(x)$ for each $i \in \mathbb{N}$. Setting $s_t = \max\{j : \Theta_j \le t\}$, define the RWREH $\{X_t\}$ by $X_t = Z_{s_t}$. In words, $\{X_t\}$ is a process which, when at site x, waits for a holding time distributed according to μ_x before, one unit of time later, jumping to one of its nearest neighbors, with jumps to the right occurring with probability ω_x. The environment $\bar{\omega}$ is chosen according to the probability measure P, and fixed thereafter. Let $M_1^s(\overline{\Omega}_\varepsilon)$ and $M_1^e(\overline{\Omega}_\varepsilon)$ denote the stationary, or stationary and ergodic, respectively, probability measures on $\overline{\Omega}_\varepsilon$, with respect to the shift $\theta : \overline{\Omega}_\varepsilon \to \overline{\Omega}_\varepsilon$ such that $(\theta\bar{\omega})_i = \bar{\omega}_{i+1}$. We will always assume that

(C0) $P \in M_1^e(\overline{\Omega}_\varepsilon)$ and $E_P(\log \mu_0([0, \kappa(\bar{\omega})])) > -\infty$ for some $\kappa(\bar{\omega})$ such that $E_P(\kappa(\bar{\omega})) < \infty$.

This condition on μ_0 is quite mild. For example, it is satisfied under the "ellipticity" condition that $\mu_0([0, \varepsilon^{-1}]) \ge \varepsilon$ for some $\varepsilon > 0$ and P-a.e. ω. It holds also in the absence of uniform ellipticity, for instance, if μ_x are atomic measures at unbounded h_x, provided $E_P(h_0) < \infty$ [take $\kappa(\bar{\omega}) = h_0$], or if μ_x are the laws of Exponential$(1/\gamma_x)$ random variables with $E_P(\log(1 + \gamma_0)) < \infty$ as in the model considered in [6].

We let $P_{\bar{\omega}}$ denote the law of the process $\{X_t\}$, conditioned on a realization $\bar{\omega} \in \overline{\Omega}_\varepsilon$ (the *quenched* law). We use $\mathbb{P}$ both for $P \times P_{\bar{\omega}}$ and for its marginal on $\mathbb{Z}^{\mathbb{R}_+}$ induced by $\{X_t\}_{t\ge 0}$, and refer to both as the *annealed* law.

The typical behavior of the RWREH is readily obtained, as in the case of the RWRE, by a hitting time decomposition. Define

$$T_n = \inf\{t \ge 0 : X_t = n\}, \qquad n \in \mathbb{Z}.$$

Using the same arguments as in [18], one has that

$$\frac{X_t}{t} \to v_P, \qquad \mathbb{P}\text{-a.e.},$$

where

$$v_P = \begin{cases} \dfrac{1}{\int E_{\bar{\omega}}(T_1)P(d\bar{\omega})}, & \displaystyle\int E_{\bar{\omega}}(T_1)P(d\bar{\omega}) < \infty, \\ -\dfrac{1}{\int E_{\bar{\omega}}(T_{-1})P(d\bar{\omega})}, & \displaystyle\int E_{\bar{\omega}}(T_{-1})P(d\bar{\omega}) < \infty, \\ 0, & \text{otherwise}. \end{cases} \tag{1.1}$$

Our interest lies in obtaining *large deviations* results, both quenched and annealed, for the RWREH. For the definition of the large deviation principle (LDP) and for background we refer to [4]. For one-dimensional RWRE, large deviations were first derived in [8] for the quenched setting, and extended into an annealed LDP by Comets, Gantert and Zeitouni [1], who also provide a variational formula relating the annealed and quenched LDPs. For RWRE in $\mathbb{Z}^d$ the quenched LDP (in the so-called *nestling* situation) was derived in [19], while the quenched LDP without a nestling assumption and the annealed LDP were recently obtained by Varadhan [17].

Our interest in the large deviations for the RWREH originated from three different sources:

1. In [2], we considered large deviations for random walks on Galton–Watson trees. We showed that, in contrast to RWRE on $\mathbb{Z}$ and in contrast to the conjectured behavior of RWRE on $\mathbb{Z}^d$, the quenched and annealed large deviation rate functions for the random walk on Galton–Watson trees coincide. We conjectured in [2] that restricting attention to a particular ray in the tree, one should recuperate the differences between quenched and annealed behavior. In Section 5 we show using our analysis of the RWREH that this is indeed the case.
2. In [1], the large deviations for the RWRE, both quenched and annealed, are considered. While preparing the notes in [18], we noted that some of the proofs do not carry over to the setup where holding times are present. Addressing this issue here, we substantially modify those parts of the proof in [1] that relied on "worst case domination." Even in the context of the standard RWRE, these new proofs have, we believe, an independent interest.
3. In [6], the authors considered a model of simple random walk on $\mathbb{Z}$ with heavy-tailed random holding times and proved that the suitably rescaled process converges to a singular diffusion. Thus, already the presence of random holding times causes a nonstandard behavior. This led us naturally to consider the more general RWREH model, where both random holding times and random drifts are present. Our assumptions on the environment, at least in the quenched setting, allow us to derive LDPs for the model of [6].

Our main goal in this paper is to study the large deviations, both quenched and annealed, of X_t/t. In doing so, we follow the basic strategy of [1]: study hitting times and then relate deviations of hitting times to deviations of the walk. More precisely, we first study the quenched large deviations of T_n/n proving that its rate function is deterministic and, in the case $\limsup_{n\to\infty} Z_n = \infty$, P-a.e., can be written as the Legendre transform of the average of the quenched logarithmic generating function of the hitting time T_1. Then, using space reversal invariance [cf. (3.12)], we obtain the quenched large deviations of T_{-n}/n (Theorem 1). The proofs in this part of the paper closely follow those in [1].

The next step involves the derivation of annealed LDPs for T_n/n (see Theorem 2). As in [1], a crucial element of the proof is the use of Varadhan's lemma to relate the quenched and annealed limits of normalized logarithmic moment generating functions (often called in this context *Lyapunov exponents*). This forces us to impose the additional restrictions (C1)–(C3) on the environment. It is here that our proofs first significantly depart from the proofs in [1]. Domination by a "worst case environment" and explicit computations provide there the integrability needed for Varadhan's lemma. Since such domination does not exist for the RWREH, due to the interaction between holding times and local drifts, we take a different approach here.

The final step in the derivation of the LDP is to transform estimates on deviations of T_n into estimates on deviations for X_t (Theorem 3 for the quenched setup and Theorem 4 for the annealed one). One direction of this transformation is trivial: namely, $\{X_t > tx\} \subseteq \{T_{\lfloor tx \rfloor} < t\}$. The other direction requires a further departure from the proofs of [1] in the absence of coupling with a worst case environment. The heart of the matter is Lemma 4, dealing with the rate of decay of the probability of the walk to backtrack *after a large time has elapsed*.

Having described our general strategy, we turn to state our results. To this end, set

$$\varphi(\lambda, \bar{\omega}) = E_{\bar{\omega}}(e^{\lambda T_1} \mathbf{1}_{T_1 < \infty}), \qquad f(\lambda, \bar{\omega}) = \log \varphi(\lambda, \bar{\omega}),$$
$$G(\lambda, P, u) = \lambda u - E_P(f(\lambda, \bar{\omega})),$$

and define $I_P^{\tau,q}(u) = \sup_{\lambda \in \mathbb{R}} G(\lambda, P, u)$. In the same way, set

$$\varphi^-(\lambda, \bar{\omega}) = E_{\bar{\omega}}(e^{\lambda T_{-1}} \mathbf{1}_{T_{-1} < \infty}), \qquad f^-(\lambda, \bar{\omega}) = \log \varphi^-(\lambda, \bar{\omega}),$$
$$G^-(\lambda, P, u) = \lambda u - E_P(f^-(\lambda, \bar{\omega})),$$

and define $I_P^{-\tau,q}(u) = \sup_{\lambda \in \mathbb{R}} G^-(\lambda, P, u)$.

THEOREM 1 (Quenched LDP for T_n/n). *Assume* (C0). *For P-almost every $\bar{\omega}$, the sequence T_n/n satisfies a weak LDP in $\mathbb{R}$ under $P_{\bar{\omega}}$ with the convex rate function $I_P^{\tau,q}(u)$, and the sequence T_{-n}/n satisfies a weak LDP in $\mathbb{R}$ under $P_{\bar{\omega}}$ with the convex rate function $I_P^{-\tau,q}(u)$. Further,*

$$I_P^{\tau,q}(u) = I_P^{-\tau,q}(u) + E_P(\log \rho_0), \tag{1.2}$$

where $\rho_x = (1 - \omega_x)/\omega_x$.

An annealed LDP for T_n/n requires additional notation and assumptions on P. Equip $M_1^{\varepsilon}(\overline{\mathbb{R}}_+)$ with the topology induced by weak convergence and S_ε with the corresponding product topology. Putting on $\overline{\Omega}_\varepsilon$ the product topology and on $M_1(\overline{\Omega}_\varepsilon)$ the corresponding topology of weak convergence, we see that S_ε, $\overline{\Omega}_\varepsilon$ and $M_1(\overline{\Omega}_\varepsilon)$ are compact metric spaces. Hereafter $\eta|_m$ denotes the restriction of $\eta \in M_1(\overline{\Omega}_\varepsilon)$ to $\{\bar{\omega}_i\}_{i=0}^{m-1}$. We say that $\eta \in M_1(\overline{\Omega}_\varepsilon)$ is *locally equivalent to the product of its marginals* if, for any $A \in S_\varepsilon^m$ and m finite, $\eta|_m(A) = 0$ if and only if $(\eta|_1)^m(A) = 0$. We consider the following assumptions:

(C1) The empirical process $R_n = n^{-1}\sum_{j=0}^{n-1}\delta_{\theta^j\bar\omega}$ satisfies under P the LDP in $M_1(\overline{\Omega}_\varepsilon)$ with good rate function $h(\cdot|P)$. Here we assume that the specific entropy $h(\eta|P)=\lim_{m\to\infty} m^{-1}H(\eta|_m|P|_m)$ with respect to P exists for any stationary η, and set $h(\eta|P)=\infty$ for nonstationary η.

(C2) P is locally equivalent to the product of its marginals. Moreover, for any stationary measure $\eta\in M_1(\overline{\Omega}_\varepsilon)$ with $h(\eta|P)<\infty$, there is a sequence $\{\eta^n\}\subset M_1^e(\overline{\Omega}_\varepsilon)$ with $\eta^n\to\eta$ and $h(\eta^n|P)\to h(\eta|P)$, such that $\eta^n|_1=\eta|_1$ for all n. There also exists a sequence of measures η^n that are locally equivalent to the product of their marginals, having all these properties, except possibly $\eta^n|_1\neq\eta|_1$.

(C3) There exist a nonrandom $b<\infty$ and a function $k(\cdot)>0$, such that P-a.e. $\mu_0([0,b))=0$ and $\mu_0([0,b+\delta])\ge k(\delta)$ for all $\delta>0$.

As noted for example in Theorems 3.10 and 4.1 and Lemma 4.8 of [5], the conditions (C1) and (C2) hold if the stationary and ergodic P corresponds to a Markov process with transition kernel $P(\bar\omega_{x+1}|\bar\omega_x)$ whose Radon–Nikodym derivative with respect to some reference probability measure on S_ε is bounded and bounded away from 0 [in particular, (C1) and (C2) hold if P is a stationary product measure]. Note that (C3) is a quantitative version of a "uniform ellipticity" condition referring to the holding times distributions.

We now have the following theorem.

THEOREM 2 (Annealed LDP for T_n/n). *Assume* (C0)–(C3). *Then the sequence T_n/n satisfies a weak LDP in $\mathbb{R}$ under $\mathbb{P}$ with the convex rate function*

$$I_P^{\tau,a}(u)=\inf_{\eta\in M_1^e(\overline{\Omega}_\varepsilon)}\bigl[I_\eta^{\tau,q}(u)+h(\eta|P)\bigr],$$

and the sequence T_{-n}/n satisfies a weak LDP in $\mathbb{R}$ under $\mathbb{P}$ with the convex rate function

$$I_P^{-\tau,a}(u)=\inf_{\eta\in M_1^e(\overline{\Omega}_\varepsilon)}\bigl[I_\eta^{-\tau,q}(u)+h(\eta|P)\bigr].$$

We next state the large deviations of the rescaled positions X_t/t.

THEOREM 3 (Quenched LDP for X_t/t). *Assume* (C0). *For P-almost every $\bar\omega$, X_t/t satisfies an LDP under $P_{\bar\omega}$ with the good convex rate function $I_P^q(\cdot)$:*

(a) *If $P(\mu_0(\{\infty\})>0)=0$, then*

$$I_P^q(v)=\begin{cases} vI_P^{\tau,q}\left(\dfrac{1}{v}\right), & v>0,\\[2ex] |v|I_P^{-\tau,q}\left(\dfrac{1}{|v|}\right), & v<0,\end{cases}\tag{1.3}$$

and $I_P^q(0):=\lim_{v\to 0}I_P^q(v)$.

(b) *If* $P(\mu_0(\{\infty\}) > 0) > 0$, *then*

$$
(1.4)\qquad I_P^q(v) = \begin{cases} \inf_{\ell\in[0,1]} vI_P^{\tau,q}\left(\frac{\ell}{v}\right), & v > 0, \\ \inf_{\ell\in[0,1]} |v|I_P^{-\tau,q}\left(\frac{\ell}{|v|}\right), & v < 0, \\ 0, & v = 0. \end{cases}
$$

The corresponding annealed statement for the positions X_t/t follows.

THEOREM 4 (Annealed LDP for X_t/t). *Assume* (C0)–(C3).

(a) *If* $P(\mu_0(\{\infty\}) > 0) = 0$, *then* X_t/t *satisfies an LDP under* $\mathbb{P}$ *with the good convex rate function* I_P^a, *where*

$$
(1.5)\qquad I_P^a(v) = \begin{cases} vI_P^{\tau,a}\left(\frac{1}{v}\right), & v > 0, \\ |v|I_P^{-\tau,a}\left(\frac{1}{|v|}\right), & v < 0, \end{cases}
$$

and $I_P^a(0) := \lim_{v\to 0} I_P^a(v)$.

(b) *If* $P(\mu_0(\{\infty\}) > 0) > 0$, *assume further that, for some* $c < \infty$, $k_0 < \infty$ *and* P*-almost-every* $\bar\omega$,

$$
(1.6)\qquad \max_{1\le j\le k} E_P\big(\mu_j(\{\infty\})|\mathcal{F}_{-k}^-\big) \ge e^{-ck} \qquad \forall k \ge k_0,
$$

where $\mathcal{F}_m^- = \sigma(\{\bar\omega_x, x \le m\})$. *Then* X_t/t *satisfies an LDP under* $\mathbb{P}$ *with the good convex rate function*

$$
(1.7)\qquad I_P^a(v) = \begin{cases} \inf_{\ell\in[0,1]} vI_P^{\tau,a}\left(\frac{\ell}{v}\right), & v > 0, \\ \inf_{\ell\in[0,1]} |v|I_P^{-\tau,a}\left(\frac{\ell}{|v|}\right), & v < 0, \\ 0, & v = 0. \end{cases}
$$

Clearly, (1.6) holds if P is a stationary product measure (and, more generally, under suitable mixing conditions).

We conclude with a discussion of the resulting rate functions. An advantage of our approach is that it yields a variational formula linking annealed and quenched rate functions (see the statement in Theorem 2) with intuitive appeal: the annealed rate function balances the exponential cost of modifying the environment, measured by an entropy term, and the quenched rate function in the new environment.

A detailed study of the properties of the rate functions for the RWRE appears in [1]. A good part of it can be transferred to the context of RWREH but we will not do so, in order to avoid boring the reader. Nevertheless, the following information on the rate functions which is immediate from our analysis, is worth noting.

PROPOSITION 1. *Assume* (C0) *for the quenched statements and* (C0)–(C3) *for the annealed ones*:

(a) *If* $P(\mu_0(\{\infty\}) > 0) = 0$ (*i.e., holding times are finite* P*-a.e.*), *then* $I_P^q(\cdot)$ *and* $I_P^a(\cdot)$ *can only vanish on the interval between* 0 *and* v_P, *and they do vanish there if* $\lambda_{\rm crit}(P) = 0$ (*see definition in Lemma* 1). *If* $\lambda_{\rm crit}(P) > 0$, *then the above mentioned rate functions vanish only at* v_P.

(b) *If* $P(\mu_0(\{\infty\}) > 0) > 0$ (*i.e., holding times are infinite with positive probability*), *then the rate functions* $I_P^q(\cdot)$ *and* $I_P^a(\cdot)$ *vanish only at the origin. If* $\bar{u}(P) < \infty$ [*see definition in* (2.3)], *then the quenched rate function is piecewise linear in a neighborhood of the origin*.

REMARK. For i.i.d. environments and a.e. finite holding times the shape of the rate functions for RWREH is similar to that of the RWRE: any nestling walk [i.e., an environment for which 0 is in the convex hull of supp$(2\omega_0 - 1)$] has $\lambda_{\rm crit}(P) = 0$ by a comparison with the embedded RWRE Z_n. Consequently, it exhibits subexponential rate of decay of slowdown probabilities if $v_P \neq 0$. In contrast with the RWRE, here one may have subexponential rate of decay of slowdown probabilities even for a nonnestling walk by having holding times with infinite exponential moments. Further, for i.i.d. environments with possibly infinite holding times, we may find the rate function vanishing only at 0 with linear pieces on both sides of 0, a situation that cannot occur in the RWRE setup.

The structure of the article is as follows: in the next section, we study key properties of the rate functions, leading to Propositions 1–3. Applying Propositions 2 and 3, we prove in Section 3 our hitting time results, Theorems 1 and 2. Section 4 provides the proofs of Theorems 3 and 4, our LDPs for the rescaled position. Throughout these sections we emphasize those elements of the proofs that differ from [1]. Section 5 is devoted to the statement and proof of our results concerning the (biased) random walk on a Galton–Watson tree. Open problems and discussion appear in Section 6.

2. Properties of the rate functions.

We begin with the following strengthening of [1], Lemma 2:

LEMMA 1. *For any* $P \in M_1^e(\overline{\Omega}_\varepsilon)$ *there exist constants* $\lambda_{\rm crit} = \lambda_{\rm crit}(P)$, $\lambda_{\rm crit}^- = \lambda_{\rm crit}^-(P) \in [0, \infty)$ *such that, for* P*-a.e.* $\bar{\omega}$,

$$\varphi(\lambda, \bar{\omega}) \begin{cases} \leq \varepsilon^{-2}, & \lambda \leq \lambda_{\rm crit}, \\ = \infty, & \lambda > \lambda_{\rm crit}, \end{cases} \qquad \varphi^-(\lambda, \bar{\omega}) \begin{cases} \leq \varepsilon^{-2}, & \lambda \leq \lambda_{\rm crit}^-, \\ = \infty, & \lambda > \lambda_{\rm crit}^-. \end{cases} \tag{2.1}$$

We will see later (see Remark 1) that $\lambda_{\rm crit}^-(P) = \lambda_{\rm crit}(P)$.

PROOF OF LEMMA 1. By the transformation $\{(\omega_x,\mu_x)\}\to\{(1-\omega_x,\mu_x)\}$, it is enough to consider $\varphi(\lambda,\bar\omega)$. By path decomposition, for each λ,

$$\begin{aligned}\varphi(\lambda,\bar\omega)&=\omega_0 e^{\lambda}E_{\mu_0}(e^{\lambda H}\mathbf{1}_{H<\infty})\\&\qquad+(1-\omega_0)E_{\mu_0}(e^{\lambda H}\mathbf{1}_{H<\infty})e^{\lambda}\varphi(\lambda,\theta^{-1}\bar\omega)\varphi(\lambda,\bar\omega),\end{aligned}\tag{2.2}$$

where H is a random variable with distribution μ_0, and E_{μ_0} denotes expectation with respect to μ_0. Thus, $\varphi(\lambda,\bar\omega)<\infty$ implies that $\varphi(\lambda,\theta^{-1}\bar\omega)<\infty$, yielding by the ergodicity of P that $\mathbf{1}_{\varphi(\lambda,\bar\omega)<\infty}$ is constant P-a.e., for all λ rational at once. This, and the monotonicity of $\varphi(\lambda,\bar\omega)$ in λ, immediately yield the existence of a deterministic λ_{crit} (possibly $\lambda_{\text{crit}}=\infty$), with $\varphi(\lambda,\bar\omega)<\infty$ for all $\lambda<\lambda_{\text{crit}}$, P-a.e. By definition, $\varphi(\lambda,\bar\omega)\le 1$ for $\lambda\le 0$, whereas for $\lambda\ge 0$, the fact that $\varphi(\lambda,\bar\omega)<\infty$ implies by (2.2) that

$$\varphi(\lambda,\theta^{-1}\bar\omega)\le\frac{1}{(1-\omega_0)E_{\mu_0}(\mathbf{1}_{H<\infty})e^{\lambda}}\le\frac{e^{-\lambda}}{\varepsilon^2},\qquad P\text{-a.e.}$$

We conclude that $\varphi(\lambda,\bar\omega)\le\varepsilon^{-2}$ for all $\lambda<\lambda_{\text{crit}}$, P-a.e. Since

$$E_{\bar\omega}[e^{\lambda T_1}\mathbf{1}_{T_1<\infty}]\ge e^{\lambda}P_{\bar\omega}(1\le T_1<\infty)\ge\omega_0\mu_0(\mathbb{R}_+)e^{\lambda}\ge\varepsilon^2e^{\lambda}\qquad P\text{-a.e.},$$

and with $\varphi(\lambda,\bar\omega)$ uniformly bounded on $(-\infty,\lambda_{\text{crit}})$, it thus follows that $\lambda_{\text{crit}}<\infty$ and by monotone convergence, also $\varphi(\lambda_{\text{crit}},\bar\omega)\le\varepsilon^{-2}<\infty$ for P-a.e. $\bar\omega$. □

Set

$$\bar u=\bar u(P)=\int\frac{E_{\bar\omega}(T_1\mathbf{1}_{T_1<\infty})}{E_{\bar\omega}(\mathbf{1}_{T_1<\infty})}P(d\bar\omega)\in[1,\infty].\tag{2.3}$$

Since $\lambda\mapsto f(\lambda,\bar\omega)=\log E_{\bar\omega}(e^{\lambda T_1}|T_1<\infty)+\log P_{\bar\omega}(T_1<\infty)$ is convex and finite for $\lambda<\lambda_{\text{crit}}$ and P-a.e. $\bar\omega$, it follows that

$$g(\lambda):=\int\frac{E_{\bar\omega}(T_1e^{\lambda T_1}\mathbf{1}_{T_1<\infty})}{E_{\bar\omega}(e^{\lambda T_1}\mathbf{1}_{T_1<\infty})}P(d\bar\omega)=\int\frac{d}{d\lambda}f(\lambda,\bar\omega)P(d\bar\omega),\tag{2.4}$$

is a nonnegative, nondecreasing function. By (2.2) and (C0) we have that, for any $\lambda\le 0$ and P-a.e. $\bar\omega$,

$$\varphi(\lambda,\bar\omega)\ge\omega_0e^{\lambda(1+\kappa(\bar\omega))}\mu_0([0,\kappa(\bar\omega)]),$$

implying that, for some $\alpha<\infty$ and all $\lambda\in\mathbb{R}$,

$$\begin{aligned}E_P(f(\lambda,\bar\omega))&\ge\log\varepsilon+(1+E_P(\kappa))\min(0,\lambda)+E_P(\log\mu_0([0,\kappa(\bar\omega)]))\\&\ge-\alpha(1+|\lambda|).\end{aligned}\tag{2.5}$$

In view of Lemma 1, it follows that $E_P(|f(\lambda,\bar\omega)|)<\infty$ hence $g(\lambda)=\frac{d}{d\lambda}E_P(f(\lambda,\bar\omega))<\infty$ and

$$u_-:=u_-(P)=\lim_{\lambda\searrow-\infty}g(\lambda)=E_P(\inf\{u\ge 0:\mu_0([0,u])>0\})+1<\infty.$$

Clearly, $u_+ = u_+(P) = \lim_{\lambda \nearrow \lambda_{\text{crit}}} g(\lambda) \geq \bar{u} = g(0)$ exists (with possible value $u_+ = +\infty$). Since $g(\lambda)$ is strictly increasing and continuous in λ, we see that, for any $u \in (u_-, u_+)$, there exists a unique $\lambda_u \in (-\infty, \lambda_{\text{crit}})$ such that $g(\lambda_u) = u$. Further, if $u < \bar{u}$, then $\lambda_u < 0$, and hence

$$I_P^{\tau,q}(u) = \sup_{\lambda \in \mathbb{R}} G(\lambda, P, u) = G(\lambda_u, P, u) = \sup_{\lambda \leq 0} G(\lambda, P, u), \qquad u \leq \bar{u}. \tag{2.6}$$

For $u \geq u_+$ we have that $\sup_{\lambda \in \mathbb{R}} G(\lambda, P, u) = G(\lambda_{\text{crit}}, P, u)$, whereas $\lambda_u > 0$ if $u_+ > u > \bar{u}$, hence also

$$I_P^{\tau,q}(u) = \sup_{\lambda \in \mathbb{R}} G(\lambda, P, u) = \sup_{\lambda \geq 0} G(\lambda, P, u), \qquad u \geq \bar{u}. \tag{2.7}$$

Further, we have the following.

PROPOSITION 2. *For any P satisfying* (C0), *the convex rate function $I_P^{\tau,q}(\cdot)$ is infinite on $(-\infty, u_-(P))$, finite on $(u_-(P), \infty)$, nonincreasing on $[u_-(P), \bar{u}(P)]$ and nondecreasing on $[\bar{u}(P), \infty)$. Moreover, if $\bar{u}(P) < \infty$, then $I_P^{\tau,q}(\bar{u}(P)) = G(0, P, \bar{u}(P))$, while if $\bar{u}(P) = \infty$, then $\lambda_{\text{crit}}(P) = 0$. Further, for all u,*

$$\sup_{\lambda \geq 0} G(\lambda, P, u) = \inf_{w \geq u} I_P^{\tau,q}(w) \tag{2.8}$$

and

$$\sup_{\lambda \leq 0} G(\lambda, P, u) = \inf_{w \leq u} I_P^{\tau,q}(w). \tag{2.9}$$

PROOF. From the definition we see that $I_P^{\tau,q}$ is convex and lower semicontinuous. Since $G(0, P, u) \geq 0$, it is also nonnegative. Suppose $u_1 < u_2 \leq \bar{u}(P)$. Then, by (2.6),

$$\begin{aligned}\sup_{\lambda \in \mathbb{R}} G(\lambda, P, u_1) &= \sup_{\lambda \leq 0} G(\lambda, P, u_1) = \sup_{\lambda \leq 0} \bigl[\lambda u_1 - E_P\bigl(f(\lambda, \bar{\omega})\bigr)\bigr] \\ &\geq \sup_{\lambda \leq 0} \bigl[\lambda u_2 - E_P\bigl(f(\lambda, \bar{\omega})\bigr)\bigr] = \sup_{\lambda \in \mathbb{R}} G(\lambda, P, u_2).\end{aligned}$$

To see that $I_P^{\tau,q}$ is nondecreasing on $[\bar{u}(P), \infty)$, use a similar argument with (2.7) instead of (2.6).

If $u < u_-(P)$, then $\sup_\xi \{u - g(\xi)\} < 0$, and since $G(0, P, u) \geq 0$ we see that $G(\lambda, P, u) = G(0, P, u) + \int_0^\lambda (u - g(\xi))\, d\xi \to \infty$ if $\lambda \to -\infty$, resulting with $I_P^{\tau,q}(u) = \infty$. In contrast, setting $\lambda_u := \lambda_{\text{crit}}$ if $u \geq u_+(P)$, it follows by (2.5) that $I_P^{\tau,q}(u) = G(\lambda_u, P, u) \leq \lambda_u u + \alpha(1 + |\lambda_u|) < \infty$ for any $u > u_-(P)$.

Recall that $P_{\bar{\omega}}(T_1 < \infty) \geq \varepsilon^2$ for all $\bar{\omega}$, and by Jensen's inequality we have that for all $\lambda \in \mathbb{R}$,

$$\begin{aligned}\lambda \int E_{\bar{\omega}}(T_1 | T_1 < \infty) P(d\bar{\omega}) &\leq E_P\bigl(\log E_{\bar{\omega}}(e^{\lambda T_1} | T_1 < \infty)\bigr) \\ &= E_P\bigl(f(\lambda, \bar{\omega})\bigr) - E_P\bigl(f(0, \bar{\omega})\bigr).\end{aligned}$$

If $\bar{u}(P) < \infty$, this implies that $G(\lambda, P, \bar{u}(P))$ is maximal at $\lambda = 0$, hence $I_P^{\tau,q}(\bar{u}(P)) = G(0, P, \bar{u}(P))$, whereas if $\bar{u}(P) = \infty$, then $E_P(f(\lambda, \bar{\omega})) = \infty$ for all $\lambda > 0$, hence $\lambda_{\text{crit}}(P) = 0$ by Lemma 1.

Turning to prove (2.8) and (2.9), consider first $\bar{u}(P) = \infty$, in which case $I_P^{\tau,q}(\cdot)$ is nonincreasing and (2.9) follows from (2.6). Further, the convex, lower semicontinuous function $\lambda \mapsto E_P(f(\lambda, \bar{\omega}))$ is then infinite if $\lambda > 0$. Hence, by duality of Fenchel–Legendre transforms, for all u,

$$\inf_{w \geq u} I_P^{\tau,q}(w) = \inf_{w \in \mathbb{R}} I_P^{\tau,q}(w) = -E_P\bigl(f(0, \bar{\omega})\bigr) = G(0, P, u) = \sup_{\lambda \geq 0} G(\lambda, P, u),$$

which amounts to (2.8). Suppose now that $\bar{u}(P) < \infty$. Since $I_P^{\tau,q}(\cdot)$ is nondecreasing on $[\bar{u}(P), \infty)$ we get (2.8) for $u \geq \bar{u}(P)$ out of (2.7). Moreover, $I_P^{\tau,q}(u)$ is nonincreasing for $u \leq \bar{u}(P)$; hence for such u the right-hand side of (2.8) equals $I_P^{\tau,q}(\bar{u}(P)) = G(0, P, \bar{u}(P))$. Further, then $G(\lambda, P, u) \leq G(\lambda, P, \bar{u}(P))$ for all $\lambda \geq 0$, with equality if $\lambda = 0$, implying the left-hand side of (2.8) also equals $G(0, P, \bar{u}(P))$, thus completing its proof. The proof of (2.9) is similar. Combining (2.6) with the monotonicity of $I_P^{\tau,q}(u)$ gives (2.9) for $u \leq \bar{u}(P)$, whereas for $u \geq \bar{u}(P)$ both sides of (2.9) equal $G(0, P, \bar{u}(P))$. $\square$

Turning to the study of the annealed rate functions, we begin with a lemma giving a characterization of $\lambda_{\text{crit}}(\eta)$ for "nice" η. The lemma corresponds to [1], Lemma 4, but in contrast to [1], Lemma 4, its proof does not use domination and explicit computations, which are not available here.

LEMMA 2. *Assume $\eta \in M_1^e(\overline{\Omega}_\varepsilon)$ is locally equivalent to the product of its marginals. Let $\Sigma := \operatorname{supp} \eta|_1 \subseteq S_\varepsilon$. Then*

$$\lambda_{\text{crit}}(\eta) = \inf_{\bar{\omega} \in \Sigma^{\mathbb{Z}}} \lambda_c(\bar{\omega}) =: \bar{\lambda} \geq 0, \tag{2.10}$$

where $\lambda_c(\bar{\omega}) := \sup\{\lambda : E_{\bar{\omega}}(e^{\lambda T_1} \mathbf{1}_{T_1 < \infty}) < \infty\}$. Moreover,

$$\varphi(\lambda, \bar{\omega}) \leq \varepsilon^{-2} \qquad \forall \lambda \leq \lambda_{\text{crit}}(\eta)\ \forall \bar{\omega} \in \Sigma^{\mathbb{Z}}. \tag{2.11}$$

Suppose η^n is a sequence in $M_1^e(\overline{\Omega}_\varepsilon)$ such that $\eta^n(\Sigma^{\mathbb{Z}}) = 1$ and all the η^n are locally equivalent to the product of their marginals. If $\eta^n \to \widehat{\eta}$ for some $\widehat{\eta} \in M_1(\overline{\Omega}_\varepsilon)$ such that $\eta \ll \widehat{\eta}$, then

$$\lambda_{\text{crit}}(\eta^n) \to \lambda_{\text{crit}}(\eta). \tag{2.12}$$

PROOF. Let $g_m(t) := \min(m - t, 1)\mathbf{1}_{[0,m]}(t)$ and

$$\varphi_m(\lambda, \bar{\omega}) := E_{\bar{\omega}}\bigl(e^{\lambda T_1} g_m(T_1)\bigr). \tag{2.13}$$

Note that $\varphi_m(\lambda, \bar{\omega})$ is continuous on $\overline{\Omega}_\varepsilon$. Indeed, $\varphi_m(\lambda, \bar{\omega})$ depends only on $(\bar{\omega}_0, \bar{\omega}_{-1}, \ldots, \bar{\omega}_{-m+1})$. Moreover, it is the sum over the contributions $\varphi_m(\lambda, \bar{\omega}, \mathbf{z})$

of the finitely many possible paths $\mathbf{z}$ of the RWRE $Z_i = z_i$, $i = 0, \dots, j$, where $z_0 = 0$, $z_j = 1$ and $z_i \le 0$ for $i < j \le m$. Fixing such a path, denote by $\widehat{\mu}_{\mathbf{z}}$ the law of $T_1 = \Theta_j$ conditional on $\{Z_0, Z_1, \dots, Z_j\}$. With $t \mapsto e^{\lambda t} g_m(t)$ bounded and continuous on $\mathbb{R}_+$,

$$\varphi_m(\lambda, \bar{\omega}, \mathbf{z}) = \prod_{i=0}^{j-1} \left[\tfrac{1}{2} + (z_{i+1} - z_i)\left(\omega_{z_i} - \tfrac{1}{2}\right)\right] \int_0^\infty e^{\lambda t} g_m(t) \widehat{\mu}_{\mathbf{z}}(dt),$$

is continuous in $(\omega_0, \omega_{-1}, \dots, \omega_{-m+1})$ and $\widehat{\mu}_{\mathbf{z}}$, hence also in $\{\bar{\omega}_x, x = 0, \dots, -m+1\}$.

Fixing $\lambda \le \lambda_{\mathrm{crit}}(\eta)$ we know from Lemma 1 that (2.11) holds for η-a.e. $\bar{\omega}$. We next show that (2.11) holds for *all* $\bar{\omega} \in \Sigma^{\mathbb{Z}}$. Suppose to the contrary that $\varphi(\lambda, \widetilde{\omega}) > \varepsilon^{-2}$ for some $\widetilde{\omega} \in \Sigma^{\mathbb{Z}}$. By monotone convergence and continuity of φ_m, there exists m large enough such that the open subset $G := \{(\bar{\omega}_0, \dots, \bar{\omega}_{-m+1}) : \varphi_m(\lambda, \bar{\omega}) > \varepsilon^{-2}\}$ of S_ε^m intersects $\mathrm{supp}(\eta|_1)^m$ at $(\widetilde{\omega}_0, \dots, \widetilde{\omega}_{-m+1})$. Clearly $(\eta|_1)^m(G) > 0$, and with η locally equivalent to the product of its marginals, also $\eta|_m(G) > 0$. Recall that $\varphi(\lambda, \bar{\omega}) \ge \varphi_m(\lambda, \bar{\omega})$, implying that $\eta(\varphi(\lambda, \bar{\omega}) > \varepsilon^{-2}) > 0$ in contradiction with Lemma 1.

If $\lambda < \bar{\lambda}$, then by definition $\varphi(\lambda, \bar{\omega}) < \infty$ for all $\bar{\omega} \in \Sigma^{\mathbb{Z}}$; hence $\lambda_{\mathrm{crit}}(\eta) \ge \lambda$ by Lemma 1. Consequently, $\lambda_{\mathrm{crit}}(\eta) \ge \bar{\lambda}$. For any $\bar{\omega} \in \Sigma^{\mathbb{Z}}$, the inequality (2.11) implies that $\lambda_c(\bar{\omega}) \ge \lambda_{\mathrm{crit}}(\eta)$; hence by definition also $\bar{\lambda} \ge \lambda_{\mathrm{crit}}(\eta)$.

Turning to prove (2.12), note that as $\mathrm{supp}\, \eta^n|_1 \subseteq \Sigma$, we have from (2.10) that $\lambda_{\mathrm{crit}}(\eta^n) \ge \bar{\lambda} = \lambda_{\mathrm{crit}}(\eta)$ and if $\lambda > \lambda_{\mathrm{crit}}(\eta)$, then $\varphi(\lambda, \widetilde{\omega}) = \infty > \varepsilon^{-2}$ for some $\widetilde{\omega} \in \Sigma^{\mathbb{Z}}$. Taking m and the open $G \subseteq S_\varepsilon^m$ as in the preceding proof of (2.11), we have that $\eta|_m(G) > 0$. Since $\eta \ll \widehat{\eta}$ and $\eta^n \to \widehat{\eta}$, also $\eta^n|_m(G) > 0$ for all n large enough. Consequently, $\eta^n(\varphi(\lambda, \bar{\omega}) > \varepsilon^{-2}) > 0$, implying that $\lambda > \lambda_{\mathrm{crit}}(\eta^n)$ for all n large enough (cf. Lemma 1). Considering $\lambda \downarrow \lambda_{\mathrm{crit}}(\eta)$ completes the proof of (2.12). □

With $M_1^P = M_1^P(\overline{\Omega}_\varepsilon) := \{\nu \in M_1(\overline{\Omega}_\varepsilon) : \mathrm{supp}\, \nu \subseteq (\mathrm{supp}\, P|_1)^{\mathbb{Z}}\}$, the next lemma is the analogue of [1], Lemma 6. This is also where we use the "uniform ellipticity" condition (C3) on the holding time distributions.

LEMMA 3. *Suppose $P \in M_1^e(\overline{\Omega}_\varepsilon)$ satisfies* (C3) *and is locally equivalent to the product of its marginals. Then, the function $(\lambda, \nu) \mapsto \int f(\lambda, \bar{\omega}) \nu(d\bar{\omega})$ is continuous on $(-\infty, \lambda_{\mathrm{crit}}(P)) \times M_1^P$.*

PROOF. Let

$$\xi_m(\lambda, \nu) := \int \log \varphi_m(\lambda, \bar{\omega}) \nu(d\bar{\omega})$$

for the bounded, continuous function $\varphi_m(\lambda, \cdot)$ of (2.13). Note that $|\varphi_m(\lambda', \bar{\omega}) - \varphi_m(\lambda, \bar{\omega})| \to 0$ as $\lambda' \to \lambda$, uniformly in $\bar{\omega}$. Considering hereafter $m \ge b + 3$, we have by (C3) that

$$\varphi_m(\lambda, \bar{\omega}) \geq \omega_0 E_{\mu_0}\big(e^{\lambda(H+1)}\mathbf{1}_{H\leq m-2}\big) \geq \varepsilon e^{-|\lambda|(b+2)}k(1) =: \frac{1}{c_\lambda}. \tag{2.14}$$

The function $\xi_m(\lambda, \nu)$ is then continuous on $\mathbb{R} \times M_1(\overline{\Omega}_\varepsilon)$. By the inequality $\log x \leq x - 1$ and the preceding lower bound on φ_m we have that

$$\begin{aligned} 0 \leq \log\left(\frac{\varphi(\lambda, \bar{\omega})}{\varphi_m(\lambda, \bar{\omega})}\right) &\leq c_\lambda\big(\varphi(\lambda, \bar{\omega}) - \varphi_m(\lambda, \bar{\omega})\big) \leq c_\lambda E_{\bar{\omega}}\big(e^{\lambda T_1}\mathbf{1}_{\infty > T_1 > m-1}\big) \\ &\leq c_\lambda e^{(\lambda-\lambda_{\text{crit}}(P))(m-1)} E_{\bar{\omega}}\big(e^{\lambda_{\text{crit}}(P)T_1}\mathbf{1}_{T_1<\infty}\big). \end{aligned}$$

Fixing $\lambda < \lambda_{\text{crit}}(P)$ and $\bar{\omega} \in (\operatorname{supp} P|_1)^{\mathbb{Z}}$, we thus deduce from (2.11) that

$$0 \leq \log\varphi(\lambda, \bar{\omega}) - \log\varphi_m(\lambda, \bar{\omega}) \leq \varepsilon^{-2}c_\lambda e^{(\lambda-\lambda_{\text{crit}}(P))(m-1)}.$$

Hence, for any $\lambda < \lambda_{\text{crit}}(P)$ and $\nu \in M_1^P$, it holds that

$$\left|\int f(\lambda, \bar{\omega})\nu(d\bar{\omega}) - \xi_m(\lambda, \nu)\right| \leq \varepsilon^{-2}c_\lambda e^{(\lambda-\lambda_{\text{crit}}(P))(m-1)}.$$

The claimed continuity follows as $\xi_m(\cdot, \cdot)$ is continuous and $|\int f(\lambda, \bar{\omega})\nu(d\bar{\omega}) - \xi_m(\lambda, \nu)| \to 0$ for $m \to \infty$, uniformly in M_1^P. □

We next provide for $I_P^{\tau,a}(\cdot)$ representations analogous to those of Proposition 2.

PROPOSITION 3. *Assuming* (C0)–(C3), *let*

$$L(\lambda) := \sup_{\eta\in M_1^P}\left[\int f(\lambda, \bar{\omega})\eta(d\bar{\omega}) - h(\eta|P)\right]. \tag{2.15}$$

Then, for any $u \in \mathbb{R}$,

$$I_P^{\tau,a}(u) = \sup_{\lambda<\lambda_{\text{crit}}(P)}[\lambda u - L(\lambda)], \tag{2.16}$$

$$\inf_{w\leq u} I_P^{\tau,a}(w) = \sup_{\lambda<0}[\lambda u - L(\lambda)] \tag{2.17}$$

and if $\lambda_{\text{crit}}(P) > 0$, *also*

$$\inf_{w\geq u} I_P^{\tau,a}(w) = \sup_{0\leq\lambda<\lambda_{\text{crit}}(P)}[\lambda u - L(\lambda)]. \tag{2.18}$$

In particular, $I_P^{\tau,a}(\cdot)$ *is a convex rate function, and is nonincreasing if* $\lambda_{\text{crit}}(P) = 0$.

PROOF. Since $\lambda \mapsto \int f(\lambda, \bar{\omega})\eta(d\bar{\omega})$ is convex, nondecreasing for any $\eta \in M_1^P$, so is $\lambda \mapsto L(\lambda)$. Note that $L(\lambda) \geq \int f(\lambda, \bar{\omega})P(d\bar{\omega}) = \infty$ for any $\lambda > \lambda_{\text{crit}}(P)$ (see Lemma 1). In contrast, $\int f(\lambda, \bar{\omega})\eta(d\bar{\omega}) \leq -2\log\varepsilon$ for all $\lambda \leq \lambda_{\text{crit}}(P)$ and $\eta \in M_1^P$ [cf. (2.11)], implying that $L(\lambda)$ is finite and bounded for such λ.

Moreover, $\lambda \mapsto \int f(\lambda, \bar{\omega})\eta(d\bar{\omega})$ is continuous on $(-\infty, \lambda_{\text{crit}}(P)]$ for any $\eta \in M_1^P$ [by Lemma 3 in the case $\lambda < \lambda_{\text{crit}}(P)$, and by monotone convergence in the case $\lambda \uparrow \lambda_{\text{crit}}(P)$ since $\varphi(\lambda_{\text{crit}}(P), \bar{\omega}) \le \varepsilon^{-2}$ for all $\bar{\omega}$ in the support of η]. Therefore, $L(\cdot)$ is lower semicontinuous and its Fenchel–Legendre transform

$$J(u) := \sup_{\lambda \in \mathbb{R}} [\lambda u - L(\lambda)] = \sup_{\lambda < \lambda_{\text{crit}}(P)} [\lambda u - L(\lambda)],$$

is convex, lower semicontinuous [and if $\lambda_{\text{crit}}(P) = 0$, also nonincreasing]. Obviously, $J(u) = \infty$ for $u < 0$. We prove below that $I_P^{\tau,a}(\cdot) = J(\cdot)$. This is all we need if $\lambda_{\text{crit}}(P) = 0$, whereas if $\lambda_{\text{crit}}(P) > 0$, then $J(u) = \max(J_-(u), J_+(u))$ with $J_-(u) := \sup_{\lambda<0}[\lambda u - L(\lambda)]$ nonincreasing and $J_+(u) := \sup_{0 \le \lambda < \lambda_{\text{crit}}(P)}[\lambda u - L(\lambda)]$ nondecreasing. By duality of Fenchel–Legendre transforms $\inf_{u \in \mathbb{R}} J(u) = -L(0) \in [0, \infty)$. Moreover, considering $\lambda \to 0$ we see that $J_+(u) \ge -L(0)$ and $J_-(u) \ge -L(0)$ for all u. With $I_P^{\tau,a}(\cdot) = J(\cdot)$, we then easily get (2.17) and (2.18) out of (2.16).

Since $\eta \mapsto G(\lambda, \eta, u) + h(\eta|P)$ is convex, lower semicontinuous on the convex, compact set M_1^P, for any $\lambda < \lambda_{\text{crit}}$, and $\lambda \mapsto G(\lambda, \eta, u)$ is concave, continuous on $(-\infty, \lambda_{\text{crit}}(P)]$, by the min–max theorem (see [15], Theorem 4.2′), we conclude that

$$\begin{aligned} J(u) &= \inf_{\eta \in M_1^P} \sup_{\lambda < \lambda_{\text{crit}}(P)} [G(\lambda, \eta, u) + h(\eta|P)] \\ &= \sup_{\lambda < \lambda_{\text{crit}}(P)} [G(\lambda, \bar{\eta}, u) + h(\bar{\eta}|P)]. \end{aligned} \tag{2.19}$$

Here, $\bar{\eta}$ is a global minimizer of the lower semicontinuous function $\eta \mapsto h(\eta|P) + \sup_{\lambda < \lambda_{\text{crit}}(P)} G(\lambda, \eta, u)$ on the compact set M_1^P. Since $h(\eta|P) = \infty$ for all $\eta \notin M_1^P$, it follows from (2.19) that, for any $u \in \mathbb{R}$,

$$J(u) \le \inf_{\eta \in M_1^e(\overline{\Omega}_\varepsilon)} \sup_{\lambda \in \mathbb{R}} [G(\lambda, \eta, u) + h(\eta|P)] = I_P^{\tau,a}(u).$$

To show the converse inequality, we assume without loss of generality that $J(u) < \infty$ and approximate the stationary $\bar{\eta}$ of (2.19) by "nice" ergodic measures. To this end, note that (C3) implies that, for all $\lambda \le 0$, $\delta > 0$ and P-a.s.,

$$e^{\lambda b} \ge E_{\mu_0}(e^{\lambda H} \mathbf{1}_{H<\infty}) \ge k(\delta) e^{\lambda(b+\delta)}.$$

Since $T_1 \ge H_1(0) + 1$ with equality whenever $Z_1 = 1$, this implies that $f(\lambda, \bar{\omega}) - \lambda(b+1) \in [\log \varepsilon k(\delta) e^{\lambda\delta}, 0]$ hence also

$$\lambda(u - b - 1 - \delta) - \log(\varepsilon k(\delta)) \ge G(\lambda, \eta, u) \ge \lambda(u - b - 1) \tag{2.20}$$

for all $\eta \in M_1^P$ and $\lambda \le 0$. In particular, since $J(u) < \infty$, by (2.19) and (2.20) we know that $u \ge (b+1)$. Fixing $u = b + 1 + 2\delta$ and $\delta > 0$, it follows from (2.20) that

$$I_\eta^{\tau,q}(u) = \sup_{\lambda \ge -K} G(\lambda, \eta, u), \tag{2.21}$$

for $K = K_u = \delta^{-1}|\log(\varepsilon k(\delta))| < \infty$ and all $\eta \in M_1^P$. Let $\bar{\eta}_\ell = (1 - \frac{1}{\ell})\bar{\eta} + \frac{1}{\ell}P \in M_1^P$, noting that $h(\bar{\eta}_\ell|P) = (1 - \frac{1}{\ell})h(\bar{\eta}|P) < \infty$. By (C2), there exist $\eta_\ell^n \in M_1^e(\overline{\Omega}_\varepsilon)$ that are locally equivalent to the product of their marginals, with $\eta_\ell^n \to \bar{\eta}_\ell$ and $h(\eta_\ell^n|P) \to h(\bar{\eta}_\ell|P)$ as $n \to \infty$. Since $P \ll \bar{\eta}_\ell$, we see by (2.12) that $\lambda_{\text{crit}}(\eta_\ell^n) \to \lambda_{\text{crit}}(P)$ as $n \to \infty$. By a diagonalization argument, we thus have $\widetilde{\eta}_\ell \in M_1^e(\overline{\Omega}_\varepsilon) \cap M_1^P$, with

$$\widetilde{\eta}_\ell \to \bar{\eta}, \qquad h(\widetilde{\eta}_\ell|P) \to h(\bar{\eta}|P), \qquad \lambda_{\text{crit}}(\widetilde{\eta}_\ell) \to \lambda_{\text{crit}}(P).$$

In particular, for any $\xi > 0$ and ℓ large enough $G(\lambda, \widetilde{\eta}_\ell, u) = -\infty$ if $\lambda > \lambda_{\text{crit}}(P) + \xi \geq \lambda_{\text{crit}}(\widetilde{\eta}_\ell)$, implying together with (2.21) that

$$\begin{aligned} I_P^{\tau,a}(u) &\leq h(\widetilde{\eta}_\ell|P) + \sup_{-K \leq \lambda \leq \lambda_{\text{crit}}(P)+\xi} G(\lambda, \widetilde{\eta}_\ell, u) \\ &\leq h(\widetilde{\eta}_\ell|P) + 2\xi u + \sup_{-K \leq \lambda \leq \lambda_{\text{crit}}(P)-\xi} G(\lambda, \widetilde{\eta}_\ell, u) \\ &\leq h(\widetilde{\eta}_\ell|P) + 3\xi u + G(\widetilde{\lambda}_\ell, \widetilde{\eta}_\ell, u), \end{aligned}$$

for some $\widetilde{\lambda}_\ell \in [-K, \lambda_{\text{crit}}(P) - \xi]$. Passing to a subsequence if needed, $\widetilde{\lambda}_\ell \to \bar{\lambda} \in [-K, \lambda_{\text{crit}}(P) - \xi]$. Considering $\ell \to \infty$ we deduce by applying Lemma 3 for $(\widetilde{\lambda}_\ell, \widetilde{\eta}_\ell) \to (\bar{\lambda}, \bar{\eta})$, that

$$I_P^{\tau,a}(u) \leq G(\bar{\lambda}, \bar{\eta}, u) + h(\bar{\eta}|P) + 3\xi u \leq J(u) + 3\xi u$$

[the rightmost inequality follows from (2.19)]. Since $\xi > 0$ and $u > b + 1$ are arbitrary, the proof of (2.16) is thus complete, except possibly at $u = b+1$. Turning to deal with this remaining case, note that P-a.e. $T_1 \geq b + 1$ by (C3). Hence, by monotone convergence for any $\eta \in M_1^P$,

$$\begin{aligned} I_\eta^{\tau,q}(b+1) &= -\inf_{\lambda \in \mathbb{R}} \int \log E_{\bar{\omega}}(e^{\lambda(T_1-b-1)}\mathbf{1}_{T_1<\infty})\eta(d\bar{\omega}) \\ &= -\int \log[\omega_0\mu_0(\{b\})]\eta|_1(d\bar{\omega}_0). \end{aligned} \tag{2.22}$$

Since it suffices to consider $\lambda \to -\infty$ in (2.22), it follows from (2.19) that

$$J(b+1) = h(\bar{\eta}|P) - \int \log[\omega_0\mu_0(\{b\})]\bar{\eta}|_1(d\bar{\omega}_0) \tag{2.23}$$

[where both sides have value $+\infty$ if $\bar{\eta}(\mu_0(\{b\}) = 0) > 0$]. Assuming without loss of generality that $J(b + 1) < \infty$ and in particular that $h(\bar{\eta}|P) < \infty$, we have by (C2) a sequence $\eta^n \in M_1^e(\overline{\Omega}_\varepsilon)$ with $\eta^n|_1 = \bar{\eta}|_1$ for all n and $h(\eta^n|P) \to h(\bar{\eta}|P)$. Noting that for all n both $\eta^n \in M_1^P$ and

$$I_{\eta^n}^{\tau,q}(b+1) = -\int \log[\omega_0\mu_0(\{b\})]\bar{\eta}|_1(d\bar{\omega}_0),$$

by (2.22), we deduce from (2.23) that

$$I_P^{\tau,a}(b+1) \leq \liminf_{n\to\infty}\{I_{\eta^n}^{\tau,q}(b+1) + h(\eta^n|P)\} = J(b+1).$$

This concludes the proof of (2.16) and with it that of the proposition. □

We conclude this section with the proof of Proposition 1.

PROOF OF PROPOSITION 1. Throughout this proof we use λ_{crit}, u_-, $\bar{u}$ and u_+, for $\lambda_{\text{crit}}(P)$, $u_-(P)$, $\bar{u}(P)$ and $u_+(P)$, respectively.

By the discussion preceding (2.6) the nonnegative function $I_P^{\tau,q}(u)$ is strictly convex on (u_-, u_+). By Proposition 2, if $\lambda_{\text{crit}} > 0$, then $u_+ > \bar{u}$ and we have that $I_P^{\tau,q}(u) > 0$ for all $u \neq \bar{u}$. In contrast, if $\lambda_{\text{crit}} = 0$, either $I_P^{\tau,q}(u) > 0$ for all $u \in \mathbb{R}$, or $I_P^{\tau,q}(u) = 0$ if and only if $u \geq \bar{u}$. By (3.12) the same applies for the nonnegative rate function $I_P^{-\tau,q}(\cdot)$. Moreover, by (1.2), if $E_P(\log \rho_0) < 0$, then $I_P^{-\tau,q}$ is strictly positive while $I_P^{\tau,q}$ is strictly positive in case $E_P(\log \rho_0) > 0$.

When (C1)–(C3) also hold, recall that $\eta \mapsto h(\eta|P)$ is a good rate function that vanishes only at $\eta = P$. Combining in this case the variational formulas of Theorem 2 and the continuity of $\eta \mapsto E_\eta(f(\lambda, \bar{\omega}))$ and $\eta \mapsto E_\eta(f^-(\lambda, \bar{\omega}))$ [using (3.12) to deduce the latter from Lemma 3], we conclude that $I_P^{\tau,a}(u) = 0$ if and only if u is such that $I_P^{\tau,q}(u) = 0$, and $I_P^{-\tau,a}(u) = 0$ if and only if $I_P^{-\tau,q}(u) = 0$.

Dealing with part (a) of the proposition, suppose that $P(\mu_0(\{\infty\}) > 0) = 0$ and $E_P(\log \rho_0) \leq 0$, in which case $T_1 < \infty$, $\mathbb{P}$-a.e. Comparing (1.1) to (2.3) we see that $\bar{u} = 1/v_P$, implying that if $E_P(\log \rho_0) = 0$, then also $\bar{u} = \infty$ and $\lambda_{\text{crit}} = 0$ (see Proposition 2). As we show in (4.1) and in (4.18) both $I_P^q(0) = \lambda_{\text{crit}}$ and $I_P^a(0) = \lambda_{\text{crit}}$. In view of (1.3) and (1.5) we see that if $\lambda_{\text{crit}} > 0$, both good rate functions $I_P^q(v)$ and $I_P^a(v)$ vanish only at $v = v_P$, whereas they vanish at $v = 0$ if $\lambda_{\text{crit}} = 0$ and if in addition $v_P > 0$, they vanish also for all $v \in [0, v_P]$. The same consideration applies in the case $E_P(\log \rho_0) > 0$: here $T_{-1} < \infty$, $\mathbb{P}$-a.e., so that $\bar{u} = -1/v_P$ and if both $\lambda_{\text{crit}} = 0$ and $v_P < 0$, then the functions $I_P^q(v)$ and $I_P^a(v)$ vanish at the interval $[v_P, 0]$.

Turning to part (b) of the proposition, whereby $P(\mu_0(\{\infty\}) > 0) > 0$, note that then for all u,

$$I_P^{\tau,q}(u) \geq G(0, P, u) = -\log \mathbb{P}(T_1 < \infty) > 0,$$

$$I_P^{-\tau,q}(u) \geq G^-(0, P, u) = -\log \mathbb{P}(T_{-1} < \infty) > 0.$$

Thus, with all four rate functions $I_P^{\tau,q}$, $I_P^{-\tau,q}$, $I_P^{\tau,a}$ and $I_P^{-\tau,a}$ being strictly positive, it follows by (1.4) and (1.7) that I_P^q and I_P^a only vanish at the origin. If $\bar{u} < \infty$, then by Proposition 2 and (1.2), both functions $I_P^{\tau,q}$ and $I_P^{-\tau,q}$ have their (positive) global minimum at $\bar{u}$, resulting by (1.4) with linear pieces for I_P^q on $[-1/\bar{u}, 1/\bar{u}]$. □

3. Proof of the LDPs for hitting times T_n/n.

PROOF OF THEOREM 1. With $T_0 = 0$ and $\tau_i = T_i - T_{i-1}$, $i = 1, 2, \dots,$ we have that conditioned on $\{\tau_i < \infty, i = 1, \dots, n\}$ the random variables $\tau_1, \dots, \tau_n$

are independent under $P_{\bar\omega}$. Hence, for any $\bar\omega$ and $\lambda \le \lambda_{\mathrm{crit}}$,

$$E_{\bar\omega}\big(e^{\lambda T_n}\mathbf{1}_{T_n<\infty}\big) = E_{\bar\omega}\big(e^{\lambda\sum_{i=1}^n \tau_i}\mathbf{1}_{\bigcap_{i=1}^n\{\tau_i<\infty\}}\big) = \prod_{i=1}^n \varphi(\lambda, \theta^i\bar\omega), \tag{3.1}$$

where the second equality is due to the Markov property. By Lemma 1 and (2.5) it follows that $E_P(|f(\lambda,\bar\omega)|) < \infty$ for all $\lambda \le \lambda_{\mathrm{crit}}$. An application of Birkhoff's pointwise ergodic theorem then yields that

$$\begin{aligned}&\frac{1}{n}\log E_{\bar\omega}\big(e^{\lambda T_n}\mathbf{1}_{T_n<\infty}\big)\\ &\quad= \frac{1}{n}\sum_{i=1}^n f(\lambda,\theta^i\bar\omega) \underset{n\to\infty}{\longrightarrow} \int f(\lambda,\bar\omega)P(d\bar\omega), \qquad P\text{-a.e.},\end{aligned} \tag{3.2}$$

first for all λ rational and then for all $\lambda \le \lambda_{\mathrm{crit}}$ by monotonicity. Fixing $u \in \mathbb{R}$, by Chebyshev's inequality, for all $\bar\omega$ and $\lambda \le 0$,

$$P_{\bar\omega}\left(\frac{T_n}{n} \le u\right) \le e^{-\lambda n u} E_{\bar\omega}\big(e^{\lambda T_n}\mathbf{1}_{T_n<\infty}\big). \tag{3.3}$$

Thus, by (3.2), P-a.e. for all u,

$$\limsup_{n\to\infty}\frac{1}{n}\log P_{\bar\omega}\left(\frac{T_n}{n}\le u\right) \le -\sup_{\lambda\le 0} G(\lambda, P, u) = -\inf_{w\le u} I_P^{\tau,q}(w), \tag{3.4}$$

where (2.9) was used in the rightmost equality. The upper bound on the upper tail is derived similarly. Indeed, using Chebyshev's inequality with $\lambda \ge 0$,

$$P_{\bar\omega}\left(\infty > \frac{T_n}{n} \ge u\right) \le e^{-\lambda n u} E_{\bar\omega}\big(e^{\lambda T_n}\mathbf{1}_{T_n<\infty}\big),$$

and hence, as in (3.4), using now (2.8), P-a.e. for all u,

$$\limsup_{n\to\infty}\frac{1}{n}\log P_{\bar\omega}\left(\infty > \frac{T_n}{n}\ge u\right) \le -\sup_{\lambda\ge 0} G(\lambda, P, u) = -\inf_{w\ge u} I_P^{\tau,q}(w). \tag{3.5}$$

Suppose $\bar u < \infty$. Any closed set $F \subseteq [1,\infty)$ is contained in $[1,u_1]\cup[u_2,\infty)$ for some $u_1 \le \bar u \le u_2$ such that $u_1 \in F$ and $u_2 \in F$ (ignoring u_2 if $F \subseteq [1,\bar u]$ and u_1 if $F \subseteq [\bar u,\infty)$). So, by the monotonicity of $I_P^{\tau,q}(\cdot)$ (proved in Proposition 2), the inequalities (3.4) and (3.5) yield the upper bound for a general closed set F. If $\bar u = \infty$ and $K \subset [1,\infty)$ is compact, then $K \subseteq [1,u_1]$ for some $u_1 \in K$ and (3.4) yields the upper bound needed for the weak LDP of Theorem 1.

Due to the continuity of $I_P^{\tau,q}(\cdot)$ in the interior of its domain, implied by Proposition 2, it suffices to prove the complementary lower bound for (small) open intervals centered at rational $u > u_-$. To this end, assume first that $u_- < u < u_+$. Define a probability measure $Q_{\bar\omega,n}$ such that

$$\frac{dQ_{\bar\omega,n}}{dP_{\bar\omega}} = \frac{1}{Z_{n,\bar\omega}}\exp(\lambda_u T_n)\mathbf{1}_{T_n<\infty}, \qquad Z_{n,\bar\omega} = E_{\bar\omega}\big(\exp(\lambda_u T_n)\mathbf{1}_{T_n<\infty}\big),$$

and let $\overline{Q}_{\bar{\omega},n}$ denote the induced law on $\{\tau_1, \dots, \tau_n\}$. Due to the Markov property, $\overline{Q}_{\bar{\omega},n}$ is an n-fold product measure, whose marginals do not depend on n, hence we write $\overline{Q}_{\bar{\omega}}$ instead of $\overline{Q}_{\bar{\omega},n}$. Note that, for any $\delta > 0$,

$$\begin{aligned} &P_{\bar{\omega}}\left(\left|\frac{T_n}{n} - u\right| < \delta\right) \\ &\quad \ge \exp\left(-nu\lambda_u - n\delta|\lambda_u| + \sum_{i=1}^{n} f(\lambda_u, \theta^i\bar{\omega})\right)\overline{Q}_{\bar{\omega}}\left(\left|\frac{T_n}{n} - u\right| < \delta\right). \end{aligned} \tag{3.6}$$

Since P is ergodic and $u < u_+$, it holds that

$$E_{\overline{Q}_{\bar{\omega}}}\left(\frac{T_n}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} E_{\overline{Q}_{\theta^i\bar{\omega}}}(\tau_1) \underset{n\to\infty}{\longrightarrow} E_P\big(E_{\overline{Q}_{\bar{\omega}}}(\tau_1)\big) = g(\lambda_u) = u, \qquad P\text{-a.e.,} \tag{3.7}$$

where we have also used (2.4). With $\lambda_u < \lambda_{\mathrm{crit}}$, it also holds that there exists a $\beta > 0$ such that

$$E_P\big(E_{\overline{Q}_{\bar{\omega}}}(e^{\beta\tau_1})\big) < \infty,$$

implying by Chebyshev's inequality and independence that

$$\overline{Q}_{\bar{\omega}}\left(\left|\frac{T_n}{n} - u\right| \ge \delta\right) \underset{n\to\infty}{\longrightarrow} 0, \qquad P\text{-a.e.,} \tag{3.8}$$

Combining (3.8) with (3.6), we get

$$\begin{aligned} &\liminf_{n\to\infty} \frac{1}{n} \log P_{\bar{\omega}}\left(\left|\frac{T_n}{n} - u\right| < \delta\right) \\ &\quad \ge -u\lambda_u - \delta|\lambda_u| + \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(\lambda_u, \theta^i\bar{\omega}) \\ &\quad = -u\lambda_u - \delta|\lambda_u| + E_P\big(f(\lambda_u, \bar{\omega})\big) \\ &\quad = -G(\lambda_u, P, u) - \delta|\lambda_u| = -I_P^{\tau,q}(u) - \delta|\lambda_u|, \qquad P\text{-a.e.} \end{aligned} \tag{3.9}$$

[the first equality is due to Birkhoff's ergodic theorem and the last one to (2.6)]. This completes the proof of the lower bound in case $u < u_+$ since $\delta > 0$ is arbitrary.

Suppose $u_+ \ge \bar{u}$ is finite. Fixing a rational $u \ge u_+$ let $\ell = \lceil (u+1)/2 \rceil$ and for $m \ge 2\ell + 1$ define

$$\xi_m(\bar{\omega}) := m + 2\sum_{i=-\ell}^{0} \kappa(\theta^i\bar{\omega}).$$

Set $f_m(\lambda, \bar{\omega}) = \log E_{\bar{\omega}}(e^{\lambda T_1}\mathbf{1}_{T_1 \le \xi_m(\bar{\omega})})$, a monotone, convex function of λ, noting that, for P-a.e. $\bar{\omega}$,

$$P_{\bar{\omega}}\big(u + 2 \le T_1 \le \xi_m(\bar{\omega})\big) \ge \varepsilon^{2\ell+1} \prod_{i=-\ell}^{0} \mu_i\big([0, \kappa(\theta^i\bar{\omega})]\big)^2 > 0.$$

Consequently, for any $\lambda \geq 0$,

$$
\begin{aligned}
\lambda \xi_m(\bar{\omega}) &\geq f_m(\lambda, \bar{\omega}) \\
&\geq \lambda(u+2) + (2\ell+1)\log \varepsilon + 2\sum_{i=-\ell}^{0} \log \mu_i\big([0, \kappa(\theta^i \bar{\omega})]\big),
\end{aligned} \tag{3.10}
$$

implying that $E_P(|f_m(\lambda, \bar{\omega})|) < \infty$ by (C0) and the stationarity of P. It follows that the concave functions $G_m(\lambda, P, u) = \lambda u - E_P(f_m(\lambda, \bar{\omega}))$ are finite and smooth in $\lambda \geq 0$, with $G_m(\lambda, P, u) \to -\infty$ for $\lambda \to \infty$ by (3.10). Thus, the monotone function $g_m(\lambda) = \frac{d}{d\lambda} G_m(\lambda, P, u)$ is negative for all λ large enough, whereas it is not hard to check that $g_m(0) \geq u - \bar{u} \geq 0$. So, for all $m \geq m_0(u)$ there exists $\lambda_{u,m} \in [0, \infty)$ such that $g_m(\lambda_{u,m}) = 0$. The proof of the lower bound proceeds similarly to that for $u < u_+$, except for truncating the variables $\{\tau_i\}$ by considering the n-fold product law $\overline{Q}_{\bar{\omega},\xi_m}$ of $\{\tau_1, \dots, \tau_n\}$ under the probability measure $Q_{\bar{\omega},n,\xi_m}$ defined by

$$
\frac{dQ_{\bar{\omega},n,\xi_m}}{dP_{\bar{\omega}}} = \frac{1}{Z_{\bar{\omega},n,\xi_m}} \prod_{i=1}^{n} e^{\lambda_{u,m}\tau_i} \mathbf{1}_{\tau_i \leq \xi_m},
$$

$$
Z_{\bar{\omega},n,\xi_m} = \prod_{i=1}^{n} E_{\theta^i \bar{\omega}}\big(e^{\lambda_{u,m} T_1} \mathbf{1}_{T_1 \leq \xi_m(\theta^i \bar{\omega})}\big).
$$

Adapting in such a manner the argument leading to (3.9), one obtains the bound

$$
\begin{aligned}
\lim_{\delta \to 0} \liminf_{n \to \infty} \frac{1}{n} \log P_{\bar{\omega}}\left(\left|\frac{T_n}{n} - u\right| < \delta\right) &\geq -G_m(\lambda_{u,m}, P, u) \\
&= -\sup_{\lambda \geq 0} G_m(\lambda, P, u) := -I_m(u), \qquad P\text{-a.e.}
\end{aligned}
$$

(for details, see [1], proof of Theorem 4). With $G_m(\lambda, P, u)$ nonincreasing in m, so are the finite, nonnegative constants $I_m(u)$. Denoting by $I_\infty(u)$ the finite, nonnegative limit of $I_m(u)$, the intersection of the nonempty, compact, nested sets $\{\lambda \geq 0 : G_m(\lambda, P, u) \geq I_\infty(u)\}$, $m \geq m_0(u)$, contains a point $\lambda_{u,\infty}$. By monotone convergence

$$
I_\infty(u) \leq \lim_{m \to \infty} G_m(\lambda_{u,\infty}, P, u) = G(\lambda_{u,\infty}, P, u) \leq I_P^{\tau,q}(u),
$$

completing the proof of the lower bound.

We conclude the proof by deriving (1.2). To this end, fixing $m < \infty$, let $T_1^{(m)}$ and $T_{-1}^{(m)}$ be the hitting times corresponding to the truncated holding times

$$
H_i^m(x) = \begin{cases} m, & m < H_i(x) < \infty, \\ H_i(x), & \text{otherwise.} \end{cases}
$$

With $Z_m^+ = \log E_{\bar\omega}(e^{\lambda T_1^{(m)}} \mathbf{1}_{T_1^{(m)}<\infty})$ and $Z_m^- := \log E_{\bar\omega}(e^{\lambda T_{-1}^{(m)}} \mathbf{1}_{T_{-1}^{(m)}<\infty})$, fixing $\lambda \in \mathbb{R}$ it follows verbatim from the proof of [18], Lemma 2.3.22, that

$$E_P(Z_m^-) = E_P(Z_m^+) + E_P(\log \rho_0) \tag{3.11}$$

(possibly with both sides being infinite if $\lambda > 0$). Recall that T_1 (or T_{-1}) is finite if and only if $T_1^{(m)}$ (or $T_{-1}^{(m)}$ resp.) is finite for some m. So, in the case $\lambda \le 0$ we have that $0 \ge Z_m^- \downarrow Z_\infty^-$ and $0 \ge Z_m^+ \downarrow Z_\infty^+$, implying by monotone convergence that

$$E_P\big(\log E_{\bar\omega}\big(e^{\lambda T_{-1}} \mathbf{1}_{T_{-1}<\infty}\big)\big) = E_P\big(\log E_{\bar\omega}\big(e^{\lambda T_1} \mathbf{1}_{T_1<\infty}\big)\big) + E_P(\log \rho_0). \tag{3.12}$$

Similarly, if $\lambda > 0$ we have that $2\log\varepsilon \le Z_m^- \uparrow Z_\infty^-$ and $2\log\varepsilon \le Z_m^+ \uparrow Z_\infty^+$, so taking $m \to \infty$ in (3.11) yields (3.12) by monotone convergence. The latter allows us to relate $I_P^{\tau,q}(\cdot)$ and $I_P^{-\tau,q}(\cdot)$, in the same way as in the case without holding times. $\square$

REMARK 1. Note that (3.12) implies that $\lambda_{\text{crit}}^-(P) = \lambda_{\text{crit}}(P)$.

PROOF OF THEOREM 2. Since the proof of the annealed weak LDP for T_{-n}/n is almost identical to that for T_n/n, we present in the sequel only the latter.

We begin the proof of the upper bound in Theorem 2 with the upper tail in case $\lambda_{\text{crit}}(P) > 0$. Integration of (3.1) yields that, for all $\lambda < \lambda_{\text{crit}}(P)$,

$$\mathbb{E}\big(e^{\lambda T_n} \mathbf{1}_{T_n<\infty}\big) = E_P\bigg(\exp\bigg(n \int f(\lambda, \bar\omega) R_n(d\bar\omega)\bigg)\bigg).$$

By (C1), $\{R_n\}$ satisfies an LDP with good rate function $h(\cdot|P)$. As $R_n \in M_1^P$ and $\{\eta : h(\eta|P) < \infty\} \subseteq M_1^P$, where $\nu \mapsto \int f(\lambda, \bar\omega)\nu(d\bar\omega)$ is bounded and continuous (by Lemma 3), it follows from Varadhan's lemma (see [4], Theorem 4.3.1) that

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{\lambda T_n} \mathbf{1}_{T_n<\infty}\big) &= \sup_{\eta \in M_1^P} \bigg(\int f(\lambda, \omega)\eta(d\bar\omega) - h(\eta|P)\bigg) \\ &= L(\lambda). \end{aligned} \tag{3.13}$$

Fix $u > 0$. Combining (3.13) and Chebyshev's inequality for each $\lambda_{\text{crit}}(P) > \lambda \ge 0$, we get the upper bound

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\bigg(\infty > \frac{T_n}{n} \ge u\bigg) &\le - \sup_{0 \le \lambda < \lambda_{\text{crit}}(P)} [\lambda u - L(\lambda)] \\ &= - \inf_{w \ge u} I_P^{\tau,a}(w), \end{aligned} \tag{3.14}$$

where the equality follows from (2.18).

Applying the same argument with $\lambda < 0$ and using (2.17), yields that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\bigg(\frac{T_n}{n} \le u\bigg) \le - \inf_{w \le u} I_P^{\tau,a}(w). \tag{3.15}$$

If $\lambda_{\text{crit}}(P) = 0$, then $I_P^{\tau,a}(\cdot)$ is nonincreasing (see Proposition 3); hence (3.15) yields the upper bound for any compact $K \subset [1, \infty)$ as needed for the weak LDP of Theorem 2. Similarly, for $\lambda_{\text{crit}}(P) > 0$, the upper bound for a general compact set follows from (3.14), (3.15) and the convexity of $I_P^{\tau,a}(\cdot)$ (proved in Proposition 3).

It suffices to prove the lower bound in Theorem 2 for $(u - \delta, u + \delta)$ with $u \in [1, \infty)$ such that $I_P^{\tau,a}(u) < \infty$ and $\delta \downarrow 0$. Fixing such u and δ there exists $\eta \in M_1^e(\overline{\Omega}_\varepsilon)$ such that $I_\eta^{\tau,q}(u) + h(\eta|P) \le I_P^{\tau,a}(u) + \delta < \infty$. In particular, $u \ge u_-(\eta)$. Applying [1], Lemma 7, as in the proof of [1], Theorem 6, but here with the measures $Q_{\bar{\omega},n} \otimes \eta(d\bar{\omega})$ if $u \in (u_-(\eta), u_+(\eta))$ and $Q_{\bar{\omega},n,m} \otimes \eta(d\bar{\omega})$ if $u \ge u_+(\eta)$ [so we can use the strong law (3.8) for η-a.e. $\bar{\omega}$], we obtain the bound

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left|\frac{T_n}{n} - u\right| < \delta\right) \ge -I_\eta^{\tau,q}(u) - h(\eta|P),$$

for all $u > u_-(\eta)$. By continuity of the convex rate function $I_\eta^{\tau,q}(u)$ as $u \downarrow u_-(\eta)$, this bound applies also for $u = u_-(\eta)$. Taking $\delta \downarrow 0$ completes the lower bound in Theorem 2 and hence finishes the proof of this theorem. $\square$

4. Proof of the LDPs for rescaled positions X_t/t.

PROOF OF THEOREM 3. (a) We start by showing that $I_P^q(\cdot)$ of (1.3) is a convex, good rate function. Recall that $u_-(P) \ge 1$; hence $I_P^q(v) = \infty$ for all $v \notin [-1, 1]$ [see (1.2) and Proposition 2]. Moreover, with $\lambda_{\text{crit}} = \lambda_{\text{crit}}(P)$, by Lemma 1, (1.2) and the definition of $I_P^{\tau,q}(\cdot)$ we have that

$$I_P^q(v) = v\mathbf{1}_{v<0} E_P(\log \rho_0) + \sup_{\lambda \le \lambda_{\text{crit}}} \{\lambda - |v| E_P(f(\lambda, \bar{\omega}))\} \tag{4.1}$$

for all $v \neq 0$. In particular, $I_P^q(\cdot)$ is convex and lower semicontinuous on $(0, \infty)$ and $(-\infty, 0)$, separately. Using the linear lower bound of (2.5) it is easy to check that $\lim_{v\downarrow 0} I_P^q(v)$ exists and equals λ_{crit}. Further, $I_P^q(\cdot)$ is continuous at 0 by (4.1). It remains to show the convexity of $I_P^q(\cdot)$ at 0, namely that, for all $v_1, v_2 > 0$,

$$v_1 I_P^q(-v_2) + v_2 I_P^q(v_1) \ge (v_1 + v_2) I_P^q(0) = (v_1 + v_2)\lambda_{\text{crit}}.$$

By (4.1) (giving a lower bound for the sup by plugging in $\lambda = \lambda_{\text{crit}}$), this follows from the inequality

$$0 \ge E_P(\log \rho_0) + 2E_P(f(\lambda_{\text{crit}}, \bar{\omega})),$$

which by (3.12), is a consequence of the fact that

$$0 \ge f^-(\lambda_{\text{crit}}, \bar{\omega}) + f(\lambda_{\text{crit}}, \theta^{-1}\bar{\omega}), \tag{4.2}$$

for P-almost every $\bar{\omega} \in \overline{\Omega}_\varepsilon$ [integrate (4.2) with respect to the stationary measure P]. Indeed, by the Markov property

$$\begin{aligned} &E_{\bar{\omega}}(e^{\lambda_{\text{crit}} T_M} \mathbf{1}_{T_{-1} < T_M < \infty}) \\ &\quad = E_{\bar{\omega}}(e^{\lambda_{\text{crit}} T_{-1}} \mathbf{1}_{T_{-1} < T_M}) E_{\theta^{-1}\bar{\omega}}(e^{\lambda_{\text{crit}} T_1} \mathbf{1}_{T_1 < \infty}) E_{\bar{\omega}}(e^{\lambda_{\text{crit}} T_M} \mathbf{1}_{T_M < \infty}). \end{aligned}$$

Recall that $E_{\bar\omega}(e^{\lambda_{\rm crit} T_M}\mathbf{1}_{T_M<\infty})<\infty$ for P-almost every $\bar\omega$ and all $M<\infty$ [see (3.1)]. Thus,

$$1\ge \frac{E_{\bar\omega}(e^{\lambda_{\rm crit} T_M}\mathbf{1}_{T_{-1}<T_M<\infty})}{E_{\bar\omega}(e^{\lambda_{\rm crit} T_M}\mathbf{1}_{T_M<\infty})}=E_{\bar\omega}\big(e^{\lambda_{\rm crit} T_{-1}}\mathbf{1}_{T_{-1}<T_M}\big)E_{\theta^{-1}\bar\omega}\big(e^{\lambda_{\rm crit} T_1}\mathbf{1}_{T_1<\infty}\big).$$

Taking the logarithm and considering $M\to\infty$, one obtains (4.2).

Because $|X_t-X_s|\le|t-s|$, it suffices to consider the LDP bounds for the sequence X_n, $n=0,1,\dots,$ which we do hereafter (without further notice), in order to simplify notation.

Starting with the lower bounds, as $|X_t-X_s|\le|t-s|$, for $v\neq0$ and $1>\delta>0$,

$$P_{\bar\omega}\bigg(\frac{X_n}{n}\in(v-2\delta,v+2\delta)\bigg)\ge P_{\bar\omega}\big((1-\delta)n<T_{[nv]}<(1+\delta)n\big),$$

and Theorem 1 implies that, P-a.e. for all $v\neq0$ and $\delta>0$,

$$\liminf_{n\to\infty}\frac1n\log P_{\bar\omega}\bigg(\frac{X_n}{n}\in(v-2\delta,v+2\delta)\bigg)\ge\begin{cases}-vI_P^{\tau,q}\Big(\dfrac1v\Big), & v>0,\\ vI_P^{-\tau,q}\Big(\dfrac1{|v|}\Big), & v<0.\end{cases}$$

Similarly, taking $1>\delta>u>0$,

$$P_{\bar\omega}\bigg(\frac{X_n}{n}\in(-2\delta,2\delta)\bigg)\ge P_{\bar\omega}\big((1-\delta)n<T_{[nu]}<(1+\delta)n\big),$$

hence by Theorem 1,

$$\liminf_{n\to\infty}\frac1n\log P_{\bar\omega}\bigg(\frac{X_n}{n}\in(-2\delta,2\delta)\bigg)\ge -uI_P^{\tau,q}\Big(\frac1u\Big),\qquad P\text{-a.e.},$$

and considering rational $u\downarrow0$ completes the proof of the LDP lower bound.

We next deal with the complementary upper bounds. Assuming without loss of generality that $E_P(\log\rho_0)\le0$, we have that $T_1<\infty$ for P-almost every $\bar\omega$ [recall that here $H_i(x)<\infty$ for all i,x], and $v_P=1/\bar u(P)\ge0$. Since $n^{-1}X_n\in[-1,1]$, it suffices to show that, P-a.e.,

$$\lim_{\zeta\downarrow0}\limsup_{n\to\infty}\frac1n\log P_{\bar\omega}\bigg(\frac{X_n}{n}\in(v-\zeta,v+\zeta)\bigg)\le -I_P^q(v)\qquad\forall\,|v|\le1 \tag{4.3}$$

(cf. [4], Theorem 4.1.11). The next lemma, whose proof is deferred, is key to the proof of (4.3).

LEMMA 4. *Assume* (C0). *Suppose* $P(\mu_0(\{\infty\})>0)=0$ *and* $E_P(\log\rho_0)\le0$. *Let* $\Gamma=\operatorname{supp}P$ *and* $S_n=\inf\{t\ge n:X_t\le0\}$. *Then*

$$\limsup_{n\to\infty}\frac1n\log\sup_{\bar\omega\in\Gamma}P_{\bar\omega}(S_n<\infty)\le-\lambda_{\rm crit}(P). \tag{4.4}$$

We now prove (4.3) for $v = 0$. For any $\bar{\omega}$ and $\Delta > 0$,

$$\varepsilon^{\Delta n} P_{\bar{\omega}}\Big(\inf_{\ell \ge n} X_\ell \le \Delta n\Big) \le P_{\bar{\omega}}\Big(\inf_{\ell \ge n} X_\ell \le 0\Big) = P_{\bar{\omega}}(S_n < \infty).$$

In particular, since $I_P^q(0) = \lambda_{\mathrm{crit}}(P)$ and

$$P_{\bar{\omega}}\big(X_n \in (-\zeta n, \zeta n)\big) \le P_{\bar{\omega}}\Big(\inf_{\ell \ge n} X_\ell \le \zeta n\Big) \le \varepsilon^{-\zeta n} P_{\bar{\omega}}(S_n < \infty), \tag{4.5}$$

(4.4) implies that (4.3) holds for $v = 0$. Considering next $v \ne 0$ and $\zeta \in (0, |v|)$ such that $u = v - \zeta \cdot \operatorname{sign} v$ is rational, note that, for any $\delta \in (0, 1)$ such that, $1/(\delta u)$ is integer,

$$\begin{aligned} &P_{\bar{\omega}}\Big(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\Big) \\ &\quad \le \varepsilon^{-2\zeta n} \sum_{k=1}^{(|u|\delta)^{-1}} P_{\bar{\omega}}\Big(\frac{T_{[nu]}}{n|u|} \in [(k-1)\delta, k\delta]\Big) P_{\theta^{[nu]}\bar{\omega}}\big(S_{[n - nk\delta|u|]} < \infty\big). \end{aligned} \tag{4.6}$$

With $\xi = k\delta|u| \le 1$, it follows from Theorem 1 and Proposition 2 that, P-a.e.,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log P_{\bar{\omega}}\Big(\frac{T_{[nu]}}{n|u|} \in [(k-1)\delta, k\delta]\Big) \\ &\quad \le -|u| I_P^{(\operatorname{sign} u)\tau, q}\Big(\frac{\xi}{|u|}\Big) + |u| w(|u|, \delta), \end{aligned} \tag{4.7}$$

for all k and rational $u, \delta > 0$, where

$$\begin{aligned} w(r, \delta) := \max\big\{ &|I_P^{\tau,q}(s) - I_P^{\tau,q}(t)| + |I_P^{-\tau,q}(s) - I_P^{-\tau,q}(t)|; \\ &s, t \in [\bar{u}(P), 1/r], |s - t| \le \delta\big\}. \end{aligned}$$

Let $\Gamma' = \{\bar{\omega} : \theta^k \bar{\omega} \in \Gamma \ \forall k \in \mathbb{Z}\}$, noting that $P(\Gamma') = 1$ by stationarity (in fact $\Gamma' = \Gamma$), whereas by (4.4),

$$\limsup_{n\to\infty} \frac{1}{n} \log \sup_{\bar{\omega}\in\Gamma'} P_{\theta^{[nu]}\bar{\omega}}\big(S_{[n - n\xi]} < \infty\big) \le -(1 - \xi)\lambda_{\mathrm{crit}}(P). \tag{4.8}$$

Substituting (4.7) and (4.8) in (4.6), and using the relation (1.3) we deduce that, P-a.e.,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log P_{\bar{\omega}}\Big(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\Big) \\ &\quad \le -\inf_{\xi\in[0,1]} \Big\{\xi I_P^q\Big(\frac{u}{\xi}\Big) + (1 - \xi) I_P^q(0)\Big\} + |u| w(|u|, \delta) - 2\zeta \log \varepsilon. \end{aligned}$$

As the finite, convex, rate function $I_P^{\tau,q}(\cdot)$ is continuous on $(u_-(P), \infty)$, the oscillation $w(r, \delta) \to 0$ for $\delta \downarrow 0$ and any fixed $r < \infty$. With $I_P^q(\cdot)$ convex and lower semicontinuous, taking $\delta \downarrow 0$ then $\zeta \downarrow 0$ we obtain the bound of (4.3) and complete the proof of the theorem in the case $P(\mu_0(\{\infty\}) > 0) = 0$.

(b) For $I_P^q(v)$ of (1.4), $v \neq 0$ we have by the same reasoning that led to (4.1), the analogous representation,

$$\begin{aligned} I_P^q(v) &= v\mathbf{1}_{v<0} E_P(\log \rho_0) + \inf_{\ell \in [0,1]} \sup_{\lambda \le \lambda_{\text{crit}}} \{\lambda \ell - |v| E_P(f(\lambda, \bar{\omega}))\} \\ &= v\mathbf{1}_{v<0} E_P(\log \rho_0) + \sup_{\lambda \le 0} \{\lambda - |v| E_P(f(\lambda, \bar{\omega}))\}, \end{aligned} \tag{4.9}$$

where the second equality follows by an application of the min–max theorem ([15], Theorem 4.2′) for the function $(\ell, \lambda) \mapsto \lambda\ell - |v|E_P(f(\lambda, \bar{\omega}))$ ($\ell \in [0, 1]$, $\lambda \in (-\infty, \lambda_{\text{crit}}]$), which is convex in ℓ and concave and continuous in λ [the continuity of $\lambda \mapsto E_P(f(\lambda, \bar{\omega}))$ follows from (2.1), (2.5) and dominated convergence]. Here too $I_P^q(v) = \infty$ for all $v \notin [-1, 1]$, whereas by (4.9), $I_P^q(\cdot)$ is convex and lower semicontinuous on $(0, \infty)$ and $(-\infty, 0)$, separately. Combining the linear lower bound (2.5) with the representation (4.9) we see that $\lim_{v\to 0} I_P^q(v) = 0$; that is, $I_P^q(\cdot)$ is continuous at 0. Since $I_P^q(\cdot) \ge 0$, its convexity at 0 trivially holds.

As for the LDP lower bounds, let $\xi > 0$ be such that $P(\mu_0(\{\infty\}) > \xi) = p > 0$. Fixing a rational $v \neq 0$, we have, for all $\ell \in [0, 1]$,

$$\begin{aligned} &P_{\bar{\omega}}\left(\frac{X_n}{n} \in (v - 2\delta, v + 2\delta)\right) \\ &\qquad \ge P_{\bar{\omega}}(T_{[nv]} \in (\ell n - \delta n, \ell n + \delta n)) P_{\theta^{[nv]}\bar{\omega}}(|X_{(1-\ell)n}| < \delta n) \end{aligned}$$

whereas

$$P_{\theta^{[nv]}\bar{\omega}}(|X_{(1-\ell)n}| < \delta n) \ge \varepsilon^{\delta n} \max_{\{j : |j-[nv]| < \delta n\}} \mu_j(\{\infty\}) := \varepsilon^{\delta n} \xi_n(\bar{\omega}).$$

We thus get the LDP lower bound with rate function (1.4) out of that of Theorem 1 (including also the case of $v = 0$), provided $\xi_n(\bar{\omega}) > \xi$ for all n large enough. By Birkhoff's pointwise ergodic theorem this holds for P-almost every $\bar{\omega}$, as

$$\left|\frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{\mu_j(\{\infty\})>\xi} - p\right| \le \frac{\delta p}{2|v|}, \qquad \left|\frac{1}{n}\sum_{j=-n}^{-1} \mathbf{1}_{\mu_j(\{\infty\})>\xi} - p\right| \le \frac{\delta p}{2|v|},$$

for all $n \ge n_0(\delta, \bar{\omega})$, whereby obviously $\xi_n(\bar{\omega}) > \xi$ whenever $n(|v| - \delta) > n_0$.

To prove the complementary upper bounds, namely, (4.3), since now $I_P^q(0) = 0$, it suffices to consider $v \neq 0$. For the same choice of $\zeta \in (0, |v|)$ and rational $u = v - \zeta \cdot \operatorname{sign} v$ we have that

$$P_{\bar{\omega}}\left(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\right) \le P_{\bar{\omega}}(T_{[nu]} \le n). \tag{4.10}$$

Considering $n \to \infty$, it thus follows from Theorem 1 and (1.4) that

$$\limsup_{n\to\infty} \frac{1}{n} \log P_{\bar{\omega}}\left(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\right) \le -I_P^q(u),$$

which holds P-a.e. for all v and ζ as above. With $I_P^q(\cdot)$ lower semicontinuous, taking $\zeta \downarrow 0$ completes the proof of (4.3) and hence that of the theorem. $\square$

PROOF OF LEMMA 4. Recall that our assumptions imply that $v_P = 1/\bar{u}(P) \geq 0$. The lemma is trivial for $v_P = 0$ as then $\lambda_{\text{crit}} = 0$. Assuming hereafter that $v_P > 0$, let $b_n(\bar{\omega}) = P_{\bar{\omega}}(S_n < \infty)$, $a_n := \sup\{b_n(\bar{\omega}) : \bar{\omega} \in \Gamma\}$, and $\Upsilon = \inf\{t \geq 1 : X_t = 0\}$. By the strong Markov property of the embedded RWRE, denoting by $P_{\bar{\omega}}^y(\cdot)$ the law of the random walk started at y in the environment $\bar{\omega}$ (where we omit y if $y = 0$), it holds that, for all k, $\bar{\omega}$ and all $y < 0$,

$$P_{\bar{\omega}}^y(S_k < \infty) \geq P_{\bar{\omega}}^y(\Upsilon \leq k) P_{\bar{\omega}}(S_k < \infty) + P_{\bar{\omega}}^y(\Upsilon > k) \geq P_{\bar{\omega}}(S_k < \infty).$$

Since

$$\begin{aligned} P_{\bar{\omega}}(S_{m+k} < \infty) &\geq \varepsilon P_{\bar{\omega}}(S_m < \infty) E_{\bar{\omega}}\bigl(P_{\bar{\omega}}^{X_{S_m}-1}(S_k < \infty) | S_m < \infty\bigr) \\ &\geq \varepsilon P_{\bar{\omega}}(S_m < \infty) P_{\bar{\omega}}(S_k < \infty), \end{aligned}$$

it follows that $b_{jk}(\bar{\omega}) \geq (\varepsilon b_k(\bar{\omega}))^j$, hence also $a_{jk} \geq (\varepsilon a_k)^j$ for all positive integers j. This and the ellipticity estimate $\varepsilon^{k+1} \leq \varepsilon a_k \leq a_{k+1} \leq a_k$ imply that $k^{-1} \log a_k \to a$, for some $a \in [\log \varepsilon, 0]$.

We next show that $n^{-1} \log b_n(\bar{\omega}) \to a$ as $n \to \infty$, for P-a.e. $\bar{\omega}$. To this end, fix $\delta > 0$ and $k < \infty$ large enough for $k^{-1} \log a_k \geq a - \delta$. There exists an $\widetilde{\omega} \in \Gamma$ such that $k^{-1} \log b_k(\widetilde{\omega}) \geq a - 2\delta$. Therefore, one may find a finite ℓ large enough such that $k^{-1} \log P_{\widetilde{\omega}}(S_k < \ell) \geq a - 3\delta$. Let $\mathbf{Z} = (Z_0, Z_1, \ldots, Z_\ell)$, $\mathbf{z} = (z_0, z_1, \ldots, z_\ell)$ and $\mathbf{\Theta} = (\Theta_1, \ldots, \Theta_\ell)$, and use the notation $P_\omega(A)$, or $P_\mu(A)$, for events A which depend on the environment only via $\omega := (\omega_{-\ell}, \ldots, \omega_\ell)$ or $\mu := (\mu_{-\ell}, \ldots, \mu_\ell)$, respectively. Note that

$$G_{\mathbf{z}} := \bigcup_{\{0 \leq j \leq \ell-1 : z_j \leq 0\}} \{\mathbf{\Theta} : \Theta_j < \ell, \Theta_{j+1} > k\}$$

are open subsets of $\mathbb{R}_+^\ell$ and

$$P_{\bar{\omega}}(S_k < \ell) = \sum_{\mathbf{z}} P_\omega(\mathbf{Z} = \mathbf{z}) P_\mu(\mathbf{\Theta} \in G_{\mathbf{z}} | \mathbf{Z} = \mathbf{z}). \tag{4.11}$$

A finite number of $\mathbf{z}$ vectors is considered in (4.11), for each of which $\bar{\omega} \mapsto P_\omega(\mathbf{Z} = \mathbf{z})$ is continuous on $\bar{\Omega}_\varepsilon$ while $\mu \mapsto \mathcal{L}(\mathbf{\Theta} | \mathbf{Z} = \mathbf{z}) : M_1^\varepsilon(\mathbb{R}_+)^{2\ell+1} \to M_1(\mathbb{R}_+^\ell)$ are also continuous [where $\mathcal{L}(\mathbf{\Theta} | \mathbf{Z} = \mathbf{z})$ denotes the conditional distribution of $\mathbf{\Theta}$ given the event $\mathbf{Z} = \mathbf{z}$]. By (4.11), we see that $\bar{\omega} \mapsto P_{\bar{\omega}}(S_k < \ell)$ is lower semicontinuous on $\overline{\Omega}_\varepsilon$. Consequently, there exists an open set $A \subseteq \overline{\Omega}_\varepsilon$ such that $P(A) > 0$ and

$$k^{-1} \log P_{\bar{\omega}}(S_k < \ell) \geq a - 4\delta, \qquad \forall \bar{\omega} \in A.$$

Now let $g(\bar\omega) \ge 0$ be the smallest integer such that $\theta^{-g(\bar\omega)}\bar\omega \in A$. Since $P(A) > 0$, it follows from ergodicity that $g(\bar\omega) < \infty$ for P-almost every $\bar\omega$, in which case

$$\begin{aligned} b_n(\bar\omega) &= P_{\bar\omega}(S_n < \infty) \\ &\ge \varepsilon^{g(\bar\omega)} P_{\theta^{-g(\bar\omega)}\bar\omega}(S_n < \infty) \ge \varepsilon^{g(\bar\omega)}\big[\varepsilon P_{\theta^{-g(\bar\omega)}\bar\omega}(S_k < \infty)\big]^{\lceil n/k\rceil} \\ &\ge \varepsilon^{g(\bar\omega)}\big[\varepsilon P_{\theta^{-g(\bar\omega)}\bar\omega}(S_k < \ell)\big]^{\lceil n/k\rceil} \ge \varepsilon^{g(\bar\omega)}\big[\varepsilon e^{k(a-4\delta)}\big]^{\lceil n/k\rceil}, \end{aligned} \tag{4.12}$$

yielding for P-almost every $\bar\omega$, the bound,

$$\liminf_{n\to\infty} n^{-1}\log b_n(\bar\omega) \ge k^{-1}\log\varepsilon - 4\delta + a \ge k^{-1}\log\varepsilon - 4\delta + \limsup_{n\to\infty} n^{-1}\log b_n(\bar\omega).$$

Taking $k \to \infty$ followed by (rational) $\delta \downarrow 0$, we conclude that

$$a = \lim_{n\to\infty} n^{-1}\log b_n(\bar\omega), \qquad P\text{-a.e.} \tag{4.13}$$

Fixing $1 > \delta > 0$, $u \in (0, v_P/(1+v_P))$, let $\mathcal{S}$ denote the finite set of integer pairs (k,ℓ) such that $1 + 1/\delta \le \min(k,\ell)$ and $(k+\ell-2)\delta u \le 1$. We have by the strong Markov property that

$$\begin{aligned} b_n(\bar\omega) &\le P_{\bar\omega}\big(T_{[nu]} \ge n(1-u)\big) \\ &\quad + \sum_{(k,\ell)\in\mathcal{S}} P_{\bar\omega}\bigg(\frac{T_{[nu]}}{nu} \in [(k-1)\delta, k\delta[\bigg) \\ &\qquad \times P_{\theta^{[nu]}\bar\omega}\bigg(\frac{T_{-[nu]}}{nu} \in [(\ell-1)\delta, \ell\delta[\bigg) \\ &\qquad \times P_{\bar\omega}\big(S_{[n-n(k+\ell)u\delta]} < \infty\big), \end{aligned} \tag{4.14}$$

where we use the convention $b_t(\bar\omega) = P_{\bar\omega}(S_t < \infty) = 1$ for $t \le 0$. Observing that

$$E_{\theta^m\bar\omega}\big(e^{\lambda T_{-m}}\mathbf{1}_{T_{-m}<\infty}\big) = \prod_{i=1}^m \varphi^-(\lambda, \theta^i\bar\omega),$$

we follow the derivation of (3.4) and (3.5) to deduce in analogy to (4.7) that, with $\gamma = \ell\delta u \le 2$,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n}\log P_{\theta^{[nu]}\bar\omega}\bigg(\frac{T_{-[nu]}}{nu} \in [(\ell-1)\delta, \ell\delta[\bigg) \\ &\qquad \le -\gamma I_P^q\Big(-\frac{u}{\gamma}\Big) + uw\Big(\frac{u}{2}, \delta\Big), \qquad P\text{-a.e.} \end{aligned} \tag{4.15}$$

By convexity of $I_P^q(\cdot)$, with $\xi = k\delta u$,

$$\xi I_P^q\Big(\frac{u}{\xi}\Big) + \gamma I_P^q\Big(-\frac{u}{\gamma}\Big) \ge (\xi+\gamma) I_P^q(0).$$

So, with $(k+\ell)\delta \geq 2$, by (4.7), (4.14) and (4.15), for P-almost every $\bar{\omega}$ and all $n > n_0(\bar{\omega})$,

$$\begin{aligned} b_n(\bar{\omega}) &\leq e^{n(2uw(u/2,\delta)+\delta)}\Bigg[e^{-n(1-u)I_P^q(u/(1-u))} \\ &\qquad + \sum_{(k,\ell)\in\mathcal{S}} e^{-n(k+\ell)\delta u I_P^q(0)}\, b_{[n-n(k+\ell)\delta u]}(\bar{\omega})\Bigg] \\ &\leq C \max_{2nu\leq j\leq n}\{e^{-jJ}\, b_{n-j}(\bar{\omega})\}, \end{aligned} \tag{4.16}$$

where $C = C(\delta, u) < \infty$ and

$$J = \min\left\{(1-u)I_P^q\left(\frac{u}{1-u}\right), I_P^q(0)\right\} - \frac{1}{2u}\left(2uw\left(\frac{u}{2},\delta\right)+\delta\right).$$

It is easy to check that, for $u > 0$, $\gamma_n \geq 0$, $C < \infty$,

$$\gamma_n \leq C \max_{2nu\leq j\leq n}\{e^{-jJ}\,\gamma_{n-j}\}, \qquad \forall n \geq n_0 \quad \Longrightarrow \quad \limsup_{n\to\infty}\frac{1}{n}\log\gamma_n \leq -J.$$

Consequently, from (4.16) we have that

$$\limsup_{n\to\infty}\frac{1}{n}\log b_n(\bar{\omega}) \leq -J, \qquad P\text{-a.e.} \tag{4.17}$$

Since $J \to I_P^q(0)$ when taking first $\delta \downarrow 0$ then $u \downarrow 0$, it follows from (4.13) and (4.17) that $a \leq -I_P^q(0)$ as stated. $\square$

PROOF OF THEOREM 4. (a) With $u_-(\eta) \geq 1$ for any $\eta \in M_1^e(\overline{\Omega}_\varepsilon)$ we have that $I_\eta^{\tau,q}(u) = I_\eta^{-\tau,q}(u) = \infty$ for all $u < 1$ [see Proposition 2 and (1.2)], hence $I_P^a(v) = \infty$ for $v \neq [-1,1]$. Since $I_P^{\tau,a}(\cdot)$ and $I_P^{-\tau,a}(\cdot)$ are rate functions, $I_P^a(\cdot)$ of (1.5) is a good rate function provided it is continuous at 0, which we show next. Denoting throughout $\lambda_{\text{crit}} = \lambda_{\text{crit}}(P)$, recall that $L(\lambda) \leq -2\log\varepsilon$ for $L(\cdot)$ of (2.15) and all $\lambda \leq \lambda_{\text{crit}}$ (see proof of Proposition 3). Hence, $I_P^{\tau,a}(u) \geq \lambda_{\text{crit}}u + 2\log\varepsilon$ by (2.16), implying that

$$\liminf_{u\downarrow 0} u I_P^{\tau,a}\left(\frac{1}{u}\right) \geq \lambda_{\text{crit}}.$$

With the same argument applying for $I_P^{-\tau,a}(\cdot)$, we get that

$$\liminf_{v\to 0} I_P^a(v) \geq \lambda_{\text{crit}}. \tag{4.18}$$

By definition, $I_P^a(v) \leq I_P^q(v)$ for $v \neq 0$. As $I_P^q(v) \to \lambda_{\text{crit}}$ for $v \to 0$ (see proof of Theorem 3), we conclude that $I_P^a(v) \to \lambda_{\text{crit}} = I_P^a(0)$ when $v \to 0$, completing the proof that $I_P^a(\cdot)$ is a good rate function. Since $I_P^{\tau,a}(\cdot)$ and $I_P^{-\tau,a}(\cdot)$ are convex, it follows immediately that $I_P^a(\cdot)$ of (1.5) is convex separately on $(0, \infty)$ and on $(-\infty, 0)$. The convexity of this function at 0 amounts to the inequality

$$v_1 I_P^a(-v_2) + v_2 I_P^a(v_1) \geq (v_1 + v_2)\lambda_{\text{crit}}, \tag{4.19}$$

which we prove next. As $P \in M_1^e(\overline{\Omega}_\varepsilon)$ is locally equivalent to the product of its marginals, the bound (2.11) results with (4.2) holding for all $\bar{\omega} \in (\text{supp}\, P|_1)^{\mathbb{Z}}$. Note that $f^-(\lambda, \bar{\omega})$ depends only on $\{\bar{\omega}_x, x \geq 0\}$ while $f(\lambda, \theta^{-1}\bar{\omega})$ depends only on $\{\bar{\omega}_x, x \leq -1\}$, so integrating (4.2) with respect to $\eta|_{(\dots,\bar{\omega}_{-2},\bar{\omega}_{-1})} \otimes \eta'|_{(\bar{\omega}_0,\bar{\omega}_1,\dots)}$ yields that

$$0 \geq E_{\eta'}\big(f^-(\lambda_{\text{crit}}, \bar{\omega})\big) + E_\eta\big(f(\lambda_{\text{crit}}, \bar{\omega})\big),$$

for any stationary $\eta, \eta' \in M_1^P$. Then, for all such η, η' and $v_1, v_2 > 0$,

$$I_{\eta'}^{-\tau,q}\left(\frac{1}{v_2}\right) + I_\eta^{\tau,q}\left(\frac{1}{v_1}\right) \geq \left(\frac{1}{v_2} + \frac{1}{v_1}\right)\lambda_{\text{crit}}.$$

With $h(\eta|P) = \infty$ for all $\eta \notin M_1^P$, also

$$I_P^{-\tau,a}\left(\frac{1}{v_2}\right) + I_P^{\tau,a}\left(\frac{1}{v_1}\right) \geq \left(\frac{1}{v_2} + \frac{1}{v_1}\right)\lambda_{\text{crit}},$$

resulting by (1.5) with (4.19).

The annealed LDP lower bounds in the case $P(\mu_0(\{\infty\}) > 0) = 0$ follow from the lower bounds of Theorem 2, by the same reasoning as in the proof of the quenched bounds in Theorem 3. Turning to the upper bounds, it suffices to show that, for any v,

$$\lim_{\zeta \downarrow 0} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\right) \leq -I_P^a(v). \tag{4.20}$$

Assume without loss of generality that $E_P(\log \rho_0) \leq 0$, in which case Lemma 4 applies. Starting with $v = 0$, we have by (4.5) that

$$\mathbb{P}\big(X_n \in (-\zeta n, \zeta n)\big) \leq \varepsilon^{-\zeta n} \sup_{\bar{\omega}\in\Gamma} P_{\bar{\omega}}(S_n < \infty),$$

and since $I_P^a(0) = \lambda_{\text{crit}}(P)$, we have (4.20) by an application of Lemma 4. Recall that $P(\Gamma') = 1$ for $\Gamma' = \{\bar{\omega} : \theta^k \bar{\omega} \in \Gamma\ \forall k \in \mathbb{Z}\}$. Hence, by (4.6) for any $v \neq 0$, $\zeta \in (0, |v|)$ and $u = v - \zeta \cdot \text{sign}\, v$,

$$\mathbb{P}\left(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\right)$$
$$\leq \varepsilon^{-2\zeta n} \sum_{k=1}^{(|u|\delta)^{-1}} \mathbb{P}\left(\frac{T_{[nu]}}{n|u|} \in [(k-1)\delta, k\delta]\right) \sup_{\bar{\omega}\in\Gamma'} P_{\theta^{[nu]}\bar{\omega}}\big(S_{[n-nk\delta|u|]} < \infty\big).$$

Thus, combining (4.8), Theorem 2 and the relation (1.5) we have that

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{X_n}{n}\in(v-\zeta,v+\zeta)\right)$$
$$\leq -2\zeta\log\varepsilon-\inf_{\xi\in[0,1]}\left\{\xi I_P^a\left(\frac{u}{\xi}\right)+(1-\xi-\delta|u|)\lambda_{\mathrm{crit}}\right\}.$$

Since $I_P^a(\cdot)$ is convex and lower semicontinuous, with $\lambda_{\mathrm{crit}}=I_P^a(0)$, taking $\delta\downarrow 0$ followed by $\zeta\downarrow 0$ we arrive at the bound (4.20).

(b) Considering $I_P^a(\cdot)$ of (1.7), note that $I_P^a(v)=vI_P^{\tau,a}(u^*\wedge 1/v)$ for any $v>0$, where $u^*\geq 1$ is a global minimizer of $I_P^{\tau,a}(u)$, setting $u^*=\infty$ in case $I_P^{\tau,a}(\cdot)$ is nonincreasing. Since $I_P^{\tau,a}(\cdot)$ is a convex rate function, the lower semicontinuity and convexity of $I_P^a(\cdot)$ on $(0,\infty)$ are easily verified. Applying the same reasoning to the convex rate function $I_P^{-\tau,a}(\cdot)$ we get the convexity and lower semicontinuity of $I_P^a(\cdot)$ of (1.7) at $(-\infty,0)$. Recall that this nonnegative function is bounded above by $I_P^q(v)$ of (1.4), which converges to 0 as $v\to 0$. The function $I_P^a(\cdot)$ of (1.7) is thus convex and continuous at 0, hence a convex good rate function on $\mathbb{R}$.

Turning to the LDP lower bounds, note that for $v>0$, $0<\delta<\ell\leq 1$ and all $n\geq k_0/\delta$ by our assumption (1.6),

$$\mathbb{P}\left(\frac{X_n}{n}\in(v-2\delta,v+2\delta)\right)$$
$$\geq\varepsilon^{2\delta n}\mathbb{P}\Big(T_{[nv]}\in\big((\ell-\delta)n,(\ell+\delta)n\big),\max_{\delta n\leq j\leq 2\delta n}H_1([nv]+j)=\infty\Big)$$
$$\geq\varepsilon^{2\delta n}E_P\Big(P_{\bar\omega}\big(T_{[nv]}\in\big((\ell-\delta)n,(\ell+\delta)n\big)\big)$$
$$\times E_P\Big(\max_{\delta n\leq j\leq 2\delta n}\mu_{[nv]+j}(\{\infty\})|\mathcal{F}_{[nv]}^-\Big)\Big)$$
$$\geq e^{-c\delta n}\varepsilon^{2\delta n}\mathbb{P}\big(T_{[nv]}\in(\ell-\delta)n,(\ell+\delta)n\big).$$

Consequently, for any $v>0$ and all $\ell\in(0,1]$, by Theorem 2,

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{X_n}{n}\in(v-2\delta,v+2\delta)\right)\geq -vI_P^{\tau,a}\left(\frac{\ell}{v}\right).$$

Optimizing over $\ell\in[0,1]$ we arrive at the stated LDP lower bound for $v>0$. The same argument applies for $T_{-[vn]}$, leading to the stated lower bound for $v<0$, and since $\mathbb{P}(X_n=0)\geq E_P(\mu_0(\{\infty\}))>0$, we have the lower bound also for $v=0$.

As for the upper bound, it suffices to consider (4.20) for $v\neq 0$, where by (4.10) we have that, with $u=v-\zeta\cdot\operatorname{sign}v$,

$$\mathbb{P}\left(\frac{X_n}{n}\in(v-\zeta,v+\zeta)\right)\leq\mathbb{P}\big(T_{[nu]}\leq n\big).$$

Considering $n \to \infty$, by Theorem 2 and the relation (1.7) we have that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{X_n}{n} \in (v-\zeta, v+\zeta)\right) \le -I_P^a(u),$$

and with $I_P^q(\cdot)$ lower semicontinuous, taking $\zeta \downarrow 0$ completes the proof of (4.20) and hence that of the theorem. □

5. Negative speed for random walks on Galton–Watson trees. Let Z be a random variable taking values on $\{1, 2, \dots\}$ with finite mean $m = E(Z) > 1$. Consider the Galton–Watson (GW) measure on rooted trees, which is the family tree of a supercritical branching process starting from the first ancestor (called the root), with each particle independently producing a random number of children according to the law of Z. The modified Galton–Watson (MGW) measure is obtained by changing the distribution of the number of children at the root to that of $Z - 1$.

The augmented Galton–Watson (AGW) measure on nonrooted trees containing a special ray $-\infty \leftrightarrow 0 \leftrightarrow \infty$ is then constructed as follows. Starting with $\mathbb{Z}$, we connect neighboring integers by an edge, and attach to each point $x \in \mathbb{Z}$ an independent MGW-tree $\mathbf{T}_x$. We write the resulting infinite, unrooted tree as $\mathbf{T} = \bigcup_{x\in\mathbb{Z}} \mathbf{T}_x$, where the roots of $\mathbf{T}_x$ and $\mathbf{T}_{x+1}$ are connected by an edge. The parent v^* of a vertex $v \in \mathbf{T} \cap \mathbf{T}_x$ is defined as the parent of v in $\mathbf{T}_x$ if v is not the root of $\mathbf{T}_x$, and as $x - 1$ if $v = x \in \mathbb{Z}$, that is, if v is the root of $\mathbf{T}_x$. An alternative construction of the AGW measure starts with a GW tree and the "rightmost" vertex v of distance n from the root, renaming it 0, while renaming the set D_m of vertices at distance m from the root as $\widetilde{D}_{m-n}$ and then taking weak limits, resulting in a measure on infinite trees with a special ray $-\infty \leftrightarrow \infty$ marked (see [10] for details). Fixing $0 < \boldsymbol{\lambda} < \infty$ and a tree ω chosen according to AGW, the $\boldsymbol{\lambda}$-biased random walk $\{S_n\}$ on ω is the Markov chain such that if j^* is the parent of a vertex j having k children $j_1, \dots, j_k$, then

$$P^v_{\boldsymbol{\lambda},\omega}[S_{n+1} = j^* | S_n = j] = \frac{\boldsymbol{\lambda}}{\boldsymbol{\lambda} + k},$$

$$P^v_{\boldsymbol{\lambda},\omega}[S_{n+1} = j_i | S_n = j] = \frac{1}{\boldsymbol{\lambda} + k}, \qquad i = 1, 2, \dots, k,$$

where $v \in \omega$ is a fixed starting point (see [11]). We denote by $P^v_{\boldsymbol{\lambda},\omega}$ the "quenched" distribution of the walk $\{S_n\}$ conditioned on the tree ω and by $P^v_{\boldsymbol{\lambda}} := \int P^v_{\boldsymbol{\lambda},\omega} \mathrm{AGW}(d\omega)$ the corresponding "annealed" measure. We write $P_{\boldsymbol{\lambda},\omega}$ for $P^0_{\boldsymbol{\lambda},\omega}$ and $P_{\boldsymbol{\lambda}}$ for $P^0_{\boldsymbol{\lambda}}$.

For x on the special ray, let $H(x) + 1$ be the first hitting time of the set $\{x - 1, x + 1\}$ [possibly $H(x) = +\infty$] and let μ_x be the distribution of $H(x)$ under $P^x_{\boldsymbol{\lambda},\omega}$. Let $\omega_x := 1/(\boldsymbol{\lambda} + 1)$. Note that ω_x is deterministic and does not depend on x. Then the projection of $\{S_n\}$ on $\mathbb{Z}$, denoted $\{X_n\}$, is a RWREH with i.i.d.

environment $\bar{\omega} = \{(\omega_x, \mu_x)\}$. Indeed, the distribution P of $\bar{\omega}$ (under the measure AGW on trees with a special ray) is a (stationary) product measure where if Z is bounded, then also $P|_1 \in M_1(S_\varepsilon)$ for some $\varepsilon > 0$, for which (C3) applies with $b = 0$. Let $P_{\bar{\omega}}$ be the distribution of $\{X_n\}$ under $P_{\boldsymbol{\lambda},\omega}$ and $\mathbb{P}$ the distribution of $\{X_n\}$ under $P_{\boldsymbol{\lambda}}$. Then we are in the RWREH model. Since P is a product measure, (C1) and (C2) are clearly satisfied. Hence we can apply our previous results. In particular, we have by Lemma 1 and (3.12) a deterministic $\lambda_{\text{crit}} \in [0, \infty)$ such that $E_{\boldsymbol{\lambda},\omega}[e^{tT_{-1}}\mathbf{1}_{T_{-1}<\infty}]$ is finite if and only if $t \le \lambda_{\text{crit}}$, for AGW-a.e. ω. Moreover, by Theorems 1 and 2 we have the weak LDP for $n^{-1}T_{-n}$ (and $n^{-1}T_n$) under $P_{\bar{\omega}}$ and $\mathbb{P}$, with quenched and annealed rate functions $I_P^{-\tau,q}$ and $I_P^{-\tau,a}$, respectively. By Theorems 3 and 4 we also have the LDP for $n^{-1}X_n$ under the measures $P_{\bar{\omega}}$ and $\mathbb{P}$, with good rate functions I_P^q and I_P^a, respectively (where $P(\mu_x(\{\infty\}) > 0) = 0$ if and only if $\boldsymbol{\lambda} \ge m$, cf. [10]). Moreover, we have seen in (3.2) that, for AGW-a.e. ω and all $t < \lambda_{\text{crit}}$,

$$(5.1) \qquad \lim_{n\to\infty} \frac{1}{n} \log E_{\boldsymbol{\lambda},\omega}\big(e^{tT_{-n}}\mathbf{1}_{T_{-n}<\infty}\big) = -G^-(t, P, 0),$$

whereas we have seen in (3.13) that

$$(5.2) \qquad \lim_{n\to\infty} \frac{1}{n} \log E_{\boldsymbol{\lambda}}\big(e^{tT_{-n}}\mathbf{1}_{T_{-n}<\infty}\big) = -\inf_{\eta\in M_1^{s,P}(\overline{\Omega}_\varepsilon)} [G^-(t, \eta, 0) + h(\eta|P)].$$

In particular, by Lemmas 1 and 3 and Varadhan's lemma, if $I_P^{-\tau,q} = I_P^{-\tau,a}$, then the limits in (5.1) and (5.2) must be equal for all $t < \lambda_{\text{crit}}$.

Let $|S_n|$ denote the distance of S_n from 0 in the tree ω. In [2] we derived the LDP for $n^{-1}|S_n|$ under both quenched and annealed measures, showing among other things that the rate function for both LDPs is the same. As announced in [2], Section 7, item 4, we show next that this is not the case for the rate functions $I_P^{-\tau,q}$ and $I_P^{-\tau,a}$ of the LDP of $n^{-1}T_{-n}$.

PROPOSITION 4. *If Z is bounded and nondegenerate, then for $t < \lambda_{\text{crit}}$ there exists an $\eta \in M_1^{s,P}(\overline{\Omega}_\varepsilon)$ such that*

$$\begin{aligned} -G^-(t, P, 0) &= E_P\big(\log E_{\bar{\omega}}\big(e^{tT_{-1}}\mathbf{1}_{T_{-1}<\infty}\big)\big) \\ &< E_\eta\big(\log E_{\bar{\omega}}\big(e^{tT_{-1}}\mathbf{1}_{T_{-1}<\infty}\big)\big) - h(\eta|P), \end{aligned}$$

except if $t = 0$ and $\mathbb{P}(T_{-1} < \infty) = 1$. That is, the limits in (5.1) *and* (5.2) *are different, and consequently $I_P^{-\tau,q} \neq I_P^{-\tau,a}$.*

PROOF. Fixing $0 < \boldsymbol{\lambda} < \infty$, recall that $T_{-n} = \sum_{k=0}^{\sigma_n-1}(H_k(Z_k) + 1)$ (in distribution), where $\sigma_n = \inf\{k \ge 0 : Z_k = -n\}$ for the biased simple random walk $\{Z_k\}$ starting at $Z_0 = 0$ such that $Z_k - Z_{k-1} = 1$ with probability $1/(1 + \boldsymbol{\lambda})$ and $Z_k - Z_{k-1} = -1$ otherwise. Recall that for fixed $\bar{\omega}$, $\{H_k(x),\ k \in \mathbb{N}\}$ are i.i.d., for each x, with distribution μ_x, and the biased simple random walk $\{Z_k\}$ is independent of $\{H_k(x),\ k \in \mathbb{N}\}$. Under the measure AGW, μ_x is an i.i.d. sequence. Fixing $t < \lambda_{\mathrm{crit}}$ let

$$V_t(x) := \log E^x_{\boldsymbol{\lambda},\omega}(e^{t(H(x)+1)}\mathbf{1}_{H(x)<\infty}) = t + \log \sum_{h=0}^{\infty} e^{th}\mu_x(\{h\}).$$

Note that

$$g(n) := \log E_{\boldsymbol{\lambda}}(e^{tT_{-n}}\mathbf{1}_{T_{-n}<\infty}) = \log E_{\mathrm{SRW}}\left(E_P\left(\exp\left(\sum_{k=0}^{\sigma_n-1} V_t(Z_k)\right)\right)\right),$$

where $E_{\mathrm{SRW}}(\cdot)$ denotes integration over all paths of the biased simple random walk $\{Z_k\}$. Since $V_t(x)$, $x \in \mathbb{Z}$, are i.i.d. random variables, they are positively correlated. This allows us to apply the FKG inequality for the increasing functions $\exp(\sum_{k=0}^{\sigma_m-1} V_t(Z_k))$ and $\exp(\sum_{k=\sigma_m}^{\sigma_{n+m}-1} V_t(Z_k))$, for each fixed path $(Z_0, Z_1, \dots, Z_{\sigma_{n+m}-1})$, yielding that

$$g(n+m) \ge \log E_{\mathrm{SRW}}\left(E_P\left(\exp\left(\sum_{k=0}^{\sigma_m-1} V_t(Z_k)\right)\right)E_P\left(\exp\left(\sum_{k=\sigma_m}^{\sigma_{n+m}-1} V_t(Z_k)\right)\right)\right).$$

Applying the strong Markov property of Z_k at the stopping time σ_m, where $Z_{\sigma_m} = -m$, it follows by the translation invariance of both the law of $\theta \mapsto \{Z_{\cdot+\theta} - Z_\theta\}$ and that of $\{V_t(\cdot)\}$, that

$$\begin{aligned} g(n+m) &\ge \log E_{\mathrm{SRW}}\left(E_P\left(\exp\left(\sum_{k=0}^{\sigma_m-1} V_t(Z_k)\right)\right)\right) \\ &\qquad + \log E_{\mathrm{SRW}}\left(E_P\left(\exp\left(\sum_{k=0}^{\sigma_n-1} V_t(Z_k)\right)\right)\right) \\ &= g(m) + g(n). \end{aligned}$$

Using the superadditivity of g and Jensen's inequality (for $\log x$), it follows that

$$\begin{aligned} \liminf_{n\to\infty} n^{-1} g(n) &\ge g(1) = \log E_{\boldsymbol{\lambda}}(e^{tT_{-1}}\mathbf{1}_{T_{-1}<\infty}) \\ &\ge \int \log E_{\boldsymbol{\lambda},\omega}(e^{tT_{-1}}\mathbf{1}_{T_{-1}<\infty})\mathrm{AGW}(d\omega) \end{aligned} \tag{5.3}$$

and the last inequality is strict as soon as $\varphi^-(t,\bar{\omega}) := E_{\boldsymbol{\lambda},\omega}(e^{tT_{-1}}\mathbf{1}_{T_{-1}<\infty})$ is a nondegenerate random variable. Note that the limits in (5.1) and (5.2) correspond to the right-hand and left-hand sides of (5.3), respectively. Thus, it suffices to show that, for Z nondegenerate, if $\varphi^-(t,\bar{\omega}) = c(t)$ for AGW-a.e. ω for some (finite) constant $c(t) > 0$, then necessarily $t = 0$ and $c(t) = 1$ [hence, $P_{\boldsymbol{\lambda},\omega}(T_{-1} < \infty) = 1$ for AGW-a.e. ω].

Turning to this task, note that we may add the ray $0 \leftrightarrow \infty$ to the MGW tree $\mathbf{T}_0$, thus making it a GW tree. With this identification, let $k_0 \geq 1$ be the number of children of 0 and let $N_0 := \sum_{k=1}^{T_{-1}} \mathbf{1}_{S_k=0}$ be the number of visits to vertex 0 by S_k prior to T_{-1}. Note that $T_{-1} = 1 + N_0 + \sum_{i=1}^{N_0} T_0^i(\omega_{r_i})$, where r_i denotes the child of 0 visited by S_k immediately after its $(i-1)$st visit of 0, with ω_{r_i} the GW tree rooted at that child and $T_0^i(\omega_{r_i})$ the time spent in this tree between the $(i-1)$st and ith visits to 0. Note that $P_{\boldsymbol{\lambda},\omega}(N_0 = \ell) = (k_0/(k_0+\boldsymbol{\lambda}))^\ell(\boldsymbol{\lambda}/(\boldsymbol{\lambda}+k_0))$ and the GW trees ω_{r_i} belong to the finite collection of k_0 trees rooted at children of 0, each being an independent realization of the same law as the original GW tree ω. Consequently, denoting by E_{k_0} expectation conditional on k_0,

$$\varphi^-(t,\bar{\omega}) = e^t E_{k_0}\Biggl(e^{tN_0}\prod_{i=1}^{N_0} E_{\boldsymbol{\lambda},\omega_{r_i}}\bigl(e^{tT_0^i}\mathbf{1}_{T_0^i<\infty}\bigr)\Biggr) = e^t E_{k_0}\Biggl(e^{tN_0}\prod_{i=1}^{N_0}\varphi^-\bigl(t,\bar{\omega}_{r_i}\bigr)\Biggr).$$

If $\varphi^-(t,\bar{\omega}) = c(t)$ for AGW-a.e. ω, then the same applies for the finite collection $\varphi^-(t,\bar{\omega}_{r_i})$ for AGW-a.e. ω, implying that $c(t)$ is a solution of the identity

$$c(t)e^{-t} = E_{k_0}\bigl((c(t)e^t)^{N_0}\bigr). \tag{5.4}$$

It is easy to verify that if Z is nondegenerate, so shall be the random variable $E_{k_0}(q^{N_0})$, provided $q \neq 1$, $0 < q < \infty$. Thus, if $\varphi^-(t,\bar{\omega}) = c(t)$ for AGW-a.e. ω and Z is nondegenerate, necessarily $c(t)e^t = q = 1$, which by (5.4) is possible only in case $t = 0$ and $c(0) = 1$, as stated. $\square$

6. Discussion and open problems.

1. We recall that CLT and stable limit laws for transient RWREs in an i.i.d. environment are derived in [9]. For recurrent RWREs, limit laws are derived in [14]. Process level limit laws of the form of singular diffusions are derived in [6] for the simple random walk with random holding times. It is natural to expect that even for i.i.d. environments the RWREH exhibits a rich spectrum of limit distributions due to the competition between traps coming from large holding times and those coming from the local drifts of the embedded RWRE. In particular, we expect a CLT to hold true whenever $\mathbb{E}(T_1^{2+\varepsilon}) < \infty$ for some $\varepsilon > 0$.
2. The study of sharp asymptotics in the slowdown regime for the RWRE has been carried out in a series of papers [3, 7, 12, 13]. Subexponential decay of slowdown probabilities is possible only for a.e. finite holding times, in which case it seems that the techniques of these papers can be extended to the RWREH. The possible subexponential rates of decay for the RWREH are influenced by the tails of the holding time distribution, and hence not limited to those present in the RWRE model.

Acknowledgments. We are grateful to Yuval Peres for helpful discussions and in particular for pointing out to us how useful the relation (4.12) really is. We also thank Martin Zerner for the argument leading to (5.3). This paper benefited from many valuable comments of the anonymous referees.

REFERENCES

[1] COMETS, F., GANTERT, N. and ZEITOUNI, O. (2000). Quenched, annealed and functional large deviations for one dimensional random walk in random environment. *Probab. Theory Related Fields* **118** 65–114. [Erratum **125** (2003) 42–44.]

[2] DEMBO, A., GANTERT, N., PERES, Y. and ZEITOUNI, O. (2002). Large deviations for random walks on Galton–Watson trees: Averaging and uncertainty. *Probab. Theory Related Fields* **122** 241–288.

[3] DEMBO, A., PERES, Y. and ZEITOUNI, O. (1996). Tail estimates for one-dimensional random walk in random environment. *Comm. Math. Phys.* **181** 667–684.

[4] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviation Techniques and Applications*, 2nd ed. Springer, New York.

[5] FÖLLMER, H. (1987). Random fields and diffusion processes. In *Ecole d'Eté de Saint Flour XV–XVII. Lecture Notes in Math.* **1362** 101–203. Springer, New York.

[6] FONTES, L. R. G., ISOPI, M. and NEWMAN, C. M. (2002). Random walks with strongly inhomogeneous rates and singular diffusions: Convergence, localization, and aging in one dimension. *Ann. Probab.* **30** 579–604.

[7] GANTERT, N. and ZEITOUNI, O. (1998). Quenched sub-exponential tail estimates for one-dimensional random walk in random environment. *Comm. Math. Phys.* **194** 177–190.

[8] GREVEN, A. and DEN HOLLANDER, F. (1994). Large deviations for a random walk in random environment. *Ann. Probab.* **22** 1381–1428.

[9] KESTEN, H., KOZLOV, M. V. and SPITZER, F. (1975). A limit law for random walk in a random environment. *Compositio Math.* **30** 145–168.

[10] LYONS, R., PEMANTLE, R. and PERES, Y. (1995). Ergodic theory on Galton–Watson trees: Speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems* **15** 593–619.

[11] LYONS, R., PEMANTLE, R. and PERES, Y. (1996). Biased random walk on Galton–Watson trees. *Probab. Theory Related Fields* **106** 249–264.

[12] PISZTORA, A. and POVEL, T. (1999). Large deviation principle for random walk in a quenched random environment in the low speed regime. *Ann. Probab.* **27** 1389–1413.

[13] PISZTORA, A., POVEL, T. and ZEITOUNI, O. (1999). Precise large deviation estimates for one-dimensional random walk in random environment. *Probab. Theory Related Fields* **113** 135–170.

[14] SINAI, YA. G. (1982). The limiting behavior of a one-dimensional random walk in random environment. *Theory Probab. Appl.* **27** 256–268.

[15] SION, M. (1958). On general minimax theorems. *Pacific J. Math.* **8** 171–176.

[16] SZNITMAN, A. S. (2001). Milieux aléatoires et petites valeurs propres. Etats de la Recherche, Luminy. In *Milieux Aléatoires*, *Panoramas et Synthèses* **12** (F. Comets and E. Pardoux, eds.). Société Mathématique de France.

[17] VARADHAN, S. R. S. (2003). Large deviations for random walks in a random environment. *Comm. Pure Appl. Math.* **56** 1222–1245.

[18] ZEITOUNI, O. (2004). Random walks in random environment. *XXXI Summer School in Probability*, *St. Flour*, *2001*. *Lecture Notes in Math.* **1873** 193–312.

[19] ZERNER, M. P. W. (1998). Lyapounov exponents and quenched large deviations for multidimensional random walk in random environment. *Ann. Probab.* **26** 1446–1476.

A. DEMBO
DEPARTMENT OF STATISTICS
AND DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: amir@stat.stanford.edu

N. GANTERT
INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT KARLSRUHE
ENGLERSTRASSE 2
76128 KARLSRUHE
GERMANY
E-MAIL: N.Gantert@math.uni-karlsruhe.de

O. ZEITOUNI
DEPARTMENT OF ELECTRICAL ENGINEERING
AND DEPARTMENT OF MATHEMATICS
TECHNION—ISRAEL INSTITUTE OF TECHNOLOGY
HAIFA 32000
ISRAEL
AND
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: zeitouni@math.umn.edu