Large Deviations for Random Walk in Random Environment with holding times

Amir Dembo,¹ Nina Gantert² and Ofer Zeitouni³

June 16, 2002.

Abstract: Suppose that the integers are assigned random variables \(\{\omega_x, \mu_x\}\) (taking values in the unit interval times 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 \(\mu_x\) and then, after one unit of time, moves one step to the right with probability \(\omega_x\), 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. As an application, we show that for random walks on Galton-Watson trees, quenched and annealed rate functions along a ray differ.

Key Words: Random walk in random environment, large deviations.

AMS (1991) subject classifications: 60J15, 60F10, 82C44, 60J80.

¹Departments of Statistics and of Mathematics, Stanford University, Stanford, CA 94305, USA. (e-mail: amir@stat.stanford.edu) Research partially supported by an NSF grant #DMS-0072331, a US-Israel BSF grant and by a Lady Davis fellowship at the Technion, Israel.

²Institut für Mathematische Stochastik, Universität Karlsruhe, Englerstr. 2, 76128 Karlsruhe (e-mail: N.Gantert@math.uni-karlsruhe.de) Research partially supported by the DFG. Part of this work was done while visiting the Dept. of Electrical Eng., Technion, Israel.

³Departments of Electrical Engineering and of Mathematics, Technion- Israel Institute of Technology, Haifa 32000, ISRAEL. (e-mail: zeitouni@ee.technion.ac.il) Partially supported by a US-Israel BSF grant.
1 Introduction and Statement of Results

We begin by giving a formal definition of the random walk in random environment with holding times (RWREH). Fix $\varepsilon > 0$, and $S_{\varepsilon} := [\varepsilon, 1 - \varepsilon] \times M_{\mu}([\mathbb{R}_+])$, where $\mathbb{R}_+ = \mathbb{R}_+ \cup \{\infty\}$ (with the usual topology of compactification at $\infty$) and $M_{\mu}([\mathbb{R}_+])$ denotes the space of Borel probability measures $\mu$ on $\mathbb{R}_+$ such that $\mu((0, 1\backslash \varepsilon]) \geq \varepsilon$. An environment $\mathfrak{w} \in S_{\varepsilon} := \mathbb{P}_\varepsilon$ has coordinates $\mathfrak{w}_x = (\omega_x, \mu_x) \in S_{\varepsilon}$. For each $\mathfrak{w} \in \mathbb{P}_\varepsilon$, we define the random walk in random environment (RWRE) $\{Z_n\}$ on $\mathbb{Z}$ as the Markov process with $Z_0 = 0$ and transition probabilities

$$\tilde{P}_\mathfrak{w}(Z_{n+1} = z + 1 | Z_n = z) = \omega_z,$$

$$\tilde{P}_\mathfrak{w}(Z_{n+1} = z - 1 | Z_n = z) = 1 - \omega_z.$$

Define next

$$\Theta_j = \sum_{i=0}^{j-1} \left( H_i(Z_i) + 1 \right)$$

where $\{H_i(x)\}_{i \in \mathbb{N}}$ are independent random variables with $\mathcal{L}(H_i(x)) = \mu_x$, for all $i \in \mathbb{N}$, and set $s_t = \max\{j : \Theta_j \leq t\}$. Define the RWRE with holding times (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 $\mu_x$ before, one unit of time later, jumping to one of its nearest neighbors, with jumps to the right occurring with probability $\omega_x$. The environment $\mathfrak{w}$ is chosen according to the probability measure $P$, and fixed thereafter. Let $M^s_1(\mathbb{P}_\varepsilon)$ and $M^f_1(\mathbb{P}_\varepsilon)$ denote the stationary, or stationary and ergodic, respectively, probability measures on $\mathbb{P}_\varepsilon$, with respect to the shift $\theta : \mathbb{P}_\varepsilon \to \mathbb{P}_\varepsilon$ such that $(\theta \mathfrak{w})_i = \mathfrak{w}_{i+1}$. We will always assume that $P \in M^f_1(\mathbb{P}_\varepsilon)$. We let $P_\mathfrak{w}$ denote the law on the process $\{X_t\}$, conditioned on a realization $\mathfrak{w} \in \mathbb{P}_\varepsilon$ (the quenched law). We use $P$ both for $P \times P_\mathfrak{w}$ and for its marginal on $\mathbb{Z}^{\mathbb{R}_+}$ induced by $\{X_t\}_{t \geq 0}$, and refer to both as the annealed law.

The study of RWRE was initiated in the mid seventies, and in the last decade there was a resurgence of interest and results for this model, see [12] and [13] for recent reviews. Our interest in the large deviations for the RWREH originated from three different sources:

1. In [2], we considered large deviation for random walks on Galton-Watson trees, and showed that in contrast to RWRE on $\mathbb{Z}$, and in contrast to the conjectured behavior of RWRE on $\mathbb{Z}^d$, quenched and annealed large deviation principles for the random walk on the Galton-Watson tree coincide. We conjectured in [2] that if one restricts attention to a particular ray in the tree, one would recuperate differences between quenched and annealed behavior. In Section 3, we show, using our analysis of the RWREH, that indeed this is the case.

2. In [1], the large deviations for the RWRE, both quenched and annealed, are considered. While preparing the notes [13], we noted that some of the proofs do not carry over to the set-up 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) holding times, and proved that the suitably rescaled process converges to a singular diffusion. They also considered aging for such processes. This led us naturally to questions concerning large deviations for their model, and eventually to the more general RWREH model.

Turning to a description of our results, define

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

Using the same arguments as in [13], one can show that

$$\frac{X_t}{t} \to v_P \quad \mathbb{P} - a.s.$$ 

where

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

As in [1], our first goal is to describe the large deviation properties of $T_n/n$, both quenched and annealed (we refer to [4] for definitions and basic properties of large deviation principles).

Set

$$\varphi(\lambda, \omega) = E_{\pi}\left(e^{\lambda T_{1} \mathbb{1}_{T_{1} < \infty}}\right), \quad f(\lambda, \omega) = \log \varphi(\lambda, \omega), \quad G(\lambda, P, u) = \lambda u - E_{P}\left(f(\lambda, \omega)\right),$$

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

$$\varphi^{-}(\lambda, \omega) = E_{\pi}\left(e^{\lambda T_{-1} \mathbb{1}_{T_{-1} < \infty}}\right), \quad f^{-}(\lambda, \omega) = \log \varphi^{-}(\lambda, \omega), \quad G^{-}(\lambda, P, u) = \lambda u - E_{P}\left(f^{-}(\lambda, \omega)\right),$$

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

**Theorem 1 Quenched LDP for $T_n/n$**

For $P$-almost every $\omega$, the sequence $T_n/n$ satisfies a weak LDP in $\mathbb{R}$ under $P_{\pi}$ with the convex rate function $I_P^{q,q}(u)$, and the sequence $T_{-n}/n$ satisfies a weak LDP in $\mathbb{R}$ under $P_{\pi}$ with the convex rate function $I_P^{-q,q}(u)$. Further,

$$I_P^{q,q}(u) = I_P^{-q,q}(u) + E_P(\log \rho_x)$$

(1.1)

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

An annealed LDP for $T_n/n$ requires few more notations and assumptions on $P$. Equip $M_1(\mathbb{P}_{+})$ with the topology induced by weak convergence and $S_x$ with the corresponding product topology. Putting on $\Omega_x$ the product topology and on $M_1(\Omega_x)$ the corresponding topology of weak convergence, we see that $S_x$, $\Omega_x$ and $M_1(\Omega_x)$ are compact metric spaces. With $R_n = n^{-1} \sum_{j=0}^{n-1} \delta_{\omega_{T_j}} \in M_1(\Omega_x)$ the empirical process determined by the environment, set
(C1) \( \{R_n\} \) satisfies under \( P \) the LDP in \( M_1(\Omega) \) with good rate function \( h(\cdot | P) \). Here we assume that the specific entropy \( h(\eta | P) = \lim_{m \to \infty} m^{-1} H(\eta | P) \) with respect to \( P \) exists for any stationary \( \eta \) (where \( \eta | P \) and \( P | \eta \) are the restrictions of \( \eta \) and \( P \) to \( \{\omega_i\}_{i=0}^m \)), and set \( h(\eta | P) = \infty \) for non-stationary \( \eta \).

(C2) \( P \) is locally equivalent to the product of its marginals (i.e., for any \( A \in S^m \) and \( m \) finite, \( P | A = 0 \) if and only if \( (P | 1)^m (A) = 0 \)). Moreover, for any stationary measure \( \eta \in M_1(\Omega) \) with \( h(\eta | P) < \infty \), there is a sequence \( \{\eta^n\} \subset M_1(\Omega) \) 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 \( \eta^n \) 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 non-random \( b < \infty \) and \( k(\delta) > 0 \), such that \( \mu_0([0, b - \delta]) = 0 \) and \( \mu_0([0, b + \delta]) \geq k(\delta) \) for all \( \delta > 0 \) and \( P \)-almost every \( \omega \) (in case \( b = 0 \), \( [0, b + \delta] \) is the empty interval).

As noted for example in [5, Theorems 3.10, 4.1 and Lemma 4.8], the conditions (C1)–(C2) hold when the stationary and ergodic \( P \) corresponds to a Markov process whose transition kernel \( P(\omega_{x+1} | \omega_x) \) has a bounded and bounded away from 0 density with respect to some reference probability measure on \( S \) (in particular, (C1)–(C2) hold when \( P \) is a stationary product measure).

We now have the

**Theorem 2** Annealed LDP for \( T_n/n \)

Assume (C1)–(C3). Then the sequence \( T_n/n \) satisfies a weak LDP in \( \mathbb{R} \) under \( P \) with the convex rate function

\[
I^\tau_P(u) = \inf_{\eta \in M_1(\Omega)} \left[ I^\tau_\eta q(u) + h(\eta | P) \right],
\]

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

\[
I^-\tau_P(u) = \inf_{\eta \in M_1(\Omega)} \left[ I^-\tau_\eta q(u) + h(\eta | P) \right].
\]

We next state the large deviations of the rescaled positions \( X_{t}/t \).

**Theorem 3** Quenched LDP for \( X_{t}/t \)

For \( P \)-almost every \( \omega \), \( X_{t}/t \) satisfies a LDP under \( P_{\omega} \) with the good convex rate function \( I^\tau_P(\cdot) \).

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

\[
I^\tau_P(v) = \begin{cases} \frac{v}{I^\tau_P(\frac{1}{P})}, & v > 0 \\ \frac{|v|}{I^-\tau_P(\frac{1}{P})}, & v < 0 \end{cases}
\]

and \( I^\tau_P(0) := \lim_{u \to 0} I^\tau_P(u) \).
(b) If \( P(\mu_0(\{\infty\}) > 0) > 0 \), then

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

(1.3)

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

**Theorem 4** Annealed LDP for \( X_t/t \)

Assume (C1)–(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

\[
I_P^a(v) = \begin{cases} 
v I_P^{\tau,a} \left( \frac{1}{t} \right), & v > 0 \\
\left| v \right| I_P^{\tau,a} \left( \frac{1}{|t|} \right), & v < 0
\end{cases}
\]  

(1.4)

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

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

\[
\max_{1 \leq j \leq k} E_P(\mu_j(\{\infty\})|\mathcal{F}_m^k) \geq e^{-ck} \quad \forall k \geq k_0,
\]  

(1.5)

where \( \mathcal{F}_m = \sigma(\{\overline{x}, x \leq m\}) \). Then \( X_t/t \) satisfies an LDP under \( \mathbb{P} \) with the good convex rate function

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

(1.6)

Clearly, (1.5) holds when \( P \) is a stationary product measure (and more generally, under suitable mixing conditions).

The structure of the article is as follows: in the next section, we present the proofs of Theorems 1–4, emphasizing in particular those elements of the proofs that differ from [1]. Section 3 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 4.

## 2 Proof of Theorems 1–4

We begin with the following strengthening of [1, Lemma 2]:
Lemma 1 There exist constants $\lambda_{\text{crit}} = \lambda_{\text{crit}}(P), \lambda_{\text{crit}}^{-} = \lambda_{\text{crit}}^{-}(P) \in [0, \infty)$ such that, $P$-a.s.

$$
\varphi(\lambda, \omega) \begin{cases} 
\leq e^{-2}, & \lambda \leq \lambda_{\text{crit}}^{-} \\
= \infty, & \lambda > \lambda_{\text{crit}}^{-}
\end{cases} \quad \varphi^{-}(\lambda, \omega) \begin{cases} 
\leq e^{-2}, & \lambda \leq \lambda_{\text{crit}}^{-} \\
= \infty, & \lambda > \lambda_{\text{crit}}^{-}
\end{cases}
$$

(2.1)

We will see later, see the remark below (2.18), that $\lambda_{\text{crit}}^{-}(P) = \lambda_{\text{crit}}(P)$.

Proof: By the transformation $\{(\omega_x, \mu_x)\} \to \{(1 - \omega_x, \mu_x)\}$, it is enough to consider $\varphi(\lambda, \omega)$. By path decomposition, for each $\lambda$,

$$
\varphi(\lambda, \omega) = \omega_0 e^{\lambda H} e^{\lambda H 1_{H_{<\infty}}} + (1 - \omega_0) E_{\mu_0} e^{\lambda H} e^{\lambda H 1_{H_{<\infty}}} \varphi(\lambda, \omega),
$$

(2.2)

where $H$ is a random variable with distribution $\mu_0$, and $E_{\mu_0}$ denotes expectation with respect to $\mu_0$. Thus, $\varphi(\lambda, \omega) < \infty$ implies that $\varphi(\lambda, \theta^{-1} \omega) < \infty$, yielding by the ergodicity of $P$ that $1_{\varphi(\lambda, \omega) < \infty}$ is constant a.s., for all $\lambda$ rational at once. This, and the monotonicity of $\varphi(\lambda, \omega)$ in $\lambda$, immediately yield the existence of a deterministic $\lambda_{\text{crit}}$ with $\varphi(\lambda, \omega) < \infty$ for all $\lambda < \lambda_{\text{crit}}$, $P$-a.s.. By definition, $\varphi(\lambda, \omega) \leq 1$ for $\lambda \leq 0$, whereas for $\lambda \geq 0$, the fact that $\varphi(\lambda, \omega) < \infty$ implies by (2.2) that $\varphi(\lambda, \theta^{-1} \omega) \leq \frac{1}{(1 - \omega_0) E_{\mu_0} 1_{H_{<\infty}}} e^\lambda \leq \frac{e^\lambda}{e^\lambda}, P$-a.s. We conclude that $\varphi(\lambda, \omega) \leq e^{-2}$ for all $\lambda < \lambda_{\text{crit}}$, $P$-a.s., and monotone convergence then implies that $\varphi(\lambda_{\text{crit}}, \omega) \leq e^{-2} < \infty, P$-a.s. \hfill \Box

Set $\overline{\mu} = \overline{\mu}(P) = \int \frac{E(T_1 1_{T_1 < \infty})}{E(1_{T_1 < \infty})} P(d\omega) \in [1, \infty]$. Define, for $\lambda < \lambda_{\text{crit}},$

$$
g(\lambda) := \int \frac{E(T_1 e^{\lambda T_1} 1_{T_1 < \infty})}{E(e^{\lambda T_1} 1_{T_1 < \infty})} P(d\omega) = \int \frac{d}{d\lambda} f(\lambda, \omega) P(d\omega) < \infty.
$$

(2.3)

Since $\lambda \to g(\lambda)$ is increasing (check that $g' \geq 0$!), it follows that $u_+ = u_+(P) = \lim_{\lambda \nearrow \lambda_{\text{crit}}} g(\lambda) \geq \overline{\mu} = g(0)$ exists (with possible value $u_+ = +\infty$). Set

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

Since $g(\lambda)$ is strictly increasing and continuous in $\lambda$, 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 $u < \overline{\mu}$ then $\lambda_u < 0$, and hence

$$
I^q_P(u) = \sup_{\lambda \in \mathbb{R}} G(\lambda, P, u) = G(\lambda_u, P, u) = \sup_{\lambda \leq 0} G(\lambda, P, u), \quad u \leq \overline{\mu}.
$$

(2.4)

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

$$
I^q_P(u) = \sup_{\lambda \in \mathbb{R}} G(\lambda, P, u) = \sup_{\lambda \geq 0} G(\lambda, P, u), \quad u \geq \overline{\mu}.
$$

(2.5)

Further, we have the following.
Lemma 2 For any $P \in M^*_1(\Omega_\epsilon)$, the convex rate function $I_P^{r,q}(\cdot)$ is infinite on $(-\infty, u_- (P))$, finite on $(u_- (P), \infty)$, non-increasing on $[u_- (P), \overline{\mu}(P)]$, and non-decreasing on $[\overline{\mu}(P), \infty)$. Moreover, $I_P^{r,q}(\overline{\mu}(P)) = G(0, P, \overline{\mu}(P))$ when $\overline{\mu}(P) < \infty$, $\lambda_{\text{crit}} (P) = 0$ whenever $\overline{\mu}(P) = \infty$, and for all $u$,

$$\sup_{\lambda \geq 0} G(\lambda, P, u) = \inf_{w \geq u} I_P^{r,q}(w), \quad (2.6)$$

and

$$\sup_{\lambda \leq 0} G(\lambda, P, u) = \inf_{w \leq u} I_P^{r,q}(w). \quad (2.7)$$

Proof: By definition $I_P^{r,q}$ is convex and lower semi-continuous. Since $G(0, P, u) \geq 0$, it is also non-negative. Suppose $u_1 < u_2 \leq \overline{\mu}(P)$. Then, by (2.4)

$$\sup_{\lambda \in \mathbb{R}} G(\lambda, P, u_1) = \sup_{\lambda \leq 0} G(\lambda, P, u_1) = \sup_{\lambda \leq 0} \left[ \lambda u_1 - E_P( f(\lambda, \overline{\omega}) \right]$$

$$\geq \sup_{\lambda \leq 0} \left[ \lambda u_2 - E_P( f(\lambda, \overline{\omega}) \right] = \sup_{\lambda \in \mathbb{R}} G(\lambda, P, u_2).$$

To see that $I_P^{r,q}$ is non-decreasing on $[\overline{\mu}(P), \infty)$, use a similar argument with (2.5) instead of (2.4).

If $u < u_- (P)$ then $\sup_{\xi} \{ u - g(\xi) \} < 0$, and since $G(0, P, u) \geq 0$ we see that $G(\lambda, P, u) = G(0, P, u) + \int_{0}^{\lambda} (u - g(\xi)) d\xi \to \infty$ when $\lambda \to -\infty$, resulting with $I_P^{r,q}(u) = \infty$. In contrast,

$$\varphi(\lambda, \overline{\omega}) \geq e^2 e^{-|\lambda| (1+\epsilon^{-1})} := 1/c_\lambda \quad \forall \lambda \in \mathbb{R}, \ \overline{\omega} \in \overline{\Omega}_\epsilon, \quad (2.8)$$

hence setting $\lambda_u := \lambda_{\text{crit}}$ when $u \geq u_+(P)$, we see that $I_P^{r,q}(u) = G(\lambda_u, P, u) \leq \lambda_u u + \log c_\lambda u < \infty$ for any $u > u_-(P)$.

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

$$\lambda \int E_\varphi(T_1 | T_1 < \infty) P(d\overline{\omega}) \leq E_P \left( \log E_P(e^{\lambda T_1} | T_1 < \infty) \right) = E_P(f(\lambda, \overline{\omega})) - E_P(f(0, \overline{\omega})).$$

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

Turning to prove (2.6) and (2.7), consider first $\overline{\mu}(P) = \infty$, in which case $I_P^{r,q}(\cdot)$ is non-increasing and (2.7) follows from (2.4). Further, the convex, lower semi-continuous function $\lambda \mapsto E_P(f(\lambda, \overline{\omega}))$ is then infinite when $\lambda > 0$. Hence, by duality of Fenchel-Legendre transforms, for all $u$,

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

which amounts to (2.6). Suppose now that $\overline{\mu}(P) < \infty$. Since $I_P^{r,q}(\cdot)$ is non-decreasing on $[\overline{\mu}(P), \infty)$ we get (2.6) for $u \geq \overline{\mu}(P)$ out of (2.5). Moreover, $I_P^{r,q}(u)$ is non-increasing for $u \leq \overline{\mu}(P)$, hence for such $u$ the right side of (2.6) equals $I_P^{r,q}(\overline{\mu}(P)) = G(0, P, \overline{\mu}(P)).$ Further, then $G(\lambda, P, u) \leq$
\(G(\lambda, P, \overline{\pi}(P))\) for all \(\lambda \geq 0\), with equality when \(\lambda = 0\), implying the left side of (2.6) also equals \(G(0, P, \overline{\pi}(P))\), thus completing its proof. The proof of (2.7) is similar. Combining (2.4) with the monotonicity of \(I_P^{\pi,q}(u)\) gives (2.7) for \(u \leq \overline{\pi}(P)\), whereas for \(u \geq \overline{\pi}(P)\) both sides of (2.7) equal \(G(0, P, \overline{\pi}(P))\). \qed

We have completed the preliminaries needed for the

**Proof of Theorem 1.** With \(T_0 = 0\) and \(\tau_i = T_i - T_{i-1}, i = 1, 2, \ldots\), we have that for any \(\overline{\omega}\) and \(\lambda \leq \lambda_{\text{crit}}\),

\[
E_{\overline{\omega}}\left(e^{\lambda T_n 1_{T_n < \infty}}\right) = E_{\overline{\omega}}\left(e^{\lambda \sum_{i=1}^{n} \tau_i 1_{\tau_i < \infty}}\right) = \prod_{i=1}^{n} \varphi(\lambda, \theta^i \overline{\omega}),
\]

(2.9)

where the second equality is due to the Markov property. By Lemma 1 and (2.8) it follows that \(|f(\lambda, \overline{\omega})| \leq c(\lambda) < \infty\) for some constant \(c(\lambda)\), all \(\lambda \leq \lambda_{\text{crit}}\) and \(P\)-a.e. \(\overline{\omega}\). An application of Birkhoff’s pointwise ergodic theorem then yields that

\[
\frac{1}{n} \log E_{\overline{\omega}}\left(e^{\lambda T_n 1_{T_n < \infty}}\right) = \frac{1}{n} \sum_{i=1}^{n} f(\lambda, \theta^i \overline{\omega}) \longrightarrow \int f(\lambda, \overline{\omega}) P(d\overline{\omega}) \quad P - \text{a.s.}
\]

(2.10)

first for all \(\lambda\) rational and then for all \(\lambda \leq \lambda_{\text{crit}}\) by monotonicity. Fixing \(u \in \mathbb{R}\), by Chebycheff’s inequality, for all \(\overline{\omega}\) and \(\lambda \leq 0\)

\[
P_{\overline{\omega}}\left(\frac{T_n}{n} \leq u\right) \leq e^{-\frac{\lambda u}{n}} E_{\overline{\omega}}\left(e^{\lambda T_n 1_{T_n < \infty}}\right).
\]

(2.11)

Thus, by (2.10), \(P\)-a.s. for all \(u\),

\[
\limsup_{n \to \infty} \frac{1}{n} \log P_{\overline{\omega}}\left(\frac{T_n}{n} \leq u\right) \leq -\sup_{\lambda \leq 0} G(\lambda, P, u) = -\inf_{w \leq u} I_P^{\pi,q}(w),
\]

(2.12)

where (2.7) was used in the rightmost equality. The upper bound on the upper tail is derived similarly. Indeed, using Chebycheff’s inequality with \(\lambda > 0\),

\[
P_{\overline{\omega}}\left(\infty > \frac{T_n}{n} \geq u\right) \leq e^{-\frac{\lambda u}{n}} E_{\overline{\omega}}\left(e^{\lambda T_n 1_{T_n < \infty}}\right),
\]

and hence, as in (2.12), using now (2.6), \(P\)-a.s. for all \(u\),

\[
\limsup_{n \to \infty} \frac{1}{n} \log P_{\overline{\omega}}\left(\infty > \frac{T_n}{n} \geq u\right) \leq -\sup_{\lambda \geq 0} G(\lambda, P, u) = -\inf_{w \geq u} I_P^{\pi,q}(w)
\]

(2.13)

Suppose \(\overline{\pi} < \infty\). Any closed set \(F \subseteq [1, \infty)\) is contained in \([1, u_1] \cup [u_2, \infty)\) for some \(u_1 \leq \overline{\pi} \leq u_2\) such that \(u_1 \in F\) and \(u_2 \in F\) (ignoring \(u_2\) if \(F \subseteq [1, \overline{\pi}]\) and \(u_1\) if \(F \subseteq [\overline{\pi}, \infty)\)). So, by the monotonicity of \(I_P^{\pi,q}(\cdot)\) (proved in Lemma 2), the inequalities (2.12) and (2.13) yield the upper bound for a general closed set \(F\). If \(\overline{\pi} = \infty\) and \(K \subseteq [1, \infty)\) is compact, then \(K \subseteq [1, u_1]\) for some \(u_1 \in K\) and (2.12) yields the upper bound needed for the weak LDP of Theorem 1.
Due to the continuity of \( T_P^u q(\cdot) \) in the interior of its domain, implied by Lemma 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 \( Q_{\pi, n} \) such that

\[
\frac{dQ_{\pi, n}}{dP_\pi} = \frac{1}{Z_n \pi} \exp(\lambda u T_n) 1_{T_n < \infty}, \quad Z_n \pi = E_\pi (\exp(\lambda u T_n) 1_{T_n < \infty})
\]

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

\[
P_\pi \left( \left| \frac{T_n}{n} - u \right| < \delta \right) \geq \exp \left( -n u \lambda_u - n \delta |\lambda_u| + \sum_{i=1}^{n} f(\lambda_u, \theta_i \omega) \right) \overline{Q}_\pi \left( \left| \frac{T_n}{n} - u \right| < \delta \right). \tag{2.14}
\]

Since \( P \) is ergodic and \( u < u_+ \), it holds that

\[
E_{\overline{Q}_\pi} \left( \frac{T_n}{n} \right) = \frac{1}{n} \sum_{i=1}^{n} E_{\overline{Q}_\pi} \tau_i \rightarrow_{n \rightarrow \infty} E_P \left( E_{\overline{Q}_\pi} \tau_1 \right) = g(\lambda_u) = u, \quad P\text{-a.s.} \tag{2.15}
\]

where we have also used (2.3). With \( \lambda_u < \lambda_{\text{crit}} \), it also holds that there exists a \( \beta > 0 \) such that

\[
E_P \left( E_{\overline{Q}_\pi} (e^{\beta \tau_1}) \right) < \infty,
\]

implying by Chebycheff’s inequality and independence that

\[
\overline{Q}_\pi \left( \left| \frac{T_n}{n} - u \right| \geq \delta \right) \rightarrow_{n \rightarrow \infty} 0, \quad P\text{-a.s.} \tag{2.16}
\]

Combining (2.16) with (2.14), we get

\[
\liminf_{n \rightarrow \infty} \frac{1}{n} \log P_\pi \left( \left| \frac{T_n}{n} - u \right| \geq \delta \right) \geq -u \lambda_u - \delta |\lambda_u| + \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{n} f(\lambda_u, \theta_i \omega)
\]

\[
= -u \lambda_u - \delta |\lambda_u| + E_P(f(\lambda_u, \omega))
\]

\[
= -G(\lambda_u, P, u) - \delta |\lambda_u| = -T_P(u) - \delta |\lambda_u| \quad P\text{-a.s.} \tag{2.17}
\]

(the first equality is due to Birkhoff’s ergodic theorem and the last one to (2.4)). This completes the proof of the lower bound when \( u < u_+ \), since \( \delta > 0 \) is arbitrary.

Fixing a rational \( u \geq u_+ \), the convex functions \( G_m(\lambda, P, u) = \lambda u - E_P(\log e^{\lambda T_1} 1_{T_1 \leq m}) \), \( m > \varepsilon^{-1} + 1 \), are finite and smooth in \( \lambda \in \mathbb{R} \). Moreover, \( P_\pi(u + 2 \leq T_1 < \infty) > 0 \) for \( P\text{-a.e.} \ \pi \), implying that for any \( m > m_0(u) \) the function \( g_m(\lambda) = \frac{d}{d\lambda} G_m(\lambda, P, u) \) is negative for \( \lambda \) large enough. For such \( m \) there exists \( \lambda_{u, m} \in [0, \infty) \) such that \( g_m(\lambda_{u, m}) = 0 \) and \( G_m(\lambda, P, u) \rightarrow -\infty \) as \( \lambda \rightarrow \infty \). 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 \( Q_{\Sigma, m} \) of \( \{ \tau_1, \ldots, \tau_n \} \) under the probability measure \( Q_{\Sigma, n, m} \) defined by

\[
\frac{dQ_{\Sigma, n, m}}{dP_\Sigma} = \frac{1}{Z_{n, m, \Sigma}} \prod_{i=1}^n e^{\lambda_{u,m} \tau_i} 1_{\tau_i \leq m}, \quad Z_{n, m, \Sigma} = \prod_{i=1}^n E_{\theta_{\Sigma}}(e^{\lambda_{u,m} \tau_i} 1_{T_i \leq m}).
\]

Adapting in such a manner the argument leading to (2.17), one obtains the bound

\[
\lim_{\delta \to 0} \lim_{n \to \infty} \inf \frac{1}{n} \log P_\Sigma \left( \left| \frac{T_n}{n} - u \right| \geq \delta \right) \geq -G_m(\lambda_{u,m}, P, u) = -\sup_{\lambda \geq 0} G_m(\lambda, P, u) := -I_m(u), \quad P\text{-a.s.}
\]

(for details, see [1, Proof of Theorem 4]). With \( G_m(\lambda, P, u) \) non-increasing in \( m \), so are the finite, non-negative constants \( I_m(u) \). Denoting by \( I_\infty(u) \) the finite, non-negative limit of \( I_m(u) \), the intersection of the non-empty, compact, nested sets \( \{ \lambda \geq 0 : G_m(\lambda, P, u) \geq I_\infty(u) \}, \ m > m_0(u), \) 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.

Finally, the proof of (1.1) follows verbatim from the proof of [13, Lemma 2.2.1] which gives

\[
E_P \left( \log E_{\Sigma}\left( e^{\lambda_{r-1} \tau_1} 1_{T_{r-1} < \infty} \right) \right) = E_P \left( \log E_{\Sigma}\left( e^{\lambda_{r} \tau_1} 1_{T_1 < \infty} \right) \right) + E_P(\log \rho_0), \quad (2.18)
\]

and allows to relate \( I_P^{\tau,q}(\cdot) \) and \( I_P^{\tau,q}(\cdot) \), in the same way as in the case without holding times. \( \blacksquare \)

**Remark:** Note that (2.18) implies that \( \lambda_{\text{crit}}(P) = \lambda_{\text{crit}}(P) \).

**Proof of Theorem 2.** Since the proof of the annealed weak LDP for \( T_{r}/n \) is almost identical to that for \( T_{r}/n \) we present in the sequel only the latter.

We begin with a lemma giving a characterization of \( \lambda_{\text{crit}}(\eta) \) for “nice” \( \eta \). 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 3** Assume \( \eta \in M_1(\Sigma_\epsilon) \) is locally equivalent to the product of its marginals. Let \( \Sigma := \text{supp } \eta 1 \subseteq S_\epsilon \). Then,

\[
\lambda_{\text{crit}}(\eta) = \inf_{\omega \in \Sigma_\epsilon} \lambda_c(\omega) =: \lambda \geq 0, \quad (2.19)
\]

where \( \lambda_c(\omega) := \sup \left\{ \lambda : E_\omega(\epsilon^{\lambda \tau_1} 1_{T_1 < \infty}) < \infty \right\} \). Moreover,

\[
\varphi(\lambda, \omega) \leq \epsilon^{-2}, \quad \forall \lambda \leq \lambda_{\text{crit}}(\eta), \quad \forall \omega \in \Sigma_\epsilon. \quad (2.20)
\]

Suppose \( \eta^n \) is a sequence in \( M_1(\Sigma_\epsilon) \) such that \( \eta^n(\Sigma_\epsilon) = 1 \) and all the \( \eta^n \) are locally equivalent to the product of their marginals. If \( \eta^n \to \tilde{\eta} \) for some \( \tilde{\eta} \in M_1(\Sigma_\epsilon) \) such that \( \eta \leq \tilde{\eta} \) then,

\[
\lambda_{\text{crit}}(\eta^n) \to \lambda_{\text{crit}}(\eta). \quad (2.21)
\]
Proof: Let $g_m(t) := \min(m - t, 1)\mathbf{1}_{[0,m]}(t)$ and
\[
\varphi_m(\lambda, \overline{\omega}) := E_{\overline{\omega}}\left( e^{\lambda T_1} g_m(T_1) \right).
\] (2.22)
Note that $\varphi_m(\lambda, \overline{\omega})$ is continuous on $\overline{\Omega}_\varepsilon$. Indeed, $\varphi_m(\lambda, \overline{\omega})$ depends only on $(\overline{\omega}_0, \overline{\omega}_1, \ldots, \overline{\omega}_m)$. Moreover, it is the sum over the contributions $\varphi_m(\lambda, \overline{\omega}, z)$ of the finitely many possible path $z$ of the RWRE $Z_i = z_i$, $i = 0, \ldots, j$, where $z_0 = 0, z_j = 1$ and $z_i \leq 0$ for $i < j \leq m$. Fixing such a path, denote by $\mu_z$ the law of $T_1 = \Theta_j$ conditional on $\{Z_0, Z_1, \ldots, Z_j\}$. With $t \mapsto e^{\lambda t} g_m(t)$ bounded and continuous,
\[
\varphi_m(\lambda, \overline{\omega}, z) = \prod_{i=0}^{j-1} \frac{1}{2} + (z_{i+1} - z_i)(\overline{\omega}_{z_i} - \frac{1}{2}) \int e^{\lambda t} g_m(t) \mu_z(dt),
\]
is continuous in $(\overline{\omega}_0, \overline{\omega}_1, \ldots, \overline{\omega}_m)$ and $\mu_z$, hence also in $(\overline{\omega}_x, x = 0, \ldots, m+1)$.

Fixing $\lambda \leq \lambda_{\text{crit}}(\eta)$ we know from Lemma 1 that (2.20) holds for $\eta$-a.e. $\overline{\omega}$. We next show that (2.20) holds for all $\overline{\omega} \in \Sigma^\varepsilon$. Suppose to the contrary that $\varphi(\lambda, \overline{\omega}) > \varepsilon^{-2}$ for some $\overline{\omega} \in \Sigma^\varepsilon$. By monotone convergence and continuity of $\varphi_m$, there exists $m$ large enough such that the open subset $G := \{ (\overline{\omega}_0, \ldots, \overline{\omega}_m) : \varphi_m(\lambda, \overline{\omega}) > \varepsilon^{-2} \}$ of $S^m_\varepsilon$ intersects $\text{supp} (\eta_1)^m$ at $(\overline{\omega}_0, \ldots, \overline{\omega}_m)$. Clearly $(\eta_1)^m(G) > 0$, and with $\eta$ locally equivalent to the product of its marginals, also $\eta_1(G) > 0$. Recall that $\varphi(\lambda, \overline{\omega}) \geq \varphi_m(\lambda, \overline{\omega})$, implying that $\eta(\varphi(\lambda, \overline{\omega}) > \varepsilon^{-2}) > 0$ in contradiction with Lemma 1.

If $\lambda < \overline{\lambda}$ then by definition $\varphi(\lambda, \overline{\omega}) < \infty$ for all $\overline{\omega} \in \Sigma^\varepsilon$, hence $\lambda_{\text{crit}}(\eta) \geq \lambda$ by Lemma 1. Consequently, $\lambda_{\text{crit}}(\eta) \geq \overline{\lambda}$. For any $\overline{\omega} \in \Sigma^\varepsilon$, the inequality (2.20) implies that $\lambda_c(\overline{\omega}) \geq \lambda_{\text{crit}}(\eta)$, hence by definition also $\overline{\lambda} \geq \lambda_{\text{crit}}(\eta)$.

Turning to prove (2.21), note that as $\text{supp} \eta_1 \subseteq \Sigma$, we have from (2.19) that $\lambda_{\text{crit}}(\eta_1) \geq \overline{\lambda} = \lambda_{\text{crit}}(\eta)$ and if $\lambda > \lambda_{\text{crit}}(\eta)$, then $\varphi(\lambda, \overline{\omega}) = \infty > \varepsilon^{-2}$ for some $\overline{\omega} \in \Sigma^\varepsilon$. Taking $m$ and the open $G \subseteq S^m_\varepsilon$ as in the preceding proof of (2.20), we have that $\eta_1(G) > 0$. Since $\eta \ll \tilde{\eta}$ and $\eta \rightarrow \tilde{\eta}$, also $\eta_1(G) > 0$ for all $n$ large enough. Consequently, $\eta(\varphi(\lambda, \overline{\omega}) > \varepsilon^{-2}) > 0$, implying that $\lambda > \lambda_{\text{crit}}(\eta_1)$ for all $n$ large enough (c.f. Lemma 1). Considering $\lambda \downarrow \lambda_{\text{crit}}(\eta)$ completes the proof of (2.21).

With $M^P_1 = M^P_1(\overline{\Omega}_\varepsilon) := \{ \nu \in M_1(\overline{\Omega}_\varepsilon) : \text{supp} \nu \subseteq (\text{supp} P_1)^2 \}$, the next lemma is the analogue of [1, Lemma 6].

Lemma 4 Assume $P \in M^P_1(\overline{\Omega}_\varepsilon)$ is locally equivalent to the product of its marginals. Then, the function $(\lambda, \nu) \mapsto \int f(\lambda, \overline{\omega}) \nu(d\overline{\omega})$ is continuous on $(-\infty, \lambda_{\text{crit}}(P)) \times M^P_1$.

Proof: Let
\[
\xi_m(\lambda, \nu) := \int \log \varphi_m(\lambda, \overline{\omega}) \nu(d\overline{\omega})
\]
for the bounded, continuous $\varphi_m(\lambda, \cdot)$ of (2.22). Note that $|\varphi_m(\lambda', \overline{\omega}) - \varphi_m(\lambda, \overline{\omega})| \rightarrow 0$ as $\lambda' \rightarrow \lambda$, uniformly in $\overline{\omega}$. Considering hereafter $m \geq \varepsilon^{-1} + 2$, we have that $\varphi_m(\lambda, \overline{\omega}) \geq 1/c_\lambda$ (for $c_\lambda$ of (2.8)).
The function $\xi_m(\lambda, \nu)$ is then continuous on $\mathbb{R} \times M_1(\overline{\Omega}_e)$. By the inequality $\log x \leq x - 1$ and the preceding lower bound on $\varphi_m$ we have that

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

Fixing $\lambda < \lambda_{\text{crit}}(P)$ and $\omega \in (\text{supp } P\big|_1)^2$, we thus deduce from (2.20) that

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

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

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

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

We next provide for $I_P^{\alpha}(\cdot)$ representations analogous to those of Lemma 2.

**Lemma 5** Assuming (C1)–(C3), let

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

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

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

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

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

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

In particular, $I_P^{\alpha}(\cdot)$ is a convex rate function and is non-increasing when $\lambda_{\text{crit}}(P) = 0.$

**Proof:** Since $\lambda \mapsto \int f(\lambda, \omega) \eta(d\omega)$ is convex, non-decreasing for any $\eta \in M_1^P$, so is $\lambda \mapsto L(\lambda)$. Note that $L(\lambda) \geq \int f(\lambda, \omega) P(d\omega) = \infty$ for any $\lambda > \lambda_{\text{crit}}(P)$ (see Lemma 1). In contrast, $\int f(\lambda, \omega) \eta(d\omega) \leq -2 \log \epsilon$ for all $\lambda \leq \lambda_{\text{crit}}(P)$ and $\eta \in M_1^P$ (c.f. (2.20)), implying that $L(\lambda)$ is finite and bounded for such $\lambda$. Moreover, $\lambda \mapsto \int f(\lambda, \omega) \eta(d\omega)$ is continuous on $(-\infty, \lambda_{\text{crit}}(P)]$ for any $\eta \in M_1^P$ (by Lemma
4 and monotone convergence), implying that $L(\cdot)$ is lower semi-continuous 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 semi-continuous (and when $\lambda_{\text{crit}}(P) = 0$, also non-increasing). Obviously, $J(u) = \infty$ for $u < 0$. We prove below that $I_P^\alpha(\cdot) = J(\cdot)$. This is all we need when $\lambda_{\text{crit}}(P) = 0$, whereas if $\lambda_{\text{crit}}(P) > 0$ then $J(u) = \max(J_-(u), J_+(u))$ with $J_- (u) := \sup_{\lambda \leq 0} [\lambda u - L(\lambda)]$ non-increasing and $J_+ (u) := \sup_{0 \leq \lambda < \lambda_{\text{crit}}(P)} [\lambda u - L(\lambda)]$ non-decreasing. 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 $J_+(u) \geq -L(0)$ and $J_-(u) \geq -L(0)$ for all $u$. With $I_P^\alpha(\cdot) = J(\cdot)$, we easily get then (2.25) and (2.26) out of (2.24).

Since $\eta \mapsto G(\lambda, \eta, u) + h(\eta|P)$ is convex, lower semi-continuous 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 [11, Theorem 4.21]), we conclude that
\[ J(u) = \inf_{\eta \in M_1^P} \sup_{\lambda < \lambda_{\text{crit}}(P)} [G(\lambda, \eta, u) + h(\eta|P)] = \sup_{\lambda < \lambda_{\text{crit}}(P)} [G(\lambda, \bar{\eta}, u) + h(\bar{\eta}|P)] . \tag{2.27} \]
Here, $\bar{\eta}$ is a global minimizer of the lower semi-continuous 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 (2.27) that for any $u \in \mathbb{R}$,
\[ J(u) \leq \inf_{\eta \in M_1^P} \sup_{\lambda \in \mathbb{R}} [G(\lambda, \eta, u) + h(\eta|P)] = I_P^\alpha(u) . \]
To show the converse inequality, we assume without loss of generality that $J(u) < \infty$ and approximate the stationary $\bar{\eta}$ of (2.27) by “nice” ergodic measures. To this end, note that (C3) implies that $J(\lambda, \varpi) - \lambda(b + 1) \in [\log \varepsilon k(\delta)e^{4\delta}, 0]$ for all $\lambda \leq 0$ and $P$-a.e. $\varpi$, hence also
\[ \lambda(u - b - 1 - \delta) - \log(\varepsilon k(\delta)) \geq G(\lambda, \eta, u) \geq \lambda(u - b - 1) \tag{2.28} \]
for all $\eta \in M_1^P$ and $\lambda \leq 0$. In particular, since $J(u) < \infty$, by (2.27) and (2.28) we know that $u \geq (b + 1)$. Fixing $u = b + 1 + 2\delta$, $\delta > 0$, it follows from (2.28) that
\[ I_P^\alpha(u) = \sup_{\lambda \geq -K} G(\lambda, \eta, u) , \tag{2.29} \]
for $K = K_u = -\delta^{-1}[\log(\varepsilon k(\delta))] < \infty$ and all $\eta \in M_1^P$. Let $\bar{\eta}_\varepsilon^u = (1 - \frac{1}{\varepsilon}) \bar{\eta} + \frac{1}{\varepsilon} P \in M_1^P$, noting that $h(\bar{\eta}_\varepsilon^u|P) = (1 - \frac{1}{\varepsilon}) h(\bar{\eta}|P) < \infty$. By (C2), there exist $\eta_n^u \in M_1^P(\bar{\eta}_\varepsilon^u)$ that are locally equivalent to the product of their marginals, with $\eta_n^u \rightarrow \bar{\eta}_\varepsilon^u$ and $h(\eta_n^u|P) \rightarrow h(\bar{\eta}_\varepsilon^u|P)$ as $n \rightarrow \infty$. Since $P \ll \bar{\eta}_\varepsilon^u$, we see by (2.21) that $\lambda_{\text{crit}}(\eta_n^u) \rightarrow \lambda_{\text{crit}}(P)$ as $n \rightarrow \infty$. By a diagonalization argument, we thus have $\tilde{\eta}_\varepsilon^u \in M_1^P(\bar{\eta}_\varepsilon^u) \cap M_1^P$, with
\[ \tilde{\eta}_\varepsilon \rightarrow \bar{\eta}_\varepsilon \quad h(\tilde{\eta}_\varepsilon|P) \rightarrow h(\bar{\eta}|P) , \quad \lambda_{\text{crit}}(\tilde{\eta}_\varepsilon) \rightarrow \lambda_{\text{crit}}(P) . \]
In particular, for any $\xi > 0$ and $\ell$ large enough $G(\lambda, \tilde{\eta}_\ell, u) = -\infty$ when $\lambda > \lambda_{\text{crit}}(P) + \xi \geq \lambda_{\text{crit}}(\tilde{\eta}_\ell)$, implying together with (2.29) that

$$I_P^\beta(u) \leq h(\tilde{\eta}_\ell|P) + \sup_{-K \leq \lambda \leq \lambda_{\text{crit}}(P) + \xi} G(\lambda, \tilde{\eta}_\ell, u)$$

$$\leq h(\tilde{\eta}_\ell|P) + 2\xi u + \sup_{-K \leq \lambda \leq \lambda_{\text{crit}}(P) - \xi} G(\lambda, \tilde{\eta}_\ell, u)$$

$$\leq h(\tilde{\eta}_\ell|P) + 3\xi u + G(\tilde{\lambda}_\ell, \tilde{\eta}_\ell, u),$$

for some $\tilde{\lambda}_\ell \in [-K, \lambda_{\text{crit}}(P) - \xi]$. Passing to a subsequence if needed, $\tilde{\lambda}_\ell \to \lambda \in [-K, \lambda_{\text{crit}}(P) - \xi]$. Considering $\ell \to \infty$ we deduce by applying Lemma 4 for $(\tilde{\lambda}_\ell, \tilde{\eta}_\ell) \to (\lambda, \eta)$, that

$$I_P^\beta(u) \leq G(\lambda, \eta, u) + h(\eta|P) + 3\xi u \leq J(u) + 3\xi u$$

(the rightmost inequality follows from (2.27)). Since $\xi > 0$ and $u > b + 1$ are arbitrary, the proof of (2.24) is thus complete, except possibly at $u = b + 1$. Turning to deal with this remaining case, note that $P$-almost surely, $T_1 \geq b + 1$ by (C3). Hence, by monotone convergence for any $\eta \in M_1^P$,

$$I_\eta^\beta(b + 1) = -\inf_{\lambda \in \mathbb{R}} \int \log E_\nu \left( e^{\lambda \left( T_1 - b - 1 \right)} \mathbf{1}_{T_1 < \infty} \right) \eta(d\nu) = -\int \log \left[ \omega_0 \mu_0(\{b\}) \right] \eta_1(d\nu) \ . \quad (2.30)$$

Since it suffices to consider $\lambda \to -\infty$ in (2.30), it follows from (2.27) that

$$J(b + 1) = h(\eta|P) - \int \log \left[ \omega_0 \mu_0(\{b\}) \right] \eta_1 (d\nu) \quad \text{(2.31)}$$

(where both sides have value $+\infty$ if $\eta(\mu_0(\{b\}) = 0) > 0$). Assuming without loss of generality that $J(b + 1) < \infty$ and in particular that $h(\eta|P) < \infty$, we have by (C2) a sequence $\eta^n \in M_1^P(\mathcal{K}_b)$ with $\eta^n_1 = \tilde{\eta}_1$ for all $n$ and $h(\eta^n|P) \to h(\eta|P)$. Noting that for all $n$ both $\eta^n \in M_1^P$ and

$$I_\eta^n(b + 1) = -\int \log \left[ \omega_0 \mu_0(\{b\}) \right] \eta_1 (d\nu) \ ,$$

by (2.30), we deduce from (2.31) that

$$I_P^\beta(b + 1) \leq \liminf_{n \to \infty} \left\{ I_\eta^n(b + 1) + h(\eta^n|P) \right\} = J(b + 1) .$$

This concludes the proof of (2.24) and with it that of the lemma. 

We begin the proof of the upper bound in Theorem 2 with the upper tail in case $\lambda_{\text{crit}}(P) > 0$. Integration of (2.9) yields that for all $\lambda < \lambda_{\text{crit}}(P)$,

$$\mathbb{E} \left( e^{\lambda T_n} \mathbf{1}_{T_n < \infty} \right) = E_P \left( \exp \left( n \int f(\lambda, \nu) R_n(d\nu) \right) \right) .$$

By (C1), $\{R_n\}$ satisfies a 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 $\nu \mapsto \int f(\lambda, \nu)(d\nu)$ is bounded and continuous (by Lemma 4), it follows from Varadhan’s lemma (see [4, Theorem 4.3.1]) that

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbb{E} \left( e^{\lambda T_n} \mathbf{1}_{T_n < \infty} \right) = \sup_{\eta \in M_1^P} \left( \int f(\lambda, \eta)(d\nu) - h(\eta|P) \right) = L(\lambda) . \quad (2.32)$$
Fix $u > 0$. Combining (2.32) and Chebychev’s inequality for each $\lambda_{\text{crit}}(P) > \lambda \geq 0$, we get the upper bound

$$
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P} \left( \infty > T_n \frac{1}{n} \geq u \right) \leq - \sup_{0 \leq \lambda < \lambda_{\text{crit}}(P)} [\lambda u - L(\lambda)] = - \inf_{w \geq u} I^\eta_P(w),
$$

(2.33)

where the equality follows from (2.26).

Applying the same argument with $\lambda < 0$ and using (2.25), yields that

$$
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P} \left( T_n \frac{1}{n} \leq u \right) \leq - \inf_{w \leq u} I^\eta_P(w).
$$

(2.34)

If $\lambda_{\text{crit}}(P) = 0$, then $I_P^\eta(\cdot)$ is non-increasing (see Lemma 5), hence (2.34) 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 (2.33), (2.34) and the convexity of $I_P^\eta(\cdot)$ (proved in Lemma 5).

Suffices to prove the lower bound in Theorem 2 for $(u - \delta, u + \delta)$ with $u \in [1, \infty)$ such that $I_P^\eta(u) < \infty$ and $\delta > 0$. Fixing such $u$ and $\delta$ there exists $\eta \in M^1(\Omega)$ such that $I_P^\eta(u) + h(\eta|P) \leq I_P^\eta(u) + \delta < \infty$. In particular, $u \geq u_-(\eta)$. Applying [1, Lemma 7] as in the proof of [1, Theorem 6] but here with the measures $Q_{\sigma,n} \otimes \eta(\text{d}w)$ if $u \in (u_-(\eta), u_+(\eta))$ and $Q_{\sigma,n,m} \otimes \eta(\text{d}w)$ if $u \geq u_+(\eta)$ (so we can use the strong law (2.16) for $\eta$-a.e. $\omega$), we obtain the bound

$$
\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P} \left( \left| \frac{T_n}{n} - u \right| \geq \delta \right) \geq - I_P^\eta(u) - h(\eta|P),
$$

for all $u > u_-(\eta)$. By continuity of the convex rate function $I_P^\eta(u)$ 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.

**Proof of Theorem 3, case (a).** We start by showing that $I_P^\eta(\cdot)$ of (1.2) is a convex, good rate function. Recall that $u_-(P) \geq 1$, hence $I_P^\eta(v) = \infty$ for all $v \in [-1, 1]$ (see (1.1) and Lemma 2). Moreover, with $\lambda_{\text{crit}} = \lambda_{\text{crit}}(P)$, by Lemma 1, (1.1) and the definition of $I_P^\eta(\cdot)$ we have that

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

(2.35)

for all $v \neq 0$. In particular, $I_P^\eta(\cdot)$ is convex and lower semi-continuous on $(0, \infty)$ and $(-\infty, 0)$, separately. The non-random bounds (2.1) and (2.8) on $\varphi(\lambda, \cdot)$ imply that $\lim_{v \downarrow 0} I_P^\eta(v)$ exists and equals to $\lambda_{\text{crit}}$. Further, $I_P^\eta(\cdot)$ is continuous at 0 by (2.35). It remains to show the convexity of $I_P^\eta(\cdot)$ at 0, namely that for all $v_1, v_2 > 0$,

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

By (2.35) (giving a lower bound for the sup by plugging in $\lambda = \lambda_{\text{crit}}$), this follows from the inequality

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

15
which by (2.18), is a consequence of the fact that
\[
0 \geq f^-(\lambda_{\text{crit}}, \omega) + f(\lambda_{\text{crit}}, \theta^{-1} \omega),
\]
for $P$-almost every $\omega \in \Omega_i$ (integrate (2.36) with respect to the stationary measure $P$). Indeed, by the Markov property
\[
E_\omega(e^{\lambda_{\text{crit}} T_M 1_{T_{i-1} < T_M < \infty}}) = E_\omega(e^{\lambda_{\text{crit}} T_1 1_{T_1 < \infty}}) E_\theta^{-1} \omega(e^{\lambda_{\text{crit}} T_1 1_{T_1 < \infty}}).
\]
Recall that $E_\omega(e^{\lambda_{\text{crit}} T_M 1_{T_M < \infty}}) < \infty$ for $P$-almost every $\omega$ and all $M < \infty$ (see (2.9)). Thus,
\[
1 \geq \frac{E_\omega(e^{\lambda_{\text{crit}} T_M 1_{T_{i-1} < T_M < \infty}})}{E_\omega(e^{\lambda_{\text{crit}} T_i 1_{T_M < \infty}})} = E_\omega(e^{\lambda_{\text{crit}} T_{i-1} 1_{T_{i-1} < T_M}}) E_\theta^{-1} \omega(e^{\lambda_{\text{crit}} T_1 1_{T_1 < \infty}}).
\]
Taking the logarithm and considering $M \to \infty$, one obtains (2.36).

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

Starting with the lower bounds, as $|X_t - X_s| \leq |t - s|$, for $v \neq 0$ and $1 > \delta > 0$,
\[
P_\omega \left( \frac{X_n}{n} \in (v - 2\delta, v + 2\delta) \right) \geq P_\omega \left( (1 - \delta)n < T_{[nv]} < (1 + \delta)n \right),
\]
and Theorem 1 implies that $P$-a.s. for all $v \neq 0$ and $\delta > 0$,
\[
\liminf_{n \to \infty} \frac{1}{n} \log P_\omega \left( \frac{X_n}{n} \in (v - 2\delta, v + 2\delta) \right) \geq \begin{cases} -v I_P^q \left( \frac{1}{v} \right), & v > 0, \\ v I_P^\tau - q \left( \frac{1}{|v|} \right), & v < 0. \end{cases}
\]
Similarly, taking $1 > \delta > u > 0$,
\[
P_\omega \left( \frac{X_n}{n} \in (-2\delta, 2\delta) \right) \geq P_\omega \left( (1 - \delta)n < T_{[nu]} < (1 + \delta)n \right),
\]
hence by Theorem 1,
\[
\liminf_{n \to \infty} \frac{1}{n} \log P_\omega \left( \frac{X_n}{n} \in (-2\delta, 2\delta) \right) \geq -u I_P^q \left( \frac{1}{u} \right), \quad P$-a.s.
\]
and considering rational $u \downarrow 0$ completes the proof of the LDP lower bound.

We next deal with the complementary upper bounds. Assuming without loss of generality that $E_P(\log \rho_0) \leq 0$, we have that $T_1 < \infty$ for $P$-almost every $\omega$ (recall that here $H_i(x) < \infty$ for all $i, x$), and $\nu P = 1/\pi(P) \geq 0$. Since $n^{-1} X_n \in [-1, 1]$, it suffices to show that $P$-a.s.
\[
\limsup_{n \to \infty} \frac{1}{n} \log P \left( \frac{X_n}{n} \in (v - \zeta, v + \zeta) \right) \leq -I_P^q(v), \quad \forall |v| \leq 1
\]
(c.f. [4, Theorem 4.1.11]). The next lemma, whose proof is deferred, is key to the proof of (2.37).
Lemma 6 Suppose $P(\mu_0(\{\infty\}) > 0 = 0$ and $E_P(\log \rho_0) \leq 0$.

Let $\Gamma = \sup P$ and $S_n = \inf \{ t \geq n : X_t \leq 0 \}$. Then,

$$\limsup_{n \to \infty} \frac{1}{n} \log \sup_{\omega \in \Gamma} P_\omega(S_n < \infty) \leq -\lambda_{\text{crit}}(P). \quad (2.38)$$

We now prove (2.37) for $v = 0$. For any $\omega$ and $\Delta > 0$,

$$\varepsilon^{\Delta n} P_{\omega}(\inf_{t \geq n} X_t \leq \Delta n) \leq P_{\omega}(\inf_{t \geq n} X_t \leq 0) = P_{\omega}(S_n < \infty).$$

In particular, since $I^P_v(0) = \lambda_{\text{crit}}(P)$ and

$$P_{\omega}(X_n \in (-\zeta n, \zeta n)) \leq P_{\omega}(\inf_{t \geq n} X_t \leq \zeta n) \leq \varepsilon^{\zeta n} P_{\omega}(S_n < \infty), \quad (2.39)$$

(2.38) implies that (2.37) holds for $v = 0$. Considering next $v \neq 0$ and $\zeta \in (0, |v|)$ such that

$$u = v - \zeta \text{sign } v \text{ is rational, note that for any } \delta \in (0, 1) \text{ such that } 1/(\delta u) \text{ is integer},$$

$$P_{\omega}\left(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\right) \leq \varepsilon^{-2\zeta n} \sum_{k=1}^{(|u|/\delta)^{-1}} P_{\omega}\left(\frac{T_{\lfloor |u|/\delta \rfloor}}{n/|u|} \in [(k - 1)\delta, k\delta]\right) P_{\theta^{|u|}v}\omega(S_{[n - n\delta|u|]} < \infty). \quad (2.40)$$

With $\xi = k\delta|u| \leq 1$, it follows from Theorem 1 and Lemma 2 that P-a.s.,

$$\limsup_{n \to \infty} \frac{1}{n} \log P_{\omega}\left(\frac{T_{\lfloor |u|/\delta \rfloor}}{n/|u|} \in [(k - 1)\delta, k\delta]\right) \leq -|u|I^P_{\theta^{|u|}v}(\xi/|u|) + |u|w(|u|, \delta), \quad (2.41)$$

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

$$w(r, \delta) := \max\{|I^P_{\theta^{|u|}v}(s) - I^P_{\theta^{|u|}v}(t)| + |I^P_{\theta^{|u|}v}(s) - I^P_{\theta^{|u|}v}(t)|; s, t \in [\pi(P), 1/r], |s - t| \leq \delta\}.$$

Let $\Gamma' = \{\omega : \theta^k\omega \in \Gamma \forall k \in \mathbb{Z}\}$, noting that $P(\Gamma') = 1$ by stationarity (in fact $\Gamma' = \Gamma$), whereas by (2.38),

$$\limsup_{n \to \infty} \frac{1}{n} \log \sup_{\omega \in \Gamma'} P_{\theta^{|u|}v}\omega(S_{[n - n\xi]} \in \infty) \leq -(1 - \xi)\lambda_{\text{crit}}(P) \quad (2.42)$$

Putting (2.41) and (2.42) in (2.40), and using the relation (1.2) we deduce that P-a.s.

$$\limsup_{n \to \infty} \frac{1}{n} \log P_{\omega}\left(\frac{X_n}{n} \in (v - \zeta, v + \zeta)\right) \leq - \inf_{\xi \in [0, 1]} \{\xi I^P_{\theta^{|u|}v}(u/\xi) + (1 - \xi)I^P_{\theta^{|u|}v}(0)\} + |u|w(|u|, \delta) - 2\zeta \log \xi$$

As the finite, convex, rate function $I^P_{\theta^{|u|}v}(\cdot)$ is continuous on $(u_{\infty}, \infty)$, the oscillation $w(r, \delta) \to 0$ for $\delta \downarrow 0$ and any fixed $r < \infty$. With $I^P_{\theta^{|u|}v}(\cdot)$ convex and lower semi-continuous, taking $\delta \downarrow 0$ then $\zeta \downarrow 0$ we obtain the bound of (2.37) and complete the proof of the theorem in case $P(\mu_0(\{\infty\}) > 0) = 0$.
Proof of Theorem 3, case (b). For $I_P^q(v)$ of (1.3), $v \neq 0$ we have by the same reasoning that led to (2.35), the analogous representation,

$$I_P^q(v) = v 1_{v<0} E_P(\log \rho_0) + \inf_{\ell \in [0,1]} \sup_{\lambda \leq \lambda_{\text{crit}}} \{ \lambda \ell - |v| E_P(f(\lambda, \omega)) \}$$

$$= v 1_{v<0} E_P(\log \rho_0) + \sup_{\lambda \leq 0} \{ \lambda - |v| E_P(f(\lambda, \omega)) \},$$

(2.43)

where the second equality follows by an application of the min-max theorem [11, Theorem 4.2'] for the function $(\ell, \lambda) \mapsto \lambda \ell - |v| E_P(f(\lambda, \omega)) (\ell \in [0,1], \lambda \in (-\infty, \lambda_{\text{crit}}])$, which is convex in $\ell$ and concave and continuous in $\lambda$ (the continuity of $\lambda \mapsto E_P(f(\lambda, \omega))$ follows from (2.1), (2.8) and dominated convergence). Here too $I_P^q(v) = \infty$ for all $v \notin [-1,1]$, whereas by (2.43), $I_P^q(\cdot)$ is convex and lower semi-continuous on $(0, \infty)$ and $(-\infty, 0)$, separately. Combining the bound (2.8) with the representation (2.43) we see that $\lim_{v \to 0} I_P^q(v) = 0$, that is, $I_P^q(\cdot)$ is continuous at 0. Since $I_P^q(\cdot) \geq 0$, 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 $\ell \in [0,1]$,

$$P^v_{\mathcal{F}} \left( \frac{X_n}{n} \in (v - 2\delta, v + 2\delta) \right) \geq P^v_{\mathcal{F}} \left( T_{[nv]} \in (\ell n - \delta n, \ell n + \delta n) \right) P_{\theta^{[nv]}(\mathcal{F})} \left( |X_{(1-\ell)n}| < \delta n \right)$$

whereas

$$P_{\theta^{[nv]}(\mathcal{F})} \left( |X_{(1-\ell)n}| < \delta n \right) \geq \epsilon^{\delta n} \max_{\{j:|j-nv|<\delta n\}} \mu_j(\{\infty\}) := \epsilon^{\delta n} \xi_n(\omega).$$

We thus get the LDP lower bound for the rate function (1.3) out of that of Theorem 1 (including also the case of $v = 0$), provided $\xi_n(\omega) > \xi$ for all $n$ large enough. By Birkhoff’s pointwise ergodic theorem this holds for $P$-almost every $\omega$, as

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

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

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

To prove the complementary upper bounds, namely, (2.37), 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 $u = v - \zeta \text{sign} v$ we have that

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

(2.44)

Considering $n \to \infty$, it thus follows from Theorem 1 and (1.3) that

$$\limsup_{n \to \infty} \frac{1}{n} \log P^v_{\mathcal{F}} \left( \frac{X_n}{n} \in (v - \zeta, v + \zeta) \right) \leq -I_P^q(u),$$

which holds $P$-a.s. for all $v$ and $\zeta$ as above. With $I_P^q(\cdot)$ lower semi-continuous, taking $\zeta \downarrow 0$ completes the proof of (2.37) and hence, that of the theorem. ☐
Proof of Lemma 6. Recall that our assumptions imply that $v_P = 1/\pi(P) \geq 0$. The lemma is trivial for $v_P = 0$ as then $\lambda_{\text{crit}} = 0$. Assuming hereafter that $v_P > 0$, let $b_n(\omega) = P_{\omega}(S_n < \infty)$, $\alpha_n := \sup\{b_n(\omega) : \omega \in \Gamma\}$, and $\Lambda = \inf\{\ell \geq 1 : X_\ell = 0\}$. By the strong Markov property of the embedded RWRE, denoting by $P_{\omega}(\cdot)$ the law of the random walk started at $y$ in the environment $\omega$ (where we omit $y$ if $y = 0$), it holds that for all $k, \omega$ and all $y < 0$,

$$P_{\omega}(S_k < \infty) \geq P_{\omega}(Y \leq k)P_{\omega}(S_k < \infty) + P_{\omega}(Y > k) \geq P_{\omega}(S_k < \infty).$$

Since

$$P_{\omega}(S_{n+k} < \infty) \geq \varepsilon P_{\omega}(S_m < \infty)E_{\omega}\left(P_{\omega}^{X_{S_m-1}}(S_k < \infty) \mid S_m < \infty\right) \geq \varepsilon P_{\omega}(S_m < \infty)P_{\omega}(S_k < \infty),$$

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

We next show that $n^{-1} \log b_n(\omega) \to a$ as $n \to \infty$, for $P$-a.e. $\omega$. To this end, fix $\delta > 0$ and let $k < \infty$ large enough for $k^{-1} \log a_k \geq a - \delta$. There exists a $\omega \in \Gamma$ such that $k^{-1} \log b_k(\omega) \geq a - 2\delta$. Therefore, one may find a finite $\ell$ large enough such that $k^{-1} \log P_{\omega}(S_\ell < \ell) \geq a - 3\delta$. Let $Z = (Z_0, Z_1, \ldots, Z_\ell)$, $z = (z_0, z_1, \ldots, z_\ell)$ and $\Theta = (\Theta_1, \ldots, \Theta_\ell)$, and use the notation $P_{\omega}(A)$, or $P_{\mu}(A)$, for events $A$ which depend on the environment only via $\omega := (\omega_0, \ldots, \omega_\ell)$ or $\mu := (\mu_0, \ldots, \mu_\ell)$, respectively. Note that

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

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

$$P_{\omega}(S_\ell < \ell) = \sum_z P_{\omega}(Z = z)P_{\mu}(\Theta \in G_z | Z = z). \tag{2.45}$$

A finite number of $z$ vectors is considered in (2.45), for each of which $\omega \mapsto P_{\omega}(Z = z)$ is continuous on $\overline{\Omega}_\varepsilon$ while $\mu \mapsto $\mathcal{L}(\Theta | Z = z) : M_\ell(\mathbb{R}_+) \mapsto M_1(\mathbb{R}_+^\ell)$ are also continuous. By (2.45), we see that $\omega \mapsto P_{\omega}(S_\ell < \ell)$ is lower semi-continuous on $\overline{\Omega}_\varepsilon$. Consequently, there exists an open set $A \subset \overline{\Omega}_\varepsilon$ such that $P(A) > 0$ and

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

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

$$b_n(\omega) = P_{\omega}(S_n < \infty) \geq e^{g(\omega)}P_{\theta^{-g(\omega)}\omega}(S_n < \infty) \geq e^{g(\omega)}[e^{P_{\theta^{-g(\omega)}\omega}(S_k < \infty)}]^{\frac{1}{n}} \geq e^{g(\omega)}[e^{\frac{1}{n}k^{-1} \log \varepsilon - 4\delta}]^{\frac{1}{k}} \geq e^{g(\omega)}[e^{(a - 4\delta)}]^{\frac{1}{k}}, \tag{2.46}$$

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

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

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

$$a = \lim_{n \to \infty} n^{-1} \log b_n(\omega), \quad P - a.s. \tag{2.47}$$
Fixing $1 > \delta > 0$, $u \in (0, v_P/(1 + v_P))$, let $\mathcal{S}$ denote the finite set of integer pairs $(k, \ell)$ such that $1 + 1/\delta \leq \min(k, \ell)$ and $(k + \ell - 2)/\delta u \leq 1$. We have by the strong Markov property that
\[
 b_n(\omega) \leq P_{\omega} \left( T_{[nu]} \geq n(1 - u) \right) + \sum_{(k, \ell) \in \mathcal{S}} P_{\omega} \left( T_{[nu]} / n u \in [(k - 1)\delta, k\delta] \right) \frac{\delta u}{nu} P_{\omega} \left( S_{[n-n(k+\ell)]} < \infty \right),
\]
(2.48)
where we use the convention $b_t(\omega) = P_{\omega}(S_t < \infty) = 1$ for $t \leq 0$. Observing that
\[
 E_{\theta n} \left( e^{\lambda T_{[m]} - m} 1_{T_{[m]} < \infty} \right) = \prod_{i=1}^{m} \varphi^{-}(\lambda, \theta^i \omega),
\]
we follow the derivation of (2.12) and (2.13) to deduce in analogy to (2.41) that, with $\gamma = \ell u \leq 2$,
\[
 \limsup_{n \to \infty} \frac{1}{n} \log P_{\omega} \left( T_{[nu]} / n u \in [(\ell - 1)\delta, \ell\delta] \right) \leq -\gamma I_P(u/\gamma) + uw(u/2, \delta) \quad P - a.s.
\]
(2.49)
By convexity of $I_P(\cdot)$, with $\xi = k\delta u$,
\[
 \xi I_P(u/\xi) + \gamma I_P(u/\gamma) \geq (\xi + \gamma) I_P(0).
\]
So, with $(k + \ell)\delta \geq 2$, by (2.41), (2.48) and (2.49), for $P$-almost every $\omega$ and all $n > n_0(\omega)$,
\[
 b_n(\omega) \leq e^{n(2w(u/2, \delta) + \delta)} \left[ e^{-n(1-u)I_P(u/1-u)} + \sum_{(k, \ell) \in \mathcal{S}} e^{-n(k+\ell)\delta u} I_P(0) b_{[n-n(k+\ell)]}(\omega) \right]
\]
\[
 \leq C \max_{2wu \leq j \leq n} \left\{ e^{-Jj} b_{n-j}(\omega) \right\},
\]
(2.50)
where $C = C(\delta, u) < \infty$ and
\[
 J = \min \left\{ \left( 1 - u \right) I_P \left( \frac{u}{1 - u} \right), I_P(0) \right\} - \frac{1}{2u} \left( 2uw(u/2, \delta) + \delta \right).
\]
It is easy to check that for $u > 0$, $\gamma_n \geq 0$, $C < \infty$,
\[
 \gamma_n \leq C \max_{2wu \leq j \leq n} \left\{ e^{-Jj} \gamma_{n-j} \right\} \quad \forall n \geq n_0 \quad \implies \limsup_{n \to \infty} \frac{1}{n} \log \gamma_n \leq -J.
\]
Consequently, from (2.50) we have that
\[
 \limsup_{n \to \infty} \frac{1}{n} \log b_n(\omega) \leq -J, \quad P - a.s.
\]
(2.51)
Since $J \to I_P(0)$ when taking first $\delta \downarrow 0$ then $u \downarrow 0$, it follows from (2.47) and (2.51) that $a \leq -I_P(0)$ as stated.
Proof of Theorem 4, case (a). With \( u_\eta(\eta) \geq 1 \) for any \( \eta \in M_1^P(\overline{\Omega}_e) \) we have that \( I_{\eta}^\tau q(u) = I_{\eta}^{\tau,q}(u) = \infty \) for all \( u < 1 \) (see Lemma 2 and (1.1)), hence \( \phi_p(\nu) = \infty \) for \( \nu \neq [-1, 1] \). Since \( I_{\nu}^{\tau,a}(\cdot) \) and \( I_{\nu}^{\tau,a}(\cdot) \) are rate functions, \( \phi_p(\cdot) \) of (1.4) 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 \( \lambda(\lambda) \leq -2\log \varepsilon \) for \( \lambda(\cdot) \) of (2.23) and all \( \lambda \leq \lambda_{\text{crit}} \) (see proof of Lemma 5). Hence, \( I_{\nu}^{\tau,a}(u) \geq \lambda_{\text{crit}} u + 2 \log \varepsilon \) by (2.24), implying that

\[
\liminf_{u \downarrow 0} u I_{\nu}^{\tau,a} \left( \frac{1}{u} \right) \geq \lambda_{\text{crit}}.
\]

With the same argument applying for \( I_{\nu}^{\tau,a}(\cdot) \), we get that

\[
\liminf_{v \to 0} I_{\nu}^p(v) \geq \lambda_{\text{crit}}.
\]

By definition, \( I_{\nu}^p(v) \leq I_{\nu}^p(v) \) for \( v \neq 0 \). As \( I_{\nu}^p(v) \to \lambda_{\text{crit}} \) for \( v \to 0 \) (see proof of Theorem 3), we conclude that \( I_{\nu}^p(v) \to \lambda_{\text{crit}} = I_{\nu}^p(0) \) when \( v \to 0 \), completing the proof that \( \phi_p(\cdot) \) is a good rate function. Since \( I_{\nu}^{\tau,a}(\cdot) \) and \( I_{\nu}^{\tau,a}(\cdot) \) are convex, it follows immediately that \( \phi_p(\cdot) \) of (1.4) is convex separately on \((0, \infty)\) and on \((-\infty, 0)\). The convexity of this function at 0 amounts to the inequality

\[
v_1 I_{\nu}^p(-v_2) + v_2 I_{\nu}^p(v_1) \geq (v_1 + v_2) \lambda_{\text{crit}}, \tag{2.52}
\]

which we prove next. As \( P \in M_1^P(\overline{\Omega}_e) \) is locally equivalent to the product of its marginals, the bound (2.20) results with (2.36) holding for all \( \omega \in (\text{supp} P_1)^2 \). Note that \( f^-(\lambda, \omega) \) depends only on \( \{\omega, x \geq 0\} \) while \( f(\lambda, \theta^{-1} \omega) \) depends only on \( \{\omega, x \leq -1\} \), so integrating (2.36) with respect to \( \eta|_{\{\ldots, \omega, \omega, \ldots\}} \otimes \eta|_{\{\omega_0, \omega_1, \ldots\}} \) yields that

\[
0 \geq E_{\eta'}(f^-(\lambda_{\text{crit}}, \omega)) + E_{\eta}(f(\lambda_{\text{crit}}, \omega)),
\]

for any stationary \( \eta, \eta' \in M_1^P \). For all such \( \eta, \eta' \) and \( v_1, v_2 > 0 \) then,

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

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

\[
I_{\nu}^{\tau,a} \left( \frac{1}{v_2} \right) + I_{\nu}^{\tau,a} \left( \frac{1}{v_1} \right) \geq \left( \frac{1}{v_2} + \frac{1}{v_1} \right) \lambda_{\text{crit}},
\]

resulting by (1.4) with (2.52).

The annealed LDP lower bounds in 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 \( v, \)

\[
\lim_{n \to \infty} \limsup_{\zeta \downarrow 0} \frac{1}{n} \log \mathbb{P} \left( \frac{X_n}{n} \in (v - \zeta, v + \zeta) \right) \leq -I_{\nu}^p(v). \tag{2.53}
\]
Assume without loss of generality that \( E_P(\log \rho_0) \leq 0 \), in which case Lemma 6 applies. Starting with \( v = 0 \), we have by (2.39) that
\[
\mathbb{P}(X_n \in (-\zeta n, \zeta n)) \leq e^{-\zeta n} \sup_{\mathbf{\varpi} \in \Gamma} P_{\mathbf{\varpi}}(S_n < \infty),
\]
and since \( I^\theta_P(0) = \lambda_{\text{crit}}(P) \), we have (2.53) by an application of Lemma 6. Recall that \( P(\Gamma^t) = 1 \) for \( \Gamma^t = \{ \mathbf{\varpi} : \theta^k \mathbf{\varpi} \in \Gamma \ \forall k \in \mathbb{Z} \} \). Hence, by (2.40) for any \( v \neq 0 \), \( \zeta \in (0, |v|) \) and \( u = v - \zeta \cdot \text{sign} \ v \),
\[
\mathbb{P}\left( X_n \cdot \frac{v - \zeta \cdot \text{sign} \ v}{n} \right) \leq e^{-2\zeta n} \sum_{k=1}^{\left( \frac{|u|}{\zeta} \right)} \mathbb{P}\left( \frac{T_{n|u|}}{n|u|} \in [(k - 1)\delta, k\delta] \right) \sup_{\mathbf{\varpi} \in \Gamma^t} P_{\mathbf{\varpi}}(S_{n - nk\delta|u|} < \infty).
\]
Thus, combining (2.42), Theorem 2 and the relation (1.4) we have that
\[
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left( X_n \cdot \frac{v - \zeta \cdot \text{sign} \ v}{n} \right) \leq -2\zeta \log \varepsilon - \inf_{\xi \in [0,1]} \{ \xi I^\theta_P(u/\xi) + (1 - \xi - \delta|u|)\lambda_{\text{crit}} \}.
\]
Since \( I^\theta_P(\cdot) \) is convex and lower semi-continuous, with \( \lambda_{\text{crit}} = I^\theta_P(0) \), taking \( \delta \downarrow 0 \) followed by \( \zeta \downarrow 0 \) we arrive at the bound (2.53).

**Proof of Theorem 4, case (b).** Considering \( I^\theta_P(\cdot) \) of (1.6), note that \( I^\theta_P(\cdot) = v I^\theta_P(u^* \wedge 1/v) \) for any \( v > 0 \), where \( u^* \geq 1 \) is a global minimizer of \( I^\theta_P(\cdot) \), setting \( u^* = \infty \) in case \( I^\theta_P(\cdot) \) is non-increasing. Since \( I^\theta_P(\cdot) \) is a convex rate function, the lower semi-continuity and convexity of \( I^\theta_P(\cdot) \) on \((0,\infty)\) are easily verified. Applying the same reasoning to the convex rate function \( I^\theta_P^{-\cdot}(\cdot) \) we get the convexity and lower semi-continuity of \( I^\theta_P(\cdot) \) of (1.6) at \((-\infty,0)\). Recall that this non-negative function is bounded above by \( I^\theta_P(\cdot) \) of (1.3) which converges to 0 as \( v \to 0 \). The function \( I^\theta_P(\cdot) \) of (1.6) 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 (1.5),
\[
\mathbb{P}\left( X_n \cdot \frac{v - 2\delta, v + 2\delta}{n} \right) \geq e^{2\delta n} \mathbb{P}\left( T_{n|v|} \in ((\ell - \delta)n, (\ell + \delta)n), \max_{\delta_n \leq j \leq 2\delta_n} H_1([nv] + j) = \infty \right) \geq e^{2\delta n} E_P\left( T_{n|v|} \in ((\ell - \delta)n, (\ell + \delta)n) \right) \cdot \max_{\delta_n \leq j \leq 2\delta_n} \mu_{[nv] + j(\{\infty\}) | \mathcal{F}_{[nv]}^-}
\geq e^{-c\delta n} e^{2\delta n} \mathbb{P}( T_{n|v|} \in ((\ell - \delta)n, (\ell + \delta)n) ) .
\]
Consequently, for any \( v > 0 \) and all \( \ell \in (0,1) \), by Theorem 2,
\[
\lim_{\delta \downarrow 0} \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left( X_n \cdot \frac{v - 2\delta, v + 2\delta}{n} \right) \geq -v I^\theta_P^{-\cdot}\left( \frac{\ell}{v} \right) .
\]
Optimizing over \( \ell \in [0,1] \) we arrive at the stated LDP lower bound for \( v > 0 \). The same argument applies for \( T_{-|v|} \), 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 (2.53) for \( v \neq 0 \), where by (2.44) we have that with \( u = v - \zeta \text{sign} \, v \),

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

Considering \( n \to \infty \), by Theorem 2 and the relation (1.6) we have that

\[
\lim_{n \to \infty} \sup_{n} \frac{1}{n} \log \mathbb{P}\left( \frac{X_n}{n} \in (v - \zeta, v + \zeta) \right) \leq -I^0_P(u),
\]

and with \( I^0_P(\cdot) \), lower semi-continuous, taking \( \zeta \downarrow 0 \) completes the proof of (2.53) and hence, that of the theorem. \( \square \)

3 Negative speed for random walks on Galton-Watson trees

Let \( Z \) be a random variable taking values on \( \{1,2,\ldots\} \) 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 non-rooted trees containing a special ray \(-\infty \leftrightarrow 0 \leftrightarrow \infty \) is then constructed as follows. Starting with \( Z \), we connect neighboring integers by an edge, and attach to each point \( x \in \mathbb{Z} \) an independent MGW-tree \( T_x \). We write the resulting infinite, unrooted tree as \( T = \cup_{x \in \mathbb{Z}} T_x \), where the roots of \( T_x \) and \( T_{x+1} \) are connected by an edge. The parent \( v^* \) of a vertex \( v \in T \cap T_x \) is defined as the parent of \( v \) in \( T_x \) if \( v \) is not the root of \( T_x \), and as \( x - 1 \) if \( v = x \in \mathbb{Z} \), i.e. if \( v \) is the root of \( T_x \). An alternative construction of 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 \( \tilde{D}_{m-n} \) and then taking weak limits, resulting with a measure on infinite trees with a special ray \(-\infty \leftrightarrow \infty \) marked (see [8] for details). Fixing \( 0 < \lambda < \infty \) and a tree \( \omega \) chosen according to AGW, the \( \lambda \)-biased random walk \{ \( S_n \) \} on \( \omega \) is the Markov chain such that if \( j^* \) is the parent of a vertex \( j \) having \( k \) children \( j_1,\ldots,j_k \), then

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

\[
P^v_{\lambda,\omega} \left[ S_{n+1} = j_i | S_n = j \right] = \frac{1}{\lambda + k}, \quad i = 1,2,\ldots,k,
\]

where \( v \in \omega \) is a fixed starting point (see [9]). We denote by \( P^v_{\lambda,\omega} \) the “quenched” distribution of the walk \{ \( S_n \) \} conditioned on the tree \( \omega \) and by \( P^v_\lambda := \int P^v_{\lambda,\omega} \text{AGW}(d\omega) \) the corresponding “annealed” measure. We write \( P_{\lambda,\omega} \) for \( P^0_{\lambda,\omega} \) and \( P_\lambda \) for \( P^0_\lambda \).

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 \( \mu_x \) be the distribution of \( H(x) \) under \( P^v_\lambda \). Let \( \omega_x := 1/(\lambda + 1) \). Note
that \( \omega_x \) 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 \( \tau = \{ (\omega_x, \mu_x) \} \). Indeed, the distribution \( P \) of \( \tau \) (under the measure \( \text{AGW} \) on trees with a special ray) is a (stationary) product measure where if \( Z \) is bounded, then also \( P_x \in \mathcal{M}_1(S) \) for some \( \varepsilon > 0 \), for which (C3) applies with \( b = 0 \). Let \( P_\tau \) be the distribution of \( \{ X_n \} \) under \( P_\lambda, \omega \) and \( \mathbb{P} \) the distribution of \( \{ X_n \} \) under \( P_\lambda \). Then, we are in the RWREH model. Since \( P \) is a product measure, (C1)-(C2) are clearly satisfied. Hence we can apply our previous results. In particular, we have by Lemma 1 and (2.18) a deterministic \( \lambda_{\text{crit}} \in [0, \infty) \) such that \( E_{\lambda, \omega}[e^{tT_{-1}}1_{T_{-1} < \infty}] \) is finite if and only if \( t \leq \lambda_{\text{crit}} \), for AGW-a.e. \( \omega \). Moreover, by Theorems 1 and 2 we have the weak LDP for \( n^{-1}T_{-n} \) (and \( n^{-1}T_n \)) under \( P_\tau \) and \( \mathbb{P} \), with quenched and annealed rate functions \( I_{\tau}^{\tau, q} \) and \( I_{\tau}^{\tau, a} \) respectively. By Theorems 3 and 4 we also have the LDP for \( n^{-1}X_n \) under the measures \( P_\tau \) and \( \mathbb{P} \), with good rate functions \( I_{\tau}^{\tau, q} \) and \( I_{\tau}^{\tau, a} \), respectively (where \( P(\mu_x(\{ \infty \}) > 0) = 0 \) if and only if \( \lambda \geq m \), c.f. [8]). Moreover, we have seen in (2.10) that for AGW-a.e. \( \omega \) and all \( t < \lambda_{\text{crit}} \),

\[
\lim_{n \to \infty} \frac{1}{n} \log E_{\lambda, \omega}(e^{tT_{-n}}1_{T_{-n} < \infty}) = -G^-(t, P, 0),
\]

whereas we have seen in (2.32) that

\[
\lim_{n \to \infty} \frac{1}{n} \log E_{\lambda}(e^{tT_{-n}}1_{T_{-n} < \infty}) = -\inf_{\eta \in \mathcal{M}_1(\mathbb{P})} \left[ G^-(t, \eta, 0) + h(\eta|P) \right].
\]

In particular, by Lemma 1, Lemma 4, and Varadhan’s lemma, if \( I_{\tau}^{\tau, q} = I_{\tau}^{\tau, a} \) then the limits in (3.1) and (3.2) must be equal for all \( t < \lambda_{\text{crit}} \).

Let \( |S_n| \) denote the distance of \( S_n \) from 0 in the tree \( \omega \). 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 \( I_{\tau}^{\tau, q} \) and \( I_{\tau}^{\tau, a} \) of the LDP of \( n^{-1}T_{-n} \).

**Proposition 1** If \( Z \) is non-degenerate, then for \( t < \lambda_{\text{crit}} \) there exists an \( \eta \in \mathcal{M}_1(\mathbb{P}) \) such that

\[
-G^-(t, P, 0) = E_{\tau} \left( \log E_{\tau}(e^{tT_{-1}}1_{T_{-1} < \infty}) \right) < E_{\eta} \left( \log E_{\tau}(e^{tT_{-1}}1_{T_{-1} < \infty}) \right) - h(\eta|P),
\]

except when \( t = 0 \) and \( \mathbb{P}(T_{-1} < \infty) = 1 \). That is, the limits in (3.1) and (3.2) are different, and consequently \( I_{\tau}^{\tau, q} \neq I_{\tau}^{\tau, a} \).

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

\[
V_t(x) := \log E_{\lambda, \omega}(e^{t(H(x)+1)}1_{H(x) < \infty}) = t + \log \sum_{h=0}^{\infty} e^{th} \mu_x(\{h\}).
\]

24
Note that
\[
g(n) := \log E_{\Lambda} \left( e^{T_n} 1_{T_n < \infty} \right) = \log E_{SRW} \left( E_P \left( \exp \left( \sum_{k=0}^{\sigma_n - 1} V_i(Z_k) \right) \right) \right),
\]
where \( E_{SRW}(\cdot) \) denotes integration over all paths of the biased simple random walk \( \{Z_k\} \). Since \( V_i(x), x \in Z \) are i.i.d. random variables, they are positively correlated. This allows us to apply the FKG inequality for the increasing functions \( \exp(\sum_{k=0}^{\sigma_n - 1} V_i(Z_k)) \) and \( \exp(\sum_{k=\sigma_m}^{\sigma_n - 1} V_i(Z_k)) \), for each fixed path \( \{Z_0, Z_1, \ldots, Z_{\sigma_n + m - 1}\} \), yielding that
\[
g(n + m) \geq \log E_{SRW} \left( E_P \left( \exp \left( \sum_{k=0}^{\sigma_m - 1} V_i(Z_k) \right) \right) \right) E_P \left( \exp \left( \sum_{k=\sigma_m}^{n + m - 1} V_i(Z_k) \right) \right) \).
\]
Applying the strong Markov property of \( Z_k \) at the stopping time \( \sigma_m \) where \( Z_{\sigma_m} = -m \), it follows by the translation invariance of both the law of \( \rho \mapsto \{Z + \rho - Z_\theta\} \) and that of \( \{V_i(\cdot)\} \), that
\[
g(n + m) \geq \log E_{SRW} \left( E_P \left( \exp \left( \sum_{k=0}^{\sigma_m - 1} V_i(Z_k) \right) \right) \right) + \log E_{SRW} \left( E_P \left( \exp \left( \sum_{k=0}^{\sigma_n - 1} V_i(Z_k) \right) \right) \right).
\]
Using the superadditivity of \( g \) and Jensen’s inequality (for \( \log x \)), it follows that
\[
\lim_{n \to \infty} \inf \, n^{-1} g(n) \geq g(1) = \log E_{\Lambda} \left( e^{T_{-1}} 1_{T_{-1} < \infty} \right) \geq \int \log E_{\Lambda, \omega} \left( e^{T_{-1}} 1_{T_{-1} < \infty} \right) AGW(d\omega)
\]
and the last inequality is strict as soon as \( \varphi^-(t, \omega) := E_{\Lambda, \omega} \left( e^{T_{-1}} 1_{T_{-1} < \infty} \right) \) is a non-degenerate random variable. Note that the limits in (3.1) and (3.2) correspond to the right and left sides of (3.3), respectively. Thus, it suffices to show that for \( Z \) non-degenerate, if \( \varphi^-(t, \omega) = c(t) \) for AGW-a.e. \( \omega \) for some (finite) constant \( c(t) > 0 \), then necessarily \( t = 0 \) and \( c(t) = 1 \) (hence, \( P_{\Lambda, \omega}(T_{-1} < \infty) = 1 \) for AGW-a.e. \( \omega \)).

Turning to this task, note that we may add the ray \( 0 \leftrightarrow \infty \) to the MGW tree \( T_0 \), thus making it a GW tree. With this identification, let \( k_0 \geq 1 \) be the number of children of \( 0 \) and \( N_0 := \sum_{k=1}^{T_{-1}} 1_{S_k = 0} \) the number of visits of 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 \( \omega_{r_i} \) 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 \( i \)-th visits of \( 0 \). Note that \( P_{\Lambda, \omega}(N_0 = \ell) = (k_0/(k_0 + \lambda))^{\ell} \Lambda/(\lambda + k_0) \) and the GW trees \( \omega_{r_i} \) 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 \( \omega \). Consequently, denoting by \( E_{k_0} \) expectation conditional on \( k_0 \),
\[
\varphi^-(t, \omega) = e^t E_{k_0} \left( e^{T_0} \prod_{i=1}^{N_0} E_{\Lambda, \omega_{r_i}} \left( e^{T_0^i} \prod_{i=1}^{N_0} \varphi^-(t, \omega_{r_i}) \right) \right) = e^t E_{k_0} \left( e^{T_0} \prod_{i=1}^{N_0} \varphi^-(t, \omega_{r_i}) \right).
\]
If \( \varphi^-(t, \omega) = c(t) \) for AGW-a.e. \( \omega \), then the same applies for the finite collection \( \varphi^-(t, \omega_{r_i}) \) for AGW-a.e. \( \omega \), implying that \( c(t) \) is a solution of the identity
\[
c(t)e^{-t} = E_{k_0} \left( (c(t)e^t)^{N_0} \right).
\]
It is easy to verify that if $Z$ is non-degenerate, so shall be the random variable $E_{k_0}(q^N_0)$, provided $q \neq 1$, $0 < q < \infty$. Thus, if $\varphi^-(t, \omega) = c(t)$ for AGW-a.e. $\omega$ and $Z$ is non-degenerate, necessarily $c(t)e^t = q = 1$, which by (3.4) is possible only in case $t = 0$ and $c(0) = 1$, as stated.

4 Discussion and open problems

1. We recall that in [7], the authors derived CLT and stable limit laws for transient RWRE’s. For recurrent RWRE’s, limit laws were derived by Sinai [10]. On the other hand, for the simple random walk with random holding times, [6] derived (process level) limit laws of the form of singular diffusion. It is natural to expect that the RWREH exhibits a rich spectrum of limiting behaviors, due to the competition between trapping mechanisms arising from large holding times and those arising from local drifts aligning in small neighborhood to create a trap, as in the standard RWRE.

2. Related to the above, we have not discussed in this paper aging phenomena for the RWREH model. In view of the results in [3, 6] one expects here to recover such phenomena.

3. As in [1], one may give a description of the shapes of the different annealed and quenched rate functions encountered in this paper. Since the techniques for doing that are similar to those in [1], we do not perform a detailed study here.

Acknowledgment We are grateful to Yuval Peres for helpful discussions and in particular for pointing out to us how useful the relation (2.46) really is. We also would like to thank Martin Zerner for the argument leading to (3.3).

References


