

## Collision local time of transient random walks and intermediate phases in interacting stochastic systems\*

Matthias Birkner<sup>1</sup>, Andreas Greven<sup>2</sup>, Frank den Hollander<sup>3 4</sup>

### Abstract

In a companion paper [6], a quenched large deviation principle (LDP) has been established for the empirical process of words obtained by cutting an i.i.d. sequence of letters into words according to a renewal process. We apply this LDP to prove that the radius of convergence of the generating function of the collision local time of two independent copies of a symmetric and strongly transient random walk on  $\mathbb{Z}^d$ ,  $d \geq 1$ , both starting from the origin, strictly increases when we condition on one of the random walks, both in discrete time and in continuous time. We conjecture that the same holds when the random walk is transient but not strongly transient. The presence of these gaps implies the existence of an *intermediate phase* for the long-time behaviour of a class of coupled branching processes, interacting diffusions, respectively, directed polymers in random environments .

**Key words:** Random walks, collision local time, annealed vs. quenched, large deviation principle, interacting stochastic systems, intermediate phase.

**AMS 2000 Subject Classification:** Primary 60G50, 60F10, 60K35, 82D60.

Submitted to EJP on December 24, 2008, resubmitted upon invitation to EJP on March 27, 2010, final version accepted January 11, 2011.

---

<sup>1</sup>Institut für Mathematik, Johannes-Gutenberg-Universität Mainz, Staudingerweg 9, 55099 Mainz, Germany.

Email: [birkner@mathematik.uni-mainz.de](mailto:birkner@mathematik.uni-mainz.de)

<sup>2</sup>Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstrasse 1 $\frac{1}{2}$ , 91054 Erlangen, Germany.

Email: [greven@mi.uni-erlangen.de](mailto:greven@mi.uni-erlangen.de)

<sup>3</sup>Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands.

Email: [denholla@math.leidenuniv.nl](mailto:denholla@math.leidenuniv.nl)

<sup>4</sup>EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

\*This work was supported in part by DFG and NWO through the Dutch-German Bilateral Research Group “Mathematics of Random Spatial Models from Physics and Biology”. MB and AG are grateful for hospitality at EURANDOM. We thank an anonymous referee for careful reading and helpful comments.

# 1 Introduction and main results

In this paper, we derive *variational representations* for the radius of convergence of the generating functions of the collision local time of two independent copies of a symmetric and transient random walk, both starting at the origin and running in discrete or in continuous time, when the average is taken w.r.t. one, respectively, two random walks. These variational representations are subsequently used to establish the existence of an intermediate phase for the long-time behaviour of a class of interacting stochastic systems.

## 1.1 Collision local time of random walks

### 1.1.1 Discrete time

Let  $S = (S_k)_{k=0}^\infty$  and  $S' = (S'_k)_{k=0}^\infty$  be two independent random walks on  $\mathbb{Z}^d$ ,  $d \geq 1$ , both starting at the origin, with an *irreducible, symmetric and transient* transition kernel  $p(\cdot, \cdot)$ . Write  $p^n$  for the  $n$ -th convolution power of  $p$ , and abbreviate  $p^n(x) := p^n(0, x)$ ,  $x \in \mathbb{Z}^d$ . Suppose that

$$\lim_{n \rightarrow \infty} \frac{\log p^{2n}(0)}{\log n} =: -\alpha, \quad \alpha \in [1, \infty). \quad (1.1)$$

Write  $\mathbb{P}$  to denote the joint law of  $S, S'$ . Let

$$V = V(S, S') := \sum_{k=1}^{\infty} 1_{\{S_k = S'_k\}} \quad (1.2)$$

be the *collision local time* of  $S, S'$ , which satisfies  $\mathbb{P}(V < \infty) = 1$  by transience, and define

$$z_1 := \sup \{z \geq 1: \mathbb{E} [z^V | S] < \infty \text{ S-a.s.}\}, \quad (1.3)$$

$$z_2 := \sup \{z \geq 1: \mathbb{E} [z^V] < \infty\}. \quad (1.4)$$

The lower indices indicate the number of random walks being averaged over. Note that, by the tail triviality of  $S$ , the range of  $z$ 's for which  $\mathbb{E}[z^V | S]$  converges is S-a.s. constant.<sup>1</sup>

Let  $E := \mathbb{Z}^d$ , let  $\tilde{E} = \cup_{n \in \mathbb{N}} E^n$  be the set of finite *words* drawn from  $E$ , and let  $\mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$  denote the shift-invariant probability measures on  $\tilde{E}^{\mathbb{N}}$ , the set of infinite *sentences* drawn from  $\tilde{E}$ . Define  $f: \tilde{E} \rightarrow [0, \infty)$  via

$$f((x_1, \dots, x_n)) = \frac{p^n(x_1 + \dots + x_n)}{p^{2\lfloor n/2 \rfloor}(0)} [2\bar{G}(0) - 1], \quad n \in \mathbb{N}, x_1, \dots, x_n \in E, \quad (1.5)$$

where  $\bar{G}(0) = \sum_{n=0}^{\infty} p^{2n}(0)$  is the Green function at the origin associated with  $p^2(\cdot, \cdot)$ , which is the transition matrix of  $S - S'$ , and  $p^{2\lfloor n/2 \rfloor}(0) > 0$  for all  $n \in \mathbb{N}$  by the symmetry of  $p(\cdot, \cdot)$ . The following variational representations hold for  $z_1$  and  $z_2$ .

<sup>1</sup>Note that  $\mathbb{P}(V = \infty) = 1$  for a symmetric and recurrent random walk, in which case trivially  $z_1 = z_2 = 1$ .

**Theorem 1.1.** Assume (1.1). Then  $z_1 = 1 + \exp[-r_1]$ ,  $z_2 = 1 + \exp[-r_2]$  with

$$r_1 \leq \sup_{Q \in \mathcal{D}^{\text{inv}}(\tilde{E}^{\mathbb{N}})} \left\{ \int_{\tilde{E}} (\pi_1 Q)(dy) \log f(y) - I^{\text{que}}(Q) \right\}, \quad (1.6)$$

$$r_2 = \sup_{Q \in \mathcal{D}^{\text{inv}}(\tilde{E}^{\mathbb{N}})} \left\{ \int_{\tilde{E}} (\pi_1 Q)(dy) \log f(y) - I^{\text{ann}}(Q) \right\}, \quad (1.7)$$

where  $\pi_1 Q$  is the projection of  $Q$  onto  $\tilde{E}$ , while  $I^{\text{que}}$  and  $I^{\text{ann}}$  are the rate functions in the quenched, respectively, annealed large deviation principle that is given in Theorem 2.2, respectively, 2.1 below with (see (2.4), (2.7) and (2.13–2.14))

$$E = \mathbb{Z}^d, \quad \nu(x) = p(x), \quad x \in E, \quad \rho(n) = p^{2\lfloor n/2 \rfloor}(0) / [2\bar{G}(0) - 1], \quad n \in \mathbb{N}. \quad (1.8)$$

Let

$$\mathcal{D}^{\text{erg,fin}}(\tilde{E}^{\mathbb{N}}) = \{Q \in \mathcal{D}^{\text{inv}}(\tilde{E}^{\mathbb{N}}) : Q \text{ is shift-ergodic, } m_Q < \infty\}, \quad (1.9)$$

where  $m_Q$  is the average word length under  $Q$ , i.e.,  $m_Q = \int_{\tilde{E}} (\pi_1 Q)(y) \tau(y)$  with  $\tau(y)$  the length of the word  $y$ . Theorem 1.1 can be improved under additional assumptions on the random walk, namely,<sup>2</sup>

$$\sum_{x \in \mathbb{Z}^d} \|x\|^\delta p(x) < \infty \quad \text{for some } \delta > 0, \quad (1.10)$$

$$\liminf_{n \rightarrow \infty} \frac{\log[p^n(S_n)/p^{2\lfloor n/2 \rfloor}(0)]}{\log n} \geq 0 \quad S - \text{a.s.}, \quad (1.11)$$

$$\inf_{n \in \mathbb{N}} \mathbb{E} [\log[p^n(S_n)/p^{2\lfloor n/2 \rfloor}(0)]] > -\infty. \quad (1.12)$$

**Theorem 1.2.** Assume (1.1) and (1.10–1.12). Then equality holds in (1.6), and

$$r_1 = \sup_{Q \in \mathcal{D}^{\text{erg,fin}}(\tilde{E}^{\mathbb{N}})} \left\{ \int_{\tilde{E}} (\pi_1 Q)(dy) \log f(y) - I^{\text{que}}(Q) \right\} \in \mathbb{R}, \quad (1.13)$$

$$r_2 = \sup_{Q \in \mathcal{D}^{\text{erg,fin}}(\tilde{E}^{\mathbb{N}})} \left\{ \int_{\tilde{E}} (\pi_1 Q)(dy) \log f(y) - I^{\text{ann}}(Q) \right\} \in \mathbb{R}. \quad (1.14)$$

In Section 6 we will exhibit classes of random walks for which (1.10–1.12) hold. We believe that (1.11–1.12) actually hold in great generality.

Because  $I^{\text{que}} \geq I^{\text{ann}}$ , we have  $r_1 \leq r_2$ , and hence  $z_2 \leq z_1$  (as is also obvious from the definitions of  $z_1$  and  $z_2$ ). We prove that strict inequality holds under the stronger assumption that  $p(\cdot, \cdot)$  is *strongly transient*, i.e.,  $\sum_{n=1}^{\infty} np^n(0) < \infty$ . This excludes  $\alpha \in (1, 2)$  and part of  $\alpha = 2$  in (1.1).

**Theorem 1.3.** Assume (1.1). If  $p(\cdot, \cdot)$  is strongly transient, then  $1 < z_2 < z_1 \leq \infty$ .

<sup>2</sup>By the symmetry of  $p(\cdot, \cdot)$ , we have  $\sup_{x \in \mathbb{Z}^d} p^n(x) \leq p^{2\lfloor n/2 \rfloor}(0)$  (see (3.15)), which implies that  $\sup_{n \in \mathbb{N}} \sup_{x \in \mathbb{Z}^d} \log[p^n(x)/p^{2\lfloor n/2 \rfloor}(0)] \leq 0$ .

Since  $\mathbb{P}(V = k) = (1 - \bar{F})\bar{F}^k$ ,  $k \in \mathbb{N} \cup \{0\}$ , with  $\bar{F} := \mathbb{P}(\exists k \in \mathbb{N}: S_k = S'_k)$ , an easy computation gives  $z_2 = 1/\bar{F}$ . But  $\bar{F} = 1 - [1/\bar{G}(0)]$  (see Spitzer [27], Section 1), and hence

$$z_2 = \bar{G}(0)/[\bar{G}(0) - 1]. \quad (1.15)$$

Unlike (1.15), no closed form expression is known for  $z_1$ . By evaluating the function inside the supremum in (1.13) at a well-chosen  $Q$ , we obtain the following upper bound.

**Theorem 1.4.** *Assume (1.1) and (1.10–1.12). Then*

$$z_1 \leq 1 + \left( \sum_{n \in \mathbb{N}} e^{-h(p^n)} \right)^{-1} < \infty, \quad (1.16)$$

where  $h(p^n) = -\sum_{x \in \mathbb{Z}^d} p^n(x) \log p^n(x)$  is the entropy of  $p^n(\cdot)$ .

There are symmetric transient random walks for which (1.1) holds with  $\alpha = 1$ . Examples are any transient random walk on  $\mathbb{Z}$  in the domain of attraction of the symmetric stable law of index 1 on  $\mathbb{R}$ , or any transient random walk on  $\mathbb{Z}^2$  in the domain of (non-normal) attraction of the normal law on  $\mathbb{R}^2$ . If in this situation (1.10–1.12) hold, then the two threshold values in (1.3–1.4) agree.

**Theorem 1.5.** *If  $p(\cdot, \cdot)$  satisfies (1.1) with  $\alpha = 1$  and (1.10–1.12), then  $z_1 = z_2$ .*

### 1.1.2 Continuous time

Next, we turn the discrete-time random walks  $S$  and  $S'$  into continuous-time random walks  $\tilde{S} = (S_t)_{t \geq 0}$  and  $\tilde{S}' = (\tilde{S}'_t)_{t \geq 0}$  by allowing them to make steps at rate 1, keeping the same  $p(\cdot, \cdot)$ . Then the collision local time becomes

$$\tilde{V} := \int_0^\infty 1_{\{\tilde{S}_t = \tilde{S}'_t\}} dt. \quad (1.17)$$

For the analogous quantities  $\tilde{z}_1$  and  $\tilde{z}_2$ , we have the following.<sup>3</sup>

**Theorem 1.6.** *Assume (1.1). If  $p(\cdot, \cdot)$  is strongly transient, then  $1 < \tilde{z}_2 < \tilde{z}_1 \leq \infty$ .*

An easy computation gives

$$\log \tilde{z}_2 = 2/G(0), \quad (1.18)$$

where  $G(0) = \sum_{n=0}^\infty p^n(0)$  is the Green function at the origin associated with  $p(\cdot, \cdot)$ . There is again no simple expression for  $\tilde{z}_1$ .

**Remark 1.7.** *An upper bound similar to (1.16) holds for  $\tilde{z}_1$  as well. It is straightforward to show that  $z_1 < \infty$  and  $\tilde{z}_1 < \infty$  as soon as  $p(\cdot)$  has finite entropy.*

<sup>3</sup>For a symmetric and recurrent random walk again trivially  $\tilde{z}_1 = \tilde{z}_2 = 1$ .

### 1.1.3 Discussion

Our proofs of Theorems 1.3–1.6 will be based on the variational representations in Theorem 1.1–1.2. Additional technical difficulties arise in the situation where the maximiser in (1.7) has infinite mean word length, which happens precisely when  $p(\cdot, \cdot)$  is transient but not strongly transient. Random walks with zero mean and finite variance are transient for  $d \geq 3$  and strongly transient for  $d \geq 5$  (Spitzer [27], Section 1).

**Conjecture 1.8.** *The gaps in Theorems 1.3 and 1.6 are present also when  $p(\cdot, \cdot)$  is transient but not strongly transient provided  $\alpha > 1$ .*

In a 2008 preprint by the authors (arXiv:0807.2611v1), the results in [6] and the present paper were announced, including Conjecture 1.8. Since then, *partial progress* has been made towards settling this conjecture. In Birkner and Sun [7], the gap in Theorem 1.3 is proved for simple random walk on  $\mathbb{Z}^d$ ,  $d \geq 4$ , and it is argued that the proof is in principle extendable to a symmetric random walk with finite variance. In Birkner and Sun [8], the gap in Theorem 1.6 is proved for a symmetric random walk on  $\mathbb{Z}^3$  with finite variance in continuous time, while in Berger and Toninelli [1] the gap in Theorem 1.3 is proved for a symmetric random walk on  $\mathbb{Z}^3$  in discrete time under a fourth moment condition.

The role of the variational representation for  $r_2$  is not to identify its value, which is achieved in (1.15), but rather to allow for a *comparison* with  $r_1$ , for which no explicit expression is available. It is an open problem to prove (1.11–1.12) under mild regularity conditions on  $S$ . Note that the gaps in Theorems 1.3–1.6 do not require (1.10–1.12).

## 1.2 The gaps settle three conjectures

In this section we use Theorems 1.3 and 1.6 to prove the existence of an *intermediate phase* for three classes of interacting particle systems where the interaction is controlled by a symmetric and transient random walk transition kernel.<sup>4</sup>

### 1.2.1 Coupled branching processes

**A.** Theorem 1.6 *proves* a conjecture put forward in Greven [17], [18]. Consider a spatial population model, defined as the Markov process  $(\eta_t)_{t \geq 0}$ , with  $\eta(t) = \{\eta_x(t) : x \in \mathbb{Z}^d\}$  where  $\eta_x(t)$  is the number of individuals at site  $x$  at time  $t$ , evolving as follows:

- (1) Each individual migrates at rate 1 according to  $a(\cdot, \cdot)$ .
- (2) Each individual gives birth to a new individual at the same site at rate  $b$ .
- (3) Each individual dies at rate  $(1 - p)b$ .
- (4) All individuals at the same site die simultaneously at rate  $pb$ .

---

<sup>4</sup>In each of these systems the case of a symmetric and recurrent random walk is trivial and no intermediate phase is present.

Here,  $a(\cdot, \cdot)$  is an irreducible random walk transition kernel on  $\mathbb{Z}^d \times \mathbb{Z}^d$ ,  $b \in (0, \infty)$  is a birth-death rate,  $p \in [0, 1]$  is a coupling parameter, while (1)–(4) occur independently at every  $x \in \mathbb{Z}^d$ . The case  $p = 0$  corresponds to a critical branching random walk, for which the average number of individuals per site is preserved. The case  $p > 0$  is challenging because the individuals descending from different ancestors are no longer independent.

A critical branching random walk satisfies the following *dichotomy* (where for simplicity we restrict to the case where  $a(\cdot, \cdot)$  is symmetric): if the initial configuration  $\eta_0$  is drawn from a shift-invariant and shift-ergodic probability distribution with a positive and finite mean, then  $\eta_t$  as  $t \rightarrow \infty$  locally dies out (“extinction”) when  $a(\cdot, \cdot)$  is *recurrent*, but converges to a non-trivial equilibrium (“survival”) when  $a(\cdot, \cdot)$  is *transient*, both irrespective of the value of  $b$ . In the latter case, the equilibrium has the same mean as the initial distribution and has all moments finite.

For the coupled branching process with  $p > 0$  there is a dichotomy too, but it is controlled by a *subtle interplay* of  $a(\cdot, \cdot)$ ,  $b$  and  $p$ : extinction holds when  $a(\cdot, \cdot)$  is *recurrent*, but also when  $a(\cdot, \cdot)$  is *transient* and  $p$  is *sufficiently large*. Indeed, it is shown in Greven [18] that if  $a(\cdot, \cdot)$  is transient, then there is a unique  $p_* \in (0, 1]$  such that survival holds for  $p < p_*$  and extinction holds for  $p > p_*$ .

Recall the critical values  $\tilde{z}_1, \tilde{z}_2$  introduced in Section 1.1.2. Then survival holds if  $\mathbb{E}(\exp[bp\tilde{V}] | \tilde{S}) < \infty$   $\tilde{S}$ -a.s., i.e., if  $p < p_1$  with

$$p_1 = 1 \wedge (b^{-1} \log \tilde{z}_1). \quad (1.19)$$

This can be shown by a size-biasing of the population in the spirit of Kallenberg [23]. On the other hand, survival with a *finite second moment* holds if and only if  $\mathbb{E}(\exp[bp\tilde{V}]) < \infty$ , i.e., if and only if  $p < p_2$  with

$$p_2 = 1 \wedge (b^{-1} \log \tilde{z}_2). \quad (1.20)$$

Clearly,  $p_* \geq p_1 \geq p_2$ . Theorem 1.6 shows that if  $a(\cdot, \cdot)$  satisfies (1.1) and is strongly transient, then  $p_1 > p_2$ , implying that there is an intermediate phase of survival with an *infinite second moment*.

**B.** Theorem 1.3 *corrects* an error in Birkner [3], Theorem 6. Here, a system of individuals living on  $\mathbb{Z}^d$  is considered subject to migration and branching. Each individual independently migrates at rate 1 according to a transient random walk transition kernel  $a(\cdot, \cdot)$ , and branches at a rate that depends on the number of individuals present at the same location. It is argued that this system has an intermediate phase in which the numbers of individuals at different sites tend to an equilibrium with a finite first moment but an infinite second moment. The proof was, however, based on a wrong rate function. The rate function claimed in Birkner [3], Theorem 6, must be replaced by that in [6], Corollary 1.5, after which the intermediate phase persists, at least in the case where  $a(\cdot, \cdot)$  satisfies (1.1) and is strongly transient. This also affects [3], Theorem 5, which uses [3], Theorem 6, to compute  $z_1$  in Section 1.1 and finds an incorrect formula. Theorem 1.4 shows that this formula actually is an upper bound for  $z_1$ .

### 1.2.2 Interacting diffusions

Theorem 1.6 *proves* a conjecture put forward in Greven and den Hollander [19]. Consider the system  $X = (X(t))_{t \geq 0}$ , with  $X(t) = \{X_x(t) : x \in \mathbb{Z}^d\}$ , of interacting diffusions taking values in  $[0, \infty)$  defined by the following collection of coupled stochastic differential equations:

$$dX_x(t) = \sum_{y \in \mathbb{Z}^d} a(x, y) [X_y(t) - X_x(t)] dt + \sqrt{bX_x(t)} dW_x(t), \quad x \in \mathbb{Z}^d, t \geq 0. \quad (1.21)$$

Here,  $a(\cdot, \cdot)$  is an irreducible random walk transition kernel on  $\mathbb{Z}^d \times \mathbb{Z}^d$ ,  $b \in (0, \infty)$  is a diffusion constant, and  $(W(t))_{t \geq 0}$  with  $W(t) = \{W_x(t) : x \in \mathbb{Z}^d\}$  is a collection of independent standard Brownian motions on  $\mathbb{R}$ . The initial condition is chosen such that  $X(0)$  is a shift-invariant and shift-ergodic random field with a positive and finite mean (the evolution preserves the mean). Note that, even though  $X$  a.s. has non-negative paths when starting from a non-negative initial condition  $X(0)$ , we prefer to write  $X_x(t)$  as  $\sqrt{X_x(t)^2}$  in order to highlight the fact that the system in (1.21) belongs to a more general family of interacting diffusions with Hölder- $\frac{1}{2}$  diffusion coefficients, see e.g. [19], Section 1, for a discussion and references.

It was shown in [19], Theorems 1.4–1.6, that if  $a(\cdot, \cdot)$  is *symmetric* and *transient*, then there exist  $0 < b_2 \leq b_*$  such that the system in (1.21) locally dies out when  $b > b_*$ , but converges to an equilibrium when  $0 < b < b_*$ , and this equilibrium has a *finite second moment* when  $0 < b < b_2$  and an *infinite second moment* when  $b_2 \leq b < b_*$ . It was shown in [19], Lemma 4.6, that  $b_* \geq b_{**} = \log \tilde{z}_1$ , and it was conjectured in [19], Conjecture 1.8, that  $b_* > b_2$ . Thus, as explained in [19], Section 4.2, if  $a(\cdot, \cdot)$  satisfies (1.1) and is strongly transient, then this conjecture is correct with

$$b_* \geq \log \tilde{z}_1 > b_2 = \log \tilde{z}_2. \quad (1.22)$$

Analogously, by Theorem 1.1 in [8] and by Theorem 1.2 in [7], the conjecture is settled for a class of random walks in dimensions  $d = 3, 4$  including symmetric simple random walk (which in  $d = 3, 4$  is transient but not strongly transient).

### 1.2.3 Directed polymers in random environments

Theorem 1.3 *disproves* a conjecture put forward in Monthus and Garel [25]. Let  $a(\cdot, \cdot)$  be a symmetric and irreducible random walk transition kernel on  $\mathbb{Z}^d \times \mathbb{Z}^d$ , let  $S = (S_k)_{k=0}^\infty$  be the corresponding random walk, and let  $\xi = \{\xi(x, n) : x \in \mathbb{Z}^d, n \in \mathbb{N}\}$  be i.i.d.  $\mathbb{R}$ -valued non-degenerate random variables satisfying

$$\lambda(\beta) := \log \mathbb{E}(\exp[\beta \xi(x, n)]) \in \mathbb{R} \quad \forall \beta \in \mathbb{R}. \quad (1.23)$$

Put

$$e_n(\xi, S) := \exp \left[ \sum_{k=1}^n \{\beta \xi(S_k, k) - \lambda(\beta)\} \right], \quad (1.24)$$

and set

$$Z_n(\xi) := \mathbb{E}[e_n(\xi, S)] = \sum_{s_1, \dots, s_n \in \mathbb{Z}^d} \left[ \prod_{k=1}^n p(s_{k-1}, s_k) \right] e_n(\xi, s), \quad s = (s_k)_{k=0}^\infty, s_0 = 0, \quad (1.25)$$

i.e.,  $Z_n(\xi)$  is the normalising constant in the probability distribution of the random walk  $S$  whose paths are reweighted by  $e_n(\xi, S)$ , which is referred to as the “polymer measure”. The  $\xi(x, n)$ ’s describe a random space-time medium with which  $S$  is interacting, with  $\beta$  playing the role of the interaction strength or inverse temperature.

It is well known that  $Z = (Z_n)_{n \in \mathbb{N}}$  is a non-negative martingale with respect to the family of sigma-algebras  $\mathcal{F}_n := \sigma(\xi(x, k), x \in \mathbb{Z}^d, 1 \leq k \leq n)$ ,  $n \in \mathbb{N}$ . Hence

$$\lim_{n \rightarrow \infty} Z_n = Z_\infty \geq 0 \quad \xi - a.s., \quad (1.26)$$

with the event  $\{Z_\infty = 0\}$  being  $\xi$ -trivial. One speaks of *weak disorder* if  $Z_\infty > 0$   $\xi$ -a.s. and of *strong disorder* otherwise. As shown in Comets and Yoshida [12], there is a unique critical value  $\beta_* \in [0, \infty]$  such that weak disorder holds for  $0 \leq \beta < \beta_*$  and strong disorder holds for  $\beta > \beta_*$ . Moreover, in the weak disorder region the paths have a Gaussian scaling limit under the polymer measure, while this is not the case in the strong disorder region. In the strong disorder region the paths are confined to a narrow space-time tube.

Recall the critical values  $z_1, z_2$  defined in Section 1.1. Bolthausen [9] observed that

$$\mathbb{E} [Z_n^2] = \mathbb{E} \left[ \exp \left[ \{\lambda(2\beta) - 2\lambda(\beta)\} V_n \right] \right], \quad \text{with } V_n := \sum_{k=1}^n 1_{\{S_k=S'_k\}}, \quad (1.27)$$

where  $S$  and  $S'$  are two independent random walks with transition kernel  $p(\cdot, \cdot)$ , and concluded that  $Z$  is  $L^2$ -bounded if and only if  $\beta < \beta_2$  with  $\beta_2 \in (0, \infty]$  the unique solution of

$$\lambda(2\beta_2) - 2\lambda(\beta_2) = \log z_2. \quad (1.28)$$

Since  $\mathbb{P}(Z_\infty > 0) \geq \mathbb{E}[Z_\infty]^2 / \mathbb{E}[Z_\infty^2]$  and  $\mathbb{E}[Z_\infty] = Z_0 = 1$  for an  $L^2$ -bounded martingale, it follows that  $\beta < \beta_2$  implies weak disorder, i.e.,  $\beta_* \geq \beta_2$ . By a stochastic representation of the size-biased law of  $Z_n$ , it was shown in Birkner [4], Proposition 1, that in fact weak disorder holds if  $\beta < \beta_1$  with  $\beta_1 \in (0, \infty]$  the unique solution of

$$\lambda(2\beta_1) - 2\lambda(\beta_1) = \log z_1, \quad (1.29)$$

i.e.,  $\beta_* \geq \beta_1$ . Since  $\beta \mapsto \lambda(2\beta) - 2\lambda(\beta)$  is strictly increasing for any non-trivial law for the disorder satisfying (1.23), it follows from (1.28–1.29) and Theorem 1.3 that  $\beta_1 > \beta_2$  when  $a(\cdot, \cdot)$  satisfies (1.1) and is strongly transient and when  $\xi$  is such that  $\beta_2 < \infty$ . In that case the weak disorder region contains a subregion for which  $Z$  is *not*  $L^2$ -bounded. This disproves a conjecture of Monthus and Garel [25], who argued that  $\beta_2 = \beta_*$ .

Camanes and Carmona [10] consider the same problem for simple random walk and specific choices of disorder. With the help of fractional moment estimates of Evans and Derrida [16], combined with numerical computation, they show that  $\beta_* > \beta_2$  for Gaussian disorder in  $d \geq 5$ , for Binomial disorder with small mean in  $d \geq 4$ , and for Poisson disorder with small mean in  $d \geq 3$ .

See den Hollander [21], Chapter 12, for an overview.

## Outline

Theorems 1.1, 1.3 and 1.6 are proved in Section 3. The proofs need only assumption (1.1). Theorem 1.2 is proved in Section 4, Theorems 1.4 and 1.5 in Section 5. The proofs need both assumptions (1.1) and (1.10–1.12)

In Section 2 we recall the LDP's in [6], which are needed for the proof of Theorems 1.1–1.2 and their counterparts for continuous-time random walk. This section recalls the minimum from [6] that is needed for the present paper. Only in Section 4 will we need some of the techniques that were used in [6].



## 2 Word sequences and annealed and quenched LDP

**Notation.** We recall the problem setting in [6]. Let  $E$  be a finite or countable set of *letters*. Let  $\tilde{E} = \cup_{n \in \mathbb{N}} E^n$  be the set of finite *words* drawn from  $E$ . Both  $E$  and  $\tilde{E}$  are Polish spaces under the discrete topology. Let  $\mathcal{P}(E^{\mathbb{N}})$  and  $\mathcal{P}(\tilde{E}^{\mathbb{N}})$  denote the set of probability measures on sequences drawn from  $E$ , respectively,  $\tilde{E}$ , equipped with the topology of weak convergence. Write  $\theta$  and  $\tilde{\theta}$  for the left-shift acting on  $E^{\mathbb{N}}$ , respectively,  $\tilde{E}^{\mathbb{N}}$ . Write  $\mathcal{P}^{\text{inv}}(E^{\mathbb{N}})$ ,  $\mathcal{P}^{\text{erg}}(E^{\mathbb{N}})$  and  $\mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$ ,  $\mathcal{P}^{\text{erg}}(\tilde{E}^{\mathbb{N}})$  for the set of probability measures that are invariant and ergodic under  $\theta$ , respectively,  $\tilde{\theta}$ .

For  $\nu \in \mathcal{P}(E)$ , let  $X = (X_i)_{i \in \mathbb{N}}$  be i.i.d. with law  $\nu$ . For  $\rho \in \mathcal{P}(\mathbb{N})$ , let  $\tau = (\tau_i)_{i \in \mathbb{N}}$  be i.i.d. with law  $\rho$  having infinite support and satisfying the *algebraic tail property*

$$\lim_{\substack{n \rightarrow \infty \\ \rho(n) > 0}} \frac{\log \rho(n)}{\log n} =: -\alpha, \quad \alpha \in [1, \infty). \quad (2.1)$$

(No regularity assumption is imposed on  $\text{supp}(\rho)$ .) Assume that  $X$  and  $\tau$  are independent and write  $\mathbb{P}$  to denote their joint law. Cut words out of  $X$  according to  $\tau$ , i.e., put (see Fig. 1)

$$T_0 := 0 \quad \text{and} \quad T_i := T_{i-1} + \tau_i, \quad i \in \mathbb{N}, \quad (2.2)$$

and let

$$Y^{(i)} := (X_{T_{i-1}+1}, X_{T_{i-1}+2}, \dots, X_{T_i}), \quad i \in \mathbb{N}. \quad (2.3)$$

Then, under the law  $\mathbb{P}$ ,  $Y = (Y^{(i)})_{i \in \mathbb{N}}$  is an i.i.d. sequence of words with marginal law  $q_{\rho, \nu}$  on  $\tilde{E}$  given by

$$q_{\rho, \nu}((x_1, \dots, x_n)) := \mathbb{P}(Y^{(1)} = (x_1, \dots, x_n)) = \rho(n) \nu(x_1) \cdots \nu(x_n), \quad n \in \mathbb{N}, x_1, \dots, x_n \in E. \quad (2.4)$$

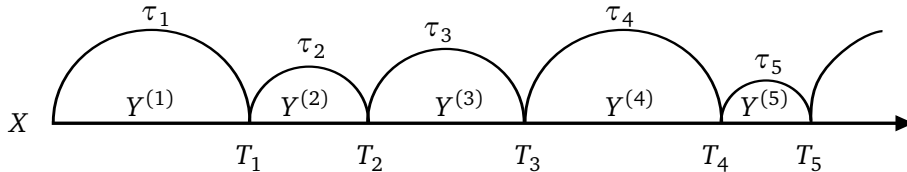


Figure 1: Cutting words from a letter sequence according to a renewal process.

**Annealed LDP** For  $N \in \mathbb{N}$ , let  $(Y^{(1)}, \dots, Y^{(N)})^{\text{per}}$  be the periodic extension of  $(Y^{(1)}, \dots, Y^{(N)})$  to an element of  $\tilde{E}^{\mathbb{N}}$ , and define

$$R_N := \frac{1}{N} \sum_{i=0}^{N-1} \delta_{\tilde{\theta}^i(Y^{(1)}, \dots, Y^{(N)})^{\text{per}}} \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}}), \quad (2.5)$$

the *empirical process of  $N$ -tuples of words*. The following large deviation principle (LDP) is standard (see e.g. Dembo and Zeitouni [14], Corollaries 6.5.15 and 6.5.17). Let

$$H(Q | q_{\rho, \nu}^{\otimes \mathbb{N}}) := \lim_{N \rightarrow \infty} \frac{1}{N} h \left( Q_{|\mathcal{F}_N} \mid (q_{\rho, \nu}^{\otimes \mathbb{N}})_{|\mathcal{F}_N} \right) \in [0, \infty] \quad (2.6)$$

be the *specific relative entropy* of  $Q$  w.r.t.  $q_{\rho, \nu}^{\otimes \mathbb{N}}$ , where  $\mathcal{F}_N = \sigma(Y^{(1)}, \dots, Y^{(N)})$  is the sigma-algebra generated by the first  $N$  words,  $Q|_{\mathcal{F}_N}$  is the restriction of  $Q$  to  $\mathcal{F}_N$ , and  $h(\cdot | \cdot)$  denotes relative entropy (defined for probability measures  $\varphi, \psi$  on a measurable space  $F$  as  $h(\varphi | \psi) = \int_F \log \frac{d\varphi}{d\psi} d\varphi$  if the density  $\frac{d\varphi}{d\psi}$  exists and as  $\infty$  otherwise).

**Theorem 2.1. [Annealed LDP]** *The family of probability distributions  $\mathbb{P}(R_N \in \cdot)$ ,  $N \in \mathbb{N}$ , satisfies the LDP on  $\mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$  with rate  $N$  and with rate function  $I^{\text{ann}}: \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}}) \rightarrow [0, \infty]$  given by*

$$I^{\text{ann}}(Q) = H(Q | q_{\rho, \nu}^{\otimes \mathbb{N}}). \quad (2.7)$$

The rate function  $I^{\text{ann}}$  is lower semi-continuous, has compact level sets, has a unique zero at  $Q = q_{\rho, \nu}^{\otimes \mathbb{N}}$ , and is affine.

**Quenched LDP.** To formulate the quenched analogue of Theorem 2.1, we need some further notation. Let  $\kappa: \tilde{E}^{\mathbb{N}} \rightarrow E^{\mathbb{N}}$  denote the *concatenation map* that glues a sequence of words into a sequence of letters. For  $Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$  such that  $m_Q := \mathbb{E}_Q[\tau_1] < \infty$  (recall that  $\tau_1$  is the length of the first word), define  $\Psi_Q \in \mathcal{P}^{\text{inv}}(E^{\mathbb{N}})$  as

$$\Psi_Q(\cdot) := \frac{1}{m_Q} \mathbb{E}_Q \left[ \sum_{k=0}^{\tau_1-1} \delta_{\theta^k \kappa(Y)}(\cdot) \right]. \quad (2.8)$$

Think of  $\Psi_Q$  as the *shift-invariant version of the concatenation of  $Y$  under the law  $Q$  obtained after randomising the location of the origin*.

For  $\text{tr} \in \mathbb{N}$ , let  $[\cdot]_{\text{tr}}: \tilde{E} \rightarrow [\tilde{E}]_{\text{tr}} := \cup_{n=1}^{\text{tr}} E^n$  denote the *word length truncation map* defined by

$$y = (x_1, \dots, x_n) \mapsto [y]_{\text{tr}} := (x_1, \dots, x_{n \wedge \text{tr}}), \quad n \in \mathbb{N}, x_1, \dots, x_n \in E. \quad (2.9)$$

Extend this to a map from  $\tilde{E}^{\mathbb{N}}$  to  $[\tilde{E}]_{\text{tr}}^{\mathbb{N}}$  via

$$[(Y^{(1)}, Y^{(2)}, \dots)]_{\text{tr}} := ([Y^{(1)}]_{\text{tr}}, [Y^{(2)}]_{\text{tr}}, \dots), \quad (2.10)$$

and to a map from  $\mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$  to  $\mathcal{P}^{\text{inv}}([\tilde{E}]_{\text{tr}}^{\mathbb{N}})$  via

$$[Q]_{\text{tr}}(A) := Q(\{z \in \tilde{E}^{\mathbb{N}}: [z]_{\text{tr}} \in A\}), \quad A \subset [\tilde{E}]_{\text{tr}}^{\mathbb{N}} \text{ measurable}. \quad (2.11)$$

Note that if  $Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$ , then  $[Q]_{\text{tr}}$  is an element of the set

$$\mathcal{P}^{\text{inv, fin}}(\tilde{E}^{\mathbb{N}}) = \{Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}}): m_Q < \infty\}. \quad (2.12)$$

**Theorem 2.2. [Quenched LDP see [6], Theorem 1.2 and Corollary 1.6]** *(a) Assume (2.1). Then, for  $\nu^{\otimes \mathbb{N}}$ -a.s. all  $X$ , the family of (regular) conditional probability distributions  $\mathbb{P}(R_N \in \cdot | X)$ ,  $N \in \mathbb{N}$ , satisfies the LDP on  $\mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$  with rate  $N$  and with deterministic rate function  $I^{\text{que}}: \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}}) \rightarrow [0, \infty]$  given by*

$$I^{\text{que}}(Q) := \begin{cases} I^{\text{fin}}(Q), & \text{if } Q \in \mathcal{P}^{\text{inv, fin}}(\tilde{E}^{\mathbb{N}}), \\ \lim_{\text{tr} \rightarrow \infty} I^{\text{fin}}([Q]_{\text{tr}}), & \text{otherwise,} \end{cases} \quad (2.13)$$

where

$$I^{\text{fin}}(Q) := H(Q | q_{\rho, \nu}^{\otimes \mathbb{N}}) + (\alpha - 1)m_Q H(\Psi_Q | \nu^{\otimes \mathbb{N}}). \quad (2.14)$$

The rate function  $I^{\text{que}}$  is lower semi-continuous, has compact level sets, has a unique zero at  $Q = q_{\rho, \nu}^{\otimes \mathbb{N}}$ , and is affine. Moreover, it is equal to the lower semi-continuous extension of  $I^{\text{fin}}$  from  $\mathcal{P}^{\text{inv, fin}}(\tilde{E}^{\mathbb{N}})$  to  $\mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$ .

(b) In particular, if (2.1) holds with  $\alpha = 1$ , then for  $\nu^{\otimes \mathbb{N}}$ -a.s. all  $X$ , the family  $\mathbb{P}(R_N \in \cdot | X)$  satisfies the LDP with rate function  $I^{\text{ann}}$  given by (2.7).

Note that the quenched rate function (2.14) equals the annealed rate function (2.7) plus an additional term that quantifies the deviation of  $\Psi_Q$  from the reference law  $\nu^{\otimes \mathbb{N}}$  on the letter sequence. This term is explicit when  $m_Q < \infty$ , but requires a truncation approximation when  $m_Q = \infty$ .

We close this section with the following observation. Let

$$\mathcal{R}_\nu := \left\{ Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}}) : \text{w-}\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{k=0}^{L-1} \delta_{\theta^k(Y)} = \nu^{\otimes \mathbb{N}} \text{ } Q - \text{a.s.} \right\}. \quad (2.15)$$

be the set of  $Q$ 's for which the concatenation of words has the same statistical properties as the letter sequence  $X$ . Then, for  $Q \in \mathcal{P}^{\text{inv, fin}}(\tilde{E}^{\mathbb{N}})$ , we have (see [6], Equation (1.22))

$$\Psi_Q = \nu^{\otimes \mathbb{N}} \iff I^{\text{que}}(Q) = I^{\text{ann}}(Q) \iff Q \in \mathcal{R}_\nu. \quad (2.16)$$

### 3 Proof of Theorems 1.1, 1.3 and 1.6

#### 3.1 Proof of Theorem 1.1

The idea is to put the problem into the framework of (2.1–2.5) and then apply Theorem 2.2. To that end, we pick

$$E := \mathbb{Z}^d, \quad \tilde{E} = \widetilde{\mathbb{Z}^d} := \cup_{n \in \mathbb{N}} (\mathbb{Z}^d)^n, \quad (3.1)$$

and choose

$$\nu(u) := p(u), \quad u \in \mathbb{Z}^d, \quad \rho(n) := \frac{p^{2\lfloor n/2 \rfloor}(0)}{2\bar{G}(0) - 1}, \quad n \in \mathbb{N}, \quad (3.2)$$

where

$$p(u) = p(0, u), \quad u \in \mathbb{Z}^d, \quad p^n(\nu - u) = p^n(u, \nu), \quad u, \nu \in \mathbb{Z}^d, \quad \bar{G}(0) = \sum_{n=0}^{\infty} p^{2n}(0), \quad (3.3)$$

the latter being the Green function of  $S - S'$  at the origin.

Recalling (1.2), and writing

$$z^V = ((z - 1) + 1)^V = 1 + \sum_{N=1}^V (z - 1)^N \binom{V}{N} \quad (3.4)$$

with

$$\binom{V}{N} = \sum_{0 < j_1 < \dots < j_N < \infty} \mathbf{1}_{\{S_{j_1} = S'_{j_1}, \dots, S_{j_N} = S'_{j_N}\}}, \quad (3.5)$$

we have

$$\mathbb{E} \left[ z^V \mid S \right] = 1 + \sum_{N=1}^{\infty} (z-1)^N F_N^{(1)}(X), \quad \mathbb{E} \left[ z^V \right] = 1 + \sum_{N=1}^{\infty} (z-1)^N F_N^{(2)}, \quad (3.6)$$

with

$$F_N^{(1)}(X) := \sum_{0 < j_1 < \dots < j_N < \infty} \mathbb{P}(S_{j_1} = S'_{j_1}, \dots, S_{j_N} = S'_{j_N} \mid X), \quad F_N^{(2)} := \mathbb{E} [F_N^{(1)}(X)], \quad (3.7)$$

where  $X = (X_k)_{k \in \mathbb{N}}$  denotes the sequence of increments of  $S$ . (The upper indices 1 and 2 indicate the number of random walks being averaged over.)

The notation in (3.1–3.2) allows us to rewrite the first formula in (3.7) as

$$\begin{aligned} F_N^{(1)}(X) &= \sum_{0 < j_1 < \dots < j_N < \infty} \prod_{i=1}^N p^{j_i - j_{i-1}} \left( \sum_{k=1}^{j_i - j_{i-1}} X_{j_{i-1} + k} \right) \\ &= \sum_{0 < j_1 < \dots < j_N < \infty} \prod_{i=1}^N \rho(j_i - j_{i-1}) \exp \left[ \sum_{i=1}^N \log \left( \frac{p^{j_i - j_{i-1}} (\sum_{k=1}^{j_i - j_{i-1}} X_{j_{i-1} + k})}{\rho(j_i - j_{i-1})} \right) \right]. \end{aligned} \quad (3.8)$$

Let  $Y^{(i)} = (X_{j_{i-1}+1}, \dots, X_{j_i})$ . Recall the definition of  $f : \widetilde{\mathbb{Z}^d} \rightarrow [0, \infty)$  in (1.5),

$$f((x_1, \dots, x_n)) = \frac{p^n (x_1 + \dots + x_n)}{p^{2\lfloor n/2 \rfloor}(0)} [2\bar{G}(0) - 1], \quad n \in \mathbb{N}, x_1, \dots, x_n \in \mathbb{Z}^d. \quad (3.9)$$

Note that, since  $\widetilde{\mathbb{Z}^d}$  carries the discrete topology,  $f$  is trivially continuous.

Let  $R_N \in \mathcal{P}^{\text{inv}}((\widetilde{\mathbb{Z}^d})^{\mathbb{N}})$  be the empirical process of words defined in (2.5), and  $\pi_1 R_N \in \mathcal{P}(\widetilde{\mathbb{Z}^d})$  the projection of  $R_N$  onto the first coordinate. Then we have

$$F_N^{(1)}(X) = \mathbb{E} \left[ \exp \left( \sum_{i=1}^N \log f(Y^{(i)}) \right) \middle| X \right] = \mathbb{E} \left[ \exp \left( N \int_{\widetilde{\mathbb{Z}^d}} (\pi_1 R_N)(dy) \log f(y) \right) \middle| X \right], \quad (3.10)$$

where  $\mathbb{P}$  is the joint law of  $X$  and  $\tau$  (recall (2.2–2.3)). By averaging (3.10) over  $X$  we obtain (recall the definition of  $F_N^{(2)}$  from (3.7))

$$F_N^{(2)} = \mathbb{E} \left[ \exp \left( N \int_{\widetilde{\mathbb{Z}^d}} (\pi_1 R_N)(dy) \log f(y) \right) \right]. \quad (3.11)$$

Without conditioning on  $X$ , the sequence  $(Y^{(i)})_{i \in \mathbb{N}}$  is i.i.d. with law (recall (2.4))

$$q_{\rho, \nu}^{\otimes \mathbb{N}} \quad \text{with} \quad q_{\rho, \nu}(x_1, \dots, x_n) = \frac{p^{2\lfloor n/2 \rfloor}(0)}{2\bar{G}(0) - 1} \prod_{k=1}^n p(x_k), \quad n \in \mathbb{N}, x_1, \dots, x_n \in \mathbb{Z}^d. \quad (3.12)$$

Next we note that

$$f \text{ in (3.9) is bounded from above.} \quad (3.13)$$

Indeed, the Fourier representation of  $p^n(x, y)$  reads

$$p^n(x) = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} dk e^{-i(k \cdot x)} \widehat{p}(k)^n \quad (3.14)$$

with  $\widehat{p}(k) = \sum_{x \in \mathbb{Z}^d} e^{i(k \cdot x)} p(0, x)$ . Because  $p(\cdot, \cdot)$  is symmetric, we have  $\widehat{p}(k) \in [-1, 1]$ , and it follows that

$$\max_{x \in \mathbb{Z}^d} p^{2n}(x) = p^{2n}(0), \quad \max_{x \in \mathbb{Z}^d} p^{2n+1}(x) \leq p^{2n}(0), \quad \forall n \in \mathbb{N}. \quad (3.15)$$

Consequently,  $f((x_1, \dots, x_n)) \leq [2\bar{G}(0) - 1]$  is bounded from above. Therefore, by applying the *annealed* LDP in Theorem 2.1 to (3.11), in combination with Varadhan's lemma (see Dembo and Zeitouni [14], Lemma 4.3.6), we get  $z_2 = 1 + \exp[-r_2]$  with

$$\begin{aligned} r_2 &:= \lim_{N \rightarrow \infty} \frac{1}{N} \log F_N^{(2)} \leq \sup_{Q \in \mathcal{P}^{\text{inv}}(\widetilde{\mathbb{Z}^d}^{\mathbb{N}})} \left\{ \int_{\widetilde{\mathbb{Z}^d}} (\pi_1 Q)(dy) \log f(y) - I^{\text{ann}}(Q) \right\} \\ &= \sup_{q \in \mathcal{P}(\widetilde{\mathbb{Z}^d})} \left\{ \int_{\widetilde{\mathbb{Z}^d}} q(dy) \log f(y) - h(q | q_{\rho, \nu}) \right\} \end{aligned} \quad (3.16)$$

(recall (1.3–1.4) and (3.6)). The second equality in (3.16) stems from the fact that, on the set of  $Q$ 's with a given marginal  $\pi_1 Q = q$ , the function  $Q \mapsto I^{\text{ann}}(Q) = H(Q | q_{\rho, \nu}^{\otimes \mathbb{N}})$  has a unique minimiser  $Q = q^{\otimes \mathbb{N}}$  (due to convexity of relative entropy). We will see in a moment that the inequality in (3.16) actually is an equality.

In order to carry out the second supremum in (3.16), we use the following.

**Lemma 3.1.** *Let  $Z := \sum_{y \in \widetilde{\mathbb{Z}^d}} f(y) q_{\rho, \nu}(y)$ . Then*

$$\int_{\widetilde{\mathbb{Z}^d}} q(dy) \log f(y) - h(q | q_{\rho, \nu}) = \log Z - h(q | q^*) \quad \forall q \in \mathcal{P}(\widetilde{\mathbb{Z}^d}), \quad (3.17)$$

where  $q^*(y) := f(y) q_{\rho, \nu}(y) / Z$ ,  $y \in \widetilde{\mathbb{Z}^d}$ .

*Proof.* This follows from a straightforward computation. □

Inserting (3.17) into (3.16), we see that the suprema are uniquely attained at  $q = q^*$  and  $Q = Q^* = (q^*)^{\otimes \mathbb{N}}$ , and that  $r_2 \leq \log Z$ . From (3.9) and (3.12), we have

$$Z = \sum_{n \in \mathbb{N}} \sum_{x_1, \dots, x_n \in \mathbb{Z}^d} p^n(x_1 + \dots + x_n) \prod_{k=1}^n p(x_k) = \sum_{n \in \mathbb{N}} p^{2n}(0) = \bar{G}(0) - 1, \quad (3.18)$$

where we use that  $\sum_{v \in \mathbb{Z}^d} p^m(u + v) p(v) = p^{m+1}(u)$ ,  $u \in \mathbb{Z}^d$ ,  $m \in \mathbb{N}$ , and recall that  $\bar{G}(0)$  is the Green function at the origin associated with  $p^2(\cdot, \cdot)$ . Hence  $q^*$  is given by

$$q^*(x_1, \dots, x_n) = \frac{p^n(x_1 + \dots + x_n)}{\bar{G}(0) - 1} \prod_{k=1}^n p(x_k), \quad n \in \mathbb{N}, x_1, \dots, x_n \in \mathbb{Z}^d. \quad (3.19)$$

Moreover, since  $z_2 = \bar{G}(0)/[\bar{G}(0) - 1]$ , as noted in (1.15), we see that  $z_2 = 1 + \exp[-\log Z]$ , i.e.,  $r_2 = \log Z$ , and so indeed equality holds in (3.16).

The *quenched* LDP in Theorem 2.2, together with Varadhan's lemma applied to (3.8), gives  $z_1 = 1 + \exp[-r_1]$  with

$$r_1 := \lim_{N \rightarrow \infty} \frac{1}{N} \log F_N^{(1)}(X) \leq \sup_{Q \in \mathcal{P}^{\text{inv}}(\widehat{\mathbb{Z}^d}^{\mathbb{N}})} \left\{ \int_{\widehat{\mathbb{Z}^d}} (\pi_1 Q)(dy) \log f(y) - I^{\text{que}}(Q) \right\} \quad X - a.s., \quad (3.20)$$

where  $I^{\text{que}}(Q)$  is given by (2.13–2.14). Without further assumptions, we are not able to reverse the inequality in (3.20). This point will be addressed in Section 4 and will require assumptions (1.10–1.12).

### 3.2 Proof of Theorem 1.3

To compare (3.20) with (3.16), we need the following lemma, the proof of which is deferred.

**Lemma 3.2.** *Assume (1.1). Let  $Q^* = (q^*)^{\otimes \mathbb{N}}$  with  $q^*$  as in (3.19). If  $m_{Q^*} < \infty$ , then  $I^{\text{que}}(Q^*) > I^{\text{ann}}(Q^*)$ .*

With the help of Lemma 3.2 we complete the proof of the existence of the gap as follows. Since  $\log f$  is bounded from above, the function

$$Q \mapsto \int_{\widehat{\mathbb{Z}^d}} (\pi_1 Q)(dy) \log f(y) - I^{\text{que}}(Q) \quad (3.21)$$

is upper semicontinuous. Therefore, by compactness of the level sets of  $I^{\text{que}}(Q)$ , the function in (3.21) achieves its maximum at some  $Q^{**}$  that satisfies

$$r_1 = \int_{\widehat{\mathbb{Z}^d}} (\pi_1 Q^{**})(dy) \log f(y) - I^{\text{que}}(Q^{**}) \leq \int_{\widehat{\mathbb{Z}^d}} (\pi_1 Q^{**})(dy) \log f(y) - I^{\text{ann}}(Q^{**}) \leq r_2. \quad (3.22)$$

If  $r_1 = r_2$ , then  $Q^{**} = Q^*$ , because the function

$$Q \mapsto \int_{\widehat{\mathbb{Z}^d}} (\pi_1 Q)(dy) \log f(y) - I^{\text{ann}}(Q) \quad (3.23)$$

has  $Q^*$  as its unique maximiser (recall the discussion immediately after Lemma 3.1). But  $I^{\text{que}}(Q^*) > I^{\text{ann}}(Q^*)$  by Lemma 3.2, and so we have a contradiction in (3.22), thus arriving at  $r_1 < r_2$ .

In the remainder of this section we prove Lemma 3.2.

*Proof.* Note that

$$q^*((\mathbb{Z}^d)^n) = \sum_{x_1, \dots, x_n \in \mathbb{Z}^d} \frac{p^n(x_1 + \dots + x_n)}{\bar{G}(0) - 1} \prod_{k=1}^n p(x_k) = \frac{p^{2n}(0)}{\bar{G}(0) - 1}, \quad n \in \mathbb{N}, \quad (3.24)$$

and hence, by assumption (1.2),

$$\lim_{n \rightarrow \infty} \frac{\log q^*((\mathbb{Z}^d)^n)}{\log n} = -\alpha \quad (3.25)$$

and

$$m_{Q^*} = \sum_{n=1}^{\infty} nq^*((\mathbb{Z}^d)^n) = \sum_{n=1}^{\infty} \frac{np^{2n}(0)}{\bar{G}(0) - 1}. \quad (3.26)$$

The latter formula shows that  $m_{Q^*} < \infty$  if and only if  $p(\cdot, \cdot)$  is strongly transient. We will show that

$$m_{Q^*} < \infty \implies Q^* = (q^*)^{\otimes \mathbb{N}} \notin \mathcal{R}_v, \quad (3.27)$$

the set defined in (2.15). This implies  $\Psi_{Q^*} \neq v^{\otimes \mathbb{N}}$  (recall (2.16)), and hence  $H(\Psi_{Q^*} | v^{\otimes \mathbb{N}}) > 0$ , implying the claim because  $\alpha \in (1, \infty)$  (recall (2.14)).

In order to verify (3.27), we compute the first two marginals of  $\Psi_{Q^*}$ . Using the symmetry of  $p(\cdot, \cdot)$ , we have

$$\Psi_{Q^*}(a) = \frac{1}{m_{Q^*}} \sum_{n=1}^{\infty} \sum_{j=1}^n \sum_{\substack{x_1, \dots, x_n \in \mathbb{Z}^d \\ x_j = a}} \frac{p^n(x_1 + \dots + x_n)}{\bar{G}(0) - 1} \prod_{k=1}^n p(x_k) = p(a) \frac{\sum_{n=1}^{\infty} np^{2n-1}(a)}{\sum_{n=1}^{\infty} np^{2n}(0)}. \quad (3.28)$$

Hence,  $\Psi_{Q^*}(a) = p(a)$  for all  $a \in \mathbb{Z}^d$  with  $p(a) > 0$  if and only if

$$a \mapsto \sum_{n=1}^{\infty} np^{2n-1}(a) \text{ is constant on the support of } p(\cdot). \quad (3.29)$$

There are many  $p(\cdot, \cdot)$ 's for which (3.29) fails, and for these (3.27) holds. However, for simple random walk (3.29) does not fail, because  $a \mapsto p^{2n-1}(a)$  is constant on the  $2d$  neighbours of the origin, and so we have to look at the two-dimensional marginal.

Observe that  $q^*(x_1, \dots, x_n) = q^*(x_{\sigma(1)}, \dots, x_{\sigma(n)})$  for any permutation  $\sigma$  of  $\{1, \dots, n\}$ . For  $a, b \in \mathbb{Z}^d$ , we have

$$\begin{aligned} m_{Q^*} \Psi_{Q^*}(a, b) &= \mathbb{E}_{Q^*} \left[ \sum_{k=1}^{\tau_1} 1_{\kappa(Y)_k = a, \kappa(Y)_{k+1} = b} \right] \\ &= \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \sum_{x_1, \dots, x_{n+n'}} q^*(x_1, \dots, x_n) q^*(x_{n+1}, \dots, x_{n+n'}) \sum_{k=1}^n 1_{(a,b)}(x_k, x_{k+1}) \\ &= q^*(x_1 = a) q^*(x_1 = b) + \sum_{n=2}^{\infty} (n-1) q^*({(a,b)} \times (\mathbb{Z}^d)^{n-2}). \end{aligned} \quad (3.30)$$

Since

$$\begin{aligned} q^*(x_1 = a) &= \frac{p(a)^2}{\bar{G}(0) - 1} + \sum_{n=2}^{\infty} \sum_{x_2, \dots, x_n \in \mathbb{Z}^d} \frac{p^n(a + x_2 + \dots + x_n)}{\bar{G}(0) - 1} p(a) \prod_{k=2}^n p(x_k) \\ &= \frac{p(a)}{\bar{G}(0) - 1} \sum_{n=1}^{\infty} p^{2n-1}(a) \end{aligned} \quad (3.31)$$

and

$$\begin{aligned}
q^* (\{(a, b)\} \times (\mathbb{Z}^d)^{n-2}) &= 1_{n=2} \frac{p(a)p(b)}{\bar{G}(0) - 1} p^2(a + b) \\
&\quad + 1_{n \geq 3} \frac{p(a)p(b)}{\bar{G}(0) - 1} \sum_{x_3, \dots, x_n \in \mathbb{Z}^d} p^n(a + b + x_3 + \dots + x_n) \prod_{k=3}^n p(x_k) \quad (3.32) \\
&= \frac{p(a)p(b)}{\bar{G}(0) - 1} p^{2n-2}(a + b),
\end{aligned}$$

we find (recall that  $\bar{G}(0) - 1 = \sum_{n=1}^{\infty} p^{2n}(0)$ )

$$\begin{aligned}
\Psi_{Q^*}(a, b) &= \frac{p(a)p(b)}{\left[ \sum_{n=1}^{\infty} p^{2n}(0) \right] \left[ \sum_{n=1}^{\infty} np^{2n}(0) \right]} \left( \left[ \sum_{n=1}^{\infty} p^{2n-1}(a) \right] \left[ \sum_{n=1}^{\infty} p^{2n-1}(b) \right] \right. \\
&\quad \left. + \left[ \sum_{n=1}^{\infty} p^{2n}(0) \right] \left[ \sum_{n=2}^{\infty} (n-1)p^{2n-2}(a+b) \right] \right). \quad (3.33)
\end{aligned}$$

Pick  $b = -a$  with  $p(a) > 0$ . Then, shifting  $n$  to  $n - 1$  in the last sum, we get

$$\frac{\Psi_{Q^*}(a, -a)}{p(a)^2} - 1 = \frac{\left[ \sum_{n=1}^{\infty} p^{2n-1}(a) \right]^2}{\left[ \sum_{n=1}^{\infty} p^{2n}(0) \right] \left[ \sum_{n=1}^{\infty} np^{2n}(0) \right]} > 0. \quad (3.34)$$

This shows that consecutive letters are not uncorrelated under  $\Psi_{Q^*}$ , and implies that (3.27) holds as claimed.  $\square$

### 3.3 Proof of Theorem 1.6

The proof follows the line of argument in Section 3.2. The analogues of (3.4–3.7) are

$$z^{\tilde{V}} = \sum_{N=0}^{\infty} (\log z)^N \frac{\tilde{V}^N}{N!}, \quad (3.35)$$

with

$$\frac{\tilde{V}^N}{N!} = \int_0^{\infty} dt_1 \cdots \int_{t_{N-1}}^{\infty} dt_N 1_{\{\tilde{s}_{t_1} = \tilde{s}'_{t_1}, \dots, \tilde{s}_{t_N} = \tilde{s}'_{t_N}\}}, \quad (3.36)$$

and

$$\mathbb{E} \left[ z^{\tilde{V}} \mid \tilde{S} \right] = \sum_{N=0}^{\infty} (\log z)^N F_N^{(1)}(\tilde{S}), \quad \mathbb{E} \left[ z^{\tilde{V}} \right] = \sum_{N=0}^{\infty} (\log z)^N F_N^{(2)}, \quad (3.37)$$

with

$$F_N^{(1)}(\tilde{S}) := \int_0^{\infty} dt_1 \cdots \int_{t_{N-1}}^{\infty} dt_N \mathbb{P} \left( \tilde{S}_{t_1} = \tilde{S}'_{t_1}, \dots, \tilde{S}_{t_N} = \tilde{S}'_{t_N} \mid \tilde{S} \right), \quad F_N^{(2)} := \mathbb{E} [F_N^{(1)}(\tilde{S})], \quad (3.38)$$



where the conditioning in the first expression in (3.37) is on the full continuous-time path  $\tilde{S} = (\tilde{S}_t)_{t \geq 0}$ . Our task is to compute

$$\tilde{r}_1 := \lim_{N \rightarrow \infty} \frac{1}{N} \log F_N^{(1)}(\tilde{S}) \quad \tilde{S} - a.s., \quad \tilde{r}_2 := \lim_{N \rightarrow \infty} \frac{1}{N} \log F_N^{(2)}, \quad (3.39)$$

and show that  $\tilde{r}_1 < \tilde{r}_2$ .

In order to do so, we write  $\tilde{S}_t = X_{J_t}^{\natural}$ , where  $X^{\natural}$  is the discrete-time random walk with transition kernel  $p(\cdot, \cdot)$  and  $(J_t)_{t \geq 0}$  is the rate-1 Poisson process on  $[0, \infty)$ , and then average over the jump times of  $(J_t)_{t \geq 0}$  while keeping the jumps of  $X^{\natural}$  fixed. In this way we reduce the problem to the one for the discrete-time random walk treated in the proof of Theorem 1.6. For the first expression in (3.38) this *partial annealing* gives an upper bound, while for the second expression it is simply part of the averaging over  $\tilde{S}$ .

Define

$$F_N^{(1)}(X^{\natural}) := \int_0^{\infty} dt_1 \cdots \int_{t_{N-1}}^{\infty} dt_N \mathbb{P}(\tilde{S}_{t_1} = \tilde{S}'_{t_1}, \dots, \tilde{S}_{t_N} = \tilde{S}'_{t_N} \mid X^{\natural}), \quad F_N^{(2)} := \mathbb{E}[F_N^{(1)}(X^{\natural})], \quad (3.40)$$

together with the critical values

$$r_1^{\natural} := \lim_{N \rightarrow \infty} \frac{1}{N} \log F_N^{(1)}(X^{\natural}) \quad (X^{\natural} - a.s.), \quad r_2^{\natural} := \lim_{N \rightarrow \infty} \frac{1}{N} \log F_N^{(2)}. \quad (3.41)$$

Clearly,

$$\tilde{r}_1 \leq r_1^{\natural} \text{ and } \tilde{r}_2 = r_2^{\natural}, \quad (3.42)$$

which can be viewed as a result of “partial annealing”, and so it suffices to show that  $r_1^{\natural} < r_2^{\natural}$ .

To this end write out

$$\begin{aligned} & \mathbb{P}(\tilde{S}_{t_1} = \tilde{S}'_{t_1}, \dots, \tilde{S}_{t_N} = \tilde{S}'_{t_N} \mid X^{\natural}) \\ &= \sum_{0 \leq j_1 \leq \dots \leq j_N < \infty} \left( \prod_{i=1}^N e^{-(t_i - t_{i-1})} \frac{(t_i - t_{i-1})^{j_i - j_{i-1}}}{(j_i - j_{i-1})!} \right) \\ & \quad \sum_{0 \leq j'_1 \leq \dots \leq j'_N < \infty} \left( \prod_{i=1}^N e^{-(t_i - t_{i-1})} \frac{(t_i - t_{i-1})^{j'_i - j'_{i-1}}}{(j'_i - j'_{i-1})!} \right) \left[ \prod_{i=1}^N p_{j'_i - j'_{i-1}} \left( \sum_{k=1}^{j_i - j_{i-1}} X_{j_{i-1} + k}^{\natural} \right) \right]. \end{aligned} \quad (3.43)$$

Integrating over  $0 \leq t_1 \leq \dots \leq t_N < \infty$ , we obtain

$$\begin{aligned} F_N^{(1)}(X^{\natural}) &= \sum_{0 \leq j_1 \leq \dots \leq j_N < \infty} \sum_{0 \leq j'_1 \leq \dots \leq j'_N < \infty} \\ & \quad \prod_{i=1}^N \left[ 2^{-(j_i - j_{i-1}) - (j'_i - j'_{i-1}) - 1} \frac{[(j_i - j_{i-1}) + (j'_i - j'_{i-1})]!}{(j_i - j_{i-1})! (j'_i - j'_{i-1})!} p_{j'_i - j'_{i-1}} \left( \sum_{k=1}^{j_i - j_{i-1}} X_{j_{i-1} + k}^{\natural} \right) \right]. \end{aligned} \quad (3.44)$$

Abbreviating

$$\Theta_n(u) = \sum_{m=0}^{\infty} p^m(u) 2^{-n-m-1} \binom{n+m}{m}, \quad n \in \mathbb{N} \cup \{0\}, u \in \mathbb{Z}^d, \quad (3.45)$$

we may rewrite (3.44) as

$$F_N^{(1)}(X^\natural) = \sum_{0 \leq j_1 \leq \dots \leq j_N < \infty} \prod_{i=1}^N \Theta_{j_i - j_{i-1}} \left( \sum_{k=1}^{j_i - j_{i-1}} X_{j_{i-1} + k}^\natural \right). \quad (3.46)$$

This expression is similar in form as the first line of (3.8), except that the order of the  $j_i$ 's is not strict. However, defining

$$\widehat{F}_N^{(1)}(X^\natural) = \sum_{0 < j_1 < \dots < j_N < \infty} \prod_{i=1}^N \Theta_{j_i - j_{i-1}} \left( \sum_{k=1}^{j_i - j_{i-1}} X_{j_{i-1} + k}^\natural \right), \quad (3.47)$$

we have

$$F_N^{(1)}(X^\natural) = \sum_{M=0}^N \binom{N}{M} [\Theta_0(0)]^M \widehat{F}_{N-M}^{(1)}(X^\natural), \quad (3.48)$$

with the convention  $\widehat{F}_0^{(1)}(X^\natural) \equiv 1$ . Letting

$$r_1^\natural = \lim_{N \rightarrow \infty} \frac{1}{N} \log \widehat{F}_N^{(1)}(X^\natural), \quad X^\natural - a.s., \quad (3.49)$$

and recalling (3.41), we therefore have the relation

$$r_1^\natural = \log \left[ \Theta_0(0) + e^{\widehat{r}_1^\natural} \right], \quad (3.50)$$

and so it suffices to compute  $\widehat{r}_1^\natural$ .

Write

$$F_N^{(1)}(X^\natural) = \mathbb{E} \left[ \exp \left( N \int_{\widetilde{\mathbb{Z}}^d} (\pi_1 R_N)(dy) \log f^\natural(y) \right) \middle| X^\natural \right], \quad (3.51)$$

where  $f^\natural: \widetilde{\mathbb{Z}}^d \rightarrow [0, \infty)$  is defined by

$$f^\natural((x_1, \dots, x_n)) = \frac{\Theta_n(x_1 + \dots + x_n)}{p^{2\lfloor n/2 \rfloor}(0)} [2\bar{G}(0) - 1], \quad n \in \mathbb{N}, x_1, \dots, x_n \in \mathbb{Z}^d. \quad (3.52)$$

Equations (3.51–3.52) replace (3.8–3.9). We can now repeat the same argument as in (3.16–3.22), with the sole difference that  $f$  in (3.9) is replaced by  $f^\natural$  in (3.52), and this, combined with Lemma 3.3 below, yields the gap  $r_1^\natural < r_2^\natural$ .

We first check that  $f^\natural$  is bounded from above, which is necessary for the application of Varadhan's lemma. To that end, we insert the Fourier representation (3.14) into (3.45) to obtain

$$\Theta_n(u) = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} dk e^{-i(k \cdot u)} [2 - \widehat{p}(k)]^{-n-1}, \quad u \in \mathbb{Z}^d, \quad (3.53)$$

from which we see that  $\Theta_n(u) \leq \Theta_n(0)$ ,  $u \in \mathbb{Z}^d$ . Consequently,

$$f_n^\natural((x_1, \dots, x_n)) \leq \frac{\Theta_n(0)}{p^{2\lfloor n/2 \rfloor}(0)} [2\bar{G}(0) - 1], \quad n \in \mathbb{N}, x_1, \dots, x_n \in \mathbb{Z}^d. \quad (3.54)$$

Next we note that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left[ 2^{-(a+b)n-1} \binom{(a+b)n}{an} \right] \begin{cases} = 0, & \text{if } a = b, \\ < 0, & \text{if } a \neq b. \end{cases} \quad (3.55)$$

From (1.1), (3.45) and (3.55) it follows that  $\Theta_n(0)/p^{2\lfloor n/2 \rfloor}(0) \leq C < \infty$  for all  $n \in \mathbb{N}$ , so that  $f^\natural$  indeed is bounded from above.

Note that  $X^\natural$  is the discrete-time random walk with transition kernel  $p(\cdot, \cdot)$ . The key ingredient behind  $\widehat{r}_1^\natural < \widehat{r}_2^\natural$  is the analogue of Lemma 3.2, this time with  $Q^* = (q^*)^{\otimes \mathbb{N}}$  and  $q^*$  given by

$$q^*(x_1, \dots, x_n) = \frac{\Theta_n(x_1 + \dots + x_n)}{\frac{1}{2}G(0) - \Theta_0(0)} \prod_{k=1}^n p(x_k), \quad (3.56)$$

replacing (3.19). The proof is deferred to the end.

**Lemma 3.3.** *Assume (1.1). Let  $Q^* = (q^*)^{\otimes \mathbb{N}}$  with  $q^*$  as in (3.56). If  $m_{Q^*} < \infty$ , then  $I^{\text{que}}(Q^*) > I^{\text{ann}}(Q^*)$ .*

This shows that  $\widehat{r}_1^\natural < \widehat{r}_2^\natural$  via the same computation as in (3.21–3.23).

The analogue of (3.18) reads

$$\begin{aligned} Z^\natural &= \sum_{n \in \mathbb{N}} \sum_{x_1, \dots, x_n \in \mathbb{Z}^d} [\Theta_n(x_1 + \dots + x_n)] \prod_{k=1}^n p(x_k) \\ &= \sum_{n \in \mathbb{N}} \sum_{m=0}^{\infty} \left\{ \sum_{x_1, \dots, x_n \in \mathbb{Z}^d} p^m(x_1 + \dots + x_n) \prod_{k=1}^n p(x_k) \right\} 2^{-n-m-1} \binom{n+m}{m} \\ &= -\Theta_0(0) + \sum_{n, m=0}^{\infty} p^{n+m}(0) 2^{-n-m-1} \binom{n+m}{m} \\ &= -\Theta_0(0) + \frac{1}{2} \sum_{k=0}^{\infty} p^k(0) = -\Theta_0(0) + \frac{1}{2}G(0). \end{aligned} \quad (3.57)$$

Consequently,

$$\log \widetilde{z}_2 = e^{-\widetilde{r}_2} = e^{-r_2^\natural} = \frac{1}{\Theta_0(0) + e^{\widetilde{r}_2^\natural}} = \frac{1}{\Theta_0(0) + Z^\natural} = \frac{2}{G(0)}, \quad (3.58)$$

where we use (3.37), (3.39), (3.42), (3.50) and (3.57).

We close by proving Lemma 3.3.

*Proof.* We must adapt the proof in Section 3.2 to the fact that  $q^*$  has a slightly different form, namely,  $p^n(x_1 + \dots + x_n)$  is replaced by  $\Theta_n(x_1 + \dots + x_n)$ , which averages transition kernels. The computations are straightforward and are left to the reader. The analogues of (3.24) and (3.26) are

$$\begin{aligned} q^*((\mathbb{Z}^d)^n) &= \frac{1}{\frac{1}{2}G(0) - \Theta_0(0)} \sum_{m=0}^{\infty} p^{n+m}(0) 2^{-n-m-1} \binom{n+m}{m}, \\ m_{Q^*} &= \sum_{n \in \mathbb{N}} n q^*((\mathbb{Z}^d)^n) = \frac{1}{4} \frac{1}{\frac{1}{2}G(0) - \Theta_0(0)} \sum_{k=0}^{\infty} k p^k(0), \end{aligned} \quad (3.59)$$

while the analogues of (3.31–3.32) are

$$q^*(x_1 = a) = \frac{p(a)}{\frac{1}{2}G(0) - \Theta_0(0)} \frac{1}{2} \sum_{k=0}^{\infty} p^k(a)[1 - 2^{-k-1}] = \frac{1}{2}p(a) \frac{G(a) - \Theta_0(a)}{\frac{1}{2}G(0) - \Theta_0(0)},$$

$$q^*(\{(a, b)\} \times (\mathbb{Z}^d)^{n-2}) = \frac{p(a)p(b)}{\frac{1}{2}G(0) - \Theta_0(0)} \sum_{m=0}^{\infty} p^{n-2+m}(a+b) 2^{-n-m-1} \binom{n+m}{m}. \quad (3.60)$$

Recalling (3.30), we find

$$\Psi_{Q^*}(a, -a) - p(a)^2 > 0, \quad (3.61)$$

implying that  $\Psi_{Q^*} \neq \nu^{\otimes \mathbb{N}}$  (recall (3.2)), and hence  $H(\Psi_{Q^*} | \nu^{\otimes \mathbb{N}}) > 0$ , implying the claim.  $\square$

## 4 Proof of Theorem 1.2

This section uses techniques from [6]. The proof of Theorem 1.2 is based on two approximation lemmas, which are stated in Section 4.1. The proof of these lemmas is given in Sections 4.2–4.3.

### 4.1 Two approximation lemmas

Return to the setting in Section 2. For  $Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})$ , let  $H(Q)$  denote the specific entropy of  $Q$ . Write  $h(\cdot | \cdot)$  and  $h(\cdot)$  to denote relative entropy, respectively, entropy. Write, and recall from (1.9),

$$\begin{aligned} \mathcal{P}^{\text{erg}}(\tilde{E}^{\mathbb{N}}) &= \{Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}}) : Q \text{ is shift-ergodic}\}, \\ \mathcal{P}^{\text{erg,fin}}(\tilde{E}^{\mathbb{N}}) &= \{Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}}) : Q \text{ is shift-ergodic, } m_Q < \infty\}. \end{aligned} \quad (4.1)$$

**Lemma 4.1.** *Let  $g : \tilde{E} \rightarrow \mathbb{R}$  be such that*

$$\liminf_{k \rightarrow \infty} \frac{g(X|_{(0,k]})}{\log k} \geq 0 \text{ for } \nu^{\otimes \mathbb{N}} \text{ - a.s. all } X \quad \text{with } X|_{(0,k]} := (X_1, \dots, X_k). \quad (4.2)$$

*Let  $Q \in \mathcal{P}^{\text{erg,fin}}(\tilde{E}^{\mathbb{N}})$  be such that  $H(Q) < \infty$  and  $G(Q) := \int_{\tilde{E}} (\pi_1 Q)(dy) g(y) \in \mathbb{R}$ . Then*

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} \left[ \exp \left( N \int_{\tilde{E}} (\pi_1 R_N)(dy) g(y) \right) \middle| X \right] \geq G(Q) - I^{\text{que}}(Q) \text{ for } \nu^{\otimes \mathbb{N}} \text{ - a.s. all } X. \quad (4.3)$$

**Lemma 4.2.** *Let  $g : \tilde{E} \rightarrow \mathbb{R}$  be such that*

$$\sup_{k \in \mathbb{N}} \int_{E^k} |g((x_1, \dots, x_k))| \nu^{\otimes k}(dx_1, \dots, dx_k) < \infty. \quad (4.4)$$

*Let  $Q \in \mathcal{P}^{\text{erg}}(\tilde{E}^{\mathbb{N}})$  be such that  $I^{\text{que}}(Q) < \infty$  and  $G(Q) \in \mathbb{R}$ . Then there exists a sequence  $(Q_n)_{n \in \mathbb{N}}$  in  $\mathcal{P}^{\text{erg,fin}}(\tilde{E}^{\mathbb{N}})$  such that*

$$\liminf_{n \rightarrow \infty} [G(Q_n) - I^{\text{que}}(Q_n)] \geq G(Q) - I^{\text{que}}(Q). \quad (4.5)$$

*Moreover, if  $E$  is countable and  $\nu$  satisfies*

$$\forall \mu \in \mathcal{P}(E): \quad h(\mu | \nu) < \infty \implies h(\mu) < \infty, \quad (4.6)$$

*then  $(Q_n)_{n \in \mathbb{N}}$  can be chosen such that  $H(Q_n) < \infty$  for all  $n \in \mathbb{N}$ .*

Lemma 4.2 immediately yields the following.

**Corollary 4.3.** *If  $g$  satisfies (4.4) and  $\nu$  satisfies (4.6), then*

$$\sup_{Q \in \mathcal{P}^{\text{inv}}(\tilde{E}^{\mathbb{N}})} \left\{ \int_{\tilde{E}} (\pi_1 Q)(dy) g(y) - I^{\text{que}}(Q) \right\} = \sup_{\substack{Q \in \mathcal{P}^{\text{erg,fin}}(\tilde{E}^{\mathbb{N}}) \\ H(Q) < \infty}} \left\{ \int_{\tilde{E}} (\pi_1 Q)(dy) g(y) - I^{\text{que}}(Q) \right\}. \quad (4.7)$$

With Corollary 4.3, we can now complete the proof of Theorem 1.2.

*Proof.* Return to the setting in Section 3.1. In Lemma 4.1, pick  $g = \log f$  with  $f$  as defined in (3.9). Then (1.11) is the same as (4.2), and so it follows that

$$\begin{aligned} & \liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} \left[ \exp \left( N \int_{\tilde{\mathbb{Z}}^d} (\pi_1 R_N)(dy) \log f(y) \right) \mid X \right] \\ & \geq \sup_{\substack{Q \in \mathcal{P}^{\text{erg,fin}}(\tilde{\mathbb{Z}}^d)^{\mathbb{N}} \\ H(Q) < \infty}} \left\{ \int_{\tilde{\mathbb{Z}}^d} (\pi_1 Q)(dy) \log f(y) - I^{\text{que}}(Q) \right\}, \end{aligned} \quad (4.8)$$

where the condition that the first term under the supremum be finite is redundant because  $g = \log f$  is bounded from above (recall (3.13)). Recalling (3.10) and (3.20), we thus see that

$$r_1 \geq \sup_{\substack{Q \in \mathcal{P}^{\text{erg,fin}}(\tilde{\mathbb{Z}}^d)^{\mathbb{N}} \\ H(Q) < \infty}} \left\{ \int_{\tilde{\mathbb{Z}}^d} (\pi_1 Q)(dy) \log f(y) - I^{\text{que}}(Q) \right\}. \quad (4.9)$$

The right-hand side of (4.9) is the same as that of (1.13), except for the restriction that  $H(Q) < \infty$ . To remove this restriction, we use Corollary 4.3. First note that, by (1.12), condition (4.4) in Lemma 4.2 is fulfilled for  $g = \log f$ . Next note that, by (1.10) and Remark 4.4 below, condition (4.6) in Lemma 4.2 is fulfilled for  $\nu = p$ . Therefore Corollary 4.3 implies that  $r_1$  equals the right-hand side of (1.13), and that the suprema in (1.13) and (1.6) agree.

Equality (1.14) follows easily from the fact that the maximiser of the right-hand side of (1.7) is given by  $Q^* = (q^*)^{\otimes \mathbb{N}}$  with  $q^*$  defined in (3.19), as discussed after Lemma 3.1: If  $m_{Q^*} < \infty$ , then we are done, otherwise we approximate  $Q^*$  via truncation.  $\square$

**Remark 4.4.** *Every  $\nu \in \mathcal{P}(\mathbb{Z}^d)$  for which  $\sum_{x \in \mathbb{Z}^d} \|x\|^\delta \nu(x) < \infty$  for some  $\delta > 0$  satisfies (4.6).*

*Proof.* Let  $\mu \in \mathcal{P}(\mathbb{Z}^d)$ , and let  $\pi_i$ ,  $i = 1, \dots, d$ , be the projection onto the  $i$ -th coordinate. Since  $h(\pi_i \mu \mid \pi_i \nu) \leq h(\mu \mid \nu)$  for  $i = 1, \dots, d$  and  $h(\mu) \leq h(\pi_1 \mu) + \dots + h(\pi_d \mu)$ , it suffices to check the claim for  $d = 1$ .

Let  $\mu \in \mathcal{P}(\mathbb{Z})$  be such that  $h(\mu \mid \nu) < \infty$ . Then

$$\begin{aligned} \sum_{x \in \mathbb{Z}} \mu(x) \log(e + |x|) &= \sum_{\substack{x \in \mathbb{Z} \\ \mu(x) \geq (e+|x|)^{\delta/2} \nu(x)}} \mu(x) \log(e + |x|) + \sum_{\substack{x \in \mathbb{Z} \\ \mu(x) < (e+|x|)^{\delta/2} \nu(x)}} \mu(x) \log(e + |x|) \\ &\leq \frac{2}{\delta} \sum_{\substack{x \in \mathbb{Z} \\ \mu(x) \geq \nu(x)}} \mu(x) \log \left( \frac{\mu(x)}{\nu(x)} \right) + \sum_{x \in \mathbb{Z}} \nu(x) (e + |x|)^{\delta/2} \log(e + |x|) \\ &\leq \frac{2}{\delta} h(\mu \mid \nu) + C \sum_{x \in \mathbb{Z}} \nu(x) |x|^\delta < \infty \end{aligned} \quad (4.10)$$

for some  $C \in (0, \infty)$ . Therefore

$$\begin{aligned} h(\mu) &= \sum_{x \in \mathbb{Z}} \mu(x) \log \left( \frac{1}{\mu(x)} \right) = \sum_{\substack{x \in \mathbb{Z} \\ \mu(x) \leq (e+|x|)^{-2}}} \mu(x) \log \left( \frac{1}{\mu(x)} \right) + \sum_{\substack{x \in \mathbb{Z} \\ \mu(x) > (e+|x|)^{-2}}} \mu(x) \log \left( \frac{1}{\mu(x)} \right) \\ &\leq \sum_{x \in \mathbb{Z}} \frac{2 \log(e + |x|)}{(e + |x|)^2} + 2 \sum_{x \in \mathbb{Z}} \mu(x) \log(e + |x|) < \infty, \end{aligned} \quad (4.11)$$

where the last inequality uses (4.10).  $\square$

## 4.2 Proof of Lemma 4.1

*Proof.* The idea is to make the first word so long that it ends in front of the first region in  $X$  that looks like the concatenation of  $N$  words drawn from  $Q$ , and after that cut  $N$  “ $Q$ -typical” words from this region. Condition (4.2) ensures that the contribution of the first word to the left-hand side of (4.3) is negligible on the exponential scale.

To formalise this idea, we borrow some techniques from [6], Section 3.1. Let  $H(\Psi_Q)$  denote the specific entropy of  $\Psi_Q$  (defined in (2.8)), and  $H_{\tau|\kappa}(Q)$  the “conditional specific entropy of word lengths under the law  $Q$  given the concatenation” (defined in [6], Lemma 1.7). We need the relation

$$H(Q | q_{\rho, \nu}^{\otimes \mathbb{N}}) = m_Q H(\Psi_Q | \nu^{\otimes \mathbb{N}}) - H_{\tau|\kappa}(Q) - \mathbb{E}_Q [\log \rho(\tau_1)]. \quad (4.12)$$

First, we note that  $H(Q) < \infty$  and  $m_Q < \infty$  imply that  $H(\Psi_Q) < \infty$  and  $H_{\tau|\kappa}(Q) < \infty$  (see [6], Lemma 1.7). Next, we fix  $\varepsilon > 0$ . Following the arguments in [6], Section 3.1, we see that for all  $N$  large enough we can find a finite set  $\mathcal{A} = \mathcal{A}(Q, \varepsilon, N) \subset \tilde{E}^N$  of “ $Q$ -typical sentences” such that, for all  $z = (y^{(1)}, \dots, y^{(N)}) \in \mathcal{A}$ , the following hold:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \log \rho(|y^{(i)}|) &\in [\mathbb{E}_Q [\log \rho(\tau_1)] - \varepsilon, \mathbb{E}_Q [\log \rho(\tau_1)] + \varepsilon], \\ \frac{1}{N} \log |\{z' \in \mathcal{A} : \kappa(z') = \kappa(z)\}| &\in [H_{\tau|\kappa}(Q) - \varepsilon, H_{\tau|\kappa}(Q) + \varepsilon], \\ \frac{1}{N} \sum_{i=1}^N g(y^{(i)}) &\in [G(Q) - \varepsilon, G(Q) + \varepsilon]. \end{aligned} \quad (4.13)$$

Put  $\mathcal{B} := \kappa(\mathcal{A}) \subset \tilde{E}$ . We can choose  $\mathcal{A}$  in such a way that the elements of  $\mathcal{B}$  have a length in  $[N(m_Q - \varepsilon), N(m_Q + \varepsilon)]$ . Moreover, we have

$$\mathbb{P}(X \text{ begins with an element of } \mathcal{B}) \geq \exp[-N\chi(Q)], \quad (4.14)$$

where we abbreviate

$$\chi(Q) := m_Q H(\Psi_Q | \nu^{\otimes \mathbb{N}}) + \varepsilon. \quad (4.15)$$

Put

$$\tau_N := \min \{i \in \mathbb{N} : \theta^i X \text{ begins with an element of } \mathcal{B}\}. \quad (4.16)$$

Then, by (4.14) and the Shannon-McMillan-Breiman theorem, we have

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \tau_N \leq \chi(Q). \quad (4.17)$$

Indeed, for each  $N$ , coarse-grain  $X$  into blocks of length  $L_N := \lfloor N(m_Q + \varepsilon) \rfloor$ . For  $i \in \mathbb{N} \cup \{0\}$ , let  $A_{N,i}$  be the event that  $\theta^{iL_N} X$  begins with an element of  $\mathcal{B}$ . Then, for any  $\delta > 0$ ,

$$\left\{ \tau_N > \exp[N(\chi(Q) + \delta)] \right\} \subset \bigcap_{i=1}^{\exp[N(\chi(Q) + \delta)]/L_N} A_{N,i}^c, \quad (4.18)$$

and hence

$$\begin{aligned} \mathbb{P}\left(\tau_N > \exp[N(\chi(Q) + \delta)]\right) &\leq \left(1 - \exp[-N\chi(Q)]\right)^{\exp[N(\chi(Q) + \delta)]/L_N} \\ &= \left\{ \left(1 - \exp[-N\chi(Q)]\right)^{\exp[N\chi(Q)]} \right\}^{e^{\delta N}/L_N} \leq \exp[-e^{\delta N}/L_N], \end{aligned} \quad (4.19)$$

which is summable in  $N$ . Thus,  $\limsup_{N \rightarrow \infty} \frac{1}{N} \log \tau_N \leq \chi(Q) + \delta$  by the first Borel-Cantelli lemma. Now let  $\delta \downarrow 0$ , to get (4.17).

Next, note that

$$\begin{aligned} &\mathbb{E} \left[ \exp \left( (N+1) \int_{\tilde{E}} (\pi_1 R_{N+1})(dy) g(y) \right) \middle| X \right] \\ &= \sum_{0 < j_1 < \dots < j_{N+1}} \prod_{i=1}^{N+1} \rho(j_i - j_{i-1}) \exp \left( \sum_{i=1}^{N+1} g(X|_{(j_{i-1}, j_i]}) \right) \\ &\geq \rho(\tau_N) \exp[g(X|_{(0, \tau_N]})] \sum_* \prod_{i=2}^{N+1} \rho(j_i - j_{i-1}) \exp \left( \sum_{i=2}^{N+1} g(X|_{(j_{i-1}, j_i]}) \right), \end{aligned} \quad (4.20)$$

where  $\sum_*$  in the last line refers to all  $(j_1, \dots, j_{N+1})$  such that  $j_1 := \tau_N < j_2 < \dots < j_{N+1}$  and  $(X|_{(j_1, j_2]}, \dots, X|_{(j_N, j_{N+1}]}) \in \mathcal{A}$ . Combining (2.1), (4.13), (4.17) and (4.20), we obtain that  $X$ -a.s.

$$\begin{aligned} &\liminf_{N \rightarrow \infty} \frac{1}{N+1} \log \mathbb{E} \left[ \exp \left[ (N+1) \int_{\tilde{E}} (\pi_1 R_{N+1})(dy) g(y) \right] \middle| X \right] \\ &\geq -\alpha \chi(Q) + \liminf_{N \rightarrow \infty} \frac{g(X|_{(0, \tau_N]})}{N} + H_{\tau|k}(Q) + \mathbb{E}_Q [\log \rho(\tau_1)] + G(Q) - 3\varepsilon. \end{aligned} \quad (4.21)$$

By Assumption (4.2),  $\liminf_{N \rightarrow \infty} N^{-1} g(X|_{(0, \tau_N]}) \geq 0$ , and so (4.21) yields that  $X$ -a.s.

$$\begin{aligned} &\liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} \left[ \exp \left( N \int_{\tilde{E}} (\pi_1 R_N)(dy) g(y) \right) \middle| X \right] \\ &\geq G(Q) - \alpha m_Q H(\Psi_Q | \nu^{\otimes \mathbb{N}}) + H_{\tau|k}(Q) + \mathbb{E}_Q [\log \rho(\tau_1)] - (3 + \alpha)\varepsilon \\ &= G(Q) - I^{\text{que}}(Q) - (3 + \alpha)\varepsilon, \end{aligned} \quad (4.22)$$

where we use (2.13–2.14), (4.12) and (4.15). Finally, let  $\varepsilon \downarrow 0$  to get the claim.  $\square$

### 4.3 Proof of Lemma 4.2

*Proof.* Without loss of generality we may assume that  $m_Q = \infty$ , for otherwise  $Q_n \equiv Q$  satisfies (4.5). The idea is to use a variation on the truncation construction in [6], Section 3. For a given truncation

level  $\text{tr} \in \mathbb{N}$ , let  $Q_{\text{tr}}^v$  be the law obtained from  $Q$  by replacing all words of length  $\geq \text{tr}$  by words of length  $\text{tr}$  whose letters are drawn independently from  $v$ . Formally, if  $Y = (Y^{(i)})_{i \in \mathbb{N}}$  has law  $Q$  and  $\tilde{Y} = (\tilde{Y}^{(i)})_{i \in \mathbb{N}}$  has law  $(v^{\otimes \text{tr}})^{\otimes \mathbb{N}}$  and is independent of  $Y$ , then  $\bar{Y} = (\bar{Y}^{(i)})_{i \in \mathbb{N}}$  defined by

$$\bar{Y}^{(i)} := \begin{cases} Y^{(i)}, & \text{if } |Y^{(i)}| < \text{tr}, \\ \tilde{Y}^{(i)}, & \text{if } |Y^{(i)}| \geq \text{tr}, \end{cases} \quad (4.23)$$

has law  $Q_{\text{tr}}^v$ .

**Lemma 4.5.** For every  $Q \in \mathcal{P}^{\text{inv,erg}}(\tilde{E}^{\mathbb{N}})$  such that  $I^{\text{que}}(Q) < \infty$  and every  $\text{tr} \in \mathbb{N}$ ,

$$\begin{aligned} H(Q_{\text{tr}}^v | q_{\rho,v}^{\otimes \mathbb{N}}) &\leq H([Q]_{\text{tr}} | q_{\rho,v}^{\otimes \mathbb{N}}), \\ H(\Psi_{Q_{\text{tr}}^v} | v^{\otimes \mathbb{N}}) &\leq H(\Psi_{[Q]_{\text{tr}}} | v^{\otimes \mathbb{N}}). \end{aligned} \quad (4.24)$$

*Proof.* The intuition is that under  $Q_{\text{tr}}^v$  all words of length  $\text{tr}$  have the same content as under  $q_{\rho,v}^{\otimes \mathbb{N}}$ , while under  $[Q]_{\text{tr}}$  they do not. The proof is straightforward but lengthy, and is deferred to Appendix A.  $\square$

Using (4.24) and noting that  $m_{Q_{\text{tr}}^v} = m_{[Q]_{\text{tr}}} < \infty$ , we obtain (recall (2.13–2.14))

$$\limsup_{\text{tr} \rightarrow \infty} I^{\text{que}}(Q_{\text{tr}}^v) \leq I^{\text{que}}(Q). \quad (4.25)$$

On the other hand, we have

$$\begin{aligned} &\int_{\tilde{E}} (\pi_1 Q_{\text{tr}}^v)(dy) g(y) \\ &= \int_{\tilde{E}} (\pi_1 Q)(dy) \mathbb{1}_{\{|y| < \text{tr}\}} g(y) + Q(\tau_1 \geq \text{tr}) \int_{E^{\text{tr}}} v^{\otimes \text{tr}}(dx_1, \dots, dx_{\text{tr}}) g((x_1, \dots, x_{\text{tr}})) \\ &\xrightarrow{\text{tr} \rightarrow \infty} G(Q) = \int_{\tilde{E}} (\pi_1 Q)(dy) g(y), \end{aligned} \quad (4.26)$$

where we use dominated convergence for the first summand and condition (4.4) for the second summand. Combining (4.25–4.26), we see that we can choose  $\text{tr} = \text{tr}(n)$  such that (4.5) holds for  $Q_n = Q_{\text{tr}(n)}^v$ .

It remains to verify that, under condition (4.6),  $H(Q_{\text{tr}}^v) < \infty$  for all  $\text{tr} \in \mathbb{N}$ . Since  $H(Q_{\text{tr}}^v) \leq h(\pi_1 Q_{\text{tr}}^v)$ , it suffices to verify that  $h(\pi_1 Q_{\text{tr}}^v) < \infty$  for all  $\text{tr} \in \mathbb{N}$ . To prove the latter, note that (we write  $\mathcal{L}_{Q_{\text{tr}}^v}(\tau_1)$  to denote the law of  $\tau_1$  under  $Q_{\text{tr}}^v$ , etc.)

$$\begin{aligned} h(\pi_1 Q_{\text{tr}}^v) &= h(\mathcal{L}_{Q_{\text{tr}}^v}(\tau_1)) + \sum_{\ell=1}^{\text{tr}} Q_{\text{tr}}^v(\tau_1 = \ell) h(\mathcal{L}_{Q_{\text{tr}}^v}(Y^{(1)} | \tau_1 = \ell)) \\ &\leq \log \text{tr} + \sum_{\ell=1}^{\text{tr}-1} \sum_{k=1}^{\ell} h(\mathcal{L}_{Q_{\text{tr}}^v}(Y_k^{(1)} | \tau_1 = \ell)) + \text{tr} h(v). \end{aligned} \quad (4.27)$$

Since  $h(\pi_1 Q | q_{\rho,v}^{\otimes \mathbb{N}}) \leq H(Q | q_{\rho,v}^{\otimes \mathbb{N}}) = I^{\text{ann}}(Q) \leq I^{\text{que}}(Q) < \infty$ , we have

$$h(\pi_1 Q | q_{\rho,v}^{\otimes \mathbb{N}}) = h(\mathcal{L}_Q(\tau_1) | \rho) + \sum_{\ell=1}^{\infty} Q(\tau_1 = \ell) h(\mathcal{L}_Q(Y^{(1)} | \tau_1 = \ell) | v^{\otimes \ell}) < \infty. \quad (4.28)$$



Moreover, for all  $\ell < \text{tr}$  and  $k = 1, \dots, \ell$ ,

$$h\left(\mathcal{L}_{Q_{\text{tr}}^v}(Y_k^{(1)}|\tau_1 = \ell) \mid v\right) \leq h\left(\mathcal{L}_{Q_{\text{tr}}^v}(Y^{(1)}|\tau_1 = \ell) \mid v^{\otimes \ell}\right) = h\left(\mathcal{L}_Q(Y^{(1)}|\tau_1 = \ell) \mid v^{\otimes \ell}\right). \quad (4.29)$$

Combine (4.28–4.29) with (4.6) to conclude that all the summands in (4.27) are finite.  $\square$

## 5 Proof of Theorems 1.4 and 1.5

*Proof of Theorem 1.4.* Let  $q \in \mathcal{P}(\widehat{\mathbb{Z}^d})$  be given by

$$q(x_1, \dots, x_n) := \bar{\rho}(n) v(x_1) \cdots v(x_n), \quad n \in \mathbb{N}, x_1, \dots, x_n \in \mathbb{Z}^d, \quad (5.1)$$

for some  $\bar{\rho} \in \mathcal{P}(\mathbb{N})$  with  $\sum_{n \in \mathbb{N}} n \bar{\rho}(n) < \infty$ , and let  $Q = q^{\otimes \mathbb{N}}$ . Then  $Q$  is ergodic,  $m_Q < \infty$ , and (recall (2.4))

$$I^{\text{que}}(Q) = H(q^{\otimes \mathbb{N}} \mid (q_{\rho, v})^{\otimes \mathbb{N}}) = h(\bar{\rho} \mid \rho) \quad (5.2)$$

because  $\Psi_Q = v^{\otimes \mathbb{N}}$ . Now pick  $\text{tr} \in \mathbb{N}$ ,  $\bar{\rho} = [\rho^*]_{\text{tr}}$  with  $\rho^*$  given by

$$\rho^*(n) := \frac{1}{Z} \exp[-h(p^n)], \quad n \in \mathbb{N}, \quad Z := \sum_{n \in \mathbb{N}} \exp[-h(p^n)], \quad (5.3)$$

$v(\cdot) = p(\cdot)$ , and compute (recall (3.2) and (3.9))

$$\begin{aligned} \int_{\widehat{\mathbb{Z}^d}} (\pi_1 Q)(dy) \log f(y) &= \int_{\widehat{\mathbb{Z}^d}} q(dy) \log f(y) \\ &= \sum_{n \in \mathbb{N}} \sum_{x_1, \dots, x_n \in \mathbb{Z}^d} \bar{\rho}(n) p(x_1) \cdots p(x_n) \log \left( \frac{p^n(x_1 + \cdots + x_n)}{\rho(n)} \right) \\ &= \sum_{n \in \mathbb{N}} \bar{\rho}(n) [-\log \rho(n) - h(p^n)] \\ &= \log Z + \sum_{n \in \mathbb{N}} \bar{\rho}(n) \log \left( \frac{\rho^*(n)}{\rho(n)} \right) \\ &= \log Z + h(\bar{\rho} \mid \rho) - h(\bar{\rho} \mid \rho^*). \end{aligned} \quad (5.4)$$

Then (1.13), (5.2) and (5.4) give the lower bound

$$r_1 \geq \log Z - h(\bar{\rho} \mid \rho^*). \quad (5.5)$$

Let  $\text{tr} \rightarrow \infty$ , to obtain  $r_1 \geq \log Z$ , which proves the claim (recall that  $z_1 = 1 + \exp[-r_1]$ ).

It is easy to see that the choice in (5.3) is optimal in the class of  $q$ 's of the form (5.1) with  $v(\cdot) = p(\cdot)$ . By using (3.15), we see that  $h(p^{2^n}) \geq -\log p^{2^n}(0)$  and  $h(p^{2^{n+1}}) \geq -\log p^{2^{n+1}}(0)$ . Hence  $Z < \infty$  by the transience of  $p(\cdot, \cdot)$ .  $\square$

*Proof of Theorem 1.5.* The claim follows from the representations (1.13–1.14) in Theorem 1.2, and the fact that  $I^{\text{que}} = I^{\text{ann}}$  when  $\alpha = 1$ .  $\square$

## 6 Examples of random walks satisfying assumptions (1.10–1.12)

In this section we exhibit classes of random walks for which (1.10–1.12) hold.

1. Let  $S$  be an irreducible symmetric random walk on  $\mathbb{Z}^d$  with  $E[\|S_1\|^3] < \infty$ . Then standard cumulant expansion techniques taken from Bhattacharya and Ranga Rao [2] can be used to show that for every  $C_1 \in (0, \infty)$  there is a  $C_2 \in (0, \infty)$  such that

$$p^n(x) = \frac{c}{n^{d/2}} \exp \left[ -\frac{1}{2n} (x, \Sigma^{-1}x) \right] \left( 1 + O \left( \frac{(\log n)^{C_2}}{n^{1/2}} \right) \right), \quad (6.1)$$

$$n \rightarrow \infty, \|x\| \leq \sqrt{C_1 n \log \log n}, p^n(x) > 0,$$

where  $\Sigma$  is the covariance matrix of  $S_1$  (which is assumed to be non-degenerate), and  $c$  is a constant that depends on  $p(\cdot)$ . The restriction  $p^n(x) > 0$  is necessary: e.g. for simple random walk  $x$  and  $n$  in (6.1) must have the same parity. The Hartman-Wintner law of the iterated logarithm (see e.g. Kallenberg [24], Corollary 14.8), which only requires  $S_1$  to have mean zero and finite variance, says that

$$\limsup_{n \rightarrow \infty} \frac{|(S_n)^i|}{\sqrt{2 \Sigma_{ii} n \log \log n}} = 1 \quad \text{a.s.}, \quad i = 1, \dots, d, \quad (6.2)$$

where  $(S_n)^i$  is the  $i$ -th component of  $S_n$ . Using  $\|S_n\| \leq \sqrt{d} \max_{1 \leq i \leq d} |(S_n)^i|$ , we obtain that there is a  $C_3 \in (0, \infty)$  such that

$$\limsup_{n \rightarrow \infty} \frac{\|S_n\|}{\sqrt{n \log \log n}} \leq C_3 \quad S - \text{a.s.} \quad (6.3)$$

Combining and (6.1) and (6.3), we find that there is a  $C_4 \in (0, \infty)$  such that

$$\log [ p^n(S_n) / p^{2\lfloor n/2 \rfloor}(0) ] \geq -C_4 \|S_n\|^2 / n \quad \forall n \in \mathbb{N} \quad S - \text{a.s.} \quad (6.4)$$

Combining (6.3) and (6.4), we get (1.11).

To get (1.12), we argue as follows. Note that  $E(S_1) = 0$  and  $E(\|S_1\|^2) < \infty$ . For  $n \in \mathbb{N}$ , we have

$$\sum_{x \in \mathbb{Z}^d} p^n(x) \log [ p^n(x) / p^{2\lfloor n/2 \rfloor}(0) ] =: \Sigma_1(n) + \Sigma_2(n), \quad (6.5)$$

where the sums run over, respectively,

$$I_1(n) := \{x \in \mathbb{Z}^d : p^n(x) / p^{2\lfloor n/2 \rfloor}(0) \geq \exp[-n^{-1}\|x\|^2 - 1]\},$$

$$I_2(n) := \{x \in \mathbb{Z}^d : p^n(x) / p^{2\lfloor n/2 \rfloor}(0) < \exp[-n^{-1}\|x\|^2 - 1]\}. \quad (6.6)$$

We have

$$\Sigma_1(n) \geq \sum_{x \in \mathbb{Z}^d} p^n(x) [-n^{-1}\|x\|^2 - 1] = -E(\|S_1\|^2) - 1. \quad (6.7)$$

Since  $u \mapsto u \log u$  is non-increasing on the interval  $[0, e)$ , we also have

$$\Sigma_2(n) \geq \sum_{x \in \mathbb{Z}^d} \{p^{2\lfloor n/2 \rfloor}(0) \exp[-n^{-1}\|x\|^2 - 1]\} [-n^{-1}\|x\|^2 - 1] \geq -p^{2\lfloor n/2 \rfloor}(0) C_5 n^{d/2} \quad (6.8)$$

for some  $C_5 \in (0, \infty)$ . By the local central limit theorem, we have  $p^{2\lfloor n/2 \rfloor}(0) \sim C_6 n^{-d/2}$  as  $n \rightarrow \infty$  for some  $C_6 \in (0, \infty)$ . Hence  $\Sigma_1(n) + \Sigma_2(n)$  is bounded away from  $-\infty$  uniformly in  $n \in \mathbb{N}$ , which proves (1.12).

**2.** Let  $S$  be a symmetric random walk on  $\mathbb{Z}$  that is in the normal domain of attraction of a symmetric stable law with index  $a \in (0, 1)$  and suppose that the one-step distribution is regularly varying, i.e.,  $P(S_1 = x) = [1 + o(1)] C |x|^{-1-a}$ ,  $|x| \rightarrow \infty$  for some  $C \in (0, \infty)$ . Then, as shown e.g. in Chover [13] and Heyde [20],

$$|S_n| \leq n^{1/a} (\log n)^{1/a+o(1)} \quad a.s. \quad n \rightarrow \infty. \quad (6.9)$$

$p$  satisfies (1.1) with  $\alpha = 1/a \in (1, \infty)$ , the standard local limit theorem gives (see e.g. Ibragimov and Linnik [22], Theorem 4.2.1)

$$p^n(x) = [1 + o(1)] n^{-1/a} f(xn^{-1/a}), \quad |x|/n^{1/a} = O(1), \quad (6.10)$$

with  $f$  the density of the stable law. The remaining region was analysed in Doney [15], Theorem A, namely,

$$p^n(x) = [1 + o(1)] C n |x|^{-1-a}, \quad |x|/n^{1/a} \rightarrow \infty. \quad (6.11)$$

In fact, the proof of (6.11) shows that for  $K$  sufficiently large there exist  $c \in (0, \infty)$  and  $n_0 \in \mathbb{N}$  such that

$$c^{-1} \leq \frac{p^n(x)}{n |x|^{-1-a}} \leq c, \quad n \geq n_0, |x| \geq Kn^{1/a}. \quad (6.12)$$

Combining (6.9–6.11), we get

$$\log [ p^n(S_n)/p^{2\lfloor n/2 \rfloor}(0) ] \geq [-(1+a)/a + o(1)] \log \log n \quad a.s., \quad (6.13)$$

which proves (1.11).

To get (1.12), we argue as follows. Pick  $K$  and  $c$  such that (6.12) holds. Obviously, it suffices to check (1.12) with the infimum over  $\mathbb{N}$  restricted to  $n \geq n_0$ . Because  $f$  is uniformly positive and bounded on  $[-K, K]$ , (6.11) gives

$$\inf_{n \geq n_0} \sum_{|x| \leq Kn^{1/a}} p^n(x) \log [ p^n(x)/p^{2\lfloor n/2 \rfloor}(0) ] \geq \log \left( \inf_{y \in [-K, K]} f(y)/2 \right) > -\infty. \quad (6.14)$$

Applying (6.10) to  $p^{2\lfloor n/2 \rfloor}(0)$  and (6.11) to  $p^n(x)$  we obtain

$$\sum_{|x| > Kn^{1/a}} p^n(x) \log [ p^n(x)/p^{2\lfloor n/2 \rfloor}(0) ] \geq -c_1 \sum_{|x| > Kn^{1/a}} \frac{1}{n^{1/a}} (|x|/n^{1/a})^{-1-a} (1+a) \log (c_2 |x|/n^{1/a}) \quad (6.15)$$

for some  $c_1, c_2 \in (0, \infty)$ . The right-hand side is an approximating Riemann sum for the integral

$$-2c_1(1+a) \int_K^\infty dy y^{-1-a} \log(c_2 y) > -\infty. \quad (6.16)$$

**3.** For an example similar to **2.** leading to  $\alpha = 1$  in (1.1), let  $p$  be a symmetric transition probability on  $\mathbb{Z}$  satisfying

$$p(x) \sim L(|x|) |x|^{-2} \quad \text{for } |x| \rightarrow \infty \quad (6.17)$$

where  $L$  is a suitable slowly varying function on  $[0, \infty)$ . Let  $\tilde{L}$  be a conjugate slowly varying function of  $1/L$ , i.e.,  $\tilde{L}$  is (up to asymptotic equivalence) determined by the requirement that  $\lim_{t \rightarrow \infty} L(t\tilde{L}(t))/\tilde{L}(t) = 1$  (see e.g. Seneta [26], Section 1.6). Transience of a  $p$ -random walk  $S = (S_n)_{n \in \mathbb{N}_0}$  is equivalent to the requirement  $\sum_{n \in \mathbb{N}} 1/(n\tilde{L}(n)) < \infty$ , as can be seen from (6.20) below. Note that in many cases, e.g. for  $L(x) = (\log(1 \vee x))^b$ ,  $x \in (0, \infty)$ , for some  $b > 1$ , we can choose  $\tilde{L} = L$ . Put  $b_n := n\tilde{L}(n)$ ,  $n \in \mathbb{N}$ . Assumption (6.17) implies that

$$\mathbb{P}(X_1 \geq m) = \mathbb{P}(X_1 \leq -m) \sim \sum_{k=m}^{\infty} \frac{L(k)}{k^2} \sim \frac{L(m)}{m} \quad \text{for } m \rightarrow \infty, \quad (6.18)$$

and hence, for any  $x > 0$ , we have

$$n\mathbb{P}(X_1 \geq b_n x) \sim n \frac{L(nx\tilde{L}(nx))}{nx\tilde{L}(nx)} \sim \frac{1}{x} \quad \text{for } n \rightarrow \infty. \quad (6.19)$$

Thus (see, e.g., Ibragimov and Linnik [22], Theorem 2.6.1)  $S_n/b_n \rightarrow Y$  in distribution as  $n \rightarrow \infty$  with  $Y$  symmetric Cauchy. The standard local limit theorem (e.g. [22], Theorem 4.2.1) gives

$$p^n(x) = [1 + o(1)] \frac{1}{n\tilde{L}(n)} f\left(\frac{x}{n\tilde{L}(n)}\right), \quad \text{when } x \text{ and } n \text{ satisfy } |x|/(n\tilde{L}(n)) = O(1), \quad (6.20)$$

with  $f$  the density of  $Y$ , in particular,  $p^n(0) \sim f(0)/(n\tilde{L}(n))$ .

We cannot literally use results from Doney [15] for an analogue of (6.11) because the case  $a = 1$  is excluded there, but the proof of Theorem 1 in [15] can easily be adapted to show that there exist  $c, K > 0$  and  $n_0 \in \mathbb{N}$  such that

$$\frac{p^n(x)}{n|x|^{-2}L(|x|)} \geq c \quad \text{for } n \geq n_0, |x| \geq Kn\tilde{L}(n). \quad (6.21)$$

By the form of the strong law presented in Kallenberg [24], Theorem 4.23 (and noting that  $p$  possesses moments of any order strictly less than 1), we have for any  $\delta > 0$

$$\limsup_{n \rightarrow \infty} \frac{|S_n|}{n^{1+\delta}} = 0 \quad \text{a.s.} \quad (6.22)$$

Combining (6.20), (6.21) and (6.22), we obtain that almost surely

$$\frac{p^n(S_n)}{p^{2\lfloor n/2 \rfloor}(0)} \geq c' \frac{nn^{-2-2\delta}L(n^{1+\delta})}{(n\tilde{L}(n))^{-1}} = c' n^{-2\delta} L(n^{1+\delta})\tilde{L}(n) \quad \text{for } n \text{ large enough,} \quad (6.23)$$

so in particular

$$\liminf_{n \rightarrow \infty} \frac{\log[p^n(S_n)/p^{2\lfloor n/2 \rfloor}(0)]}{\log n} \geq -2\delta \quad S - \text{a.s.} \quad (6.24)$$

Let  $\delta \downarrow 0$  to conclude that (1.11) holds.

The fact that (1.12) is satisfied for this example can be checked analogously to (6.14–6.16), where now the sum is split according to whether  $|x| \leq Kn\tilde{L}(n)$  or  $|x| > Kn\tilde{L}(n)$ .

**4.** A(n admittedly rather restricted) class of two-dimensional symmetric transient random walks with  $\alpha = 1$  in (1.1) can be obtained as follows. Let  $q$  be a symmetric transition probability on  $\mathbb{Z}$

such that  $q(x) \sim L(|x|)|x|^{-3}$  for a suitable slowly varying function  $L$  that grows to  $\infty$ , in particular,  $q$  is in the non-normal domain of attraction of the one-dimensional normal law. Then

$$\mu_q(x) := \sum_{|k| \leq x} k^2 q(k) \sim 2 \sum_{k=1}^x \frac{L(|k|)}{k} \sim 2 \log(x) L(x) \quad \text{for } x \rightarrow \infty. \quad (6.25)$$

Let  $\tilde{L}$  be a conjugate of the slowly varying function  $1/(2L(x^{1/2})\log(x^{1/2}))$ , i.e.,  $\tilde{L}$  is defined (up to asymptotic equivalence) by the requirement  $\lim_{x \rightarrow \infty} (2L((x\tilde{L}(x))^{1/2}) \log((x\tilde{L}(x))^{1/2}))/\tilde{L}(x) = 1$  (again, in many examples, e.g.  $L(x) = \log(x)^c$  for  $c > 0$ , we can choose  $\tilde{L}(x) = 2L(x^{1/2})\log(x^{1/2})$ ).

Put  $b_n := (n\tilde{L}(n))^{1/2}$ . One can easily check that  $nb_n^{-2}\mu_q(b_n) \rightarrow 1$  as  $n \rightarrow \infty$  and that  $nq([b_n x, \infty)) \rightarrow 0$  for any  $x > 0$  as  $n \rightarrow \infty$ . Hence (see, e.g., Ibragimov and Linnik [22], Theorem 2.6.2 and its proof) for a  $q$ -random walk  $\tilde{S} = (\tilde{S}_n)_{n \in \mathbb{N}_0}$ ,  $\tilde{S}_n/(n\tilde{L}(n))^{1/2} \rightarrow Z$  in distribution, where  $Z$  is centred normal, and a local form of this limit holds as well. Analogous to (6.21), one can show that there exist  $c, K > 0$  and  $n_0 \in \mathbb{N}$  such that

$$\frac{q^n(x)}{n|x|^{-3}L(|x|)} \geq c \quad \text{for } n \geq n_0, |x| \geq K\sqrt{n\tilde{L}(n)}, \quad (6.26)$$

and the analogue of (6.22) is  $\limsup_{n \rightarrow \infty} n^{-\delta-1/2}|\tilde{S}_n| = 0$  for any  $\delta > 0$ , as follows again from Kallenberg [24], Theorem 4.23. Thus, we see that  $\liminf_{n \rightarrow \infty} \log[q^n(\tilde{S}_n)/q^{2\lfloor n/2 \rfloor}(0)]/\log n \geq 0$  a.s. by arguing as in 3. The fact that  $q$  satisfies (1.12) can be checked analogously to (6.14–6.16), this time splitting the sum according to whether  $|x|$  is smaller or larger than  $K(n\tilde{L}(n))^{1/2}$ .

Finally, assume that  $L$  is chosen so that  $\sum_{n \in \mathbb{N}} 1/(n\tilde{L}(n)) < \infty$ . For  $x = (x_1, x_2) \in \mathbb{Z}^2$ , put  $p(x) = q(x_1)q(x_2)$ . By the discussion above, the  $p$ -random walk is transient, satisfies (1.1) with  $\alpha = 1$  (we have  $p^n(0) = (q^n(0))^2 \sim 1/(n\tilde{L}(n))$ ) and (1.10–1.12) holds.

## A Appendix: Proof of Lemma 4.5

For the first inequality in (4.24), apply Lemma A.1 below with  $F = \tilde{E}$ ,  $G = E^{\text{tr}}$ ,  $\nu = q_{\rho, \nu}$ ,  $q = \pi_n[Q]_{\text{tr}}$ , where  $\pi_n$  denotes the projection onto the first  $n$  words. This yields

$$h(\pi_n Q_{\text{tr}}^\nu | q_{\rho, \nu}^{\otimes n}) \leq h(\pi_n [Q]_{\text{tr}} | q_{\rho, \nu}^{\otimes n}), \quad n \in \mathbb{N}, \quad (A.1)$$

implying  $H(Q_{\text{tr}}^\nu | q_{\rho, \nu}^{\otimes \mathbb{N}}) \leq H([Q]_{\text{tr}} | q_{\rho, \nu}^{\otimes \mathbb{N}})$ .

**Lemma A.1.** *Let  $F$  be countable,  $G \subset F$ ,  $\nu \in \mathcal{P}(F)$ ,  $n \in \mathbb{N}$ ,  $q \in \mathcal{P}(F^n)$ . Define  $q' \in \mathcal{P}(F^n)$  via*

$$q'(x) = q(\xi_G(x)) \prod_{i \in I_G(x)} \nu_G(x_i), \quad x = (x_1, \dots, x_n) \in F^n, \quad (A.2)$$

where  $I_G(x) = \{1 \leq i \leq n: x_i \in G\}$ ,  $\xi_G(x) = \{y \in F^n: y_i \in G \text{ if } i \in I_G(x), y_i = x_i \text{ if } i \notin I_G(x)\}$ ,  $\nu_G(\cdot) = \nu(\cdot \cap G)/\nu(G)$ , i.e., a  $q'$ -draw arises from a  $q$ -draw by replacing the coordinates in  $G$  by an independent draw from  $\nu$  conditioned to be in  $G$ . Then

$$h(q' | \nu^{\otimes n}) \leq h(q | \nu^{\otimes n}). \quad (A.3)$$

*Proof.* For  $I \subset \{1, \dots, n\}$ , we write  $I^c := \{1, \dots, n\} \setminus I$ . For  $y \in (F \setminus G)^{I^c}, z \in G^I$ , we denote by  $(y; z)$  the element of  $F^{\{1, \dots, n\}}$  defined by  $(y; z)_i = z_i$  if  $i \in I$ ,  $(y; z)_i = y_i$  if  $i \in I^c$ . Put  $q_{I,y}(z) := q(y; z)/q(\xi_{G,I}(y))$ , where  $\xi_{G,I}(y) = \{(y; z') : z' \in G^I\}$ , i.e.,  $q_{I,y} \in \mathcal{P}(G^I)$  is the law of the coordinates in  $I$  under  $q$  given that these take values in  $G$  and that the coordinates in  $I^c$  are equal to  $y$ .

Fix  $I \subset \{1, \dots, n\}, y \in (F \setminus G)^{I^c}$ . We first verify that

$$\sum_{z \in G^I} q'(y; z) \log \left( \frac{q'(y; z)}{\nu^{\otimes n}(y; z)} \right) \leq \sum_{z \in G^I} q(y; z) \log \left( \frac{q(y; z)}{\nu^{\otimes n}(y; z)} \right). \quad (\text{A.4})$$

By definition, the left-hand side of (A.4) equals

$$q(\xi_{G,I}(y)) \sum_{z \in G^I} \left( \prod_{i \in I} \nu_G(z_i) \right) \log \left( \frac{q(\xi_{G,I}(y))}{\nu(G)^{|I|} \prod_{j \in I^c} \nu(y_j)} \right) = q(\xi_{G,I}(y)) \log \left( \frac{q(\xi_{G,I}(y))}{\nu(G)^{|I|} \prod_{j \in I^c} \nu(y_j)} \right), \quad (\text{A.5})$$

whereas the right-hand side of (A.4) is equal to

$$q(\xi_{G,I}(y)) \sum_{z \in G^I} q_{I,y}(z) \log \left( \frac{q(\xi_{G,I}(y)) q_{I,y}(z)}{\prod_{i \in I} \nu(z_i) \times \prod_{j \in I^c} \nu(y_j)} \right). \quad (\text{A.6})$$

Thus, the right-hand side of (A.4) minus the left-hand side of (A.4) equals

$$q(\xi_{G,I}(y)) \sum_{z \in G^I} q_{I,y}(z) \log \left( \frac{q_{I,y}(z)}{\prod_{i \in I} \nu_G(z_i)} \right) = q(\xi_{G,I}(y)) h \left( q_{I,y}(\cdot) \mid \nu_G^{\otimes |I|} \right) \geq 0. \quad (\text{A.7})$$

The claim follows from (A.4) by observing that

$$h(q' \mid \nu^{\otimes n}) = \sum_{I \subset \{1, \dots, n\}} \sum_{y \in (F \setminus G)^{I^c}} \sum_{z \in G^I} q'(y; z) \log \left( \frac{q'(y; z)}{\nu^{\otimes n}(y; z)} \right), \quad (\text{A.8})$$

and analogously for  $h(q \mid \nu^{\otimes n})$ .  $\square$

For the proof of the second inequality in (4.24), i.e.,

$$H(\Psi_{Q_{\text{tr}}^y} \mid \nu^{\otimes \mathbb{N}}) \leq H(\Psi_{[Q]_{\text{tr}}} \mid \nu^{\otimes \mathbb{N}}), \quad (\text{A.9})$$

we need some further notation. Let  $\text{tr} \in \mathbb{N}$  be a given truncation level,  $*$  a new symbol,  $* \notin E$ ,  $E_* := E \cup \{*\}$ ,  $\tilde{E}_* := \bigcup_{n=0}^{\infty} (E_*)^n$ , where  $\tilde{E}_*^0 := \{\varepsilon\}$  with  $\varepsilon$  the empty word (i.e., the neutral element of  $\tilde{E}_*$  viewed as a semigroup under concatenation). For  $y \in \tilde{E}$ , let

$$\tilde{E}_* \ni [y]_{\text{tr},*} := \begin{cases} y, & \text{if } |y| < \text{tr}, \\ *^{\text{tr}}, & \text{if } |y| \geq \text{tr}, \end{cases} \quad (\text{A.10})$$

where  $*^{\text{tr}} = * \cdots *$  denotes the word in  $\tilde{E}_*$  consisting of  $\text{tr}$  times  $*$ , and

$$E^{\text{tr}} \cup \{\varepsilon\} \ni [y]_{\text{tr},\sim} := \begin{cases} \varepsilon, & \text{if } |y| < \text{tr}, \\ [y]_{\text{tr}}, & \text{if } |y| \geq \text{tr}. \end{cases} \quad (\text{A.11})$$

Let  $Q \in \mathcal{P}^{\text{erg}}(\tilde{E}^{\mathbb{N}})$  satisfy  $H([Q]_{\text{tr}}) < \infty$ . For  $Y = (Y^{(i)})_{i \in \mathbb{N}}$  with law  $Q$  and  $N \in \mathbb{N}$ , let

$$\begin{aligned} K^{(N, \text{tr})} &:= \kappa([Y^{(1)}]_{\text{tr}}, \dots, [Y^{(N)}]_{\text{tr}}), \\ K^{(N, \text{tr}, *)} &:= \kappa([Y^{(1)}]_{\text{tr}, *}, \dots, [Y^{(N)}]_{\text{tr}, *}), \\ K^{(N, \text{tr}, \sim)} &:= \kappa([Y^{(1)}]_{\text{tr}, \sim}, \dots, [Y^{(N)}]_{\text{tr}, \sim}). \end{aligned} \quad (\text{A.12})$$

Thus,  $K^{(N, \text{tr}, *)}$  consists of the letters in the first  $N$  words from  $[Y]_{\text{tr}}$  such that letters in words of length exactly equal to  $\text{tr}$  are masked by  $*$ 's, while  $K^{(N, \text{tr}, \sim)}$  consists of the letters in words of length  $\text{tr}$  among the first  $N$  words of  $[Y]_{\text{tr}}$ . Note that by construction there is a deterministic function  $\Xi: \tilde{E}_* \times \tilde{E} \rightarrow \tilde{E}$  such that  $K^{(N, \text{tr})} = \Xi(K^{(N, \text{tr}, *)}, K^{(N, \text{tr}, \sim)})$ . We assume that  $Q(\tau_1 \geq \text{tr}) > 0$ , otherwise  $K^{(N, \text{tr}, \sim)}$  is trivially equal to  $\varepsilon$  for all  $N$ .

Extend  $[\cdot]_{\text{tr}, *}$  and  $[\cdot]_{\text{tr}, \sim}$  in the obvious way to a map on  $\tilde{E}^{\mathbb{N}}$  and  $\mathcal{P}(\tilde{E}^{\mathbb{N}})$ . Then  $[Q]_{\text{tr}}, [Q]_{\text{tr}, *}, [Q]_{\text{tr}, \sim} \in \mathcal{P}^{\text{erg}}(\tilde{E}_*^{\mathbb{N}})$ ,  $m_{[Q]_{\text{tr}}} = m_{[Q]_{\text{tr}, *}} \leq \text{tr}$ ,  $m_{[Q]_{\text{tr}, \sim}} = \text{tr}$ ,  $\Psi_{[Q]_{\text{tr}}}, \Psi_{[Q]_{\text{tr}, *}}, \Psi_{[Q]_{\text{tr}, \sim}} \in \mathcal{P}^{\text{erg}}(E_*^{\mathbb{N}})$ . By ergodicity of  $Q$ , we have (see [5], Section 3.1, for analogous arguments)

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Q(K^{(N, \text{tr})}) = -m_{[Q]_{\text{tr}}} H(\Psi_{[Q]_{\text{tr}}}) \quad \text{a.s.} \quad (\text{A.13})$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Q(K^{(N, \text{tr}, *)}) = -m_{[Q]_{\text{tr}}} H(\Psi_{[Q]_{\text{tr}, *}}) \quad \text{a.s.} \quad (\text{A.14})$$

Since  $Q(K^{(N, \text{tr})}) = Q(K^{(N, \text{tr}, *)}, K^{(N, \text{tr}, \sim)}) = Q(K^{(N, \text{tr}, *)} | K^{(N, \text{tr}, \sim)})$ , we see from (A.13–A.14) that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Q(K^{(N, \text{tr}, \sim)} | K^{(N, \text{tr}, *)}) = -m_{[Q]_{\text{tr}}} (H(\Psi_{[Q]_{\text{tr}}}) - H(\Psi_{[Q]_{\text{tr}, *}})) =: -H_{\text{tr}, \sim | *}(Q) \quad \text{a.s.} \quad (\text{A.15})$$

The assumption  $H([Q]_{\text{tr}}) < \infty$  guarantees that all the quantities appearing in (A.13–A.15) are proper. Note that  $H_{\text{tr}, \sim | *}(Q)$  can be interpreted as the conditional specific relative entropy of the letters in the “long” words of  $[Y]_{\text{tr}}$  given the letters in the “short” words (see Lemma A.2 below). Note that  $H_{\text{tr}, \sim | *}(Q)$  in (A.15) is defined as a “per word” quantity. Since the fraction of long words in  $[Y]_{\text{tr}}$  is  $Q(\tau_1 \geq \text{tr})$  and each of these words contains  $\text{tr}$  letters, the corresponding conditional specific relative entropy “per letter” is  $H_{\text{tr}, \sim | *}(Q) / [Q(\tau_1 \geq \text{tr}) \text{tr}]$ , as it appears in (A.22) below.

*Proof of (A.9).* Without loss of generality we may assume that  $Q(\tau_1 \geq \text{tr}) \in (0, 1)$ . Indeed, if  $Q(\tau_1 \geq \text{tr}) = 0$ , then  $Q_{\text{tr}}^v = [Q]_{\text{tr}}$ , while if  $Q(\tau_1 \geq \text{tr}) = 1$ , then  $\Psi_{Q_{\text{tr}}^v} = v^{\otimes \mathbb{N}}$ . In both cases (A.9) obviously holds.

**Step 1.** We will first assume that  $|E| < \infty$ . Then  $H([Q]_{\text{tr}}) < \infty$  is automatic. Since  $v^{\otimes \mathbb{N}}$  is a product measure, we have, for any  $\Psi \in \mathcal{P}^{\text{inv}}(E^{\mathbb{N}})$ ,

$$H(\Psi | v^{\otimes \mathbb{N}}) = -H(\Psi) - \sum_{x \in E} \Psi(\{x\} \times E^{\mathbb{N}}) \log v(x), \quad (\text{A.16})$$

where  $H(\Psi)$  denotes the specific entropy of  $\Psi$ . We have

$$\begin{aligned} H(\Psi_{[Q]_{\text{tr}}} | v^{\otimes \mathbb{N}}) &= -H(\Psi_{[Q]_{\text{tr}}}) - \frac{1}{m_{[Q]_{\text{tr}}}} \mathbb{E}_Q \left[ \sum_{j=1}^{\tau_1 \wedge \text{tr}} \log v(Y_j^{(1)}) \right], \\ H(\Psi_{Q_{\text{tr}}^v} | v^{\otimes \mathbb{N}}) &= -H(\Psi_{Q_{\text{tr}}^v}) - \frac{1}{m_{[Q]_{\text{tr}}}} \left( \mathbb{E}_Q \left[ \sum_{j=1}^{\tau_1 \wedge \text{tr}} \log v(Y_j^{(1)}); \tau_1 < \text{tr} \right] - Q(\tau_1 \geq \text{tr}) \text{tr} h(v) \right), \end{aligned} \quad (\text{A.17})$$

where  $h(\nu) = -\sum_{x \in E} \nu(x) \log \nu(x)$  is the entropy of  $\nu$ . Hence

$$\begin{aligned} & H(\Psi_{[Q]_{\text{tr}}} | \nu^{\otimes \mathbb{N}}) - H(\Psi_{Q_{\text{tr}}^{\nu}} | \nu^{\otimes \mathbb{N}}) \\ &= -[H(\Psi_{[Q]_{\text{tr}}}) - H(\Psi_{Q_{\text{tr}}^{\nu}})] - \frac{1}{m_{[Q]_{\text{tr}}}} \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log \nu(Y_j^{(1)}); \tau_1 \geq \text{tr} \right] - \frac{Q(\tau_1 \geq \text{tr}) \text{tr}}{m_{[Q]_{\text{tr}}}} h(\nu). \end{aligned} \quad (\text{A.18})$$

By (A.15) applied to  $Q$  and to  $Q_{\text{tr}}^{\nu}$  (note that  $[Q_{\text{tr}}^{\nu}]_{\text{tr}} = Q_{\text{tr}}^{\nu}$ ), we have

$$H(\Psi_{[Q]_{\text{tr}}}) = H(\Psi_{[Q]_{\text{tr},*}}) + \frac{1}{m_{[Q]_{\text{tr}}}} H_{\text{tr}, \sim |*}(Q), \quad (\text{A.19})$$

$$H(\Psi_{Q_{\text{tr}}^{\nu}}) = H(\Psi_{[Q_{\text{tr}}^{\nu}]_{\text{tr},*}}) + \frac{1}{m_{[Q_{\text{tr}}^{\nu}]_{\text{tr}}}} H_{\text{tr}, \sim |*}(Q_{\text{tr}}^{\nu}). \quad (\text{A.20})$$

By construction,  $m_{[Q_{\text{tr}}^{\nu}]_{\text{tr}}} = m_{[Q]_{\text{tr}}}$ ,  $[Q_{\text{tr}}^{\nu}]_{\text{tr},*} = [Q]_{\text{tr},*}$ ,  $H_{\text{tr}, \sim |*}(Q_{\text{tr}}^{\nu}) = Q(\tau_1 \geq \text{tr}) \text{tr} h(\nu)$ . Combining (A.18–A.20), we obtain

$$H(\Psi_{[Q]_{\text{tr}}} | \nu^{\otimes \mathbb{N}}) - H(\Psi_{Q_{\text{tr}}^{\nu}} | \nu^{\otimes \mathbb{N}}) = \frac{1}{m_{[Q]_{\text{tr}}}} \left( -H_{\text{tr}, \sim |*}(Q) - \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log \nu(Y_j^{(1)}); \tau_1 \geq \text{tr} \right] \right). \quad (\text{A.21})$$

Finally, we observe that

$$\begin{aligned} & \frac{1}{Q(\tau_1 \geq \text{tr}) \text{tr}} \left( -H_{\text{tr}, \sim |*}(Q) - \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log \nu(Y_j^{(1)}); \tau_1 \geq \text{tr} \right] \right) \\ &= \frac{-H_{\text{tr}, \sim |*}(Q)}{Q(\tau_1 \geq \text{tr}) \text{tr}} - \frac{1}{\text{tr}} \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log \nu(Y_j^{(1)}) \middle| \tau_1 \geq \text{tr} \right] \end{aligned} \quad (\text{A.22})$$

is the “specific relative entropy of the law of letters in the concatenation of long words given the concatenation of short words in  $[Q]_{\text{tr}}$  with respect to  $\nu^{\otimes \mathbb{N}}$ ”, which is  $\geq 0$  (see Lemma A.2 below).

**Step 2.** We extend (A.9) to a general letter space  $E$  by using the *coarse-graining* construction from [6], Section 8. Let  $\mathcal{A}_c = \{A_{c,1}, \dots, A_{c,n_c}\}$ ,  $c \in \mathbb{N}$ , be a sequence of nested finite partitions of  $E$ , and let  $\langle \cdot \rangle_c : E \rightarrow \langle E \rangle_c$  be the coarse-graining map as defined in [6], Section 8. Since  $\langle E \rangle_c$  is finite and the word length truncation  $[\cdot]_{\text{tr}}$  and the letter coarse-graining  $\langle \cdot \rangle_c$  commute, we have

$$H(\langle \Psi_{Q_{\text{tr}}^{\nu}} \rangle_c | \langle \nu^{\otimes \mathbb{N}} \rangle_c) \leq H(\langle \Psi_{[Q]_{\text{tr}}} \rangle_c | \langle \nu^{\otimes \mathbb{N}} \rangle_c) \quad \text{for all } c \in \mathbb{N} \quad (\text{A.23})$$

by Step 1. This implies (A.9) by taking  $c \rightarrow \infty$  (see the arguments in [6], Lemma 8.1 and the second part of (8.13)).  $\square$

**Lemma A.2.** Assume  $|E| < \infty$ . Let  $\text{tr} \in \mathbb{N}$ ,  $Q \in \mathcal{P}^{\text{erg}}(\tilde{E}^{\mathbb{N}})$  with  $Q(\tau_1 \geq \text{tr}) > 0$ . For  $N \in \mathbb{N}$ , put  $\tilde{L}_N := |K^{(N, \text{tr}, \sim)}|$ . Then a.s.

$$\begin{aligned} 0 &\leq \lim_{N \rightarrow \infty} \frac{1}{\tilde{L}_N} h(Q(K^{(N, \text{tr}, \sim)} \in \cdot | K^{(N, \text{tr}, *)}) | \nu^{\otimes \tilde{L}_N}) \\ &= \frac{-H_{\text{tr}, \sim |*}(Q)}{Q(\tau_1 \geq \text{tr}) \text{tr}} - \frac{1}{\text{tr}} \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log \nu(Y_j^{(1)}) \middle| \tau_1 \geq \text{tr} \right]. \end{aligned}$$



*Proof.* Note that, by construction,  $\tilde{L}_N = \tilde{L}_N(K^{(N,\text{tr},*)})$  is a deterministic function of  $K^{(N,\text{tr},*)}$  (namely, the number of  $*$ 's in  $K^{(N,\text{tr},*)}$ ), and

$$\lim_{N \rightarrow \infty} \tilde{L}_N/N = \text{tr}Q(\tau_1 \geq \text{tr}) \quad \text{a.s.} \quad (\text{A.24})$$

by ergodicity of  $Q$ . Fix  $\epsilon > 0$ . By ergodicity of  $Q$ , there exists a random  $N_0 < \infty$  such that for all  $N \geq N_0$  there is a finite (random) set  $B_{N,\epsilon} = B_{N,\epsilon}(K^{(N,\text{tr},*)}) \subset E^{\tilde{L}_N}$  such that  $Q(K^{(N,\text{tr},\sim)} \in B_{N,\epsilon} \mid K^{(N,\text{tr},*)}) \geq 1 - \epsilon$ ,

$$\frac{1}{\tilde{L}_N} \log Q(K^{(N,\text{tr},\sim)} = b \mid K^{(N,\text{tr},*)}) \in [-H_{\text{tr},\sim|*}(Q) - \epsilon, -H_{\text{tr},\sim|*}(Q) + \epsilon] \quad (\text{A.25})$$

and

$$\frac{1}{\tilde{L}_N} \sum_{j=1}^{\tilde{L}_N} \log v(b_j) \in [\chi - \epsilon, \chi + \epsilon] \quad \text{with} \quad \chi = \frac{1}{\text{tr}} \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log v(Y_j^{(1)}) \mid \tau_1 \geq \text{tr} \right] \quad (\text{A.26})$$

for all  $b = (b_1, \dots, b_{\tilde{L}_N}) \in B_{N,\epsilon}$ . Here, (A.25) follows from (A.15), while for (A.26) we note that

$$\begin{aligned} \lim_{N \rightarrow \infty} N^{-1} \sum_{j=1}^{|K^{(N,\text{tr})}|} \log v(K_j^{(N,\text{tr})}) &= \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr} \wedge \tau_1} \log v(Y_j^{(1)}) \right], \\ \lim_{N \rightarrow \infty} N^{-1} \sum_{j=1}^{|K^{(N,\text{tr},\sim)}|} \log v(K_j^{(N,\text{tr},\sim)}) &= \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log v(Y_j^{(1)}); \tau_1 \geq \text{tr} \right], \end{aligned} \quad (\text{A.27})$$

and recall (A.24). It follows that

$$\begin{aligned} &\frac{1}{\tilde{L}_N} h(Q(K^{(N,\text{tr},\sim)} \in \cdot \mid K^{(N,\text{tr},*)}) \mid v^{\otimes \tilde{L}_N}) \\ &= \frac{1}{\tilde{L}_N} \sum_{b \in E^{\tilde{L}_N}} Q(K^{(N,\text{tr},\sim)} = b \mid K^{(N,\text{tr},*)}) \log \left( \frac{Q(K^{(N,\text{tr},\sim)} = b \mid K^{(N,\text{tr},*)})}{\prod_{j=1}^{\tilde{L}_N} v(b_j)} \right) =: \Sigma_1 + \Sigma_2, \end{aligned} \quad (\text{A.28})$$

where  $\Sigma_1$  runs over  $b \in B_{N,\epsilon}$  and  $\Sigma_2$  over  $b \in E^{\tilde{L}_N} \setminus B_{N,\epsilon}$ . We have

$$\Sigma_1 \in [\chi' - \epsilon', \chi' + \epsilon'] \quad \text{with} \quad \chi' = -\frac{H_{\text{tr},\sim|*}(Q)}{\text{tr}Q(\tau_1 \geq \text{tr})} - \frac{1}{\text{tr}} \mathbb{E}_Q \left[ \sum_{j=1}^{\text{tr}} \log v(Y_j^{(1)}) \mid \tau_1 \geq \text{tr} \right] \quad (\text{A.29})$$

for  $N \geq N_0$  by (A.24–A.26), where  $\epsilon' = \epsilon'(Q, \epsilon)$  tends to zero as  $\epsilon \downarrow 0$ . Multiplying and dividing by  $Q(K^{(N,\text{tr},\sim)} \notin B_{N,\epsilon} \mid K^{(N,\text{tr},*)})$ , we see that

$$\begin{aligned} |\Sigma_2| &\leq Q(K^{(N,\text{tr},\sim)} \notin B_{N,\epsilon} \mid K^{(N,\text{tr},*)}) \max_{b \in E} \log \left( \frac{1}{v(b)} \right) \\ &\quad - Q(K^{(N,\text{tr},\sim)} \notin B_{N,\epsilon} \mid K^{(N,\text{tr},*)}) \log Q(K^{(N,\text{tr},\sim)} \notin B_{N,\epsilon} \mid K^{(N,\text{tr},*)}) \\ &\quad + Q(K^{(N,\text{tr},\sim)} \notin B_{N,\epsilon} \mid K^{(N,\text{tr},*)}) \log |E|, \end{aligned} \quad (\text{A.30})$$

which tends to zero as  $\epsilon \downarrow 0$  because  $Q(K^{(N,\text{tr},\sim)} \notin B_{N,\epsilon} \mid K^{(N,\text{tr},*)}) \leq \epsilon$ .  $\square$

## References

- [1] Q. Berger and F.L. Toninelli, On the critical point of the random walk pinning model in dimension  $d = 3$ , *Electron. J. Probab.* 15 (2010) 654–683. MR2650777
- [2] R.N. Bhattacharya and R. Ranga Rao, *Normal Approximation and Asymptotic Expansions* (corrected ed.), Robert E. Krieger Publishing Co., Inc., Melbourne, FL, 1986. MR0855460
- [3] M. Birkner, *Particle Systems with Locally Dependent Branching: Long-Time Behaviour, Genealogy and Critical Parameters*, Dissertation, Johann Wolfgang Goethe-Universität Frankfurt am Main, 2003. <http://publikationen.uni-frankfurt.de/volltexte/2003/314/>
- [4] M. Birkner, A condition for weak disorder for directed polymers in random environment, *Electron. Comm. Probab.* 9 (2004) 22–25. MR2041302
- [5] M. Birkner, Conditional large deviations for a sequence of words, *Stochastic Process. Appl.* 118 (2008) 703–729. MR2411517
- [6] M. Birkner, A. Greven, F. den Hollander, Quenched large deviation principle for words in a letter sequence, *Probab. Theory Relat. Fields* 148 (2010) 403–456. MR2678894
- [7] M. Birkner and R. Sun, Annealed vs quenched critical points for a random walk pinning model, *Ann. Inst. Henri Poincaré* 46 (2010), 414–441. MR2667704
- [8] M. Birkner and R. Sun, Disorder relevance for the random walk pinning model in dimension 3, *Ann. Inst. Henri Poincaré* 47 (2011), 259–293.
- [9] E. Bolthausen, A note on the diffusion of directed polymers in a random environment, *Commun. Math. Phys.* 123 (1989) 529–534. MR1006293
- [10] A. Camanes and P. Carmona, The critical temperature of a directed polymer in a random environment, *Markov Proc. Relat. Fields* 15 (2009) 105–116. MR2509426
- [11] F. Comets, T. Shiga and N. Yoshida, Directed polymers in random environment: path localization and strong disorder, *Bernoulli* 9 (2003) 705–723. MR1996276
- [12] F. Comets and N. Yoshida, Directed polymers in random environment are diffusive at weak disorder, *Ann. Probab.* 34 (2006) 1746–1770. MR2271480
- [13] J. Chover, A law of the iterated logarithm for stable summands, *Proc. Amer. Math. Soc.* 17 (1966) 441–443. MR0189096
- [14] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (2nd ed.), Springer, New York, 1998. MR1619036
- [15] R.A. Doney, One-sided local large deviation and renewal theorems in the case of infinite mean, *Probab. Theory Relat. Fields* 107 (1997) 451–465. MR1440141
- [16] M.R. Evans and B. Derrida, Improved bounds for the transition temperature of directed polymers in a finite-dimensional random medium, *J. Stat. Phys.* 69 (1992) 427–437.

- [17] A. Greven, Phase transition for the coupled branching process, Part I: The ergodic theory in the range of finite second moments, *Probab. Theory Relat. Fields* 87 (1991) 417–458. MR1085176
- [18] A. Greven, On phase transitions in spatial branching systems with interaction, in: *Stochastic Models* (L.G. Gorostiza and B.G. Ivanoff, eds.), CMS Conference Proceedings 26 (2000) 173–204. MR1765010
- [19] A. Greven and F. den Hollander, Phase transitions for the long-time behavior of interacting diffusions, *Ann. Probab.* 35 (2007) 1250–1306. MR2330971
- [20] C.C. Heyde, A note concerning behaviour of iterated logarithm type, *Proc. Amer. Math. Soc.* 23 (1969) 85–90. MR0251772
- [21] F. den Hollander, *Random polymers*, Lectures from the 37th Probability Summer School held in Saint-Flour, 2007. *Lecture Notes in Mathematics* 1974. Springer, Heidelberg, 2009. MR2504175
- [22] I.A. Ibragimov and Yu.V. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen, 1971. MR0322926
- [23] O. Kallenberg, Stability of critical cluster fields, *Math. Nachr.* 77 (1977) 7–43. MR0443078
- [24] O. Kallenberg, *Foundations of Modern Probability* (2nd ed.), Springer, New York, 2002. MR1876169
- [25] C. Monthus and T. Garel, Freezing transition of the directed polymer in a 1+d random medium: Location of the critical temperature and unusual critical properties, *Phys. Rev. E* 74 (2006), 011101.
- [26] E. Seneta, *Regularly Varying Functions*, *Lecture Notes in Mathematics* 508, Springer, Berlin-New York, 1976. MR0453936
- [27] F. Spitzer, *Principles of Random Walk* (2nd. ed.), Springer, New York, 1976. MR0388547