Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2011, Vol. 47, No. 1, 214–242
DOI: 10.1214/10-AIHP369

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Process-level quenched large deviations for random walk in random environment

Firas Rassoul-Agha[a,1] and Timo Seppäläinen[b,2]

[a]*Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, UT 84109, USA.*
E-mail: firas@math.utah.edu
[b]*Department of Mathematics, University of Wisconsin-Madison, 419 Van Vleck Hall, Madison, WI 53706, USA.*
E-mail: seppalai@math.wisc.edu

Received 7 September 2009; revised 9 April 2010; accepted 16 April 2010

Abstract. We consider a bounded step size random walk in an ergodic random environment with some ellipticity, on an integer lattice of arbitrary dimension. We prove a level 3 large deviation principle, under almost every environment, with rate function related to a relative entropy.

Résumé. Nous considérons une marche aléatoire en environnement aléatoire ergodique. La marche est elliptique et à pas bornés. Nous prouvons un principe de grandes déviations au niveau 3, sous presque tout environnement, avec une fonctionnelle d'action liée à une entropie relative.

MSC: 60K37; 60F10; 82D30; 82C44

Keywords: Random walk; Random environment; RWRE; Large deviation; Environment process; Relative entropy; Homogenization

1. Introduction

We describe the standard model of random walk in random environment (RWRE) on $\mathbb{Z}^d$. Let Ω be a Polish space and $\mathfrak{S}$ its Borel σ-algebra. Let $\{T_z\colon z \in \mathbb{Z}^d\}$ be a group of continuous commuting bijections on Ω: $T_{x+y} = T_x T_y$ and T_0 is the identity. Let $\mathbb{P}$ be a $\{T_z\}$-invariant probability measure on $(\Omega, \mathfrak{S})$ that is ergodic under this group. In other words, the σ-algebra of Borel sets invariant under $\{T_z\}$ is trivial under $\mathbb{P}$.

Denote the space of probability distributions on $\mathbb{Z}^d$ by $\mathcal{P} = \{(p_z)_{z\in\mathbb{Z}^d} \in [0,1]^{\mathbb{Z}^d}\colon \sum_z p_z = 1\}$ and give it the weak topology or, equivalently, the restriction of the product topology. Let $\omega \mapsto (p_z(\omega))_{z\in\mathbb{Z}^d}$ be a continuous mapping from Ω to $\mathcal{P}$. For $x, y \in \mathbb{Z}^d$ define $\pi_{x,y}(\omega) = p_{y-x}(T_x\omega)$. We call ω and also $(\pi_{x,y}(\omega))_{x,y\in\mathbb{Z}^d}$ an environment because it determines the transition probabilities of a Markov chain.

The set of admissible steps is denoted by $\mathscr{R} = \{z\colon \mathbb{E}[\pi_{0,z}] > 0\}$. One can then redefine $\mathcal{P} = \{(p_z)_{z\in\mathscr{R}} \in [0,1]^{\mathscr{R}}\colon \sum_z p_z = 1\}$ and transition probabilities $\pi_{x,y}$ are defined only for $x, y \in \mathbb{Z}^d$ such that $y - x \in \mathscr{R}$.

Given ω and a starting point $x \in \mathbb{Z}^d$, let P_x^ω be the law of the Markov chain $X_{0,\infty} = (X_n)_{n\ge0}$ on $\mathbb{Z}^d$, starting at $X_0 = x$ and having transition probabilities $(\pi_{y,y+z}(\omega))$. That is,

$$P_x^\omega\{X_{n+1} = y + z \mid X_n = y\} = \pi_{y,y+z}(\omega) \quad \text{for all } y, z \in \mathbb{Z}^d.$$

---

[1] Supported in part by NSF Grant DMS-0747758.
[2] Supported in part by NSF Grant DMS-0701091 and by the Wisconsin Alumni Research Foundation.

$X_{0,\infty}$ is called a random walk in environment ω and P_x^ω is called the *quenched* distribution. The *joint* distribution is $P_x(\mathrm{d}x_{0,\infty}, \mathrm{d}\omega) = P_x^\omega(\mathrm{d}x_{0,\infty})\mathbb{P}(\mathrm{d}\omega)$. Its marginal on $(\mathbb{Z}^d)^{\mathbb{Z}_+}$ is also denoted by P_x and called the *averaged* (or *annealed*) distribution since ω is averaged out:

$$P_x(A) = \int P_x^\omega(A)\mathbb{P}(\mathrm{d}\omega) \quad \text{for a measurable } A \subset \left(\mathbb{Z}^d\right)^{\mathbb{Z}_+}.$$

The canonical case of the above setting is $\Omega = \mathcal{P}^{\mathbb{Z}^d}$ and $p_z(\omega) = (\omega_0)_z$.

Next a quick description of the problem we are interested in. Assume given a sequence of probability measures Q_n on a Polish space $(\mathcal{X}, \mathscr{B}_{\mathcal{X}})$ and a lower semicontinuous function $I : \mathcal{X} \to [0, \infty]$. Then the *large deviation upper bound* holds with *rate function* I if

$$\varlimsup_{n\to\infty} n^{-1} \log Q_n(C) \le -\inf_C I \quad \text{for all closed sets } C \subset \mathcal{X}.$$

Similarly, rate function I governs the *large deviation lower bound* if

$$\varliminf_{n\to\infty} n^{-1} \log Q_n(O) \ge -\inf_O I \quad \text{for all open sets } O \subset \mathcal{X}.$$

If both hold with the same rate function I, then the *large deviation principle* (LDP) holds with rate I. We shall use basic, well known features of large deviation theory and relative entropy without citing every instance. The reader can consult references [3–5,15,21].

If the upper bound (resp., lower bound, resp., LDP) holds with some function $I : \mathcal{X} \to [0, \infty]$, then it also holds with the *lower semicontinuous regularization* I_{lsc} of I defined by

$$I_{\mathrm{lsc}}(x) = \sup\Big\{\inf_O I : x \in O \text{ and } O \text{ is open}\Big\}.$$

Thus the rate function can be required to be lower semicontinuous, and then it is unique.

Large deviations arrange themselves more or less naturally in three levels. Most of the work on quenched large deviations for RWRE has been at *level 1*, that is, on large deviations for $P_0^\omega\{X_n/n \in \cdot\}$. Greven and den Hollander [10] considered the product one-dimensional nearest-neighbor case, Comets, Gantert and Zeitouni [2] the ergodic one-dimensional nearest-neighbor case, Yilmaz [24] the ergodic one-dimensional case with bounded step size, Zerner [25] the multi-dimensional product nestling case, and Varadhan [22] the general ergodic multidimensional case with bounded step size. Rosenbluth [17] gave a variational formula for the rate function in [22]. *Level 2* quenched large deviations appeared in the work of Yilmaz [24] for the distributions $P_0^\omega\{n^{-1}\sum_{k=0}^{n-1} \delta_{T_{X_k}\omega, Z_{k+1}} \in \cdot\}$. Here $Z_k = X_k - X_{k-1}$ denotes the step of the walk.

Our object of study, *level 3* or *process level* large deviations concerns the *empirical process*

$$R_n^{1,\infty} = n^{-1} \sum_{k=0}^{n-1} \delta_{T_{X_k}\omega, Z_{k+1,\infty}}, \tag{1.1}$$

where $Z_{k+1,\infty} = (Z_i)_{i\ge k+1}$ denotes the entire sequence of future steps. Quenched distributions $P_0^\omega\{R_n^{1,\infty} \in \cdot\}$ are probability measures on the space $\mathcal{M}_1(\Omega \times \mathscr{R}^{\mathbb{N}})$. This is the space of Borel probability measures on $\Omega \times \mathscr{R}^{\mathbb{N}}$ endowed with the weak topology generated by bounded continuous functions.

The levels do form a hierarchy: higher level LDPs can be projected down to give LDPs at lower levels. Such results are called *contraction principles* in large deviation theory.

The main technical contribution of this work is the extension of a homogenization argument that proves the upper bound to the multivariate level 2 setting. This idea goes back to Kosygina, Rezakhanlou and Varadhan [12] in the context of diffusions with random drift, and was used by both Rosenbluth [17] and Yilmaz [24] to prove their LDPs.

Before turning to specialized assumptions and notation, here are some general conventions. $\mathbb{Z}_+$, $\mathbb{Z}_-$ and $\mathbb{N}$ denote, respectively, the set of non-negative, non-positive and positive integers. $|\cdot|$ denotes the ℓ^∞-norm on $\mathbb{R}^d$. $\{e_1, \dots, e_d\}$ is the canonical basis of $\mathbb{R}^d$. In addition to $\mathcal{M}_1(\mathcal{X})$ for the space of probability measures on $\mathcal{X}$, we write $\mathcal{Q}(\mathcal{X})$ for the set of Markov transition kernels on $\mathcal{X}$. Our spaces are Polish and the σ-algebras Borel. Given $\mu \in \mathcal{M}_1(\mathcal{X})$ and $q \in \mathcal{Q}(\mathcal{X})$, $\mu \times q$ is the probability measure on $\mathcal{X} \times \mathcal{X}$ defined by

$$\mu \times q(A \times B) = \int \mathbb{1}_A(x) q(x, B) \mu(\mathrm{d}x)$$

and μq is its second marginal. For a probability measure P, E^P denotes the corresponding expectation operator. Occasionally $P(f)$ may replace $E^P[f]$.

2. Main result

Fix a dimension $d \geq 1$. Following are the hypotheses for the level 3 LDP. In Section 3 we refine these to state precisely what is used by different parts of the proof.

$$\mathscr{R} \text{ is finite and } \Omega \text{ is a compact metric space.} \tag{2.1}$$

$$\forall x \in \mathbb{Z}^d, \exists m \in \mathbb{N} \text{ and } z_1, \ldots, z_m \in \mathscr{R} \text{ such that } x = z_1 + \cdots + z_m. \tag{2.2}$$

$$\exists p > d \text{ such that } \mathbb{E}[|\log \pi_{0,z}|^p] < \infty \ \forall z \in \mathscr{R}. \tag{2.3}$$

When $\mathscr{R}$ is finite the canonical $\Omega = \mathcal{P}^{\mathbb{Z}^d}$ is compact. The commonly used assumption of uniform ellipticity, namely the existence of $\kappa > 0$ such that $\mathbb{P}\{\pi_{0,z} \geq \kappa\} = 1$ for $z \in \mathscr{R}$ and $\mathscr{R}$ contains the $2d$ unit vectors, implies assumptions (2.2) and (2.3).

We need notational apparatus for backward, forward and bi-infinite paths. The increments of a bi-infinite path $(x_i)_{i \in \mathbb{Z}}$ in $\mathbb{Z}^d$ with $x_0 = 0$ are denoted by $z_i = x_i - x_{i-1}$. The sequences (x_i) and (z_i) are in 1–1 correspondence. Segments of sequences are denoted by $z_{i,j} = (z_i, z_{i+1}, \ldots, z_j)$, also for $i = -\infty$ or $j = \infty$, and also for random variables: $Z_{i,j} = (Z_i, Z_{i+1}, \ldots, Z_j)$.

In general $\eta_{i,j}$ denotes the pair $(\omega, z_{i,j})$, but when i and j are clear from the context we write simply η. We will also sometimes abbreviate $\eta_- = \eta_{-\infty,0}$. The spaces to which elements η belong are $\boldsymbol{\Omega}_- = \Omega \times \mathscr{R}^{\mathbb{Z}_-}$, $\boldsymbol{\Omega}_+ = \Omega \times \mathscr{R}^{\mathbb{N}}$ and $\boldsymbol{\Omega} = \Omega \times \mathscr{R}^{\mathbb{Z}}$. Their relevant shift transformations are

$$S_z^- : \boldsymbol{\Omega}_- \to \boldsymbol{\Omega}_- : (\omega, z_{-\infty,0}) \mapsto (T_z\omega, z_{-\infty,0}, z),$$
$$S^+ : \boldsymbol{\Omega}_+ \to \boldsymbol{\Omega}_+ : (\omega, z_{1,\infty}) \mapsto (T_{z_1}\omega, z_{2,\infty}),$$
$$S : \boldsymbol{\Omega} \to \boldsymbol{\Omega} : (\omega, z_{-\infty,\infty}) \mapsto (T_{z_1}\omega, \bar{z}_{-\infty,\infty}),$$

where $\bar{z}_i = z_{i+1}$. We use the same symbols S_z^-, S^+ and S to act on $z_{-\infty,0}$, $z_{1,\infty}$ and $z_{-\infty,\infty}$ in the same way.

The empirical process (1.1) lives in $\boldsymbol{\Omega}_+$ but the rate function is best defined in terms of backward paths. Invariance allows us to pass conveniently between theses settings. If $\mu \in \mathcal{M}_1(\boldsymbol{\Omega}_+)$ is S^+-invariant, it has a unique S-invariant extension $\bar{\mu}$ on $\boldsymbol{\Omega}$. Let $\mu_- = \bar{\mu}_{|\boldsymbol{\Omega}_-}$, the restriction of $\bar{\mu}$ to its marginal on $\boldsymbol{\Omega}_-$. There is a unique kernel q_μ on $\boldsymbol{\Omega}_-$ that fixes μ_- (that is, $\mu_- q_\mu = \mu_-$) and satisfies

$$q_\mu\bigl(\eta_-, \bigl\{S_z^- \eta_- : z \in \mathscr{R}\bigr\}\bigr) = 1 \quad \text{for } \mu_-\text{-a.e. } \eta_-. \tag{2.4}$$

Namely

$$q_\mu\bigl(\eta_-, S_z^- \eta_-\bigr) = \bar{\mu}\bigl\{Z_1 = z | (\omega, Z_{-\infty,0}) = \eta_-\bigr\}.$$

(Uniqueness here is μ_--a.s.) Indeed, on the one hand, the above q_μ does leave μ_- invariant. On the other hand, if q is a kernel supported on shifts and leaves μ_- invariant, and if f is a bounded measurable function on $\boldsymbol{\Omega}_-$, then

$$\begin{aligned}
&\int q\bigl(\eta_-, S_z^- \eta_-\bigr) f(\eta_-) \bar{\mu}(\mathrm{d}\eta) \\
&\quad = \sum_{z'} \int q\bigl(\eta_-, S_{z'}^- \eta_-\bigr) f(T_{z'-z}\omega, z_{-\infty,0}) \mathbb{1}\bigl\{z' = z\bigr\} \mu_-(\mathrm{d}\eta_-)
\end{aligned}$$

$$= \int f(T_{-z}\omega, z_{-\infty,-1})\mathbb{1}\{z_0 = z\}\mu_-(\mathrm{d}\eta_-)$$

$$= \int f(\eta_-)\mathbb{1}\{z_1 = z\}\bar{\mu}(\mathrm{d}\eta)$$

$$= \int q_\mu\big(\eta_-, S_z^-\eta_-\big) f(\eta_-)\bar{\mu}(\mathrm{d}\eta).$$

The RWRE transition gives us the kernel $p^- \in \mathcal{Q}(\boldsymbol{\Omega}_-)$ defined by

$$p^-\big(\eta_-, S_z^-\eta_-\big) = \pi_{0,z}(\omega) \quad \text{for } \eta_- = (\omega, z_{-\infty,0}) \in \boldsymbol{\Omega}_-.$$

If $q \in \mathcal{Q}(\boldsymbol{\Omega}_-)$ satisfies $\mu_- \times q \ll \mu_- \times p^-$, then $q(\eta_-, \{S_z^-\eta_-\colon z \in \mathscr{R}\}) = 1$ μ_--a.s. and their relative entropy is given by

$$H\big(\mu_- \times q|\mu_- \times p^-\big) = \int \sum_{z\in\mathscr{R}} q\big(\eta_-, S_z^-\eta_-\big) \log \frac{q(\eta_-, S_z^-\eta_-)}{p^-(\eta_-, S_z^-\eta_-)}\mu_-(\mathrm{d}\eta_-). \tag{2.5}$$

Let μ_0 denote the marginal of μ on Ω. Our main theorem is the following.

Theorem 2.1. *Let $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z\})$ be an ergodic system. Assume (2.1), (2.2) and (2.3). Then, for $\mathbb{P}$-a.e. ω, the large deviation principle holds for the laws $P_0^\omega\{R_n^{1,\infty} \in \cdot\}$, with rate function $H_{\mathrm{quen}} : \mathcal{M}_1(\boldsymbol{\Omega}_+) \to [0,\infty]$ equal to the lower semicontinuous regularization of the convex function*

$$H(\mu) = \begin{cases} H\big(\mu_- \times q_\mu|\mu_- \times p^-\big) & \text{if } \mu \text{ is } S^+\text{-invariant and } \mu_0 \ll \mathbb{P}, \\ \infty & \text{otherwise.} \end{cases} \tag{2.6}$$

We make next some observations about the rate function H_{quen}.

Remark 2.1. *As is often the case for process level LDPs, the rate function is affine. This follows because we can replace q_μ with a "universal" kernel $\bar{q}$ whose definition is independent of μ. Namely, define*

$$U : \boldsymbol{\Omega}_- \to \boldsymbol{\Omega}_- : (\omega, z_{-\infty,0}) \mapsto (T_{-z_0}\omega, z_{-\infty,-1}).$$

Then, on the event where $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\delta_{U^k\eta_-}$ exists define

$$\bar{q}\big(\eta_-, S_z^-\eta_-\big) = q_\mu\big(\eta_-, S_z^-\eta_-\big) \quad \text{for } \mu = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\delta_{U^k\eta_-}. \tag{2.7}$$

On the complement, set $\bar{q}(\eta_-, S_z^-\eta_-) = \delta_{z_0}(z)$, for some fixed $z_0 \in \mathscr{R}$.

Remark 2.2. *Let us also recall the convex analytic characterization of l.s.c. regularization. Let $\mathscr{C}_b(\mathcal{X})$ denote the space of bounded continuous functions on $\mathcal{X}$. Given a function $J : \mathcal{M}_1(\mathcal{X}) \to [0,\infty]$, let $J^* : \mathscr{C}_b(\mathcal{X}) \to \mathbb{R}$ be its* convex conjugate *defined by*

$$J^*(f) = \sup_{\mu\in\mathcal{M}_1(\mathcal{X})}\big\{E^\mu[f] - J(\mu)\big\}$$

and let $J^{**} : \mathcal{M}_1(\mathcal{X}) \to \mathbb{R}$ be its convex biconjugate defined by

$$J^{**}(\mu) = \sup_{f\in\mathscr{C}_b(\mathcal{X})}\big\{E^\mu[f] - J^*(f)\big\}.$$

If J is convex and not identically infinite, J^{**} is the same as its lower semicontinuous regularization J_{lsc}; see Propositions 3.3 *and* 4.1 *of* [8] *or Theorem* 5.18 *of* [15]. *Thus the rate function in Theorem 2.1 is $H_{\mathrm{quen}} = H^{**}$.*

As expected, rate function H has in fact an alternative representation as a specific relative entropy. For a probability measure ν on Ω, define the probability measure $\nu \times P_0^{\cdot}$ on $\boldsymbol{\Omega}_+$ by

$$\int_{\boldsymbol{\Omega}_+} f \,\mathrm{d}\big(\nu \times P_0^{\cdot}\big) = \int_{\Omega} \left[\int_{\mathscr{R}^{\mathbb{N}}} f(\omega, z_{1,\infty}) P_0^{\omega}(\mathrm{d}z_{1,\infty}) \right] \nu(\mathrm{d}\omega).$$

On any of the product spaces of environments and paths, define the σ-algebras $\mathcal{G}_{m,n} = \sigma\{\omega, z_{m,n}\}$. Let $H_{\mathcal{G}_{m,n}}(\alpha|\beta)$ denote the relative entropy of the restrictions of the probability measures α and β to the σ-algebra $\mathcal{G}_{m,n}$. Let Π be the kernel of the environment chain $(T_{X_n}\omega)$, defined as $\Pi f(\omega) = E_0^{\omega}[f(T_{X_1}\omega)] = \sum_z \pi_{0,z}(\omega) f(T_z\omega)$.

Lemma 2.2. *Let $\mu \in \mathcal{M}_1(\Omega_+)$ be S^+-invariant. Then the limit*

$$h\big(\mu|\mu_0 \times P_0^{\cdot}\big) = \lim_{n\to\infty} \frac{1}{n} H_{\mathcal{G}_{1,n}}\big(\mu|\mu_0 \times P_0^{\cdot}\big) \tag{2.8}$$

exists and equals $H(\mu_- \times q_\mu|\mu_- \times p^-)$.

Proof. Fix μ. Let $\mu_i^{\omega,z_{1,i-1}}(\cdot)$ denote the conditional distribution of Z_i under μ, given $\mathcal{G}_{1,i-1}$. Then by the S-invariance,

$$\bar{\mu}[Z_1 = u|\mathcal{G}_{2-i,0}](\omega, z_{2-i,0}) = \mu_i^{T_{x_{1-i}}\omega, z_{2-i,0}}(u).$$

For $i = 1$ we must interpret $\mathcal{G}_{1,0} = \sigma\{\omega\} = \mathfrak{S}$ and $(\omega, z_{1,0})$ simply as ω. Observe also that the conditional distribution of Z_i under $\mu_0 \times P_0^{\cdot}$, given $\mathcal{G}_{1,i-1}$, is $\pi_{0,\boldsymbol{\cdot}}(T_{x_{i-1}}\omega)$.

By two applications of the conditional entropy formula (Lemma 10.3 of [21] or Exercise 6.14 of [15]),

$$\begin{aligned}
H_{\mathcal{G}_{1,n}}\big(\mu|\mu_0 \times P_0^{\cdot}\big) &= \sum_{i=1}^{n} \int H\big(\mu_i^{\omega,z_{1,i-1}} | \pi_{0,\boldsymbol{\cdot}}(T_{x_{i-1}}\omega)\big) \mu(\mathrm{d}\omega, \mathrm{d}z_{1,\infty}) \\
&= \sum_{i=1}^{n} \int H\big(\bar{\mu}[Z_1 = \cdot|\mathcal{G}_{2-i,0}](T_{x_{i-1}}\omega, z_{1,i-1}) | \pi_{0,\boldsymbol{\cdot}}(T_{x_{i-1}}\omega)\big) \mu(\mathrm{d}\omega, \mathrm{d}z_{1,\infty}) \\
&= \sum_{i=1}^{n} \int H\big(\bar{\mu}[Z_1 = \cdot|\mathcal{G}_{2-i,0}](\omega, z_{2-i,0}) | \pi_{0,\boldsymbol{\cdot}}(\omega)\big) \mu_-(\mathrm{d}\omega, \mathrm{d}z_{-\infty,0}) \\
&= \sum_{i=1}^{n} H_{\mathcal{G}_{2-i,1}}\big(\mu_- \times q_\mu | \mu_- \times p^-\big).
\end{aligned} \tag{2.9}$$

As $k \to \infty$, the σ-algebras $\mathcal{G}_{-k,1}$ generate the σ-algebra $\mathcal{G}_{-\infty,1} = \sigma\{\omega, z_{-\infty,1}\}$, and consequently

$$H_{\mathcal{G}_{2-i,1}}\big(\mu_- \times q_\mu|\mu_- \times p^-\big) \nearrow H\big(\mu_- \times q_\mu|\mu_- \times p^-\big) \quad \text{as } i \nearrow \infty. \tag{2.10}$$

We have taken some liberties with notation and regarded $\mu_- \times q_\mu$ and $\mu_- \times p^-$ as measures on the variables $(\omega, z_{-\infty,1})$, instead of on pairs $((\omega, z_{-\infty,0}), (\omega', z'_{-\infty,0}))$. This is legitimate because the simple structure of the kernels q_μ and p^-, namely (2.4) implies that $z'_{-\infty,0} = z_{-\infty,1}$ and $\omega' = T_{z_1}\omega$ almost surely under these measures.

The claim follows by dividing through (2.9) by n and letting $n \to \infty$. $\square$

Note that the specific entropy in (2.8) is not an entropy between two S^+-invariant measures unless μ_0 is Π-invariant. The next lemma exploits the previous one to say something about the zeros of H_{quen}.

Lemma 2.3. *If $H_{\mathrm{quen}}(\mu) = 0$ then $\mu(\mathrm{d}\omega, \mathrm{d}z_{1,\infty}) = \mu_0(\mathrm{d}\omega) P_0^{\omega}(\mathrm{d}z_{1,\infty})$ for some Π-invariant μ_0.*

Note that it is not necessarily true that $\mu_0 \ll \mathbb{P}$ in the above lemma.

Remark 2.3. *One can show that under (2.1) and (2.2) there is at most one $\mathbb{P}_\infty \in \mathcal{M}_1(\Omega)$ that is Π-invariant and such that $\mathbb{P}_\infty \ll \mathbb{P}$; see, for example, [13]. In fact, in this case $\mathbb{P}_\infty \sim \mathbb{P}$. The above lemma shows that the zeros of H_{quen} consist of $P_0^\infty = \mathbb{P}_\infty \times P_0^{\cdot}$ (if $\mathbb{P}_\infty \ll \mathbb{P}$ exists) and possibly measures of the form $\mu_0 \times P_0^{\cdot}$, with μ_0 being Π-invariant but such that $\mu_0 \not\ll \mathbb{P}$.*

Proof of Lemma 2.3. There is a sequence of S^+-invariant probability measures $\mu^{(m)} \to \mu$ such that $H(\mu^{(m)}) \to 0$ and $\mu_0^{(m)} \ll \mathbb{P}$. (If $\mu_0 \ll \mathbb{P}$ then we can take $\mu^{(m)} = \mu$.) Let $\mu_1^{(m)}$ denote the marginal distribution on $(\omega, z_{-\infty,1})$ which can be identified with $\mu_-^{(m)} \times q_{\mu^{(m)}}$ and converges to the corresponding marginal μ_1. By the continuity of the kernel $\pi_{0,z}(\omega)$, $\mu_-^{(m)} \times p^- \to \mu_- \times p^-$. From these limits and the lower semicontinuity of relative entropy,

$$H\big(\mu_1 | \mu_- \times p^-\big) = \lim_{m\to\infty} H\big(\mu_1^{(m)} | \mu_-^{(m)} \times p^-\big) = 0.$$

This tells us that μ_- is p^--invariant, which in turn implies that μ_0 is Π-invariant, and together with the S^+-invariance of μ implies also that $\mu = \mu_0 \times P_0^{\cdot}$. (The last point can also be seen from (2.9) and (2.10).) □

We close this section with some examples.

Let $\Omega = \mathcal{P}^{\mathbb{Z}^d}$ with $\mathcal{P} = \{(p_z)_{z\in\mathscr{R}} \in (0,1)^{\mathscr{R}}\colon \sum_z p_z = 1\}$. Let $\nu \in \boldsymbol{\Omega}_+$ be the law of a classical random walk; i.e. $\nu = \nu_0 \times P_0^{\cdot}$ with $\nu_0 = \delta_\alpha^{\otimes\mathbb{Z}^d}$, for some $\alpha \in \mathcal{P}$. Then $H(\nu_- \times q_\nu | \nu_- \times p^-) = 0$. However, if $\sum_z z\alpha_z$ is not in the set $\mathcal{N} = \{E^\mu[Z_1]\colon H_{\text{quen}}(\mu) = 0\}$, then, $H_{\text{quen}}(\nu) > 0$. Note that by the contraction principle, $\mathcal{N}$ is the zero set of the level-1 rate function. Hence if $\mathbb{P}$ is product, by [22] $\mathcal{N}$ consists of a singleton or a line segment. Thus we can pick α so that the mean $\sum z\alpha_z$ does not lie in $\mathcal{N}$, and consequently we have measures ν for which $H_{\text{quen}}(\nu) > 0 = H(\nu_- \times q_\nu | \nu_- \times p^-)$. That is, the rate H_{quen} does not have to pick up the entropy value.

Lower semicontinuity of relative entropy implies $H_{\text{quen}}(\mu) = H(\mu)$ when $\mu_0 \ll \mathbb{P}$. This equality can still happen when $\mu_0 \not\ll \mathbb{P}$; i.e. the l.s.c. regularization can bring the rate H_{quen} down from infinity all the way to the entropy. Here is a somewhat singular example. Assume $\mathbb{P}\{\pi_{0,0}(\omega) > 0\} = 1$ and let $\zeta = (0,0,0,\ldots)$ be the constant sequence of 0-steps in $\mathbb{Z}^d$. For each $\bar\omega \in \Omega$ define the (trivially S^+-invariant) probability measure $\nu^{\bar\omega} = \delta_{(\bar\omega,\zeta)}$ on $\boldsymbol{\Omega}_+$. Then, for all $\bar\omega$ in the (minimal closed) support of $\mathbb{P}$,

$$H_{\text{quen}}\big(\nu^{\bar\omega}\big) = H\big(\nu_-^{\bar\omega} \times q_{\nu_-^{\bar\omega}} | \nu_-^{\bar\omega} \times p^-\big) = -\log \pi_{0,0}(\bar\omega). \tag{2.11}$$

The case $\pi_{0,0}(\bar\omega) = 0$ is allowed here, which can of course happen if uniform ellipticity is not assumed.

The second equality in (2.11) is clear from definitions, because the kernel is trivial: $q_{\nu_-^{\bar\omega}}(\eta_-, \eta_-) = 1$. Since $H_{\text{quen}}(\nu^{\bar\omega})$ is defined by l.s.c. regularization and entropy itself is l.s.c., entropy always gives a lower bound for H_{quen}. If $\mathbb{P}\{\bar\omega\} > 0$ then $\nu_0^{\bar\omega} = \delta_{\bar\omega} \ll \mathbb{P}$ and the first equality in (2.11) is true by definition. If $\mathbb{P}\{\bar\omega\} = 0$ pick a sequence of open neighborhoods $G_j \searrow \bar\omega$. The assumption that $\bar\omega$ lies in the support of $\mathbb{P}$ implies $\mathbb{P}(G_j) > 0$. Define a sequence of approximating measures by $\mu^j = \frac{1}{\mathbb{P}(G_j)} \int_{G_j} \nu^\omega \mathbb{P}(d\omega)$ with entropies

$$H\big(\mu^j \times q_{\mu^j} | \mu^j \times p^-\big) = -\frac{1}{\mathbb{P}(G_j)} \int_{G_j} \log \pi_{0,0}(\omega) \mathbb{P}(d\omega).$$

The above entropies converge to $-\log \pi_{0,0}(\bar\omega)$ by continuity of $\pi_{0,0}(\cdot)$. We have verified (2.11).

3. Multivariate level 2 and setting the stage for the proofs

The assumptions made for the main result are the union of all the assumptions used in this paper. To facilitate future work, we next list the different assumption that are needed for different parts of the proof.

The lower bounds in Theorem 2.1 above and Theorem 3.1 below do not require Ω compact nor $\mathscr{R}$ finite. They hold under the assumption that $\mathbb{P}$ is ergodic for $\{T_z\colon z \in \mathscr{R}\}$ and the following two conditions are satisfied.

$$\mathbb{P}\{\pi_{0,z} > 0\} \in \{0,1\} \quad \text{for all } z \in \mathbb{Z}^d. \tag{3.1}$$

$$\begin{aligned}&\text{Either } \mathbb{E}[|\log \pi_{0,z}|] < \infty \text{ holds for all } z \in \mathscr{R} \text{ or there exists a}\\&\text{probability measure } \mathbb{P}_\infty \text{ on } (\Omega, \mathfrak{S}) \text{ with } \mathbb{P}_\infty \Pi = \mathbb{P}_\infty \text{ and } \mathbb{P}_\infty \ll \mathbb{P}.\end{aligned} \tag{3.2}$$

Note that (3.1) is a regularity condition that says that either all environments allow the move or all prohibit it.

Our proof of the upper bound uses stricter assumptions. The upper bound holds if $\mathbb{P}$ is ergodic for $\{T_z\colon z \in \mathscr{R}\}$, $\mathscr{R}$ is finite, Ω is compact, the moment assumption (2.3) holds, and

$$\forall x \in \mathscr{R}, \exists m \in \mathbb{N}, \exists z_1, \dots, z_m \in \mathscr{R} \text{ such that } x + z_1 + \cdots + z_m = 0. \tag{3.3}$$

On its own, (3.3) is weaker than (2.2). However, since the additive group generated by $\mathscr{R}$ is isomorphic to $\mathbb{Z}^{d'}$ for some $d' \le d$, we always assume, without any loss of generality, that

$$\mathbb{Z}^d \text{ is the smallest additive group containing } \mathscr{R}. \tag{3.4}$$

Then, under (3.4), (3.3) is equivalent to (2.2).

The only place where the condition $p > d$ (in (2.3)) is needed is for Lemma 5.1 to hold. See Remark 5.3. The only place where (2.2) (or (3.3)) is needed is in the proof of (5.6) in Lemma 5.5. This is the only reason that our result does not cover the so-called *forbidden direction* case. A particularly interesting special case is the space–time, or dynamic, environment; i.e. when $\mathscr{R} \subset \{z\colon z \cdot e_1 = 1\}$. A level 1 quenched LDP can be proved for space–time RWRE through the subadditive ergodic theorem, as was done for elliptic walks in [22]. Yilmaz [23] has shown that for i.i.d. space–time RWRE in 4 and higher dimensions the quenched and averaged level 1 rate functions coincide in a neighborhood of the limit velocity. In contrast with large deviations, the functional central limit theorem of i.i.d. space–time RWRE is completely understood; see [14], and also [1] for a different proof for steps that have exponential tails.

Next we turn to the strategy of the proof of Theorem 2.1. The process level LDP comes by the familiar projective limit argument from large deviation theory. The intermediate steps are multivariate quenched level 2 LDPs. For each $\ell \in \mathbb{N}$ define the multivariate empirical measure

$$R_n^{1,\ell} = n^{-1} \sum_{k=0}^{n-1} \delta_{T_{X_k}\omega, Z_{k+1,k+\ell}}.$$

This empirical measure lives on the space $\boldsymbol{\Omega}_\ell = \Omega \times \mathscr{R}^\ell$ whose generic element is now denoted by $\eta = (\omega, z_{1,\ell})$.

We can treat $R_n^{1,\ell}$ as the position level (level 2) empirical measure of a Feller-continuous Markov chain. Denote by P_η (with expectation E_η) the law of the Markov chain $(\eta_k)_{k\ge 0}$ on $\boldsymbol{\Omega}_\ell$ with initial state η and transition kernel

$$p^+\big(\eta, S_z^+\eta\big) = \pi_{x_\ell, x_\ell+z}(\omega) = \pi_{0,z}(T_{x_\ell}\omega) \quad \text{for } \eta = (\omega, z_{1,\ell}) \in \boldsymbol{\Omega}_\ell,$$

where

$$S_z^+ : \boldsymbol{\Omega}_\ell \to \boldsymbol{\Omega}_\ell : (\omega, z_{1,\ell}) \mapsto (T_{z_1}\omega, z_{2,\ell}, z).$$

This Markov chain has empirical measure

$$L_n = n^{-1} \sum_{k=0}^{n-1} \delta_{\eta_k}$$

that satisfies the following LDP. Define an entropy H_ℓ on $\mathcal{M}_1(\boldsymbol{\Omega}_\ell)$ by

$$H_\ell(\mu) = \begin{cases} \inf\big\{H\big(\mu \times q \,|\, \mu \times p^+\big)\colon q \in \mathcal{Q}(\boldsymbol{\Omega}_\ell) \text{ with } \mu q = \mu\big\} & \text{if } \mu_0 \ll \mathbb{P}, \\ \infty & \text{otherwise.} \end{cases} \tag{3.5}$$

H_ℓ is convex by an argument used below at the end of Section 4. Recall Remark 2.2 about l.s.c. regularization.

Theorem 3.1. *Same assumptions as in Theorem 2.1. For any fixed $\ell \geq 1$, for $\mathbb{P}$-a.e. ω, and for all $z_{1,\ell} \in \mathscr{R}^\ell$, the large deviation principle holds for the sequence of probability measures $P_\eta\{L_n \in \cdot\}$ on $\mathcal{M}_1(\boldsymbol{\Omega}_\ell)$ with convex rate function H_ℓ^{**}.*

The lower bound in Theorem 3.1 follows from a change of measure and the ergodic theorem, and hints at the correct rate function. Donsker and Varadhan's [6] general Markov chain argument gives the upper bound but without the absolute continuity restriction in (3.5). Thus the main issue is to deal with the case when the rate is infinite. This is nontrivial because the set of measures with $\mu_0 \not\ll \mathbb{P}$ is dense in the set of probability measures with the same support as $\mathbb{P}$. This is where the homogenization argument from [12,17,24] comes in.

We conclude this section with a lemma that contains the projective limit step.

Lemma 3.2. *Assume $\mathbb{P} \in \mathcal{M}_1(\Omega)$ is invariant for the shifts $\{T_z\colon z \in \mathscr{R}\}$ and satisfies the regularity assumption (3.1). Assume that for each fixed $\ell \geq 1$ there exists a rate function $I_\ell : \mathcal{M}_1(\boldsymbol{\Omega}_\ell) \to [0, \infty]$ that governs the large deviation lower bound for the laws $P_\eta\{L_n \in \cdot\}$, for $\mathbb{P}$-almost-every ω and all $z_{1,\ell} \in \boldsymbol{\Omega}_\ell$. Then, for $\mathbb{P}$-a.e. ω, the large deviation lower bound holds for $P_0^\omega\{R_n^{1,\infty} \in \cdot\}$ with rate function $I(\mu) = \sup_{\ell \geq 1} I_\ell(\mu_{|\boldsymbol{\Omega}_\ell})$, for $\mu \in \mathcal{M}_1(\boldsymbol{\Omega}_+)$.*

When $\mathscr{R}$ is finite and Ω is compact the same statement holds for the upper bound and the large deviation principle.

Proof. Observe first that P_η is the law of $(T_{X_k}\omega, Z_{k+1,k+\ell})_{k\geq 0}$ under P_0^ω, conditioned on $Z_{1,\ell} = z_{1,\ell}$. Since $P_0^\omega\{Z_{1,\ell} = z_{1,\ell}\} > 0$ $\mathbb{P}$-a.s. we have for all open sets $O \subset \mathcal{M}_1(\boldsymbol{\Omega}_\ell)$,

$$\begin{aligned}\varliminf_{n\to\infty} n^{-1}\log P_0^\omega\big\{R_n^{1,\ell} \in O\big\} &\geq \varliminf_{n\to\infty} n^{-1}\log\big[P_0^\omega\{Z_{1,\ell} = z_{1,\ell}\}P_0^\omega\big\{R_n^{1,\ell} \in O | Z_{1,\ell} = z_{1,\ell}\big\}\big]\\ &= \varliminf_{n\to\infty} n^{-1}\log P_0^\omega\big\{R_n^{1,\ell} \in O | Z_{1,\ell} = z_{1,\ell}\big\} \geq -\inf_O I_\ell.\end{aligned}$$

Similarly, in the case of the upper bound, and when $\mathscr{R}$ is finite, we have for all closed sets $C \subset \mathcal{M}_1(\boldsymbol{\Omega}_\ell)$,

$$\begin{aligned}\varlimsup_{n\to\infty} n^{-1}\log P_0^\omega\big\{R_n^{1,\ell} \in C\big\} &\leq \varlimsup_{n\to\infty} \max_{z_{1,\ell}\in\mathscr{R}^\ell} n^{-1}\log P_0^\omega\big\{R_n^{1,\ell} \in C | Z_{1,\ell} = z_{1,\ell}\big\}\\ &\leq \max_{z_{1,\ell}\in\mathscr{R}^\ell} \varlimsup_{n\to\infty} n^{-1}\log P_0^\omega\big\{R_n^{1,\ell} \in C | Z_{1,\ell} = z_{1,\ell}\big\}\\ &\leq -\inf_C I_\ell.\end{aligned}$$

We conclude that conditioning is immaterial and, $\mathbb{P}$-a.s., the laws of $R_n^{1,\ell}$ induced by P_0^ω satisfy a large deviation lower (resp., upper) bound governed by I_ℓ. The lemma now follows from the Dawson–Gärtner projective limit theorem (see Theorem 4.6.1 in [3]). $\square$

The next two sections prove Theorem 3.1: lower bound in Section 4 and upper bound in Section 5. Section 6 finishes the proof of the main Theorem 2.1.

4. Lower bound

We now prove the large deviation lower bound in Theorem 3.1. This section is valid for a general $\mathscr{R}$ that can be infinite and a general Polish Ω. Lemmas 4.1 and 4.2 are valid under (3.1) only while the lower bound proof also requires (3.2). Recall that assumption (3.4) entails no loss of generality.

We start with some ergodicity properties of the measures involved in the definition of the function H_ℓ. Recall that $\boldsymbol{\Omega}_\ell = \Omega \times \mathscr{R}^\ell$ and that for a measure $\mu \in \mathcal{M}_1(\boldsymbol{\Omega}_\ell)$, μ_0 is its marginal on Ω. Denote by $P_0^{(\ell)}$ the law of $(\omega, Z_{1,\ell})$ under P_0.

Lemma 4.1. *Let $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z\})$ be ergodic and assume (3.1) and (3.4) hold. Fix $\ell \geq 1$ and let $\mu \in \mathcal{M}_1(\boldsymbol{\Omega}_\ell)$ be such that $\mu \ll P_0^{(\ell)}$. Let q be a Markov transition kernel on $\boldsymbol{\Omega}_\ell$ such that*:

(a) μ *is* q*-invariant* (*i.e.*, $\mu q = \mu$);
(b) $q(\eta, S_z^+\eta) > 0$ *for all* $z \in \mathscr{R}$ *and* μ*-a.e.* $\eta \in \boldsymbol{\Omega}_\ell$;
(c) $\sum_{z\in\mathscr{R}} q(\eta, S_z^+\eta) = 1$, *for* μ*-a.e.* $\eta \in \boldsymbol{\Omega}_\ell$.

Then, $\mu \sim P_0^{(\ell)}$ *and the Markov chain* $(\eta_k)_{k\ge 0}$ *on* $\boldsymbol{\Omega}_\ell$ *with kernel* q *and initial distribution* μ *is ergodic. In particular, we have for all* $F \in L^1(\mu)$

$$\lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} E^{Q_\eta}\big[F(\eta_k)\big] = E^\mu[F] \quad \textit{for } \mu\textit{-a.e. } \eta. \tag{4.1}$$

Here, Q_η *is the Markov chain with transition kernel* q *and initial state* η.

Proof. First, let us prove mutual absolute continuity. Let $f = \frac{\mathrm{d}\mu}{\mathrm{d}P_0^{(\ell)}}$. Then, by assumptions (a) and (c),

$$0 = \int \mathbb{1}\{f=0\} f \,\mathrm{d}P_0^{(\ell)} = \int \mathbb{1}\{f=0\}\,\mathrm{d}\mu = \sum_{z\in\mathscr{R}} \int q\big(\eta, S_z^+\eta\big)\mathbb{1}\big\{f\big(S_z^+\eta\big)=0\big\}\mu(\mathrm{d}\eta).$$

By assumption (b), this implies that for $z \in \mathscr{R}$

$$0 = \int \mathbb{1}\big\{f\big(S_z^+\eta\big)=0\big\}\mu(\mathrm{d}\eta) = \int \mathbb{1}\big\{f\big(S_z^+\eta\big)=0\big\} f(\eta) P_0^{(\ell)}(\mathrm{d}\eta).$$

By regularity (3.1) we conclude that $\mathbb{1}\{f(\eta)>0\} \le \mathbb{1}\{f(S_z^+\eta)>0\}$, for all $z\in\mathscr{R}$, $z_{1,\ell}\in\mathscr{R}^\ell$, and $\mathbb{P}$-a.e. ω.

By first following the path $z_{1,\ell}$, then taking an increment of $z\in\mathscr{R}$, then following a path $\tilde z_{1,\ell}\in\mathscr{R}^\ell$, one sees that for all $z_{1,\ell}, \tilde z_{1,\ell} \in \mathscr{R}^\ell$, all $z\in\mathscr{R}$, and $\mathbb{P}$-a.e. ω,

$$\mathbb{1}\big\{f(\omega, z_{1,\ell})>0\big\} \le \mathbb{1}\big\{f(T_{x_\ell+z}\omega, \tilde z_{1,\ell})>0\big\}. \tag{4.2}$$

Now pick a finite subset $\hat z_1, \ldots, \hat z_M \in \mathscr{R}$ that generates $\mathbb{Z}^d$ as an additive group; e.g., take the elements needed for generating the canonical basis $e_1, \ldots, e_d$. Note that $M > d$ can happen; e.g., take $d=1$ and $\mathscr{R} = \{2, 5\}$.

Applying (4.2) repeatedly, one can arrange for z to be any point of the form $\sum_{i=1}^M k_i \hat z_i$ with $k_i \in \mathbb{Z}_+$. Furthermore, the ergodicity of $\mathbb{P}$ under shifts $\{T_z\}$ implies its ergodicity under shifts $\{T_{\hat z_1}, \ldots, T_{\hat z_M}\}$, since the latter generate the former. We can thus average over $k = (k_1, \ldots, k_M) \in [0,n]^M$, take $n\to\infty$, and invoke the multidimensional ergodic theorem (see, for example, Appendix 14.A of [9]). This shows that for all $z_{1,\ell}, \tilde z_{1,\ell}\in\mathscr{R}^\ell$ and $\mathbb{P}$-a.e. ω

$$\mathbb{1}\big\{f(\omega, z_{1,\ell})>0\big\} \le \mathbb{P}\big\{\omega\colon\ f(\omega, \tilde z_{1,\ell})>0\big\}.$$

Since f integrates to 1 there exists a $z_{1,\ell}\in\mathscr{R}^\ell$ with $\mathbb{P}\{f(\omega, z_{1,\ell})>0\}>0$. This implies that $\mathbb{P}\{f(\omega,\tilde z_{1,\ell})>0\}=1$ for all $\tilde z_{1,\ell}\in\mathscr{R}^\ell$ and hence $\mu \sim P_0^{(\ell)}$.

Next, we address the ergodicity issue. By Corollary 2 of Section IV.2 of [16], we have that for any $F\in L^1(\mu)$ and μ-a.e. $\eta\in\boldsymbol{\Omega}_\ell$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} E^{Q_\eta}\big[F(\eta_k)\big] = E^\mu[F|\mathcal{I}_{\mu,q}].$$

Here, $\mathcal{I}_{\mu,q}$ is the σ-algebra of *q-invariant sets*:

$$\left\{A \text{ measurable: } \int q(\eta, A)\mathbb{1}_{A^c}(\eta)\mu(\mathrm{d}\eta) = \int q\big(\eta, A^c\big)\mathbb{1}_A(\eta)\mu(\mathrm{d}\eta) = 0\right\}.$$

Ergodicity would thus follow from showing that $\mathcal{I}_{\mu,q}$ is μ-trivial. To this end, let A be $\mathcal{I}_{\mu,q}$-measurable. By assumptions (b) and (c) and mutual absolute continuity we have that for all $z\in\mathscr{R}$

$$\int \mathbb{1}_A\big(S_z^+\eta\big)\mathbb{1}_{A^c}(\eta) P_0^{(\ell)}(\mathrm{d}\eta) = 0.$$

Replacing the set $\{f > 0\}$ by A^c, in the above proof of mutual absolute continuity, one concludes that $P_0^{(\ell)}(A) \in \{0, 1\}$. The same holds under μ and the lemma is proved. □

We are now ready to derive the lower bound. We first prove a slightly weaker version.

Lemma 4.2. *Let $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z\})$ be ergodic and assume (3.1) and (3.4) hold. Fix $\ell \geq 1$. Then, for $\mathbb{P}$-a.e. ω, for all $z_{1,\ell} \in \mathscr{R}^\ell$, and for any open set $O \subset \mathcal{M}_1(\boldsymbol{\Omega}_\ell)$*

$$\varliminf_{n\to\infty} n^{-1} \log P_\eta\{L_n \in O\} \geq -\inf\bigl\{H\bigl(\mu \times q|\mu \times p^+\bigr)\colon \mu \in O,\ \mu_0 \ll \mathbb{P},\ q \in \mathcal{Q}(\boldsymbol{\Omega}_\ell),\ \mu q = \mu,$$

$$\text{and } \forall z \in \mathscr{R}, q\bigl(\eta, S_z^+\eta\bigr) > 0, \mu\text{-a.s.}\bigr\}.$$

Proof. Fix $\mu \in O$ and q as in the above display. We can also assume that $H(\mu \times q|\mu \times p^+) < \infty$. Then $q(\eta, \{S_z^+\eta\colon z \in \mathscr{R}\}) = 1$ μ-a.s. We can find a weak neighborhood such that $\mu \in B \subset O$. That is, we can find $\varepsilon > 0$, a positive integer m, and bounded continuous functions $f_k : \boldsymbol{\Omega}_\ell \to \mathbb{R}$, such that

$$B = \bigl\{\nu \in \mathcal{M}_1(\boldsymbol{\Omega}_\ell)\colon \forall k \leq m, \bigl|E^\nu[f_k] - E^\mu[f_k]\bigr| < \varepsilon\bigr\}.$$

Let $\mathscr{F}_n$ be the σ-algebra generated by $\eta_0, \ldots, \eta_n$. Recall that Q_η is the law of the Markov chain with initial state η and transition kernel q. Then

$$\begin{aligned}
n^{-1} \log P_\eta\{L_n \in O\} &\geq n^{-1} \log P_\eta\{L_n \in B\} \\
&\geq n^{-1} \log \frac{E^{Q_\eta}[(\mathrm{d}Q_{\eta|\mathscr{F}_{n-1}}/\mathrm{d}P_{\eta|\mathscr{F}_{n-1}})^{-1}\mathbb{1}\{L_n \in B\}]}{Q_\eta\{L_n \in B\}} + n^{-1} \log Q_\eta\{L_n \in B\}
\end{aligned}$$

(by Jensen's inequality, applied to $\log x$)

$$\begin{aligned}
&\geq \frac{-n^{-1} E^{Q_\eta}[\log(\mathrm{d}Q_{\eta|\mathscr{F}_{n-1}}/\mathrm{d}P_{\eta|\mathscr{F}_{n-1}})\mathbb{1}\{L_n \in B\}]}{Q_\eta\{L_n \in B\}} + n^{-1} \log Q_\eta\{L_n \in B\} \\
&= \frac{-n^{-1} E^{Q_\eta}[\log(\mathrm{d}Q_{\eta|\mathscr{F}_{n-1}}/\mathrm{d}P_{\eta|\mathscr{F}_{n-1}})]}{Q_\eta\{L_n \in B\}} + \frac{n^{-1} E^{Q_\eta}[\log(\mathrm{d}Q_{\eta|\mathscr{F}_{n-1}}/\mathrm{d}P_{\eta|\mathscr{F}_{n-1}})\mathbb{1}\{L_n \notin B\}]}{Q_\eta\{L_n \in B\}} \\
&\quad + n^{-1} \log Q_\eta\{L_n \in B\} \\
&= \frac{-n^{-1} H(Q_{\eta|\mathscr{F}_{n-1}}|P_{\eta|\mathscr{F}_{n-1}})}{Q_\eta\{L_n \in B\}} + \frac{n^{-1} E_\eta[\mathrm{d}Q_{\eta|\mathscr{F}_{n-1}}/\mathrm{d}P_{\eta|\mathscr{F}_{n-1}} \log(\mathrm{d}Q_{\eta|\mathscr{F}_{n-1}}/\mathrm{d}P_{\eta|\mathscr{F}_{n-1}})\mathbb{1}\{L_n \notin B\}]}{Q_\eta\{L_n \in B\}} \\
&\quad + n^{-1} \log Q_\eta\{L_n \in B\} \\
&\geq \frac{-n^{-1} H(Q_{\eta|\mathscr{F}_{n-1}}|P_{\eta|\mathscr{F}_{n-1}})}{Q_\eta\{L_n \in B\}} - \frac{n^{-1}\mathrm{e}^{-1}}{Q_\eta\{L_n \in B\}} + n^{-1} \log Q_\eta\{L_n \in B\}.
\end{aligned}$$

In the last inequality we used $x \log x \geq -\mathrm{e}^{-1}$. Observe next that μ and q satisfy the assumptions of Lemma 4.1. Thus, $Q_\eta\{L_n \in B\}$ converges to 1 for μ-a.e. η. Furthermore, if we define

$$F(\eta) = \sum_{z\in\mathscr{R}} q\bigl(\eta, S_z^+\eta\bigr) \log \frac{q(\eta, S_z^+\eta)}{p^+(\eta, S_z^+\eta)} \geq 0 \quad \text{(by Jensen's inequality)}$$

then $E^\mu[F] = H(\mu \times q|\mu \times p^+) < \infty$ and (4.1) implies that for μ-a.e. η

$$\lim_{n\to\infty} n^{-1} H(Q_{\eta|\mathscr{F}_{n-1}}|P_{\eta|\mathscr{F}_{n-1}}) = \lim_{n\to\infty} E^{Q_\eta}\Biggl[n^{-1} \sum_{k=0}^{n-1} F(\eta_k)\Biggr] = E^\mu[F] = H\bigl(\mu \times q|\mu \times p^+\bigr).$$

We have thus shown that

$$\varliminf_{n\to\infty} n^{-1}\log P_\eta\{L_n\in O\}\ge -H\big(\mu\times q|\mu\times p^+\big)$$

for μ-a.e. η. By Lemma 4.1, this is also true $P_0^{(\ell)}$-a.s. □

To prove the lower bound in Theorem 3.1 we next need to remove the positivity restriction on q. This is a simple consequence of convexity.

Proof of the lower bound in Theorem 3.1. Recall our assumption (3.2). If an invariant measure $\mathbb{P}_\infty$ exists, then let $\hat q=p^+$ and $\hat\mu(\mathrm{d}\omega,\mathrm{d}z_{1,\ell})=\mathbb{P}_\infty(\mathrm{d}\omega)P_0^\omega(\mathrm{d}x_{0,\ell})$. If, alternatively, $\mathbb{E}[|\log\pi_{0,z}|]<\infty$, for all $z\in\mathscr{R}$, then set $\hat\pi_z=c\mathrm{e}^{-|z|}/(\mathbb{E}[|\log\pi_{0,z}|]\vee 1)$, where c is chosen so that $\sum_{z\in\mathscr{R}}\hat\pi_z=1$. This ensures that

$$\sum_z\hat\pi_z\mathbb{E}\bigg[\log\frac{\hat\pi_z}{\pi_{0,z}}\bigg]<\infty.$$

In this case, define $\hat\mu(\mathrm{d}\omega,\mathrm{d}z_{1,\ell})=\mathbb{P}(\mathrm{d}\omega)P(\mathrm{d}z_{1,\ell})$, where P is an i.i.d. probability measure with $P\{Z_i=z\}=\hat\pi_z$. Let $\hat q(\eta,S_z^+\eta)=\hat\pi_z$.

Observe that in either case, $\hat\mu\ll P_0^{(\ell)}$, $\hat\mu\hat q=\hat\mu$ and $H(\hat\mu\times\hat q|\hat\mu\times p^+)<\infty$.

Let $\mu\in O$ be such that $\mu_0\ll\mathbb{P}$. By (3.1), $\mu\ll P_0^{(\ell)}$. Let q be such that μ is q-invariant and $H(\mu\times q|\mu\times p^+)<\infty$.

Fix $\varepsilon\in(0,1)$ and define $\mu_\varepsilon=\varepsilon\hat\mu+(1-\varepsilon)\mu$. For $\varepsilon>0$ small enough, this measure belongs to the open set O. It is also clear that $\mu_\varepsilon\ll P_0^{(\ell)}$. Let $f_\varepsilon=\frac{\mathrm{d}\mu}{\mathrm{d}\mu_\varepsilon}$ and $\hat f_\varepsilon=\frac{\mathrm{d}\hat\mu}{\mathrm{d}\mu_\varepsilon}$. Note that Lemma 4.1 implies that $\hat\mu\sim P_0^{(\ell)}$. Thus, $\hat\mu\sim\mu_\varepsilon$ and $\mu_\varepsilon\{\hat f_\varepsilon>0\}=1$. Next, define the kernel

$$q_\varepsilon\big(\eta,S_z^+\eta\big)=\varepsilon\hat f_\varepsilon(\eta)\hat q\big(\eta,S_z^+\eta\big)+(1-\varepsilon)f_\varepsilon(\eta)q\big(\eta,S_z^+\eta\big).$$

Then, μ_ε-a.s., $\sum_{z\in\mathscr{R}}q_\varepsilon(\eta,S_z^+\eta)=1$ and $q_\varepsilon(\eta,S_z^+\eta)>0$ for all $z\in\mathscr{R}$. Furthermore, $\mu_\varepsilon q_\varepsilon=\mu_\varepsilon$. Indeed,

$$\begin{aligned}
&\sum_{z\in\mathscr{R}}\int G\big(S_z^+\eta\big)\big[\varepsilon\hat f_\varepsilon(\eta)\hat q\big(\eta,S_z^+\eta\big)+(1-\varepsilon)f_\varepsilon(\eta)q\big(\eta,S_z^+\eta\big)\big]\mu_\varepsilon(\mathrm{d}\zeta)\\
&\quad=\varepsilon\sum_{z\in\mathscr{R}}\int G\big(S_z^+\eta\big)\hat q\big(\eta,S_z^+\eta\big)\hat\mu(\mathrm{d}\eta)+(1-\varepsilon)\sum_{z\in\mathscr{R}}\int G\big(S_z^+\eta\big)q\big(\eta,S_z^+\eta\big)\mu(\mathrm{d}\eta)\\
&\quad=\varepsilon\int G\,\mathrm{d}\hat\mu+(1-\varepsilon)\int G\,\mathrm{d}\mu=\int G\,\mathrm{d}\mu_\varepsilon.
\end{aligned}$$

On the other hand, Jensen's inequality (applied to $x\log x$) implies

$$\begin{aligned}
H\big(\mu_\varepsilon\times q_\varepsilon|\mu_\varepsilon\times p^+\big)&=\sum_z\int q_\varepsilon\big(\eta,S_z^+\eta\big)\log\frac{q_\varepsilon(\eta,S_z^+\eta)}{p^+(\eta,S_z^+\eta)}\mu_\varepsilon(\mathrm{d}\eta)\\
&\le\sum_z\int\varepsilon\hat f_\varepsilon(\eta)\hat q\big(\eta,S_z^+\eta\big)\log\frac{\hat q(\eta,S_z^+\eta)}{p^+(\eta,S_z^+\eta)}\mu_\varepsilon(\mathrm{d}\eta)\\
&\quad+\sum_z\int(1-\varepsilon)f_\varepsilon(\eta)q\big(\eta,S_z^+\eta\big)\log\frac{q(\eta,S_z^+\eta)}{p^+(\eta,S_z^+\eta)}\mu_\varepsilon(\mathrm{d}\eta)\\
&=\varepsilon\sum_z\int\hat q\big(\eta,S_z^+\eta\big)\log\frac{\hat q(\eta,S_z^+\eta)}{p^+(\eta,S_z^+\eta)}\hat\mu(\mathrm{d}\eta)\\
&\quad+(1-\varepsilon)\sum_z\int q\big(\eta,S_z^+\eta\big)\log\frac{q(\eta,S_z^+\eta)}{p^+(\eta,S_z^+\eta)}\mu(\mathrm{d}\eta)\\
&=\varepsilon H\big(\hat\mu\times\hat q|\hat\mu\times p^+\big)+(1-\varepsilon)H\big(\mu\times q|\mu\times p^+\big).
\end{aligned}$$

Since $H(\hat\mu\times\hat q\,|\,\hat\mu\times p^+)<\infty$, applying Lemma 4.2 and then taking $\varepsilon\to 0$ proves the lower bound in Theorem 3.1 with function H_ℓ. The argument above can also be used to show that H_ℓ is convex. Thus the lower bound also holds with H_ℓ^{**}. $\square$

5. Upper bound

To motivate the complicated upper bound proof we first present a simple version of it that works for a finite Ω, which is the case of a periodic environment. In this case, the upper bound only requires the regularity assumption (3.1). Note also that the finiteness of Ω implies the existence of $\mathbb{P}_\infty$ as in assumption (3.2), and hence the lower bound (and, consequently, the large deviation principle) also holds under only (3.1).

Fix $\ell\ge 1$. Given bounded continuous functions h and f on $\boldsymbol{\Omega}_\ell$ define

$$K_{\ell,h}(f)=\mathbb{P}\text{-}\operatorname*{ess\,sup}_{\omega}\sup_{z_{1,\ell}}\log\sum_z p^+\bigl(\eta,S_z^+\eta\bigr)\mathrm{e}^{f(\eta)-h(\eta)+h(S_z^+\eta)}.$$

Define $\underline{K}_\ell:\mathscr{C}_b(\boldsymbol{\Omega}_\ell)\to\mathbb{R}$ by

$$\underline{K}_\ell(f)=\inf_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)}K_{\ell,h}(f).$$

A small modification of Donsker and Varadhan's argument in [6], given below in Lemma 5.2, shows that for $\mathbb{P}$-a.e. ω and all $z_{1,\ell}\in\mathscr{R}^\ell$ one has, for all compact sets $C\subset\mathcal{M}_1(\boldsymbol{\Omega}_\ell)$,

$$\overline{\lim_{n\to\infty}}\,n^{-1}\log P_\eta\{L_n\in C\}\le-\inf_{\mu\in C}\underline{K}_\ell^*(\mu),$$

where $\underline{K}_\ell^*(\mu)=\sup_{f\in\mathscr{C}_b(\boldsymbol{\Omega}_-)}\{E^\mu[f]-\underline{K}_\ell(f)\}$ is the convex conjugate of $\underline{K}_\ell$. Now we observe what it takes to turn this rate function $\underline{K}_\ell^*$ into H_ℓ^{**} and thereby match the upper and lower bounds.

First

$$\begin{aligned}\underline{K}_\ell(f)&=\inf_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)}\mathbb{P}\text{-}\operatorname*{ess\,sup}_{\omega}\sup_{z_{1,\ell}}\log\sum_z p^+\bigl(\eta,S_z^+\eta\bigr)\mathrm{e}^{f(\eta)-h(\eta)+h(S_z^+\eta)}\\&=\inf_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)}\sup_{\mu:\,\mu_0\ll\mathbb{P}}\bigl\{E^\mu[f]-E^\mu\bigl[h-\log p^+\bigl(\mathrm{e}^h\bigr)\bigr]\bigr\}.\end{aligned}\tag{5.1}$$

On the other hand, given $\mu,\nu\in\mathcal{M}_1(\boldsymbol{\Omega}_\ell)$, we have this variational formula:

$$\inf\bigl\{H\bigl(\alpha|\alpha_1\times p^+\bigr):\ \alpha\in\mathcal{M}_1\bigl(\boldsymbol{\Omega}_\ell^2\bigr),\alpha_1=\mu,\alpha_2=\nu\bigr\}=\sup_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)}\bigl\{E^\nu[h]-E^\mu\bigl[\log p^+\bigl(\mathrm{e}^h\bigr)\bigr]\bigr\},$$

where α_1 and α_2 are the first and second marginals of α (see Theorem 2.1 of [7], Lemma 2.19 of [19] or Theorem 13.1 of [15]). Out of this we get

$$\begin{aligned}H_\ell^*(f)&=\sup_{\mu:\,\mu_0\ll\mathbb{P}}\bigl\{E^\mu[f]-\inf\bigl\{H\bigl(\mu\times q|\mu\times p^+\bigr):\ \mu q=\mu\bigr\}\bigr\}\\&=\sup_{\mu:\,\mu_0\ll\mathbb{P}}\Bigl\{E^\mu[f]-\inf_{\alpha\in\mathcal{M}_1(\boldsymbol{\Omega}_\ell^2)}\bigl\{H\bigl(\alpha|\alpha_1\times p^+\bigr):\ \alpha_1=\alpha_2=\mu\bigr\}\Bigr\}\\&=\sup_{\mu:\,\mu_0\ll\mathbb{P}}\Bigl\{E^\mu[f]-\sup_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)}E^\mu\bigl[h-\log p^+\bigl(\mathrm{e}^h\bigr)\bigr]\Bigr\}\\&=\sup_{\mu:\,\mu_0\ll\mathbb{P}}\inf_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)}\bigl\{E^\mu[f]-E^\mu\bigl[h-\log p^+\bigl(\mathrm{e}^h\bigr)\bigr]\bigr\}.\end{aligned}\tag{5.2}$$

Comparison of (5.1) and (5.2) shows that matching $\underline{K}_\ell$ and H_ℓ^*, and thereby completing the upper bound of Theorem 3.1, boils down to an application of a minimax theorem (such as König's theorem, see [11] or [15]). However, the set $\{\mu:\ \mu_0\ll\mathbb{P}\}$ is compact if, and only if, $\mathbb{P}$ has finite support.

To get around this difficulty we abandon the attempt to prove the equality of $\underline{K}_\ell$ and H^*_ℓ. Instead, we redefine $\underline{K}_\ell$ by taking infimum over a larger set of functions. This decreases $\underline{K}_\ell$ and makes it possible to prove $H^*_\ell \geq \underline{K}_\ell$. We will still be able to prove that $H^*_\ell \leq \underline{K}_\ell$ and that $\underline{K}^*_\ell$ governs the large deviation upper bound. The new definition extends the class of functions to include weak limits of $h_k(S^+_z\eta) - h_k(\eta)$, which may lose this form. Such limits are the so-called "corrector functions," familiar from quenched central limit theorems for random walk in random environment (see, for example, [14] and the references therein). Let us introduce this class of functions and redefine $\underline{K}_\ell$.

Definition 5.1. *A measurable function $F : \boldsymbol{\Omega}_\ell \times \mathscr{R} \to \mathbb{R}$ is in class $\mathcal{K}^p(\boldsymbol{\Omega}_\ell \times \mathscr{R})$ if it satisfies the following three conditions*:

(i) *Moment*: *for each $z_{1,\ell} \in \mathscr{R}^\ell$ and $z \in \mathscr{R}$, $\mathbb{E}[|F(\omega, z_{1,\ell}, z)|^p] < \infty$.*

(ii) *Mean zero*: *for all $n \geq \ell$ and $\{a_i\}_{i=1}^n \in \mathscr{R}^n$ the following holds. If $\eta_0 = (\omega, a_{n-\ell+1,n})$ and $\eta_i = S^+_{a_i}\eta_{i-1}$ for $i = 1, \ldots, n$, then*

$$\mathbb{E}\left[\sum_{i=0}^{n-1} F(\eta_i, a_{i+1})\right] = 0.$$

In other words, expectation vanishes whenever the sequence of moves $S^+_{a_1}, \ldots, S^+_{a_n}$ takes $(\omega, z_{1,\ell})$ to $(T_x\omega, z_{1,\ell})$ for all ω, for fixed x and $z_{1,\ell}$.

(iii) *Closed loop*: *for $\mathbb{P}$-a.e. ω and any two paths $\{\eta_i\}_{i=0}^n$ and $\{\bar\eta_j\}_{j=0}^m$ with $\eta_0 = \bar\eta_0 = (\omega, z_{1,\ell})$, $\eta_n = \bar\eta_m$, $\eta_i = S^+_{a_i}\eta_{i-1}$, and $\bar\eta_j = S^+_{\bar a_j}\bar\eta_{j-1}$, for $i, j > 0$ and some $\{a_i\}_{i=1}^n \in \mathscr{R}^n$ and $\{\bar a_j\}_{j=1}^m \in \mathscr{R}^m$, we have*

$$\sum_{i=0}^{n-1} F(\eta_i, a_{i+1}) = \sum_{j=0}^{m-1} F(\bar\eta_j, \bar a_{j+1}).$$

Remark 5.1. *In* (iii) *above, if one has a loop ($\eta_0 = \eta_n$), then one can take $m = 0$ and the right-hand side in the above display vanishes.*

Remark 5.2. *Note that functions $F(\eta, z) = h(S^+_z\eta) - h(\eta)$ belong to this class.*

The following sublinear growth property is crucial. We postpone its proof to the Appendix.

Lemma 5.1. *Let $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z\})$ be ergodic. Assume $\mathbb{P}$ satisfies assumptions* (2.1) *and* (3.4). *Let $F \in \mathcal{K}^p(\boldsymbol{\Omega}_\ell \times \mathscr{R})$ with $p > d$ being the same as in assumption* (2.3). *Then, for $\mathbb{P}$-a.e. ω*

$$\lim_{n\to\infty} n^{-1} \sup_{z_{1,\ell}\in\mathscr{R}^\ell} \sup_{\substack{(a_1,\ldots,a_n)\in\mathscr{R}^n \\ \eta_0=(\omega,z_{1,\ell}),\eta_i=S^+_{a_i}\eta_{i-1}}} \left|\sum_{k=0}^{n-1} F(\eta_k, a_{k+1})\right| = 0.$$

Remark 5.3. *The above lemma clarifies why the method we use requires the condition $p > d$. Indeed, consider the case $\ell = 0$, $\Omega = \mathcal{P}^{\mathbb{Z}^d}$, $\mathbb{P}$ a product measure, and $F(\omega, z) = h(T_z\omega) - h(\omega)$ with h being a function of just ω_0. Then, the conclusion of the lemma is that $n^{-1}\sup_{|x|\leq n}|h(\omega_x)|$ vanishes at the limit. For this to happen one needs more than d moments for h.*

Now, for $F \in \mathcal{K}^p(\boldsymbol{\Omega}_\ell \times \mathscr{R})$ and $f \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)$, redefine

$$K_{\ell,F}(f) = \mathbb{P}\text{-ess}\sup_\omega \sup_{z_{1,\ell}} \log \sum_z p^+(\eta, S^+_z\eta)\mathrm{e}^{f(\eta)+F(\eta,z)}.$$

Redefine $\underline{K}_\ell : \mathscr{C}_b(\boldsymbol{\Omega}_\ell) \to \mathbb{R}$ by

$$\underline{K}_\ell(f) = \inf_{F\in\mathcal{K}^p(\boldsymbol{\Omega}_\ell\times\mathscr{R})} K_{\ell,F}(f).$$

Lemma 5.2. *Assume the conclusion of Lemma 5.1 holds. For $\mathbb{P}$-a.e. ω and all $z_{1,\ell}$, for all compact sets $C \subset \mathcal{M}_1(\boldsymbol{\Omega}_\ell)$,*

$$\varlimsup_{n\to\infty} n^{-1} \log P_\eta\{L_n \in C\} \le -\inf_{\mu\in C} \underline{K}^*_\ell(\mu),$$

where $\underline{K}^*_\ell(\mu) = \sup_{f\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)}\{E^\mu[f] - \underline{K}_\ell(f)\}$ is the convex conjugate of $\underline{K}_\ell$.

Proof. Fix $\mu \in C$ and $c < \inf_C \underline{K}^*_\ell$. There exist $f \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)$ and $F \in \mathcal{K}^p(\boldsymbol{\Omega}_\ell \times \mathscr{R})$ such that $E^\mu[f] - K_{\ell,F}(f) > c$. Fix $\varepsilon > 0$ and define the neighborhood

$$B_\varepsilon(\mu) = \bigl\{\nu \in \mathcal{M}_1(\boldsymbol{\Omega}_\ell)\colon \bigl|E^\nu[f] - E^\mu[f]\bigr| < \varepsilon\bigr\}.$$

Lemma 5.1 implies that for $\mathbb{P}$-a.e. ω there exists a finite $c_\varepsilon(\omega) > 0$ such that for all n and $z_{1,\ell} \in \mathscr{R}^\ell$,

$$\sum_{k=0}^{n-1} F(\eta_k, Z_{k+\ell+1}) \ge -c_\varepsilon - n\varepsilon, \qquad P_\eta\text{-a.s.}$$

Therefore, for all $z_{1,\ell} \in \mathscr{R}^\ell$ and $\mathbb{P}$-a.e. ω,

$$\begin{aligned}
P_\eta\{L_n \in B_\varepsilon\} &= E_\eta\bigl[e^{nL_n(f)}e^{-nL_n(f)}\mathbb{1}\{L_n \in B_\varepsilon\}\bigr]\\
&\le e^{c_\varepsilon+n\varepsilon}e^{-n\inf_{\nu\in B_\varepsilon} E^\nu[f]}E_\eta\Biggl[\exp\Biggl\{-c_\varepsilon - n\varepsilon + \sum_{k=0}^{n-1} f(\eta_k)\Biggr\}\Biggr]\\
&\le e^{c_\varepsilon+n\varepsilon}e^{-nE^\mu[f]}e^{n\varepsilon}E_\eta\Biggl[\exp\Biggl\{\sum_{k=0}^{n-1}\bigl(f(\eta_k) + F(\eta_k, Z_{k+\ell+1})\bigr)\Biggr\}\Biggr]\\
&= e^{c_\varepsilon+n\varepsilon}e^{-nE^\mu[f]}e^{n\varepsilon}E_\eta\Biggl[E_\eta\Biggl[\exp\Biggl\{\sum_{k=0}^{n-1}\bigl(f(\eta_k) + F(\eta_k, Z_{k+\ell+1})\bigr)\Biggr\}\Biggm|\eta_i\colon i \le n-1\Biggr]\Biggr]\\
&= e^{c_\varepsilon+n\varepsilon}e^{-nE^\mu[f]}e^{n\varepsilon}E_\eta\Biggl[\exp\Biggl\{\sum_{k=0}^{n-2}\bigl(f(\eta_k) + F(\eta_k, Z_{k+\ell+1})\bigr)\Biggr\}E_{\eta_{n-1}}\bigl[e^{f(\eta_0)+F(\eta_0,Z_{\ell+1})}\bigr]\Biggr]\\
&\le e^{c_\varepsilon+n\varepsilon}e^{-nE^\mu[f]}e^{n\varepsilon}e^{K_{\ell,F}(f)}E_\eta\Biggl[\exp\Biggl\{\sum_{k=0}^{n-2}\bigl(f(\eta_k) + F(\eta_k, Z_{k+\ell+1})\bigr)\Biggr\}\Biggr]\\
&\le \cdots \le e^{c_\varepsilon+n\varepsilon}e^{-nE^\mu[f]}e^{n\varepsilon}e^{nK_{\ell,F}(f)} \le e^{c_\varepsilon+2n\varepsilon-cn}.
\end{aligned}$$

Since C is compact, it can be covered by a finite collection of $B_\varepsilon(\mu_i)$'s and

$$\varlimsup_{n\to\infty} n^{-1}\log P_\eta\{L_n \in C\} \le -c + 2\varepsilon.$$

Thus, taking $\varepsilon \to 0$ and c to $\inf_C \underline{K}^*_\ell$ proves the lemma for a compact C. □

Our next theorem gives the connection between $\underline{K}_\ell$ and H_ℓ.

Theorem 5.3. *Same assumptions as in Theorem 3.1. Then, $H^*_\ell \equiv \underline{K}_\ell$ for all $\ell \ge 1$.*

We are now ready to prove the above theorem and finish the proof of Theorem 3.1.

Proof of Theorem 5.3 and the upper bound in Theorem 3.1. It suffices to prove that for bounded continuous functions f,

$$\underline{K}_\ell(f) \le H^*_\ell(f) = \sup_\mu\bigl\{E^\mu[f] - H_\ell(\mu)\bigr\}. \tag{5.3}$$

Indeed, this would imply that $H_\ell^{**} \le \underline{K}_\ell^*$ and Lemma 5.2 implies then the upper bound in Theorem 3.1. Furthermore, due to the lower bound and the uniqueness of the rate function (see Theorem 2.18 of [15]), we in fact have that $H_\ell^{**} = \underline{K}_\ell^*$. This implies that $H_\ell^* = \underline{K}_\ell^{**}$ and since $\underline{K}_\ell$ is convex and continuous in the uniform norm, we have that $H_\ell^* = \underline{K}_\ell$.

Let us now prove (5.3). This is trivial when $H_\ell^*(f) = \infty$. Assume thus that $H_\ell^*(f) < \infty$.

Let $\mathfrak{S}_k$ be an increasing sequence of finite σ-algebras on Ω, generating $\mathfrak{S}$. Assume that for all $k \ge 1$ and $y \in \mathscr{R}$, $T_y \mathfrak{S}_{k-1} \subset \mathfrak{S}_k$. Let $\mathcal{M}_1^k = \mathcal{M}_1^k(\boldsymbol{\Omega}_\ell)$ be the set of probability measures μ on $\boldsymbol{\Omega}_\ell$ such that $\mu_0 \ll \mathbb{P}$ and $\frac{\mathrm{d}\mu_0}{\mathrm{d}\mathbb{P}}$ is $\mathfrak{S}_k$-measurable. Now write

$$H_\ell^*(f) = \sup_{\mu:\ \mu_0 \ll \mathbb{P}} \bigl\{ E^\mu[f] - H_\ell(\mu) \bigr\} \ge \sup_{\mu \in \mathcal{M}_1^k} \bigl\{ E^\mu[f] - H_\ell(\mu) \bigr\}.$$

To conclude the proof of (5.3), one invokes the following lemma.

Lemma 5.4. *Same assumptions as in Theorem 3.1. Fix $\ell \ge 1$ and let $f \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)$ and $A < \infty$ be such that*

$$A \ge \sup_{\mu \in \mathcal{M}_1^k} \bigl\{ E^\mu[f] - H_\ell(\mu) \bigr\}$$

for all $k \ge 1$. Then, $A \ge \underline{K}_\ell(f)$.

Proof. Let $\mathcal{M}_1^{k,2}$ be the set of probability measures α on $\boldsymbol{\Omega}_\ell^2$ such that the first $\boldsymbol{\Omega}_\ell$-marginal $\alpha_1 \in \mathcal{M}_1^k$. Observe next that if $\alpha \in \mathcal{M}_1(\boldsymbol{\Omega}_\ell^2)$ is such that $\alpha_1 \ne \alpha_2$, then

$$\inf_{h \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \bigl\{ E^{\alpha_2}[h] - E^{\alpha_1}[h] \bigr\} = -\infty.$$

Write

$$\begin{aligned}
A &\ge \sup_{\mu \in \mathcal{M}_1^k} \bigl\{ E^\mu[f] - \inf\bigl\{ H\bigl(\mu \times q | \mu \times p^+\bigr):\ \mu q = \mu \bigr\}\bigr\} \\
&= \sup\bigl\{ E^{\alpha_1}[f] - H\bigl(\alpha | \alpha_1 \times p^+\bigr):\ \alpha \in \mathcal{M}_1^{k,2}, \alpha_1 = \alpha_2 \bigr\} \\
&\ge \sup_{\alpha \in \mathcal{M}_1^{k,2}} \inf_{h \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \bigl\{ E^{\alpha_1}[f] + E^{\alpha_2}[h] - E^{\alpha_1}[h] - H\bigl(\alpha | \alpha_1 \times p^+\bigr) \bigr\}.
\end{aligned}$$

Since the quantity in braces is linear (and hence continuous and convex) in h and concave and upper semicontinuous in α, and since $\mathcal{M}_1^{k,2}$ is compact, we can apply König's minimax theorem; see [11]. Then

$$\begin{aligned}
A &\ge \inf_{h \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \sup_{\alpha \in \mathcal{M}_1^{k,2}} \bigl\{ E^{\alpha_1}[f] + E^{\alpha_2}[h] - E^{\alpha_1}[h] - H\bigl(\alpha | \alpha_1 \times p^+\bigr) \bigr\} \\
&= \inf_{h \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \sup_{\mu \in \mathcal{M}_1^k} \sup_{q \in \mathcal{Q}(\boldsymbol{\Omega}_\ell)} \int \bigl[f(\eta) + qh(\eta) - h(\eta) - H\bigl(q(\eta,\cdot) | p^+(\eta,\cdot)\bigr) \bigr] \mu(\mathrm{d}\eta) \\
&= \inf_{h \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \sup_{\mu \in \mathcal{M}_1^k} E^\mu\bigl[f - h + \log p^+\bigl(\mathrm{e}^h\bigr) \bigr] \\
&\ge \inf_{h \in \mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \sup_{z_{1,\ell} \in \mathscr{R}^\ell} \sup \biggl\{ E^\mu\bigl[f - h + \log p^+\bigl(\mathrm{e}^h\bigr) \bigr]:\ \mu = \mu_0 \otimes \delta_{z_{1,\ell}} \text{ and } \frac{\mathrm{d}\mu_0}{\mathrm{d}\mathbb{P}} \text{ is } \mathfrak{S}_k \text{ measurable} \biggr\}.
\end{aligned}$$

In the last equality above we passed the sup under the integral, since the integrand is a function of $q(\eta,\cdot)$ and one can maximize for each η separately. Then we used the variational characterization of relative entropy; see Lemma 10.1

in [21] or Theorem 6.7 in [15]. We thus have

$$A \ge \inf_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \sup_{z_{1,\ell}\in\mathscr{R}^\ell} \mathbb{P}\text{-ess}\sup_\omega \mathbb{E}\bigl[f-h+\log p^+\bigl(\mathrm{e}^h\bigr)|\mathfrak{S}_k\bigr]$$

$$= \inf_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \sup_{z_{1,\ell}\in\mathscr{R}^\ell} \mathbb{P}\text{-ess}\sup_\omega \mathbb{E}\Bigl[\log\sum_z p^+\bigl(\eta, S_z^+\eta\bigr)\mathrm{e}^{f(\eta)-h(\eta)+h(S_z^+\eta)}|\mathfrak{S}_k\Bigr].$$

Let $\nu\in\mathcal{M}_1(\mathscr{R})$ with $\nu(z)>0$ for all $z\in\mathscr{R}$. Write the last conditional expectation as

$$\mathbb{E}\Bigl[\log\sum_z \nu(z)\exp\bigl\{\log\bigl[\nu(z)^{-1}p^+\bigl(\eta, S_z^+\eta\bigr)\bigr]+f(\eta)-h(\eta)+h\bigl(S_z^+\eta\bigr)\bigr\}|\mathfrak{S}_k\Bigr].$$

An application of an infinite-dimensional version of Jensen's inequality (see Lemma A.1) and cancelling the $\nu(z)$-factors gives

$$A \ge \inf_{h\in\mathscr{C}_b(\boldsymbol{\Omega}_\ell)} \sup_{z_{1,\ell}\in\mathscr{R}^\ell} \mathbb{P}\text{-ess}\sup_\omega \Bigl\{\log\sum_z \mathrm{e}^{\mathbb{E}[\log p^+(\eta,S_z^+\eta)+f(\eta)-h(\eta)+h(S_z^+\eta)|\mathfrak{S}_k]}\Bigr\}.$$

The above means that for $\varepsilon>0$ and $k\ge 1$ there exists a bounded continuous function $h_{k,\varepsilon}$ such that for all $z_{1,\ell}\in\mathscr{R}^\ell$ and $\mathbb{P}$-a.s.

$$A+\varepsilon \ge \log\sum_z \mathrm{e}^{\mathbb{E}[f(\eta)+\log p^+(\eta,S_z^+\eta)-h_{k,\varepsilon}(\eta)+h_{k,\varepsilon}(S_z^+\eta)|\mathfrak{S}_k]}. \tag{5.4}$$

Next, we show that the sequence

$$F_{k,\varepsilon}(\eta,z)=\mathbb{E}\bigl[h_{k,\varepsilon}\bigl(S_z^+\eta\bigr)-h_{k,\varepsilon}(\eta)|\mathfrak{S}_{k-1}\bigr] \tag{5.5}$$

is uniformly bounded in $L^p(\mathbb{P})$, for any fixed $z_{1,\ell}$ and z. Hence, along a subsequence, $F_{k,\varepsilon}$ converges in the $L^p(\mathbb{P})$ weak topology to some $F_\varepsilon\in L^p(\mathbb{P})$. We can in fact use the same subsequence for all $z_{1,\ell}$ and z. We will still call this subsequence $(F_{k,\varepsilon})$. One can also directly check that $F_\varepsilon\in\mathcal{K}^p(\boldsymbol{\Omega}_\ell\times\mathscr{R})$. In order not to interrupt the flow we postpone the proof of these two facts to Lemma 5.5 below.

On the other hand,

$$M_k(\eta,z)=\mathbb{E}\bigl[f(\eta)+\log p^+\bigl(\eta,S_z^+\eta\bigr)|\mathfrak{S}_{k-1}\bigr]$$

is a martingale whose $L^p(\mathbb{P})$-norm is uniformly bounded. It thus converges in $L^p(\mathbb{P})$ (as well as almost-surely) to $f(\eta)+\log p^+(\eta,S_z^+\eta)$, for all $z_{1,\ell}$ and z. Thus, by Theorem 3.13 of [18], for each fixed $z_{1,\ell}$ and z, there exists a sequence of random variables $g_{k,\varepsilon}(\eta,z)$ that converges strongly in L^p (and thus a subsequence converges $\mathbb{P}$-a.s.) to $f(\eta)+\log p^+(\eta,S_z^+\eta)+F_\varepsilon(\eta,z)$ and such that $g_{k,\varepsilon}$ is a convex combination of $\{M_j+F_{j,\varepsilon}\colon j\le k\}$. One can then extract a further subsequence that converges $\mathbb{P}$-a.s. for all $z_{1,\ell}$ and z.

By Jensen's inequality, we have for all $z_{1,\ell}\in\mathscr{R}^\ell$ and $\mathbb{P}$-a.s.

$$\mathrm{e}^{A+\varepsilon}\ge\sum_{z\in\mathscr{R}}\mathbb{E}\bigl[\mathrm{e}^{\mathbb{E}[f(\eta)+\log p^+(\eta,S_z^+\eta)-h_{k,\varepsilon}(\eta)+h_{k,\varepsilon}(S_z^+\eta)|\mathfrak{S}_k]}|\mathfrak{S}_{k-1}\bigr]\ge\sum_{z\in\mathscr{R}}\mathrm{e}^{M_k(\eta,z)+F_{k,\varepsilon}(\eta,z)}.$$

Since this is valid for all $k\ge 1$, another application of Jensen's inequality gives

$$\mathrm{e}^{A+\varepsilon}\ge\sum_{z\in\mathscr{R}}\mathrm{e}^{g_{k,\varepsilon}(\eta,z)}.$$

Taking $k\to\infty$ implies, for $\mathbb{P}$-a.e. ω and all $z_{1,\ell}\in\mathscr{R}^\ell$,

$$A+\varepsilon\ge f(\eta)+\log\sum_{z\in\mathscr{R}}p^+\bigl(\eta,S_z^+\eta\bigr)\mathrm{e}^{F_\varepsilon(\eta,z)}$$

and thus

$$A+\varepsilon \geq \inf_{F\in\mathcal{K}^p(\boldsymbol{\Omega}_\ell\times\mathscr{R})} \sup_{z_{1,\ell}\in\mathscr{R}^\ell} \mathbb{P}\text{-ess}\sup_\eta \left\{ f(\eta)+\log\sum_{z\in\mathscr{R}} p^+\big(\eta, S_z^+\eta\big)\mathrm{e}^{F(\eta,z)}\right\}.$$

Taking $\varepsilon\to 0$ proves that $A\geq \underline{K}_\ell$. □

Lemma 5.5. *Assume $(\Omega,\mathfrak{S},\mathbb{P},\{T_z\})$ is ergodic. Assume $\mathbb{P}$ satisfies assumptions (2.1), (2.2) and (2.3). Then, for $\varepsilon>0$, $z_{1,\ell}\in\mathscr{R}^\ell$ and $z\in\mathscr{R}$,*

$$\sup_k \mathbb{E}\big[\big|F_{k,\varepsilon}(\omega,z_{1,\ell},z)\big|^p\big]<\infty. \tag{5.6}$$

Moreover, if a subsequence converges (in weak $L^p(\mathbb{P})$-topology), for each $z_{1,\ell}$ and z, to a limit F_ε, then F_ε belongs to class $\mathcal{K}^p(\boldsymbol{\Omega}_\ell\times\mathscr{R})$.

Proof. Since f is bounded (5.4) implies that for $\mathbb{P}$-a.e. ω and for all $z_{1,\ell}\in\mathscr{R}^\ell$

$$F_{k,\varepsilon}(\eta,z)\leq C-\mathbb{E}\big[\log p^+\big(\eta,S_z^+\eta\big)|\mathfrak{S}_{k-1}\big].$$

The $L^p(\mathbb{P})$-norm of the right-hand side is bounded by $C+\mathbb{E}[|\log\pi_{0,z}|^p]$, which is finite by assumption (2.3).

By assumption (2.2), there exist $a_1,\dots,a_m\in\mathscr{R}$ such that $x_\ell+z+a_1+\cdots+a_{m-\ell}=0$ and $a_{m-\ell+1,m}=z_{1,\ell}$. Then, letting $\eta_0=S_z^+\eta$ and $\eta_{i+1}=S_{a_{i+1}}^+\eta_i$ and defining $(y_i,z_{1,\ell}^i)$ such that $\eta_i=(T_{y_i}\omega,z_{1,\ell}^i)$, $0\leq i\leq m-1$, we have

$$\begin{aligned}
&\sum_{i=0}^{m-1}\mathbb{E}\big[h_{k,\varepsilon}\big(S_{a_{i+1}}^+\eta_i\big)-h_{k,\varepsilon}(\eta_i)|T_{-y_i}\mathfrak{S}_k\big]\\
&\quad=\sum_{i=0}^{m-1}\mathbb{E}\big[h_{k,\varepsilon}\big(T_{y_i+z_1^i}\omega,S_{a_{i+1}}^+z_{1,\ell}^i\big)-h_{k,\varepsilon}\big(T_{y_i}\omega,z_{1,\ell}^i\big)|T_{-y_i}\mathfrak{S}_k\big]\\
&\quad=\sum_{i=0}^{m-1}\mathbb{E}\big[h_{k,\varepsilon}\big(T_{z_1^i}\omega,S_{a_{i+1}}^+z_{1,\ell}^i\big)-h_{k,\varepsilon}\big(\omega,z_{1,\ell}^i\big)|\mathfrak{S}_k\big]\circ T_{y_i}\\
&\quad=\sum_{i=0}^{m-1}\mathbb{E}\big[h_{k,\varepsilon}\big(S_{a_{i+1}}^+\big(\omega,z_{1,\ell}^i\big)\big)-h_{k,\varepsilon}\big(\omega,z_{1,\ell}^i\big)|\mathfrak{S}_k\big]\circ T_{y_i}\\
&\quad\leq Cm-\sum_{i=0}^{m-1}\mathbb{E}\big[\log p^+\big(\big(\omega,z_{1,\ell}^i\big),S_{a_{i+1}}^+\big(\omega,z_{1,\ell}^i\big)\big)|\mathfrak{S}_k\big]\circ T_{y_i}.
\end{aligned}$$

The last inequality is a result of (5.4). Taking conditional expectations given $\mathfrak{S}_{k-1}$ one has

$$\begin{aligned}
-F_{k,\varepsilon}(\eta,z)&=\mathbb{E}\big[h_{k,\varepsilon}(\eta)-h_{k,\varepsilon}\big(S_z^+\eta\big)|\mathfrak{S}_{k-1}\big]\\
&=\sum_{i=0}^{m-1}\mathbb{E}\big[h_{k,\varepsilon}\big(S_{a_{i+1}}^+\eta_i\big)-h_{k,\varepsilon}(\eta_i)|\mathfrak{S}_{k-1}\big]\\
&=\sum_{i=0}^{m-1}\mathbb{E}\big[\mathbb{E}\big[h_{k,\varepsilon}\big(S_{a_{i+1}}^+\eta_i\big)-h_{k,\varepsilon}(\eta_i)|T_{-y_i}\mathfrak{S}_k\big]|\mathfrak{S}_{k-1}\big]\\
&\leq Cm-\sum_{i=0}^{m-1}\mathbb{E}\big[\mathbb{E}\big[\log p^+\big(\big(\omega,z_{1,\ell}^i\big),S_{a_{i+1}}^+\big(\omega,z_{1,\ell}^i\big)\big)|\mathfrak{S}_k\big]\circ T_{y_i}|\mathfrak{S}_{k-1}\big].
\end{aligned}$$

The $L^p(\mathbb{P})$-norm of the right-hand side is bounded by $(C+\mathbb{E}[|\log\pi_{0,z}|^p])m$, which is finite by assumption (2.3).

Consider next a weakly convergent subsequence. We will still denote it by $F_{k,\varepsilon}$. Let F_ε be its limit. Clearly, $F_\varepsilon \in L^p(\mathbb{P})$ and the moment condition (i) in Definition 5.1 is satisfied. Also, since the mean zero property (ii), in Definition 5.1, is satisfied for each $F_{k,\varepsilon}$, it is satisfied for F_ε.

Furthermore, weak convergence in $L^p(\mathbb{P})$ and finiteness of the σ-algebras $\mathfrak{S}_j$ imply that for any fixed j, $\mathbb{E}[F_{k,\varepsilon}|\mathfrak{S}_j]$ converges to $\mathbb{E}[F_\varepsilon|\mathfrak{S}_j]$ for every $z_{1,\ell}\in\mathscr{R}^\ell$ and $\mathbb{P}$-a.e. ω. Since the closed loop property holds for every $F_{k,\varepsilon}$, we have that for any two paths $\{\eta_i\}_{i=0}^n$ and $\{\bar\eta_j\}_{j=0}^m$ as in (iii) of Definition 5.1,

$$\mathbb{E}\Bigg[\sum_{i=0}^{n-1}F_\varepsilon(\eta_i,a_{i+1})|\mathfrak{S}_j\Bigg]=\mathbb{E}\Bigg[\sum_{j=0}^{m-1}F_\varepsilon(\bar\eta_j,\bar a_{j+1})|\mathfrak{S}_j\Bigg].$$

Taking $j\to\infty$ and using the martingale convergence theorem proves the closed loop property holds for F_ε. $\square$

The proof of Theorems 5.3 and 3.1 is thus complete. $\square$

6. Proof of Theorem 2.1

We will now present the proof of the main theorem. Note first that for all $k\ge 0$ and $\mathbb{P}$-a.e. ω,

$$P_0^\omega\big\{S^+(T_{X_k}\omega,Z_{k+1,\infty})=(T_{X_{k+1}}\omega,Z_{k+2,\infty})\big\}=1.$$

Thus, the empirical measure $R_n^{1,\infty}$ comes deterministically close to the set of S^+-invariant measures and every non-S^+-invariant measure has a neighborhood that has zero probability for all large enough n. Since the set of such measures is open and function H in Theorem 2.1 is infinite on it, we need not be concerned with them.

Recall definitions (3.5) of H_ℓ and (2.6) of H. Now, Lemma 3.2 and Theorem 3.1 imply that an almost-sure level 3 large deviation principle holds with rate function $\sup_{\ell\ge1}H_\ell^{**}$. It remains to identify this rate function with the one in the statement of Theorem 2.1. This is shown in the next lemma.

Lemma 6.1. *Assume $\mathbb{P}$ is invariant for the shifts $\{T_z\}$ and satisfies assumptions (3.1) and (3.4). If $\mu\in\mathcal{M}_1(\boldsymbol{\Omega}_+)$ is S^+-invariant, then*

$$\sup_{\ell\ge1}H_\ell(\mu_{|\boldsymbol{\Omega}_\ell})=H(\mu). \tag{6.1}$$

In particular, H is convex. If, furthermore, the compactness assumption (2.1) holds then

$$\sup_{\ell\ge1}H_\ell^{**}(\mu_{|\boldsymbol{\Omega}_\ell})=H^{**}(\mu). \tag{6.2}$$

Proof. Let us start with the first identity. Assume $\mu_0\ll\mathbb{P}$ since otherwise the equality holds trivially. Let $\mu_-^{(\ell)}$ be the law of $(\omega,Z_{1-\ell,0})$ under μ_-. Then, the S^+-invariance of μ implies that

$$\begin{aligned}H_\infty(\mu)&\stackrel{\text{def}}{=}\sup_{\ell\ge1}H_\ell(\mu_{|\boldsymbol{\Omega}_\ell})\\&=\sup_{\ell\ge1}\inf\big\{H\big(\mu_-^{(\ell)}\times q\big|\mu_-^{(\ell)}\times p^-\big)\colon q\in\mathcal{Q}(\boldsymbol{\Omega}_\ell)\text{ and }\mu_-^{(\ell)}q=\mu_-^{(\ell)}\big\}.\end{aligned}$$

Recall the universal kernel $\bar q$ that corresponds to all S^+-invariant measures $\mu\in\mathcal{M}_1(\boldsymbol{\Omega}_+)$. Let

$$\bar q^{(\ell)}\big(\zeta,S_z^-\zeta\big)=E^{\mu_-}\big[\bar q\big(\eta,S_z^-\eta\big)|\eta_{1-\ell,0}=\zeta\big].$$

Then, $\mu_-^{(\ell)}$ is $\bar q^{(\ell)}$-invariant. Moreover, $\mu_-^{(\ell)}\times\bar q^{(\ell)}$ is the restriction of $\mu_-\times\bar q$ to $\boldsymbol{\Omega}_\ell^2$. Thus, $H(\mu_-^{(\ell)}\times\bar q^{(\ell)}|\mu_-^{(\ell)}\times p^-)\le H(\mu_-\times\bar q|\mu_-\times p^-)$. This shows that $H_\infty(\mu)\le H(\mu)$.

The other direction is trivial if $H_\infty(\mu)=\infty$. On the other hand, if $H_\infty(\mu)=h<\infty$, then there exists a sequence $q_-^{(\ell)}\in\mathcal{Q}(\boldsymbol{\Omega}_\ell)$ such that $\mu_-^{(\ell)}$ is $q_-^{(\ell)}$-invariant, and

$$H\bigl(\mu_-^{(\ell)}\times q_-^{(\ell)}\big|\mu_-^{(\ell)}\times p^-\bigr)\le h+\ell^{-1}.$$

This implies that, for $\mu_-^{(\ell)}$-a.e. $\eta\in\boldsymbol{\Omega}_\ell$, $q_-^{(\ell)}(\eta,\{S_z^-\eta\colon z\in\mathscr{R}\})=1$.

For $\ell\ge\ell'$, measures $\mu_-^{(\ell)}\times q_-^{(\ell)}$ have marginals $\mu_-^{(\ell')}$. Thus, for ℓ' fixed, measures $\mu_-^{(\ell)}\times q_-^{(\ell)}$ restricted to $\boldsymbol{\Omega}_{\ell'}^2$ are tight. We can use the diagonal trick to extract one sequence that converges weakly on all spaces $\boldsymbol{\Omega}_{\ell'}^2$ simultaneously. By Kolmogorov's extension theorem one can find a limit point $Q\in\mathcal{M}_1(\boldsymbol{\Omega}_-^2)$. The marginals of Q are equal to μ_- and hence the conditional distribution of the second coordinate ζ under Q, given the first coordinate η, defines a kernel $q_-(\eta,\mathrm{d}\zeta)$ that leaves μ_- invariant. The following entropy argument shows that q_- is still supported on S_z^--shifts; that is

$$Q\biggl\{(\eta,\zeta)\in\boldsymbol{\Omega}_-\colon\ \zeta\in\bigcup_z\{S_z^-\eta\}\biggr\}=1. \tag{6.3}$$

For any $\varepsilon>0$, there exists a compact subset $K_\varepsilon\subset\Omega$ such that $\mu_-(K_\varepsilon\times\mathscr{R}^{\mathbb{Z}_-})\ge1-\varepsilon$. On the other hand, for any finite $\mathcal{A}\subset\mathscr{R}$ the function $\omega\mapsto F(\omega,\mathcal{A})=\sum_{z\in\mathcal{A}}\pi_{0,z}(\omega)$ is continuous. Furthermore, this function increases up to 1, for all ω, as $\mathcal{A}$ increases to $\mathscr{R}$. Thus, for each ω in K_ε choose a set $\mathcal{A}$ so that $F(\omega,\mathcal{A})\ge1-\varepsilon/2$ and pick an open neighborhood G of ω so that for $\omega'\in G$, $F(\omega',\mathcal{A})\ge1-\varepsilon$. Since K_ε is compact, it can be covered with finitely many such neighborhoods. Let $\mathcal{A}_\varepsilon$ be the union of the corresponding sets $\mathcal{A}$. Then, $\mathcal{A}_\varepsilon$ is finite and $F(\omega,\mathcal{A}_\varepsilon)\ge1-\varepsilon$ for all $\omega\in K_\varepsilon$. In fact, we can and will choose $\mathcal{A}_\varepsilon$ to increase to $\mathscr{R}$ as ε decreases to 0.

Now recall the variational characterization of relative entropy (see Lemma 10.1 in [21] or Theorem 6.7 in [15]) and write

$$\begin{aligned}
h+1&\ge H\bigl(\mu_-^{(\ell)}\times q_-^{(\ell)}\big|\mu_-^{(\ell)}\times p^-\bigr)\\
&=E^{\mu_-^{(\ell)}}\biggl[\sup_f\biggl\{\sum_z q_-^{(\ell)}\bigl(\eta,S_z^-\eta\bigr)f\bigl(S_z^-\eta\bigr)-\log\sum_z\pi_{0,z}(\omega)\mathrm{e}^{f(S_z^-\eta)}\biggr\}\biggr]\\
&\ge E^{\mu_-^{(\ell)}}\biggl[\sup_f\biggl\{\sum_z q_-^{(\ell)}\bigl(\eta,S_z^-\eta\bigr)f\bigl(S_z^-\eta\bigr)-\log\sum_z\pi_{0,z}(\omega)\mathrm{e}^{f(S_z^-\eta)}\biggr\}\mathbb{1}_{K_\varepsilon\times\mathscr{R}^\ell}(\eta)\biggr]\\
&\ge E^{\mu_-^{(\ell)}}\biggl[\biggl\{C\sum_{z\notin\mathcal{A}_\varepsilon}q_-^{(\ell)}\bigl(\eta,S_z^-\eta\bigr)-\log\sum_z\pi_{0,z}(\omega)\mathrm{e}^{C\mathbb{1}\{z\notin\mathcal{A}_\varepsilon\}}\biggr\}\mathbb{1}_{K_\varepsilon\times\mathscr{R}^\ell}(\eta)\biggr]\\
&=E^{\mu_-^{(\ell)}}\biggl[\biggl\{C\sum_{z\notin\mathcal{A}_\varepsilon}q_-^{(\ell)}\bigl(\eta,S_z^-\eta\bigr)-\log\biggl(1+\bigl(\mathrm{e}^C-1\bigr)\sum_{z\notin\mathcal{A}_\varepsilon}\pi_{0,z}(\omega)\biggr)\biggr\}\mathbb{1}_{K_\varepsilon\times\mathscr{R}^\ell}(\eta)\biggr]\\
&\ge E^{\mu_-^{(\ell)}}\biggl[\biggl\{C\sum_{z\notin\mathcal{A}_\varepsilon}q_-^{(\ell)}\bigl(\eta,S_z^-\eta\bigr)-\log\bigl(1+\bigl(\mathrm{e}^C-1\bigr)\varepsilon\bigr)\biggr\}\mathbb{1}_{K_\varepsilon\times\mathscr{R}^\ell}(\eta)\biggr].
\end{aligned}$$

In the third inequality we used $f(\eta)=C\mathbb{1}\{z_0\notin\mathcal{A}_\varepsilon\}$. Now, fix a $\delta>0$ and choose C large such that $(1+h)/C<\delta/2$. Then choose $\varepsilon>0$ small such that $C^{-1}\log(1+(\mathrm{e}^C-1)\varepsilon)+\varepsilon<\delta/2$. The above inequalities then become

$$\begin{aligned}
&E^{\mu_-^{(\ell)}}\biggl[\sum_{z\notin\mathcal{A}_\varepsilon}q_-^{(\ell)}\bigl(\eta_{1-\ell,0},S_z^-\eta_{1-\ell,0}\bigr)\biggr]\\
&\quad\le C^{-1}(1+h)+C^{-1}\log\bigl(1+\bigl(\mathrm{e}^C-1\bigr)\varepsilon\bigr)+1-\mu_-^{(\ell)}\bigl(K_\varepsilon\times\mathscr{R}^\ell\bigr)<\delta.
\end{aligned}$$

Since $\{(\eta,\zeta)\in\boldsymbol{\Omega}^2_-:\ \zeta_{1-\ell',0}=S_z^-\eta_{1-\ell',0}\text{ and }z\in\mathcal{A}_\varepsilon\}$ is closed it follows that

$$
\begin{aligned}
&Q\bigl\{(\eta,\zeta)\in\boldsymbol{\Omega}^2_-:\ \zeta_{1-\ell',0}=S_z^-\eta_{1-\ell',0}\text{ and }z\in\mathcal{A}_\varepsilon\bigr\}\\
&\quad\ge\varlimsup_{\ell\to\infty}\mu_-^{(\ell)}\times q_-^{(\ell)}\bigl\{(\eta_{1-\ell,0},\zeta_{1-\ell,0})\in\boldsymbol{\Omega}^2_\ell:\ \zeta_{1-\ell',0}=S_z^-\eta_{1-\ell',0}\text{ and }z\in\mathcal{A}_\varepsilon\bigr\}\\
&\quad\ge\varlimsup_{\ell\to\infty}\mu_-^{(\ell)}\times q_-^{(\ell)}\bigl\{\bigl(\eta_{1-\ell,0},S_z^-\eta_{1-\ell,0}\bigr):\ \eta_{1-\ell,0}\in\boldsymbol{\Omega}_\ell,z\in\mathcal{A}_\varepsilon\bigr\}\ge 1-\delta.
\end{aligned}
$$

But $\{(\eta,\zeta):\ \eta\in\boldsymbol{\Omega}_-,\ \zeta=S_z^-\eta\text{ and }z\in\mathcal{A}_\varepsilon\}$ is equal to the decreasing limit

$$
\bigcap_{\ell'\ge1}\bigl\{(\eta,\zeta)\in\boldsymbol{\Omega}^2_-:\ \zeta_{1-\ell',0}=S_z^-\eta_{1-\ell',0}\text{ and }z\in\mathcal{A}_\varepsilon\bigr\}.
$$

Now, taking $\delta\to0$ then $\varepsilon\to0$ proves (6.3). Since there is a unique kernel that leaves μ_- invariant and is supported on S_z^--shifts, $Q=\mu_-\times\bar q$ is the only possible limit point. Lower semicontinuity of the entropy implies that

$$
\begin{aligned}
H\bigl((\mu_-\times\bar q)_{|\boldsymbol{\Omega}^2_{\ell'}}\big|(\mu_-\times p^-)_{|\boldsymbol{\Omega}^2_{\ell'}}\bigr)&\le\varliminf_{\ell\to\infty}H\bigl((\mu_-^{(\ell)}\times q_-^{(\ell)})_{|\boldsymbol{\Omega}^2_{\ell'}}\big|(\mu_-^{(\ell)}\times p^-)_{|\boldsymbol{\Omega}^2_{\ell'}}\bigr)\\
&\le\varliminf_{\ell\to\infty}H\bigl(\mu_-^{(\ell)}\times q_-^{(\ell)}\big|\mu_-^{(\ell)}\times p^-\bigr)\le h.
\end{aligned}
$$

Taking $\ell'\to\infty$ proves that $H(\mu)\le H_\infty(\mu)$ and (6.1) holds.

Next, we prove (6.2). First, we show that for $\mu\in\mathcal{M}_1(\boldsymbol{\Omega}_\ell)$,

$$
H_\ell(\mu)=\inf\bigl\{H(\nu):\ \nu\text{ is }S^+\text{-invariant and }\nu_{|\boldsymbol{\Omega}_\ell}=\mu\bigr\}. \tag{6.4}
$$

If $\mu_0\not\ll\mathbb{P}$ then both sides are infinite. Suppose $\mu_0\ll\mathbb{P}$. Write temporarily $I(\nu)=\sup_\ell H_\ell^{**}(\nu_{|\boldsymbol{\Omega}_\ell})$ for the level 3 rate function. If ν is S^+-invariant and $\nu_0\ll\mathbb{P}$ then by (6.1)

$$
I(\nu)=\sup_\ell H_\ell^{**}(\nu_{|\boldsymbol{\Omega}_\ell})=\sup_\ell H_\ell(\nu_{|\boldsymbol{\Omega}_\ell})=H(\nu). \tag{6.5}
$$

By the level 3 to level 2 contraction,

$$
H_\ell(\mu)=H_\ell^{**}(\mu)=\inf\bigl\{I(\nu):\ \nu\text{ is }S^+\text{-invariant and }\nu_{|\boldsymbol{\Omega}_\ell}=\mu\bigr\}.
$$

Since $I\le H$, to prove (6.4) it suffices to consider the case $H_\ell(\mu)<\infty$. Only S^+-invariant measures have finite level 3 rate, hence there exists at least one S^+-invariant ν such that $\nu_{|\boldsymbol{\Omega}_\ell}=\mu$. Furthermore, the measures ν that appear in the contraction satisfy $\nu_0=\mu_0\ll\mathbb{P}$, and so by (6.5) equation (6.4) follows.

Now, consider S^+-invariant measures ν. By (6.1) $I\le H$, and since I is a l.s.c. convex function, also $I\le H^{**}$. By (6.4) and the basic Lemma A.2,

$$
H_\ell^{**}(\mu)=\inf\bigl\{H^{**}(\nu):\ \nu\text{ is }S^+\text{-invariant and }\nu_{|\boldsymbol{\Omega}_\ell}=\mu\bigr\}.
$$

Outside S^+-invariant measures $H^{**}\equiv\infty$ so whether or not the invariance condition is included in the infimum is immaterial.

Let $c>I(\nu)$. For each ℓ use above to find $\mu^{(\ell)}$ such that $\mu^{(\ell)}_{|\boldsymbol{\Omega}_\ell}=\nu_{|\boldsymbol{\Omega}_\ell}$ and $H^{**}(\mu^{(\ell)})<c$. $\mu^{(\ell)}\to\nu$ and so by lower semicontinuity $H^{**}(\nu)\le\varliminf H^{**}(\mu^{(\ell)})\le c$. This shows $H^{**}\le I$. $\square$

Appendix A: Technical lemmas

Lemma A.1. *Let g be a bounded measurable function on a product space $\mathcal{X}\times\mathcal{Y}$, μ a probability measure on $\mathcal{X}$ and ρ a probability measure on $\mathcal{Y}$. Then*

$$
\log\int_{\mathcal{X}}e^{\int_{\mathcal{Y}}g(x,y)\rho(dy)}\mu(dx)\le\int_{\mathcal{Y}}\biggl[\log\int_{\mathcal{X}}e^{g(x,y)}\mu(dx)\biggr]\rho(dy).
$$

Proof. The inequality can be thought of as an infinite-dimensional Jensen's inequality, applied to the convex functional $\Psi(f)=\log\int_{\mathcal{X}}\mathrm{e}^{f(x)}\mu(\mathrm{d}x)$. Proof is immediate from the variational characterization of relative entropy; see Lemma 10.1 in [21] or Theorem 6.7 in [15]. First for an arbitrary probability measure γ on $\mathcal{X}$,

$$\begin{aligned}\int_{\mathcal{Y}}\biggl[\log\int_{\mathcal{X}}\mathrm{e}^{g(x,y)}\mu(\mathrm{d}x)\biggr]\rho(\mathrm{d}y) &\geq \int_{\mathcal{Y}}\biggl[\int_{\mathcal{X}}g(x,y)\gamma(\mathrm{d}x)-H(\gamma|\mu)\biggr]\rho(\mathrm{d}y)\\ &= \int_{\mathcal{X}}\biggl[\int_{\mathcal{Y}}g(x,y)\rho(\mathrm{d}y)\biggr]\gamma(\mathrm{d}x)-H(\gamma|\mu)=\log\int_{\mathcal{X}}\mathrm{e}^{\int_{\mathcal{Y}}g(x,y)\rho(\mathrm{d}y)}\mu(\mathrm{d}x)\end{aligned}$$

where the last equality comes from taking

$$\gamma(\mathrm{d}x)=\biggl(\int_{\mathcal{X}}\mathrm{e}^{\int_{\mathcal{Y}}g(z,y)\rho(\mathrm{d}y)}\mu(\mathrm{d}z)\biggr)^{-1}\mathrm{e}^{\int_{\mathcal{Y}}g(x,y)\rho(\mathrm{d}y)}\mu(\mathrm{d}x).\qquad\square$$

Lemma A.2. *Let $\mathbb{S}$ and $\mathbb{T}$ be compact metric spaces and $\pi:\mathbb{S}\to\mathbb{T}$ continuous. Let $f:\mathbb{S}\to[0,\infty]$ be an arbitrary function and $f_{\mathrm{lsc}}(s)=\lim_{r\searrow 0}\inf_{x\in B(s,r)}f(x)$ its lower semicontinuous regularization. Let $g(t)=\inf_{\pi(s)=t}f(s)$. Then $g_{\mathrm{lsc}}(t)=\inf_{\pi(s)=t}f_{\mathrm{lsc}}(s)$.*

Proof. Immediately $g_{\mathrm{lsc}}(t)\geq\inf_{\pi(s)=t}f_{\mathrm{lsc}}(s)$ because the function on the right is at or below $g(t)$ and on a compact metric space it is l.s.c.

Let $c>\inf_{\pi(s)=t}f_{\mathrm{lsc}}(s)$. Fix s so that $\pi(s)=t$ and $f_{\mathrm{lsc}}(s)<c$. Find $s_j\to s$ so that $f(s_j)<c$ (constant sequence $s_j=s$ is a legitimate choice). Then $\pi(s_j)\to t$, and consequently

$$g_{\mathrm{lsc}}(t)\leq\underline{\lim}\,g\bigl(\pi(s_j)\bigr)\leq\underline{\lim}\,f(s_j)\leq c.\qquad\square$$

Appendix B: Proof of Lemma 5.1

In what follows, C denotes a chameleon constant which can change values from line to line. The only values it depends upon are $|\mathscr{R}|$, ℓ and d. C_r is again a chameleon constant but its value also depends on r. Finally, $C_r(\omega)$ also depends on ω. Note that $\ell\geq 1$ is a fixed integer, throughout this section.

Recall that $\tilde{x}_\ell=\tilde{z}_1+\cdots+\tilde{z}_\ell$. Similarly, $\bar{x}_\ell=\bar{z}_1+\cdots+\bar{z}_\ell$. Under (3.4) there always exists a path from $(y,\tilde{z}_{1,\ell})$ to $(x,z_{1,\ell})$ in the sense that there exist $m\geq\ell$ and $a_1,\dots,a_{m-\ell}\in\mathscr{R}$ such that

$$y+\tilde{x}_\ell+a_1+\cdots+a_{m-\ell}=x.$$

The definition is independent of $z_{1,\ell}$ but for symmetry of language it seems sensible to keep it in the statement. The case $m=\ell$ is admissible also and then $y+\tilde{x}_\ell=x$. Then if we set $a_{m-\ell+1,m}=z_{1,\ell}$, the composition $S^+_{a_m}\circ\cdots\circ S^+_{a_1}$ takes $(T_y\omega,\tilde{z}_{1,\ell})$ to $(T_x\omega,z_{1,\ell})$ for all $\omega\in\Omega$.

Paths can be concatenated. If there is a path from $(y,\tilde{z}_{1,\ell})$ to $(x,z_{1,\ell})$ and from $(u,\bar{z}_{1,\ell})$ to $(y,\tilde{z}_{1,\ell})$, then we have

$$y+\tilde{x}_\ell+a_1+\cdots+a_{m-\ell}=x\quad\text{and}\quad u+\bar{x}_\ell+b_1+\cdots+b_{n-\ell}=y.$$

Taking $b_{n-\ell+1,n}=\tilde{z}_{1,\ell}$ we then have

$$u+\bar{x}_\ell+b_1+\cdots+b_n+a_1+\cdots+a_{m-\ell}=x$$

and there is a path from $(u,\bar{z}_{1,\ell})$ to $(x,z_{1,\ell})$.

For any two points $(x,z_{1,\ell})$ and $(\bar{x},\bar{z}_{1,\ell})$ and any $\tilde{z}_{1,\ell}$ there exists a point $y\in\mathbb{Z}^d$ such that from $(y,\tilde{z}_{1,\ell})$ there is a path to both $(x,z_{1,\ell})$ and $(\bar{x},\bar{z}_{1,\ell})$. For this, find first $\bar{a}_1,\dots,\bar{a}_{m-\ell}$ and $a_1,\dots,a_{n-\ell}\in\mathscr{R}$ such that

$$\bar{x}-x=(\bar{a}_1+\cdots+\bar{a}_{m-\ell})-(a_1+\cdots+a_{n-\ell})$$

so that

$$y'=\bar{x}-(\bar{a}_1+\cdots+\bar{a}_{m-\ell})=x-(a_1+\cdots+a_{n-\ell})$$

and then take $y = y' - \tilde{x}_\ell$. By induction, there is a common starting point for paths to any finite number of points.

Now fix $F \in \mathcal{K}^p(\boldsymbol{\Omega}_\ell \times \mathscr{R})$. If there is a path from $(y, \tilde{z}_{1,\ell})$ to $(x, z_{1,\ell})$, set $\eta_0 = (T_y\omega, \tilde{z}_{1,\ell})$, $\eta_i = S^+_{a_i}\eta_{i-1}$ for $i = 1, \dots, m$ so that $\eta_m = (T_x\omega, z_{1,\ell})$, and then

$$L\big(\omega, (y, \tilde{z}_{1,\ell}), (x, z_{1,\ell})\big) = \sum_{i=0}^{m-1} F(\eta_i, a_{i+1}). \tag{B.1}$$

By the closed loop property $L(\omega, (y, \tilde{z}_{1,\ell}), (x, z_{1,\ell}))$ is independent of the path chosen. If $a_1, \dots, a_{m-\ell}$ work for $(y, \tilde{z}_{1,\ell})$ and $(x, z_{1,\ell})$, then these steps work also for $(y + u, \tilde{z}_{1,\ell})$ and $(x + u, z_{1,\ell})$. The effect on the right-hand side of (B.1) is simply to shift ω by u, and consequently

$$L\big(T_u\omega, (y, \tilde{z}_{1,\ell}), (x, z_{1,\ell})\big) = L\big(\omega, (y + u, \tilde{z}_{1,\ell}), (x + u, z_{1,\ell})\big). \tag{B.2}$$

Next define $f : \Omega \times \mathscr{R}^{2\ell} \times \mathbb{Z}^d \to \mathbb{R}$ by

$$f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, x) = L\big(\omega, (y, \tilde{z}_{1,\ell}), (x, \bar{z}_{1,\ell})\big) - L\big(\omega, (y, \tilde{z}_{1,\ell}), (0, z_{1,\ell})\big) \tag{B.3}$$

for any $(y, \tilde{z}_{1,\ell})$ with a path to both $(0, z_{1,\ell})$ and $(x, \bar{z}_{1,\ell})$. This definition is independent of the choice of $(y, \tilde{z}_{1,\ell})$, again by the closed loop property.

Here are some basic properties of f. We postpone the proof of this lemma to the end of this section.

Lemma B.1. *Same setting as Lemma 5.1.*

(a) *There exists a constant C depending only on d, ℓ and $R = \max\{|z|: z \in \mathscr{R}\}$, such that we have for all $z_{1,\ell}, \bar{z}_{1,\ell} \in \mathscr{R}^\ell$, $x \in \mathbb{Z}^d$ and $\mathbb{P}$-a.e. ω,*

$$\big|f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, x)\big| \le \max_{\substack{\tilde{z}_{1,\ell} \in \mathscr{R}^\ell \\ z \in \mathscr{R}}} \sum_{b:\ |b| \le C|x|} \big|F(T_b\omega, \tilde{z}_{1,\ell}, z)\big|.$$

(b) *The closed loop property of F implies that for any $z_{1,\ell}, \bar{z}_{1,\ell} \in \mathscr{R}^\ell$, $x, \bar{x} \in \mathbb{Z}^d$, and $\mathbb{P}$-a.e. ω,*

$$\begin{aligned} f(T_x\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, \bar{x} - x) &= f(\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, \bar{x}) - f(\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, x) \\ &= f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, \bar{x}) - f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, x). \end{aligned}$$

(c) *The mean zero property of F implies that for any $\bar{z}_{1,\ell} \in \mathscr{R}^\ell$ and $x \in \mathbb{Z}^d$, $\mathbb{E}[f(\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, x)] = 0$.*

Next, extend f to a continuous function of $\xi \in \mathbb{R}^d$ by linear interpolation. Here is one way to do that. Recall that $\{e_1, \dots, e_d\}$ is the canonical basis of $\mathbb{R}^d$. Introduce the following notation: for $p \in [0, 1]$ and $i \in \{1, \dots, d\}$, let $B_i(p)$ be a Bernoulli random variable with parameter p. For a vector $p = (p_1, \dots, p_d) \in [0, 1]^d$, let $B(p) = \sum_{i=1}^d B_i(p_i)e_i$ with $(B_i(p_i))$ independent.

Now, for given η, $\bar{z}_{1,\ell}$ and $\xi = \sum_{i=1}^d \xi_i e_i$, let $[\xi] = \sum_{i=1}^d [\xi_i]e_i$, where $[\xi_i]$ is the largest integer smaller than or equal to ξ_i, and define

$$f(\eta, \bar{z}_{1,\ell}, \xi) = E\big[f\big(\eta, \bar{z}_{1,\ell}, [\xi] + B\big(\xi - [\xi]\big)\big)\big].$$

Think of f as a collection of functions of (ω, ξ). The idea is to homogenize these functions by showing that, for fixed $z_{1,\ell}$ and $\bar{z}_{1,\ell}$ and for $\mathbb{P}$-a.e. ω, $g_n(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, \xi) = n^{-1} f((\omega, z_{1,\ell}), \bar{z}_{1,\ell}, n\xi)$ is equicontinuous and hence converges, uniformly on compacts and along a subsequence, to a function $g(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, \xi)$. Next, one shows that g has to be constant and since $g(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, 0) = 0$ we conclude that g_n converges uniformly on compacts to 0. Observe now that if $\eta_0 = (\omega, z_{1,\ell})$ and $\eta_{k+1} = S^+_{a_{k+1}}\eta_k$ for $0 \le k \le n - 1$ and $a_k \in \mathscr{R}$, then

$$n^{-1} \sum_{k=0}^{n-1} F(\eta_k, a_{k+1}) = g_n(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, \xi),$$

where $\xi = (x_\ell + a_1 + \cdots + a_{n-\ell})/n$ and $\bar{z}_{1,\ell} = (a_{n-\ell+1}, \ldots, a_n)$. Thus,

$$\max_{(a_1,\ldots,a_n)\in\mathscr{R}^n} \left| n^{-1} \sum_{k=0}^{n-1} F(\eta_k, a_{k+1}) \right| \le \max_{\bar{z}_{1,\ell}\in\mathscr{R}^\ell} \sup_{\xi:\ |\xi|\le R} \bigl| g_n(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, \xi) \bigr|,$$

where $R = \max\{|z|:\ z \in \mathscr{R}\}$. This completes the proof of the lemma.

The above strategy was introduced by Kosygina, Rezakhanlou and Varadhan [12] in the context of diffusions with random drift, then carried out by Rosenbluth [17] for random walk in random environment in the case $\ell = 0$. Equicontinuity follows from an application of the Garsia–Rodemich–Rumsey theorem (see [20]) which requires the moment assumption on F. The fact that g is constant follows from an application of the ergodic theorem along with the mean 0 property of F. We present the proof, adapted to our setting, for the sake of completeness.

Let us start with equicontinuity. This will be shown by breaking the space into two parts. Each of the following two lemmas covers one part. Let us denote $B_r(\xi) = \{\zeta \in \mathbb{R}^d:\ |\zeta - \xi| \le r\}$.

Lemma B.2. *Same assumptions on $\mathbb{P}$ and F as in Lemma 5.1. Then, for any $r \ge 1$ and any $\gamma \in (0, d+p)$, one has that for $\mathbb{P}$-a.e. ω and all $z_{1,\ell}, \bar{z}_{1,\ell} \in \mathscr{R}^\ell$*

$$\varlimsup_{n\to\infty} \int_{B_r(0)} \int_{B_{2d/n}(\xi)\cap B_r(0)} \frac{|g_n(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, \xi) - g_n(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, \zeta)|^p}{|\xi - \zeta|^\gamma} \,\mathrm{d}\zeta\,\mathrm{d}\xi = 0.$$

Proof. Changing variables, the above integral can be rewritten as

$$\begin{aligned} &\frac{1}{n^p} \int_{B_r(0)} \int_{B_{2d/n}(\xi)\cap B_r(0)} \frac{|f(n\xi) - f(n\zeta)|^p}{|\xi - \zeta|^\gamma} \,\mathrm{d}\zeta\,\mathrm{d}\xi \\ &\quad = \frac{1}{n^{2d+p-\gamma}} \int_{B_{rn}(0)} \int_{B_{2d}(\xi)\cap B_{rn}(0)} \frac{|f(\xi) - f(\zeta)|^p}{|\xi - \zeta|^\gamma} \,\mathrm{d}\zeta\,\mathrm{d}\xi, \end{aligned} \tag{B.4}$$

where we dropped ω, $z_{1,\ell}$, and $\bar{z}_{1,\ell}$ from the arguments of f for the moment.

Observe next that if ζ is on the boundary of a $\mathbb{Z}^d$-cell, i.e. $\zeta_i \in \mathbb{Z}$ for some $i \in \{1, \ldots, d\}$, then the fact that $\zeta_i + B_i(0)$ has the same distribution as $\zeta_i - 1 + B_i(1)$ shows that one can set $[\zeta_i]$ to be either ζ_i or $\zeta_i - 1$ and the value of f at ζ would not be affected.

Therefore, if ξ and ζ belong to the same $\mathbb{Z}^d$-cell, we can assume that $[\xi_i] = [\zeta_i] = x$, the lower left corner of the cell. Abbreviate $p_i = \zeta_i - x_i$ and $q_i = \xi_i - x_i$. Then

$$\begin{aligned} \bigl|f(\zeta) - f(\xi)\bigr| &= \bigl|E\bigl[f\bigl(x + B(\zeta - x)\bigr) - f\bigl(x + B(\xi - x)\bigr)\bigr]\bigr| \\ &= \left| \sum_{(b_i)\in\{0,1\}^d} \left[\prod_i p_i^{b_i}(1-p_i)^{1-b_i} - \prod_i q_i^{b_i}(1-q_i)^{1-b_i} \right] f\left(x + \sum_i b_i e_i\right) \right| \\ &= \left| \sum_{(b_i)\in\{0,1\}^d} \left[\prod_i p_i^{b_i}(1-p_i)^{1-b_i} - \prod_i q_i^{b_i}(1-q_i)^{1-b_i} \right] \left[f\left(x + \sum_i b_i e_i\right) - f(x) \right] \right| \\ &\le \sum_{(b_i)\in\{0,1\}^d} \left| \prod_i p_i^{b_i}(1-p_i)^{1-b_i} - \prod_i q_i^{b_i}(1-q_i)^{1-b_i} \right| \cdot \left| f\left(x + \sum_i b_i e_i\right) - f(x) \right| \\ &\le C|\zeta - \xi| \sum_{(b_i)\in\{0,1\}^d} \left| f\left(x + \sum_i b_i e_i\right) - f(x) \right|, \end{aligned}$$

where we have used the fact that for $a, b, c, d \in [0, 1]$,

$$|ab - cd| \le \bigl|(a-c)b\bigr| + \bigl|(b-d)c\bigr| \le |a-c| + |b-d|.$$

If, on the other hand, ξ and ζ are in two different $\mathbb{Z}^d$-cells then, since $|\xi-\zeta|\le 2d$, there exist points $\zeta_0,\dots,\zeta_m$, with $m\le C$, such that $\zeta_0=\xi$, $\zeta_m=\zeta$, each two consecutive ones belong to the same $\mathbb{Z}^d$-cell, and $|\zeta_{k+1}-\zeta_k|\le C|\zeta-\xi|$. One can then write

$$\begin{aligned}\bigl|f(\zeta)-f(\xi)\bigr| &\le C\sum_{k=0}^{m-1}|\zeta_{k+1}-\zeta_k|\sum_{(b_i)\in\{0,1\}^d}\left|f\left([\zeta_k]+\sum_i b_ie_i\right)-f\bigl([\zeta_k]\bigr)\right|\\ &\le C|\zeta-\xi|\sum_{k=0}^{m-1}\sum_{(b_i)\in\{0,1\}^d}\left|f\left([\zeta_k]+\sum_i b_ie_i\right)-f\bigl([\zeta_k]\bigr)\right|\\ &= C|\zeta-\xi|\sum_{k=0}^{m-1}\sum_{(b_i)\in\{0,1\}^d}\left|f\left(T_{[\zeta_k]}\omega,\bar z_{1,\ell},\bar z_{1,\ell},\sum_i b_ie_i\right)\right|,\end{aligned}$$

where we have used part (b) of Lemma B.1. Furthermore, using part (a) of the same lemma, and that $|\zeta_k-[\xi]|\le C|\xi-\zeta|\le C$, we have

$$\bigl|f(\zeta)-f(\xi)\bigr|\le C|\zeta-\xi|\max_{\tilde z_{1,\ell}\in\mathscr{R}^\ell}\max_{z\in\mathscr{R}}\max_{x:\,|x-[\xi]|\le C}\bigl|F(T_x\omega,\tilde z_{1,\ell},z)\bigr|.$$

Since $d+p>\gamma$, one has that $\int_{B_{2d}(\xi)}|\zeta-\xi|^{p-\gamma}\,\mathrm{d}\zeta<\infty$. Setting,

$$G(\omega)=\max_{\tilde z_{1,\ell}\in\mathscr{R}^\ell}\max_{z\in\mathscr{R}}\max_{x:\,|x|\le C}\bigl|F(T_x\omega,\tilde z_{1,\ell},z)\bigr|^p\in L^1(\mathbb{P}),\tag{B.5}$$

integral (B.4) is then bounded by

$$Cn^{\gamma-d-p}\left(n^{-d}\sum_{y:\,|y|\le rn}G(T_y\omega)\right).$$

The lemma follows since $d+p>\gamma$ and, by the ergodic theorem (see, for example, Theorem 14.A8 in [9]), the quantity in parentheses converges to a finite constant. □

Lemma B.3. *Same assumptions on $\mathbb{P}$ and F as in Lemma 5.1. Then, for any $r\ge 1$ and any $\gamma\in(d+p-1,d+p)$, there exists a constant C_r such that for $\mathbb{P}$-a.e. ω and all $z_{1,\ell},\bar z_{1,\ell}\in\mathscr{R}^\ell$*

$$\varlimsup_{n\to\infty}\int_{B_r(0)}\int_{B_r(0)\setminus B_{2d/n}(\xi)}\frac{|g_n(\omega,z_{1,\ell},\bar z_{1,\ell},\xi)-g_n(\omega,z_{1,\ell},\bar z_{1,\ell},\zeta)|^p}{|\xi-\zeta|^\gamma}\,\mathrm{d}\zeta\,\mathrm{d}\xi\le C_r.$$

Proof. Once again, changing variables the above integral becomes

$$\frac{1}{n^{2d+p-\gamma}}\int_{B_{rn}(0)}\int_{B_{rn}(0)\setminus B_{2d}(\xi)}\frac{|f(\xi)-f(\zeta)|^p}{|\xi-\zeta|^\gamma}\,\mathrm{d}\zeta\,\mathrm{d}\xi.\tag{B.6}$$

Write

$$\begin{aligned}\bigl|f(\xi)-f(\zeta)\bigr|\le{}&\bigl|f\bigl([\xi]\bigr)-f\bigl([\zeta]\bigr)\bigr|+\sum_{(b_i)\in\{0,1\}^d}\left|f\left([\xi]+\sum_i b_ie_i\right)-f\bigl([\xi]\bigr)\right|\\ &+\sum_{(b_i)\in\{0,1\}^d}\left|f\left([\zeta]+\sum_i b_ie_i\right)-f\bigl([\zeta]\bigr)\right|.\end{aligned}$$

Observing that $\gamma > 1$, $\gamma < d + p$, $|\xi - \zeta| \ge 2d$ and $|[\zeta]| \le C|\zeta| \le C_r n$, the second and third terms above are dealt with exactly as in the previous lemma (using the ergodic theorem). For example,

$$\begin{aligned}
&\frac{1}{n^{2d+p-\gamma}} \int_{B_{rn}(0)} \int_{B_{rn}(0)\setminus B_{2d}(\xi)} \frac{|f([\xi]+\sum_i b_i e_i) - f([\xi])|^p}{|\xi-\zeta|^\gamma} \, d\zeta \, d\xi \\
&\quad \le \frac{C}{n^{2d+p-\gamma}} \int_{B_{rn}(0)} \left| f\Bigl([\xi] + \sum_i b_i e_i\Bigr) - f\bigl([\xi]\bigr) \right|^p d\xi \\
&\quad \le C n^{\gamma-d-p} \Biggl(n^{-d} \sum_{y:\ |y|\le rn} G(T_y\omega) \Biggr).
\end{aligned}$$

Observe next that since $|\xi - \zeta| \ge 2d$, $1 \le |[\xi] - [\zeta]| \le C|\xi - \zeta|$ and we are reduced to bounding the sum

$$\frac{1}{n^{2d+p-\gamma}} \sum_{\substack{x,y:\ x\ne y \\ |x|,|y|\le rn}} \frac{|f(x) - f(y)|^p}{|x-y|^\gamma}.$$

Now, $|f(x) - f(y)|^p \le C m^{p-1} \sum_{i=1}^m G(T_{x_i}\omega)$, where G was defined in (B.5) and (x_i) is any path in $\mathbb{Z}^d$ from x to y, with length $m \le C|x - y|$. If one chooses canonical paths that go from each x to each y and that stay as close as possible to the line connecting x and y, e.g., staying at distance less than d from the line, then the above sum is bounded by

$$\frac{1}{n^{2d+p-\gamma}} \sum_{s:\ |s|\le C_r n} A_{s,n,\gamma} G(T_s\omega),$$

where

$$A_{s,n,\gamma} = \sum_{\substack{x,y:\ x\ne y \\ |x|,|y|\le rn}} |x-y|^{p-1-\gamma} \mathbb{1}\{s \text{ is on the canonical path from } x \text{ to } y\}.$$

Consider a fixed s. For a given integer ρ_1, there are at most $C_r \rho_1^{d-1}$ x's such that $|x - s| = \rho_1$. Fix such an x. See Fig. 1. Because the line joining x and y has to be within a bounded distance of s, radius R is bounded by

$$R \le \rho_2 \sin(\theta + \varphi) \le \rho_2(\sin\theta + \sin\varphi) \le C(1 + \rho_2/\rho_1).$$

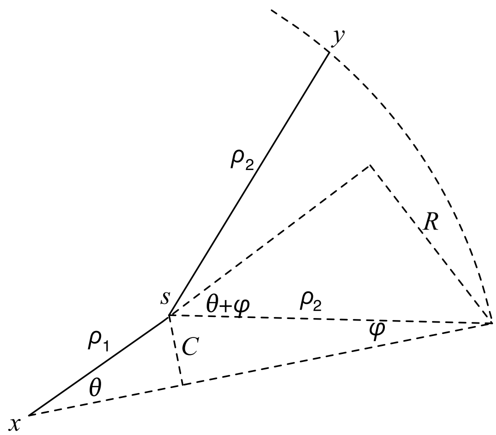

Fig. 1. y count.

Hence, there can be at most $C_r(1+\rho_2/\rho_1)^{d-1}$ possible y's with $|y-s|=\rho_2$ being a given integer. Thus, there are at most $C_r(\rho_1+\rho_2)^{d-1}$ pairs (x,y) that have s on the canonical path joining them. Furthermore, $\rho_1+\rho_2\le C_r|x-y|$. Therefore,

$$A_{s,n,\gamma}\le C_r\sum_{\rho_1,\rho_2=1}^{C_r n}(\rho_1+\rho_2)^{d+p-2-\gamma}\le C_r n^{d+p-\gamma}.$$

This allows us to bound the above sum by

$$C_r n^{-d}\sum_{s\in\mathbb{Z}^d:\ |s|\le C_r n}G(T_s\omega),$$

which, by the ergodic theorem, converges to a constant. □

We have shown that for a fixed $r\ge 1$, if $d+p-1<\gamma<d+p$, then for all $z_{1,\ell}$ and $\bar z_{1,\ell}$ and $\mathbb{P}$-a.e. ω

$$\sup_n\int_{B_r(0)}\int_{B_r(0)}\frac{|g_n(\omega,z_{1,\ell},\bar z_{1,\ell},\xi)-g_n(\omega,z_{1,\ell},\bar z_{1,\ell},\zeta)|^p}{|\xi-\zeta|^\gamma}\,\mathrm{d}\zeta\,\mathrm{d}\xi\le C_r(\omega).$$

Next, we apply an extension of Theorem 2.1.3 in [20]; see Exercise 2.4.1 therein.

Garsia–Rodemich–Rumsey's theorem. *Let $g:\mathbb{R}^d\to\mathbb{R}$ be a continuous function on $B_r(0)$ for some $r>0$. Let $\gamma>2d$. If*

$$\int_{B_r(0)}\int_{B_r(0)}\frac{|g(\xi)-g(\zeta)|^p}{|\xi-\zeta|^\gamma}\,\mathrm{d}\zeta\,\mathrm{d}\xi\le C_r,$$

then for $\xi,\zeta\in B_{r/2}(0)$,

$$\bigl|g(\xi)-g(\zeta)\bigr|\le C_r'|\xi-\zeta|^{(\gamma-2d)/p},$$

where C_r' depends on C_r and the dimension d.

From this theorem it follows that

$$\sup_n\bigl|g_n(\omega,z_{1,\ell},\bar z_{1,\ell},\xi)-g_n(\omega,z_{1,\ell},\bar z_{1,\ell},\zeta)\bigr|\le C_r(\omega)|\xi-\zeta|^{(\gamma-2d)/p}.$$

Since $2d<d+p$, there exists a suitable γ such that $\gamma-2d>0$. This shows that $\{g_n(\omega,z_{1,\ell},\bar z_{1,\ell},\xi)\}$ is equicontinuous in $\xi\in B_r(0)$, for all $r\ge 1$. Let $g(\omega,z_{1,\ell},\bar z_{1,\ell},\xi)$ be a uniform (on compacts) limit point, for fixed ω.

Now compute, for any fixed $i_0\in\{1,\dots,d\}$ and $\xi=\sum_{i=1}^d\xi_i e_i$ with $\xi_i\ge 0$,

$$\begin{aligned}
&\Biggl|n^{-(d-1)}\sum_{\substack{0\le k_i<[n\xi_i]\\1\le i\le d,i\ne i_0}}g_n\biggl(\frac{[n\xi_{i_0}]}{n}e_{i_0}+\sum_{j\ne i_0}\frac{k_j}{n}e_j\biggr)-\int\prod_{i\ne i_0}\mathbb{1}\{0\le\zeta_i\le\xi_i\}g\biggl(\xi_{i_0}e_{i_0}+\sum_{j\ne i_0}\zeta_je_j\biggr)\prod_{i\ne i_0}\mathrm{d}\zeta_i\Biggr|\\
&\quad\le\Biggl|n^{-(d-1)}\sum_{\substack{0\le k_i<[n\xi_i]\\1\le i\le d,i\ne i_0}}g_n\biggl(\frac{[n\xi_{i_0}]}{n}e_{i_0}+\sum_{j\ne i_0}\frac{k_j}{n}e_j\biggr)-n^{-(d-1)}\sum_{\substack{0\le k_i<[n\xi_i]\\1\le i\le d,i\ne i_0}}g\biggl(\frac{[n\xi_{i_0}]}{n}e_{i_0}+\sum_{j\ne i_0}\frac{k_j}{n}e_j\biggr)\Biggr| &\text{(B.7)}\\
&\qquad+\Biggl|n^{-(d-1)}\sum_{\substack{0\le k_i<[n\xi_i]\\1\le i\le d,i\ne i_0}}g\biggl(\frac{[n\xi_{i_0}]}{n}e_{i_0}+\sum_{j\ne i_0}\frac{k_j}{n}e_j\biggr)\\
&\qquad-\int\prod_{i\ne i_0}\mathbb{1}\{0\le\zeta_i\le\xi_i\}g\biggl(\xi_{i_0}e_{i_0}+\sum_{j\ne i_0}\zeta_je_j\biggr)\prod_{i\ne i_0}\mathrm{d}\zeta_i\Biggr|. &\text{(B.8)}
\end{aligned}$$

The term on line (B.7) converges to 0 because of the uniform convergence of g_n to g and the term on line (B.8) converges to 0 because g is continuous and the sum is a Riemann sum. Similarly,

$$\left| n^{-(d-1)} \sum_{\substack{0\le k_i<[n\xi_i]\\ 1\le i\le d, i\ne i_0}} g_n\Big(\sum_{j\ne i_0}\frac{k_j}{n}e_j\Big) - \int \prod_{i\ne i_0} \mathbb{1}\{0\le \zeta_i\le \xi_i\} g\Big(\sum_{j\ne i_0}\zeta_j e_j\Big)\prod_{i\ne i_0}\mathrm{d}\zeta_i\right|$$

converges to 0, as $n\to\infty$.

On the other hand,

$$\begin{aligned}
&n^{-(d-1)} \sum_{\substack{0\le k_i<[n\xi_i]\\ 1\le i\le d, i\ne i_0}} g_n\Big(\frac{[n\xi_{i_0}]}{n}e_{i_0} + \sum_{j\ne i_0}\frac{k_j}{n}e_j\Big) - n^{-(d-1)} \sum_{\substack{0\le k_i<[n\xi_i]\\ 1\le i\le d, i\ne i_0}} g_n\Big(\sum_{j\ne i_0}\frac{k_j}{n}e_j\Big)\\
&\quad= n^{-d} \sum_{\substack{0\le k_i<[n\xi_i]\\ 1\le i\le d, i\ne i_0}} f\Big([n\xi_{i_0}]e_{i_0} + \sum_{j\ne i_0}k_j e_j\Big) - n^{-d} \sum_{\substack{0\le k_i<[n\xi_i]\\ 1\le i\le d, i\ne i_0}} f\Big(\sum_{j\ne i_0}k_j e_j\Big)\\
&\quad= n^{-d} \sum_{\substack{0\le k_i<[n\xi_i]\\ 1\le i\le d}} \Big\{ f\Big(e_{i_0} + \sum_{j}k_j e_j\Big) - f\Big(\sum_{j}k_j e_j\Big)\Big\}\\
&\quad= n^{-d}\sum_{x\in V_n} G'(T_x\omega),
\end{aligned}$$

where $G'(\omega) = f(\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, e_{i_0}) \in L^1(\mathbb{P})$ and

$$V_n = \Big\{ x = \sum_{i=1}^d k_i e_i \colon\ 0\le k_i < [n\xi_i], \forall i\Big\}.$$

For the last equality above we used (b) of Lemma B.1. By (c) of Lemma B.1 we have $\mathbb{E}[G']=0$ and the ergodic theorem implies that the above converges to 0.

We have thus shown that

$$\int \prod_{i\ne i_0} \mathbb{1}\{0\le \zeta_i\le \xi_i\} g\Big(\xi_{i_0}e_{i_0} + \sum_{j\ne i_0}\zeta_j e_j\Big)\prod_{i\ne i_0}\mathrm{d}\zeta_i = \int \prod_{i\ne i_0} \mathbb{1}\{0\le \zeta_i\le \xi_i\} g\Big(\sum_{j\ne i_0}\zeta_j e_j\Big)\prod_{i\ne i_0}\mathrm{d}\zeta_i$$

which implies that

$$\int \prod_{i\ne i_0} \mathbb{1}\{\xi_i'\le \zeta_i\le \xi_i\} g\Big(\xi_{i_0}e_{i_0} + \sum_{j\ne i_0}\zeta_j e_j\Big)\prod_{i\ne i_0}\mathrm{d}\zeta_i = \int \prod_{i\ne i_0} \mathbb{1}\{\xi_i'\le \zeta_i\le \xi_i\} g\Big(\xi_{i_0}'e_{i_0} + \sum_{j\ne i_0}\zeta_j e_j\Big)\prod_{i\ne i_0}\mathrm{d}\zeta_i$$

and hence g is independent of ξ_{i_0}, for all $i_0\in\{1,\dots,d\}$. This means $g(\omega, z_{1,\ell}, \bar z_{1,\ell}, \xi) = g(\omega, z_{1,\ell}, \bar z_{1,\ell}, 0) = 0$. In other words, g_n converges uniformly (on compacts) to $g=0$. Lemma 5.1 is thus proved.

Proof of Lemma B.1. Recall that $\{e_1,\dots,e_d\}$ be the canonical basis of $\mathbb{R}^d$. For each $1\le i\le d$, there exist n_i, m_i, $(a_{i,j})_{j=1}^{n_i}$ and $(\bar a_{i,j})_{j=1}^{m_i}$ from $\mathscr{R}$ such that

$$e_i = \bar a_{i,1} + \cdots + \bar a_{i,m_i} - a_{i,1} - \cdots - a_{i,n_i}.$$

Write $x = \sum_{i=1}^d b_i e_i$. Then,

$$x = \sum_{i=1}^d\sum_{j=1}^{m_i} b_i \bar a_{i,j} - \sum_{i=1}^d\sum_{j=1}^{n_i} b_i a_{i,j}.$$

One can thus find a y that has a path to both 0 and x and such that $|y| \le C|x|$. This proves (a).

To prove (b), let $(y, \tilde z_{1,\ell})$ have a path to both $(x, \bar z_{1,\ell})$ and $(\bar x, \bar z_{1,\ell})$. Find $(y', \tilde z'_{1,\ell})$ that has a path to both $(y, \tilde z_{1,\ell})$ and $(0, z_{1,\ell})$. Then, from (B.3),

$$\begin{aligned}
&f(\omega, z_{1,\ell}, \bar z_{1,\ell}, \bar x) - f(\omega, z_{1,\ell}, \bar z_{1,\ell}, x)\\
&\quad= \bigl[L\bigl(\omega, \bigl(y', \tilde z'_{1,\ell}\bigr), (y, \tilde z_{1,\ell})\bigr) + L\bigl(\omega, (y, \tilde z_{1,\ell}), (\bar x, \bar z_{1,\ell})\bigr) - L\bigl(\omega, \bigl(y', \tilde z'_{1,\ell}\bigr), (0, z_{1,\ell})\bigr)\bigr]\\
&\qquad - \bigl[L\bigl(\omega, \bigl(y', \tilde z'_{1,\ell}\bigr), (y, \tilde z_{1,\ell})\bigr) + L\bigl(\omega, (y, \tilde z_{1,\ell}), (x, \bar z_{1,\ell})\bigr) - L\bigl(\omega, \bigl(y', \tilde z'_{1,\ell}\bigr), (0, z_{1,\ell})\bigr)\bigr]\\
&\quad= L\bigl(\omega, (y, \tilde z_{1,\ell}), (\bar x, \bar z_{1,\ell})\bigr) - L\bigl(\omega, (y, \tilde z_{1,\ell}), (x, \bar z_{1,\ell})\bigr).
\end{aligned}$$

The last line above is independent of $z_{1,\ell}$ so we can substitute $\bar z_{1,\ell}$ for $z_{1,\ell}$ and get the second equality of part (b). For the first equality, by the definition of f (B.3), the shift property (B.2), and the second equality in (b) just proved, we have for a new $(y, \tilde z_{1,\ell})$

$$\begin{aligned}
&f(T_x\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, \bar x - x)\\
&\quad= L\bigl(T_x\omega, (y, \tilde z_{1,\ell}), (\bar x - x, \bar z_{1,\ell})\bigr) - L\bigl(T_x\omega, (y, \tilde z_{1,\ell}), (0, \bar z_{1,\ell})\bigr)\\
&\quad= L\bigl(\omega, (y + x, \tilde z_{1,\ell}), (\bar x, \bar z_{1,\ell})\bigr) - L\bigl(\omega, (y + x, \tilde z_{1,\ell}), (x, \bar z_{1,\ell})\bigr)\\
&\quad= f(\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, \bar x) - f(\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, x).
\end{aligned}$$

For (c), by the earlier observation, we can choose y so that from $(y, \bar z_{1,\ell})$ there is a path to both $(x, \bar z_{1,\ell})$ and $(0, \bar z_{1,\ell})$. Then

$$f(\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, x) = L\bigl(\omega, (y, \bar z_{1,\ell}), (x, \bar z_{1,\ell})\bigr) - L\bigl(\omega, (y, \bar z_{1,\ell}), (0, \bar z_{1,\ell})\bigr).$$

Both L-terms above equal sums $\sum_{i=0}^{m-1} F(\eta_i, a_{i+1})$ where $\eta_0 = (T_y\omega, \bar z_{1,\ell})$ and $\eta_m = (T_u\omega, \bar z_{1,\ell})$ with $u = x$ or $u = 0$. Both have $\mathbb{E}$-mean zero by property (ii) of Definition 5.1. □

References

[1] C. Boldrighini, R. A. Minlos and A. Pellegrinotti. Random walks in quenched i.i.d. space–time random environment are always a.s. diffusive. *Probab. Theory Related Fields* **129** (2004) 133–156. MR2052866

[2] F. Comets, N. Gantert and O. Zeitouni. Quenched, annealed and functional large deviations for one-dimensional random walk in random environment. *Probab. Theory Related Fields* **118** (2000) 65–114. MR1785454

[3] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, 2nd edition. *Applications of Mathematics* **38**. Springer, New York, 1998. MR1619036

[4] F. den Hollander. *Large Deviations*. *Fields Institute Monographs* **14**. Amer. Math. Soc., Providence, RI, 2000. MR1739680

[5] J.-D. Deuschel and D. W. Stroock. *Large Deviations*. *Pure and Applied Mathematics* **137**. Academic Press, Boston, MA, 1989. MR0997938

[6] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. *Comm. Pure Appl. Math.* **28** (1975) 1–47. MR0386024

[7] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. III. *Comm. Pure Appl. Math.* **29** (1976) 389–461. MR0428471

[8] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*, English edition. *Classics in Applied Mathematics* **28**. SIAM, Philadelphia, PA, 1999. MR1727362

[9] H.-O. Georgii. *Gibbs Measures and Phase Transitions*. *de Gruyter Studies in Mathematics* **9**. Walter de Gruyter, Berlin, 1988. MR0956646

[10] A. Greven and F. den Hollander. Large deviations for a random walk in random environment. *Ann. Probab.* **22** (1994) 1381–1428. MR1303649

[11] G. Kassay. A simple proof for König's minimax theorem. *Acta Math. Hungar.* **63** (1994) 371–374. MR1261480

[12] E. Kosygina, F. Rezakhanlou and S. R. S. Varadhan. Stochastic homogenization of Hamilton–Jacobi–Bellman equations. *Comm. Pure Appl. Math.* **59** (2006) 1489–1521. MR2248897

[13] F. Rassoul-Agha. The point of view of the particle on the law of large numbers for random walks in a mixing random environment. *Ann. Probab.* **31** (2003) 1441–1463. MR1989439

[14] F. Rassoul-Agha and T. Seppäläinen. An almost sure invariance principle for random walks in a space–time random environment. *Probab. Theory Related Fields* **133** (2005) 299–314. MR2198014

[15] F. Rassoul-Agha and T. Seppäläinen. A course on large deviation theory with an introduction to Gibbs measures. Preprint, 2009.

[16] M. Rosenblatt. *Markov Processes. Structure and Asymptotic Behavior*. Springer, New York, 1971. MR0329037

[17] J. Rosenbluth. Quenched large deviations for multidimensional random walk in random environment: A variational formula. Thesis dissertation, New York University, 2006. Available at http://arxiv.org/abs/0804.1444. MR2623687
[18] W. Rudin. *Functional Analysis*, 2nd edition. McGraw-Hill, New York, 1991. MR1157815
[19] T. Seppäläinen. Large deviations for lattice systems. I. Parametrized independent fields. *Probab. Theory Related Fields* **96** (1993) 241–260. MR1227034
[20] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. Springer, Berlin, 2006. MR2190038
[21] S. R. S. Varadhan. *Large Deviations and Applications*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **46**. SIAM, Philadelphia, PA, 1984. MR0758258
[22] S. R. S. Varadhan. Large deviations for random walks in a random environment. *Comm. Pure Appl. Math.* **56** (2003) 1222–1245. Dedicated to the memory of Jürgen K. Moser. MR1989232
[23] A. Yilmaz. Large deviations for random walk in a space–time product environment. *Ann. Probab.* **37** (2009a) 189–205. MR2489163
[24] A. Yilmaz. Quenched large deviations for random walk in a random environment. *Comm. Pure Appl. Math.* **62** (2009b) 1033–1075. MR2531552
[25] M. P. W. Zerner. Lyapounov exponents and quenched large deviations for multidimensional random walk in random environment. *Ann. Probab.* **26** (1998) 1446–1476. MR1675027