Quenched Free Energy and Large Deviations for Random Walks in Random Potentials

FIRAS RASSOUL-AGHA
University of Utah

TIMO SEPPÄLÄINEN
University of Wisconsin-Madison

AND

ATILLA YILMAZ
University of California-Berkeley

Abstract

We study quenched distributions on random walks in a random potential on integer lattices of arbitrary dimension and with an arbitrary finite set of admissible steps. The potential can be unbounded and can depend on a few steps of the walk. Directed, undirected, and stretched polymers, as well as random walk in random environment, are covered. The restriction needed is on the moment of the potential, in relation to the degree of mixing of the ergodic environment. We derive two variational formulas for the limiting quenched free energy and prove a process-level quenched large deviation principle (LDP) for the empirical measure. As a corollary we obtain LDPs for types of random walks in random environments not covered by earlier results. © 2012 Wiley Periodicals, Inc.

1 Introduction

This paper investigates the limiting free energy and large deviations for several much-studied lattice models of random motion in a random medium. These include walks in random potentials, also called polymer models, and the standard random walk in a random environment (RWRE). We derive variational formulas for the free energy and process-level large deviations for the empirical measure.

1.1 Walks in Random Potentials and Environments

We call our basic model *random walk in a random potential* (RWRP). A special case is random walk in a random environment (RWRE). Fix a dimension $d \in \mathbb{N}$. There are three ingredients to the model: (i) a reference random walk on $\mathbb{Z}^d$, (ii) an environment, and (iii) a potential.

(i) Fix a finite subset $\mathscr{R} \subset \mathbb{Z}^d$. Let P_x denote the distribution of the discrete-time random walk on $\mathbb{Z}^d$ that starts at x and has jump probability $\widehat{p}(z) = 1/|\mathscr{R}|$ for $z \in \mathscr{R}$ and $\widehat{p}(z) = 0$ otherwise. E_x is expectation under P_x. The walk is

Communications on Pure and Applied Mathematics, Vol. LXVI, 0202–0244 (2013)

denoted by $X_{0,\infty} = (X_n)_{n\geq 0}$. Let $\mathcal{G}$ be the additive subgroup of $\mathbb{Z}^d$ generated by $\mathscr{R}$.

(ii) An *environment* ω is a sample point from a probability space $(\Omega, \mathfrak{S}, \mathbb{P})$. Ω comes equipped with a group $\{T_z : z \in \mathcal{G}\}$ of measurable commuting bijections that satisfy $T_{x+y} = T_x T_y$, and T_0 is the identity. $\mathbb{P}$ is a $\{T_z : z \in \mathcal{G}\}$–invariant probability measure on $(\Omega, \mathfrak{S})$ that is ergodic under this group. In other words, if $A \in \mathfrak{S}$ satisfies $T_z A = A$ for all $z \in \mathcal{G}$, then $\mathbb{P}(A) = 0$ or 1. $\mathbb{E}$ will denote expectation relative to $\mathbb{P}$. We call $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z : z \in \mathcal{G}\})$ a *measurable ergodic dynamical system*.

(iii) A *potential* is a measurable function $V : \Omega \times \mathscr{R}^\ell \to \mathbb{R}$ for some integer $\ell \geq 0$.

Given an environment ω and a starting point $x \in \mathbb{Z}^d$, for $n \geq 1$ define the *quenched polymer measures*

$$Q_{n,x}^{V,\omega}\{X_{0,\infty} \in A\} = \frac{1}{Z_{n,x}^{V,\omega}} E_x\big[e^{-\sum_{k=0}^{n-1} V(T_{X_k}\omega, Z_{k+1,k+\ell})} \mathbb{1}_A(X_{0,\infty})\big] \tag{1.1}$$

normalized by the *quenched partition function*

$$Z_{n,x}^{V,\omega} = E_x\big[e^{-\sum_{k=0}^{n-1} V(T_{X_k}\omega, Z_{k+1,k+\ell})}\big] = \sum_{z_{1,n+\ell-1} \in \mathscr{R}^{n+\ell-1}} |\mathscr{R}|^{-n-\ell+1}\, e^{-\sum_{k=0}^{n-1} V(T_{x_k}\omega, z_{k+1,k+\ell})}.$$

$Z_k = X_k - X_{k-1}$ is a step of the walk and vectors are $X_{i,j} = (X_i, X_{i+1}, \dots, X_j)$. $Q_{n,x}^{V,\omega}$ represents the evolution of the polymer in a "frozen" environment ω. (The picture is that of a heated sword quenched in water.) Let us mention two models of special importance.

Example 1.1 ($k+1$–Dimensional Directed Polymer in a Random Environment). Take the canonical setting: product space $\Omega = \Gamma^{\mathbb{Z}^d}$ with generic points $\omega = (\omega_x)_{x\in\mathbb{Z}^d}$ and translations $(T_x\omega)_y = \omega_{x+y}$. Then let $d = k+1$, $V(\omega) = -\beta\omega_0$ with inverse temperature parameter β, $\mathscr{R} = \{e_i + e_{k+1} : 1 \leq i \leq k\}$, and the coordinates $\{\omega_x\}$ i.i.d. under $\mathbb{P}$. Thus the projection of the walk on $\mathbb{Z}^k$ is a simple random walk, and at every step the walk sees a fresh environment.

Example 1.2 (Random Walk in a Random Environment). RWRE is a Markov chain X_n on $\mathbb{Z}^d$ whose transition probabilities are determined by an environment $\omega \in \Omega$. Let $\mathcal{P} = \{(\rho_z)_{z\in\mathscr{R}} \in [0,1]^{\mathscr{R}} : \sum_z \rho_z = 1\}$ be the set of probability distributions on $\mathscr{R}$ and $p : \Omega \to \mathcal{P}$ a measurable function with $p(\omega) = (p_z(\omega))_{z\in\mathscr{R}}$. A transition probability matrix is defined by

$$\pi_{x,y}(\omega) = \begin{cases} p_{y-x}(T_x\omega), & y-x \in \mathscr{R}, \\ 0, & y-x \notin \mathscr{R}, \end{cases} \quad \text{for } x, y \in \mathbb{Z}^d.$$

Given ω and $x \in \mathbb{Z}^d$, P_x^ω is the law of the Markov chain $X_{0,\infty} = (X_n)_{n\ge 0}$ on $\mathbb{Z}^d$ with initial point $X_0 = x$ and transition probabilities $\pi_{y,z}(\omega)$. That is, P_x^ω satisfies $P_x^\omega\{X_0 = x\} = 1$ and

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

P_x^ω is called the *quenched* distribution of the walk X_n. The *averaged* (or *annealed*) distribution is the path marginal $P_x(\cdot) = \int P_x^\omega(\cdot)\,\mathbb{P}(d\omega)$ of the *joint* distribution $P_x(dx_{0,\infty}, d\omega) = P_x^\omega(dx_{0,\infty})\mathbb{P}(d\omega)$.

RWRE is a special case of (1.1) with $V(\omega, z_{1,\ell}) = -\log \pi_{0,z_1}(\omega)$. (Note the abuse of notation: for RWRE P_0 is the averaged measure while in RWRP P_0 is the reference random walk. This should cause no confusion.)

Of particular interest are RWREs where 0 lies outside the convex hull of $\mathscr{R}$. These are *strictly directed* in the sense that for some $\hat{u} \in \mathbb{R}^d$, $z \cdot \hat{u} > 0$ for each admissible step $z \in \mathscr{R}$. General large deviation theory for these walks is covered for the first time in the present paper.

1.2 Results

We have two types of results. First we prove the $\mathbb{P}$-a.s. existence of the *quenched free energy*

$$\lim_{n\to\infty} n^{-1} \log Z_{n,0}^{V,\omega} = \lim_{n\to\infty} n^{-1} \log E_0\big[e^{-\sum_{k=0}^{n-1} V(T_{X_k}\omega, Z_{k+1,k+\ell})}\big] \tag{1.2}$$

and derive two variational formulas for the limit. The assumption we need combines moment bounds on V with the degree of mixing in $\mathbb{P}$: if $\mathbb{P}$ is merely ergodic we require a bounded V, while with independence or exponential mixing L^p for $p > d$ is sufficient. The existence of the limit is not entirely new because in some cases it follows from subadditive methods and concentration inequalities. In Example 1.1, [8] proved the limit under an exponential moment assumption and [47] with the tail assumption under which greedy lattice animals are known to have linear growth. Our variational descriptions of the free energy are new.

The second results are large deviation principles (LDPs) for the quenched distributions $Q_{n,0}^{V,\omega}\{R_n^\infty \in \cdot\}$ of the *empirical process*

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

$T_{X_k}\omega$ is the environment seen from the current position of the walk, and $Z_{k+1,\infty} = (Z_i)_{k+1\le i<\infty}$ is the entire sequence of future steps. We assume Ω separable metric with Borel σ-algebra $\mathfrak{S}$. Distributions $Q_{n,0}^{V,\omega}\{R_n^\infty \in \cdot\}$ are probability measures on $\mathcal{M}_1(\Omega \times \mathscr{R}^{\mathbb{N}})$, the space of Borel probability measures on $\Omega \times \mathscr{R}^{\mathbb{N}}$ endowed with the weak topology generated by bounded continuous functions.

The LDP takes this standard form. There is a lower-semicontinuous convex rate function $I^V_{q,3} : \mathcal{M}_1(\Omega \times \mathscr{R}^{\mathbb{N}}) \to [0,\infty]$ such that these bounds hold:

$$\overline{\lim_{n\to\infty}}\, n^{-1} \log Q^{V,\omega}_{n,0}\{R^\infty_n \in C\} \le -\inf_{\mu\in C} I^V_{q,3}(\mu) \quad \text{for all compact sets } C,$$

$$\underline{\lim_{n\to\infty}}\, n^{-1} \log Q^{V,\omega}_{n,0}\{R^\infty_n \in O\} \ge -\inf_{\mu\in O} I^V_{q,3}(\mu) \quad \text{for all open sets } O.$$

Large deviations of R^∞_n are called *level 3* or *process level* large deviations. For basic large deviation theory we refer the reader to [11, 12, 14, 31, 45].

Since we prove the upper bound only for compact sets, the result is technically known as a weak LDP. In the important special case of strictly directed walk in an i.i.d. environment, we strengthen the result to a full LDP where the upper bound is valid for all closed sets. Often Ω is compact and then this issue vanishes. As a corollary we obtain large deviations for RWRE.

This paper does not investigate models that allow $V = \infty$. An example in RWRE would be a walk on a supercritical percolation cluster.

1.3 Overview of Literature and Predecessors of This Work

Random walk in a random environment was introduced by Chernov [6] in 1967 and Temkin [44] in 1972 as a model for DNA replication. Random walk in random potential appeared in the work of Huse and Henley [24] in 1985 on impurity-induced domain-wall roughening in the two-dimensional Ising model. The seminal mathematical work on RWRE was Solomon in 1975 [39], and on RWRP, Imbrie and Spencer in 1988 [25] and Bolthausen in 1989 [3]. Despite a few decades of effort many basic questions on (1) recurrence, transience, and zero-one laws, (2) fluctuation behavior, and (3) large deviations remain only partially answered. Accounts of parts of the state of the art can be found in the lectures [4, 23, 43, 50] on RWRE, and in [9, 13, 21, 36, 42] on RWRP.

Our LDP, Theorem 3.1, specialized to RWRE covers the quenched level 1 LDPs for RWRE that have been established over the last two decades. In the one-dimensional case, Greven and den Hollander [22] considered the i.i.d. nearest-neighbor case; Comets, Gantert, and Zeitouni [7] the ergodic nearest-neighbor case; and Yilmaz [49] the ergodic case with $\mathscr{R} = \{z : |z| \le M\}$ for some M. In the multidimensional setting Zerner [52] looked at the i.i.d. nearest-neighbor nestling case, and Varadhan [46] the general ergodic case with bounded step size and $\{z : |z| = 1\} \subset \mathscr{R}$. All these works, with the exception of [52], required uniform ellipticity at least on part of $\mathscr{R}$, i.e., $\pi_{0,z} \ge \kappa$ for a fixed $\kappa > 0$ and all z with $|z| = 1$. [52] needs $\mathbb{E}[|\log \pi_{0,z}|^d] < \infty$, still for all $|z| = 1$. Rosenbluth [35] gave a variational formula for the rate function in [46] under an assumption of $p > d$ moments on $\log \pi_{0,z}$, $|z| = 1$.

Article [32] proved a quenched level 3 LDP for RWRE under a general ergodic environment, subject to bounded steps, $p > d$ moments on $\log \pi_{0,z}$, and an irreducibility assumption that required the origin to be accessible from every $x \in \mathbb{Z}^d$.

(See Remark 2.7 for a more technical explanation of the scope of [32] compared to the present paper.) Level 3 large deviations for RWRE have not appeared in other works. [49] gave a quenched univariate level 2 LDP. This means that the path component in the empirical measure has only one step: $n^{-1}\sum_{k=0}^{n-1}\delta_{T_{X_k}\omega,Z_{k+1}}$.

One goal of the present paper is to eliminate the unsatisfactory irreducibility assumption of [32, 49]. This is important because the irreducibility assumption excluded several basic and fruitful models, such as directed polymers, RWRE in a space-time, or dynamical, environment (the case $\mathscr{R}\subset\{x : x\cdot e_1 = 1\}$) and RWRE with a forbidden direction (the case $\mathscr{R}\subset\{x : x\cdot\hat{u}\geq 0\}$ for some $\hat{u}\neq 0$). Corollary A.3 shows that the forbidden direction is the only case not covered by [32], but [32] did not address the more general polymer model.

Our results cover the quenched level 1 LDPs for space-time RWRE derived in [48] for an i.i.d. environment in a neighborhood of the asymptotic velocity and by Avena, den Hollander, and Redig [2] for a space-time random environment given by a mixing attractive spin flip particle system. Our results can also be adapted to continuous time to cover the quenched level 1 LDP by Drewitz and others [15] for a random walk among a Poisson system of moving traps.

On the RWRP side, Theorems 2.3 and 3.1 cover, respectively, the existence of free energy and the quenched level 1 LDPs for simple random walk in random potential proved by Zerner [51] and the corresponding results for directed simple random walk in random potential proved by Carmona and Hu [5] and Comets, Shiga, and Yoshida [8]. (See [41] for an earlier continuous counterpart of [51].) We also give an entropy interpretation for the rate function and two variational formulas for the free energy, while earlier descriptions of these objects came in terms of Lyapunov functions and subadditivity arguments. As far as we know, level 2 or 3 large deviations have not been established in the past for RWRP.

The technical heart of [32] was a multidimensional extension of a homogenization argument that goes back to Kosygina, Rezakhanlou, and Varadhan [27] in the context of diffusion in time-independent random potential. This argument was used by Rosenbluth [35] and Yilmaz [49] to prove LDPs for RWRE.

The main technical contribution of the current work is a new approach to the homogenization argument that allows us to drop the aforementioned irreducibility requirement. One comment to make is that the construction that we undertake in Appendix C does use the invertibility of the transformations T_z assumed in Section 1.1. This is the only place where that is needed.

The homogenization method of [27] was sharpened by Kosygina and Varadhan [28] to handle time-dependent but bounded random potentials. The results in [27, 28] concerned homogenization of stochastic Hamilton-Jacobi-Bellman equations and yielded variational formulas for the effective Hamiltonian. For a special case of the random Hamiltonian one can convert these results into quenched large deviations for the velocity of a diffusion in a random potential, with variational formulas for the quenched free energy. Using different methods, [29] and [1] obtain homogenization results similar to [27, 28], respectively. Furthermore, [1] allows unbounded potentials and requires mixing to compensate for the unboundedness; compare with part (d) of our Lemma A.4. It is noteworthy that when $d = 1$, an ergodic L^1 potential is in fact enough; see [18] and compare with part (b) of our Lemma A.4.

We end this section with some conventions for easy reference. For a measurable space $(\mathcal{X}, \mathscr{B})$, $\mathcal{M}_1(\mathcal{X})$ is the space of probability measures on $\mathcal{X}$ and $\mathcal{Q}(\mathcal{X})$ the set of Markov transition kernels on $\mathcal{X}$. 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 $\mu \times q(A \times B) = \int \mathbb{1}_A(x) q(x, B)\, \mu(dx)$, and μq is its second marginal. $E^\mu[f]$ denotes expectation of f under probability measure μ. The increments of a path (x_i) in $\mathbb{Z}^d$ are denoted by $z_i = x_i - x_{i-1}$. Segments of sequences are denoted by $z_{i,j} = (z_i, z_{i+1}, \dots, z_j)$, also for $j = \infty$.

2 Variational Representations for Free Energy

Standing assumptions in this section are that $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z : z \in \mathcal{G}\})$ is a measurable ergodic dynamical system and, as throughout the paper, $\mathscr{R}$ is an arbitrary finite subset of $\mathbb{Z}^d$ that generates the additive group $\mathcal{G}$. These will not be repeated in the statements of lemmas and theorems. Most of the time we also assume that $\mathfrak{S}$ is countably generated; this will be mentioned. The relevant Markov process for this analysis is $(T_{X_n}\omega, Z_{n+1,n+\ell})$ with state space $\mathbf{\Omega}_\ell = \Omega \times \mathscr{R}^\ell$. The evolution goes via the transformations $S_z^+(\omega, z_{1,\ell}) = (T_{z_1}\omega, (z_{2,\ell}, z))$ on $\mathbf{\Omega}_\ell$ where the step z is chosen randomly from $\mathscr{R}$ as stipulated by the kernel $\hat{p}$. Elements of $\mathbf{\Omega}_\ell$ are abbreviated $\eta = (\omega, z_{1,\ell})$.

We first look at the limiting logarithmic moment-generating function (1.2), also called the *pressure* or the *free energy*. To cover much-studied directed polymer models it is important to go beyond bounded continuous potentials. To achieve this, and at the same time provide a succinct statement of a key hypothesis for Lemma 2.8 below, we introduce class $\mathcal{L}$ in the next definition. Let

$$D_n = \{z_1 + \cdots + z_n \in \mathbb{Z}^d : z_{1,n} \in \mathscr{R}^n\} \tag{2.1}$$

denote the set of points accessible from the origin in exactly n steps from $\mathscr{R}$.

DEFINITION 2.1. A function $g : \Omega \to \mathbb{R}$ *is in class* $\mathcal{L}$ if $g \in L^1(\mathbb{P})$ and for any nonzero $z \in \mathscr{R}$

$$\varlimsup_{\varepsilon \to 0} \varlimsup_{n \to \infty} \max_{x \in \bigcup_{k=0}^n D_k} \frac{1}{n} \sum_{0 \le i \le \varepsilon n} |g \circ T_{x+iz}| = 0 \quad \mathbb{P}\text{-a.s.}$$

Similarly, a function g *on* $\mathbf{\Omega}_\ell$ *is a member of* $\mathcal{L}$ if $g(\,\cdot\,, z_{1,\ell}) \in \mathcal{L}$ for each $z_{1,\ell} \in \mathscr{R}^\ell$.

A bounded g is in $\mathcal{L}$ under an arbitrary ergodic $\mathbb{P}$, and so is any $g \in L^1(\mathbb{P})$ if $d = 1$. In general, there is a tradeoff between the degree of mixing in $\mathbb{P}$ and the moment of g required. For example, if sufficiently separated shifts of g are i.i.d. or there is exponential mixing, then $g \in L^p(\mathbb{P})$ for some $p > d$ guarantees $g \in \mathcal{L}$.

Under polynomial mixing a higher moment is needed. Lemma A.4 in Appendix A collects sufficient conditions for membership in $\mathcal{L}$.

We have two variational formulas for the free energy. One is duality in terms of entropy. The other involves a functional $K_\ell(g)$ defined by a minimization over gradientlike auxiliary functions. Class $\mathcal{K}_\ell$ below is a generalization of a class of functions previously introduced by [35].

DEFINITION 2.2. A measurable function $F : \mathbf{\Omega}_\ell \times \mathscr{R} \to \mathbb{R}$ *is in class* $\mathcal{K}_\ell$ if it satisfies the following three conditions:

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

(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 $i = 1, \dots, n$, then

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

In other words, expectation vanishes whenever the sequence of moves $S^+_{a_1}, \dots, S^+_{a_n}$ 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 $\{\overline{\eta}_j\}_{j=0}^m$ with $\eta_0 = \overline{\eta}_0 = (\omega, z_{1,\ell})$, $\eta_n = \overline{\eta}_m$, $\eta_i = S^+_{a_i}\eta_{i-1}$, and $\overline{\eta}_j = S^+_{\overline{a}_j}\overline{\eta}_{j-1}$, for $i, j > 0$ and some $\{a_i\}_{i=1}^n \in \mathscr{R}^n$ and $\{\overline{a}_j\}_{j=1}^m \in \mathscr{R}^m$, we have

$$\sum_{i=0}^{n-1} F(\eta_i, a_{i+1}) = \sum_{j=0}^{m-1} F(\overline{\eta}_j, \overline{a}_{j+1}).$$

In case of a loop ($\eta_0 = \eta_n$) in (iii) above, one can take $m = 0$ and the right-hand side in the display vanishes. The simplest members of $\mathcal{K}_\ell$ are gradients $F(\eta, z) = h(S^+_z\eta) - h(\eta)$ with bounded measurable $h : \mathbf{\Omega}_\ell \to \mathbb{R}$. Lemma C.3 in the appendix shows that $\mathcal{K}_\ell$ is the $L^1(\mathbb{P})$-closure of such gradients.

For $F \in \mathcal{K}_\ell$ and $g : \mathbf{\Omega}_\ell \to \mathbb{R}$ such that $g(\cdot, z_{1,\ell}) \in L^1(\mathbb{P})$ for all $z_{1,\ell} \in \mathscr{R}^\ell$, define

$$K_{\ell,F}(g) = \mathbb{P}\text{-}\operatorname*{ess\,sup}_{\omega} \max_{z_{1,\ell}} \log \sum_{z \in \mathscr{R}} \frac{1}{|\mathscr{R}|} e^{g(\eta) + F(\eta, z)}$$

and then

$$K_\ell(g) = \inf_{F \in \mathcal{K}_\ell} K_{\ell,F}(g).$$

The reference walk $\widehat{p}$ with uniform steps from $\mathscr{R}$ defines a Markov kernel $\widehat{p}_\ell$ on $\mathbf{\Omega}_\ell$ by

$$\widehat{p}_\ell(\eta, S^+_z\eta) = \frac{1}{|\mathscr{R}|} \quad \text{for } z \in \mathscr{R} \text{ and } \eta = (\omega, z_{1,\ell}) \in \mathbf{\Omega}_\ell. \tag{2.2}$$

Let μ_0 denote the Ω-marginal of a measure $\mu \in \mathcal{M}_1(\mathbf{\Omega}_\ell)$. Define an entropy $H_{\ell,\mathbb{P}}$ on $\mathcal{M}_1(\mathbf{\Omega}_\ell)$ by

$$(2.3) \qquad H_{\ell,\mathbb{P}}(\mu) = \begin{cases} \inf\{H(\mu \times q \mid \mu \times \widehat{p}_\ell) : q \in \mathcal{Q}(\mathbf{\Omega}_\ell) \text{ with } \mu q = \mu\} \\ \quad \text{if } \mu_0 \ll \mathbb{P}, \\ \infty \quad \text{otherwise.} \end{cases}$$

Inside the braces the familiar relative entropy is

$$H(\mu \times q \mid \mu \times \widehat{p}_\ell) = \int \sum_{z \in \mathscr{R}} q(\eta, S_z^+ \eta) \log \frac{q(\eta, S_z^+ \eta)}{\widehat{p}_\ell(\eta, S_z^+ \eta)} \, \mu(d\eta).$$

$H_{\ell,\mathbb{P}} : \mathcal{M}_1(\mathbf{\Omega}_\ell) \to [0, \infty]$ is convex. (The argument for this can be found at the end of section 4 in [32].) For measurable functions g on $\mathbf{\Omega}_\ell$ define

$$(2.4) \qquad H^{\#}_{\ell,\mathbb{P}}(g) = \sup_{\mu \in \mathcal{M}_1(\mathbf{\Omega}_\ell),\, c>0} \{E^{\mu}[\min(g, c)] - H_{\ell,\mathbb{P}}(\mu)\}.$$

For g from the space of bounded measurable functions (or bounded continuous functions if Ω comes with a metric) $H^{\#}_{\ell,\mathbb{P}}(g)$ is the convex dual of $H_{\ell,\mathbb{P}}$, and then we write $H^{*}_{\ell,\mathbb{P}}(g)$. The constant $R = \max\{|z| : z \in \mathscr{R}\}$ also appears frequently in the results.

For the rest of the section we fix $\ell \geq 0$ and consider measurable functions $g : \mathbf{\Omega}_\ell \to \mathbb{R}$.

THEOREM 2.3. *Assume $\mathfrak{S}$ is countably generated. Let $g \in \mathcal{L}$. Then, for $\mathbb{P}$-a.e. ω the limit*

$$\Lambda_\ell(g) = \lim_{n\to\infty} n^{-1} \log E_0\big[e^{\sum_{k=0}^{n-1} g(T_{X_k}\omega, Z_{k+1,k+\ell})}\big]$$

exists, is deterministic, and satisfies $\Lambda_\ell(g) = K_\ell(g) = H^{\#}_{\ell,\mathbb{P}}(g)$.

Remark 2.4. The limit $\Lambda_\ell(g)$ satisfies these bounds:

$$(2.5) \qquad \begin{aligned} \mathbb{E}[\min_{z_{1,\ell} \in \mathscr{R}^\ell} g(\omega, z_{1,\ell})] &\leq \Lambda_\ell(g) \\ &\leq \varlimsup_{n\to\infty} \max_{\substack{x_i - x_{i-1} \in \mathscr{R} \\ 1 \leq i \leq n}} n^{-1} \sum_{k=0}^{n-1} \max_{\tilde{z}_{1,\ell} \in \mathscr{R}^\ell} g(T_{x_k}\omega, \tilde{z}_{1,\ell}). \end{aligned}$$

The upper bound is nonrandom by invariance. The lower bound comes from ergodicity of the Markov chain $T_{X_n}\omega$ [32, lemma 4.1] and Jensen's inequality:

$$n^{-1} \log E_0\big[e^{\sum_{k=0}^{n-1} g(T_{X_k}\omega, Z_{k+1,k+\ell})}\big] \geq n^{-1} \sum_{k=0}^{n-1} E_0[\min_{z_{1,\ell}} g(T_{X_k}\omega, z_{1,\ell})].$$

If g is unbounded from above and $\mathscr{R}$ allows the walk to revisit sites, then a situation where $\Lambda_\ell(g) = \infty$ can be easily created. Under some independence and moment assumptions, the limit on the right in (2.5) is known to be a.s. finite.

Remark 2.5. Suppose Ω is a product space with i.i.d. coordinates $\{\omega_x\}$ under $\mathbb{P}$, the walk is strictly directed (0 does not lie in the convex hull of $\mathcal{R}$), and $g(\cdot, z_{1,\ell})$ is a local function on Ω. Then the assumption $\mathbb{E}|g(\cdot, z_{1,\ell})|^p < \infty$ for some $p > d$ and all $z_{1,\ell} \in \mathcal{R}^\ell$ is sufficient for the above Theorem 2.3 and the finiteness of the limit $\Lambda_\ell(g)$. That such $g \in \mathcal{L}$ is proved in Lemma A.4 in Appendix A. Under this moment bound, finiteness of the upper bound in (2.5) follows from lattice animal bounds [10, 19, 30].

Remark 2.6. If $\mathcal{R} = \{\pm e_1, \dots, \pm e_d\}$ or $\mathcal{R} = \{e_1 \pm e_2, \dots, e_1 \pm e_d\}$ and if we take g to be a function of ω only, then $\Lambda_0(g) = K_0(g) = H^\#_{0,\mathbb{P}}(g)$ corresponds to a discretization of the variational formula for the effective Hamiltonian $\overline{H}$ of the homogenized stochastic Hamilton-Jacobi-Bellman equation considered in [27, 28, 29]. It is also related to the variational formula for the exponential decay rate of the Green's function of Brownian motion in a periodic potential; see (1.1) in [38].

Here is an outline of the proof of Theorem 2.3. Introduce the empirical measure $R_n^\ell = n^{-1} \sum_{k=0}^{n-1} \delta_{T_{X_k}\omega, Z_{k+1,k+\ell}}$ so that $nR_n^\ell(g) = \sum_{k=0}^{n-1} g(T_{X_k}\omega, Z_{k+1,k+\ell})$ gives convenient compact notation for the sum in the exponent. Let

$$\overline{\Lambda}_\ell(g, \omega) = \varlimsup_{n\to\infty} n^{-1} \log E_0\big[e^{nR_n^\ell(g)}\big],$$

$$\underline{\Lambda}_\ell(g, \omega) = \varliminf_{n\to\infty} n^{-1} \log E_0\big[e^{nR_n^\ell(g)}\big].$$

The existence of $\Lambda_\ell(g)$ and the variational formulas are established through the inequalities

$$\overline{\Lambda}_\ell(g) \overset{\text{(i)}}{\le} K_\ell(g) \overset{\text{(ii)}}{\le} H^\#_{\ell,\mathbb{P}}(g) \overset{\text{(iii)}}{\le} \underline{\Lambda}_\ell(g). \tag{2.6}$$

Inequality (2.6.i) is proved in Lemma 2.11. This is the only step that requires $g \in \mathcal{L}$ rather than just $L^1(\mathbb{P})$. Inequality (2.6.ii) is proved in Lemma 2.12. This is where the main technical effort of the paper lies, in order to relax the irreducibility assumption on $\mathcal{R}$ used in [32]. Bound (2.6.iii) is proved with the usual change-of-measure argument. It follows as a special case from Lemma 2.15 below. The proof of Theorem 2.3 comes at the end of this section after the lemmas. To improve the readability of this section, some lemmas are proved in an appendix at the end of the paper.

Remark 2.7. Suppose 0 lies in the relative interior of the convex hull of $\mathcal{R}$. Then for every $x \in \mathcal{G}$ there exists $z_{1,n} \in \mathcal{R}^n$ with $x_n = x$ (Corollary A.3). Under this irreducibility the approach of [32] becomes available and can be used to prove our results under the assumption that $g(\cdot, z_{1,\ell}) \in L^p(\mathbb{P})$ for some $p > d$ and all $z_{1,\ell} \in \mathcal{R}^\ell$. In this case (2.6.i) is proved via a slight variation of [32, lemma 5.2] rather than our Lemma 2.11. This relies crucially on [32, lemma 5.1], which is where $p > d$ moments are required. We replace this with the much weaker Lemma 2.9, which only requires one moment, but then we need Lemma 2.11, which requires $g \in \mathcal{L}$.

We turn to developing inequalities (2.6). Decomposing the free energy according to asymptotic directions ξ turns out to be useful. Let $\mathcal{U}$ be the (compact) convex hull of $\mathscr{R}$ in $\mathbb{R}^d$. For each rational point $\xi \in \mathcal{U}$ fix a positive integer $b(\xi)$ such that $b(\xi)\xi \in D_{b(\xi)}$ (recall definition (2.1) of D_n). The existence of $b(\xi)$ follows from Lemma A.1 in Appendix A. Then fix a path $\{\hat{x}_n(\xi)\}_{n\in\mathbb{Z}_+}$, starting at $\hat{x}_0(\xi) = 0$, with admissible steps $\hat{x}_n(\xi) - \hat{x}_{n-1}(\xi) \in \mathscr{R}$ and such that $\hat{x}_{jb(\xi)}(\xi) = jb(\xi)\xi$ for all $j \in \mathbb{Z}_+$. Even though stationarity and ergodicity are standing assumptions in this section, the next lemma actually needs no assumptions on $\mathbb{P}$.

LEMMA 2.8. *Let $g \in \mathcal{L}$. Then for $\mathbb{P}$-a.e. ω*

$$\overline{\lim_{n\to\infty}}\, n^{-1}\log E_0\big[e^{nR_n^\ell(g)}\big] \le \sup_{\xi\in\mathcal{U}\cap\mathbb{Q}^d} \overline{\lim_{n\to\infty}}\, n^{-1}\log E_0\big[e^{nR_n^\ell(g)}\mathbb{1}\{X_n = \hat{x}_n(\xi)\}\big]. \tag{2.7}$$

PROOF. Fix a small $\varepsilon > 0$, an integer $k \ge |\mathscr{R}|\varepsilon^{-1}$, and a nonzero $\hat{z} \in \mathscr{R}$. For $x \in D_n$ write $x = \sum_{z\in\mathscr{R}} a_z z$ with $a_z \in \mathbb{Z}_+$ and $\sum_{z\in\mathscr{R}} a_z = n$. Let $m_n = \lceil n/(k(1-2\varepsilon))\rceil$ and $s_z^{(n)} = \lceil k(1-2\varepsilon)a_z/n\rceil$. Then $k^{-1}\sum_z s_z^{(n)} \le 1-\varepsilon$ and $m_n s_z^{(n)} \ge a_z$ for each $z \in \mathscr{R}$. Let

$$\xi(n,x) = k^{-1}\sum_{z\in\mathscr{R}} s_z^{(n)} z + \Big(1 - k^{-1}\sum_{z\in\mathscr{R}} s_z^{(n)}\Big)\hat{z}. \tag{2.8}$$

Then $\xi(n,x) \in k^{-1}D_k$. With ε fixed small enough and considering $n > k/\varepsilon$, we constructed an admissible path of $m_n k - n \le 4n\varepsilon$ steps from x to $m_n k\xi(n,x)$. This path has at least $m_n(k - \sum s_z^{(n)}) \ge m_n k\varepsilon \ge n\varepsilon/(1-2\varepsilon)$ $\hat{z}$-steps. Consequently, at least a fixed fraction δ of the steps of the path are $\hat{z}$-steps for all $x \in D_n$ and all n.

Let b be the least common multiple of the (finitely many) integers $\{b(\xi) : \xi \in k^{-1}D_k\}$. Now we take another bounded number of additional steps to get from $m_n k\xi(n,x)$ to the path $\hat{x}_\cdot(\xi(n,x))$. Pick ℓ_n such that $(\ell_n - 1)b < m_n \le \ell_n b$. Then by repeating the steps of $k\xi(n,x)$ in (2.8) $\ell_n b - m_n \le b$ times, we go from $m_n k\xi(n,x)$ to $\ell_n k b\xi(n,x) = \hat{x}_{\ell_n kb}(\xi(n,x))$. The duration of this last leg is bounded independently of n and $x \in D_n$ because k was fixed at the outset and b is determined by k. Thus the total number of steps from $x \in D_n$ to $\hat{x}_{\ell_n kb}(\xi(n,x))$ is $r_n = \ell_n kb - n \le 5n\varepsilon$ for large enough n. Let $\mathbf{u}(n,x) = (u_1, u_2, \dots, u_{r_n})$ denote this sequence of steps. Again we note that at least a fixed fraction δ of the u_i's are $\hat{z}$-steps.

Develop an estimate:

$$\begin{aligned}
\frac{1}{n}\log E_0\big[e^{nR_n^\ell(g)}\big] &= \frac{1}{n}\log\sum_{x\in D_n} E_0\big[e^{nR_n^\ell(g)}, X_n=x\big]\\
&\le \max_{x\in D_n}\frac{1}{n}\log E_0\big[e^{(n-\ell)R_{n-\ell}^\ell(g)}, X_n=x\big]\\
&\quad+\max_{w\in D_{n-\ell}}\max_{y\in\cup_{s=0}^{\ell}D_s}\frac{\ell}{n}\,\overline{g}(T_{w+y}\omega)+\frac{C\log n}{n}\\
&\le \max_{x\in D_n}\frac{1}{n}\log E_0\big[e^{\ell_n bkR_{\ell_n bk}^\ell(g)}, X_{\ell_n bk}=\hat{x}_{\ell_n bk}(\xi(n,x))\big]\\
&\quad+\frac{C\log n}{n}+\max_{w\in D_{n-\ell}}\max_{y\in\bigcup_{s=0}^{\ell}D_s}\frac{2\ell}{n}\,\overline{g}(T_{w+y}\omega)\\
&\quad+\max_{x\in D_n}\frac{1}{n}\sum_{i=1}^{r_n}\overline{g}(T_{x+u_1+\cdots+u_i}\omega)+\frac{r_n}{n}\log|\mathscr{R}|.
\end{aligned}$$

Above, $\overline{g}(\omega)=\max_{z_{1,\ell}\in\mathscr{R}^\ell}|g(\omega,z_{1,\ell})|$. The third-last line of the above display is bounded above by

$$\max_{\xi\in k^{-1}D_k}\frac{1}{n}\log E_0\big[e^{\ell_n bkR_{\ell_n bk}^\ell(g)}, X_{\ell_n bk}=\hat{x}_{\ell_n bk}(\xi)\big],$$

and so its limsup is almost surely at most

$$(1+5\varepsilon)\sup_{\xi\in\mathcal{U}\cap\mathbb{Q}^d}\varlimsup_{n\to\infty} n^{-1}\log E_0\big[e^{nR_n^\ell(g)}\mathbb{1}\{X_n=\hat{x}_n(\xi)\}\big].$$

The proof of (2.7) is complete once we show that a.s.

$$\begin{aligned}
&\varlimsup_{\varepsilon\to0}\varlimsup_{n\to\infty}\max_{x\in D_n}\frac{1}{n}\sum_{i=1}^{r_n}\overline{g}(T_{x+u_1+\cdots+u_i}\omega)=0,\\
&\varlimsup_{\varepsilon\to0}\varlimsup_{n\to\infty}\max_{w\in D_{n-\ell}}\max_{y\in\bigcup_{s=0}^{\ell}D_s}\frac{1}{n}\overline{g}(T_{w+y}\omega)=0.
\end{aligned}\tag{2.9}$$

To this end, observe that the ordering of the steps of $\mathbf{u}(n,x)$ was so far immaterial. Because Definition 2.1 cannot handle zero steps, we need to be careful about them. The ratio of zero steps to $\hat{z}$-steps is at most $t=\lceil\delta^{-1}\rceil$. We begin $\mathbf{u}(n,x)$ by alternating $\hat{z}$-steps with blocks of at most t zero steps, until the $\hat{z}$-steps and the zero steps are exhausted. After that, order the remaining nonzero steps of $\mathscr{R}$ in any fashion $z_1,z_2,\dots$, and have $\mathbf{u}(n,x)$ take first all its z_1-steps, then all its z_2-steps, and so on. Since zero steps do not shift ω but simply repeat the same $\overline{g}$-value at most t times, we get the bound

$$\sum_{i=1}^{r_n}\overline{g}(T_{x+u_1+\cdots+u_i}\omega)\le t|\mathscr{R}|\max_{y\in x+\mathbf{u}(n,x)}\max_{z\in\mathscr{R}\setminus\{0\}}\sum_{i=0}^{r_n}\overline{g}(T_{y+iz}\omega).$$

By $y \in x + \mathbf{u}(n,x)$ we mean y is on the path starting from x and taking steps in $\mathbf{u}(n,x)$. A similar bound develops for the second line of (2.9), and the limits in (2.9) follow from membership in $\mathcal{L}$. $\square$

The next step is to show (2.6.i): for $g \in \mathcal{L}$ and $\mathbb{P}$-a.e. ω, $\overline{\Lambda}_\ell(g,\omega) \le K_\ell(g)$. The following ergodic property is crucial. Recall the definition of the path $\hat{x}_\cdot(\xi)$ above Lemma 2.8. For $\xi \in \mathbb{Q}^d \cap \mathcal{U}$ and $z_{1,\ell}, \bar{z}_{1,\ell} \in \mathscr{R}^\ell$, define

$$\mathcal{A}_n(\xi, z_{1,\ell}, \bar{z}_{1,\ell}) = \{(a_1,\dots,a_n) \in \mathscr{R}^n : \\ z_1 + \cdots + z_\ell + a_1 + \cdots + a_{n-\ell} = \hat{x}_n(\xi), \\ a_{n-\ell+1,n} = \bar{z}_{1,\ell}\}.$$

The vectors of steps $(a_1,\dots,a_n)$ take $\eta_0 = (\omega, z_{1,\ell})$ to $\eta_n = (T_{\hat{x}_n(\xi)}\omega, \bar{z}_{1,\ell})$ via $\eta_i = S^+_{a_i}\eta_{i-1}$, $1 \le i \le n$.

LEMMA 2.9. *Let $F \in \mathcal{K}_\ell$. Then, for each $\xi \in \mathbb{Q}^d \cap \mathcal{U}$ and $z_{1,\ell}, \bar{z}_{1,\ell} \in \mathscr{R}^\ell$,*

$$\lim_{n\to\infty} \max_{(a_1,\dots,a_n)\in\mathcal{A}_n(\xi,z_{1,\ell},\bar{z}_{1,\ell})} \left| \frac{1}{n}\sum_{i=0}^{n-1} F(\eta_i, a_{i+1}) \right| = 0$$

in $L^1(\mathbb{P})$ and for $\mathbb{P}$-a.e. ω.

Remark 2.10. Due to the closed loop property (iii) in Definition 2.2, the sum above is independent of $(a_1,\dots,a_n) \in \mathcal{A}_n(\xi, z_{1,\ell}, \bar{z}_{1,\ell})$. In other words, there actually is no maximum. Also, Lemma 2.9 holds regardless of the choices made in the definition of $\hat{x}_\cdot(\xi)$.

We postpone the proof of Lemma 2.9 to Appendix C.

LEMMA 2.11. *Let $g \in \mathcal{L}$. Then $\overline{\Lambda}_\ell(g,\omega) \le K_\ell(g)$ for $\mathbb{P}$-a.e. ω.*

PROOF. By Lemma 2.8 it is enough to show that

$$\varlimsup_{n\to\infty} n^{-1} \log E_0\big[e^{nR_n^\ell(g)} \mathbb{1}\{X_n = \hat{x}_n(\xi)\}\big] \le K_{\ell,F}(g) \quad \mathbb{P}\text{-a.s.}$$

for fixed $\xi \in \mathbb{Q}^d \cap \mathcal{U}$ and $F \in \mathcal{K}_\ell$. Abbreviate $\eta_k = (T_{X_k}\omega, Z_{k+1,k+\ell})$. Fix $\varepsilon > 0$. Lemma 2.9 implies that for $\mathbb{P}$-a.e. ω there exists a finite $c_\varepsilon(\omega)$ such that for all n, on the event $\{X_n = \hat{x}_n(\xi)\}$,

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

Therefore, for $\mathbb{P}$-a.e. ω

$$
\begin{aligned}
&-n^{-1}c_\varepsilon - \varepsilon + n^{-1}\log E_0\big[e^{nR_n^\ell(g)}\mathbb{1}\{X_n = \hat{x}_n(\xi)\}\big]\\
&\quad\le n^{-1}\log E_0\Big[\exp\Big\{\sum_{k=0}^{n-1}(g(\eta_k) + F(\eta_k, Z_{k+\ell+1}))\Big\}\mathbb{1}\{X_n = \hat{x}_n(\xi)\}\Big]\\
&\quad\le n^{-1}\log E_0\Big[\exp\Big\{\sum_{k=0}^{n-1}(g(\eta_k) + F(\eta_k, Z_{k+\ell+1}))\Big\}\Big]\\
&\quad= n^{-1}\log E_0\Big[\exp\Big\{\sum_{k=0}^{n-2}(g(\eta_k) + F(\eta_k, Z_{k+\ell+1}))\Big\}\\
&\qquad\qquad\times E_0\big[e^{g(\eta_0)+F(\eta_0,Z_{\ell+1})} \mid \eta_{n-1}\big]\Big]\\
&\quad\le n^{-1}K_{\ell,F}(g) + n^{-1}\log E_0\Big[\exp\Big\{\sum_{k=0}^{n-2}(g(\eta_k) + F(\eta_k, Z_{k+\ell+1}))\Big\}\Big]\\
&\quad\le \cdots \le K_{\ell,F}(g).
\end{aligned}
$$

The claim follows by taking $n \nearrow \infty$ and then $\varepsilon \searrow 0$. □

We have shown (2.6.i) and next in line is (2.6.ii). The following lemma is the most laborious step in the paper.

LEMMA 2.12. *In addition to ergodicity, assume now that $\mathfrak{S}$ is countably generated. Assume $g(\cdot, z_{1,\ell}) \in L^1(\mathbb{P})$ is bounded above. Then*

$$K_\ell(g) \le H^\#_{\ell,\mathbb{P}}(g) = \sup_{\mu\in\mathcal{M}_1(\mathbf{\Omega}_\ell)}\{E^\mu[g] - H_{\ell,\mathbb{P}}(\mu)\}.$$

PROOF. We can assume $H^\#_{\ell,\mathbb{P}}(g) < \infty$. The first technical issue is to find some compactness to control the supremum on the right. Assume Ω compact would not be helpful because the problem is the absolute continuity condition in the definition of $H_{\ell,\mathbb{P}}(\mu)$.

Fix a sequence of increasing finite algebras $\mathfrak{S}_k$ on Ω that satisfy $T_{\pm z}\mathfrak{S}_{k-1} \subset \mathfrak{S}_k$ for all $k \in \mathbb{N}$ and $z \in \mathscr{R}$, and whose union generates $\mathfrak{S}$. Let $\mathcal{M}_1^k = \mathcal{M}_1^k(\mathbf{\Omega}_\ell)$ be the set of probability measures μ on $\mathbf{\Omega}_\ell$ for which there exist $\mathfrak{S}_k$-measurable Radon-Nikodym derivatives $\phi_{z_{1,\ell}}$ on Ω (with respect to $\mathbb{P}$) such that for bounded measurable G

$$\int_{\mathbf{\Omega}_\ell} G\,d\mu = \sum_{z_{1,\ell}\in\mathscr{R}^\ell}\int_\Omega \phi_{z_{1,\ell}}(\omega)G(\omega, z_{1,\ell})\mathbb{P}(d\omega).$$

Such μ satisfy $\mu_0 \ll \mathbb{P}$ and so

$$H^{\#}_{\ell,\mathbb{P}}(g) = \sup_{\mu:\mu_0\ll\mathbb{P}} \{E^{\mu}[g] - H_{\ell,\mathbb{P}}(\mu)\} \geq \sup_{\mu\in\mathcal{M}_1^k} \{E^{\mu}[g] - H_{\ell,\mathbb{P}}(\mu)\}.$$

Abbreviate $A = H^{\#}_{\ell,\mathbb{P}}(g)$. The proof is completed by verifying this statement:

$$\text{If } A \geq \sup_{\mu\in\mathcal{M}_1^k}\{E^{\mu}[g] - H_{\ell,\mathbb{P}}(\mu)\} \text{ for all } k \geq 1, \\ \text{then } A \geq K_\ell(g). \tag{2.10}$$

Let α denote a generic probability measure on $\mathbf{\Omega}_\ell^2$ with marginals α_1 and α_2, and let $b\mathbf{\Omega}_\ell$ denote the space of bounded measurable functions on $\mathbf{\Omega}_\ell$.

$$\begin{aligned} A &\geq \sup_{\mu\in\mathcal{M}_1^k,\, q:\mu q=\mu} \{E^{\mu}[g] - H(\mu\times q \mid \mu\times\widehat{p}_\ell)\} \\ &= \sup_{\alpha:\alpha_1\in\mathcal{M}_1^k,\alpha_1=\alpha_2} \{E^{\alpha_1}[g] - H(\alpha \mid \alpha_1\times\widehat{p}_\ell)\} \\ &= \sup_{\alpha:\alpha_1\in\mathcal{M}_1^k} \inf_{h\in b\mathbf{\Omega}_\ell} \{E^{\alpha_1}[g] + E^{\alpha_2}[h] - E^{\alpha_1}[h] - H(\alpha \mid \alpha_1\times\widehat{p}_\ell)\}. \end{aligned} \tag{2.11}$$

Let F denote a bounded measurable test function on $\mathbf{\Omega}_\ell^2$. $\mathcal{M}_2^k$ is the set of probability measures α on $\mathbf{\Omega}_\ell^2$ of the form

$$\int_{\mathbf{\Omega}_\ell^2} F\,d\alpha = \sum_{z\in\mathscr{R}} \int_{\mathbf{\Omega}_\ell} \alpha_1(d\eta)\, q(\eta, S_z^+\eta)\, F(\eta, S_z^+\eta)$$

where $\alpha_1 \in \mathcal{M}_1^k$ and kernel $q(\eta, S_z^+\eta) = q((\omega, z_{1,\ell}), (T_{z_1}\omega, z_{2,\ell}z))$ is $\mathfrak{S}_k$-measurable as a function of ω for each fixed $(z_{1,\ell}, z)$. A measure $\alpha \in \mathcal{M}_2^k$ is uniquely represented by a finite sequence $(\phi_{i,z_{1,\ell}}, q_{i,z_{1,\ell},z})$ via the identity

$$\int_{\mathbf{\Omega}_\ell^2} F\,d\alpha = \sum_{i,z_{1,\ell},z} \phi_{i,z_{1,\ell}} q_{i,z_{1,\ell},z} \int_{A_i} F((\omega, z_{1,\ell}), (T_{z_1}\omega, z_{2,\ell}z))\,\mathbb{P}(d\omega) \tag{2.12}$$

where $\{A_i\}$ is the finite set of atoms of $\mathfrak{S}_k$ such that $\mathbb{P}(A_i) > 0$, $\phi_{i,z_{1,\ell}}$ is the value of $\phi_{z_{1,\ell}}(\omega)$ for $\omega \in A_i$, and $q_{i,z_{1,\ell},z}$ is the value of $q(\eta, S_z^+\eta)$ for $\omega \in A_i$. Thus $\mathcal{M}_2^k$ is in bijective correspondence with a compact subset of a euclidean space, and (2.12) shows that via this identification the integral is continuous in α for any F that is suitably integrable under $\mathbb{P}$. Similarly, the entropy

$$H(\alpha \mid \alpha_1\times\widehat{p}_\ell) = \sum_{i,z_{1,\ell},z} \phi_{i,z_{1,\ell}} q_{i,z_{1,\ell},z} \mathbb{P}(A_i) \log(|\mathscr{R}| q_{i,z_{1,\ell},z})$$

is continuous and convex in α.

We turn our attention back to (2.11). Once we restrict α to the compact Hausdorff space $\mathcal{M}_2^k$ the expression in braces is upper-semicontinuous and concave in α

and convex in h. We can apply König's minimax theorem ([26] or [31]) and continue as follows:

$$\begin{aligned}
A &\geq \sup_{\alpha\in\mathcal{M}_2^k} \inf_{h\in b\mathbf{\Omega}_\ell} \{E^{\alpha_1}[g] + E^{\alpha_2}[h] - E^{\alpha_1}[h] - H(\alpha \mid \alpha_1 \times \widehat{p}_\ell)\} \\
&= \inf_{h\in b\mathbf{\Omega}_\ell} \sup_{\alpha\in\mathcal{M}_2^k} \{E^{\alpha_1}[g] + E^{\alpha_2}[h] - E^{\alpha_1}[h] - H(\alpha \mid \alpha_1 \times \widehat{p}_\ell)\} \\
&= \inf_{h\in b\mathbf{\Omega}_\ell} \sup_{\alpha\in\mathcal{M}_2^k} \sum_{z_{1,\ell}} \int_{\Omega} \mathbb{P}(d\omega)\phi_{z_{1,\ell}}(\omega) \\
&\quad \times \Big\{\sum_z q(\eta, S_z^+\eta)(g(\eta) - h(\eta) + h(S_z^+\eta)) - H(q(\eta,\cdot) \mid \widehat{p}_\ell(\eta,\cdot))\Big\} \\
&= \inf_{h\in b\mathbf{\Omega}_\ell} \sup_{\alpha\in\mathcal{M}_2^k} \sum_{z_{1,\ell}} \int_{\Omega} \mathbb{P}(d\omega)\phi_{z_{1,\ell}}(\omega) \\
&\quad \times \Big\{\sum_z q(\eta, S_z^+\eta)\mathbb{E}[g(\eta) - h(\eta) + h(S_z^+\eta) \mid \mathfrak{S}_k] - H(q(\eta,\cdot) \mid \widehat{p}_\ell(\eta,\cdot))\Big\} \\
&= \inf_{h\in b\mathbf{\Omega}_\ell} \sup_{\mu\in\mathcal{M}_1^k} \sum_{z_{1,\ell}} \int_{\Omega} \mathbb{P}(d\omega)\phi_{z_{1,\ell}}(\omega) \log \sum_z \frac{1}{|\mathscr{R}|} e^{\mathbb{E}[g(\eta)-h(\eta)+h(S_z^+\eta)\mid\mathfrak{S}_k]}.
\end{aligned}$$

Above we introduced the densities $\phi_{z_{1,\ell}}(\omega)$ and the kernel q that correspond to $\alpha \in \mathcal{M}_2^k$, used $\mathfrak{S}_k$-measurability to take conditional expectation, and then took the supremum over the kernels q with the first marginal $\alpha_1 = \mu$ fixed. This supremum is a finite case of the convex duality of relative entropy:

$$\sup_q \Big\{\sum_z q(z)v(z) - \sum_z q(z) \log \frac{q(z)}{p(z)}\Big\} = \log \sum_x p(x) e^{v(x)},$$

and the maximizing probability is $q(z) = (\sum_x p(x)e^{v(x)})^{-1} p(z)e^{v(z)}$. In our case $v(z) = \mathbb{E}[g(\eta) - h(\eta) + h(S_z^+\eta) \mid \mathfrak{S}_k]$, so the maximizing kernel is $\mathfrak{S}_k$-measurable in ω and thus admissible under the condition $\alpha \in \mathcal{M}_2^k$.

Performing the last supremum over $\mu \in \mathcal{M}_1^k$ gives

$$A \geq \inf_{h\in b\mathbf{\Omega}_\ell} \max_{z_{1,\ell}\in\mathscr{R}^\ell} \ \mathbb{P}\text{-}\operatorname{ess\,sup}_\omega \Big\{\log \sum_z \frac{1}{|\mathscr{R}|} e^{\mathbb{E}[g(\eta)-h(\eta)+h(S_z^+\eta)\mid\mathfrak{S}_k]}\Big\}.$$

Consequently, for $\varepsilon > 0$ and $k \geq 1$ there exists a bounded measurable function $h_{k,\varepsilon}$ on $\mathbf{\Omega}_\ell$ such that for all $z_{1,\ell} \in \mathscr{R}^\ell$ and $\mathbb{P}$-a.s.

$$A + \log|\mathscr{R}| + \varepsilon \geq \log \sum_z e^{\mathbb{E}[g(\eta)-h_{k,\varepsilon}(\eta)+h_{k,\varepsilon}(S_z^+\eta)\mid\mathfrak{S}_k]}. \tag{2.13}$$

For integers $0 \leq i \leq k$ define

$$F_{k,\varepsilon}^{(i)}(\eta, z) = \mathbb{E}[h_{k,\varepsilon}(S_z^+\eta) - h_{k,\varepsilon}(\eta) \mid \mathfrak{S}_{k-i}]. \tag{2.14}$$

We next extract a limit point in $\mathcal{K}_\ell$. The proof of the following lemma is given in Appendix C.

LEMMA 2.13. *Assume $\mathfrak{S}$ is countably generated and $A < \infty$. Construct $F^{(i)}_{k,\varepsilon}$ as in* (2.14). *Fix $\ell \ge 0$ and let $g(\omega, z_{1,\ell}) \in L^1(\mathbb{P})$ for all $z_{1,\ell} \in \mathscr{R}^\ell$. Fix $\varepsilon > 0$. Then, as $k \to \infty$, along a subsequence that works simultaneously for all $z_{1,\ell} \in \mathscr{R}^\ell$, $z \in \mathscr{R}$, and $i \ge 0$, one can write*

$$F^{(i)}_{k,\varepsilon} = \widehat{F}^{(i)}_{k,\varepsilon} - R^{(i)}_{k,\varepsilon}$$

with $\widehat{F}^{(i)}_{k,\varepsilon}(\eta, z)$ converging in weak $L^1(\mathbb{P})$ to a limit $\widehat{F}^{(i)}_\varepsilon$ and the error terms $\omega \mapsto R^{(i)}_{k,\varepsilon}(\omega, z_{1,\ell}, z) \ge 0$ $\mathfrak{S}_{k-i}$-measurable and converging to 0 $\mathbb{P}$-a.s. Furthermore, as $i \to \infty$, $\widehat{F}^{(i)}_\varepsilon$ converges strongly in $L^1(\mathbb{P})$ to a limit $\widehat{F}_\varepsilon$,

$$c(z) = \mathbb{E}[\widehat{F}_\varepsilon(\omega, (z, \dots, z), z)] \ge 0$$

for all $z \in \mathscr{R}$, and $F_\varepsilon(\eta, z) = \widehat{F}_\varepsilon(\eta, z) - c(z_1)$ belongs to class $\mathcal{K}_\ell$.

Fix $i \ge 0$ for the moment. As a uniformly integrable martingale, $M_k(\eta) = \mathbb{E}[g(\eta) \mid \mathfrak{S}_{k-i}]$ converges as $k \to \infty$ to $g(\eta)$, both a.s. and in $L^1(\mathbb{P})$ for all $z_{1,\ell}$ and z.

Fix $z_{1,\ell}$ and z. The weak-$L^1(\mathbb{P})$ closure of the convex hull of $\{M_j + \widehat{F}^{(i)}_{j,\varepsilon} : j \ge k\}$ is equal to its strong closure [37, theorem 3.12] . Since $g(\eta) + \widehat{F}^{(i)}_\varepsilon(\eta, z)$ is in this closure, there exist finite convex combinations

$$\widehat{G}^{(i)}_{k,\varepsilon} = \sum_{j \ge k} \alpha_{j,k}\big(M_j + \widehat{F}^{(i)}_{j,\varepsilon}\big)$$

such that

$$\mathbb{E}\big|g(\eta) + \widehat{F}^{(i)}_\varepsilon(\eta, z) - \widehat{G}^{(i)}_{k,\varepsilon}(\eta, z)\big| \le \frac{1}{k}.$$

Along a subsequence (that we again index by k) $\widehat{G}^{(i)}_{k,\varepsilon}(\eta, z)$ converges $\mathbb{P}$-a.s. to $g(\eta) + \widehat{F}^{(i)}_\varepsilon(\eta, z)$. Consequently, also

$$G^{(i)}_{k,\varepsilon} = \sum_{j \ge k} \alpha_{j,k}\big(M_j + F^{(i)}_{j,\varepsilon}\big) \underset{k\to\infty}{\longrightarrow} g + \widehat{F}^{(i)}_\varepsilon \quad \mathbb{P}\text{-a.s.}$$

Along a further subsequence this holds simultaneously for all $z_{1,\ell}$ and z.

By (2.13) and Jensen's inequality, we have for all $z_{1,\ell} \in \mathscr{R}^\ell$ and $\mathbb{P}$-a.s.

$$\begin{aligned} e^{A+\log|\mathscr{R}|+\varepsilon} &\ge \sum_{z\in\mathscr{R}} \mathbb{E}\big[e^{\mathbb{E}[g(\eta)-h_{k,\varepsilon}(\eta)+h_{k,\varepsilon}(S^+_z\eta)|\mathfrak{S}_k]} \bigm| \mathfrak{S}_{k-i}\big] \\ &\ge \sum_{z\in\mathscr{R}} e^{M_k(\eta,z)+F^{(i)}_{k,\varepsilon}(\eta,z)}. \end{aligned}$$

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

$$e^{A+\log|\mathscr{R}|+\varepsilon} \geq \sum_{z\in\mathscr{R}} e^{G^{(i)}_{k,\varepsilon}(\eta,z)}.$$

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

$$A + \varepsilon \geq g(\eta) + \log \sum_{z\in\mathscr{R}} \frac{1}{|\mathscr{R}|} e^{\widehat{F}^{(i)}_\varepsilon(\eta,z)}.$$

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

$$A + \varepsilon \geq g(\eta) + \log \sum_{z\in\mathscr{R}} \frac{1}{|\mathscr{R}|} e^{\widehat{F}_\varepsilon(\eta,z)}.$$

Since $c(z_1) \geq 0$ the above inequality still holds if $\widehat{F}_\varepsilon$ is replaced with F_ε. Thus

$$A + \varepsilon \geq \inf_{F\in\mathcal{K}_\ell} \max_{z_{1,\ell}\in\mathscr{R}^\ell} \mathbb{P}\text{-}\operatorname*{ess\,sup}_\eta \Big\{ g(\eta) + \log \sum_{z\in\mathscr{R}} \frac{1}{|\mathscr{R}|} e^{F(\eta,z)} \Big\}.$$

Taking $\varepsilon \to 0$ gives $A \geq K_\ell(g)$. (2.10) is verified and thereby the proof of Lemma 2.12 is complete. □

Next for technical purposes is a Fatou-type lemma for K_ℓ.

LEMMA 2.14. *Let* $g_k(\cdot, z_{1,\ell}) \underset{k\to\infty}{\longrightarrow} g(\cdot, z_{1,\ell})$ *in* $L^1(\mathbb{P})$ *for each* $z_{1,\ell} \in \mathscr{R}^\ell$. *Then*

$$K_\ell(g) \leq \varliminf_{k\to\infty} K_\ell(g_k).$$

PROOF. We can assume $\varliminf_{k\to\infty} K_\ell(g_k) = A < \infty$. Fix $\varepsilon > 0$. There exists a subsequence, denoted again by g_k, such that $K_\ell(g_k) < A + \varepsilon$ for all k. Pick $F_k \in \mathcal{K}_\ell$ such that

$$g_k(\eta) + \log \sum_{z\in\mathscr{R}} \frac{1}{|\mathscr{R}|} e^{F_k(\eta,z)} < A + \varepsilon \tag{2.15}$$

for all k, $z_{1,\ell} \in \mathscr{R}^\ell$, and $\mathbb{P}$-a.e. ω. Out of this we can produce an $F \in \mathcal{K}_\ell$ such that

$$g(\eta) + \log \sum_{z\in\mathscr{R}} \frac{1}{|\mathscr{R}|} e^{F(\eta,z)} \leq A + \varepsilon. \tag{2.16}$$

This implies $K_\ell(g) \leq A + \varepsilon$ and taking $\varepsilon \to 0$ finishes the proof.

The construction of F is a simplified version of the argument to realize a limit point in $\mathcal{K}_\ell$ in the proof of Lemma 2.12. We sketch the steps. The reader who aims to master the proof may find it useful to fill in the details.

For each k, $z_{1,\ell}$, and z,

$$F_k(\eta, z) \leq A + \varepsilon - g_k(\eta) + \log|\mathscr{R}|.$$

Thus F_k^+ is uniformly integrable. Controlling F_k^- is indirect. Set $\eta_0 = (\omega, z_{1,\ell})$, $z_0 = z$, $a_i = z_{i-1}$, and $\eta_i = S_{a_i}^+ \eta_{i-1}$ for $i = 1, 2, \ldots, \ell + 1$. By the mean-zero property of F_k (part (ii) in Definition 2.2),

$$\mathbb{E}[F_k^-(\eta, z)] \le \sum_{i=0}^{\ell} \mathbb{E}[F_k^-(\eta_i, a_{i+1})] = \sum_{i=0}^{\ell} \mathbb{E}[F_k^+(\eta_i, a_{i+1})],$$

and so $\mathbb{E}[F_k^-]$ is bounded uniformly in k. Apply Lemma C.4 to write $F_k^- = \widetilde{F}_k + R_k$ such that along a subsequence $\widetilde{F}_k$ is uniformly integrable and $R_k \ge 0$ converges to 0 in $\mathbb{P}$-probability, for each $z_{1,\ell}$. Along a further subsequence $\widehat{F}_k \equiv F_k^+ - \widetilde{F}_k$ converges weakly in $L^1(\mathbb{P})$ to a limit $\widehat{F}$ and the limits $R_k \to 0$ and $g_k \to g$ hold almost surely.

In (2.15) write $F_k = \widehat{F}_k - R_k$. As done in the proof of Lemma 2.12, take almost surely convergent convex combinations of $\widehat{F}_k$, R_k, and g_k and substitute these into (2.15). Taking the limit now yields (2.16) but with $\widehat{F}$ in place of F.

Almost sure convergence of convex combinations ensures that $\widehat{F}$ satisfies the closed-loop property. But it may fail the mean-zero property. To remedy this, let $c(z) = \mathbb{E}[\widehat{F}(\omega, (z, \ldots, z), z)]$. By the weak convergence $c(z)$ is a limit of $\mathbb{E}[\widehat{F}_k(\omega, (z, \ldots, z), z)]$, which is nonnegative due to $R_k \ge 0$ and the mean-zero property for F_k. Since $c(z) \ge 0$, (2.16) holds with $F(\eta, z) = \widehat{F}(\eta, z) - c(z_1)$. That F satisfies both the mean-zero and the closed-loop property is verified with the argument given between equations (C.10) and (C.11) in Appendix C. The point is that the closed-loop property of $\widehat{F}$ allows us to define the path integral $\widehat{f}$ that is used in that argument. This verifies that $F \in \mathcal{K}_\ell$ and completes the proof. □

Next is a large deviation lower bound lemma that gives us (2.6.iii) and serves again to prove Theorem 3.1 below.

LEMMA 2.15. *Let $g(\cdot, z_{1,\ell}) \in L^1(\mathbb{P})$ be bounded above. Then for $\mathbb{P}$-a.e. ω*

$$\underline{\lim}_{n\to\infty} n^{-1} \log E_0\big[e^{nR_n^\ell(g)}\big] \ge \sup_\mu \{E^\mu[g] - H_{\ell,\mathbb{P}}(\mu)\}. \tag{2.17}$$

Assume additionally that Ω is a separable metric space. Then for $\mathbb{P}$-a.e. ω this lower bound holds for all open $O \subset \mathcal{M}_1(\mathbf{\Omega}_\ell)$:

$$\underline{\lim}_{n\to\infty} n^{-1} \log E_0\big[e^{nR_n^\ell(g)} \mathbb{1}\{R_n^\ell \in O\}\big] \ge -\inf_{\mu\in O} \{H_{\ell,\mathbb{P}}(\mu) - E^\mu[g]\}. \tag{2.18}$$

PROOF. This proof proceeds along the familiar lines of Markov chain lower-bound arguments, and we refer to [32, sec. 4] for further details.

Switch to the $\mathbf{\Omega}_\ell$-valued Markov chain $\eta_k = (T_{X_k}\omega, Z_{k+1,k+\ell})$ with transition kernel $\widehat{p}_\ell$ defined in (2.2). Then R_n^ℓ is the position level empirical measure $L_n = n^{-1} \sum_{k=0}^{n-1} \delta_{\eta_k}$. Denote by P_η (with expectation E_η) the distribution of the Markov chain $(\eta_k)_{k\ge 0}$ with initial state η. Starting at $\eta = (\omega, z_{1,\ell})$ is the same as conditioning our original process on $Z_{1,\ell}$:

$$E_0[G((T_{X_k}\omega,\ Z_{k+1,k+\ell})_{0\le k\le n})\mathbb{1}\{Z_{1,\ell}=z_{1,\ell}\}] = \frac{1}{|\mathscr{R}|^{\ell}}\, E_\eta[G(\eta_0,\eta_1,\dots,\eta_n)].$$

Consequently, for any $z_{1,\ell}\in\mathscr{R}^\ell$ and $\mathbb{P}$-almost every ω, with $\eta=(\omega,z_{1,\ell})$,

$$\varliminf_{n\to\infty} n^{-1}\log E_0\big[e^{nR_n^\ell(g)}\mathbb{1}\{R_n^\ell\in O\}\big] \ge \varliminf_{n\to\infty} n^{-1}\log E_\eta\big[e^{nL_n(g)}\mathbb{1}\{L_n\in O\}\big].$$

Next we reduce the right-hand sides of (2.17) and (2.18) to nice measures. A convexity argument shows that the supremum/infimum is not altered by restricting it to measures μ with these properties: $\mu_0\ll\mathbb{P}$ and there exists a kernel $q\in\mathcal{Q}(\mathbf{\Omega}_\ell)$ such that $\mu q=\mu$, $H(\mu\times q\mid\mu\times\widehat{p}_\ell)<\infty$, $q(\eta,\cdot)$ is supported on shifts $S_z^+\eta$, and $q(\eta,S_z^+\eta)>0$ for all $z\in\mathscr{R}$ and μ-a.e. η. We omit this argument. It can be patterned after the lower-bound proof of [32, theorem 3.1, p. 224]. This step needs the integrability of $g(\omega,z_{1,\ell})$ under $\mathbb{P}$.

These properties of μ imply the equivalence $\mu_0\sim\mathbb{P}$ and the ergodicity of the Markov chain Q_η with initial state η and transition kernel q [32, lemma 4.1].

Next follows a standard change-of-measure argument. Let $\mathscr{F}_n$ be the σ-algebra generated by $(\eta_0,\eta_1,\dots,\eta_n)$. Then

$$\begin{aligned}
&n^{-1}\log E_\eta[e^{nL_n(g)}\mathbb{1}\{L_n\in O\}]\\
&\quad\ge n^{-1}\log\frac{E^{Q_\eta}\Big[\Big(\frac{dQ_{\eta|\mathscr{F}_{n-1}}}{dP_{\eta|\mathscr{F}_{n-1}}}\Big)^{-1}e^{nL_n(g)}\mathbb{1}\{L_n\in O\}\Big]}{Q_\eta\{L_n\in O\}}+n^{-1}\log Q_\eta\{L_n\in O\}\\
&\quad\ge\frac{-n^{-1}E^{Q_\eta}\Big[\log\Big(\frac{dQ_{\eta|\mathscr{F}_{n-1}}}{dP_{\eta|\mathscr{F}_{n-1}}}\Big)\Big]}{Q_\eta\{L_n\in O\}}+\frac{E^{Q_\eta}[L_n(g)]}{Q_\eta\{L_n\in O\}}+n^{-1}\log Q_\eta\{L_n\in O\}\\
&\qquad+\frac{n^{-1}E^{Q_\eta}\Big[\log\Big(\frac{dQ_{\eta|\mathscr{F}_{n-1}}}{dP_{\eta|\mathscr{F}_{n-1}}}\Big)\mathbb{1}\{L_n\notin O\}\Big]}{Q_\eta\{L_n\in O\}}-\frac{E^{Q_\eta}[L_n(g)\mathbb{1}\{L_n\notin O\}]}{Q_\eta\{L_n\in O\}}\\
&\quad=\frac{-n^{-1}H(Q_{\eta|\mathscr{F}_{n-1}}\mid P_{\eta|\mathscr{F}_{n-1}})}{Q_\eta\{L_n\in O\}}+\frac{E^{Q_\eta}[L_n(g)]}{Q_\eta\{L_n\in O\}}+n^{-1}\log Q_\eta\{L_n\in O\}\\
&\qquad+\frac{n^{-1}E_\eta\Big[\frac{dQ_{\eta|\mathscr{F}_{n-1}}}{dP_{\eta|\mathscr{F}_{n-1}}}\log\Big(\frac{dQ_{\eta|\mathscr{F}_{n-1}}}{dP_{\eta|\mathscr{F}_{n-1}}}\Big)\mathbb{1}\{L_n\notin O\}\Big]}{Q_\eta\{L_n\in O\}}-\frac{E^{Q_\eta}[L_n(g)\mathbb{1}\{L_n\notin O\}]}{Q_\eta\{L_n\in O\}}\\
&\quad\ge-\frac{E^{Q_\eta}\big[n^{-1}\sum_{k=0}^{n-1}F(\eta_k)\big]}{Q_\eta\{L_n\in O\}}+\frac{E^{Q_\eta}[L_n(g)]}{Q_\eta\{L_n\in O\}}+n^{-1}\log Q_\eta\{L_n\in O\}\\
&\qquad-\frac{n^{-1}e^{-1}}{Q_\eta\{L_n\in O\}}-\frac{(\sup g)Q_\eta\{L_n\notin O\}}{Q_\eta\{L_n\in O\}},
\end{aligned}$$

where we used $x \log x \geq -e^{-1}$ and

$$F(\eta) = \sum_{z \in \mathscr{R}} q(\eta, S_z^+ \eta) \log \frac{q(\eta, S_z^+ \eta)}{\widehat{p}_\ell(\eta, S_z^+ \eta)}.$$

Since $F \geq 0$ by Jensen's inequality and g is bounded above, ergodicity gives the limits for μ_0-a.e. ω:

$$\varliminf_{n \to \infty} n^{-1} \log E_0\big[e^{nR_n^\ell(g)} \mathbb{1}\big\{R_n^\ell \in O\big\}\big] \geq E^\mu[g] - H(\mu \times q \mid \mu \times \widehat{p}_\ell).$$

(For the details of $Q_\eta\{L_n \in O\} \to 1$, see the proof of [32, lemma 4.2].) By $\mu_0 \sim \mathbb{P}$ this also holds $\mathbb{P}$-a.s. □

We are ready for the proof of the theorem.

PROOF OF THEOREM 2.3. Assume first that $g \in \mathcal{L}$ is bounded above. Then Lemmas 2.11, 2.12, and 2.15 give these $\mathbb{P}$-a.s. inequalities:

$$\overline{\Lambda}_\ell(g, \omega) \leq K_\ell(g) \leq H^\#_{\ell,\mathbb{P}}(g) \leq \underline{\Lambda}_\ell(g, \omega).$$

Existence of the limit $\Lambda_\ell(g)$ and $\Lambda_\ell(g) = K_\ell(g) = H^\#_{\ell,\mathbb{P}}(g)$ follows.

Next, consider $g \in \mathcal{L}$. Lemma 2.14 implies that $K_\ell(g) = \sup_c K_\ell(\min(g, c))$. Existence of the limit $\Lambda_\ell(\min(g, c))$ combined with Lemma 2.11 implies

$$\begin{aligned} K_\ell(g) &= \sup_c K_\ell(\min(g, c)) = \sup_c \Lambda_\ell(\min(g, c)) = \sup_c \underline{\Lambda}_\ell(\min(g, c)) \\ &\leq \underline{\Lambda}_\ell(g) \leq \overline{\Lambda}_\ell(g) \leq K_\ell(g). \end{aligned}$$

Existence of the limit and the equality $\Lambda_\ell(g) = K_\ell(g)$ follow again. For the other variational formula write

$$\begin{aligned} K_\ell(g) &= \sup_c K_\ell(\min(g, c)) \\ &= \sup_c \sup_{\mu \in \mathcal{M}_1(\mathbf{\Omega}_\ell)} \{E^\mu[\min(g, c)] - H_{\ell,\mathbb{P}}(\mu)\} = H^\#_{\ell,\mathbb{P}}(g). \end{aligned}$$ □

3 Large Deviations under Quenched Polymer Measures

As before, we continue to assume that $\mathscr{R}$ is finite and $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z : z \in \mathcal{G}\})$ is a measurable ergodic system where $\mathcal{G}$ is the additive subgroup of $\mathbb{Z}^d$ generated by $\mathscr{R}$. Now assume additionally that Ω is a separable metric space and $\mathfrak{S}$ is its Borel σ-algebra.

Since our limiting logarithmic moment-generating functions $\Lambda_\ell(g)$ are defined only $\mathbb{P}$-a.s., we need a separable function space that generates the weak topology of probability measures. Give $\mathbf{\Omega}_\ell$ a totally bounded metric and let $\mathfrak{U}_b(\mathbf{\Omega}_\ell)$ be the space of uniformly continuous functions under this metric. These functions are bounded. The space $\mathfrak{U}_b(\mathbf{\Omega}_\ell)$ is separable under the supremum norm and generates the same topology on $\mathcal{M}_1(\mathbf{\Omega}_\ell)$ as does the space of bounded continuous functions.

Given a real-valued function V on $\mathbf{\Omega}_\ell$, define the quenched polymer measures

$$Q_{n,0}^{V,\omega}(A) = \frac{1}{Z_{n,0}^{V,\omega}} E_0\big[e^{-\sum_{k=0}^{n-1} V(T_{X_k}\omega, Z_{k+1,k+\ell})} \mathbb{1}_A(\omega, X_{0,\infty})\big],$$

where A is an event on environments and paths and

$$Z_{n,0}^{V,\omega} = E_0\big[e^{-\sum_{k=0}^{n-1} V(T_{X_k}\omega, Z_{k+1,k+\ell})}\big].$$

Theorem 2.3 gives the a.s. limit $\Lambda_\ell(-V) = \lim n^{-1} \log Z_{n,0}^{V,\omega}$. Next we prove an LDP for the quenched distributions $Q_{n,0}^{V,\omega}\{R_n^\ell \in \cdot\}$ of the empirical measure

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

THEOREM 3.1. *Fix $\ell \geq 0$. Let V be a measurable function on $\mathbf{\Omega}_\ell$, $V \in \mathcal{L}$, and $\Lambda_\ell(-V) < \infty$. Then for $\mathbb{P}$-a.e. ω the weak large deviation principle holds for the sequence of probability distributions $Q_{n,0}^{V,\omega}\{R_n^\ell \in \cdot\}$ on $\mathcal{M}_1(\mathbf{\Omega}_\ell)$ with convex rate function*

$$I_{q,2,\ell}^V(\mu) = \sup_{g \in \mathfrak{U}_b(\mathbf{\Omega}_\ell)} \{E^\mu[g] - \Lambda_\ell(g - V)\} + \Lambda_\ell(-V). \tag{3.1}$$

Rate $I_{q,2,\ell}^V$ is also equal to the lower-semicontinuous regularization of

$$H_{\ell,\mathbb{P}}^V(\mu) = \inf_{c<0} \{H_{\ell,\mathbb{P}}(\mu) + E^\mu[\max(V, c)] + \Lambda_\ell(-V)\}. \tag{3.2}$$

PROOF OF THEOREM 3.1. We show an upper bound for compact sets A,

$$\varlimsup_{n\to\infty} n^{-1} \log Q_{n,0}^{V,\omega}\{R_n^\ell \in A\} \leq -\inf_{\mu \in A} I_{q,2,\ell}^V(\mu), \tag{3.3}$$

a lower bound for open sets G,

$$\varliminf_{n\to\infty} n^{-1} \log Q_{n,0}^{V,\omega}\{R_n^\ell \in G\} \geq -\inf_{\mu \in G} H_{\ell,\mathbb{P}}^V(\mu), \tag{3.4}$$

and then match the rates.

By Theorem 2.3 and the separability of $\mathfrak{U}_b(\mathbf{\Omega}_\ell)$, we have $\mathbb{P}$-a.s. these finite limits for all $g \in \mathfrak{U}_b(\mathbf{\Omega}_\ell)$:

$$\lim_{n\to\infty} n^{-1} \log E^{Q_{n,0}^{V,\omega}}\big[e^{nR_n^\ell(g)}\big] = \Lambda_\ell(g - V) - \Lambda_\ell(-V).$$

Inequality (3.3) follows by a convex duality argument (see [11, theorem 4.5.3] or [31, theorem 5.24]).

Lower bound (3.4) follows from Lemma 2.15 and a truncation: for $-\infty < c < 0$

$$n^{-1} \log Q_{n,0}^{V,\omega}\{R_n^\ell \in O\} \geq n^{-1} \log E_0\big[e^{-nR_n^\ell(\max(V,c))} \mathbb{1}\{R_n^\ell \in O\}\big] - n^{-1} \log E_0\big[e^{-nR_n^\ell(V)}\big].$$

(3.4) continues to hold if $H^V_{\ell,\mathbb{P}}$ is replaced with its lower-semicontinuous regularization $H^{V,**}_{\ell,\mathbb{P}}(\mu) = \sup_B \inf_{\nu\in B} H^V_{\ell,\mathbb{P}}(\nu)$, where the supremum is over open neighborhoods B of μ.

Theorem 2.3 implies that for $g \in \mathfrak{U}_b(\mathbf{\Omega}_\ell)$

$$\begin{aligned}
&\Lambda_\ell(g - V) - \Lambda_\ell(-V)\\
&\quad= \sup_{\mu\in\mathcal{M}_1(\mathbf{\Omega}_\ell),\,c>0} \{E^\mu[\min(g - V, c)] - H_{\ell,\mathbb{P}}(\mu) - \Lambda_\ell(-V)\}\\
&\quad= \sup_{\mu\in\mathcal{M}_1(\mathbf{\Omega}_\ell),\,c<0} \{E^\mu[g - \max(V, c)] - H_{\ell,\mathbb{P}}(\mu) - \Lambda_\ell(-V)\}\\
&\quad= \sup_{\mu\in\mathcal{M}_1(\mathbf{\Omega}_\ell)} \{E^\mu[g] - H^V_{\ell,\mathbb{P}}(\mu)\}.
\end{aligned}$$

Another convex duality gives $I^V_{q,2,\ell}(\mu) = H^{V,**}_{\ell,\mathbb{P}}(\mu)$ because the lower-semicontinuous regularization $H^{V,**}_{\ell,\mathbb{P}}$ is also equal to the double convex dual of $H^V_{\ell,\mathbb{P}}$. $\square$

Next we record the LDP for the quenched distributions of the empirical process $R^\infty_n = n^{-1}\sum_{k=0}^{n-1} \delta_{T_{X_k}\omega,\, Z_{k+1,\infty}}$.

THEOREM 3.2. *Let V be a measurable function on some $\mathbf{\Omega}_{\ell_0}$ with $V \in \mathcal{L}$ and $\Lambda_{\ell_0}(-V) < \infty$. Then for $\mathbb{P}$-a.e. ω the weak large deviation principle holds for the sequence of probability distributions $Q^{V,\omega}_{n,0}\{R^\infty_n \in \cdot\}$ on $\mathcal{M}_1(\Omega \times \mathscr{R}^{\mathbb{N}})$ with convex rate function $I^V_{q,3}(\mu) = \sup_{\ell\ge\ell_0} I^V_{q,2,\ell}(\mu|_{\mathbf{\Omega}_\ell})$.*

PROOF. This comes from a projective limit. Formula (3.1) shows that $I^V_{q,2,\ell}(\mu \circ \gamma^{-1}_{\ell+1,\ell}) \le I^V_{q,2,\ell+1}(\mu)$ for $\mu \in \mathcal{M}_1(\mathbf{\Omega}_{\ell+1})$ where $\gamma_{\ell+1,\ell} : \mathbf{\Omega}_{\ell+1} \to \mathbf{\Omega}_\ell$ is the natural projection. Since weak topology of $\mathcal{M}_1(\Omega \times \mathscr{R}^{\mathbb{N}})$ can be generated by uniformly continuous functions, a base for the topology can be created from inverse images of open sets from the spaces $\mathcal{M}_1(\mathbf{\Omega}_\ell)$. Apply Theorem B.1. $\square$

In one of the most basic situations, namely for strictly directed walks in i.i.d. environments, we can upgrade the weak LDPs into full LDPs. This means that the upper bound is valid for all closed sets. "Strictly directed" means that there is a vector $\hat{u} \in \mathbb{R}^d$ such that $z \cdot \hat{u} > 0$ for all $z \in \mathscr{R}$. Equivalently, 0 does not lie in the convex hull of $\mathscr{R}$.

Here is the setting. Let Γ be a Polish space. Set $\Omega = \Gamma^{\mathbb{Z}^d}$ with generic elements $\omega = (\omega_x)_{x\in\mathbb{Z}^d}$ and shift maps $(T_x\omega)_y = \omega_{x+y}$. Assume that the coordinates $\{\omega_x\}$ are i.i.d. under $\mathbb{P}$.

THEOREM 3.3. *As described above, let $\mathbb{P}$ be an i.i.d. product measure on a Polish product space Ω. Assume that 0 does not lie in the convex hull of $\mathscr{R}$. Let V be a measurable function on some $\mathbf{\Omega}_\ell$, $V \in \mathcal{L}$, and assume that $\Lambda_\ell(-\beta V) < \infty$ for some $\beta > 1$. Then for $\mathbb{P}$-a.e. ω the full LDP holds for the sequence of probability distributions $Q^{V,\omega}_{n,0}\{R^\infty_n \in \cdot\}$ on $\mathcal{M}_1(\Omega \times \mathscr{R}^{\mathbb{N}})$ with convex rate function $I^V_{q,3}$ described in Theorem 3.2.*

PROOF. $\Lambda_\ell(V) < \infty$ by Jensen's inequality. Due to Theorem 3.2 it suffices to show that the distributions $Q_{n,0}^{V,\omega}\{R_n^\infty \in \cdot\}$ are exponentially tight for $\mathbb{P}$-a.e. ω. Suppose we can show that

$$\text{distributions } P_0\{R_n^\infty \in \cdot\} \text{ are exponentially tight for } \mathbb{P}\text{-a.e. } \omega. \tag{3.5}$$

From the lower bound in (2.5) and the hypotheses on V we have constants $0 < c_0, c_1 < \infty$ such that, for $\mathbb{P}$-a.e. ω,

$$E_0[e^{-nR_n^\ell(V)}] \ge e^{-c_0 n} \quad \text{and} \quad E_0[e^{-nR_n^\ell(\beta V)}] \le e^{c_1 \beta n}$$

for large enough n. Fix ω so that these bounds and (3.5) hold. Given $c < \infty$, pick a compact $A \subset \mathcal{M}_1(\Omega \times \mathscr{R}^{\mathbb{N}})$ such that $P_0\{R_n^\infty \in A^c\} \le e^{-\beta(c_0+c_1+c)n/(\beta-1)}$ for large n. Then

$$\begin{aligned} Q_{n,0}^{V,\omega}\{R_n^\infty \in A^c\} &\le E_0[e^{-nR_n^\ell(V)}]^{-1} E_0[e^{-nR_n^\ell(\beta V)}]^{\beta^{-1}} P_0\{R_n^\infty \in A^c\}^{1-\beta^{-1}} \\ &\le e^{-cn}. \end{aligned}$$

Thus it suffices to check (3.5).

Next observe from

$$\mathbb{P}\{\omega : P_0(R_n^\infty \in A^c) \ge e^{-cn}\} \le e^{cn}\, \overline{P}(R_n^\infty \in A^c)$$

and the Borel-Cantelli lemma that we only need exponential tightness under the averaged measure $\overline{P} = \mathbb{P} \otimes P_0$. As the last reduction, note that by the compactness of $\mathscr{R}^{\mathbb{N}}$ it is enough to have the exponential tightness of the $\overline{P}$-distributions of $R_n^0 = n^{-1}\sum_{k=0}^{n-1} \delta_{T_{X_k}\omega}$.

The exponential tightness that is part of Sanov's theorem gives compact sets $\{U_{m,x} : m \in \mathbb{N}, x \in \mathbb{Z}^d\}$ in the state space Γ of the ω_x such that

$$\mathbb{P}\left\{n^{-1}\sum_{k=0}^{n-1} \mathbb{1}_{U_{m,x}^c}(\omega_{y_k}) > e^{-m-|x|}\right\} \le e^{-n(m+|x|)}.$$

Here $\{y_k\}$ are any distinct sites. Define

$$H_m = \{Q \in \mathcal{M}_1(\Omega) : \forall x \in \mathbb{Z}^d \;\; Q\{\omega : \omega_x \notin U_{m,x}\} \le e^{-(m+|x|)}\}$$

and compact sets

$$K_b = \bigcap_{m \ge \ell(b)} H_m$$

where $\ell = \ell(b)$ is chosen for $b \in \mathbb{N}$ so that

$$\sum_{m \ge \ell - b} \sum_x e^{-(m+|x|)} \le 1.$$

Now

$$\begin{aligned}\overline{P}(R_n^0 \in K_b^c) &\le \sum_{m\ge\ell(b)} \overline{P}(R_n^0 \in H_m^c)\\ &\le \sum_{m\ge\ell(b)} \sum_x \overline{P}\big(R_n^0\{\omega_x \notin U_{m,x}\} > e^{-m-|x|}\big)\\ &\le \sum_{m\ge\ell(b)} \sum_x \overline{P}\Big\{n^{-1}\sum_{k=0}^{n-1} \mathbb{1}_{U_{m,x}^c}(\omega_{x+X_k}) > e^{-m-|x|}\Big\} \le e^{-bn}.\end{aligned}$$

The crucial point used above was that under the assumption on $\mathscr{R}$ the points $\{X_n\}$ of the walk are distinct (Corollary A.2), and so the variables $\{\omega_{x+X_n}\}$ are i.i.d. under $\overline{P}$. This gives the exponential tightness of the $\overline{P}$-distributions of $R_n^0 = n^{-1}\sum_{k=0}^{n-1}\delta_{T_{X_k}\omega}$. □

Remark 3.4. For exponential tightness the theorem above is in some sense the best possible. Theorem 3.3 can fail if 0 lies in the convex hull of $\mathscr{R}$. Then a loop is possible (Corollary A.2). Suppose the distribution of ω_0 is not supported on any compact set. Then, given any compact set U in Γ, wait until the walk finds an environment $\omega_x \notin U$, and then forever after execute a loop at x.

4 Large Deviations for Random Walk in a Random Environment

This final section before the appendices is a remark about adapting the results of Section 3 to RWRE described in Example 1.2. Continue with the assumptions on $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z : z \in \mathcal{G}\})$ from Section 3. Fix any $\ell \ge 1$ and let $V(\omega, z_{1,\ell}) = -\log \pi_{0,z_1}(\omega)$ to put RWRE in the polymer framework. Then $\Lambda_1(-V) = -\log|\mathscr{R}|$. The necessary assumption is now

$$|\log \pi_{0,z}| \in \mathcal{L} \quad \text{for each } z \in \mathscr{R}. \tag{4.1}$$

The commonly used RWRE assumption of *uniform ellipticity*, namely the existence of $\kappa > 0$ such that $\mathbb{P}\{\pi_{0,z} \ge \kappa\} = 1$ for $z \in \mathscr{R}$, implies (4.1).

Under assumption (4.1) Theorems 3.1 and 3.2 are valid for RWRE and give quenched weak LDPs for the distributions $P_0^\omega\{R_n^\ell \in \cdot\}$ and $P_0^\omega\{R_n^\infty \in \cdot\}$. Note though that for $\ell \ge 2$, $Q_{n,0}^{V,\omega}\{R_n^\ell \in B\}$ is not exactly equal to $P_0^\omega\{R_n^\ell \in B\}$ because under $Q_{n,0}^{V,\omega}$ steps Z_k for $k > n$ are taken from kernel $\widehat{p}$. This difference vanishes in the limit due to $\log \pi_{0,z}(\omega) \in L^1(\mathbb{P})$. These LDPs take care of cases of RWRE not covered by [32], namely those walks for which 0 does not lie in the relative interior of the convex hull $\mathcal{U}$ of $\mathscr{R}$.

For RWRE the rate function $I_{q,2,\ell}^V$ in Theorem 3.1 can be expressed directly as the lower-semicontinuous regularization of an entropy. Indeed, let $\overline{V}(\omega, z_{1,\ell}) = -\log \pi_{0,z_\ell}(T_{x_{\ell-1}}\omega)$. The difference between using potential $\overline{V}$ and potential V is only in finitely many terms in the exponent. Thus $\Lambda_\ell(g - V) = \Lambda_\ell(g - \overline{V})$

for all $g \in \mathfrak{U}_b(\boldsymbol{\Omega}_\ell)$. Then (3.1) shows that $I^V_{q,2,\ell} = I^{\overline{V}}_{q,2,\ell}$. The latter rate is the lower-semicontinuous regularization of $H^{\overline{V}}_{\ell,\mathbb{P}}$ in (3.2), which itself equals $H_{\ell,\mathbb{P}}$ from (2.3) with $\widehat{p}_\ell$ replaced with the kernel $p^+(\eta, S^+_z\eta) = \pi_{0,z}(T_{x_\ell}\omega)$ of the Markov chain $(T_{X_k}\omega, Z_{k+1,k+\ell})$ under P^ω_0. By [32, lemma 6.1] the same is true of the level 3 rate $I^V_{q,3}$ under the additional assumption that Ω is a compact space. We refer to [32] for this and some other properties of $I^V_{q,3}$.

If Ω is compact, these weak LDPs are of course full LDPs, that is, the upper bound holds for all closed sets. For RWRE with finite $\mathscr{R}$ the natural canonical choice of Ω is compact: in the setting of Example 1.2, take $\Omega = \mathcal{P}^{\mathbb{Z}^d}$ with generic elements $\omega = (\omega_x)_{x\in\mathbb{Z}^d}$ and $p(\omega) = \omega_0$ the projection at the origin.

If Ω is compact we can project the LDP of Theorem 3.1 to the level of the walk to obtain the following statements: The limiting logarithmic moment-generating function

$$\lambda(t) = \lim_{n\to\infty} \frac{1}{n} \log E^\omega_0[e^{t\cdot X_n}], \quad t \in \mathbb{R}^d, \tag{4.2}$$

exists a.s. Its convex conjugate

$$\lambda^*(\zeta) = \sup_{t\in\mathbb{R}^d} \{\zeta \cdot t - \lambda(t)\}, \quad \zeta \in \mathbb{R}^d,$$

is the rate function for the LDP of the distributions $P^\omega_0\{n^{-1}X_n \in \cdot\}$ on $\mathbb{R}^d$. For walks without ellipticity, in particular for walks with $0 \notin \mathcal{U}$, even this quenched position-level LDP is new. It has been proved in the past only in a neighborhood of the limiting velocity [48].

In the following appendices we invoke the ergodic theorem a few times. By that we mean the multidimensional ergodic theorem; see, for example, [20, theorem 14.A8].

Appendix A Some Auxiliary Lemmas

In this appendix $\mathscr{R}$ is a finite subset of $\mathbb{Z}^d$, $\mathcal{G}$ the additive subgroup of $\mathbb{Z}^d$ generated by $\mathscr{R}$, and $\mathcal{U}$ the convex hull of $\mathscr{R}$ in $\mathbb{R}^d$.

LEMMA A.1. *Let $\xi \in \mathbb{Q}^d \cap \mathcal{U}$. Then there exist rational coefficients $\alpha_z \geq 0$ such that $\sum_{z\in\mathscr{R}} \alpha_z = 1$ and $\xi = \sum_{z\in\mathscr{R}} \alpha_z z$.*

PROOF. Suppose first that $\mathscr{R} = \{\hat{z}_0, \dots, \hat{z}_n\}$ for affinely independent points $\hat{z}_0, \dots, \hat{z}_n$. This means that the vectors $\hat{z}_1 - \hat{z}_0, \dots, \hat{z}_n - \hat{z}_0$ are linearly independent in $\mathbb{R}^d$, and then necessarily $n \leq d$. Augment this set to a basis $\{b_1 = \hat{z}_1 - \hat{z}_0, \dots, b_n = \hat{z}_n - \hat{z}_0, b_{n+1}, \dots, b_d\}$ of $\mathbb{R}^d$ where $b_{n+1}, \dots, b_d$ are also integer vectors (for example, by including a suitable set of $d - n$ standard basis vectors). Let A be the unique invertible linear transformation such that $Ab_i = e_i$ for $1 \leq i \leq d$. In the standard basis the matrix of A is the inverse of the matrix $B = [b_1, \dots, b_d]$; hence this matrix has rational entries.

Now let $\xi = \sum_{i=0}^{n} \alpha_i \hat{z}_i$ be a representation of ξ as a convex combination of $\hat{z}_0, \dots, \hat{z}_n$. Then $\xi - \hat{z}_0 = \sum_{i=1}^{n} \alpha_i (\hat{z}_i - \hat{z}_0)$, and after an application of A, $A\xi - A\hat{z}_0 = \sum_{i=1}^{n} \alpha_i e_i$. The vector on the left has rational coordinates by the assumptions and by what was just said about A. The vector on the right is $[\alpha_1, \dots, \alpha_n, 0, \dots, 0]^{\mathsf{T}}$. Hence the coefficients $\alpha_1, \dots, \alpha_n$ are rational, and so is $\alpha_0 = 1 - \sum_{i=1}^{n} \alpha_i$.

Now consider the case of a general $\mathscr{R}$. By Carathéodory's theorem, every point in the convex hull of $\mathscr{R}$ is a convex combination of $d + 1$ or fewer affinely independent points of $\mathscr{R}$ [34, cor. 17.1.1]. Thus the argument given above covers the general case. □

The next simple corollary characterizes the existence of a loop.

COROLLARY A.2. *The existence of a loop* (*i.e.*, $z_{1,m} \in \mathscr{R}^m$ *with* $z_1 + \cdots + z_m = 0$) *is equivalent to* $0 \in \mathcal{U}$.

This corollary expresses the irreducibility assumption used in [32] in terms of the convex hull of $\mathscr{R}$.

COROLLARY A.3. *There is a path from* 0 *to each* $y \in \mathcal{G}$ *with steps from* $\mathscr{R}$ *if and only if* 0 *is in the relative interior of* $\mathcal{U}$.

PROOF. Each $y \in \mathcal{G}$ is reachable from 0 if and only if $-x$ is reachable from 0 for each $x \in \mathscr{R}$. This is equivalent to the existence of an identity $0 = x_1 + \cdots + x_m$ where each x_i is in $\mathscr{R}$ and each $z \in \mathscr{R}$ appears at least once among the x_i's. Equivalently, we can write 0 as a convex combination of $\mathscr{R}$ so that each $z \in \mathscr{R}$ has a positive rational coefficient. By using Lemma A.1, this in turn is equivalent to the following statement: for each $z \in \mathscr{R}$, $-\varepsilon z \in \mathcal{U}$ for small enough $\varepsilon > 0$. By theorem 6.4 in [34] this is the same as $0 \in \operatorname{ri} \mathcal{U}$. □

This lemma gives sufficient conditions for membership in class $\mathcal{L}$ of Definition 2.1.

LEMMA A.4. *Let* $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_x : x \in \mathcal{G}\})$ *be a measurable ergodic dynamical system. Let* $0 \le g \in L^1(\mathbb{P})$. *Assume one of the four conditions below.*

(a) g *is bounded.*
(b) $d = 1$.
(c) $d \ge 2$. *There exist* $r \in (0, \infty)$ *and* $p > d$ *such that* $\mathbb{E}[g^p] < \infty$ *and* $\{g \circ T_{x_i} : i = 1, \dots, m\}$ *are i.i.d. whenever* $|x_i - x_j| \ge r$ *for all* $i \ne j$.
(d) $d \ge 2$. *There exist* $a > d$ *and* $p > ad/(a - d)$ *such that* $\mathbb{E}[g^p] < \infty$, *and for each* $z \in \mathscr{R} \setminus \{0\}$ *and large* $k \in \mathbb{N}$,

$$\sup_{\substack{A \in \sigma(g \circ T_x : x \cdot z \le 0) \\ B \in \sigma(g \circ T_x : x \cdot z \ge k)}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \le k^{-a}. \tag{A.1}$$

Then, for each $z \in \mathscr{R} \setminus \{0\}$,

$$\lim_{\varepsilon\to 0}\varlimsup_{n\to\infty}\max_{x\in\mathcal{G}:|x|\le n}\frac{1}{n}\sum_{i=0}^{\varepsilon n} g\circ T_{x+iz}=0\quad \mathbb{P}\text{-}a.s. \tag{A.2}$$

PROOF. Part (a) is immediate.

For (b) let $s\in\mathbb{N}$ be such that $\mathcal{G}=\{ns:n\in\mathbb{Z}\}$. Fix $z=as$ and let $\overline{g}=g-\mathbb{E}(g\mid\mathcal{I}_z)$ where $\mathcal{I}_z$ is the σ-algebra of events invariant under T_z. By T_{as}-invariance

$$\max_{-n\le j\le n}\frac{1}{n}\sum_{i=0}^{\varepsilon n}\mathbb{E}(g\mid\mathcal{I}_z)\circ T_{js+ias}\le\left(\varepsilon+\frac{1}{n}\right)\max_{0\le j<a}\mathbb{E}(g\mid\mathcal{I}_z)\circ T_{js}\quad\mathbb{P}\text{-a.s.}$$

By the ergodic theorem

$$\varlimsup_{n\to\infty}\max_{|\ell|\le n}\left|\frac{1}{n}\sum_{i=0}^{n}\overline{g}\circ T_{\ell z+iz}\right|=0\quad\mathbb{P}\text{-a.s.}$$

This limit is not changed by taking a finite maximum over the shifts by T_{js}, $0\le j<a$.

Part (c) follows from part (d).

Fix z for part (d). First two reductions.

(i) The maximum over x in (A.2) can be restricted to a set A_n of size $|A_n|\le Cn^{d-1}\varepsilon^{-1}$ at the expense of doubling ε in the upper summation limit. The reason is that $g\ge 0$ and if $x'=x+jz$ for some $1\le j<n\varepsilon/2$, then the $2n\varepsilon$-sum started at x covers the $n\varepsilon$-sum started at x'.

(ii) It suffices to consider a subsequence $n_m=m^\gamma$ for any fixed $\gamma>0$ because $n_{m+1}/n_m\to 1$ and $g\ge 0$.

Since constants satisfy (A.2) we can replace g with $\overline{g}=g-\mathbb{E}[g]$. Let $S_n^x=\sum_{i=0}^n\overline{g}\circ T_{x+iz}$. Equation (A.1) and the translation invariance of $\mathbb{P}$ imply strong mixing as defined by [33]. Then applying theorem 6 therein with $u=n^{-b}$, r large enough, and $t=\delta n/(cr)$, we get a generalization of the Fuk-Nagaev inequality to square-integrable, mean-zero, strongly mixing random variables. This implies that for fixed $\varepsilon,\delta>0$, $\mathbb{P}\{|S_{n\varepsilon}^x|>n\delta\}\le C(\varepsilon,\delta)n^{1-b}$ with $b=ap/(a+p)>d$. By a straightforward union bound

$$\mathbb{P}\left\{\max_{x\in A_n}\left|\sum_{i=0}^{n\varepsilon}\overline{g}\circ T_{x+iz}\right|>n\delta\right\}\le Cn^{d-1}\varepsilon^{-1}\mathbb{P}\{|S_{n\varepsilon}^0|>n\delta\}\le C(\varepsilon,\delta)n^{d-b}.$$

Along the subsequence $n_m=m^\gamma$ for $\gamma>(b-d)^{-1}$ the last bound is summable. We get $\mathbb{P}$-a.s. convergence to 0 for each fixed $\varepsilon>0$ by the Borel-Cantelli lemma. □

For a general ergodic system (a) cannot be improved. For example, take $d=2$, an i.i.d. sequence $\{\omega_{i,0}\}_{i\in\mathbb{Z}}$, and then set $\omega_{i,j}=\omega_{i,0}$. For $z=e_2$, we have

$n^{-1}\sum_{j=0}^{n\varepsilon}|\omega_{x+(0,j)}| \geq \varepsilon|\omega_x|$ and consequently the limit in n in (A.2) blows up unless $\omega_{i,j}$ is a bounded process.

If the mixing in part (d) above is faster than any polynomial, then we can take $a \to \infty$ and the condition becomes $p > d$. Part (c) is close to optimal. If $\mathbb{E}[g^d] = \infty$ then $n^{-1}\max_{|x|\leq n} g \circ T_x$ blows up by the second Borel-Cantelli lemma. Currently we do not know if $p \geq d$ is sufficient in (c).

Appendix B Weak LDP through a Projective Limit

We describe a small alteration of the projective limit LDP. Let $\mathcal{X}$ and $\mathcal{X}_j$, $j \in \mathbb{N}$, be metric spaces with continuous maps $g_j : \mathcal{X} \to \mathcal{X}_j$ and $g_{j,i} : \mathcal{X}_j \to \mathcal{X}_i$ for $i < j$ such that $g_i = g_{j,i} \circ g_j$ and $g_{k,i} = g_{j,i} \circ g_{k,j}$. Let $\{\mu_n\}$ be a sequence of Borel probability measures on $\mathcal{X}$, and define $\mu_n^j = \mu_n \circ g_j^{-1}$ on $\mathcal{X}_j$. Let $I_j : \mathcal{X}_j \to [0,\infty]$ be lower-semicontinuous. Define $I(x) = \sup_j I_j(g_j(x))$ for $x \in \mathcal{X}$.

THEOREM B.1.

(i) *Suppose that for all j, $I_j \circ g_{j+1,j} \leq I_{j+1}$ and I_j satisfies the large deviation upper bound for compact sets in $\mathcal{X}_j$. Then I satisfies the large deviation upper bound for compact sets in $\mathcal{X}$.*

(ii) *Assume that $\mathcal{U} = \{g_j^{-1}(U_j) : j \in \mathbb{N}, U_j \subseteq \mathcal{X}_j \text{ open}\}$ is a base for the topology of $\mathcal{X}$. Suppose that for all j, I_j satisfies the large deviation lower bound for open sets in $\mathcal{X}_j$. Then I satisfies the large deviation lower bound for open sets in $\mathcal{X}$.*

PROOF. Part (ii) is straightforward. We prove part (i). Let $A \subseteq \mathcal{X}$ be compact. Since $g_j(A)$ is compact in $\mathcal{X}_j$ and $g_j^{-1}(g_j(A)) \supseteq A$,

$$\overline{\lim}\, n^{-1}\log\mu_n(A) \leq \overline{\lim}\, n^{-1}\log\mu_n^j(g_j(A)) \leq -\inf_{y\in g_j(A)} I_j(y) = -\inf_{x\in A} I_j(g_j(x)),$$

from which

$$\overline{\lim}\, n^{-1}\log\mu_n(A) \leq -\sup_j \inf_{x\in A} I_j(g_j(x)).$$

Next we claim a minimax property from the assumption of monotonicity:

$$\sup_j \inf_{x\in A} I_j(g_j(x)) = \inf_{x\in A}\sup_j I_j(g_j(x)) \equiv \inf_{x\in A} I(x). \tag{B.1}$$

Inequality $\leq$ is obviously true. To show $\geq$, let $c < \inf_{x\in A}\sup_j I_j(g_j(x))$. Then each $x \in A$ has an index $j(x)$ such that $I_{j(x)}(g_{j(x)}(x)) > c$. The set $D_x = \{z \in \mathcal{X} : I_{j(x)}(g_{j(x)}(z)) > c\}$ is open by the continuity of g_j and lower semicontinuity of I_j. Cover A with finitely many sets: $A \subseteq D_{x_1} \cup \cdots \cup D_{x_k}$. Fix $j \geq j(x_1) \vee \cdots \vee j(x_k)$. Then if $x \in A$ pick ℓ such that $x \in D_{x_\ell}$, and we have

$$I_j(g_j(x)) \geq I_{j(x_\ell)}\big(g_{j,j(x_\ell)}(g_j(x))\big) = I_{j(x_\ell)}\big(g_{j(x_\ell)}(x)\big) > c.$$

Thus $\inf_{x\in A} I_j(g_j(x)) \ge c$. We have proved (B.1) and thereby the upper large deviation bound for A. $\square$

Appendix C Proofs of Lemmas 2.9 and 2.13

Standing assumptions in this section are the same as in Section 2: $(\Omega, \mathfrak{S}, \mathbb{P}, \{T_z : z \in \mathcal{G}\})$ is a measurable ergodic dynamical system and, as throughout the paper, $\mathscr{R}$ is an arbitrary finite subset of $\mathbb{Z}^d$ that generates the additive group $\mathcal{G}$. Throughout this section $\ell \ge 0$ is a fixed integer. C denotes a chameleon constant that can change from term to term and only depends on $\mathscr{R}$, ℓ, and d. In order to avoid working on a sublattice, we will assume throughout this appendix that $\mathscr{R}$ generates $\mathbb{Z}^d$ as a group. This does not cause any loss of generality. The additive group $\mathcal{G}$ generated by $\mathscr{R}$ is linearly isomorphic to $\mathbb{Z}^{d'}$ for some $d' \le d$ [40, pp. 65–66] and we can transport the model to $\mathbb{Z}^{d'}$.

A crucial tool will be the path integral of a function $F \in \mathcal{K}_\ell$. The main idea is that due to the closed loop property these functions are gradientlike.

For ℓ-tuples $\tilde{z}_{1,\ell}, \bar{z}_{1,\ell} \in \mathscr{R}^\ell$ we write $\tilde{x}_\ell = \tilde{z}_1 + \cdots + \tilde{z}_\ell$ and $\bar{x}_\ell = \bar{z}_1 + \cdots + \bar{z}_\ell$. We say that there exists a path from $(y, \tilde{z}_{1,\ell})$ to $(x, z_{1,\ell})$ when there exist $a_1, \dots, a_m \in \mathscr{R}$ such that 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 $\omega \in \Omega$. This is equivalent to the pair of equations

$$y + \tilde{x}_\ell + a_1 + \cdots + a_{m-\ell} = x \quad \text{and} \quad a_{m-\ell+1,m} = z_{1,\ell}.$$

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, for any finite number of points there is a common starting point from which there exists a path to each of the chosen points.

Now fix a measurable function $F : \mathbf{\Omega}_\ell \times \mathscr{R} \to \mathbb{R}$ that satisfies the closed loop property (iii) of Definition 2.2. If there is a path $(a_i)_{i=1}^m$ 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

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

By the closed loop property, $L(\omega, (y, \tilde{z}_{1,\ell}), (x, z_{1,\ell}))$ is independent of the path chosen. We also admit an empty path that gives

$$L(\omega, (x, z_{1,\ell}), (x, z_{1,\ell})) = 0.$$

If $a_1, \dots, a_m$ 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 (C.1) is to shift ω by u, and consequently

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

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

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

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.

LEMMA C.1. *Let $F(\cdot, z_{1,\ell}, z) \in L^1(\mathbb{P})$ for each $(z_{1,\ell}, z)$ and satisfy the closed loop property* (iii) *of Definition* 2.2.

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

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

In particular, $f \in L^1(\mathbb{P})$ for all $(z_{1,\ell}, \bar{z}_{1,\ell}, x)$.

(b) *For $z_{1,\ell}, \bar{z}_{1,\ell}, \tilde{z}_{1,\ell} \in \mathscr{R}^\ell$, $x, \bar{x} \in \mathbb{Z}^d$, and $\mathbb{P}$-a.e. ω,*

$$f(\omega, z_{1,\ell}, \tilde{z}_{1,\ell}, \bar{x}) = f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, x) + f(T_x\omega, \bar{z}_{1,\ell}, \tilde{z}_{1,\ell}, \bar{x} - x).$$

(c) *Assume additionally that F satisfies the mean-zero property* (ii) *of Definition* 2.2. *Then for any $\bar{z}_{1,\ell} \in \mathscr{R}^\ell$ and $x \in \mathbb{Z}^d$, $\mathbb{E}[f(\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, x)] = 0$.*

PROOF. Let $e_1, \dots, e_d$ be the canonical basis of $\mathbb{R}^d$. For each $1 \le i \le d$, there exist nonnegative integers $n_i^\pm$ and $(a_{i,j}^\pm)_{j=1}^{n_i^\pm}$ from $\mathscr{R}$ such that

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

Write $x = \sum_{i=1}^d b_i \epsilon_i e_i$ with $b_i \ge 0$ and $\epsilon_i \in \{-1, +1\}$. Then,

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

One can thus find a y that has paths to both 0 and x that stay inside a ball of radius $C(|x| + 1)$. This proves (a).

To prove (b), let $(y, \hat{z}_{1,\ell})$ have paths to $(-x, z_{1,\ell})$, $(0, \bar{z}_{1,\ell})$, and $(\bar{x} - x, \tilde{z}_{1,\ell})$. Use the definition of f (C.3) and the shift property (C.2) to write

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

For (c), let y be 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

$$f(\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, x) = L(\omega, (y, \bar{z}_{1,\ell}), (x, \bar{z}_{1,\ell})) - L(\omega, (y, \bar{z}_{1,\ell}), (0, \bar{z}_{1,\ell})).$$

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 zero $\mathbb{E}$-mean by property (ii) of Definition 2.2. □

Remark C.2. Part (b) above shows that f is a path integral of F or, alternatively, that F is a gradient of f. More precisely,

$$F(\omega, z_{1,\ell}, z) = f(\omega, \bar{z}_{1,\ell}, S_z^+ z_{1,\ell}, z_1) - f(\omega, \bar{z}_{1,\ell}, z_{1,\ell}, 0)$$

for all $z_{1,\ell}, \bar{z}_{1,\ell} \in \mathscr{R}^\ell$, $z \in \mathscr{R}$, and $\mathbb{P}$-a.e. ω. (S_z^+ acts on $\mathscr{R}^\ell$ in the obvious way.)

LEMMA C.3. *Let $F \in \mathcal{K}_\ell$. Then there exists a sequence of bounded measurable functions $h_k : \mathbf{\Omega}_\ell \to \mathbb{R}$ such that $\mathbb{E}[|h_k(S_z^+\eta) - h_k(\eta) - F(\eta, z)|] \to 0$ for all $z_{1,\ell} \in \mathscr{R}^\ell$ and $z \in \mathscr{R}$.*

PROOF OF LEMMA C.3. Starting with F, denote its path integral by f as above. Define

$$g_n(\omega, z_{1,\ell}) = -|\mathscr{R}|^{-\ell}(2n + 1)^{-d} \sum_{\bar{z}_{1,\ell} \in \mathscr{R}^\ell} \sum_{|x| \le n} f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, x).$$

By part (b) of Lemma C.1

$$\begin{aligned} & f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, x) + f(T_x\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, z_1) \\ &\quad = f(\omega, z_{1,\ell}, \bar{z}_{1,\ell}, x + z_1) \\ &\quad = f(\omega, z_{1,\ell}, S_z^+ z_{1,\ell}, z_1) + f(T_{z_1}\omega, S_z^+ z_{1,\ell}, \bar{z}_{1,\ell}, x). \end{aligned}$$

Consequently, from the closed loop property alone,

$$
\begin{aligned}
g_n(S_z^+\eta) - g_n(\eta) &= |\mathscr{R}|^{-\ell}(2n+1)^{-d} \sum_{\bar z_{1,\ell}\in\mathscr{R}^\ell} \sum_{|x|\le n} \big[f(\omega, z_{1,\ell}, \bar z_{1,\ell}, x) \\
&\qquad\qquad - f(T_{z_1}\omega, S_z^+ z_{1,\ell}, \bar z_{1,\ell}, x)\big] \\
&= |\mathscr{R}|^{-\ell}(2n+1)^{-d} \sum_{\bar z_{1,\ell}\in\mathscr{R}^\ell} \sum_{|x|\le n} \big[f(\omega, z_{1,\ell}, S_z^+ z_{1,\ell}, z_1) \\
&\qquad\qquad - f(T_x\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, z_1)\big]
\end{aligned}
$$

$$
= F(\eta, z) - |\mathscr{R}|^{-\ell}(2n+1)^{-d} \sum_{\bar z_{1,\ell}\in\mathscr{R}^\ell} \sum_{|x|\le n} f(T_x\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, z_1). \tag{C.4}
$$

By parts (a) and (c) of Lemma C.1 and by the ergodic theorem we see that $F(\eta, z)$ is the $L^1(\mathbb{P})$-limit of $g_n(S_z^+\eta) - g_n(\eta)$ for each $z_{1,\ell} \in \mathscr{R}^\ell$ and $z \in \mathscr{R}$. Finally, approximate the integrable g_n with a bounded h_n in $L^1(\mathbb{P})$. □

PROOF OF LEMMA 2.9. The $L^1(\mathbb{P})$ convergence to 0 follows from Lemma C.3. Next, observe that for any $a_{1,n}$ that satisfies the properties in braces in the statement of the lemma, the F-sum satisfies

$$
\sum_{i=0}^{n-1} F(\eta_i, a_{i+1}) = f(\omega, z_{1,\ell}, \bar z_{1,\ell}, \hat x_n(\xi)).
$$

Consequently, the task is to show that $n^{-1} f(\omega, z_{1,\ell}, \bar z_{1,\ell}, \hat x_n(\xi))$ has a limit $\mathbb{P}$-a.s.

Recall that the definition of the path $\hat x_\centerdot(\xi)$ given above Lemma 2.8 involved an integer $b = b(\xi)$ such that $b\xi \in \mathbb{Z}^d$ and $\hat x_{mb}(\xi) = mb\xi$ for all m. Using (b) of Lemma C.1 we have

$$
\begin{aligned}
(mb)^{-1} f(\omega, z_{1,\ell}, \bar z_{1,\ell}, mb\xi) &= (mb)^{-1} f(\omega, z_{1,\ell}, \bar z_{1,\ell}, 0) \\
&\quad + (mb)^{-1} \sum_{j=0}^{m-1} f(T_{jb\xi}\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, b\xi),
\end{aligned}
$$

and by the ergodic theorem, the right-hand side has a $\mathbb{P}$-a.s. limit.

Given n choose m_n so that $m_n b \le n < (m_n+1)b$. By (a) and (b) of Lemma C.1,

$$
\begin{aligned}
&n^{-1}|f(\omega, z_{1,\ell}, \bar z_{1,\ell}, \hat x_n(\xi)) - f(\omega, z_{1,\ell}, \bar z_{1,\ell}, m_n b\xi)| \\
&\quad = n^{-1}|f(T_{m_n b\xi}\omega, \bar z_{1,\ell}, \bar z_{1,\ell}, \hat x_n(\xi) - m_n b\xi)| \\
&\quad \le (m_n b)^{-1} G(T_{m_n b\xi}\omega) \underset{n\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-a.s.},
\end{aligned}
$$

where

$$G(\omega) = \sum_{x:|x| \le C(b|\xi|+1)} \max_{\widetilde{z}_{1,\ell} \in \mathscr{R}^\ell} \max_{z \in \mathscr{R}} |F(T_x\omega, \widetilde{z}_{1,\ell}, z)|$$

is in $L^1(\mathbb{P})$. □

PROOF OF LEMMA 2.13. Fix $\varepsilon > 0$ for the rest of the proof. From (2.13) we have that for $\mathbb{P}$-a.e. ω and for all $z_{1,\ell} \in \mathscr{R}^\ell$ and $z \in \mathscr{R}$

$$F^{(0)}_{k,\varepsilon}(\eta, z) \le C - \mathbb{E}[g(\eta) \mid \mathfrak{S}_k].$$

By the fact that $g(\omega, z_{1,\ell}) \in L^1(\mathbb{P})$ for all $z_{1,\ell} \in \mathscr{R}^\ell$, the right-hand side is uniformly integrable. Thus so is $F^{(0),+}_{k,\varepsilon} = \max(F^{(0)}_{k,\varepsilon}, 0)$.

Let $F^{(0),-}_{k,\varepsilon} = \max(-F^{(0)}_{k,\varepsilon}, 0)$. Observe that by the T_z-invariance of $\mathbb{P}$

$$\begin{aligned}\mathbb{E}\big[F^{(0)}_{k,\varepsilon}(\omega, z_{1,\ell}, z)\big] &= \mathbb{E}[h_{k,\varepsilon}(T_{z_1}\omega, S^+_z z_{1,\ell}) - h_{k,\varepsilon}(\omega, z_{1,\ell})] \\ &= \mathbb{E}[h_{k,\varepsilon}(\omega, S^+_z z_{1,\ell})] - \mathbb{E}[h_{k,\varepsilon}(\omega, z_{1,\ell})].\end{aligned}$$

Thus $F^{(0)}_{k,\varepsilon}$ satisfies the mean-zero property (ii) in Definition 2.2. Letting $\eta_0 = (\omega, z_{1,\ell})$, $z_0 = z$, $a_i = z_{i-1}$, and $\eta_i = S^+_{a_i}\eta_{i-1}$ for $i = 1, \dots, \ell+1$, one has that

$$\mathbb{E}\big[F^{(0),-}_{k,\varepsilon}(\eta, z)\big] \le \sum_{i=0}^{\ell} \mathbb{E}\big[F^{(0),-}_{k,\varepsilon}(\eta_i, a_{i+1})\big] = \sum_{i=0}^{\ell} \mathbb{E}\big[F^{(0),+}_{k,\varepsilon}(\eta_i, a_{i+1})\big]$$

is bounded uniformly in k. We apply the following lemma to extract a uniformly integrable part leaving a small error.

LEMMA C.4 (Lemma 4.3 of [28]). *Let $\{g_n\}_{n\ge1}$ be a sequence of nonnegative functions such that $\sup_n E[g_n] \le C$. Then there is a subsequence $\{n_j\}_{j\ge1}$ and an increasing sequence $a_j \nearrow \infty$ such that $g_{n_j}\mathbb{1}\{g_{n_j} \le a_j\}$ is uniformly integrable and $g_{n_j}\mathbb{1}\{g_{n_j} > a_j\}$ converges to 0 in probability.*

By the above lemma we can write $F^{(0),-}_{k,\varepsilon} = \widetilde{F}^{(0)}_{k,\varepsilon} + R^{(0)}_{k,\varepsilon}$ such that, along a subsequence, $\widetilde{F}^{(0)}_{k,\varepsilon}$ is uniformly integrable and $R^{(0)}_{k,\varepsilon} \ge 0$ is $\mathfrak{S}_k$-measurable and converges to 0 in $\mathbb{P}$-probability. One can then take a further subsequence along which $\widehat{F}^{(0)}_{k,\varepsilon} = F^{(0),+}_{k,\varepsilon} - \widetilde{F}^{(0)}_{k,\varepsilon}$ converges in weak $L^1(\mathbb{P})$ to $\widehat{F}^{(0)}_{\varepsilon}$ and $R^{(0)}_{k,\varepsilon} \to 0$ $\mathbb{P}$-a.s. (Uniform integrability gives sequential compactness in weak L^1; see [16, theorem 9, p. 292].) We will always keep indexing subsequences by k. Now we have the decomposition

$$F^{(0)}_{k,\varepsilon} = \widehat{F}^{(0)}_{k,\varepsilon} - R^{(0)}_{k,\varepsilon}. \tag{C.5}$$

An attempt to check that the limit $\widehat{F}^{(0)}_{\varepsilon}$ satisfies the closed loop property runs into difficulty because we have very weak control of the errors $R^{(0)}_{k,\varepsilon}$, and the conditioning in definition (2.14) damages the closed loop property of the function

$h_{k,\varepsilon}(S_z^+\eta) - h_{k,\varepsilon}(\eta)$. To get around this we defined the family indexed by i in (2.14). In the next lemma we develop a hierarchy of errors obtained by successive applications of Lemma C.4. We give the proof after the current proof is done. Recall that given $z_{1,j} \in \mathscr{R}^j$, $x_j = z_1 + \cdots + z_j$. We will use the notation $\varnothing$ for a path of length $j = 0$ and then $x_0 = 0$.

LEMMA C.5. *There exist nonnegative random variables on $\mathbf{\Omega}_\ell \times \mathscr{R}$, denoted by $\widetilde{R}_{k,\varepsilon}^{(i,j,z_{1,j})}$, $\widehat{R}_{k,\varepsilon}^{(i,j,z_{1,j+1})}$, and $R_{k,\varepsilon}^{(i,j,z_{1,j})}$, with $0 \le j \le i \le k$ and $z_{1,j+1} \in \mathscr{R}^{j+1}$, such that the following properties are satisfied*:

(a) $R_{k,\varepsilon}^{(0,0,\varnothing)} = R_{k,\varepsilon}^{(0)}$.

(b) $\widetilde{R}_{k,\varepsilon}^{(i,j,z_{1,j})}$, $\widehat{R}_{k,\varepsilon}^{(i,j,z_{1,j+1})}$, *and* $R_{k,\varepsilon}^{(i,j,z_{1,j})}$ *are* $T_{x_j}\mathfrak{S}_{k-i}$*-measurable.*

(c) $\mathbb{E}[R_{k,\varepsilon}^{(i,0,\varnothing)} \mid \mathfrak{S}_{k-i-1}] = \widetilde{R}_{k,\varepsilon}^{(i+1,0,\varnothing)} + R_{k,\varepsilon}^{(i+1,0,\varnothing)}$ *for all* $i \ge 0$.

(d) $\mathbb{E}[R_{k,\varepsilon}^{(i,j,z_{1,j})} \mid T_{x_{j-1}}\mathfrak{S}_{k-i-1}] = \widehat{R}_{k,\varepsilon}^{(i+1,j-1,z_{1,j})} + R_{k,\varepsilon}^{(i+1,j-1,z_{1,j-1})}$ *for all* $i \ge j \ge 1$ *and* $z_{1,j} \in \mathscr{R}^j$.

(e) $\mathbb{E}[R_{k,\varepsilon}^{(i,j,z_{1,j})} \mid T_{x_{j+1}}\mathfrak{S}_{k-i-1}] = \widetilde{R}_{k,\varepsilon}^{(i+1,j+1,z_{1,j+1})} + R_{k,\varepsilon}^{(i+1,j+1,z_{1,j+1})}$ *for all* $i \ge j \ge 0$ *and* $z_{1,j+1} \in \mathscr{R}^{j+1}$.

(f) *As* $k \to \infty$, $\widetilde{R}_{k,\varepsilon}^{(i,j,z_{1,j})}$ *and* $\widehat{R}_{k,\varepsilon}^{(i,j,z_{1,j+1})}$ *are uniformly integrable and converge in weak* $L^1(\mathbb{P})$ *to a limit* $\widetilde{R}_{\varepsilon}^{(i,j,z_{1,j})}$ *and* $\widehat{R}_{\varepsilon}^{(i,j,z_{1,j+1})}$, *respectively.*

(g) $R_{k,\varepsilon}^{(i,j,z_{1,j})}$ *converges to* 0 $\mathbb{P}$*-a.s. as* $k \to \infty$.

(h) *One has for* $j \ge 0$, $z_{1,j+1} \in \mathscr{R}^{j+1}$, *and* $s \ge 1$

$$\begin{aligned}
&\widetilde{R}_{\varepsilon}^{(s,0,\varnothing)} + \widetilde{R}_{\varepsilon}^{(s+1,0,\varnothing)} + \widetilde{R}_{\varepsilon}^{(s+2,1,z_1)} \\
&\quad + \cdots + \widetilde{R}_{\varepsilon}^{(j+s,j-1,z_{1,j-1})} + \widetilde{R}_{\varepsilon}^{(j+s+1,j,z_{1,j})} \\
&\qquad = \widetilde{R}_{\varepsilon}^{(s,1,z_1)} + \cdots + \widetilde{R}_{\varepsilon}^{(j+s,j+1,z_{1,j+1})} + \widehat{R}_{\varepsilon}^{(j+s+1,j,z_{1,j+1})}.
\end{aligned}$$

(i) *For any fixed* $j \ge 0$ *and* $z_{1,j+1} \in \mathscr{R}^{j+1}$ *both* $\widetilde{R}_{\varepsilon}^{(i,j,z_{1,j})}$ *and* $\widehat{R}_{\varepsilon}^{(i,j,z_{1,j+1})}$ *converge to* 0 *strongly in* $L^1(\mathbb{P})$ *as* $i \to \infty$.

The limits as $k \to \infty$ *are to be understood in the sense that there exists one subsequence along which all the countably many claimed limits hold simultaneously.*

Fix $i \ge 0$ and let $k \ge i$. Starting with (C.5), using (a) of the above lemma and applying (c) repeatedly, we have the decomposition $F_{k,\varepsilon}^{(i)} = \widehat{F}_{k,\varepsilon}^{(i)} - R_{k,\varepsilon}^{(i)}$ with

$$\widehat{F}_{k,\varepsilon}^{(i)} = \mathbb{E}\big[\widehat{F}_{k,\varepsilon}^{(0)} - \widetilde{R}_{k,\varepsilon}^{(1,0,\varnothing)} - \cdots - \widetilde{R}_{k,\varepsilon}^{(i,0,\varnothing)} \mid \mathfrak{S}_{k-i}\big] \quad \text{and} \quad R_{k,\varepsilon}^{(i)} = R_{k,\varepsilon}^{(i,0,\varnothing)}.$$

$R_{k,\varepsilon}^{(i)}$ is $\mathfrak{S}_{k-i}$-measurable and $\widehat{F}_{k,\varepsilon}^{(i)}$ uniformly integrable. (The proof of theorem 5.1 in chapter 4 of [17] applies to the uniformly integrable sequence $\widehat{F}_{k,\varepsilon}^{(0)} - \widetilde{R}_{k,\varepsilon}^{(1,0,\varnothing)} - \cdots - \widetilde{R}_{k,\varepsilon}^{(i,0,\varnothing)}$.) One can check by a standard π-λ or monotone class argument that any weak $L^1(\mathbb{P})$ limit coincides with the weak limit without the conditioning, namely $\widehat{F}_\varepsilon^{(i)} = \widehat{F}_\varepsilon^{(0)} - \widetilde{R}_\varepsilon^{(1,0,\varnothing)} - \cdots - \widetilde{R}_\varepsilon^{(i,0,\varnothing)}$. Furthermore, since $\mathbb{E}[R_{k,\varepsilon}^{(0)}]$ is uniformly bounded in k, we have

$$\mathbb{E}\big[\widetilde{R}_{k,\varepsilon}^{(1,0,\varnothing)} + \cdots + \widetilde{R}_{k,\varepsilon}^{(i,0,\varnothing)}\big] \le \mathbb{E}\big[R_{k,\varepsilon}^{(0)}\big] \le C.$$

Taking $i \to \infty$ we see that $\widehat{F}_\varepsilon^{(i)}$ decreases, converging strongly in $L^1(\mathbb{P})$ to

$$\widehat{F}_\varepsilon = \widehat{F}_\varepsilon^{(0)} - \sum_{i\ge 1} \widetilde{R}_\varepsilon^{(i,0,\varnothing)}.$$

Then, $\widehat{F}_\varepsilon \in L^1(\mathbb{P})$ satisfies (i) of Definition 2.2.

Fix a path $x_{0,\infty}$ with increments in $\mathscr{R}$. Fix integers $k \ge b \ge j \ge 0$. Recall that $T_{\pm z}\mathfrak{S}_{s-1} \subset \mathfrak{S}_s$ for all $z \in \mathscr{R}$ and $s \ge 1$. In particular, $\mathfrak{S}_{k-b+j} \supset T_{x_1}\mathfrak{S}_{k-b+j-1} \supset \cdots \supset T_{x_j}\mathfrak{S}_{k-b}$. Applying (e) and (b) of Lemma C.5 repeatedly, one has

$$\mathbb{E}\big[R_{k,\varepsilon}^{(b-j)} \bigm| T_{x_j}\mathfrak{S}_{k-b}\big] = \mathbb{E}\Big[\sum_{s=1}^{j} \widetilde{R}_{k,\varepsilon}^{(b-j+s,s,z_{1,s})} \bigm| T_{x_j}\mathfrak{S}_{k-b}\Big] + R_{k,\varepsilon}^{(b,j,z_{1,j})}. \tag{C.6}$$

Thus,

$$\begin{aligned}
&\mathbb{E}\big[h_{k,\varepsilon}(T_{x_{j+1}}\omega, z_{j+2,j+\ell+1}) - h_{k,\varepsilon}(T_{x_j}\omega, z_{j+1,j+\ell}) \bigm| \mathfrak{S}_{k-b}\big] && \text{(C.7)}\\
&\quad= \mathbb{E}\big[\mathbb{E}[h_{k,\varepsilon}(T_{x_{j+1}}\omega, z_{j+2,j+\ell+1})\\
&\qquad\qquad - h_{k,\varepsilon}(T_{x_j}\omega, z_{j+1,j+\ell}) \mid T_{-x_j}\mathfrak{S}_{k-b+j}] \bigm| \mathfrak{S}_{k-b}\big]\\
&\quad= \mathbb{E}\big[F_{k,\varepsilon}^{(b-j)}(T_{x_j}\omega, z_{j+1,j+\ell}, z_{j+\ell+1}) \bigm| \mathfrak{S}_{k-b}\big]\\
&\quad= \mathbb{E}\big[\widehat{F}_{k,\varepsilon}^{(b-j)}(T_{x_j}\omega, z_{j+1,j+\ell}, z_{j+\ell+1}) \bigm| \mathfrak{S}_{k-b}\big]\\
&\qquad - \mathbb{E}\big[R_{k,\varepsilon}^{(b-j)}(T_{x_j}\omega, z_{j+1,j+\ell}, z_{j+\ell+1}) \bigm| \mathfrak{S}_{k-b}\big]\\
&\quad= \mathbb{E}\big[\widehat{F}_{k,\varepsilon}^{(b-j)}(T_{x_j}\omega, z_{j+1,j+\ell}, z_{j+\ell+1}) \bigm| \mathfrak{S}_{k-b}\big] && \text{(C.8)}\\
&\qquad - \mathbb{E}\big[\widetilde{R}_{k,\varepsilon}^{(b-j+1,1,z_1)}(\cdot, z_{j+1,j+\ell}, z_{j+\ell+1})\\
&\qquad\qquad + \cdots + \widetilde{R}_{k,\varepsilon}^{(b,j,z_{1,j})}(\cdot, z_{j+1,j+\ell}, z_{j+\ell+1}) \mid T_{x_j}\mathfrak{S}_{k-b}\big](T_{x_j}\omega) && \text{(C.9)}\\
&\qquad - R_{k,\varepsilon}^{(b,j,z_{1,j})}(T_{x_j}\omega, z_{j+1,j+\ell}, z_{j+\ell+1}).
\end{aligned}$$

The last equality used (C.6) and the formula $\mathbb{E}[g \mid \mathfrak{S}_k] \circ T_x = \mathbb{E}[g \circ T_x \mid T_{-x}\mathfrak{S}_k]$. The two sequences in (C.8) and (C.9) are uniformly integrable and converge weakly in $L^1(\mathbb{P})$ (along a subsequence) to

$$\widehat{F}_\varepsilon^{(b-j)}(T_{x_j}\omega, z_{j+1,j+\ell}, z_{j+\ell+1})$$

and

$$\big(\widetilde{R}_\varepsilon^{(b-j+1,1,z_{1,j})} + \cdots + \widetilde{R}_\varepsilon^{(b,j,z_{1,j})}\big)(T_{x_j}\omega, z_{j+1,j+\ell}, z_{j+\ell+1}),$$

respectively.

For any two paths $\{\eta_i\}_{i=0}^n$ and $\{\overline{\eta}_j\}_{j=0}^m$ as in (iii) of Definition 2.2

$$\sum_{i=0}^{n-1} \mathbb{E}\big[h_{k,\varepsilon}\big(S^+_{a_{i+1}}\eta_i\big) - h_{k,\varepsilon}(\eta_i) \bigm| \mathfrak{S}_{k-b}\big] = \sum_{j=0}^{m-1} \mathbb{E}\big[h_{k,\varepsilon}\big(S^+_{\overline{a}_{j+1}}\overline{\eta}_j\big) - h_{k,\varepsilon}(\overline{\eta}_j) \bigm| \mathfrak{S}_{k-b}\big].$$

Now, taking $b > \max(m,n)$ and further subsequences of (C.7), we arrive at

$$\sum_{i=0}^{n-1} \Big(\widehat{F}_\varepsilon^{(b-i)}(\eta_i, a_{i+1}) - \sum_{s=1}^{i} \widetilde{R}_\varepsilon^{(b-i+s,s,z_{1,s})}(\eta_i, a_{i+1})\Big) = \sum_{j=0}^{m-1} \Big(\widehat{F}_\varepsilon^{(b-j)}(\overline{\eta}_j, \overline{a}_{j+1}) - \sum_{s=1}^{j} \widetilde{R}_\varepsilon^{(b-j+s,s,z'_{1,s})}(\overline{\eta}_j, \overline{a}_{j+1})\Big).$$

Here, $z_{1,n}$ and $z'_{1,m}$ denote the steps of the two paths corresponding to $\{\eta_i\}$ and $\{\overline{\eta}_j\}$. Taking $b \to \infty$ and applying part (i) of Lemma C.5, we conclude that $\widehat{F}_\varepsilon$ satisfies the closed loop property (iii) of Definition 2.2. Next, we work on the mean-zero property.

Abbreviate $\hat{z}_{1,\ell} = (z, \dots, z) \in \mathscr{R}^\ell$. Then,

$$\begin{aligned} c(z) = \mathbb{E}[\widehat{F}_\varepsilon(\omega, \hat{z}_{1,\ell}, z)] &= \inf_i \mathbb{E}[\widehat{F}_\varepsilon^{(i)}(\omega, \hat{z}_{1,\ell}, z)] \\ &= \inf_i \lim_{k\to\infty} \mathbb{E}[\widehat{F}_{k,\varepsilon}^{(i)}(\omega, \hat{z}_{1,\ell}, z)] \ge \inf_i \lim_{k\to\infty} \mathbb{E}[F_{k,\varepsilon}^{(i)}(\omega, \hat{z}_{1,\ell}, z)] \\ &= \lim_{k\to\infty} \mathbb{E}[h_{k,\varepsilon}(T_z\omega, \hat{z}_{1,\ell}) - h_{k,\varepsilon}(\omega, \hat{z}_{1,\ell})] = 0. \end{aligned}$$

Since $\widehat{F}_\varepsilon$ satisfies the closed loop property, one can define its path integral $\widehat{f}_\varepsilon$ as above and use (C.3) to write

$$\begin{aligned} &\widehat{f}_\varepsilon(\omega, \overline{z}_{1,\ell}, \overline{z}_{1,\ell}, z) \\ &\quad = \widehat{F}_\varepsilon(T_{-\ell z}\omega, \hat{z}_{1,\ell}, z) \\ &\qquad + \widehat{F}_\varepsilon(T_{-(\ell-1)z}\omega, \hat{z}_{1,\ell}, \overline{z}_1) + \cdots + \widehat{F}_\varepsilon(\omega, (\hat{z}_\ell, \overline{z}_{1,\ell-1}), \overline{z}_\ell) \\ &\qquad - \widehat{F}_\varepsilon(T_{-\ell z}\omega, \hat{z}_{1,\ell}, \overline{z}_1) - \cdots - \widehat{F}_\varepsilon(T_{-z}\omega, (\hat{z}_\ell, \overline{z}_{1,\ell-1}), \overline{z}_\ell). \end{aligned} \tag{C.10}$$

Thus, we have $c(z) = \mathbb{E}[\widehat{f}_\varepsilon(\omega, \overline{z}_{1,\ell}, \overline{z}_{1,\ell}, z)]$ for all $\overline{z}_{1,\ell} \in \mathscr{R}^\ell$ and $z \in \mathscr{R}$. Hence

$$c(z) = |\mathscr{R}|^{-\ell} \sum_{\overline{z}_{1,\ell} \in \mathscr{R}^\ell} \mathbb{E}[\widehat{f}_\varepsilon(\omega, \overline{z}_{1,\ell}, \overline{z}_{1,\ell}, z)].$$

Integrating (C.4) out (with $F = \widehat{F}_\varepsilon$) one sees that

$$\mathbb{E}[g_n(S_z^+\eta) - g_n(\eta)] = \mathbb{E}[\widehat{F}_\varepsilon(\eta, z) - c(z_1)].$$

Since $g_n(S_z^+\eta) - g_n(\eta)$ has the mean-zero property (ii) of Definition 2.2, we conclude that $\widehat{F}_\varepsilon(\eta, z) - c(z_1)$ does too. Let $\bar{z}_{1,\ell} \in \mathscr{R}^\ell$, $z_{1,n} \in \mathscr{R}^n$, $x = z_1 + \cdots + z_n$, and apply the mean-zero property of $\widehat{F}_\varepsilon(\eta, z) - c(z_1)$ to the path that takes steps $(z_{1,n}, \bar{z}_{1,\ell})$ to go from $(0, \bar{z}_{1,\ell})$ to $(x + \bar{x}_\ell, \bar{z}_{1,\ell})$. This gives

$$\mathbb{E}[\widehat{f}_\varepsilon(\omega, \bar{z}_{1,\ell}, \bar{z}_{1,\ell}, x + \bar{x}_\ell)] = c(\bar{z}_1) + \cdots + c(\bar{z}_\ell) + \sum_{i=1}^n c(z_i). \tag{C.11}$$

Since the left-hand side does not depend on $z_{1,n}$ as long as the increments add up to x, we see that $c(z_1) + \cdots + c(z_n)$ only depends on $z_1 + \cdots + z_n$. Consequently, $\widehat{F}_\varepsilon(\eta, z) - c(z_1)$ also has the closed loop property and thus belongs to $\mathcal{K}_\ell$. This completes the proof of Lemma 2.13. $\square$

PROOF OF LEMMA C.5. In what follows, decomposing a sequence $R_k \geq 0$ means applying Lemma C.4 to it. The leftmost term in the decomposition is the one that converges in weak $L^1(\mathbb{P})$ topology along a subsequence. Its limit is denoted by the same symbol, with k omitted. The rightmost term in the decomposition is the one converging to 0 $\mathbb{P}$-a.s. Subsequences are chosen to work for all $z_{1,j}$, $j \geq 1$, at once, and are still indexed by k. Once a subsequence has been given to suit a decomposition, subsequent decompositions go along this subsequence, and so on. Induction will be repeatedly used in our proof and once an induction is complete, the diagonal trick is used to obtain one subsequence that works for all the terms simultaneously. Recall that $R_{k,\varepsilon}^{(0)} \geq 0$ and $\mathbb{E}[R_{k,\varepsilon}^{(0)}]$ is bounded uniformly in k.

The following diagram may be instructive to the reader during the course of the proof. Index the columns from left to right by $i = 0, 1, \ldots, k$ and the rows from top to bottom by $j = 0, 1, \ldots, k$.

$$\begin{array}{cccccccc}
\mathfrak{S}_k & \mathfrak{S}_{k-1} & \mathfrak{S}_{k-2} & \cdots & \cdots & \mathfrak{S}_{k-i} & \cdots & \mathfrak{S}_0 \\
 & T_{x_1}\mathfrak{S}_{k-1} & T_{x_1}\mathfrak{S}_{k-2} & \cdots & \cdots & T_{x_1}\mathfrak{S}_{k-i} & \cdots & T_{x_1}\mathfrak{S}_0 \\
 & & T_{x_2}\mathfrak{S}_{k-2} & & & \vdots & & \\
 & & & \ddots & & T_{x_j}\mathfrak{S}_{k-i} & & \vdots \\
 & & & & \ddots & \vdots & & \vdots \\
 & & & & & T_{x_i}\mathfrak{S}_{k-i} & & \\
 & & & & & & \ddots & \\
 & & & & & & & T_{x_k}\mathfrak{S}_0
\end{array}$$

The algebra on row j and column $i \geq j$ is $T_{x_j}\mathfrak{S}_{k-i}$. Each algebra in the diagram includes the one down and to the right of it, the one up and to the right of it, and the one immediately to the right of it. The decomposition in (d) corresponds to a step up and to the right in the diagram, while the decomposition in (e) corresponds to a step down and to the right.

We will define $\widetilde{R}_{k,\varepsilon}^{(i,j,z_{1,j})}$, $\widehat{R}_{k,\varepsilon}^{(i,j,z_{1,j+1})}$, and $R_{k,\varepsilon}^{(i,j,z_{1,j})}$ by induction on $s = i - j \geq 0$. On the above diagram, this corresponds to the s^{th} diagonal starting at $\mathfrak{S}_{k-s}$ and going down to $T_{x_1}\mathfrak{S}_{k-s-1}$ and so on. We check property (i) after the whole induction process is complete.

Induction Assumption for s. There exist nonnegative random variables on $\mathbf{\Omega}_\ell \times \mathscr{R}$, denoted by $\widetilde{R}_{k,\varepsilon}^{(i,j,z_{1,j})}$, $\widehat{R}_{k,\varepsilon}^{(i,j,z_{1,j+1})}$, and $R_{k,\varepsilon}^{(i,j,z_{1,j})}$, with $0 \leq j \leq i \leq k$, $i-j \leq s$, and $z_{1,j+1} \in \mathscr{R}^{j+1}$, such that properties (a)–(h) are satisfied (whenever the terms involved have already been defined).

Set $R_{k,\varepsilon}^{(0,0,\varnothing)} = R_{k,\varepsilon}^{(0)}$ and $\widetilde{R}_{k,\varepsilon}^{(0,0,\varnothing)} = 0$. For $k > j \geq 0$, $z_{1,j+1} \in \mathscr{R}^{j+1}$, observe that $T_{x_{j+1}}\mathfrak{S}_{k-j-1} \subset T_{x_j}\mathfrak{S}_{k-j}$ and decompose inductively

$$\mathbb{E}\big[R_{k,\varepsilon}^{(j,j,z_{1,j})} \mid T_{x_{j+1}}\mathfrak{S}_{k-j-1}\big] = \widetilde{R}_{k,\varepsilon}^{(j+1,j+1,z_{1,j+1})} + R_{k,\varepsilon}^{(j+1,j+1,z_{1,j+1})}.$$

For $j \geq 0$ and $z_{1,j+1} \in \mathscr{R}^{j+1}$ set $\widehat{R}_{k,\varepsilon}^{(j,j,z_{1,j+1})} = 0$. These are actually never used in properties (a)–(i) of the lemma. This settles the case $s = 0$.

Next, for $k > 0$, decompose

$$\mathbb{E}\big[R_{k,\varepsilon}^{(0,0,\varnothing)} \mid \mathfrak{S}_{k-1}\big] = \widetilde{R}_{k,\varepsilon}^{(1,0,\varnothing)} + R_{k,\varepsilon}^{(1,0,\varnothing)},$$

and for $k > j \geq 1$ decompose inductively

$$\mathbb{E}\big[R_{k,\varepsilon}^{(j+1,j,z_{1,j})} \mid T_{x_{j+1}}\mathfrak{S}_{k-j-2}\big] = \widetilde{R}_{k,\varepsilon}^{(j+2,j+1,z_{1,j+1})} + R_{k,\varepsilon}^{(j+2,j+1,z_{1,j+1})}.$$

Set $\widehat{R}_{k,\varepsilon}^{(j+1,j,z_{1,j+1})} = 0$ for all $k > j \geq 0$ and $z_{1,j+1} \in \mathscr{R}^{j+1}$. These are again not used in properties (a)–(i) of the lemma. This settles the case $s = 1$.

Now fix $s \geq 1$ and assume the induction assumption for this s. We will define $\widetilde{R}_{k,\varepsilon}^{(j+s+1,j,z_{1,j})}$, $\widehat{R}_{k,\varepsilon}^{(j+s+1,j,z_{1,j+1})}$, and $R_{k,\varepsilon}^{(j+s+1,j,z_{1,j})}$ by induction on $j \geq 0$. On the above diagram, this corresponds to going along the fixed s^{th} diagonal.

Induction Assumption for j with $s \geq 1$ Fixed. We have defined $\widetilde{R}_{k,\varepsilon}^{(j+s+1,j,z_{1,j})}$, $\widehat{R}_{k,\varepsilon}^{(j+s+1,j,z_{1,j+1})}$, and $R_{k,\varepsilon}^{(j+s+1,j,z_{1,j})}$ such that properties (a)–(h) are satisfied (whenever the terms involved have already been defined).

Observe that $\mathfrak{S}_{k-s-1} \subset T_{z_1}\mathfrak{S}_{k-s}$ and temporarily decompose

$$\mathbb{E}\big[R_{k,\varepsilon}^{(s,0,\varnothing)} \mid \mathfrak{S}_{k-s-1}\big] = \widetilde{R}_{k,\varepsilon}^{(s+1,0,\varnothing)} + R_{k,\varepsilon}^{(s+1,0,\varnothing)} \quad \text{and}$$
$$\mathbb{E}\big[R_{k,\varepsilon}^{(s,1,z_1)} \mid \mathfrak{S}_{k-s-1}\big] = \widehat{R}_{k,\varepsilon}^{(s+1,0,z_1)} + \overline{R}_{k,\varepsilon}^{(s+1,0,z_1)}.$$

Let R be the smallest of $\overline{R}_{k,\varepsilon}^{(s+1,0,z_1)}$, $z_1 \in \mathscr{R}$, and $R_{k,\varepsilon}^{(s+1,0,\varnothing)}$. Use (c) and (e) with $i = s-1$ and $j = 0$ to write

$$\begin{aligned}
&\mathbb{E}\big[\widetilde{R}_{k,\varepsilon}^{(s,0,\varnothing)} \mid \mathfrak{S}_{k-s-1}\big] + \widetilde{R}_{k,\varepsilon}^{(s+1,0,\varnothing)} + R_{k,\varepsilon}^{(s+1,0,\varnothing)} \\
&\quad = \mathbb{E}\big[R_{k,\varepsilon}^{(s-1,0,\varnothing)} \mid \mathfrak{S}_{k-s-1}\big] \\
&\quad = \mathbb{E}\big[\widetilde{R}_{k,\varepsilon}^{(s,1,z_1)} \mid \mathfrak{S}_{k-s-1}\big] + \widehat{R}_{k,\varepsilon}^{(s+1,0,z_1)} + \overline{R}_{k,\varepsilon}^{(s+1,0,z_1)}.
\end{aligned}$$

The above display shows that the differences $\overline{R}_{k,\varepsilon}^{(s+1,0,z_1)} - R$ and $R_{k,\varepsilon}^{(s+1,0,\varnothing)} - R$ are uniformly integrable. Redefine all the terms $\overline{R}_{k,\varepsilon}^{(s+1,0,z_1)}$, $z_1 \in \mathscr{R}$, and $R_{k,\varepsilon}^{(s+1,0,\varnothing)}$ to equal R and redefine $\widetilde{R}_{k,\varepsilon}^{(s+1,0,\varnothing)}$ to equal

$$\widetilde{R}_{k,\varepsilon}^{(s+1,0,\varnothing)} + R_{k,\varepsilon}^{(s+1,0,\varnothing)} - R$$

and $\widehat{R}_{k,\varepsilon}^{(s+1,0,z_1)}$ to equal

$$\widehat{R}_{k,\varepsilon}^{(s+1,0,z_1)} + \overline{R}_{k,\varepsilon}^{(s+1,0,z_1)} - R.$$

The upshot is that one can assume that $\overline{R}_{k,\varepsilon}^{(s+1,0,z_1)} = R_{k,\varepsilon}^{(s+1,0,\varnothing)}$ for all $z_1 \in \mathscr{R}$. Taking $k \to \infty$ in the above display verifies (h) for $j = 0$. This starts the induction at $j = 0$.

Now we go from j to $j+1$. Temporarily decompose

$$\mathbb{E}\big[R_{k,\varepsilon}^{(j+s+1,j,z_{1,j})} \mid T_{x_{j+1}}\mathfrak{S}_{k-j-s-2}\big] = \widetilde{R}_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+1})} + R_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+1})} \tag{C.12}$$

and

$$\mathbb{E}\big[R_{k,\varepsilon}^{(j+s+1,j+2,z_{1,j+2})} \mid T_{x_{j+1}}\mathfrak{S}_{k-j-s-2}\big] = \widehat{R}_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+2})} + \overline{R}_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+2})}. \tag{C.13}$$

Then, one has

$$\begin{aligned}
&\mathbb{E}\big[\widetilde{R}_{k,\varepsilon}^{(s,0,\varnothing)} + \widetilde{R}_{k,\varepsilon}^{(s+1,0,\varnothing)} + \widetilde{R}_{k,\varepsilon}^{(s+2,1,z_1)} + \cdots \\
&\quad + \widetilde{R}_{k,\varepsilon}^{(j+s+1,j,z_{1,j})} \mid T_{x_{j+1}}\mathfrak{S}_{k-j-s-2}\big] \\
&\quad + \widetilde{R}_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+1})} + R_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+1})} \\
&\qquad = \mathbb{E}\big[R_{k,\varepsilon}^{(s-1,0,\varnothing)} \mid T_{x_{j+1}}\mathfrak{S}_{k-j-s-2}\big] \\
&\qquad = \mathbb{E}\big[\widetilde{R}_{k,\varepsilon}^{(s,1,z_1)} + \cdots + \widetilde{R}_{k,\varepsilon}^{(j+s+1,j+2,z_{1,j+2})} \mid T_{x_{j+1}}\mathfrak{S}_{k-j-s-2}\big] \\
&\qquad\quad + \widehat{R}_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+2})} + \overline{R}_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+2})}.
\end{aligned}$$

An explanation: Use (c) first twice to condition $R^{(s-1,0,\varnothing)}$ on $\mathfrak{S}_{k-s}$ and then on $\mathfrak{S}_{k-s-1}$. Next, use (e) conditioning $R^{(s+1,0,\varnothing)}$ on $T_{x_1}\mathfrak{S}_{k-s-2}$, then $R^{(s+2,1,z_1)}$ on $T_{x_2}\mathfrak{S}_{k-s-3}$, and so on, until conditioning $R^{(s+j+1,j,z_{1,j})}$ on $T_{x_{j+1}}\mathfrak{S}_{k-s-j-2}$. On the other side, use (e) conditioning $R^{(s-1,0,\varnothing)}$ on $T_{x_1}\mathfrak{S}_{k-s}$, then $R^{(s,1,x_1)}$ conditioned on $T_{x_2}\mathfrak{S}_{k-s-1}$, and so on until $R^{(s+j,j+1,z_{1,j+1})}$ is conditioned on $T_{x_{j+2}}\mathfrak{S}_{k-s-j-1}$. Then use (C.12) and (C.13) and condition $R^{(s+j+1,j+2,z_{1,j+2})}$ on $T_{x_{j+1}}\mathfrak{S}_{k-s-j-2}$.

Now, repeating what we have done for the case $j = 0$, we can assume that $\overline{R}_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+2})} = R_{k,\varepsilon}^{(j+s+2,j+1,z_{1,j+1})}$ for all $z_{1,j+2} \in \mathscr{R}^{j+2}$. Taking $k \to \infty$ verifies (h).

We have achieved the induction on j and thus also the induction on s. Our construction is thus complete once we prove it satisfies (i). Using (h), this follows easily by induction on $j \geq 0$ once one shows that $\widetilde{R}_\varepsilon^{s,0,\varnothing} \to 0$ strongly in $L^1(\mathbb{P})$, which itself follows from the fact that $\mathbb{E}[\widetilde{R}_{k,\varepsilon}^{(1,0,\varnothing)} + \cdots + \widetilde{R}_{k,\varepsilon}^{(s,0,\varnothing)}] \leq \mathbb{E}[R_{k,\varepsilon}^{(0)}]$ is uniformly bounded in k. The lemma is proved. □

Acknowledgment. We thank Elena Kosygina and Fraydoun Rezakhanlou for valuable discussions. Rassoul-Agha was supported in part by National Science Foundation Grant DMS-0747758. Seppäläinen was supported in part by National Science Foundation Grant DMS-1003651 and by the Wisconsin Alumni Research Foundation.

Bibliography

[1] Armstrong, S. N.; Souganidis, P. E. Stochastic homogenization of Hamilton-Jacobi and degenerate Bellman equations in unbounded environments. Preprint, 2011. arXiv:1103.2814 [math.AP]

[2] Avena, L.; den Hollander, F.; Redig, F. Large deviation principle for one-dimensional random walk in dynamic random environment: attractive spin-flips and simple symmetric exclusion. *Markov Process. Related Fields* **16** (2010), no. 1, 139–168.

[3] Bolthausen, E. A note on the diffusion of directed polymers in a random environment. *Comm. Math. Phys.* **123** (1989), no. 4, 529–534.

[4] Bolthausen, E.; Sznitman, A.-S. *Ten lectures on random media*. DMV Seminar, 32. Birkhäuser, Basel, 2002.

[5] Carmona, P.; Hu, Y. Fluctuation exponents and large deviations for directed polymers in a random environment. *Stochastic Process. Appl.* **112** (2004), no. 2, 285–308.

[6] Chernov, A. A. Replication of a multicomponent chain, by the "lightning mechanism". *Biophysics* **12** (1967), 336–341.

[7] Comets, F.; Gantert, N.; Zeitouni, O. Quenched, annealed and functional large deviations for one-dimensional random walk in random environment. *Probab. Theory Related Fields* **118** (2000), no. 1, 65–114.

[8] Comets, F.; Shiga, T.; Yoshida, N. Directed polymers in a random environment: path localization and strong disorder. *Bernoulli* **9** (2003), no. 4, 705–723.

[9] Comets, F.; Shiga, T.; Yoshida, N. Probabilistic analysis of directed polymers in a random environment: a review. *Stochastic analysis on large scale interacting systems*, 115–142. Advanced Studies in Pure Mathematics, 39. Mathematical Society of Japan, Tokyo, 2004.

[10] Cox, J. T.; Gandolfi, A.; Griffin, P. S.; Kesten, H. Greedy lattice animals. I. Upper bounds. *Ann. Appl. Probab.* **3** (1993), no. 4, 1151–1169.

[11] Dembo, A.; Zeitouni, O. *Large deviations techniques and applications*. 2nd ed. Applications of Mathematics (New York), 38. Springer, New York, 1998.

[12] den Hollander, F. *Large deviations*. Fields Institute Monographs, 14. American Mathematical Society, Providence, R.I., 2000.

[13] den Hollander, F. *Random polymers: Lectures from the 37th Probability Summer School held in Saint-Flour, 2007*. Lecture Notes in Mathematics, 1974. Springer, Berlin, 2009.

[14] Deuschel, J.-D.; Stroock, D. W. *Large deviations*. Pure and Applied Mathematics, 137. Academic Press, Boston, 1989.

[15] Drewitz, A.; Gärtner, J.; Ramírez, A. F.; Sun, R. Survival probability of a random walk among a Poisson system of moving traps. Preprint, 2010. arXiv:1010.3958 [math.PR]

[16] Dunford, N.; Schwartz, J. T. *Linear operators. I. General theory*. Pure and Applied Mathematics, 7. Interscience, New York–London, 1958.

[17] Durrett, R. *Probability: theory and examples*. 2nd ed. Duxbury Press, Belmont, Calif., 1996.

[18] E, W.; Wehr, J.; Xin, J. Breakdown of homogenization for the random Hamilton-Jacobi equations. *Commun. Math. Sci.* **6** (2008), no. 1, 189–197.

[19] Gandolfi, A.; Kesten, H. Greedy lattice animals. II. Linear growth. *Ann. Appl. Probab.* **4** (1994), no. 1, 76–107.

[20] Georgii, H.-O. *Gibbs measures and phase transitions*. de Gruyter Studies in Mathematics, 9. Walter de Gruyter, Berlin, 1988.

[21] Giacomin, G. *Random polymer models*. Imperial College Press, London, 2007.

[22] Greven, A.; den Hollander, F. Large deviations for a random walk in random environment. *Ann. Probab.* **22** (1994), no. 3, 1381–1428.

[23] Havlin, S.; Ben-Avraham, D. Diffusion in disordered media. *Advances in Physics* **51** (2002), no. 1, 187–292.

[24] Huse, D. A.; Henley, C. L. Pinning and roughening of domain walls in Ising systems due to random impurities. *Phys. Rev. Lett.* **54** (1985), no. 25, 2708–2711.

[25] Imbrie, J. Z.; Spencer, T. Diffusion of directed polymers in a random environment. *J. Statist. Phys.* **52** (1988), no. 3-4, 609–626.

[26] Kassay, G. A simple proof for König's minimax theorem. *Acta Math. Hungar.* **63** (1994), no. 4, 371–374.

[27] Kosygina, E.; Rezakhanlou, F.; Varadhan, S. R. S. Stochastic homogenization of Hamilton-Jacobi-Bellman equations. *Comm. Pure Appl. Math.* **59** (2006), no. 10, 1489–1521.

[28] Kosygina, E.; Varadhan, S. R. S. Homogenization of Hamilton-Jacobi-Bellman equations with respect to time-space shifts in a stationary ergodic medium. *Comm. Pure Appl. Math.* **61** (2008), no. 6, 816–847.

[29] Lions, P.-L.; Souganidis, P. E. Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media. *Comm. Partial Differential Equations* **30** (2005), no. 1-3, 335–375.

[30] Martin, J. B. Linear growth for greedy lattice animals. *Stochastic Process. Appl.* **98** (2002), no. 1, 43–66.

[31] Rassoul-Agha, F.; Seppäläinen, T. *A course on large deviation theory with an introduction to Gibbs measures*. Preprint, 2010. Available at: `http://www.math.utah.edu/~firas/Papers/rassoul-seppalainen-ldp.pdf`

[32] Rassoul-Agha, F.; Seppäläinen, T. Process-level quenched large deviations for random walk in random environment. *Ann. Inst. H. Poincaré Probab. Statist.* **47** (2011), no. 1, 214–242.

[33] Rio, E. The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* **23** (1995), no. 3, 1188–1203.

[34] Rockafellar, R. T. *Convex analysis*. Princeton Mathematical Series, 28. Princeton University Press, Princeton, N.J., 1970.

[35] Rosenbluth, J. M. *Quenched large deviation for multidimensional random walk in random environment: A variational formula*. Doctoral dissertation, New York University, 2006. Available at: ProQuest LLC.

[36] Rubinstein, M.; Colby, R. H. *Polymer physics*. Oxford University Press, Oxford–New York, 2003.

[37] Rudin, W. *Functional analysis*. 2nd ed. International Series in Pure and Applied Mathematics. McGraw-Hill, New York, 1991.

[38] Schroeder, C. Green's functions for the Schrödinger operator with periodic potential. *J. Funct. Anal.* **77** (1988), no. 1, 60–87.

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

[40] Spitzer, F. *Principles of random walks*. 2nd ed. Graduate Texts in Mathematics, 34. Springer, New York, 1976.

[41] Sznitman, A.-S. Shape theorem, Lyapounov exponents, and large deviations for Brownian motion in a Poissonian potential. *Comm. Pure Appl. Math.* **47** (1994), no. 12, 1655–1688.

[42] Sznitman, A.-S. *Brownian motion, obstacles and random media*. Springer Monographs in Mathematics. Springer, Berlin, 1998.

[43] Sznitman, A.-S. Topics in random walks in random environment. *School and Conference on Probability Theory*, 203–266. International Centre for Theoretical Physics Lecture Notes, XVII. Abdus Salam International Centre for Theoretical Physics, Trieste, 2004.

[44] Temkin, D. E. One-dimensional random walks in a two-component chain. *Soviet Math. Dokl.* **13** (1972), 1172–1176.

[45] Varadhan, S. R. S. *Large deviations and applications*. CBMS-NSF Regional Conference Series in Applied Mathematics, 46. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1984.

[46] Varadhan, S. R. S. Large deviations for random walks in a random environment. *Comm. Pure Appl. Math.* **56** (2003), no. 8, 1222–1245.

[47] Vargas, V. Strong localization and macroscopic atoms for directed polymers. *Probab. Theory Related Fields* **138** (2007), no. 3-4, 391–410.

[48] Yilmaz, A. Large deviations for random walk in a space-time product environment. *Ann. Probab.* **37** (2009), no. 1, 189–205.

[49] Yilmaz, A. Quenched large deviations for random walk in a random environment. *Comm. Pure Appl. Math.* **62** (2009), no. 8, 1033–1075.

[50] Zeitouni, O. Random walks in random environment. *Lectures on probability theory and statistics*, 189–312. *Lecture Notes in Mathematics*, 1837. Springer, Berlin, 2004.

[51] Zerner, M. P. W. Directional decay of the Green's function for a random nonnegative potential on $\mathbf{Z}^d$. *Ann. Appl. Probab.* **8** (1998), no. 1, 246–280.

[52] Zerner, M. P. W. Lyapounov exponents and quenched large deviations for multidimensional random walk in random environment. *Ann. Probab.* **26** (1998), no. 4, 1446–1476.

FIRAS RASSOUL-AGHA
Department of Mathematics
University of Utah
155 South 1400 East
Salt Lake City, UT 84109
E-mail: firas@math.utah.edu

TIMO SEPPÄLÄINEN
Department of Mathematics
University of Wisconsin-Madison
419 Van Vleck Hall
Madison, WI 53706
E-mail: seppalai@math.wisc.edu

ATILLA YILMAZ
Department of Mathematics
Boğaziçi University
TR-34342 Bebek
Istanbul
TURKEY
E-mail: atilla.yilmaz@boun.edu.tr

Received April 2011.
Revised November 2011.