Zurich Open Repository and Archive
University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 1997

Localization transition for a polymer near an interface

Bolthausen, E ; den Hollander, F

DOI: https://doi.org/10.1214/aop/1024404516

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-22231
Journal Article
Published Version

Originally published at:
Bolthausen, E; den Hollander, F (1997). Localization transition for a polymer near an interface. The Annals of Probability, 25(3):1334-1366.
DOI: https://doi.org/10.1214/aop/1024404516

The Annals of Probability
1997, Vol. 25, No. 2, 531–572

ON SELF-ATTRACTING d-DIMENSIONAL RANDOM WALKS[1]

BY ERWIN BOLTHAUSEN AND UWE SCHMOCK

University of Zürich and ETH Zürich

Let $\{X_t\}_{t \geq 0}$ be a symmetric, nearest-neighbor random walk on $\mathbb{Z}^d$ with exponential holding times of expectation $1/d$, starting at the origin. For a potential $V\colon \mathbb{Z}^d \to [0, \infty)$ with finite and nonempty support, define transformed path measures by $d\hat{\mathbb{P}}_T \equiv \exp(T^{-1}\int_0^T\int_0^T V(X_s - X_t)\,ds\,dt)\,d\mathbb{P}/Z_T$ for $T > 0$, where Z_T is the normalizing constant. If $d = 1$ or if the self-attraction is sufficiently strong, then $\|X_t\|_\infty$ has an exponential moment under $\hat{\mathbb{P}}_T$ which is uniformly bounded for $T > 0$ and $t \in [0, T]$. We also prove that $\{X_t\}_{t \geq 0}$ under suitable subsequences of $\{\hat{\mathbb{P}}_T\}_{T > 0}$ behaves for large T asymptotically like a mixture of space-inhomogeneous ergodic random walks. For special cases like a sufficiently strong Dirac-type interaction, we even prove convergence of the transformed path measures and the law of X_T as well as of the law of the empirical measure L_T under $\{\hat{\mathbb{P}}_T\}_{T > 0}$.

1. Introduction. Let $\Omega \equiv D([0, \infty), \mathbb{Z}^d)$ be the set of right-continuous paths from $[0, \infty)$ to $\mathbb{Z}^d$ having left-hand limits. For every $t \geq 0$ let X_t with $X_t(\omega) \equiv \omega(t)$ for $\omega \in \Omega$ denote the evaluation map. The space Ω is equipped with the σ-algebra F generated by $\{X_t\}_{t \geq 0}$. Let $\mathbb{P}$ be the unique path measure on (Ω, F) such that $\{X_t\}_{t \geq 0}$ is a symmetric, nearest-neighbor random walk on $\mathbb{Z}^d$ with exponential holding times of expectation $1/d$, starting at the origin. For $t \geq 0$ the empirical distribution process $\{L_{t,T}\}_{T \geq t}$ after time t is defined by

$$\Omega \times (t, \infty) \ni (\omega, T) \mapsto L_{t,T}(\omega) \equiv \frac{1}{T - t}\int_{[t,T)} \delta_{X_s(\omega)}\,ds \in M_1(\mathbb{Z}^d) \tag{1.1}$$

and $L_{t,t}(\omega) = \delta_{X_t(\omega)}$, where $M_1(\mathbb{Z}^d)$ denotes the set of probability measures on the d-dimensional cubic lattice $\mathbb{Z}^d$. If $t = 0$, then we write L_T instead of $L_{0,T}$.

Let $V\colon \mathbb{Z}^d \to [0, \infty)$ be a function, which is not identically zero, such that the radius $R \equiv \sup\{\|x\|_1\colon x \in \mathbb{Z}^d, V(x) \neq 0\}$ of its support is finite. Define a "Hamiltonian" $H\colon M_1(\mathbb{Z}^d) \to [0, \infty)$ by

$$H(\mu) \equiv \sum_{x, y \in \mathbb{Z}^d} V(x - y)\mu(x)\mu(y), \qquad \mu \in M_1(\mathbb{Z}^d).$$

Received January 1995; revised June 1996.
[1]Research supported by the Swiss National Foundation Contract 21-298333.90.
Key words and phrases. d-dimensional random walk, attractive interaction, large deviations, weak convergence, maximum entropy principle, Dirac-type interaction.
AMS 1991 *subject classifications*. Primary 60F05; secondary 60F10, 60K35.

Without loss of generality we may and will assume in the following that V is a symmetric function in the sense that $V(x) = V(-x)$ for all $x \in \mathbb{Z}^d$. Note that

$$TH(L_T) = \frac{1}{T}\int_0^T\int_0^T V(X_s - X_t)\,ds\,dt, \qquad T > 0. \tag{1.2}$$

Our aim is to investigate the limiting behavior of the transformed probability measures

$$\hat{\mathbb{P}}_T(A) \equiv \mathbb{E}\left[1_A \exp(TH(L_T))\right]/Z_T, \qquad A \in F, T \geq 0, \tag{1.3}$$

as $T \to \infty$, where $Z_T \equiv \mathbb{E}[\exp(TH(L_T))]$ is the normalizing constant.

If the self-attraction is sufficiently strong, then it is intuitively clear that under $\hat{\mathbb{P}}_T$ the paths tend to clump together much more than for the free walk. An interesting consequence, which we derive near the end of Section 4 during the proof of our main result, is the following theorem; it excludes diffusive behavior.

THEOREM 1.4. *Assume that the self-attraction is sufficiently strong to satisfy Condition* 1.10 *below*, *which is certainly the case for* $d = 1$ *or for* $V = \beta 1_{\{0\}}$ *with* $\beta > d$. *Then there exists an* $\alpha_0 > 0$ *such that*

$$\sup_{T\geq 0}\ \sup_{t\in[0,T]} \hat{\mathbb{E}}_T\left[\exp(\alpha_0 \|X_t\|_\infty)\right] < \infty,$$

where $\hat{\mathbb{E}}_T$ *denotes expectation with respect to* $\hat{\mathbb{P}}_T$.

REMARK 1.5. Brydges and Slade [3], who work on the discrete-time random walk with the Dirac-type potential $V = \beta 1_{\{0\}}$, recently proved that for two and more dimensions and sufficiently small $\beta > 0$ the diffusive behavior persists in the sense that $\{\hat{\mathbb{P}}_T(X_T/\sqrt{T})^{-1}\}_{T\in\mathbb{N}}$ converges to a nontrivial distribution as T tends to infinity, which is Gaussian for $d \geq 3$. (They actually prove convergence of the rescaled process and describe the limiting process explicitly.) As far as we know, it is still an open problem whether the diffusive behavior persists for all coupling strengths $\beta > 0$ which are too small to satisfy our Condition 1.10.

Somewhat related models have been investigated recently in [1], [16] and [17]. The self-attraction in these models is different from the one of reinforced random walks (see, e.g., [5] and [13]), which have a more Markovian structure. Our model is more in the spirit of the widely studied self-repellent random walks which have a minus sign in the exponent (see, e.g., [10], Chapter 10.1 and the references given there), but no $1/T$ factor. It is easy to see that if this $1/T$ factor in (1.2) would be absent in our model, then the interaction would be too strong for an interesting result and, for example, for $V = 1_{\{0\}}$, the path measures $\{\hat{\mathbb{P}}_T\}_{T>0}$ would completely collapse to the Dirac measure on the zero function as $T \to \infty$.

The Donsker–Varadhan large deviation theory plays a crucial role in this paper. We define the rate function J: $M_1(\mathbb{Z}^d) \to (0, d]$ by

$$(1.6)\quad J(\mu) \equiv \frac{1}{2} \sum_{\substack{\{x,y\}\subset\mathbb{Z}^d\\ \|x-y\|_1=1}} \left(\sqrt{\mu(x)} - \sqrt{\mu(y)}\right)^2 = d - \sum_{\substack{\{x,y\}\subset\mathbb{Z}^d\\ \|x-y\|_1=1}} \sqrt{\mu(x)\mu(y)}$$

for all $\mu \in M_1(\mathbb{Z}^d)$. Using these two representations for J and Fatou's lemma, it follows that J is continuous in the weak topology on $M_1(\mathbb{Z}^d)$, which is metrizable by the total-variation distance $\|\mu - \nu\| \equiv \frac{1}{2}\sum_{x\in\mathbb{Z}^d}|\mu(x) - \nu(x)|$. The measures $\{\mathbb{P}L_T^{-1}\}_{T>0}$ satisfy a weak large deviation principle with rate function J. This means

$$(1.7)\qquad \limsup_{T\to\infty} \frac{1}{T} \log \mathbb{P}(L_T \in C) \le -\inf_{\mu\in C} J(\mu)$$

for every compact subset C of $M_1(\mathbb{Z}^d)$, and, for every open subset U of $M_1(\mathbb{Z}^d)$,

$$(1.8)\qquad \liminf_{T\to\infty} \frac{1}{T} \log \mathbb{P}(L_T \in U) \ge -\inf_{\mu\in U} J(\mu).$$

See [8], Theorem 8.1, for the lower bound (1.8), which we will use in the proof of Proposition 1.12, and [7], Theorem 5, for the identification of the rate function.

As an abbreviation, we define Λ: $M_1(\mathbb{Z}^d) \to \mathbb{R}$ by $\Lambda \equiv H - J$. Let

$$(1.9)\qquad b \equiv \sup_{\mu\in M_1(\mathbb{Z}^d)} \Lambda(\mu)$$

be the lowest upper bound for Λ. With these preparations, we can make precise what we consider as a sufficiently strong self-attraction. In addition to the properties of V stated above, we assume throughout this article (with the exception of Lemma 2.1, Lemma 2.2 and the last section) the following condition:

CONDITION 1.10. Let V be chosen such that $b > 0$.

Lemma 2.1 shows that this condition is always satisfied in one dimension. Furthermore, if there exists a $\mu \in M_1(\mathbb{Z}^d)$ with $H(\mu) > d$, then Condition 1.10 is satisfied because the rate function J is bounded by d and, therefore, $b \ge H(\mu) - J(\mu) > 0$.

Note that Λ is shift-invariant, which means that $\Lambda \circ \theta_x = \Lambda$ for all $x \in \mathbb{Z}^d$, where the shift transformation θ_x is defined by $\theta_x(\mu)(y) = \mu(y - x)$ for all $\mu \in M_1(\mathbb{Z}^d)$ and $y \in \mathbb{Z}^d$. Let

$$K \equiv \left\{\mu \in M_1(\mathbb{Z}^d) : \Lambda(\mu) = b\right\}$$

be the set of optimal measures where the supremum in (1.9) is attained. It is not immediately clear whether $K \neq \emptyset$. Furthermore, due to the shift-invariance of Λ, the set K is shift-invariant too and cannot be compact unless K is empty. Therefore, we introduce the subset

$$K(0) \equiv \{ \mu \in K: \mu(0) \geq H(\mu)/\|V\|_1 \},$$

where $\|V\|_1 \equiv \sum_{x \in \mathbb{Z}^d} V(x)$, of those optimal measures which have a considerable amount of their mass at the origin.

PROPOSITION 1.11. *The following statements hold*:

(a) *K is a nonvoid closed subset of* $M_1(\mathbb{Z}^d)$.
(b) *If $\mu \in K$, then $\mu(x) > 0$ for all $x \in \mathbb{Z}^d$.*
(c) *The set $K(0)$ is compact and $K = \{\theta_x(\mu): \mu \in K(0), x \in \mathbb{Z}^d\}$.*
(d) *If $\varepsilon > 0$, then* $\sup\{\Lambda(\mu): \mu \in M_1(\mathbb{Z}^d) \setminus U_\varepsilon(K)\} < b$, *where $U_\varepsilon(K)$ is the ε-neighborhood of K with respect to the total-variation distance.*

The following result concerning the asymptotic behavior of the partition function will be proved at the end of Section 2.

PROPOSITION 1.12. $\lim_{T \to \infty}(1/T)\log Z_T = b$.

In Lemma 4.4 we will show among other things that for every $\mu \in K$ there exists a unique family $\{\mathbb{Q}_x^\mu\}_{x \in \mathbb{Z}^d}$ of Markovian path measures on (Ω, F) with $\mathbb{Q}_x^\mu(X_0 = x) = 1$ for all $x \in \mathbb{Z}^d$, whose conservative infinitesimal generator $Q^\mu \equiv (q_{x,y}^\mu)_{x, y \in \mathbb{Z}^d}$ is determined by

$$q_{x,y}^\mu \equiv \begin{cases} \frac{1}{2}\sqrt{\mu(y)/\mu(x)}, & \text{if } \|x-y\|_1 = 1, \\ 0, & \text{if } \|x-y\|_1 > 1. \end{cases} \tag{1.13}$$

Note that $\{\mathbb{Q}_x^\mu\}_{x \in \mathbb{Z}^d}$ is reversible with respect to the measure μ.

Considering the equivalence relation on $M_1(\mathbb{Z}^d)$ given by the shift transformation, we denote by $[\mu] \equiv \{\theta_x(\mu): x \in \mathbb{Z}^d\}$ the equivalence class of $\mu \in M_1(\mathbb{Z}^d)$, by $\tilde{M}_1(\mathbb{Z}^d) \equiv \{[\mu]: \mu \in M_1(\mathbb{Z}^d)\}$ the set of all equivalence classes and, finally, by $\tilde{K} \equiv \{[\mu]: \mu \in K\}$ the optimal ones. We equip $\tilde{M}_1(\mathbb{Z}^d)$ with the metric

$$\|[\mu] - [\nu]\| \equiv \inf_{x \in \mathbb{Z}^d} \|\mu - \theta_x(\nu)\|, \qquad \mu, \nu \in M_1(\mathbb{Z}^d). \tag{1.14}$$

Note that the infimum is attained because $\|\mu - \theta_x(\nu)\| \to 1$ as $\|x\|_1 \to \infty$. Since the canonical projection $M_1(\mathbb{Z}^d) \ni \mu \mapsto [\mu]$ is continuous, Proposition 1.11(c) shows that $\tilde{K}$ is compact. This is the substitute for the missing compactness of K. For $\varrho \in \tilde{K}$ define $\zeta_\varrho = \sum_{\mu \in \varrho}\sqrt{\mu(0)}$. If $\mu \in K$, then $\zeta_{[\mu]} = \sum_{x \in \mathbb{Z}^d}\sqrt{\mu(x)}$, and we will show in Lemma 4.12 that $\mu(x)$ decays exponentially fast as $\|x\|_\infty$ tends to infinity, hence $\zeta_{[\mu]} < \infty$. For every $\mu \in K$ define $\tilde{\mu} \in M_1(\mathbb{Z}^d)$ by $\tilde{\mu}(x) = \sqrt{\mu(x)}/\zeta_{[\mu]}$ for all $x \in \mathbb{Z}^d$. Finally, let id_Ω denote the identity on Ω. The main result of this article is the following theorem.

THEOREM 1.15. *The set* $\{\hat{\mathbb{P}}_T[L_T]^{-1}\}_{T\geq 0}$ *is relatively compact in* $M_1(\tilde{M}_1(\mathbb{Z}^d))$ *with respect to the weak topology. Every accumulation point* Σ *of* $\{\hat{\mathbb{P}}_T[L_T]^{-1}\}_{T\geq 0}$ *as* $T\to\infty$ *is concentrated on* $\tilde{K}$. *If*

$$\lim_{k\to\infty}\hat{\mathbb{P}}_{T_k}\big[L_{T_k}\big]^{-1}=\Sigma \tag{1.16}$$

for a sequence $\{T_k\}_{k\in\mathbb{N}}$ *tending to infinity, then*

$$\lim_{k\to\infty}\hat{\mathbb{P}}_{T_k}\big(L_{T_k},\mathrm{id}_\Omega,X_{T_k}\big)^{-1}=\int_{\tilde{K}}\sum_{\mu\in\varrho}\tilde{\mu}(0)\,\delta_\mu\otimes\mathbb{Q}_0^\mu\otimes\tilde{\mu}\,\Sigma(d\varrho) \tag{1.17}$$

with respect to the weak topology on $M_1(M_1(\mathbb{Z}^d)\times\Omega\times\mathbb{Z}^d)$.

We always consider the path space Ω equipped with the standard Skorohod metric ([9], Chapter 3, (5.2)), which turns Ω into a Polish space ([9], Chapter 3, Theorem 5.6) with Borel σ-algebra F ([9], Chapter 3, Proposition 7.1).

If there exists only one accumulation point Σ in Theorem 1.15, then we obtain convergence in (1.17) for the full sequence and the right-hand side of (1.17) simplifies. A sufficient criterion for this to happen is $|\tilde{K}|=1$.

COROLLARY 1.18. *If* $\tilde{K}=\{\varrho\}$ *for some* $\varrho\in\tilde{M}_1(\mathbb{Z}^d)$, *then*

$$\lim_{T\to\infty}\hat{\mathbb{P}}_T(L_T,\mathrm{id}_\Omega,X_T)^{-1}=\sum_{\mu\in\varrho}\tilde{\mu}(0)\,\delta_\mu\otimes\mathbb{Q}_0^\mu\otimes\tilde{\mu}.$$

As an illustration of this corollary, consider $A\subset\mathbb{Z}^d$. Then, for every $\nu\in\varrho$,

$$\lim_{T\to\infty}\hat{\mathbb{P}}_T(X_T\in A)=\sum_{\mu\in\varrho}\tilde{\mu}(0)\tilde{\mu}(A)=\frac{1}{\zeta_\varrho^2}\sum_{x\in\mathbb{Z}^d}\sqrt{\nu(x)}\sum_{y\in A}\sqrt{\nu(x+y)}\,.$$

To decide whether $\tilde{K}$ contains just one element or not is quite delicate. For a Dirac-type interaction we will prove the following result:

THEOREM 1.19. *For* $\beta>0$ *define* $V=\beta 1_{\{0\}}$ *on* $\mathbb{Z}^d$. *If* $\beta\geq 2d$, *then* $|\tilde{K}|=1$.

REMARK 1.20. The corresponding variational problem for the one-dimensional Brownian motion is given by

$$\sup\left\{\beta\int_{\mathbb{R}}g^4(x)\,dx-\tfrac{1}{2}\int_{\mathbb{R}}|\nabla g|^2\,dx\colon g\in H^1(\mathbb{R}),\,\|g\|_{L^2}=1\right\}.$$

For every $\beta>0$ this variational expression has solutions which can easily be determined explicitly (see [11]). Uniqueness (up to translations) follows from a symmetrization argument. The delicacy of the variational problem on $\mathbb{Z}^d$ is that no symmetrization argument seems to be available. We successfully tried to lower the bound for β to the integer $2d$ (see Lemma 6.3) and with more work a small additional improvement would be possible, but our method does not allow us to reach zero. Indeed, numerical results for the one-dimensional case suggest that uniqueness does not hold for all $\beta\in(0,2)$.

In the remaining part of this introduction we briefly outline the method we adopt to prove our results. In Section 2 we first show that for every $\mu \in M_1(\mathbb{Z}^d)$ with a large value of the Hamiltonian there exists a point $x \in \mathbb{Z}^d$ where a considerable amount of the mass of μ is concentrated, in the sense that $\mu(x) \geq H(\mu)/\|V\|_1$. If μ is nearly optimal in the sense that $\Lambda(\mu)$ is close to the supremum b in (1.9), then we prove in Lemma 2.3 that most of the mass of μ is concentrated in the vicinity of the abovementioned x. It is crucial for our results that the size of this vicinity depends on μ only via the distance of $\Lambda(\mu)$ to the supremum b of Λ. This will imply that a sequence $\{\mu_k\}_{k \in \mathbb{N}}$ is tight if $\lim_{k \to \infty} \Lambda(\mu_k) = b$ and every μ_k has considerable mass at x. Using this observation, we can prove Proposition 1.11.

Since we do not have the large deviation upper bound (1.7) for all closed subsets of $M_1(\mathbb{Z}^d)$ and since $\tilde{M}_1(\mathbb{Z}^d)$ is not compact either, we project the random walk onto a large discrete torus $\mathbb{Z}_l^d \equiv \mathbb{Z}^d/l\mathbb{Z}^d$, where a full large deviation principle for the empirical measures is available. On $\mathbb{Z}_l^d$ we have the torus analogue Λ_l of Λ and the corresponding supremum b_l. For sufficiently large l the uniform distribution on $\mathbb{Z}_l^d$ cannot maximize Λ_l. Instead, the situation resembles the one in $\mathbb{Z}^d$; namely, for every $\mu \in M_1(\mathbb{Z}_l^d)$ with $\Lambda_l(\mu)$ close to b_l, there exists a "d-dimensional octahedron" in $\mathbb{Z}_l^d$, where most of the mass of μ is concentrated, and the size of this octahedron depends on μ only via $b_l - \Lambda_l(\mu)$. Using this observation, we can show that every optimal measure μ on $\mathbb{Z}^d$, when projected onto a large discrete torus $\mathbb{Z}_l^d$, looks very similar to an optimal measure on $\mathbb{Z}_l^d$. On the other hand, if μ is an optimal measure on a large torus, then we can find suitable seams to cut the torus apart such that, after identification with a cube in $\mathbb{Z}^d$, the trivially extended measure $\overline{\mu}$ on $\mathbb{Z}^d$ looks very similar to an optimal one on $\mathbb{Z}^d$. This turns Proposition 1.12 into an easy corollary.

In Section 3 we want to prove that $\limsup_{T \to \infty}(1/T)\log \hat{\mathbb{P}}_T(L_T \notin U_\varepsilon(K)) < 0$ for every $\varepsilon > 0$. The full large deviation principle for the torus immediately implies that the empirical measure L_T^l on $\mathbb{Z}_l^d$ has a high $\hat{\mathbb{P}}_T$-probability of being close to the projections of the optimal measures constituting K. Unfortunately, this does not imply that L_T has to be close to an optimal measure on $\mathbb{Z}^d$, because a priori the mass of L_T might be distributed among several, widely separated humps in $\mathbb{Z}^d$, and these humps might fall on top of each other when projected onto $\mathbb{Z}_l^d$. In the special case $d = 1$ and $|\tilde{K}| = 1$, the set K is a discrete line and we could visualize $U_\varepsilon(K)$ as a tube centered around K, which explains why Section 3 bears the title "The tube problem." To solve this problem, we devise a suitable way to fold the abovementioned annoying paths of the random walk such that the humps of the corresponding empirical measures cannot fall on top of each other during the projection. For this to work we have to keep the probabilistic "cost" of the folding operation small (with respect to $\hat{\mathbb{P}}_T$) and, on the other hand, have to shift a considerable part of the mass of L_T. We divide the troublesome paths into a T-dependent number of subsets, such that for each of these sets, we can find a slab of fixed width $3w$ which separates the mass of L_T and in which the corresponding paths spend only a small amount of time. Folding this slab to obtain a slab of width w yields the desired estimates.

In Section 4 we prove the tightness of $\{\hat{\mathbb{P}}_T L_T^{-1}\}_{T \geq 0}$. Note that tightness is not an immediate consequence of Section 3, because the shift-invariant set K of optimal measures is not compact. We start by considering, for every μ in $M_1(\mathbb{Z}^d)$, an affine function $\langle h^\mu, \cdot \rangle$: $M_1(\mathbb{Z}^d) \to \mathbb{R}$ which approximates H at μ. If $\Lambda(\mu) > 0$, then there exists a unique measure π^μ maximizing $\langle h^\mu, \cdot \rangle - J(\cdot)$ and, via a Feynman–Kac-like formula, we can define time-homogeneous Markovian probability measures $\{\mathbb{Q}_y^\mu\}_{y \in \mathbb{Z}^d}$ on (Ω, F) which are reversible with respect to π^μ. If μ is optimal, then $\pi^\mu = \mu$. If $\Lambda(\mu)$ is sufficiently close to b, then we can derive nontrivial upper and lower bounds for the exponential decay of π^μ, *uniformly* for all μ satisfying $\Lambda(\mu) > (1 - \varepsilon_1)b$ for a specific $\varepsilon_1 > 0$. Furthermore, *uniformly* for these μ, we obtain the convergence of $\mathbb{Q}_y^\mu X_t^{-1}$ to π^μ as $t \to \infty$, with an explicit dependence of the convergence rate on the starting point $y \in \mathbb{Z}^d$. Lemma 4.23, the main one in Section 4, then states the following: If L_T is in a neighborhood of an optimal measure μ, which has a considerable amount of its mass at x, then the $\hat{\mathbb{P}}_T$-probability for a corresponding path to be far away from x at a given time $t \in [0, T]$ is negligible, uniformly for large T. We prove this lemma by a "partial path exchange" argument. This means that we compare the paths with a far reaching excursion at time t to similar ones which hang around x during the period of this excursion. For the latter paths the value of $H(L_T)$ is considerably bigger, giving them a higher $\hat{\mathbb{P}}_T$-probability. Using Lemma 4.23, we can derive Theorem 1.4 and the tightness of $\{\hat{\mathbb{P}}_T L_T^{-1}\}_{T \geq 0}$.

We start the proof of our main theorem in Section 5 by showing that certain quantities and measures, like π^μ and $\mathbb{Q}_y^\mu$ for $y \in \mathbb{Z}^d$, depend continuously on μ as long as μ is nearly optimal in the sense that $\Lambda(\mu) > (1 - \varepsilon_1)b$. To prove weak convergence on $M_1(\Omega)$, it suffices to consider continuous functions f: $\Omega \to [0, 1]$, which depend only on a finite part of the paths, say $[0, s]$ for a given f. By the results of Sections 3 and 4, we can reduce our convergence problem from $\mathbb{Z}^d$ to various big cubes in $\mathbb{Z}^d$. It follows from the uniform convergence results in Section 4 that we can find a $t \geq 2s$ such that $\mathbb{Q}_y^\mu X_{t/2}^{-1}$ and $\mathbb{Q}_y^\mu X_t^{-1}$ are close to the corresponding equilibrium distribution π^μ, uniformly for all nearly optimal μ, which are essentially concentrated in one cube, and all starting points y in another, larger cube. The abovementioned continuity results allow us, for sufficiently large T, to replace L_T by $L_{t,T-t}$ defined as in (1.1). Furthermore, we may express the term $TH(L_T)$ by $(T - 2t)H(L_{t,T-t})$ plus an affine correction $2t\langle h^{L_{t,T-t}}, \cdot \rangle$. If L_T satisfies the recently stated condition for μ, then so does $L_{t,T-t}$. Under the measures $\mathbb{Q}_y^{L_{t,T-t}}$ the intervals $[0, t/2]$, $[t/2, t]$ and $[T - t, T]$ are long enough for convergence close to equilibrium; the dependence on the starting point y thereby fades away. This indicates why we obtain a product measure in (1.17).

A heuristic explanation why the normalized square root $\tilde{\mu}$ and not μ itself (as the stationary distribution of $\{\mathbb{Q}_x^\mu\}_{x \in \mathbb{Z}^d}$) determines the distribution of the final point X_T and the mixture within the equivalence class $[\mu] \in \tilde{K}$ might be the following: At time T the paths do not have to be prepared to build up empirical mass according to μ. Instead, under $\hat{\mathbb{P}}_T$, they behave like the free walk after T. Therefore, the distribution of X_T is more spread out than μ, which is one property of $\tilde{\mu}$. Similarly, there was no need to build up empirical mass according to μ before time zero. Technically speaking, $\tilde{\mu}$ is an eigenvector of the generator of the semigroup given by the Feynman–Kac formula (4.9), which includes the affine approximation $\langle h^{\mu}, \cdot \rangle$.

In the last section we prove our uniqueness result for a sufficiently strong Dirac-type interaction, namely, Theorem 1.19. The main work is to show that nearly all the mass of an optimal measure is concentrated at one point of $\mathbb{Z}^d$. To derive uniqueness from this fact, we basically use the concavity of the function $[0, \pi/2] \ni \alpha \mapsto H(\cos^2(\alpha)\delta_0 + \sin^2(\alpha)\delta_1)$ for small α. To reach the lower bound $2d$ for β, our method actually requires some numerical work.

2. Proofs of Propositions 1.11 and 1.12. We first show that the self-attraction is always strong enough in one dimension.

LEMMA 2.1. *If* $d \in \mathbb{N}$, *then* $b < \|V\|_\infty \equiv \max_{x \in \mathbb{Z}^d} V(x)$. *If* $d = 1$, *then* $b > 0$.

PROOF. The upper bound follows from $H \leq \|V\|_\infty$ and $J > 0$. Consider now the case $d = 1$. Then there exists $k \in \mathbb{N}_0$ with $V(k) > 0$. Choose $n \in \mathbb{N}$ even such that $(n - k)V(k) \geq 48$. Define $\mu(i) \equiv N^{-1}(\max\{0, 1 - |i|/n\})^2$ for all $i \in \mathbb{Z}$, where $N \equiv 1 + 2\sum_{i=1}^{n-1} i^2/n^2 \leq 1 + 2\sum_{i=1}^{n-1} i/n = n$. Then

$$H(\mu) \geq V(k) \sum_{i=-n/2}^{n/2-k} \mu(i + k)\mu(i) \geq V(k)(n - k)\left(\frac{1}{4N}\right)^2 \geq \frac{3}{nN}$$

and $J(\mu) = 2/(nN)$. Hence, $b \geq H(\mu) - J(\mu) \geq 1/(nN) > 0$. □

LEMMA 2.2. *For each* $\mu \in M_1(\mathbb{Z}^d)$ *there exists* $x \in \mathbb{Z}^d$ *with* $\mu(x) \geq H(\mu)/\|V\|_1$.

PROOF. Define $c = H(\mu)/\|V\|_1$. If $\mu(x) < c$ for all $x \in \mathbb{Z}^d$, then we obtain the contradiction $H(\mu) < c\sum_{y \in \mathbb{Z}^d} \mu(y) \sum_{x \in \mathbb{Z}^d} V(x - y) \leq c\|V\|_1 = H(\mu)$. □

The following technical-looking but important lemma will be used to show tightness in the proof of Proposition 1.11. It is also the main tool to prove the exponential decay of the stationary measures in Lemma 4.12. We use $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ to denote rounding to the next higher and lower integer, respectively.

LEMMA 2.3. *If* $\varepsilon \in [0, b/(2\|V\|_1 + 2b)]$ *and* $\mu \in M_1(\mathbb{Z}^d)$ *satisfy* $\Lambda(\mu) \geq (1 - \varepsilon)b$ *and if* $x \in \mathbb{Z}^d$ *satisfies* $\mu(x) \geq \Lambda(\mu)/\|V\|_1$, *then*

$$\mu(\{y \in \mathbb{Z}^d: \|x - y\|_1 > n\}) \leq \varepsilon + \frac{4\|V\|_\infty + 2d}{b}\sqrt{\frac{R + 1}{n}} \tag{2.4}$$

for all $n > n_0 \equiv \max\{R(R + 1), \lceil 4(R + 1)\|V\|_1^2(4\|V\|_\infty + 2d)^2/b^4 \rceil\}$.

PROOF. Replacing μ by $\theta_{-x}(\mu)$ if necessary, we may assume that $x = 0$. For every $j \in \mathbb{N}$ define $A_j \equiv \{y \in \mathbb{Z}^d\colon (j-1)r < \|y\|_1 \le jr\}$, where $r \equiv \max\{1, R\}$. Since the sets $\{A_j\}_{j\in\mathbb{N}}$ are disjoint, there exists $k \in \{1, \dots, \lfloor n/r \rfloor\}$ such that $\mu(A_k) \le 1/\lfloor n/r \rfloor$. Define $A = A_k$, $B = \{y \in \mathbb{Z}^d\colon \|y\|_1 \le kr\}$ and

$$\tilde{\mu}(y) = \begin{cases} \mu(y), & \text{for } y \in \mathbb{Z}^d \setminus (A \cup \{0\}), \\ 0, & \text{for } y \in A, \\ \mu(0) + \mu(A), & \text{for } y = 0. \end{cases} \tag{2.5}$$

Let $\lambda \equiv \tilde{\mu}(B^c)$. If $\lambda = 0$, then estimate (2.4) holds. Since $\tilde{\mu}(0) \ge \mu(0) > 0$, the case $\lambda = 1$ is excluded. Therefore, it remains to consider the case $\lambda \in (0, 1)$. If $\mu_0 \equiv \tilde{\mu}(\cdot|B)$ and $\mu_1 \equiv \tilde{\mu}(\cdot|B^c)$, then $\tilde{\mu} = (1-\lambda)\mu_0 + \lambda\mu_1$. Since $\tilde{\mu}(A) = 0$, it follows from the definition of r that $H(\tilde{\mu}) = (1-\lambda)^2 H(\mu_0) + \lambda^2 H(\mu_1)$ and $J(\tilde{\mu}) = (1-\lambda)J(\mu_0) + \lambda J(\mu_1) \ge (1-\lambda)^2 J(\mu_0) + \lambda^2 J(\mu_1)$. Hence,

$$\begin{aligned} \Lambda(\tilde{\mu}) &\le (1-\lambda)^2 \Lambda(\mu_0) + \lambda^2 \Lambda(\mu_1) \\ &\le \big((1-\lambda)^2 + \lambda^2\big) b \le b \max\{\lambda, 1-\lambda\}. \end{aligned} \tag{2.6}$$

In the remaining part of the proof we want to show that $\Lambda(\tilde{\mu})$ is close to $\Lambda(\mu)$ and thereby close to b. Since λ turns out to be substantially smaller than 1, this will imply that $1-\lambda$ is the maximum in (2.6), hence λ has to be small.

First note that $|H(\mu) - H(\tilde{\mu})| \le 4\|V\|_\infty \mu(A)$. It follows from the second representation in (1.6) that

$$J(\tilde{\mu}) - J(\mu) = \sum_{\substack{\{y,z\}\subset\mathbb{Z}^d \\ \{y,z\}\cap(A\cup\{0\})\neq\varnothing \\ \|y-z\|_1 = 1}} \left(\sqrt{\mu(y)\mu(z)} - \sqrt{\tilde{\mu}(y)\tilde{\mu}(z)}\right) \tag{2.7}$$

If $k = 1$, then A contains all neighbors of the origin and we may replace $A \cup \{0\}$ by A in (2.7). If $k \ge 2$, then $\mu(z) = \tilde{\mu}(z)$ for all $z \in \mathbb{Z}^d$ with $\|z\|_1 = 1$. Since $\mu(0) \le \tilde{\mu}(0)$, we obtain an upper estimate when we replace $A \cup \{0\}$ by A in (2.7). Dropping all remaining terms of the form $-\sqrt{\tilde{\mu}(y)\tilde{\mu}(z)}$ in (2.7), and adding $\sqrt{\mu(y)\mu(z)}$ for $\{y, z\} \subset A$ with $\|y - z\|_1 = 1$, it follows that

$$J(\tilde{\mu}) - J(\mu) \le \sum_{y\in A} \sum_{\substack{z\in\mathbb{Z}^d \\ \|y-z\|_1 = 1}} \sqrt{\mu(y)}\sqrt{\mu(z)}.$$

By the Cauchy–Schwarz inequality,

$$J(\tilde{\mu}) - J(\mu) \le \sqrt{2d\mu(A)}\sqrt{2d\mu(\mathbb{Z}^d)} = 2d\sqrt{\mu(A)}. \tag{2.8}$$

Hence, since $\mu(A) \le \sqrt{\mu(A)} \le \sqrt{1/\lfloor n/r \rfloor} \le \sqrt{(R+1)/n}$,

$$\Lambda(\tilde{\mu}) \ge \Lambda(\mu) - (4\|V\|_\infty + 2d)\sqrt{(R+1)/n}. \tag{2.9}$$

Note that $\lambda \le 1 - \tilde{\mu}(0) \le 1 - \mu(0) \le 1 - \Lambda(\mu)/\|V\|_1$. If $\max\{\lambda, 1-\lambda\} = \lambda$ in (2.6), then the last estimate, (2.6) and (2.9) together show that

$$\left(1 + \frac{b}{\|V\|_1}\right)\Lambda(\mu) \le b + (4\|V\|_\infty + 2d)\sqrt{\frac{R+1}{n}}. \tag{2.10}$$

By assumption,

$$\Lambda(\mu) \geq b(1-\varepsilon) \geq b\left(1 - \frac{b}{2\|V\|_1 + 2b}\right) = b\left(1 + \frac{b}{2\|V\|_1}\right) \Big/ \left(1 + \frac{b}{\|V\|_1}\right).$$

Together with (2.10) this leads to a contradiction for every $n > n_0$. Therefore, $\max\{\lambda, 1-\lambda\} = 1-\lambda$ and (2.4) follows from the estimates (2.6), (2.9) and $\Lambda(\mu) \geq (1-\varepsilon)b$. □

PROOF OF PROPOSITION 1.11. (a) Let $\{\mu_k\}_{k\in\mathbb{N}}$ be a sequence in $M_1(\mathbb{Z}^d)$ such that $\lim_{k\to\infty}\Lambda(\mu_k) = b$. By the shift-invariance of Λ and Lemma 2.2 we may assume that $\mu_k(0) \geq H(\mu_k)/\|V\|_1$ for all $k \in \mathbb{N}$. Then it follows from Lemma 2.3 that $\{\mu_k\}_{k\in\mathbb{N}}$ is tight. By Prohorov's theorem, we may assume that $\{\mu_k\}_{k\in\mathbb{N}}$ converges to some $\mu \in M_1(\mathbb{Z}^d)$. Since H and J are continuous, K is closed and $\Lambda(\mu) = b$, hence $\mu \in K$.

(b) Suppose that there exists $\mu \in K$ satisfying $\mu(x) = 0$ for some $x \in \mathbb{Z}^d$. Define the neighborhood of x by $N_x = \{y \in \mathbb{Z}^d: \|x-y\|_1 = 1\}$. Without loss of generality we may assume that $\mu(y) > 0$ for at least one $y \in N_x$. For $t \in [0,1]$ let $\mu_t \equiv (1-t^2)\mu + t^2\delta_x$, where $\delta_x \in M_1(\mathbb{Z}^d)$ satisfies $\delta_x(x) = 1$. Then

$$\frac{d}{dt}H(\mu_t)\bigg|_{t=0} = 0 \quad \text{and} \quad \frac{d}{dt}J(\mu_t)\bigg|_{t=0} = -\sum_{y\in N_x}\sqrt{\mu(y)} < 0,$$

which is a contradiction to $\mu \in K$.

(c) Using the proof of (a), compactness of $K(0)$ follows. The representation of K follows from Lemma 2.2.

(d) Assume that there exists a sequence $\{\mu_k\}_{k\in\mathbb{N}}$ in $M_1(\mathbb{Z}^d) \setminus U_\varepsilon(K)$ with $\lim_{k\to\infty}\Lambda(\mu_k) = b$. By the proof of (a) and the shift-invariance of the total-variation distance, we may then assume that $\{\mu_k\}_{k\in\mathbb{N}}$ converges to some $\mu \in K$, but this is a contradiction to the choice of $\{\mu_k\}_{k\in\mathbb{N}}$. □

To prepare the proof of Proposition 1.12 and the treatment of the tube problem in Section 3, we need to pass to a large discrete torus in order to have a full large deviation principle available. Furthermore, we have to study the connections between the optimal measures on $\mathbb{Z}^d$ and those on the discrete torus.

For $l \in \mathbb{N} \setminus \{1\}$ let $\mathbb{Z}_l^d \equiv \mathbb{Z}^d/l\mathbb{Z}^d$ be the discrete torus and let $\pi_l: \mathbb{Z}^d \to \mathbb{Z}_l^d$ be the canonical projection. Then $X_t^l \equiv \pi_l(X_t)$ for $t \geq 0$ is the ordinary symmetric random walk on $\mathbb{Z}_l^d$. Naturally, we equip the set $M_1(\mathbb{Z}_l^d)$ of probability measures on $\mathbb{Z}_l^d$ with the total-variation distance. Let $L_T^l \equiv L_T\pi_l^{-1}$ denote the empirical distribution of $\{X_t^l\}_{t\geq 0}$ up to time $T \geq 0$. It follows from [6], Theorem 4.2.58, that the measures $\{P(L_T^l)^{-1}\}_{T\geq 0}$ satisfy a full large deviation principle as $T \to \infty$ with the good rate function

$$(2.11)\quad J_l(\mu) \equiv \tfrac{1}{2}\sum_{x\in\mathbb{Z}_l^d}\sum_{i=1}^{d}\left(\sqrt{\mu(x)} - \sqrt{\mu(x+e_i)}\right)^2, \qquad \mu \in M_1(\mathbb{Z}_l^d),$$

where $e_i \equiv (0,\dots,0,1,0,\dots,0) \in \mathbb{Z}_l^d$ with the 1 at position i. We define the potential $V_l\colon \mathbb{Z}_l^d \to [0,\infty)$ on the discrete torus by

$$V_l(x) = \sum_{y \in \pi_l^{-1}(x)} V(y), \qquad x \in \mathbb{Z}_l^d,$$

and the corresponding Hamiltonian by

$$H_l(\mu) = \sum_{x,y \in \mathbb{Z}_l^d} V_l(x-y)\mu(x)\mu(y), \qquad \mu \in M_1(\mathbb{Z}_l^d).$$

Again, we will use the abbreviation $\Lambda_l \equiv H_l - J_l$. The large deviation principle implies that

$$\lim_{T\to\infty} \frac{1}{T} \log \mathbb{E}\left[\exp\left(TH_l\left(L_T^l\right)\right)\right] = b_l,$$

where $b_l \equiv \sup_{\mu \in M_1(\mathbb{Z}_l^d)} \Lambda_l(\mu)$. Since Λ_l is continuous and $M_1(\mathbb{Z}_l^d)$ is compact, it follows that $K_l \equiv \{\mu \in M_1(\mathbb{Z}_l^d)\colon \Lambda_l(\mu) = b_l\}$ is nonvoid and compact. If μ is in $M_1(\mathbb{Z}^d)$, then $\mu^l \equiv \mu\pi_l^{-1}$ is in $M_1(\mathbb{Z}_l^d)$,

$$\begin{aligned} H_l(\mu^l) &= \sum_{\substack{x,y \in \{1,\dots,l\}^d \\ r,s,t \in \mathbb{Z}^d}} V((x+lr)-(y+ls)+lt)\mu(x+lr)\mu(y+ls) \\ &\geq H(\mu) \end{aligned} \tag{2.12}$$

and

$$J_l(\mu^l) \leq J(\mu), \tag{2.13}$$

since

$$\left(\sqrt{u+v} - \sqrt{x+y}\right)^2 \leq \left(\sqrt{u} - \sqrt{x}\right)^2 + \left(\sqrt{v} - \sqrt{y}\right)^2$$

for all $u,v,x,y \in [0,\infty)$. These two estimates show that $b_l \geq b$.

If $\mu \in M_1(\mathbb{Z}_l^d)$ is the uniform distribution on $\mathbb{Z}_l^d$, then $H_l(\mu) = \|V\|_1/l^d$ and $J_l(\mu) = 0$, hence $b_l \geq \|V\|_1/l^d$ and $\mu \notin K_l$ for $l > (\|V\|_1/b)^{1/d}$. This already indicates that, for large l, the optimal measures are essentially concentrated on small regions. Using $b_l \geq b$ and the arguments which led to Lemma 2.3, one can indeed prove a corresponding result for measures on $\mathbb{Z}_l^d$.

LEMMA 2.14. *Let $l, n \in \mathbb{N}$ satisfy $l \geq 2n+1$ and $n > n_0$ with n_0 defined as in Lemma 2.3. If $\varepsilon \in [0, b/(2\|V\|_1 + 2b)]$ and $\mu \in M_1(\mathbb{Z}_l^d)$ satisfy $\Lambda_l(\mu) \geq (1-\varepsilon)b_l$, then there exists $x \in \mathbb{Z}^d$ such that*

$$\begin{aligned} &\mu\left(\mathbb{Z}_l^d \setminus \pi_l\left(\{y \in \mathbb{Z}^d\colon \|x-y\|_1 \leq n\}\right)\right) \\ &\qquad \leq \varepsilon + \frac{4\|V\|_\infty + 2d}{b}\sqrt{\frac{R+1}{n}}. \end{aligned} \tag{2.15}$$

For $\mu \in M_1(\mathbb{Z}_l^d)$ and $x = (x_1,\dots,x_d) \in \mathbb{Z}^d$ we define $\bar\mu_x \in M_1(\mathbb{Z}^d)$ by

$$\bar\mu_x(y) = \begin{cases} \mu(\pi_l(y)), & \text{if } y \in \mathbb{Z}^d \cap \prod_{i=1}^d [x_i - \lfloor l/2 \rfloor, x_i + \lfloor (l-1)/2 \rfloor], \\ 0, & \text{otherwise.} \end{cases}$$

There exists a minimal $z \in \{1, \dots, l\}^d$ with respect to the lexicographic order such that $\Lambda(\overline{\mu}_z) = \max_{x \in \mathbb{Z}^d} \Lambda(\overline{\mu}_x)$. Let $\overline{\mu}$ denote this $\overline{\mu}_z$. Note that $\overline{\mu}^l = \mu$ for every $\mu \in M_1(\mathbb{Z}_l^d)$. The next lemma establishes the close relationship between the nearly optimal measures on $\mathbb{Z}^d$ and those on $\mathbb{Z}_l^d$ for large l. For $\varepsilon > 0$ and $l \in \mathbb{N} \setminus \{1\}$ define $U_{\varepsilon, l} = U_\varepsilon(\{\mu^l\}_{\mu \in K})$.

LEMMA 2.16. *Let* $\varepsilon_0 \equiv \min\{b/(2\|V\|_1 + 2b), 1/(2\|V\|_\infty)\}$. *For every* ε *in* $(0, \varepsilon_0]$ *there exists* $l_0 \in \mathbb{N} \setminus \{1\}$ *such that for every* $l \geq l_0$ *the following statements hold*:

(a) *If* $\mu \in M_1(\mathbb{Z}_l^d)$ *satisfies* $\Lambda_l(\mu) \geq (1 - \varepsilon) b_l$, *then*

$$\Lambda(\overline{\mu}) \geq \Lambda_l(\mu) - 2(d+1)\varepsilon;$$

(b) $b_l \leq b + 2(d+1)\varepsilon$;

(c) $\sup\{\Lambda_l(\mu): \mu \in M_1(\mathbb{Z}_l^d) \setminus U_{\varepsilon, l}\} < b$;

(d) $K_l \subset U_{\varepsilon, l}$.

PROOF. (a) Take any $n > n_0$, with n_0 as in Lemma 2.3, such that the right-hand side of (2.15) is less than 2ε. Let $l \geq 2n + 2R + 3$ be given. Take any $\mu \in M_1(\mathbb{Z}_l^d)$ satisfying $\Lambda_l(\mu) \geq (1 - \varepsilon) b_l$. By Lemma 2.14 there exists $x \in \mathbb{Z}^d$ such that $\mu(A) \leq 2\varepsilon$ with $A \equiv \mathbb{Z}_l^d \setminus \pi_l(\{y \in \mathbb{Z}^d: \|x - y\|_1 \leq n\})$. Define $A' = \mathbb{Z}_l^d \setminus \pi_l(\{y \in \mathbb{Z}^d: \|x - y\|_1 \leq n + 1\})$. Note that $\|V_l\|_\infty = \|V\|_\infty$. Since $\varepsilon \leq 1/(2\|V\|_\infty)$, it follows that

$$H_l(\mu) - H(\overline{\mu}_x) \leq \sum_{y, z \in A} V_l(y - z)\mu(y)\mu(z) \leq \|V_l\|_\infty (\mu(A))^2 \leq 2\varepsilon.$$

Using (1.6), (2.11) and the Cauchy–Schwarz inequality, it follows that

$$J(\overline{\mu}_x) - J_l(\mu) \leq \sum_{y \in A'} \sum_{i=1}^{d} \sqrt{\mu(y)\mu(y + e_i)} \leq \sqrt{d\mu(A')}\sqrt{d\mu(A)} \leq 2d\varepsilon.$$

(b) Apply part (a) to any $\mu \in K_l$.

(c) Assume that (c) does not hold. Then there exist a strictly increasing sequence $\{l_k\}_{k \in \mathbb{N}}$ in $\mathbb{N} \setminus \{1\}$ and a sequence $\{\mu^{(k)}\}_{k \in \mathbb{N}}$ such that every $\mu^{(k)}$ is in the compact set $M_1(\mathbb{Z}_{l_k}^d) \setminus U_{\varepsilon, l_k}$ and satisfies $\Lambda_{l_k}(\mu^{(k)}) \geq b$. Remember that $b_l \geq b$ for all $l \in \mathbb{N} \setminus \{1\}$. Using part (b) and choosing a subsequence if necessary, we may assume that $(1 - \varepsilon_0/k) b_{l_k} \leq b$ for all $k \in \mathbb{N}$. It follows from part (a) that

$$\Lambda(\overline{\mu}^{(k)}) \geq \left(1 - \frac{\varepsilon_0}{k}\right) b_{l_k} - \frac{2(d+1)\varepsilon_0}{k} \geq \left(1 - \frac{b + 2d + 2}{bk}\varepsilon_0\right) b \tag{2.17}$$

for all $k \in \mathbb{N}$. By Lemma 2.2 there exists, for every $k \in \mathbb{N}$, a point $x_k \in \mathbb{Z}^d$ with $\overline{\mu}^{(k)}(x_k) \geq \Lambda(\overline{\mu}^{(k)})/\|V\|_1$. Since every single measure is tight and since Lemma 2.3 applies to $\overline{\mu}^{(k)}$ for every $k \geq (b + 2d + 2)/b$, it follows that $\{\theta_{-x_k} \overline{\mu}^{(k)}\}_{k \in \mathbb{N}}$ is a tight subset of $M_1(\mathbb{Z}^d)$. Using Prohorov's theorem and choosing a sub-subsequence if necessary, we may assume that $\{\theta_{-x_k} \overline{\mu}^{(k)}\}_{k \in \mathbb{N}}$ converges to some $\mu \in M_1(\mathbb{Z}^d)$. Since Λ is continuous, it follows from (2.17) that $\Lambda(\mu) \geq b$, hence $\mu \in K$ and $\theta_{x_k}\mu \in K$ for all $k \in \mathbb{N}$. Furthermore,

$$\left\| \mu^{(k)} - \left(\theta_{x_k}\mu\right)^{l_k} \right\| \leq \| \overline{\mu}^{(k)} - \theta_{x_k}\mu \| = \| \theta_{-x_k}\overline{\mu}^{(k)} - \mu \| \to 0 \quad \text{as } k \to \infty.$$

Thus, $\mu^{(k)} \in U_{\varepsilon, l_k}$ for all sufficiently large $k \in \mathbb{N}$. This is a contradiction to the choice of the sequence $\{\mu^{(k)}\}_{k \in \mathbb{N}}$. Hence, there exists $l_0 \geq 2n + 2R + 3$ such that (c) holds.

(d) If part (c) holds for $l \in \mathbb{N} \setminus \{1\}$ and if $\mu \in K_l$, then $\Lambda_l(\mu) = b_l \geq b$, hence $\mu \notin \mathcal{M}_1(\mathbb{Z}_l^d) \setminus U_{\varepsilon, l}$ by part (c). □

PROOF OF PROPOSITION 1.12. The lower bound in Proposition 1.12 follows from (1.8) as in the proof of [6], Lemma 2.17. To show the upper bound, choose any $\varepsilon > 0$. By (2.12) and the full large deviation principle for $\{\mathbb{P}(L_T^l)^{-1}\}_{T>0}$,

$$\limsup_{T\to\infty} \frac{1}{T} \log Z_T \leq \lim_{T\to\infty} \frac{1}{T} \log \mathbb{E}\left[\exp\left(TH_l\left(L_T^l\right)\right)\right] = b_l \leq b + 2(d+1)\varepsilon,$$

where the last inequality follows from (Lemma 2.16(b) for all large l. □

3. The tube problem. As explained in the introduction, the tube problem is to show that L_T stays in $\hat{\mathbb{P}}_T$-law inside a "tube" around K as $T \to \infty$, that is, to prove the following proposition:

PROPOSITION 3.1. *For any $\varepsilon > 0$,*

$$\limsup_{T\to\infty} \frac{1}{T} \log \hat{\mathbb{P}}_T\left(L_T \notin U_\varepsilon(K)\right) < 0.$$

The proof of this proposition together with a corollary concerning the existence of an uniformly bounded exponential moment is given at the end of this section. The difficulty in proving Proposition 3.1 is coming from the fact that we have only a weak large deviation principle for $\{\mathbb{P}L_T^{-1}\}_{T>0}$ at our disposal. Also, the monotonicity argument based on (2.12), which we used in the above proof of Proposition 1.12, does not work here.

LEMMA 3.2. *If $\varepsilon > 0$, then there exists $l_0 \in \mathbb{N}$ such that*

$$\sup_{l \geq l_0} \limsup_{T\to\infty} \frac{1}{T} \log \hat{\mathbb{P}}_T\left(L_T^l \notin U_{\varepsilon, l}\right) < 0.$$

PROOF. The large deviation principle for $\{\mathbb{P}(L_T^l)^{-1}\}_{T>0}$ and [6], Exercise 2.1.24, show that

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{E}\left[\exp\left(TH_l\left(L_T^l\right)\right); L_T^l \notin U_{\varepsilon, l}\right] \leq \sup\{\Lambda_l(\mu): \mu \notin U_{\varepsilon, l}\}.$$

By Lemma 2.16(c) there exists $l_0 \in \mathbb{N}$ with $\sup\{\Lambda_l(\mu): \mu \notin U_{\varepsilon, l}\} < b$ for all $l \geq l_0$. This together with Proposition 1.12 proves the claim. □

Already this seems to be very close to Proposition 3.1, except for one very annoying point. We know by Lemma 2.14 that the elements in K_l are essentially concentrated on small sets, namely, d-dimensional octahedrons, uniformly in l. Therefore, Lemma 3.2 says that L_T is essentially concentrated on the union of the l-translates of such a small set. The delicacy is to exclude the possibility that L_T has substantial mass on more than one of these translated sets.

To explain the key idea for the solution of this problem, we need to introduce some additional notation. Given a coordinate direction $\kappa \in \{1, \dots, d\}$ and an integer $i \in \mathbb{Z}$, let $h_{i,\kappa} \equiv \{(x_1, \dots, x_d) \in \mathbb{Z}^d \colon x_\kappa = i\}$ denote the corresponding (discrete) hyperplane. Such a hyperplane divides $\mathbb{Z}^d$ into two half-spaces given by $h^+_{i,\kappa} \equiv \{(x_1, \dots, x_d) \in \mathbb{Z}^d \colon |x_\kappa - 1/2| \le |2i - x_\kappa - 1/2|\}$ and $h^-_{i,\kappa} \equiv \mathbb{Z}^d \setminus h^+_{i,\kappa}$. Note that $h^+_{i,\kappa}$ always contains the origin, the d unit vectors $e_j \equiv (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{Z}^d$, where the 1 is at place $j \in \{1, \dots, d\}$, and the hyperplane $h_{i,\kappa}$ itself (this justifies the plus sign). We have to use $1/2$ instead of 0 in the definition of $h^+_{i,\kappa}$ to handle the case $i = 0$ conveniently. For $i \in \mathbb{Z}$ and $w \in \mathbb{N}_0$ define

$$i_w = \begin{cases} i + w, & \text{if } i \ge 1, \\ i - w, & \text{if } i \le 0, \end{cases}$$

and let $s_{i,\kappa,w} \equiv (h^+_{i_w,\kappa} \cap h^-_{i,\kappa}) \cup h_{i,\kappa}$ be the slab of width w.

The main idea to prove Proposition 3.1 is the following. If the empirical measure L_T has substantial mass in more than one of the translated d-dimensional octahedrons, then we choose a slab $s_{i,\kappa,3w}$, which is visited seldom, such that L_T has substantial mass in the half-spaces $h^-_{i_{3w},\kappa}$ and $h^+_{i,\kappa} \setminus h_{i,\kappa}$. With the help of two reflections at the hyperplanes $h_{i_w,\kappa}$ and $h_{i_{2w},\kappa}$ we fold up the path inside the slab $s_{i,\kappa,3w}$ such that it fits into the slab $s_{i,\kappa,w}$. The empirical distribution of the new path, when projected to $\mathbb{Z}^d_l$, turns out not to be essentially concentrated on one "d-dimensional octahedron" of $\mathbb{Z}^d_l$. Hence Lemma 3.2 applies. Of course, we have to show that the probabilistic "cost" of these two reflections is less than the "cost" for L^l_T to substantially deviate from K_l.

In one dimension we could prove Proposition 3.1 using one reflection, because $V(x) = V(-x)$ for all $x \in \mathbb{Z}$. In higher dimensions we need the more complicated construction with two reflections to prove Lemma 3.6(c) below. This is due to the fact that, for a general potential, $V(x_1, \dots, x_d) \ne V(x_1, \dots, x_{\kappa-1}, -x_\kappa, x_{\kappa+1}, \dots, x_d)$.

For $\kappa \in \{1, \dots, d\}$ and $i \in \mathbb{Z}$ let $\sigma_{i,\kappa,0} \equiv 0$. Define the arrival and departure times of the random walk for the hyperplane $h_{i,\kappa}$, recursively for every $k \in \mathbb{N}_0$, by $\tau_{i,\kappa,k} = \inf\{t > \sigma_{i,\kappa,k} \colon X_t \in h_{i,\kappa}\}$ and $\sigma_{i,\kappa,k+1} = \inf\{t > \tau_{i,\kappa,k} \colon X_t \notin h_{i,\kappa}\}$. For $T > 0$ let $n_{i,\kappa,T} \equiv \max\{k \in \mathbb{N}_0 \colon \sigma_{i,\kappa,k} < T\}$ be the number of excursions starting from the hyperplane $h_{i,\kappa}$ before T. Let $\bar\tau_{i,\kappa,T} \equiv \tau_{i,\kappa,n_{i,\kappa,T}}$ be the end of the last excursion which started from $h_{i,\kappa}$ before T. Furthermore, we denote by $M_T \equiv \max_{s,t \in [0,T]} \|X_s - X_t\|_\infty$ the maximal spread up to time T. The following lemma contains some useful large deviation estimates.

LEMMA 3.3. (a) *If* $\varepsilon > 0$, $i \in \mathbb{Z}$, $\kappa \in \{1, \dots, d\}$, $T > 0$ *and* $\delta \le \varepsilon \exp(-1 - (\|V\|_\infty + 1)/\varepsilon)$, *then* $\hat{\mathbb{P}}_T(L_T(h_{i,\kappa}) \le \delta,\, n_{i,\kappa,T} > \varepsilon T) \le \exp(-T)$.

(b) *If* $T \ge \exp(\|V\|_\infty + 2)$, *then* $\hat{\mathbb{P}}_T(M_T \ge T^2) \le d\exp(-T^2)$.

(c) *If* $\alpha, T \ge 0$, *then* $\hat{\mathbb{E}}_T[\exp(\alpha M_T);\, M_T \ge T\exp(\|V\|_\infty + 2 + \alpha)] \le d$.

PROOF. (a) Since the κ-component of $\{X_t\}_{t\ge 0}$ is recurrent, $\mathbb{P}(\tau_{i,\kappa,k} < \infty) = 1$ for all $k \in \mathbb{N}_0$. The random variables $\xi_k \equiv \inf\{t > 0\colon X_{t+\tau_{i,\kappa,k}} \notin h_{i,\kappa}\}$ for $k \in \mathbb{N}_0$ are independent and exponentially distributed with expectation 1. They describe the duration of the visits of $\{X_t\}_{t\ge 0}$ at the hyperplane $h_{i,\kappa}$. By the derivation of the one-dimensional version of Cramér's theorem ([6], Section 1.2),

$$\mathbb{P}(L_T(h_{i,\kappa}) \le \delta,\, n_{i,\kappa,T} > \varepsilon T) \le \mathbb{P}\left(\sum_{k=0}^{\lfloor \varepsilon T\rfloor} \xi_k \le \delta T\right) \le \exp(-mh(\delta T/m))$$

with $m \equiv \lfloor \varepsilon T \rfloor + 1 \ge \delta T$, where the rate function is given by

$$h(x) \equiv \sup_{\lambda\in\mathbb{R}}\left(\lambda x - \log\int_0^\infty e^{\lambda y}e^{-y}\,dy\right) = \begin{cases}\infty, & \text{for } x \le 0,\\ x - 1 - \log x, & \text{for } x > 0.\end{cases}$$

Note that the rate function h is decreasing within the interval [0, 1]. Since $\varepsilon T \le m$ and $-h(x) \le 1 + \log x$ for $x > 0$, it follows that

$$-mh(\delta T/m) \le -\varepsilon Th(\delta T/m) \le \varepsilon T(1 + \log \delta/\varepsilon) \le -(\|V\|_\infty + 1)T.$$

Since $0 \le H(\mu) \le \|V\|_\infty$ for all $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$, it follows that $\exp(TH(L_T)) \le \exp(T\|V\|_\infty)$ and $\mathbb{E}[\exp(TH(L_T))] \ge 1$. Using (1.3), part (a) follows.

(b) For every $\kappa \in \{1, \dots, d\}$ let $\{\xi_{\kappa,k}\}_{k\in\mathbb{N}}$ be the times between successive jumps of the random walk in coordinate direction κ. These times are independent and exponentially distributed with expectation 1. If $M_T \ge T^2$, then there is a direction κ in which the random walk jumped at least $\lceil T^2 \rceil$ times. Hence

$$\mathbb{P}(M_T \ge T^2) \le \sum_{\kappa=1}^{d} \mathbb{P}\left(\sum_{k=1}^{\lceil T^2\rceil} \xi_{\kappa,k} \le T\right) \le d\exp(-T^2 h(1/T)).$$

Using the inequalities $-h(1/T) \le 1 + \log(1/T) \le -\|V\|_\infty - 1$ and $T \le T^2$, it follows that $\mathbb{P}(M_T \ge T^2) \le d\exp(-\|V\|_\infty T - T^2)$. Similarly as in the proof of part (a), part (b) follows from the last estimate.

(c) If $T > 0$ and $M_T = m$ with $m \ge T\exp(\|V\|_\infty + 2 + \alpha)$, then there exists a coordinate direction in which the random walk $\{X_t\}_{t\ge 0}$ jumped at least m times during $[0, T]$. Similarly as in part (b), it follows that

$$\mathbb{P}(M_T = m) \le d\exp(-mh(T/m)) \le d\exp(-m(\|V\|_\infty + 1 + \alpha))$$

and

$$\hat{\mathbb{P}}_T(M_T = m) \le de^{-m-\alpha m}.$$

Therefore,

$$\hat{\mathbb{E}}_T[\exp(\alpha M_T);\, M_T \ge T\exp(\|V\|_\infty + 2 + \alpha)] \le d\textstyle\sum_{m\in\mathbb{N}} e^{-m} \le d. \qquad \square$$

For $i \in \mathbb{Z}$ and $\kappa \in \{1, \dots, d\}$ let

$$\varrho_{i,\kappa}(x) \equiv \begin{cases} x, & \text{for } x \in h^+_{i,\kappa}, \\ x - 2(x_\kappa - i)e_\kappa, & \text{for } x = (x_1, \dots, x_d) \in h^-_{i,\kappa}, \end{cases}$$

be the map which reflects the half-space $h^-_{i,\kappa}$ into $h^+_{i,\kappa}$, and, for $w \in \mathbb{N}$, let

$$\varphi_{i,\kappa,w}(x) \equiv \begin{cases} \varrho_{i_w,\kappa}(x), & \text{for } x \in h^+_{i_{2w},\kappa}, \\ x - (i_{2w} - i)e_\kappa, & \text{for } x \in h^-_{i_{2w},\kappa}, \end{cases} \tag{3.4}$$

be the map which folds up the slab $s_{i,\kappa,3w}$. The second case in (3.4) corresponds to reflecting $\varrho_{i_w,\kappa}(x)$ from $h^+_{i,\kappa} \setminus h_{i,\kappa}$ back into $h^-_{i,\kappa}$. For $T > 0$ define the accompanying map $\overline{\varphi}_{i,\kappa,w,T}\colon \Omega \to \Omega$, which folds up the paths, by

$$\overline{\varphi}_{i,\kappa,w,T}(\omega)(t) = \begin{cases} \varphi_{i,\kappa,w}(\omega(t)), & \text{for } t < \overline{\tau}_{i_w,\kappa,T} \wedge \overline{\tau}_{i_{2w},\kappa,T}, \\ \varrho_{i_w,\kappa}(\omega(t)), & \text{for } \overline{\tau}_{i_w,\kappa,T} \wedge \overline{\tau}_{i_{2w},\kappa,T} \le t < \overline{\tau}_{i_w,\kappa,T}, \\ \omega(t), & \text{for } t \ge \overline{\tau}_{i_w,\kappa,T}. \end{cases}$$

LEMMA 3.5. *If $i \in \mathbb{Z}$, $\kappa \in \{1, \dots, d\}$, $T > 0$ and $w \in \mathbb{N}$, then $\mathbb{P}\overline{\varphi}^{-1}_{i,\kappa,w,T} \ll \mathbb{P}$ and $d\mathbb{P}\,\overline{\varphi}^{-1}_{i,\kappa,w,T}/d\mathbb{P} \le 2^{n_{i,\kappa,T} + n_{i_w,\kappa,T}}$.*

PROOF. Since the random walk is symmetric, it suffices to give, for every $\omega \in \Omega$, a crude upper estimate for the number of paths $\tilde{\omega} \in \Omega$ with $\overline{\varphi}_{i,\kappa,w,T}(\tilde{\omega}) = \omega$. If the path ω leaves one of the hyperplanes $h_{i,\kappa}$ and $h_{i_w,\kappa}$ before T, then $\tilde{\omega}$ may have gone (if at all possible) into the other direction with respect to the κ-coordinate. Since ω leaves the hyperplanes $h_{i,\kappa}$ and $h_{i_w,\kappa}$ before T exactly $n_{i,\kappa,T} + n_{i_w,\kappa,T}$ times, we get the claimed estimate. □

For every integer $i \in \mathbb{Z}$, coordinate direction $\kappa \in \{1, \dots, d\}$ and width $w \in \mathbb{N}$, we define a "folding operator" $\hat{\varphi}_{i,\kappa,w}\colon\ M_1(\mathbb{Z}^d) \to M_1(\mathbb{Z}^d)$, which corresponds to $\varphi_{i,\kappa,w}$ given in (3.4), by $\hat{\varphi}_{i,\kappa,w}(\mu) = \mu\varphi^{-1}_{i,\kappa,w}$ for all $\mu \in M_1(\mathbb{Z}^d)$. Note that $\hat{\varphi}_{i,\kappa,w} \circ L_T = L_T \circ \overline{\varphi}_{i,\kappa,w,T}$ for every $T > 0$.

LEMMA 3.6. *Let ε_0 be defined as in Lemma 2.16. For every choice of $\varepsilon > 0$ and $\delta \in (0, \min\{\varepsilon_0, \varepsilon/(45d)\}]$ there exist $l_1 \in \mathbb{N}$ and a width $w \in \mathbb{N}$ such that, for every $l \ge l_1$ and every $\mu \in M_1(\mathbb{Z}^d) \setminus U_\varepsilon(K)$ satisfying $\mu^l \in U_{\delta,l}$, there exist an integer $i \in \mathbb{Z}$ and a coordinate direction $\kappa \in \{1, \dots, d\}$ such that*:

(a) $\mu(s_{i,\kappa,3w}) \le 2\delta$,
(b) $(\hat{\varphi}_{i,\kappa,w}(\mu))^l \notin U_{\varepsilon/(7d),l}$,
(c) $|H(\hat{\varphi}_{i,\kappa,w}(\mu)) - H(\mu)| \le 4\|V\|_\infty \delta^2$.

PROOF. Define n_0 as in Lemma 2.3. Choose an integer $n > n_0$ such that the second term on the right-hand side of (2.4) is less than δ. Since $\delta \le \varepsilon_0$, Lemma 2.16(d) applies. Hence there exists $l_1 \ge 8(2n + 1)$ such that $K_l \subset U_{\delta,l}$

for all $l \geq l_1$. Define $w = 2n + 1$. Take $l \geq l_1$ and $\mu \in M_1(\mathbb{Z}^d) \setminus U_\varepsilon(K)$ with $\mu^l \in U_{\delta, l}$. Since $U_{\delta, l} \equiv U_\delta(\{\nu^l\}_{\nu \in K})$, there exists $\nu \in K$ with $\| \mu^l - \nu^l \| \leq \delta$. By Lemma 2.3 there exists a d-dimensional octahedron $O \equiv \{y \in \mathbb{Z}^d: \|x - y\|_1 \leq n\}$, centered at some $x \in \mathbb{Z}^d$, such that $\nu(O) \geq 1 - \delta$. For every $z \in \mathbb{Z}^d$ define the shifted octahedron $O_z = \{y + lz: y \in O\}$. Let $C \equiv \bigcup_{z \in \mathbb{Z}^d} O_z$ be the collection of all these octahedrons. Since $C = \pi_l^{-1}(\pi_l(O))$ and $\| \mu^l - \nu^l \| \leq \delta$, it follows that

$$\begin{aligned} \mu(C) &= \mu^l(\pi_l(O)) \geq \nu^l(\pi_l(O)) - \delta \\ &= \nu(C) - \delta \geq \nu(O) - \delta \geq 1 - 2\delta. \end{aligned} \tag{3.7}$$

Assume that there exists $z \in \mathbb{Z}^d$ with $\mu(O_z) \geq 1 - 4\varepsilon/9$. Let $\bar{\mu} \equiv \mu(\cdot | O_z)$ and $\tilde{\nu} \equiv \nu(\cdot | O)$. Then $\| \mu^l - \tilde{\mu}^l \| \leq \| \mu - \tilde{\mu} \| \leq 4\varepsilon/9$ and $\| \nu^l - \tilde{\nu}^l \| \leq \| \nu - \tilde{\nu} \| \leq \delta$. Furthermore, $\| \tilde{\mu} - \theta_{lz} \tilde{\nu} \| = \| \tilde{\mu}^l - (\theta_{lz} \tilde{\nu})^l \| = \| \tilde{\mu}^l - \tilde{\nu}^l \| \leq \| \tilde{\mu}^l - \mu^l \| + \| \mu^l - \nu^l \| + \| \nu^l - \tilde{\nu}^l \| \leq 4\varepsilon/9 + 2\delta$. Hence

$$\| \mu - \theta_{lz} \nu \| \leq \| \mu - \tilde{\mu} \| + \| \tilde{\mu} - \theta_{lz} \tilde{\nu} \| + \| \theta_{lz} \tilde{\nu} - \theta_{lz} \nu \| \leq 8\varepsilon/9 + 3\delta < \varepsilon.$$

This is a contradiction to $\theta_{lz} \nu \in K$ and $\mu \notin U_\varepsilon(K)$.

For every $j \in \mathbb{Z}$ and $\kappa \in \{1, \dots, d\}$ define

$$A_{j, \kappa} = \bigcup_{\substack{(z_1, \dots, z_d) \in \mathbb{Z}^d \\ z_\kappa \leq j}} O_{(z_1, \dots, z_d)}.$$

Assume that

$$\mu(A_{j, \kappa}) \leq \frac{\varepsilon}{5d} \quad \text{or} \quad \mu(A_{j, \kappa}) \geq \mu(C) - \frac{\varepsilon}{5d}$$

for all $j \in \mathbb{Z}$ and $\kappa \in \{1, \dots, d\}$. For every $\kappa \in \{1, \dots, d\}$ we know that $A_{j, \kappa} \downarrow \emptyset$ as $j \to -\infty$ and $A_{j, \kappa} \uparrow C$ as $j \to \infty$. Hence, there exists $j_\kappa \in \mathbb{Z}$ such that $A_\kappa \equiv A_{j_\kappa, \kappa} \setminus A_{j_\kappa - 1, \kappa}$ satisfies $\mu(A_\kappa) \geq \mu(C) - 2\varepsilon/(5d)$. Using (3.7), it follows that $\mu(A_\kappa) \geq 1 - 2\delta - 2\varepsilon/(5d) \geq 1 - 4\varepsilon/(9d)$. If we define $z = (j_1, \dots, j_d)$, then $O_z = \bigcap_{\kappa=1}^d A_\kappa$ and $\mu(O_z) \geq 1 - 4\varepsilon/9$. According to the previous paragraph, this leads to a contradiction. Hence, there exist $j \in \mathbb{Z}$ and $\kappa \in \{1, \dots, d\}$ such that

$$\frac{\varepsilon}{5d} < \mu(A_{j, \kappa}) < \mu(C) - \frac{\varepsilon}{5d}.$$

With $x = (x_1, \dots, x_d)$ as above, define $i = jl + x_\kappa + 4w$. Then $s_{i, \kappa, 3w} \cap C = \emptyset$, hence (a) follows from (3.7). Define $A = \varphi_{i, \kappa, w}(A_{j, \kappa})$ and $B = \varphi_{i, \kappa, w}(C \setminus A_{j, \kappa})$. Then $\hat{\varphi}_{i, \kappa, w}(\mu)(A) > \varepsilon/(5d)$ and $\hat{\varphi}_{i, \kappa, w}(\mu)(B) > \varepsilon/(5d)$. Furthermore, since $\varphi_{i, \kappa, w}$ shifts either $A_{j, \kappa}$ or $C \setminus A_{j, \kappa}$ by $2w$ in the κ-direction and since the $\| \cdot \|_1$-diameter of O equals $w - 1$, it follows that, for every d-dimensional octahedron $O' \equiv \{y \in \mathbb{Z}^d: \|y - z\|_1 \leq n\}$ with $z \in \mathbb{Z}^d$, the intersection of $\pi_l^{-1}(\pi_l(O'))$ with either A or B is empty, hence

$$\hat{\varphi}_{i, \kappa, w}(\mu)\left(\pi_l^{-1}(\pi_l(O'))\right) < 1 - \frac{\varepsilon}{5d}. \tag{3.8}$$

Assume that $(\hat{\varphi}_{i, \kappa, w}(\mu))^l$ is in $U_{\varepsilon/(7d), l}$. Then there exists a measure $\hat{\mu} \in K$ with $\|(\hat{\varphi}_{i, \kappa, w}(\mu))^l - \hat{\mu}^l\| < \varepsilon/(7d)$ and, by Lemma 2.3, there exists $\hat{z} \in \mathbb{Z}^d$

such that $O' \equiv \{y \in \mathbb{Z}^d: \|y - \hat{z}\|_1 \le n\}$ satisfies $\hat{\mu}(O') \ge 1 - \delta$. Hence,

$$\hat{\varphi}_{i,\kappa,w}(\mu)\big(\pi_l^{-1}(\pi_l(O'))\big) > \hat{\mu}(O') - \frac{\varepsilon}{7d} \ge 1 - \delta - \frac{\varepsilon}{7d} > 1 - \frac{\varepsilon}{5d},$$

which is a contradiction to (3.8). This proves part (b).

Let $D \equiv \{(y_1, \dots, y_d) \in \mathbb{Z}^d: i_{-R} \le y_\kappa \le i_{R+3w}\}$ be the region where the terms of H may be distorted. Since $n > n_0 \ge R$ by the definition of n_0 in Lemma 2.3, it follows that $D \cap C = \varnothing$, hence $\mu(D) \le 2\delta$ by (3.7). Therefore,

$$\begin{aligned}
&\big|H(\hat{\varphi}_{i,\kappa,w}(\mu)) - H(\mu)\big| \\
&\quad = \bigg|\sum_{y,z\in D}\big(V(\varphi_{i,\kappa,w}(y) - \varphi_{i,\kappa,w}(z)) - V(y-z)\big)\mu(y)\mu(z)\bigg| \\
&\quad \le \|V\|_\infty(\mu(D))^2 \le 4\|V\|_\infty\delta^2,
\end{aligned}$$

because V is nonnegative. This proves part (c). □

PROOF OF PROPOSITION 3.1. Choose any $\varepsilon > 0$. According to Lemma 3.2 there exists $l_0 \in \mathbb{N}$ such that

$$\gamma \equiv -\sup_{l \ge l_0}\limsup_{T\to\infty}\frac{1}{T}\log\hat{\mathbb{P}}_T\big(L_T^l \notin U_{\varepsilon/(7d),l}\big) > 0. \tag{3.9}$$

Let ε_0 be given as in Lemma 2.16, define

$$\delta = \min\left\{\varepsilon_0, \frac{\varepsilon}{45d}, \sqrt{\frac{\gamma}{12\|V\|_\infty}}, \frac{\gamma}{24}\exp\left(-1 - 12\frac{\|V\|_\infty + 1}{\gamma}\right)\right\}$$

and let the corresponding $l_1 \in \mathbb{N}$ and $w \in \mathbb{N}$ be determined by Lemma 3.6. According to Lemma 3.2 there exists $l_2 \in \mathbb{N}$ such that

$$\sup_{l\ge l_2}\limsup_{T\to\infty}\frac{1}{T}\log\hat{\mathbb{P}}_T\big(L_T^l \notin U_{\delta,l}\big) < 0. \tag{3.10}$$

Let $l \equiv \max\{l_0, l_1, l_2\}$ and $T > 0$. Then

$$\hat{\mathbb{P}}_T(L_T \notin U_\varepsilon(K)) \le \hat{\mathbb{P}}_T\big(L_T^l U_{\delta,l}\big) + \hat{\mathbb{P}}_T\big(L_T \notin U_\varepsilon(K), L_T^l \in U_{\delta,l}\big).$$

The first probability on the right-hand side is estimated by (3.10).

To estimate the second probability, define $f_{i,\kappa}(\mu) = |H(\hat{\varphi}_{i,\kappa,w}(\mu)) - H(\mu)|$ for every $i \in \mathbb{Z}$, $\kappa \in \{1, \dots, d\}$ and $\mu \in M_1(\mathbb{Z}^d)$. Using Lemma 3.6, it follows that

$$\begin{aligned}
&\hat{\mathbb{P}}_T\big(L_T \notin U_\varepsilon(K), L_T^l \in U_{\delta,l}\big) \\
&\quad \ge \hat{\mathbb{P}}_T\big(M_T \ge T^2\big) + \sum_{\kappa=1}^{d}\sum_{i=-\lfloor T^2\rfloor}^{\lfloor T^2\rfloor}\hat{\mathbb{P}}_T\Big(\hat{\varphi}_{i,\kappa,w}(L_T)^l \notin U_{\varepsilon/(7d),l}, \\
&\qquad\qquad L_T(s_{i,\kappa,3w}) \le 2\delta, f_{i,\kappa}(L_T) \le 4\|V\|_\infty\delta^2\Big).
\end{aligned}$$

The term $\hat{\mathbb{P}}_T(M_T \geq T^2)$ is estimated in Lemma 3.3(b). Due to the choice of δ, it follows from Lemma 3.3(a) that

$$\hat{\mathbb{P}}_T\left(L_T(s_{i,\kappa,3w}) \leq 2\delta,\ \max_{k\in\{0,1,2,3\}} n_{i_{kw},\kappa,T} \geq \frac{\gamma T}{12}\right) \geq \sum_{k\in\{0,1,2,3\}} \hat{\mathbb{P}}_T\big(L_T(h_{i_{kw},\kappa}) \leq 2\delta,\, n_{i_{kw},\kappa,T} \geq \gamma T/12\big) \leq 4\exp(-T)$$

for all $i \in \mathbb{Z}$ and $\kappa \in \{1,\dots,d\}$. Since $\lim_{T\to\infty}(1/T)\log(2dT^2+d) = 0$, it therefore suffices to show that

$$\lim_{T\to\infty} \sup_{\substack{i\in\mathbb{Z}\\ \kappa\in\{1,\dots,d\}}} \frac{1}{T}\log\hat{\mathbb{P}}_T\left(\hat{\varphi}_{i,\kappa,w}(L_T)^l \notin U_{\varepsilon/(7d),l},\, f_{i,\kappa}(L_T) \leq \frac{\gamma}{3},\ \max_{k\in\{0,1,2,3\}} n_{i_{kw},\kappa,T} \leq \frac{\gamma T}{12}\right) \leq -\frac{\gamma}{3}. \tag{3.11}$$

Since $\hat{\varphi}_{i,\kappa,w} \circ L_T = L_T \circ \overline{\varphi}_{i,\kappa,w,T}$ as well as $n_{i,\kappa,T} \circ \overline{\varphi}_{i,\kappa,w,T} = n_{i,\kappa,T} + n_{i_{2w},\kappa,T}$ and $n_{i_w,\kappa,T} \circ \overline{\varphi}_{i,\kappa,w,T} = n_{i_w,\kappa,T} + n_{i_{3w},\kappa,T}$, it follows from (1.3) and Lemma 3.5 that, for all $i \in \mathbb{Z}$, $\kappa \in \{1,\dots,d\}$ and $T > 0$,

$$\begin{aligned}
&\hat{\mathbb{P}}_T\left(\hat{\varphi}_{i,\kappa,w}(L_T)^l \notin U_{\varepsilon/(7d),l},\, f_{i,\kappa}(L_T) \leq \frac{\gamma}{3},\ \max_{k\in\{0,1,2,3\}} n_{i_{kw},\kappa,T} \leq \frac{\gamma T}{12}\right)\\
&\quad\leq \frac{e^{\gamma T/3}}{Z_T}\mathbb{E}\Big[\exp\big(TH(L_T\circ\overline{\varphi}_{i,\kappa,w,T})\big);(L_T\circ\overline{\varphi}_{i,\kappa,w,T})^l \notin U_{\varepsilon/(7d),l},\\
&\qquad\qquad (n_{i,\kappa,T}+n_{i_w,\kappa,T})\circ\overline{\varphi}_{i,\kappa,w,T} \leq \gamma T/3\Big]\\
&\quad\leq e^{2\gamma T/3}\hat{\mathbb{P}}_T\big(L_T^l \notin U_{\varepsilon/(7d),l}\big).
\end{aligned}$$

Using (3.9), the estimate (3.11) follows. $\square$

COROLLARY 3.12. *For $T \geq 0$ let $M_T \equiv \max_{s,t\in[0,T]}\|X_s - X_t\|_\infty$ denote the maximal spread up to time T. Given $\varepsilon > 0$, there exists $\tilde{\alpha} > 0$ such that*

$$\sup_{T\geq 0}\hat{\mathbb{E}}_T\big[\exp(\tilde{\alpha}M_T);\, L_T \notin U_\varepsilon(K)\big] < \infty. \tag{3.13}$$

PROOF. Proposition 3.1 implies the existence of $\tilde{\alpha} \in (0,1]$ and $c > 0$ such that $\hat{\mathbb{P}}_T(L_T \notin U_\varepsilon(K)) \leq c\exp(-\tilde{\alpha}T\exp(\|V\|_\infty + 3))$ for all $T \geq 0$. Hence

$$\hat{\mathbb{E}}_T\big[\exp(\tilde{\alpha}M_T);\, M_T \leq T\exp(\|V\|_\infty+3),\, L_T \notin U_\varepsilon(K)\big] \leq c.$$

Using Lemma 3.3(c), the corollary follows. $\square$

4. Tightness. In this section we prove the tightness of $\{\hat{\mathbb{P}}_T L_T^{-1}\}_{T\geq 0}$ and also Theorem 1.4. For $x \in \mathbb{Z}^d$ let $\mathbb{P}_x$ denote the path measure on (Ω, F) of a symmetric, nearest-neighbor random walk on $\mathbb{Z}^d$ with exponential holding times of expectation $1/d$, starting at x. Let $\mathbb{E}_x$ denote the corresponding expectation. Note that $\mathbb{P}_0 = \mathbb{P}$ and $\mathbb{E}_0 = \mathbb{E}$. For every $\mu \in M_1(\mathbb{Z}^d)$ define $H'(\mu)\colon \mathbb{Z}^d \to [0,\infty)$ and $h^\mu\colon \mathbb{Z}^d \to \mathbb{R}$ by

$$H'(\mu)(x) = 2 \sum_{y \in \mathbb{Z}^d} V(x-y)\mu(y), \qquad x \in \mathbb{Z}^d, \tag{4.1}$$

and

$$h^\mu(x) = H'(\mu)(x) - H(\mu), \qquad x \in \mathbb{Z}^d. \tag{4.2}$$

For a bounded function $f\colon \mathbb{Z}^d \to \mathbb{R}$ let $\langle f, \mu \rangle \equiv \sum_{x \in \mathbb{Z}^d} f(x)\mu(x)$. Furthermore, define

$$\lambda^\mu = \sup_{\nu \in M_1(\mathbb{Z}^d)} \left(\langle h^\mu, \nu \rangle - J(\nu) \right) \tag{4.3}$$

and note that $2\|V\|_\infty - H(\mu) \geq \lambda^\mu \geq \langle h^\mu, \mu \rangle - J(\mu) = H(\mu) - J(\mu) = \Lambda(\mu)$. We start with a lemma which shows that, in particular for all nearly optimal measures μ, there exists a unique solution π^μ of the linearized variational expression in (4.3). This π^μ is the stationary distribution of a certain ergodic random walk.

LEMMA 4.4. *Let $\mu \in M_1(\mathbb{Z}^d)$ with $\lambda^\mu > 0$ be given.*

(a) *For every $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that*

$$\nu\left(\{x \in \mathbb{Z}^d\colon \|x\|_\infty > 2n\}\right) \leq 4\varepsilon \tag{4.5}$$

for all $\nu \in M_1(\mathbb{Z}^d)$ satisfying $\langle h^\mu, \nu \rangle - J(\nu) \geq (1-\varepsilon)\lambda^\mu$.

(b) *There exists a unique $\pi^\mu \in M_1(\mathbb{Z}^d)$ with $\lambda^\mu = \langle h^\mu, \pi^\mu \rangle - J(\pi^\mu)$. It satisfies $\pi^\mu(x) > 0$ for all $x \in \mathbb{Z}^d$.*

(c) *If $x \in \mathbb{Z}^d$, then $\pi^{\theta_x(\mu)} = \theta_x(\pi^\mu)$.*

(d) *There exists a set $\{\mathbb{Q}_x^\mu\}_{x \in \mathbb{Z}^d}$ of time-homogeneous Markovian probability measures on (Ω, F) such that, for every $x \in \mathbb{Z}^d$, $t \geq 0$ and $A \in F_t$,*

$$\mathbb{Q}_x^\mu(A) = \frac{\exp(-\lambda^\mu t)}{\sqrt{\pi^\mu(x)}} \mathbb{E}_x\left[1_A \exp(t\langle h^\mu, L_t \rangle)\sqrt{\pi^\mu(X_t)}\right]. \tag{4.6}$$

(e) *The conservative generator $Q^\mu \equiv (q_{x,y}^\mu)_{x,y \in \mathbb{Z}^d}$ corresponding to $\{\mathbb{Q}_x^\mu\}_{x \in \mathbb{Z}^d}$ is determined by*

$$q_{x,y}^\mu = \begin{cases} \frac{1}{2}\sqrt{\pi^\mu(y)/\pi^\mu(x)}, & \text{if } \|x-y\|_1 = 1, \\ 0, & \text{if } \|x-y\|_1 > 1. \end{cases}$$

(f) *The measure π^μ is the reversible distribution of $\{\mathbb{Q}_x^\mu\}_{x \in \mathbb{Z}^d}$.*

(g) *If $\mu \in K$, then $\lambda^\mu = b$ and $\pi^\mu = \mu$.*

PROOF. (a) It suffices to consider the case $\varepsilon < \frac{1}{4}$. Define $\delta = \varepsilon\lambda^\mu$. Using the finite support of V, the tightness of μ and (4.1), it follows that there exists $n \in \mathbb{N}$ with $n \geq (2d/\delta)^2$ such that $H'(\mu)(x) \leq \delta$ for all $x \in \mathbb{Z}^d$ with $\|x\|_\infty > n$.

For every $j \in \mathbb{N}$ define $A_j = \{x \in \mathbb{Z}^d: \|x\|_\infty = j\}$. Since the sets $\{A_j\}_{j\in\mathbb{N}}$ are disjoint, there exists $k \in \{n+1, \dots, 2n\}$ such that $\nu(A_k) \le 1/n$. Define $A = A_k$ and $B = \{x \in \mathbb{Z}^d: \|x\|_\infty \le k\}$. Use (2.5) with ν instead of μ to define the measure $\tilde{\nu}$. Note that $\langle h^\mu, \tilde{\nu}\rangle \ge \langle h^\mu, \nu\rangle - \delta$. The arguments leading to (2.8) show that $J(\tilde{\nu}) - J(\nu) \le 2d\sqrt{\nu(A)} \le \delta$. Hence

$$\lambda^\mu - 3\delta \le \langle h^\mu, \nu\rangle - J(\nu) - 2\delta \le \langle h^\mu, \tilde{\nu}\rangle - J(\tilde{\nu}). \tag{4.7}$$

Define $\gamma = \tilde{\nu}(B^c)$. If $\gamma = 0$, then we are done. If $\gamma = 1$, then $\nu(B^c) = 1$ and $(1-\varepsilon)\lambda^\mu \le \langle h^\mu, \nu\rangle - J(\nu) < \delta = \varepsilon\lambda^\mu$, which is a contradiction to $\varepsilon < \frac{1}{4}$. It remains to consider $\gamma \in (0,1)$. Defining $\nu_0 = \tilde{\nu}(\cdot|B)$ and $\nu_1 = \tilde{\nu}(\cdot|B^c)$, it follows that $J(\tilde{\nu}) = (1-\gamma)J(\nu_0) + \gamma J(\nu_1) \ge (1-\gamma)J(\nu_0)$. Since $\langle h^\mu, \cdot\,\rangle$ is linear and $\langle h^\mu, \nu_1\rangle \le \delta$, it follows that $\langle h^\mu, \tilde{\nu}\rangle - J(\tilde{\nu}) \le (1-\gamma)\lambda^\mu + \delta$. Using (4.7) and solving for γ, we obtain $\gamma \le 4\delta/\lambda^\mu = 4\varepsilon$.

(b) To prove the existence of π^μ, let $\{\nu_k\}_{k\in\mathbb{N}}$ be a sequence in $M_1(\mathbb{Z}^d)$ with $\lim_{k\to\infty}(\langle h^\mu, \nu_k\rangle - J(\nu_k)) = \lambda^\mu$. It follows from part (a) that $\{\nu_k\}_{k\in\mathbb{N}}$ is tight. Hence we may assume that the sequence converges to some $\pi^\mu \in M_1(\mathbb{Z}^d)$. Since J is continuous, $\langle h^\mu, \pi^\mu\rangle - J(\pi^\mu) = \lambda^\mu$. By the same method as in the proof of Proposition 1.11(b) it follows that $\pi^\mu(x) > 0$ for all $x \in \mathbb{Z}^d$.

To prove the uniqueness of π^μ, it suffices to show that J is strictly mid-convex on the set of all $\nu \in M_1(\mathbb{Z}^d)$ satisfying $\nu(x) > 0$ for all $x \in \mathbb{Z}^d$. Let ν and $\tilde{\nu}$ be such measures, define $\hat{\nu} = \frac{1}{2}\nu + \frac{1}{2}\tilde{\nu}$ and assume that $J(\tilde{\nu}) = \frac{1}{2}J(\nu) + \frac{1}{2}J(\tilde{\nu})$. Using the second expression for J in (1.6), a short computation shows that this is equivalent to $(\sqrt{\nu(x)\tilde{\nu}(y)} - \sqrt{\nu(y)\tilde{\nu}(x)}\,)^2 = 0$, hence $\nu(x)/\nu(y) = \tilde{\nu}(x)/\tilde{\nu}(y)$, for all $x, y \in \mathbb{Z}^d$ with $\|x-y\|_1 = 1$. Therefore, $\nu = \tilde{\nu}$.

(c) Since $\langle h^{\theta_x(\mu)}, \theta_x(\nu)\rangle = \langle h^\mu, \nu\rangle$ for all $\nu \in M_1(\mathbb{Z}^d)$, it follows from (4.3) that $\lambda^{\theta_x(\mu)} = \lambda^\mu$. Furthermore, $\langle h^{\theta_x(\mu)}, \theta_x(\pi^\mu)\rangle - J(\theta_x(\pi^\mu)) = \langle h^\mu, \pi^\mu\rangle - J(\pi^\mu) = \lambda^\mu = \lambda^{\theta_x(\mu)}$. Hence $\pi^{\theta_x(\mu)} = \theta_x(\pi^\mu)$ by the uniqueness from part (b).

(d) Since $\pi^\mu(x) > 0$ by part (b), we can define

$$\lambda_x^\mu = h^\mu(x) - d + \frac{1}{2\sqrt{\pi^\mu(x)}} \sum_{\substack{y\in\mathbb{Z}^d \\ \|x-y\|_1 = 1}} \sqrt{\pi^\mu(y)}\,, \qquad x \in \mathbb{Z}^d. \tag{4.8}$$

Take $x \in \mathbb{Z}^d \setminus \{0\}$. For t in the interval $(-\pi^\mu(0), \pi^\mu(x))$ define $\nu_{x,t} \in M_1(\mathbb{Z}^d)$ by $\nu_{x,t} = \pi^\mu + t\delta_0 - t\delta_x$. By part (b), π^μ maximizes the variational expression in (4.3). Hence $0 = (d/dt)(\langle h^\mu, \nu_{x,t}\rangle - J(\nu_{x,t}))|_{t=0} = \lambda_0^\mu - \lambda_x^\mu$. By (4.8), (1.6) and part (b), $\sum_{x\in\mathbb{Z}^d}\lambda_x^\mu\,\pi^\mu(x) = \langle h^\mu, \pi^\mu\rangle - J(\pi^\mu) = \lambda^\mu$, hence $\lambda^\mu = \lambda_x^\mu$ for all $x \in \mathbb{Z}^d$.

Define a semigroup of transition kernels $\{P_t^\mu\}_{t\ge 0}$ by

$$P_t^\mu(x, A) = \mathbb{E}_x\big[\exp(t\langle h^\mu, L_t\rangle)1_A(X_t)\big], \qquad x \in \mathbb{Z}^d,\ A \subset \mathbb{Z}^d,\ t \ge 0. \tag{4.9}$$

The corresponding operator semigroup on the space of bounded functions on $\mathbb{Z}^d$ is denoted by $\{P_t^\mu\}_{t\ge 0}$, too. Let $L^\mu = (L_{x,y}^\mu)_{x,y\in\mathbb{Z}^d}$ be the generator of this operator semigroup. Its components are given by

$$(4.10)\quad L^{\mu}_{x,y} = \lim_{t\downarrow 0} \frac{P^{\mu}_t(x,\{y\}) - \delta_{xy}}{t} = \begin{cases} 1/2, & \text{for } \|x-y\|_1 = 1,\\ h^{\mu}(x) - d, & \text{for } x = y,\\ 0, & \text{for } \|x-y\|_1 \geq 2. \end{cases}$$

Furthermore, $\|L^{\mu}\|_{\text{op}} \leq 2\|V\|_{\infty} + 2d$.

If $\sqrt{\pi^{\mu}}$ denotes the vector $(\sqrt{\pi^{\mu}(x)})_{x\in\mathbb{Z}^d}$, then it follows with the help of (4.8) that $L^{\mu}\sqrt{\pi^{\mu}} = \lambda^{\mu}\sqrt{\pi^{\mu}}$, hence $P^{\mu}_t\sqrt{\pi^{\mu}} = \exp(\lambda^{\mu}t)\sqrt{\pi^{\mu}}$. Therefore, (4.6) defines a probability measure on (Ω, F_t) for every $t \geq 0$ and $x \in \mathbb{Z}^d$. Furthermore, for every $x \in \mathbb{Z}^d$, these measures are consistent. Hence, they can be uniquely extended to measures $\{\mathbb{Q}^{\mu}_x\}_{x\in\mathbb{Z}^d}$ on (Ω, F) ([14], Chapter V, Theorem 4.2). The other properties of $\{\mathbb{Q}^{\mu}_x\}_{x\in\mathbb{Z}^d}$ follow from those of $\{\mathbb{P}_x\}_{x\in\mathbb{Z}^d}$.

(e) The generator follows from (4.6) using (4.8) and (4.10).

(f) Check that $\pi^{\mu}(x)q^{\mu}_{x,y} = \pi^{\mu}(y)q^{\mu}_{y,x}$ for all $x, y \in \mathbb{Z}^d$.

(g) Remember that $L^{\mu}\sqrt{\pi^{\mu}} = \lambda^{\mu}\sqrt{\pi^{\mu}}$ by the proof of part (d). Due to Proposition 1.11(b), the proof of (d) also works when λ^{μ} from (4.3) and π^{μ} are replaced by b and μ, provided that the variational expression in (1.9) and $\Lambda(\nu_{x,t})$ are used instead of the one in (4.3) and $\langle h^{\mu}, \nu_{x,t}\rangle - J(\nu_{x,t})$. Hence, $L^{\mu}\sqrt{\mu} = b\sqrt{\mu}$. Since the matrix L^{μ} is symmetric, $b\langle\sqrt{\mu}, \sqrt{\pi^{\mu}}\rangle = \langle L^{\mu}\sqrt{\mu}, \sqrt{\pi^{\mu}}\rangle = \langle\sqrt{\mu}, L^{\mu}\sqrt{\pi^{\mu}}\rangle = \lambda^{\mu}\langle\sqrt{\mu}, \sqrt{\pi^{\mu}}\rangle$. Therefore, $b = \lambda^{\mu}$ and $\pi^{\mu} = \mu$ by the uniqueness in part (b). □

Define $\varepsilon_1 = b/(4\|V\|_1)$. Note that $\varepsilon_1 < \frac{1}{4}$. For every $x \in \mathbb{Z}^d$ let

$$(4.11)\quad K(x, \varepsilon_1) \equiv \{\mu \in M_1(\mathbb{Z}^d) : \Lambda(\mu) > (1-\varepsilon_1)b,\ \mu(x) > \Lambda(\mu)/\|V\|_1\}$$

denote the set of all nearly optimal measures with considerable mass at x. The following three lemmas show that, *uniformly* for all μ in a neighborhood of the set K of optimal measures, the corresponding invariant measures π^{μ} given by Lemma 4.4 have an exponential decay without decaying too fast and that $\{X_t\}_{t\geq 0}$ under $\{\mathbb{Q}^{\mu}_x\}_{x\in\mathbb{Z}^d}$ "forgets" its starting point sufficiently fast.

LEMMA 4.12. *There exist constants $c_1, c_2 \in (0,\infty)$ such that*

$$\pi^{\mu}(y) \leq c_1 \exp(-c_2\|x-y\|_{\infty})$$

for all $x, y \in \mathbb{Z}^d$ and $\mu \in K(x, \varepsilon_1)$.

PROOF. It suffices to consider $x, y \in \mathbb{Z}^d$ with $\|x-y\|_{\infty} \geq 10n_0$, where n_0 is as in Lemma 2.3, and prove the estimate with the constants $\tilde{c}_1 \equiv (d+1)^2$ and $c_2 \equiv b/(2\max\{e^2, 2\|V\|_{\infty}\})$ for these x, y. Define $j = \lceil \|x-y\|_{\infty}/2\rceil$ and $t = 2c_2 j/b$. Using (4.6) for $A \equiv \Omega$ and $\lambda^{\mu} \geq \Lambda(\mu) \geq 3b/4$, it follows that

$$\sqrt{\pi^{\mu}(y)} \leq \exp(-3bt/4)\mathbb{E}_y[\exp(t\langle h^{\mu}, L_t\rangle)].$$

Note that $\langle h^{\mu}, L_t\rangle \leq 2\|V\|_{\infty}$ by (4.1). If $z \in \mathbb{Z}^d$ satisfies $\|z-x\|_1 \geq 5n_0 > 4n_0 + R$, then (4.1) and an application of Lemma 2.3 with $n \equiv 4n_0$ show that

$$H'(\mu)(z) \leq 2\|V\|_{\infty}\mu(\{\tilde{y} \in \mathbb{Z}^d : \|\tilde{y} - x\|_1 > 4n_0\}) \leq b\frac{\|V\|_{\infty}}{\|V\|_1} \leq b.$$

Hence $h^{\mu}(z) \le b/4$ because $H(\mu) \ge \Lambda(\mu) \ge 3b/4$. Let $N_t \equiv \max_{s\in[0,t]}\|X_s - y\|_\infty$. If the walk starts at y and if $N_t \le j$, then $\langle h^{\mu}, L_t\rangle \le b/4$. Therefore,

$$(4.13)\qquad \sqrt{\pi^{\mu}(y)} \le \exp(-bt/2) + \exp(-3bt/4)\exp(2t\|V\|_\infty)\mathbb{P}_y(N_t > j).$$

Applying Cramér's theorem as in the proof of Lemma 3.3, it follows that $\mathbb{P}_y(N_t > j) \le d\exp(-jh(t/j))$. Using $t/j \le e^{-2}$ and $h(e^{-2}) \ge 1$ for the first step and $j \ge 2t\|V\|_\infty$ for the second one, it follows that $\mathbb{P}_y(N_t > j) \le d\exp(-j) \le d\exp(-2t\|V\|_\infty)$. Substituting this estimate into (4.13) proves the lemma. □

LEMMA 4.14. *There exists a constant* $c_3 \in (0,\infty)$ *such that*

$$\pi^{\mu}(y) \ge c_3(4\|V\|_\infty + 2d)^{-2\|x-y\|_1}$$

for all $x, y \in \mathbb{Z}^d$ *and* $\mu \in K(x, \varepsilon_1)$.

PROOF. Define $n_1 = \max\{n_0 + 1, \lceil (R+1)(4\|V\|_\infty + 2d)^2/(b\varepsilon_1)^2\rceil\}$ with n_0 as in Lemma 2.3. Since $\mu(\{z \in \mathbb{Z}^d\colon \|x - z\|_1 > n_1\}) \le 2\varepsilon_1$ by Lemma 2.3, it follows from (4.1) and (4.2) that $H'(\mu)(z) \le 4\varepsilon_1\|V\|_\infty \le b$ and, therefore, $h^{\mu}(z) \le b/4$ for all $z \in \mathbb{Z}^d$ with $\|x - z\|_1 > n_1 + R$, because $H(\mu) \ge \Lambda(\mu) \ge 3b/4$. Since $\|h^{\mu}\|_\infty \le 2\|V\|_\infty$ and

$$\begin{aligned} b/4 + 2\|V\|_\infty(2n_1 + 2R + 1)^d \max\{\pi^{\mu}(z)\colon \|x - z\|_1 \le n_1 + R\} \\ \ge \langle h^{\mu}, \pi^{\mu}\rangle \ge \langle h^{\mu}, \pi^{\mu}\rangle - J(\pi^{\mu}) = \lambda^{\mu} \ge \Lambda(\mu) \ge 3b/4, \end{aligned}$$

where the equality follows from Lemma 4.4(b), there exists a point $\tilde{y} \in \mathbb{Z}^d$ with $\|x - \tilde{y}\|_1 \le n_1 + R$ such that

$$(4.15)\qquad \pi^{\mu}(\tilde{y}) \ge \frac{b}{4\|V\|_\infty}(2n_1 + 2R + 1)^{-d}.$$

From the proof of Lemma 4.4(d), in particular from (4.8), it follows that $\frac{1}{2}\sqrt{\pi^{\mu}(\tilde{z})/\pi^{\mu}(z)} - d \le \lambda^{\mu} - h^{\mu}(z) \le 2\|V\|_\infty$ for all $z, \tilde{z} \in \mathbb{Z}^d$ with $\|z - \tilde{z}\|_1 = 1$. Hence $\pi^{\mu}(z) \ge (4\|V\|_\infty + 2d)^{-2}\pi^{\mu}(\tilde{z})$. Given $y \in \mathbb{Z}^d$, there exists a path from $\tilde{y}$ via x to y with a length not exceeding $n_1 + R + \|x - y\|_1$. Applying the last estimate to every bond of this path and using (4.15), the lemma follows. □

LEMMA 4.16. *There exists an increasing function* $c\colon \mathbb{N}_0 \to [0,\infty)$ *such that*

$$(4.17)\qquad \|\mathbb{Q}_x^{\mu}X_t^{-1} - \mathbb{Q}_y^{\mu}X_t^{-1}\| \le \frac{c(\|x - y\|_1)}{t}$$

for all $x, y \in \mathbb{Z}^d$, $\mu \in K(x, \varepsilon_1)$ *and* $t > 0$. *Furthermore*,

$$(4.18)\qquad \lim_{t\to\infty}\ \sup_{\mu\in K(x,\varepsilon_1)} \|\mathbb{Q}_x^{\mu}X_t^{-1} - \pi^{\mu}\| = 0.$$

PROOF. Define random walks $\{X'_s\}_{s\ge 0}$ and $\{Y'_s\}_{s\ge 0}$ by $X'_s(\omega_1, \omega_2) = \omega_1(s)$ and $Y'_s(\omega_1, \omega_2) = \omega_2(s)$ for all $s \ge 0$ and $(\omega_1, \omega_2) \in \Omega^2$. Let $\tau \equiv \inf\{s \ge 0\colon$

$X'_s = Y'_s\}$ denote the first time they meet. Then $T\colon \Omega^2 \to \Omega^2$, defined by

$$T(\omega_1, \omega_2)(s) = \begin{cases} (\omega_1(s), \omega_2(s)), & \text{for } s \in [0, \tau(\omega_1, \omega_2)), \\ (\omega_1(s), \omega_1(s)), & \text{for } s \in [\tau(\omega_1, \omega_2), \infty), \end{cases}$$

gives the coalescent random walks and $\mathbb{Q}^{\mu,c}_{x,y} \equiv (\mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y) T^{-1}$ is the corresponding coupling measure (to verify this, show that $\mathbb{Q}^{\mu,c}_{x,y}$ is a time-homogeneous Markovian measure on $(\Omega^2, F \otimes F)$ with the classical coupling generator given in [4], Example 5.11). For every $A \subset \mathbb{Z}^d$ the coupling inequality yields

$$\begin{aligned} \left|\mathbb{Q}^\mu_x(X_t \in A) - \mathbb{Q}^\mu_y(X_t \in A)\right| &= \left|\mathbb{Q}^{\mu,c}_{x,y}(X'_t \in A) - \mathbb{Q}^{\mu,c}_{x,y}(Y'_t \in A)\right| \\ &\le \mathbb{Q}^{\mu,c}_{x,y}(\tau > t). \end{aligned}$$

Let $\mathbb{E}^\mu_{x,y}$ denote the expectation with respect to $\mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y$. Since $\tau \circ T = \tau$, it follows that $\mathbb{Q}^{\mu,c}_{x,y}(\tau > t) = (\mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y)(\tau > t) \le \mathbb{E}^\mu_{x,y}[\tau]/t$. To prove (4.17), it suffices to show for $x \ne y$ that $\mathbb{E}^\mu_{x,y}[\tau] \le c(j)$ with $j \equiv \|x - y\|_1$ and

$$c(j) \equiv \frac{(4\|V\|_\infty + 2d)^{2j}}{c_3^2} \frac{2^j j!}{d} \exp(4\|V\|_\infty + 2d), \tag{4.19}$$

where c_3 is the constant from Lemma 4.14.

Define $\sigma_0 = 0$ and $\sigma_n = \min\{i \in \mathbb{N}\colon i > \sigma_{n-1}, (X'_i, Y'_i) = (x, y)\}$ for every $n \in \mathbb{N}$. The idea is to show that after each stopping time σ_n the two walks have a new, independent chance to meet within the next time unit. Let $\{\Theta_s\}_{s \ge 0}$ be the time-shift operators on Ω^2. Then $\tau \le \sum_{n=0}^\infty \sigma_1 \circ \Theta_{\sigma_n} 1_{\{\tau > \sigma_n\}}$. Note that $\{\tau > \sigma_n\} = \{\tau \circ \Theta_{\sigma_{n-1}} > \sigma_1 \circ \Theta_{\sigma_{n-1}}\} \cap \{\tau > \sigma_{n-1}\}$ and $\{\tau > \sigma_{n-1}\} \in F_{\sigma_{n-1}}$ for every $n \in \mathbb{N}$. Hence, by the strong Markov property,

$$\mathbb{E}^\mu_{x,y}\left[\sigma_1 \circ \Theta_{\sigma_n} 1_{\{\tau > \sigma_n\}} \middle| F_{\sigma_n}\right] = 1_{\{\tau > \sigma_n\}} \mathbb{E}^\mu_{x,y}[\sigma_1]$$

and

$$\mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y\left(\tau > \sigma_n \middle| F_{\sigma_{n-1}}\right) = 1_{\{\tau > \sigma_{n-1}\}} \mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y(\tau > \sigma_1).$$

Therefore, using the last equality recursively,

$$\mathbb{E}^\mu_{x,y}[\tau] \le \mathbb{E}^\mu_{x,y}[\sigma_1] \sum_{n=0}^\infty \left(\mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y(\tau > \sigma_1)\right)^n = \frac{\mathbb{E}^\mu_{x,y}[\sigma_1]}{\mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y(\tau \le \sigma_1)}. \tag{4.20}$$

By Lemma 4.4(f), the unique invariant distribution of $\mathbb{Q}^\mu_x \otimes \mathbb{Q}^\mu_y$ is given by $\pi^\mu \otimes \pi^\mu$, hence the Markov chain $\{(X'_n, Y'_n)\}_{n \in \mathbb{N}_0}$ is positive recurrent ([12], page 74). It follows from [12], Example 5.1(a), that $\mathbb{E}^\mu_{x,y}[\sigma_1] = (\pi^\mu(x)\pi^\mu(y))^{-1}$. Using (4.20) and Lemma 4.14, the first quotient of (4.19) follows.

To estimate the denominator in (4.20), we only consider the case $\pi^\mu(x) \ge \pi^\mu(y)$ and use $\{X'_1 = x, Y'_1 = x\} \subset \{\tau \le \sigma_1\}$, because the case $\pi^\mu(x) < \pi^\mu(y)$ using $\{X'_1 = y, Y'_1 = y\} \subset \{\tau \le \sigma_1\}$ is similar. It follows from (4.6) and $\lambda^\mu - \langle h^\mu, L_1\rangle \le 2\|V\|_\infty$ that $\mathbb{Q}^\mu_x(X_1 = x) \ge \exp(-2\|V\|_\infty)\mathbb{P}_x(X_1 = x)$ and $\mathbb{Q}^\mu_y(X_1 = x) \ge \exp(-2\|V\|_\infty)\mathbb{P}_y(X_1 = x)$. Note that $\mathbb{P}_x(X_1 = x) \ge e^{-d}$, because the walk may stay at x during $[0, 1]$. To estimate $\mathbb{P}_y(X_1 = x)$, note that there exists at least one path from y to x of length $j \equiv \|x - y\|_1$, hence the walk can reach x from y in time $u \equiv 1$ using j steps if it goes along this path and then stays at x. Since the distribution of the time for j jumps is given by the gamma density $f_{d,j}(s) \equiv d^j s^{j-1} e^{-ds}/(j-1)!$, it follows that

$$\mathbb{P}_y(X_u = x) \geq \frac{1}{(2d)^j} \int_0^u f_{d,j}(s)\, d e^{-d(u-s)}\, ds = \frac{f_{d,j+1}(u)}{(2d)^j}. \tag{4.21}$$

Hence $\mathbb{P}_y(X_1 = x) \geq 2^{-j} d e^{-d}/j!$. Combining this estimate with the ones given above, the remaining factors in (4.19) follow.

To prove (4.18), note that due to Lemma 4.4(f) and (4.17),

$$\begin{aligned} \|\mathbb{Q}_x^\mu X_t^{-1} - \pi^\mu\| &\leq \sum_{y \in \mathbb{Z}^d} \pi^\mu(x - y) \|\mathbb{Q}_x^\mu X_t^{-1} - \mathbb{Q}_{x-y}^\mu X_t^{-1}\| \\ &\leq \sum_{y \in \mathbb{Z}^d} \Pi(y) \min\{1, c(\|y\|_1)/t\} \end{aligned} \tag{4.22}$$

for all $\mu \in K(x, \varepsilon_1)$ and $t > 0$, where $\Pi(y) \equiv \sup\{\pi^\mu(x - y)\colon \mu \in K(x, \varepsilon_1)\}$ for every $y \in \mathbb{Z}^d$. Lemma 4.12 implies that $\sum_{y \in \mathbb{Z}^d} \Pi(y) < \infty$. Hence (4.18) follows from (4.22) using the dominated convergence theorem for the limit $t \to \infty$. □

We need the following lemma for $t = 0$ to prove the tightness of $\{\hat{\mathbb{P}}_T L_T^{-1}\}_{T \geq 0}$, and we will need its full strength to reduce the convergence in (1.17) from $\mathbb{Z}^d$ to various big cubes. Furthermore, it will enable us to prove Theorem 1.4.

LEMMA 4.23. *There exist* $\alpha, \varepsilon_0, T_0 > 0$ *such that, for every* $\mu \in K(0)$,

$$\sup_{T \geq T_0} \sup_{t \in [0,T]} \sum_{x \in \mathbb{Z}^d} \hat{\mathbb{E}}_T\big[\exp(\alpha \|X_t - x\|_\infty);\, L_T \in U_{\varepsilon_0}(\theta_x(\mu))\big] < \varepsilon. \tag{4.24}$$

REMARK 4.25. Let us briefly explain how we are going to prove this important lemma. If L_T is in an ε_0-neighborhood of $\theta_x(\mu)$, then, due to Lemma 2.3, the process $\{X_s\}_{s \in [0,T]}$ spends most of its time in the vicinity of x. When the process is far away from x at time t, then it must be on an excursion from the main bulk of $\theta_x(\mu)$ during a time interval $[u, v]$, where $u \in [0, t]$ and $v \in [t, T]$. Since we do not have Lemma 4.16 for random times, we need to discretize time. As in (1.1) define, for $(u, v) \neq (0, T)$,

$$L_{u,v,T} = \frac{1}{u + T - v} \int_{[0,u) \cup [v,T)} \delta_{X_s}\, ds. \tag{4.26}$$

Splitting L_T as

$$L_T = \frac{v - u}{T} L_{u,v} + \left(1 - \frac{v - u}{T}\right) L_{u,v,T},$$

we can use (4.1) and (4.2) to decompose the Hamiltonian in the following way:

$$TH(L_T) = (T - v + u)H(L_{u,v,T}) + (v-u)\langle h^{L_{u,v,T}}, L_{u,v}\rangle + \frac{(v-u)^2}{T}\big(H(L_{u,v}) - \langle h^{L_{u,v,T}}, L_{u,v}\rangle\big). \tag{4.27}$$

Since $v - u$ is small compared with T, the contribution of $H(L_{u,v})$ is small. Furthermore, since $L_{u,v}$ has its support outside the main bulk of $L_{u,v,T}$ and since V has only finite support, the contribution of $\langle H'(L_{u,v,T}), L_{u,v}\rangle$ is small. If $v - u \le s_0$ for an appropriate $s_0 > 0$ and if $\|X_t - x\|_\infty$ is sufficiently large, then the process must have jumped very often during $[u, v]$. Using a large deviation argument concerning the sum of the holding times, we can show that the probability of such a far-reaching excursion during $[u, v]$ decreases exponentially with $\|X_t - x\|_\infty$. If $v - u > s_0$, then we consider a partially exchanged path, which is identical to the original one during $[0, T]\setminus[u, v]$ but hangs around the main bulk of $\theta_x(\mu)$ during $[u, v]$. For this modified path, $\langle H'(L_{u,v,T}), L_{u,v}\rangle$ and therefore $H(L_T)$ are considerably increased to surpass $\alpha\|X_t - x\|_\infty$ without paying too much "entropy." Hence, the partially exchanged path has a substantially higher probability with respect to the transformed measure $\hat{\mathbb{P}}_T$. That is, the $\hat{\mathbb{P}}_T$-probability of the original path was small enough to balance $\exp(\alpha\|X_t - x\|_\infty)$. See [1] for a more involved application of such a "partial path exchange" argument.

Proof of Lemma 4.23. With b given by (1.9), define $\alpha = \cosh^{-1}(1 + b/3)$, $\varepsilon_0 = \varepsilon$ with

$$\varepsilon \equiv \min\left\{\frac{b}{\|V\|_1}, \frac{b}{168\|V\|_\infty}, \frac{1}{18}\inf_{\mu\in K(0)} f(\mu)\right\}, \tag{4.28}$$

where $f(\mu) \equiv \inf_{x\in\mathbb{Z}^d\setminus\{0\}}\|\theta_x(\mu) - \mu\|$ for all $\mu \in M_1(\mathbb{Z}^d)$, and $T_0 = 1/\varepsilon$. To prove that $\varepsilon > 0$, first note that $b > 0$ by Condition 1.10 and that $K(0)$ is compact by Proposition 1.11(c). Since $\lim_{\|x\|_1\to\infty}\|\theta_x(\mu) - \mu\| = 1$ for every $\mu \in M_1(\mathbb{Z}^d)$, there always exists $x_\mu \in \mathbb{Z}^d\setminus\{0\}$ with $f(\mu) = \|\theta_{x_\mu}(\mu) - \mu\|$. Furthermore, $f(\mu) > 0$, because there is no shift-invariant μ. To see that f is continuous, notice that for all $\mu, \nu \in M_1(\mathbb{Z}^d)$,

$$\begin{aligned} f(\nu) &\le \big\|\theta_{x_\mu}(\nu) - \nu\big\| \\ &\le \big\|\theta_{x_\mu}(\nu) - \theta_{x_\mu}(\mu)\big\| + \big\|\theta_{x_\mu}(\mu) - \mu\big\| + \|\mu - \nu\| \\ &= f(\mu) + 2\|\mu - \nu\|. \end{aligned}$$

Given $\mu \in K(0)$, there exists $m \in \mathbb{N}$ such that $m > R$ and

$$\mu(B^c_{m-R,0}) \le \varepsilon, \tag{4.29}$$

where $B_{n,x} \equiv \{y \in \mathbb{Z}^d : \|x - y\|_\infty < n\}$ for $n \in \mathbb{N}$ and $x \in \mathbb{Z}^d$. By Proposition 1.11(b) the constant

$$c_\mu = \max_{y,z\in S_{m,0}} \frac{1}{\sqrt{\mu(y)\mu(z)}} \tag{4.30}$$

is well defined, where $S_{m,x} \equiv \{y \in \mathbb{Z}^d\colon \|x-y\|_\infty = m\}$ for $x \in \mathbb{Z}^d$. Using Lemma 4.4(g) and Lemma 4.16, it follows that there exists $s_0 \ge 1$ such that

$$\min_{y,z\in S_{m,0}} \frac{\mathbb{Q}_y^\mu(X_s = z)}{\mu(z)} \ge \frac{1}{2} \quad \text{for every } s \ge s_0. \tag{4.31}$$

As a convenient abbreviation, choose $n \in \mathbb{N}$ such that $n \ge m+1+\max\{s_0, 2dm\}$ and $(k/e)^k \ge (k-1)!$ for all $k \ge n-m-1$.

For the remaining part of the proof fix a time $T \ge T_0$. For all $t \in [0,T]$ and $x \in \mathbb{Z}^d$ define the random time $\sigma_{t,x} = \max\{0, \sup\{u \le t\colon X_{u-} \in S_{m,x}\}\}$ and the stopping time $\tau_{t,x} = \min\{T, \inf\{v \ge t\colon X_v \in S_{m,x}\}\}$. We say that a path with $X_t \in B_{n,x}^c$ is on its excursion from the main bulk of $\theta_x(\mu)$ during $[\sigma_{t,x}, \tau_{t,x})$. We define the time spans, for which the walk rests before and after its excursion, by $\xi_{t,x}^\sigma = \min\{\sigma_{t,x}, \inf\{u \in (0,\sigma_{t,x}]\colon X(\sigma_{x,t}-u) \ne X(\sigma_{x,t}-)\}\}$ and $\xi_{t,x}^\tau = \min\{T-\tau_{t,x}, \inf\{v \in (0, T-\tau_{t,x}]\colon X(\tau_{t,x}+v) \ne X(\tau_{t,x})\}\}$. Furthermore, for all $\varrho > 0$ and $x \in \mathbb{Z}^d$, we define the two events $A_{\varrho,\sigma,t,x} = \{\xi_{t,x}^\sigma \le \varrho, \xi_{t,x}^\sigma \ne \sigma_{t,x}\}$ and $A_{\varrho,\tau,t,x} = \{\xi_{t,x}^\tau \le \varrho, \xi_{t,x}^\tau \ne T-\tau_{t,x}\}$. We want to show that, for every $\varrho \in (0,1]$,

$$\begin{aligned}\sum_{x\in\mathbb{Z}^d} \hat{\mathbb{E}}_T\Big[&\exp(\alpha\|X_t-x\|_\infty); \{L_T \in U_\varepsilon(\theta_x(\mu)), X_t \in B_{n,x}^c\}\\ &\cap A_{\varrho,\sigma,t,x}^c \cap A_{\varrho,\tau,t,x}^c\Big] \le c,\end{aligned} \tag{4.32}$$

where the constant c is explicitly expressible in terms of α, b, c_μ, d, ε, m, s_0 and $\|V\|_\infty$; it does not depend on ϱ, t or T. Since $\Omega = D([0,\infty), \mathbb{Z}^d)$, it follows that $A_{\varrho,\sigma,t,x} \downarrow \varnothing$ and $A_{\varrho,\tau,t,x} \downarrow \varnothing$ for $\varrho \downarrow 0$. Hence we can apply Fatou's lemma to (4.32) for the limit $\varrho \downarrow 0$. Since $U_\varepsilon(\theta_x(\mu))$ and $U_\varepsilon(\theta_y(\mu))$ are disjoint for all $x, y \in \mathbb{Z}^d$ with $x \ne y$, we obtain that the series of expectations in (4.24) is bounded by $c + e^{\alpha n}$ for every $T \ge t_0$ and $t \in [0,T]$.

As a reduction step, let us show that (4.32) for $t \in [0, T/2)$ follows from (4.32) for $t \in [T/2, T]$. For this purpose, we define the map $\mathcal{T}_T\colon \Omega \to \Omega$, which reverses the time in $[0,T]$, by

$$\mathcal{T}_T(\omega)(s) = \begin{cases} \omega_-(T-s) + \omega(0) - \omega_-(T), & \text{for } s \in [0,T),\\ 2\omega(0) - \omega(s), & \text{for } s \in [T,\infty),\end{cases}$$

where ω_- is the left-continuous version of $\omega \in \Omega$. Since $\{X_s\}_{s\ge 0}$ is a time-homogeneous process with independent, symmetric increments under $\mathbb{P}$ and $\mathbb{P}(X_s \ne X_{s-}) = 0$ for all $s > 0$, it follows that $\mathbb{P} = \mathbb{P}\mathcal{T}_T^{-1}$. Hence $\mathcal{T}_T$ is measure-preserving. Note that $L_T \circ \mathcal{T}_T = \theta_{X_0 - X_{T-}}(L_T)$. Hence $H(L_T) = H(L_T \circ \mathcal{T}_T)$ and

$$\{L_T \circ \mathcal{T}_T \in U_\varepsilon(\theta_x(\mu))\} = \Big\{L_T \in U_\varepsilon\big(\theta_{x - X_0 + X_{T-}}(\mu)\big)\Big\}.$$

Furthermore, $t - \sigma_{t,x} \circ \mathcal{T}_T = \tau_{T-t, x-X_0+X_{T-}} - (T-t)$. In addition, $\xi_{t,x}^\sigma \circ \mathcal{T}_T = \xi_{T-t, x-X_0+X_{T-}}^\tau$. Since the ε-neighborhoods in (4.32) are disjoint, the series and the expectation can be exchanged, and the above relations can be used to rewrite (4.32) with $T-t$ in place of t.

In the following we fix $t \in [T/2, T]$ and $x \in \mathbb{Z}^d$ and show that

$$\begin{aligned}\mathbb{E}\big[\exp(\alpha\|X_t - x\|_\infty + TH(L_T)); \{L_T \in U_\varepsilon(\theta_x(\mu)), X_t \in B^c_{n,x}\} \\ \cap A^c_{\varrho,\sigma,t,x} \cap A^c_{\varrho,\tau,t,x}\big] \\ \le c\mathbb{E}\big[\exp(TH(L_T)); L_T \in U_{9\varepsilon}(\theta_x(\mu))\big],\end{aligned} \tag{4.33}$$

which implies (4.32) because the 9ε-neighborhoods are disjoint. We will drop the indices t and x at various places.

If $L_T \in U_\varepsilon(\theta_x(\mu))$, then, by (4.29),

$$L_T(B^c_{m-R,x}) \le \varepsilon + \mu(B^c_{m-R,0}) \le 2\varepsilon. \tag{4.34}$$

If in addition $X_t \in B^c_{n,x}$, then $L_{u,v}(B^c_{m,x}) = 1$ for every $u \in [\sigma - \xi^\sigma, \sigma]$ and $v \in [\tau, \tau + \xi^\tau]$. Since $L_{u,v,T}(B^c_{m-R,x}) \le L_T(B^c_{m-R,x}) \le 2\varepsilon$ by (4.34), it follows from (4.1) and the definition of the radius R that $\langle H'(L_{u,v,T}), L_{u,v}\rangle \le 4\varepsilon\|V\|_\infty$ on $\{L_T \in U_\varepsilon(\theta_x(\mu)), X_t \in B^c_{n,x}\}$. It follows from (4.34) and $L_{u,v}(B^c_{m,x}) = 1$ that $v - u \le 2\varepsilon T$. Since $0 \le H(\nu) \le \|V\|_\infty$ for all $\nu \in M_1(\mathbb{Z}^d)$, it follows from (4.2) and (4.27) that $TH(L_T) \le (T - v + u)H(L_{u,v,T}) + 8\varepsilon(v-u)\|V\|_\infty$.

To discretize time, define the sets $I_\varrho = \{t - j\varrho: j \in \{1, 2, \dots, \lfloor 2\varepsilon T/\varrho\rfloor\}\}$ and $I'_\varrho = \{(t + j\varrho) \wedge T: j \in \{1, 2, \dots, \lfloor 2\varepsilon T/\varrho\rfloor\}\}$. Since $t \ge T/2$ and $\varepsilon \le 1/8$, it follows that $I_\varrho \subset [T/4, T)$. If $L_T \in U_\varepsilon(\theta_x(\mu))$ and $(u,v) \in I_\varrho \times I'_\varrho$, then $L_{u,v,T} \in U_{5\varepsilon}(\theta_x(\mu))$. For the next inequality we use the results of the previous paragraph and the Markov property, partially applied in the form that $\mathbb{P}(C|\sigma(\{X_s: s \in [u,v]\})) = P(C|X_u, X_v)$ for $C \in \sigma(\{X_s: s \in [0,u] \cup [v,\infty)\})$ $\mathbb{P}$-almost surely. Since the walk may be outside of $B_{m,x}$ during $[t,T]$ if t is close to T, we have to consider two different terms (one of them may be zero):

$$\begin{aligned}&\mathbb{E}\big[\exp(\alpha\|X_t - x\|_\infty + TH(L_T)); \{L_T \in U_\varepsilon(\theta_x(\mu)), X_t \in B^c_{n,x}\} \cap A^c_{\varrho,\sigma} \cap A^c_{\varrho,\tau}\big] \\ &\quad\le \sum_{u\in I_\varrho}\sum_{v\in I'_\varrho\setminus\{T\}}\sum_{y,z\in S_{m,x}} \exp(8\varepsilon(v-u)\|V\|_\infty)\mathbb{P}(X_u = y) \\ &\qquad\times\mathbb{E}\big[\exp((T-v+u)H(L_{u,v,T})); \\ &\qquad\qquad L_{u,v,T} \in U_{5\varepsilon}(\theta_x(\mu))|X_u = y, X_v = z\big] \\ &\qquad\times\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty); X_\varrho \ne y, X_{t-u} \in B^c_{n,x}, \\ &\qquad\qquad L_{v-u}(B^c_{m,x}) = 1, X_{v-u-\varrho} \ne z, X_{v-u} = z\big] \\ &\qquad+ \sum_{u\in I_\varrho}\sum_{v\in I'_\varrho\cap\{T\}}\sum_{y\in S_{m,x}} \exp(8\varepsilon(T-u)\|V\|_\infty) \\ &\qquad\quad\times\mathbb{E}\big[\exp(uH(L_u)); L_u \in U_{5\varepsilon}(\theta_x(\mu)), X_u = y\big] \\ &\qquad\quad\times\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty); \\ &\qquad\qquad L_{T-u}(B^c_{m,x}) = 1, X_\varrho \ne y, X_{t-u} \in B^c_{n,x}\big].\end{aligned} \tag{4.35}$$

We will handle the terms with $v - u \le s_0$ by a large deviations argument; for the other ones we will use a "partial path exchange" argument.

Consider $y, z \in S_{m,x}$ and $(u,v) \in I_\varrho \times (I'_\varrho \setminus \{T\})$. For every $s \in [0,\infty)$ and $e \in \mathbb{Z}^d$ with $\|e\|_1 = 1$ let N^e_s denote the number of jumps of $\{X_w\}_{w\ge 0}$ during $[0,s]$ of size e. The $2d$ processes $\{N^e_s\}_{s\ge 0}$ are independent Poisson processes under $\mathbb{P}_y$ with intensity $1/2$. Since $\mathbb{P}_y(N^e_\varrho \geq 1 | N^e_s = k) \leq k\varrho/s$ for all $s \geq \varrho$ and $k \in \mathbb{N}$ (see, e.g., [15], Chapter 5, Theorem 3.2), it follows that $\mathbb{P}_y(N^e_\varrho \geq 1, N^e_{t-u} = k) \leq (\varrho/2)\mathbb{P}_y(N^e_{t-u} = k-1)$. Since $\mathbb{P}_y(X_{v-u-\varrho} \neq X_{v-u}) \leq 1 - e^{-d\varrho} \leq d\varrho$, it follows by considering the direction of the first jump of $\{X_s\}_{s\geq 0}$ that

$$\begin{aligned}\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty);\, X_\varrho \neq y,\, X_{t-u} \in B^c_{n,x},\, X_{v-u-\varrho} \neq z,\, X_{v-u} = z\big]\\ \leq d^2\varrho^2 e^\alpha \mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty);\, X_{t-u} \in B^c_{n-1,x}\big].\end{aligned} \tag{4.36}$$

Similarly, for every $u \in I_\varrho$ and $y \in S_{m,x}$,

$$\begin{aligned}\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty);\, X_\varrho \neq y,\, X_{t-u} \in B^c_{n,x}\big]\\ \leq d\varrho e^\alpha \mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty);\, X_{t-u} \in B^c_{n-1,x}\big].\end{aligned} \tag{4.37}$$

Consider $y, z \in S_{m,x}$ and $(u, v) \in I_\varrho \times I'_\varrho$ satisfying $v - u \leq s_0$. If the event $\{X_0 = y, \|X_{t-u} - y\|_\infty = k\}$ occurs for some $k \geq n - m - 1$, then there exists at least one coordinate direction in which the walk has jumped at least k times during $[0, v-u]$. As in the proof of Lemma 3.3, it follows that

$$\begin{aligned}\mathbb{P}_y(\|X_{t-u} - y\|_\infty = k) &\leq d\exp(-kh((v-u)/k))\\ &\leq d\left(\frac{e(v-u)}{k}\right)^k \leq d\frac{s_0^k}{(k-1)!}\left(\frac{v-u}{s_0}\right)^{2dm},\end{aligned} \tag{4.38}$$

because $v - u \leq s_0 \leq k$, $2dm \leq k$ and $(k/e)^k \geq (k-1)!$ by the above choice of n, and because $-h(r) \leq 1 + \log r$ for all $r > 0$. It follows that

$$\begin{aligned}&\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty);\, X_{t-u} \in B^c_{n-1,x}\big]\\ &\quad\leq e^{\alpha m}\sum_{k=n-m-1}^\infty e^{\alpha k}\mathbb{P}_y(\|X_{t-u} - y\|_\infty = k)\\ &\quad\leq d\exp(\alpha(m+1))s_0\exp(s_0 e^\alpha)\left(\frac{v-u}{s_0}\right)^{2dm}.\end{aligned} \tag{4.39}$$

Define $g\colon \mathbb{Z}^d \to \mathbb{R}$ by $g(\tilde{z}) = \max\{-\|V\|_\infty, h^{\theta_\mu(x)}(\tilde{z}) - 40\varepsilon\|V\|_\infty\}$. If $v - u \leq s_0$, it follows with (4.21) applied to $\mathbb{P}_y(X_{u-v} = z)$ with $j \equiv \|y - z\|_1 \leq 2dm$ that

$$\begin{aligned}&\mathbb{E}_y\big[\exp((v-u)\langle g, L_{v-u}\rangle);\, X_{v-u} = z\big]\\ &\quad\geq \frac{d}{(2dm)!}\left(\frac{v-u}{s_0}\right)^{2dm}\exp(-s_0(d + \|V\|_\infty)),\end{aligned}$$

because $s_0 \geq 1$. Comparison with (4.39) shows that

$$\begin{aligned}&\exp(8\varepsilon(v-u)\|V\|_\infty)\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty);\, X_{t-u} \in B^c_{n-1,x}\big]\\ &\quad\leq (2dm)!s_0 e^{\alpha(m+1)}\exp(s_0 e^\alpha)\exp(s_0 d + s_0(1+8\varepsilon)\|V\|_\infty)\\ &\qquad\times \mathbb{E}_y\big[\exp((v-u)\langle g, L_{v-u}\rangle);\, X_{v-u} = z\big].\end{aligned} \tag{4.40}$$

This estimate is also valid without the event $\{X_{v-u} = z\}$ in the last expectation.

We now consider $y, z \in S_{m,x}$ and $(u, v) \in I_\varrho \times I'_\varrho$ which satisfy $v - u > s_0$. Lemma 4.4(d) and (g) together with (4.30) and (4.31) show that

$$\begin{aligned}(4.41)\qquad &\mathbb{E}_y\big[\exp\big((v-u)\langle h^{\theta_x(\mu)}, L_{v-u}\rangle\big); X_{v-u} = z\big]\\ &\quad = e^{b(v-u)}\sqrt{\frac{\theta_x(\mu)(y)}{\theta_x(\mu)(z)}}\,\mathbb{Q}_y^{\theta_x(\mu)}(X_{v-u} = z) \ge \frac{e^{b(v-u)}}{2c_\mu}.\end{aligned}$$

For $s \ge 0$ let X'_s denote the first component of X_s. Since $f(s) \equiv \mathbb{E}[\exp(\alpha X'_s)]$ satisfies $f'(s) = f(s)(\cosh(\alpha) - 1)$, it follows that $f(s) = \exp((\cosh(\alpha) - 1)s)$. Hence, $\mathbb{E}[\exp(\alpha\|X_s\|_\infty)] \le 2d\exp((\cosh(\alpha) - 1)s)$ for all $s \ge 0$. In particular,

$$\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty\big) \le 2de^{\alpha m}\exp\big((\cosh(\alpha) - 1)(v-u)\big).$$

Since $\alpha = \cosh^{-1}(1 + b/3)$ and $56\varepsilon\|V\|_\infty \le b/3$ by the choice of α and ε, a comparison with (4.41) yields

$$\begin{aligned}&\exp(16\varepsilon(v-u)\|V\|_\infty)\mathbb{E}_y\big[\exp(\alpha\|X_{t-u} - x\|_\infty)\big]\\ (4.42)\qquad &\quad \le 4c_\mu de^{\alpha m}\exp(-b(v-u)/3)\\ &\qquad \times\mathbb{E}_y\big[\exp((v-u)\langle g, L_{v-u}\rangle); X_{v-u} = z\big],\end{aligned}$$

where we may again drop the event $\{X_{v-u} = z\}$ in the last expectation.

Finally, if $(u, v) \in I_\varrho \times I'_\varrho$, then $v - u \le 4\varepsilon T$ and, by (4.1) and (4.2),

$$(4.43)\qquad \frac{(v-u)^2}{T}\big(H(L_{u,v}) - \langle h^{L_{u,v,T}}, L_{u,v}\rangle\big) + 8\varepsilon(v-u)\|V\|_\infty \ge 0.$$

Furthermore, if $L_{u,v,T} \in U_{5\varepsilon}(\theta_x(\mu))$, then $\|h^{L_{u,v,T}} - h^{\theta_x(\mu)}\|_\infty \le 40\varepsilon\|V\|_\infty$. Using $h^{L_{u,v,T}} \ge -\|V\|_\infty$, it follows that

$$(4.44)\qquad \langle h^{L_{u,v,T}}, L_{u,v}\rangle \ge \langle g, L_{u,v}\rangle.$$

We now have all the ingredients to derive (4.33) from (4.35). There are at most $\lfloor s_0/\varrho\rfloor^2$ terms in (4.35) with $(u, v) \in I_\varrho \times (I'_\varrho \setminus \{T\})$ satisfying $v - u \le s_0$ and at most $\lfloor s_0/\varrho\rfloor$ terms with $(u, v) \in I_\varrho \times (I'_\varrho \cap \{T\})$ satisfying $v - u \le s_0$. We can use (4.36) and (4.37), respectively, and then (4.40) to obtain an upper estimate for these. We then use the Markov property to put the factors in (4.35) together, use (4.44) to replace g by $h^{L_{u,v,T}}$, insert the exponential of the left-hand side of (4.43) and use the decomposition (4.27). Finally, since $v - u \le 4\varepsilon T$, we can replace $L_{u,v,T} \in U_{5\varepsilon}(\theta_x(\mu))$ by $L_T \in U_{9\varepsilon}(\theta_x(\mu))$.

To estimate the other terms in (4.35) with $v - u > s_0$, we first use (4.36), if $v < T$, or (4.37), if $v = T$. The next step is to use (4.42), where we drop the event $\{X_{v-u} = z\}$ in the case $v = T$. Then, as above, we use the Markov property to put the factors in (4.35) together, use (4.44) to replace g by $h^{L_{u,v,T}}$, insert the exponential of (4.43) and use the decomposition (4.27). Then, since $v - u \le 4\varepsilon T$, we can replace $L_{u,v,T} \in U_{5\varepsilon}(\theta_x(\mu))$ by $L_T \in U_{9\varepsilon}(\theta_x(\mu))$. We are left with two sums, which we regard as Riemann sums and estimate as follows:

$$\sum_{\substack{(u,v)\in I_\varrho\times(I'_\varrho\setminus\{T\})\\ v-u>s_0}} \varrho^2 \exp\biggl(-b\frac{v-u}{3}\biggr) + \sum_{\substack{(u,v)\in I_\varrho\times(I'_\varrho\cap\{T\})\\ v-u>s_0}} \varrho \exp\biggl(-b\frac{v-u}{3}\biggr)$$
$$\le \int_0^t\int_t^T \exp\biggl(-b\frac{v-u}{3}\biggr)\,dv\,du + \int_0^T \exp\biggl(-b\frac{T-u}{3}\biggr)\,du \le \frac{3b+9}{b^2}.$$

Using this, we finally obtain (4.33) from (4.35). □

PROOF OF THEOREM 1.4. Let $\alpha, \varepsilon_0, T_0 > 0$ be given by Lemma 4.23. Since $K(0)$ is compact and nonvoid by Proposition 1.11, there exists a finite nonvoid subset M of $K(0)$ such that $\bigcup_{\mu\in M} U_{\varepsilon_0/2}(\mu)$ covers $K(0)$. According to Corollary 3.12, there exists $\tilde{\alpha} > 0$ such that (3.13) holds with ε replaced by $\varepsilon_0/2$. Define $\alpha_0 = \frac{1}{2}\min\{\alpha, \tilde{\alpha}\}$. Note that $U_{\varepsilon_0/2}(K) \subset \bigcup_{\mu\in M}\bigcup_{x\in\mathbb{Z}^d} U_{\varepsilon_0}(\theta(\mu))$ by Proposition 1.11(c). In view of Lemma 3.3(c), it suffices to show that

$$\text{(4.45)}\qquad \sup_{T\ge T_0}\sup_{t\in[0,T]}\sum_{\mu\in M}\sum_{x\in\mathbb{Z}^d}\hat{\mathbb{E}}_T\bigl[\exp(\alpha_0\|X_t\|_\infty);\, L_T\in U_{\varepsilon_0}(\theta_x(\mu))\bigr] < \infty.$$

Since $\exp(\alpha_0\|X_t - X_0\|_\infty) \le \exp(\alpha_0\|X_t - x\|_\infty)\exp(\alpha_0\|X_0 - x\|_\infty)$ for all $t \ge 0$ and $x \in \mathbb{Z}^d$, an application of the Cauchy–Schwarz inequality shows that

$$\sup_{T\ge T_0}\sup_{t\in[0,T]}\sum_{\mu\in M}\sum_{x\in\mathbb{Z}^d}\hat{\mathbb{E}}_T\bigl[\exp(2\alpha_0\|X_t - x\|_\infty);\, L_T\in U_{\varepsilon_0}(\theta_x(\mu))\bigr] < \infty$$

is sufficient for (4.45). Since $2\alpha_0 \le \alpha$, this follows from Lemma 4.23. □

We are now ready to prove the tightness of the set $\{\hat{\mathbb{P}}_T L_T^{-1}\}_{T\ge 0}$ by combining Proposition 3.1 with Lemma 4.23. For use in the proof of Theorem 1.15, we formulate (4.47) of the following proposition with the supremum over all t in $[0,T]$. For part (b) we only need the case $t = 0$ and use the fact that $\hat{\mathbb{P}}_T(X_0 = 0) = 1$. For $n \in \mathbb{N}$ define

$$K(n) = \bigl\{\theta_x(\mu)\colon \mu\in K(0),\, x\in\mathbb{Z}^d,\, \|x\|_\infty \le n\bigr\}.$$

PROPOSITION 4.46. *Let ε_0 be defined as in Lemma* 4.23.

(a) *For every $\eta > 0$ there exists $n \in \mathbb{N}$ such that, for every $\varepsilon \in (0, \varepsilon_0/2]$, there exists $T_1 > 0$ satisfying*

$$\text{(4.47)}\qquad \sup_{T\ge T_1}\sup_{t\in[0,T]}\hat{\mathbb{P}}_T\bigl(L_T\in M_1(\mathbb{Z}^d)\setminus U(\varepsilon, n, X_t)\bigr) \le \eta,$$

where $U(\varepsilon, n, x) \equiv \{\theta_x(\mu)\colon \mu\in U_\varepsilon(K(n))\}$ for all $x\in\mathbb{Z}^d$.

(b) *The set $\{\hat{\mathbb{P}}_T L_T^{-1}\}_{T\ge 0}$ is a tight subset of $M_1(M_1(\mathbb{Z}^d))$.*

PROOF. (a) Since $K(0)$ is compact and nonvoid by Proposition 1.11, there exists a finite nonvoid subset M of $K(0)$ such that $\bigcup_{\mu\in M} U_{\varepsilon_0/2}(\mu)$ covers $K(0)$. Let α and T_0 be given by Lemma 4.23. For each $\mu\in M$ let c_μ denote the left-hand side of (4.24). Define $c_0 = \max_{\mu \in M} c_\mu$. Choose $n \in \mathbb{N}$ satisfying $\eta e^{\alpha n} \geq 2c_0|M|$. By a Chebyshev-type estimate, for all $\mu \in M$, $T \geq T_0$ and $t \in [0,T]$,

$$\sum_{x \in \mathbb{Z}^d} \hat{\mathbb{P}}_T\big(L_T \in U_{\varepsilon_0}(\theta_x(\mu)), \|X_t - x\|_\infty > n\big)$$
$$\leq \frac{\eta}{2c_0|M|} \sum_{x \in \mathbb{Z}^d} \hat{\mathbb{E}}_T\big[\exp(\alpha\|X_t - x\|_\infty); L_T \in U_{\varepsilon_0}(\theta_x(\mu))\big] \leq \frac{\eta}{2|M|}.$$

Given $\varepsilon \in (0, \varepsilon_0/2]$, Proposition 3.1 guarantees the existence of $T_1 \geq T_0$ such that $\sup_{T \geq T_1} \hat{\mathbb{P}}_T(L_T \notin U_\varepsilon(K)) \leq \eta/2$. Note that $\bigcup_{\mu \in M} U_{\varepsilon_0}(\mu)$ covers $U_\varepsilon(K(0))$. Hence, using Proposition 1.11(c),

$$\big\{\theta_{-X_t}(L_T) \notin U_\varepsilon(K(n))\big\}$$
$$\subset \{L_T \notin U_\varepsilon(K)\} \cup \bigcup_{x \in \mathbb{Z}^d} \bigcup_{\mu \in M} \big\{L_T \in U_{\varepsilon_0}(\theta_x(\mu)), \|X_t - x\|_\infty > n\big\}.$$

Therefore, (4.47) follows from the two estimates above.

(b) Given $\eta > 0$ and $\varepsilon \in (0, \varepsilon_0/2]$, choose $n \in \mathbb{N}$ and $T_1 > 0$ according to part (a) such that $\hat{\mathbb{P}}_T(L_T \notin U_\varepsilon(K(n))) \leq \eta$ for all $T \geq T_1$. Since $K(0)$ is compact by Proposition 1.11(c), the set $K(n)$ is compact, too. Since $[0, T_1] \ni T \mapsto \hat{\mathbb{P}}_T L_T^{-1} \in M_1(M_1(\mathbb{Z}^d))$ is continuous, $\{\hat{\mathbb{P}}_T L_T^{-1}\}_{T \in [0,T_1]}$ is compact. Hence, by Prohorov's theorem, there is a compact $C \subset M_1(\mathbb{Z}^d)$ with $\hat{\mathbb{P}}_T(L_T \notin C) \leq \eta$ for all $T \in [0, T_1]$. Therefore, $\hat{\mathbb{P}}_T(L_T \notin U_\varepsilon(C')) \leq \eta$ for all $T \geq 0$, where $C' \equiv C \cup K(n)$. Part (b) now follows from [9], Chapter 3, Theorem 2.2. □

5. Proof of the main theorem. For every $\varrho \in \tilde{M}_1(\mathbb{Z}^d)$ with $\lambda^\mu > 0$ for $\mu \in \varrho$ define $\zeta_\varrho = \sum_{\mu \in \varrho} \sqrt{\pi^\mu(0)}$. Using $[\mu] = \{\theta_x(\mu): x \in \mathbb{Z}^d\}$ and Lemma 4.4(c), it follows that $\zeta_{[\mu]} = \sum_{x \in \mathbb{Z}^d} \sqrt{\pi^\mu(x)}$ for every $\mu \in M_1(\mathbb{Z}^d)$ with $\lambda^\mu > 0$. For every $\varepsilon > 0$ define $K(\varepsilon) = \{\mu \in M_1(\mathbb{Z}^d): \Lambda(\mu) > (1-\varepsilon)b\}$ and $\tilde{K}(\varepsilon) = \{[\mu]: \mu \in K(\varepsilon)\}$. Note that $K(\varepsilon) = \bigcup_{x \in \mathbb{Z}^d} K(x, \varepsilon)$ by (4.11) and Lemma 2.2. It follows from Lemma 4.12 that $\zeta_\varrho < \infty$ for every $\varrho \in \tilde{K}(\varepsilon_1)$, where $\varepsilon_1 = b/(4\|V\|_1)$. Therefore, for every $\mu \in K(\varepsilon_1)$, we can define $\tilde{\pi}^\mu \in M_1(\mathbb{Z}^d)$ by $\tilde{\pi}^\mu(y) = \sqrt{\pi^\mu(y)}/\zeta_{[\mu]}$ for all $y \in \mathbb{Z}^d$. According to Lemma 4.4(g) these definitions are compatible with the ones given before Theorem 1.15. To prove the weak convergence stated in our main theorem, we need the continuity of several maps in a neighborhood of the optimal measures. Note that the set $\{\mu \in M_1(\mathbb{Z}^d): \lambda^\mu > 0\}$ contains $K(\varepsilon_1)$.

LEMMA 5.1. *As in Section* 4, *let* $\varepsilon_1 = b/(4\|V\|_1)$.

(a) *The maps* $\mu \mapsto h^\mu \in l^\infty(\mathbb{Z}^d)$ *and* $\mu \mapsto \lambda^\mu$ *are continuous on* $M_1(\mathbb{Z}^d)$.

(b) *The map* $\{\mu \in M_1(\mathbb{Z}^d): \lambda^\mu > 0\} \ni \mu \mapsto \pi^\mu \in M_1(\mathbb{Z}^d)$ *is continuous*.

(c) *The quotient topology on* $\tilde{M}_1(\mathbb{Z}^d)$ *coincides with the topology generated by the metric* (1.14).

(d) *For every* $\varepsilon > 0$ *the set* $\tilde{K}(\varepsilon)$ *is open in* $\tilde{M}_1(\mathbb{Z}^d)$.

(e) *The map* $\tilde{K}(\varepsilon_1) \ni \varrho \mapsto \zeta_\varrho \in [1, \infty)$ *is bounded and continuous*.

(f) *The map* $K(\varepsilon_1) \ni \mu \mapsto \tilde{\pi}^\mu \in M_1(\mathbb{Z}^d)$ *is continuous*.

(g) *The map* $K(\varepsilon_1) \ni \mu \mapsto \mathbb{Q}_y^\mu \in M_1(\Omega)$ *is continuous for every* $y \in \mathbb{Z}^d$.

PROOF. (a) It follows from (4.3) that

$$|\lambda^{\mu} - \lambda^{\nu}| \le \|h^{\mu} - h^{\nu}\|_{\infty} \le 8\|V\|_{\infty}\|\mu - \nu\|$$

for all $\mu, \nu \in M_1(\mathbb{Z}^d)$.

(b) Let $\{\mu_k\}_{k\in\mathbb{N}} \subset \{\mu \in M_1(\mathbb{Z}^d)\colon \lambda^{\mu} > 0\}$ converge to some μ with $\lambda^{\mu} > 0$. By Lemma 4.4(b) there exist π^{μ} and $\pi_k \equiv \pi^{\mu_k}$ for all $k \in \mathbb{N}$, and they satisfy

$$\begin{aligned}\lambda^{\mu} &\ge \langle h^{\mu}, \pi_k\rangle - J(\pi_k)\\ &\ge \langle h^{\mu_k}, \pi_k\rangle - J(\pi_k) - \|h^{\mu_k} - h^{\mu}\|_{\infty}\\ &= \lambda^{\mu_k} - \|h^{\mu_k} - h^{\mu}\|_{\infty}.\end{aligned}$$

Part (a) implies that $\lim_{k\to\infty}(\langle h^{\mu}, \pi_k\rangle - J(\pi_k)) = \lambda^{\mu}$. Hence, according to Lemma 4.4(a), the set $\{\pi_k\}_{k\in\mathbb{N}}$ is tight. Since J is continuous, every accumulation point π of $\{\pi_k\}_{k\in\mathbb{N}}$ satisfies $\langle h^{\mu}, \pi\rangle - J(\pi) = \lambda^{\mu}$. By Lemma 4.4(b), π^{μ} is the unique solution of this equation, hence $\lim_{k\to\infty}\pi_k = \pi^{\mu}$.

(c) Let τ be the quotient topology on $\tilde{M}_1(\mathbb{Z}^d)$ and let τ' be the topology induced by (1.14). The canonical projection from $M_1(\mathbb{Z}^d)$ to $\tilde{M}_1(\mathbb{Z}^d)$ is continuous with respect to τ', hence $\tau' \subset \tau$. The other way round, if $A \in \tau$, then the set $B \equiv \{\mu \in M_1(\mathbb{Z}^d)\colon [\mu] \in A\}$ is open. If $\mu \in B$, then there exists $\varepsilon > 0$ such that $U_{\varepsilon}(\mu) \subset B$. Furthermore, $U_{\varepsilon}([\mu]) = \{[\nu]\colon \nu \in U_{\varepsilon}(\mu)\} \subset A$. Hence, $\tau \subset \tau'$.

(d) Use the shift-invariance of Λ and part (c).

(e) According to part (c) it suffices to show that $K(x, \varepsilon_1) \ni \mu \mapsto \zeta_{[\mu]}$ is continuous for every $x \in \mathbb{Z}^d$. It follows from Lemma 4.12 that there exists a constant $C > 0$, independent of $x \in \mathbb{Z}^d$, such that

$$\zeta_{[\mu]} = \sum_{y\in\mathbb{Z}^d}\sqrt{\pi^{\mu}(y)} \le \sum_{y\in\mathbb{Z}^d}\sup_{\mu\in K(x,\varepsilon_1)}\sqrt{\pi^{\mu}(y)} \le C < \infty.$$

Hence, the series converges uniformly in $\mu \in K(x, \varepsilon_1)$ and the continuity follows from part (b).

(f) Combine (b) and (e).

(g) Let $x, y \in \mathbb{Z}^d$. By [2], Lemma 2.21, it suffices to consider an arbitrary continuous function $f\colon \Omega \to [0,1]$, which is F_t-measurable for some $t \ge 0$, and to show that $K(x, \varepsilon_1) \ni \mu \mapsto \mathbb{E}_y^{\mu}[f]$ is continuous. Using (4.6) and the continuity of $\mu \mapsto \langle h^{\mu}, L_t(\omega)\rangle$ for every $\omega \in \Omega$, this follows from (a) and (b). □

PROOF OF THEOREM 1.15. Proposition 4.46(b) and the continuity of the projection from $M_1(\mathbb{Z}^d)$ to $\tilde{M}_1(\mathbb{Z}^d)$ imply that $\{\hat{\mathbb{P}}_T[L_T]^{-1}\}_{T\ge 0}$ is tight and, therefore, relatively compact. According to Proposition 3.1, every accumulation point of this sequence as $T \to \infty$ is concentrated on $\tilde{K}$. Let $\Sigma \in M_1(\tilde{M}_1(\mathbb{Z}^d))$ denote such an accumulation point and let $\{T_k\}_{k\in\mathbb{N}}$ be a sequence tending to infinity and satisfying (1.16).

Let $C_{\mathbb{Z}^d}$ be the set of all functions $g\colon \mathbb{Z}^d \to [0,1]$ and let $C_{M_1(\mathbb{Z}^d)}$ be the set of all uniformly continuous functions $\psi\colon M_1(\mathbb{Z}^d) \to [0,1]$. Finally, let C_{Ω} denote the set of all continuous functions $f\colon \Omega \to [0,1]$, which are F_s-measurable for some $s \ge 0$. Note that $C_{M_1(\mathbb{Z}^d)}$ and C_{Ω} are convergence determining by [9], Theorem 3.1, and [2], Lemma 2.21, respectively. For $(\psi, f, g) \in C_{M_1(\mathbb{Z}^d)} \times C_{\Omega} \times C_{\mathbb{Z}^d}$ define the map $\Psi_{(\psi,f,g)}(\mu, \omega, x) = \psi(\mu)f(\omega)g(x)$. By [9], Proposition 4.6(b), the set of all these $\Psi_{(\psi,f,g)}$ is convergence determining for the weak convergence on $M_1(M_1(\mathbb{Z}^d) \times \Omega \times \mathbb{Z}^d)$.

To prove (1.17), choose any $(\psi, f, g) \in C_{M_1(\mathbb{Z}^d)} \times C_\Omega \times C_{\mathbb{Z}^d}$, where f is F_s-measurable for some $s \geq 0$. For $\mu \in M_1(\mathbb{Z}^d)$ with $\lambda^\mu > 0$ and $x \in \mathbb{Z}^d$, let $\mathbb{E}_x^\mu$ denote the expectation with respect to $\mathbb{Q}_x^\mu$. Define φ: $\tilde{K}(\varepsilon_1) \to [0,1]$ by

$$\varphi(\varrho) = \sum_{\mu \in \varrho} \tilde{\pi}^\mu(0)\psi(\mu)\mathbb{E}^\mu[f]\langle g, \tilde{\pi}^\mu\rangle. \tag{5.2}$$

Let χ: $\tilde{M}_1(\mathbb{Z}^d) \to [0,1]$ be a continuous function which is equal to 1 on $\tilde{K}(\varepsilon_1/2)$ and vanishes outside of the set $\tilde{K}(\varepsilon_1)$; according to Lemma 5.1(c) we may choose $\chi([\mu]) \equiv \min\{1, \max\{0, 2(\Lambda(\mu)/b - 1 + \varepsilon_1)/\varepsilon_1\}\}$. Defining $\tilde{\varphi}(\varrho) = \chi(\varrho)\varphi(\varrho)$ for all $\varrho \in \tilde{K}(\varepsilon_1)$ and $\tilde{\varphi}(\varrho) = 0$ otherwise, we obtain a function $\tilde{\varphi}$: $\tilde{M}_1(\mathbb{Z}^d) \to [0,1]$. In view of Lemma 4.4(g), it suffices for the proof of (1.17) to show that

$$\lim_{k \to \infty} \hat{\mathbb{E}}_{T_k}\left[\psi(L_{T_k})g(X_{T_k})f\right] = \int_{\tilde{K}} \tilde{\varphi}\, d\Sigma. \tag{5.3}$$

By Lemma 5.1(f) and (g), each term of the series in (5.2) is continuous on $K(\varepsilon_1)$. Note that $\zeta_{[\mu]} \geq 1$ for $\mu \in K(\varepsilon_1)$. Using $[\mu] = \{\theta_y(\mu)\colon y \in \mathbb{Z}^d\}$ and Lemma 4.4(c) to rewrite the series in (5.2), it follows from Lemma 4.12 that, for every $x \in \mathbb{Z}^d$, this series converges uniformly for $\mu \in K(x, \varepsilon_1)$. Therefore, the function $K(\varepsilon_1) \ni \mu \mapsto \varphi([\mu])$ is continuous. Since χ vanishes outside $\tilde{K}(\varepsilon_1)$, the function $M_1(\mathbb{Z}^d) \ni \mu \mapsto \tilde{\varphi}([\mu])$ is continuous, too. Hence, by Lemma 5.1(c), $\tilde{K}(\varepsilon_1) \ni \varrho \mapsto \varphi(\varrho)$ and $\tilde{M}_1(\mathbb{Z}^d) \ni \varrho \mapsto \tilde{\varphi}(\varrho)$ are continuous. Choose $\eta \in (0,1]$. Since $\{\hat{\mathbb{P}}_{T_k}[L_{T_k}]^{-1}\}_{k \in \mathbb{N}}$ satisfies (1.16), it is sufficient for the proof of (5.3) to show that, for all sufficiently large T,

$$\left|\hat{\mathbb{E}}_T[\psi(L_T)g(X_T)f] - \hat{\mathbb{E}}_T[\tilde{\varphi}([L_T])]\right| \leq 100\eta. \tag{5.4}$$

Let us now determine all relevant epsilons, cube sizes and time intervals for the proof of (5.4). Since $\tilde{\varphi}$ is continuous, there exists, for every $\varrho \in \tilde{K}$, a radius $\varepsilon_\varrho > 0$ such that $|\tilde{\varphi}(\varrho) - \tilde{\varphi}(\varrho')| \leq \eta/2$ for all $\varrho' \in U_{\varepsilon_\varrho}(\varrho)$. Since $\tilde{K}$ is compact, there exists a finite subset M of $\tilde{K}$ such that $\tilde{K}$ is covered by $\bigcup_{\varrho \in M} U_{\varepsilon_\varrho/3}(\varrho)$. Define $\varepsilon_\varphi = \min_{\varrho \in M} \varepsilon_\varrho/3$. Since $\bigcup_{\varrho \in M} U_{2\varepsilon_\varrho/3}(\varrho)$ covers $U_{\varepsilon_\varphi}(\tilde{K})$, it follows that

$$|\tilde{\varphi}(\varrho) - \tilde{\varphi}(\varrho')| \leq \eta \quad \text{for all } \varrho, \varrho' \in U_{\varepsilon_\varphi}(\tilde{K}) \text{ with } \|\varrho - \varrho'\| \leq \varepsilon_\varphi. \tag{5.5}$$

By the uniform continuity of ψ, there exists $\varepsilon_\psi > 0$ such that $|\psi(\mu) - \psi(\nu)| \leq \eta$ for all $\mu, \nu \in M_1(\mathbb{Z}^d)$ with $\|\mu - \nu\| \leq \varepsilon_\psi$. The set $K(0, \varepsilon_1)$ is open because Λ is continuous. The set $K(0)$ is compact by Proposition 1.11(c) and contained in $K(0, \varepsilon_1)$. Hence there exists $\varepsilon_\Lambda > 0$ such that $U_{\varepsilon_\Lambda}(K(0)) \subset K(0, \varepsilon_1/2)$. The shift-invariance of H and Λ then implies that $U_{\varepsilon_\Lambda}(K) \subset K(\varepsilon_1)$. With ε_0 as in Lemma 4.23, define $\varepsilon = \min\{\varepsilon_0, \varepsilon_\psi, \varepsilon_\varphi, \varepsilon_\Lambda\}$.

Since $K(0, \varepsilon_1)$ may contain shift-equivalent measures, define $n_1 = n_0 + 2$ with n_0 as in Lemma 2.3. Then, if $\mu \in K(0, \varepsilon_1)$ and $y \in \mathbb{Z}^d$ satisfy $\theta_y(\mu) \in K(0, \varepsilon_1)$, it follows that $\mu(-y) > (1 - \varepsilon_1)b/\|V\|_1 \geq 3b/(4\|V\|_1)$ by (4.11) and, on the other hand, $\mu(B^c_{n_1, 0}) \leq \varepsilon_1 + b/(2\|V\|_1) < \mu(-y)$ by Lemma 2.3. Hence $\|y\|_\infty < n_1$.

According to Proposition 4.46(a) there exist $n_2 \in \mathbb{N}$ and $T' > 0$ such that

$$\sup_{T \geq T'} \sup_{t \in [0,T]} \hat{\mathbb{P}}_T\big(L_T \notin U(\varepsilon/2, n_2, X_t)\big) \leq \eta/5. \tag{5.6}$$

Note that $\zeta_{[\mu]} \geq 1$ for $\mu \in K(\varepsilon_1)$. Hence, by Lemma 4.12, there exists $n_3 \geq n_1 + 2n_2$ such that

$$\max_{x \in B_{n_2,0}} \sup_{\mu \in K(x, \varepsilon_1)} \tilde{\pi}^{\mu}\big(B^c_{n_3,0}\big) \leq \eta. \tag{5.7}$$

Furthermore, there exists $n_4 \geq n_2 + n_3$ such that

$$\max_{x \in B_{n_2+n_3,0}} \sup_{\mu \in K(x, \varepsilon_1)} \pi^{\mu}\big(B^c_{n_4,0}\big) \leq \eta. \tag{5.8}$$

By Lemma 4.14, Lemma 4.16 and shift-invariance, there exists $t \geq 2s$ such that

$$\max_{u \in \{t/2, t\}} \sup_{x \in \mathbb{Z}^d} \sup_{\mu \in K(x, \varepsilon_1)} \max_{y, z \in B_{2n_4, x}} \left| \frac{\mathbb{Q}^{\mu}_y(X_u = z)}{\pi^{\mu}(z)} - 1 \right| \leq \eta \tag{5.9}$$

and

$$\sup_{x \in B_{n_3,0}} \sup_{\mu \in K(x, \varepsilon_1)} \|\mathbb{Q}^{\mu}_0 X^{-1}_{t/2} - \pi^{\mu}\| \leq \eta. \tag{5.10}$$

Using (4.26) and the splitting $L_T = (2t/T)L_{t,T-t,T} + (1 - 2t/T)L_{t,T-t}$, it follows in a similar way as (4.27) that

$$|TH(L_T) - Y_{t,T}| = \frac{4t^2}{T}\left|H(L_{t,T-t,T}) - \langle h^{L_{t,T-t}}, L_{t,T-t,T}\rangle\right| \leq \frac{8t^2}{T}\|V\|_\infty,$$

where $Y_{t,T} \equiv (T - 2t)H(L_{t,T-t}) + 2t\langle h^{L_{t,T-t}}, L_{t,T-t,T}\rangle$. Choose $T'' > T'$ such that $\exp(8t^2\|V\|_\infty/T'') \leq 1 + \eta$ and $T'' \geq 4t/\varepsilon$. Then, for all $T \geq T''$,

$$\left|1 - \frac{\mathbb{E}[\exp(Y_{t,T})]}{Z_T}\right| \leq \eta \tag{5.11}$$

and

$$\|L_{t,T-t} - L_T\| \leq \frac{2t}{T} \leq \frac{\varepsilon}{2}. \tag{5.12}$$

To show that (5.4) holds for all $T \geq T''$, fix any such T for the remaining part of the proof. We are now going to reduce our problem to various big cubes and decouple the time intervals $[0, t]$, $[t, T - t]$ and $[T - t, T]$.

If $x, y \in \mathbb{Z}^d$ satisfy $\|x - y\|_\infty \geq n_1 + 2n_2$, then $U(\varepsilon/2, n_2, x)$ and $U(\varepsilon/2, n_2, y)$ are disjoint by the argument which led to the choice of n_1. Hence, it follows by using $\hat{\mathbb{P}}_T(X_0 = 0) = 1$ and applying (5.6) for the five intermediate times 0, $t/2$, t, $T - t$ and T that

$$\hat{\mathbb{P}}_T\big(L_T \in U(\varepsilon/2, n_2, X_t), A_T\big) \geq 1 - \eta, \tag{5.13}$$

where

$$A_T \equiv \left\{X_{t/2} \in B_{n_4,0},\, X_t \in B_{n_3,0},\, \|X_{T-t} - X_t\|_\infty < n_1 + 2n_2,\, \|X_T - X_t\|_\infty < n_3\right\}.$$

By (5.13) it is sufficient for the proof of (5.4) to show that

$$\begin{aligned}\Big|\hat{\mathbb{E}}_T[\psi(L_T)g(X_T)f;\, L_T\in U(\varepsilon/2,n_2,X_t),A_T]\\ -\hat{\mathbb{E}}_T[\varphi([L_T]);\, L_T\in U(\varepsilon/2,n_2,X_t),A_T]\Big|\le 98\eta.\end{aligned}\tag{5.14}$$

It follows from (5.12) and (5.6) that

$$\begin{aligned}0&\le\hat{\mathbb{P}}_T(L_{t,T-t}\in U(\varepsilon,n_2,X_t))-\hat{\mathbb{P}}_T(L_T\in U(\varepsilon/2,n_2,X_t))\\ &\le\hat{\mathbb{P}}_T(L_T\notin U(\varepsilon/2,n_2,X_t))\le\eta.\end{aligned}\tag{5.15}$$

To further reduce (5.14), we first use (5.12), (5.5), $\varepsilon\le\varepsilon_\psi$ and (5.15) to replace L_T by $L_{t,T-t}$; an application of (5.11) then shows that we have to prove that

$$\begin{aligned}&\Big|\mathbb{E}[\psi(L_{t,T-t})g(X_T)f\exp(Y_{t,T});\, L_{t,T-t}\in U(\varepsilon,n_2,X_t),A_T]\\ &\quad-\mathbb{E}[\varphi([L_{t,T-t}])\exp(Y_{t,T});\, L_{t,T-t}\in U(\varepsilon,n_2,X_t),A_T]\Big|\\ &\le 92\eta Z_T.\end{aligned}\tag{5.16}$$

It remains to show that the two expectations in (5.16) are essentially the same, using the fact that the time intervals $[t/2,t]$ and $[T-t,T]$ are long enough for the new ergodic random walks to converge close to their equilibrium distributions. Using the definition of $Y_{t,T}$, the Markov property and (4.6), it follows that the first expectation in (5.16) can be rewritten as

$$\begin{aligned}&\sum_{y\in B_{n_3,0}}\int_{U(\varepsilon,n_2,y)\times B_{n_1+2n_2,y}}\psi(\mu)\exp((T-2t)H(\mu)+2t\lambda^\mu)\sqrt{\frac{\pi^\mu(0)\pi^\mu(z)}{\pi^\mu(y)}}\\ &\quad\times\mathbb{E}_0^\mu[f;\,X_{t/2}\in B_{n_4,0},X_t=y]\,\mathbb{E}_z^\mu\left[\frac{g(X_t)}{\sqrt{\pi^\mu(X_t)}};\,X_t\in B_{n_3,y}\right]\\ &\quad\times\mathbb{P}_y(L_{T-2t},X_{T-2t})^{-1}(d\mu,dz).\end{aligned}\tag{5.17}$$

It follows from (5.11) that the two expectations in (5.16) and, therefore, (5.17) are bounded above by $(1+\eta)Z_T$. Using (5.7) and (5.9), it follows that

$$\begin{aligned}&\left|\mathbb{E}_z^\mu\left[\frac{g(X_t)}{\zeta_{[\mu]}\sqrt{\pi^\mu(X_t)}};\,X_t\in B_{n_3,y}\right]-\langle g,\tilde{\pi}^\mu\rangle\right|\\ &\quad\le\tilde{\pi}^\mu(B_{n_3,y}^c)+\eta\tilde{\pi}^\mu(B_{n_3,y})\le 2\eta\end{aligned}\tag{5.18}$$

for all $y\in B_{n_3,0}$, $\mu\in U(\varepsilon,n_2,y)$ and $z\in B_{n_1+2n_2,y}$. According to Lemma 4.4(d) the measures $\{\mathbb{Q}_x^\mu\}_{x\in\mathbb{Z}^d}$ with $\mu\in U_\varepsilon(K)$ are Markovian. Using $f\le 1$, (5.8), (5.9) and (5.10), it follows that, for all $y\in B_{n_3,0}$ and $\mu\in U(\varepsilon,n_2,y)$,

$$\begin{aligned}&\left|\frac{\mathbb{E}_0^\mu[f;\,X_{t/2}\in B_{n_4,0},X_t=y]}{\pi^\mu(y)}-\mathbb{E}_0^\mu[f]\right|\\ &\le\mathbb{Q}_0^\mu(X_{t/2}\in B_{n_4,0}^c)+\sum_{x\in B_{n_4,0}}\mathbb{Q}_0^\mu(X_{t/2}=x)\left|\frac{\mathbb{Q}_x^\mu(X_{t/2}=y)}{\pi^\mu(y)}-1\right|\\ &\le\|\mathbb{Q}_0^\mu X_{t/2}^{-1}-\pi^\mu\|+\pi^\mu(B_{n_4,0}^c)+\eta\mathbb{Q}_0^\mu(X_{t/2}\in B_{n_4,0})\le 3\eta.\end{aligned}\tag{5.19}$$

Using (5.18) and (5.19) to rewrite (5.17), we obtain

$$\begin{aligned}
&\bigg|\mathbb{E}[\psi(L_{t,T-t})g(X_T)f\exp(Y_{t,T});\,L_{t,T-t}\in U(\varepsilon,n_2,X_t),\,A_T] \\
&\quad-\sum_{y\in B_{n_3,0}}\int_{U(\varepsilon,n_2,y)\times B_{n_1+2n_2,y}}\psi(\mu)\mathbb{E}_0^{\mu}[f]\langle g,\tilde{\pi}^{\mu}\rangle\tilde{\pi}^{\mu}(0) \\
&\qquad\times\zeta_{[\mu]}^4\tilde{\pi}^{\mu}(y)\tilde{\pi}^{\mu}(z)\exp((T-2t)H(\mu)+2t\lambda^{\mu}) \\
&\qquad\qquad\times\mathbb{P}_y(L_{T-2t},X_{T-2t})^{-1}(d\mu,dz)\bigg| \\
&\quad\le(5+6\eta)\eta(1+\eta)Z_T,
\end{aligned}\tag{5.20}$$

where we used that the first expectation is bounded by $(1+\eta)Z_T$. Note that the sum in (5.20) is bounded by $(1+\eta)(1+2\eta)(1+3\eta)Z_T$. Since $\mu\mapsto H(\mu)$, $\mu\mapsto\lambda^{\mu}$ and $\mu\mapsto\zeta_{[\mu]}$ are shift-invariant, and since $\tilde{\pi}^{\theta_y(\mu)}(y)=\tilde{\pi}^{\mu}(0)$ and $\tilde{\pi}^{\theta_y(\mu)}(y+z)=\tilde{\pi}^{\mu}(z)$ by Lemma 4.4(c), the sum in (5.20) can be rewritten as

$$\begin{aligned}
&\int_{U(\varepsilon,n_2,0)\times B_{n_1+2n_2,0}}\sum_{y\in B_{n_3,0}}\psi(\theta_y(\mu))\mathbb{E}_0^{\theta_y(\mu)}[f]\langle g,\tilde{\pi}^{\theta_y(\mu)}\rangle\tilde{\pi}^{\theta_y(\mu)}(0) \\
&\quad\times\zeta_{[\mu]}^4\tilde{\pi}^{\mu}(0)\tilde{\pi}^{\mu}(z)\exp((T-2t)H(\mu)+2t\lambda^{\mu}) \\
&\quad\times\mathbb{P}_0(L_{T-2t},X_{T-2t})^{-1}(d\mu,dz).
\end{aligned}\tag{5.21}$$

Using (5.2), Lemma 4.4(c) and (5.7), it follows that

$$\begin{aligned}
0&\le\varphi([\mu])-\sum_{y\in B_{n_3,0}}\psi(\theta_y(\mu))\mathbb{E}_0^{\theta_y(\mu)}[f]\langle g,\tilde{\pi}^{\theta_y(\mu)}\rangle\tilde{\pi}^{\theta_y(\mu)}(0) \\
&\le\tilde{\pi}^{\mu}(B_{n_3,0}^c)\le\eta
\end{aligned}\tag{5.22}$$

for all $\mu\in U(\varepsilon,n_2,0)$. Using (5.21) and (5.22) to rewrite (5.20), we obtain

$$\begin{aligned}
&\bigg|\mathbb{E}[\psi(L_{t,T-t})g(X_T)f\exp(Y_{t,T});\,L_{t,T-t}\in U(\varepsilon,n_2,X_t),\,A_T] \\
&\quad-\int_{U(\varepsilon,n_2,0)\times B_{n_1+2n_2,0}}\varphi([\mu])\zeta_{[\mu]}^4\tilde{\pi}^{\mu}(0)\tilde{\pi}^{\mu}(z) \\
&\qquad\times\exp((T-2t)H(\mu)+2t\lambda^{\mu}) \\
&\qquad\times\mathbb{P}_0(L_{T-2t},X_{T-2t})^{-1}(d\mu,dz)\bigg| \\
&\quad\le\eta(1+\eta)(6+11\eta+6\eta^2)Z_T.
\end{aligned}\tag{5.23}$$

The calculations leading from the first expectation in (5.16) to the estimate (5.23) are also valid for $\psi=\varphi([\cdot])$ with φ given by (5.2), when we set $g=1_{\mathbb{Z}^d}$ and $f=1_{\Omega}$ in these calculations. Therefore, (5.23) also holds with the expectation in (5.23) replaced by the second one from (5.16). Since we chose $\eta\le1$, the estimate (5.16) follows from the two versions of (5.23). □

6. Proof of Theorem 1.19. For $\beta > 0$ define a function $V\colon \mathbb{Z}^d \to [0,\infty)$, which models a Dirac-type interaction, by $V = \beta 1_{\{0\}}$. In this section we write $\mu_x \equiv \mu(x)$ for $\mu \in M_1(\mathbb{Z}^d)$ and $x \in \mathbb{Z}^d$. Define

$$K_\beta = \{\, \mu \in K\colon \mu_0 = \max_{x \in \mathbb{Z}^d} \mu_x \}.$$

LEMMA 6.1. *If $a \in [1/2, 1]$ and $\mu \in M_1(\mathbb{Z}^d)$ satisfy $\mu_x \le a$ for all $x \in \mathbb{Z}^d$, then $\sum_{x \in \mathbb{Z}^d} \mu_x^2 \le a^2 + (1-a)^2$.*

PROOF. There exist $I \subset \mathbb{Z}^d$ and $y \in \mathbb{Z}^d \setminus I$ such that $r \equiv \sum_{x \in I} \mu_x \le a$ and $s \equiv r + \mu_y \ge a$. Note that $s - a \le s - \mu_y = r$. Hence $\mu_y^2 = ((s-a) + (a-r))^2 \le (s-a)^2 + (a-r)^2 + 2r(a-r)$ and

$$\begin{aligned}\sum_{x \in \mathbb{Z}^d} \mu_x^2 &= \sum_{x \in I} \mu_x^2 + \mu_y^2 + \sum_{x \in \mathbb{Z}^d \setminus (I \cup \{y\})} \mu_x^2 \\ &\le r^2 + (s-a)^2 + (a-r)^2 + 2r(a-r) + (1-s)^2 \\ &\le a^2 + (1-a)^2. \qquad \square\end{aligned}$$

LEMMA 6.2. *If $\beta \ge 2d$ and $\mu \in K_\beta$, then $\mu_0 \ge \left(1 + \sqrt{1 - (2d/\beta)^2}\right)/2$.*

PROOF. Define $\nu \in M_1(\mathbb{Z}^d)$ by $\nu_0 = 1$ and $\nu_x = 0$ for all $x \in \mathbb{Z}^d \setminus \{0\}$. Then $H(\nu) = \beta$ and $J(\nu) = d$. Define

$$a_n = \tfrac{1}{2}\left(1 + \sqrt{1 - (2d/\beta)^{2(1-2^{-n})}}\right)$$

for all $n \in \mathbb{N}_0$. We show by induction that $\mu_0 > a_n$ for all $n \in \mathbb{N}_0$.

Assume that $\mu_0 \le a_0$. Then $H(\mu) \le \beta/2$ by Lemma 6.1. Since $J(\mu) > 0$, it follows that $\Lambda(\mu) < \beta/2$, which contradicts $\Lambda(\mu) \ge \Lambda(\nu) = \beta - d \ge \beta/2$.

Assume now that $\mu_0 \le a_{n+1}$. Then, by Lemma 6.1,

$$\begin{aligned}H(\mu) &\le 2\beta\left(a_{n+1} - \frac{1}{2}\right)^2 + \frac{\beta}{2} \\ &= \frac{\beta}{2}\left(1 - \left(\frac{2d}{\beta}\right)^{2(1-2^{-(n+1)})}\right) + \frac{\beta}{2} = \beta - d\left(\frac{2d}{\beta}\right)^{1-2^{-n}}.\end{aligned}$$

According to the induction hypothesis, $\mu_0 > a_n$. Hence $\mu_y < 1 - a_n$ for all $y \in \mathbb{Z}^d$ with $\|y\|_1 = 1$. Restricting the sum in (1.6) to all $\{0, y\} \subset \mathbb{Z}^d$ with $\|y\|_1 = 1$, it follows that

$$J(\mu) > d\left(\sqrt{a_n} - \sqrt{1 - a_n}\right)^2 = d - 2d\sqrt{a_n(1-a_n)} = d - d(2d/\beta)^{1-2^{-n}},$$

hence $\Lambda(\mu) < \beta - d$. Again, this contradicts $\Lambda(\mu) \ge \Lambda(\nu) = \beta - d$. $\square$

The result of Lemma 6.2 would be sufficient to prove Theorem 1.19 for $\beta \ge 3.1766d$. To prove $|K_\beta| = 1$ for all $\beta \ge 2d$, we need a refinement of Lemma 6.2.

LEMMA 6.3. *If* $\beta \ge 2d$ *and* $\mu \in K_\beta$, *then* $\mu_0 \ge \left(1 + \sqrt{1 - (1.19d/\beta)^2}\right)/2$.

PROOF. Define the measure $\nu \in M_1(\mathbb{Z}^d)$ by $\nu_0 = \left(1 + \sqrt{1 - d/(2\beta^2)}\right)/2$ and $\nu_x = (1 - \nu_0)/(2d)$ for all $x \in \mathbb{Z}^d$ satisfying $\|x\|_1 = 1$. Hence $\nu_x = 0$ for $\|x\|_1 > 1$. A lengthy but elementary calculation shows that

$$\Lambda(\nu) = \frac{\beta}{4d}\left(2d + 1 + (2d - 1)\sqrt{1 - \frac{d}{2\beta^2}}\right) + \frac{6d - 1}{16\beta} - d.$$

Define $\alpha_1 = 2$ and, recursively for $n \in \{1, 2, \dots, 10\}$,

$$\alpha_{n+1} = \frac{1}{\sqrt{8}}\sqrt{48 - 8\sqrt{14} + \frac{4\sqrt{14} + 8\sqrt{2d'}\,\alpha_n - 30}{d'} + \frac{1}{(d')^2} - \left(8 - \frac{4}{d'}\right)\sqrt{4 - \alpha_n^2}}\,,$$

where $d' \equiv \min\{d, 2\}$. Evaluating this numerically shows that $\alpha_1 > \alpha_2 > \alpha_3 > \cdots > \alpha_{11}$. Furthermore, $\alpha_{11} \approx 1.18075 < 1.19$, if $d = 1$, and $\alpha_{11} \approx 1.10491$, if $d \ge 2$. For every $n \in \{1, 2, \dots, 11\}$ define

$$a_n = \left(1 + \sqrt{1 - (\alpha_n d/\beta)^2}\right)/2.$$

Lemma 6.2 shows that $\mu_0 \ge a_1$.

Assume that there exists $n \in \{1, \dots, 10\}$ with $\mu_0 \in [a_n, a_{n+1}]$. Then, by Lemma 6.1,

$$H(\mu) \le 2\beta\left(a_{n+1} - \frac{1}{2}\right)^2 + \frac{\beta}{2} = \beta - \frac{\alpha_{n+1}^2 d^2}{2\beta}.$$

Since $\mu_0 \ge a_n$,

$$J(\mu) \ge \tfrac{1}{2} \sum_{\substack{x \in \mathbb{Z}^d \\ \|x\|_1 = 1}} \left(\sqrt{a_n} - \sqrt{\mu_x}\right)^2.$$

Under the restriction $\sum_{\|x\|_1 = 1} \mu_x \le 1 - a_n$, this lower bound is minimal when $\mu_x = (1 - a_n)/(2d)$ for all $x \in \mathbb{Z}^d$ with $\|x\|_1 = 1$. Therefore,

$$\begin{aligned} J(\mu) &\ge d\left(\sqrt{a_n} - \sqrt{\frac{1 - a_n}{2d}}\right)^2 \\ &= \frac{2d + 1}{4} + \frac{2d - 1}{4}\sqrt{1 - \left(\frac{\alpha_n d}{\beta}\right)^2} - \frac{1}{\sqrt{2d}}\,\frac{\alpha_n d^2}{\beta}, \end{aligned}$$

hence

$$\begin{aligned} \Lambda(\mu) - \Lambda(\nu) \le \frac{1}{16\beta}\Bigg(&\frac{4(2d - 1)\beta^2}{d}\left(1 - \sqrt{1 - \frac{d}{2\beta^2}}\right) - 8d^2\alpha_{n+1}^2 \\ &+ 8d\sqrt{2d}\,\alpha_n - 6d + 1 + 4(2d - 1)\beta\left(1 - \sqrt{1 - \left(\frac{\alpha_n d}{\beta}\right)^2}\right)\Bigg). \end{aligned}$$

Since the functions

$$\beta \mapsto \beta^2\Big(1 - \sqrt{1 - d/(2\beta^2)}\Big)$$

and

$$\beta \mapsto \beta\Big(1 - \sqrt{1 - (\alpha_n d/\beta)^2}\Big)$$

are decreasing on $[2d, \infty)$, we get an upper estimate by setting $\beta = 2d$ in these expressions. Furthermore, $\mathbb{N} \ni d \mapsto 1 - \sqrt{1 - 1/(8d)}$ is maximal for $d = 1$. Hence, $\Lambda(\mu) - \Lambda(\nu) \le d^2 f(d, n)/(16\beta)$, where

$$f(d, n) \equiv 48 - 8\sqrt{14} - 8\alpha_{n+1}^2 - \left(8 - \frac{4}{d}\right)\sqrt{4 - \alpha_n^2} + \frac{4\sqrt{14} + 8\sqrt{2d}\,\alpha_n - 30}{d} + \frac{1}{d^2}.$$

In order to show that $f(d, n) \le f(d', n)$, it suffices to show that $f(d+1, n) \le f(d, n)$ for all $d \ge 2$. Using $1/(d+1)^2 \le 1/d^2$ and

$$\frac{1}{\sqrt{d}} - \frac{1}{\sqrt{d+1}} = \frac{1}{\sqrt{d(d+1)}\,(\sqrt{d} + \sqrt{d+1})} \ge \frac{1}{2(d+1)\sqrt{d}} \ge \frac{1}{\sqrt{2}\,(d+1)d},$$

it suffices to show that $0 \le 4\sqrt{14} + 8\alpha_n + 4\sqrt{4 - \alpha_n^2} - 30$, which is in fact true for $\alpha_n \in [1.0186, 2]$.

Since $f(d', n) = 0$ by the definition of α_{n+1}, it follows that $\Lambda(\mu) - \Lambda(\nu) \le 0$. Since $\nu \notin K_\beta$ by Proposition 1.11(b), we thereby obtain a contradiction. □

Proof of Theorem 1.19. If $\beta > d$, then $H(\nu) = \beta > d$ for every Dirac measure ν. Therefore, Condition 1.10 is satisfied and $|K_\beta| \ge 1$ by Proposition 1.11(a). Assume that there exist $\mu, \tilde{\mu} \in K_\beta$ with $\mu \ne \tilde{\mu}$. Define $\varphi, \tilde{\varphi} \in l_2(\mathbb{Z}^d)$ by $\varphi_x = \sqrt{\mu_x}$ and $\tilde{\varphi}_x = \sqrt{\tilde{\mu}_x}$ for all $x \in \mathbb{Z}^d$. Define

$$\chi = \frac{\tilde{\varphi} - \langle \tilde{\varphi}, \varphi \rangle_{l_2} \varphi}{\|\tilde{\varphi} - \langle \tilde{\varphi}, \varphi \rangle_{l_2} \varphi\|_{l_2}} \quad \text{and} \quad \varepsilon_0 = \arcsin\|\tilde{\varphi} - \langle \tilde{\varphi}, \varphi \rangle_{l_2} \varphi\|_{l_2}$$

as well as $\psi\colon [0, \varepsilon_0] \to l_2(\mathbb{Z}^d)$ by $\psi(\varepsilon) = \varphi \cos \varepsilon + \chi \sin \varepsilon$ and $\nu\colon [0, \varepsilon_0] \to M_1(\mathbb{Z}^d)$ by $\nu_x(\varepsilon) = (\varphi_x \cos \varepsilon + \chi_x \sin \varepsilon)^2$ for all $x \in \mathbb{Z}^d$ and $\varepsilon \in [0, \varepsilon_0]$. Note that $\psi(\varepsilon_0) = \tilde{\varphi}$. Define $\lambda(\varepsilon) = \Lambda(\nu(\varepsilon))$ for all $\varepsilon \in [0, \varepsilon_0]$. Then

$$\lambda' = 4\beta \sum_{x \in \mathbb{Z}^d} \psi_x^3 \psi_x' - \sum_{\substack{\{x, y\} \subset \mathbb{Z}^d \\ \|x - y\|_1 = 1}} (\psi_x - \psi_y)(\psi_x' - \psi_y')$$

and, using $\psi'' = -\psi$,

$$\lambda'' = 4\beta \sum_{x\in\mathbb{Z}^d} \psi_x^2\Big(3(\psi_x')^2 - \psi_x^2\Big) + 2J\circ\nu - \sum_{\substack{\{x,y\}\subset\mathbb{Z}^d\\ \|x-y\|_1=1}} \big(\psi_x' - \psi_y'\big)^2.$$

To prove the theorem, it suffices to show that $\lambda''(\varepsilon) < 0$ for all $\varepsilon \in [0, \varepsilon_0]$.

Since $\langle \psi', \psi\rangle_{l_2} = 0$ and $\|\psi\|_{l_2} = \|\psi'\|_{l_2} = 1$, it follows that, for every $x \in \mathbb{Z}^d$,

$$|\psi_x'\psi_x| = \bigg|\sum_{y\in\mathbb{Z}^d\setminus\{x\}} \psi_y'\psi_y\bigg| \le \sqrt{1-(\psi_x')^2}\sqrt{1-\psi_x^2}\,.$$

Squaring and solving for $(\psi_x')^2$ yields $(\psi_x')^2 \le 1 - \psi_x^2$. Since $J \le d$ on $M_1(\mathbb{Z}^d)$, it follows that $\lambda'' < 2d + 12\beta - 16H\circ\nu$. Since the estimate in Lemma 6.3 holds for μ_0 and $\tilde{\mu}_0$, it is also valid for $\nu_0(\varepsilon)$ with $\varepsilon \in [0, \varepsilon_0]$. Therefore,

$$\begin{aligned}\lambda'' &< 2d + 12\beta - 16\beta\nu_0^2\\ &\le 2d + 4(1.19^2d^2)/\beta - 4\beta\Big(2\sqrt{1-(1.19d/\beta)^2} - 1\Big).\end{aligned}$$

For $\beta \ge 2.38d/\sqrt{3} \approx 1.3741d$ this upper bound is obviously decreasing in β and it is negative for $\beta = 2d$. Hence, $\lambda'' < 0$ for all $\beta \ge 2d$. □

Acknowledgments. We would like to thank D. C. Brydges, J.-D. Deuschel and G. Slade for discussions about our model and related ones.

REFERENCES

[1] BOLTHAUSEN, E. (1994). Localization of a two-dimensional random walk with an attractive path interaction. *Ann. Probab.* **22** 875–918.

[2] BOLTHAUSEN, E., DEUSCHEL, J.-D. and SCHMOCK, U. (1993). Convergence of path measures arising from a mean field or polaron type interaction. *Probab. Theory Related Fields* **95** 283–310.

[3] BRYDGES, D. C. and SLADE, G. (1995). The diffusive phase of a model of self-interacting walks. *Probab. Theory Related Fields* **103** 285–315.

[4] CHEN, M. F. (1992). *From Markov Chains to Non-Equilibrium Particle Systems*. World Scientific, Singapore.

[5] DAVIS, B. (1990). Reinforced random walk. *Probab. Theory Related Fields* **84** 203–229.

[6] DEUSCHEL, J.-D. and STROOCK, D. W. (1989). *Large Deviations*. Academic Press, San Diego.

[7] DONSKER, M. D. and VARADHAN, S. R. S. (1975). Asymptotic evaluation of certain Markov process expectations for large time. I. *Comm. Pure Appl. Math.* **28** 1–47.

[8] DONSKER, M. D. and VARADHAN, S. R. S. (1976). Asymptotic evaluation of certain Markov process expectations for large time. III. *Comm. Pure Appl. Math.* **29** 389–461.

[9] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes, Characterization and Convergence*. Wiley, New York.

[10] MADRAS, N. and SLADE, G. (1993). *The Self-Avoiding Walk*. Birkhäuser, Boston.

[11] MANSMANN, U. (1991). The free energy of the Dirac polaron, an explicit solution. *Stochastics* **34** 93–125.

[12] NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge Univ. Press.

[13] PEMANTLE, R. (1988). Phase transition in reinforced random walk and RWRE on trees. *Ann. Probab.* **16** 1229–1241.

[14] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces*. Academic Press, New York.

[15] ROSS, S. M. (1985). *Introduction to Probability Models*. Academic Press, San Diego.

[16] SCHMOCK, U. (1989). Convergence of one-dimensional Wiener sausage path measures to a mixture of Brownian taboo processes. *Stochastics* **29** 203–220.

[17] SZNITMAN, A.-S. (1991). On the confinement property of two-dimensional Brownian motion among Poissonian obstacles. *Comm. Pure Appl. Math.* **44** 1137–1170.

ANGEWANDTE MATHEMATIK
UNIVERSITÄT ZÜRICH
WINTERTHURER STRASSE 190
CH-8057 ZÜRICH
SWITZERLAND
E-MAIL: eb@amath.unizh.ch

DEPARTEMENT MATHEMATIK
ETH-ZENTRUM
CH-8092 ZÜRICH
SWITZERLAND
E-MAIL: schmock@math.ethz.ch