ON SELF-ATTRACTING RANDOM WALKS

Erwin Bolthausen and Uwe Schmock

Universität Zürich

Abstract. In this survey paper we mainly discuss the results contained in two of our recent articles [2] and [5]. Let $\{X_t\}_{t\geq 0}$ be a continuous-time, symmetric, nearest-neighbour random walk on $\mathbb{Z}^d$. For every $T > 0$ we define the transformed path measure $\mathrm{d}\widehat{\mathbb{P}}_T = (1/Z_T)\exp(H_T)\,\mathrm{d}\mathbb{P}$, where $\mathbb{P}$ is the original one and Z_T is the appropriate normalizing constant. The Hamiltonian H_T imparts the self-attracting interaction of the paths up to time T. We consider the case where H_T is given by a potential function V on $\mathbb{Z}^d$ with finite support, and the case $H_T = -N_T$, where N_T denotes the number of points visited by the random walk up to time T. In both cases the typical paths under $\widehat{\mathbb{P}}_T$ as $T \to \infty$ clump together much more than those of the free random walk and give rise to localization phenomena.

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 $t \in [0,\infty)$ let $X_t(\omega) \equiv \omega(t)$ for $\omega \in \Omega$ be the evaluation map. The space Ω is equipped with the σ-algebra $\mathcal{A}$ generated by $\{X_t\}_{t\geq 0}$. Let $\mathbb{P}$ be the unique probability measure on $(\Omega, \mathcal{A})$ such that $\{X_t\}_{t\geq 0}$ is a continuous-time, symmetric, nearest-neighbour random walk on $\mathbb{Z}^d$, starting at the origin, with exponential holding times of expectation $1/d$.

The interactions we consider are special cases of the following type: For every $T > 0$ let $H_T : \Omega \to \mathbb{R}$ denote a "Hamiltonian", where H_T only depends on the restriction of the path to $[0,T]$. We consider the transformed probability measures

$$\widehat{\mathbb{P}}_T(A) = \mathbb{E}[1_A \exp(H_T)]/Z_T, \quad A \in \mathcal{A},\, T > 0, \tag{1.1}$$

where Z_T is the appropriate normalizing constant given by

$$Z_T \equiv \mathbb{E}[\exp(H_T)]. \tag{1.2}$$

The main interest is in the limiting behaviour as T tends to infinity. In this context the mean square displacement $\widehat{\mathbb{E}}_T[\|X_T\|_2^2]$, where $\widehat{\mathbb{E}}_T$ denotes the expectation with respect to $\widehat{\mathbb{P}}_T$, indicates whether there is a transition to superdiffusive behaviour

1991 *Mathematics Subject Classification.* 60F10 (primary), 60F05 and 60K35 (secondary).

Key words and phrases. large deviations – attractive interaction – d-dimensional random walk – maximum entropy principle – weak convergence – localization – confinement.

Research supported by the Swiss National Foundation, Contract Nr. 21-298333.90.

This is a survey paper. Details of the proofs of the results are published in [2] and [5].

or to localization phenomena. Note that the process $\{X_t\}_{t\in[0,T]}$ under $\widehat{\mathbb{P}}_T$ is, in general, not Markovian.

The best-known example is the weakly self-avoiding (or self-repellent) random walk with the Hamiltonian

$$H_T \equiv -\int_0^T \int_0^T V(X_s - X_t)\,\mathrm{d}s\,\mathrm{d}t, \tag{1.3}$$

where the Dirac-type potential $V : \mathbb{Z}^d \to [0,\infty)$ for some coupling constant $\beta > 0$ is given by

$$V(x) \equiv \begin{cases} \beta & \text{for } x = 0, \\ 0 & \text{for } x \in \mathbb{Z}^d \setminus \{0\}. \end{cases} \tag{1.4}$$

This model has been intensively investigated, see the monograph by Madras and Slade [15] and the extensive bibliography therein. For an analog one-dimensional discrete-time model there are also recent results by König [13, 14]. The model with the Hamiltonian (1.3) becomes self-attracting when we simply change the sign. However, it is easy to see that in this case the attraction is too strong for an interesting result. As $T \to \infty$, the measures $\{\widehat{\mathbb{P}}_T\}_{T>0}$ converge to the one-point measure at the path which remains at the starting point. In order to get an interesting limiting behaviour, one has to divide the interaction by T, see (1.7) below.

In Section 2 we model the self-attracting interaction by a more general potential function $V : \mathbb{Z}^d \to [0,\infty)$ such that the radius

$$R_V \equiv \sup\{\, \|x\|_1 : x \in \mathbb{Z}^d, V(x) \neq 0 \,\} \tag{1.5}$$

of its support is finite. To exclude the trivial case, we assume that $V \neq 0$. We define the Hamiltonian corresponding to V by

$$H(\mu) \equiv \sum_{x,y\in\mathbb{Z}^d} V(x-y)\mu(x)\mu(y), \quad \mu \in \mathcal{M}_1(\mathbb{Z}^d), \tag{1.6}$$

where $\mathcal{M}_1(\mathbb{Z}^d)$ denotes the set of probability measures on $\mathbb{Z}^d$. Without loss of generality we may and will assume in the following that V is symmetric in the sense that $V(x) = V(-x)$ for all $x \in \mathbb{Z}^d$. Note that

$$TH(L_T) = \frac{1}{T}\int_0^T \int_0^T V(X_s - X_t)\,\mathrm{d}s\,\mathrm{d}t, \quad T > 0, \tag{1.7}$$

where $\{L_T\}_{T>0}$ denotes the empirical distribution process of $\{X_t\}_{t\geq 0}$ defined by

$$\Omega \times (0,\infty) \ni (\omega, T) \mapsto L_T(\omega) = \frac{1}{T}\int_{[0,T)} \delta_{X_t(\omega)}\,\mathrm{d}t \in \mathcal{M}_1(\mathbb{Z}^d). \tag{1.8}$$

Our aim is to investigate the limiting behaviour as $T \to \infty$ of the transformed probability measures $\{\widehat{\mathbb{P}}_T\}_{T>0}$ arising from (1.1) with $H_T \equiv TH(L_T)$.

Our results in [5], which are discussed in Section 2, depend on the value

$$b \equiv \liminf_{T\to\infty} \frac{1}{T}\log \mathbb{E}[\exp(TH(L_T))]. \tag{1.9}$$

Since H is nonnegative, the relation $b \geq 0$ is clear. If $b > 0$, then the attraction is strong enough for a collapse transition. We prove in [5] that the sets $\{\widehat{\mathbb{P}}_T L_T^{-1}\}_{T\geq 1}$ and $\{\widehat{\mathbb{P}}_T X_T^{-1}\}_{T\geq 1}$ as well as $\{\widehat{\mathbb{P}}_T\}_{T\geq 1}$ are tight. Furthermore, we characterize the possible accumulation points of these measures as $T \to \infty$. For the special Dirac-type potential given by (1.4) with $\beta \geq 2d$, we can even prove convergence of the above sequences as $T \to \infty$ and we can quite explicitely describe the limiting measures.

For $d \geq 2$ there seems to be a phase transition in the model with the Dirac-type potential given by (1.4). For a sufficiently small coupling constant $\beta > 0$ in (1.4), Brydges and Slade [6], who actually work on the discrete-time random walk, show that $b = 0$ in two or more dimensions. If $d \geq 3$ and if the coupling constant $\beta > 0$ is sufficiently small (which implies that $b = 0$), they can prove a central limit theorem for $\{\widehat{\mathbb{P}}_T(X_T/\sqrt{T})^{-1}\}_{T>0}$ as $T \to \infty$, which means that the diffusive behaviour persists. Their approach for $d \geq 3$ is based on the lace expansion. For the potential (1.4) the limit in (1.9) exists by a subadditivity argument. As far as we know, it is still an open problem to prove that the diffusive behaviour persists for *all* coupling constants $\beta > 0$ which yield $b = 0$ via (1.9).

In Section 3 we model the self-attracting interaction in a slightly different way. Let

$$D_T \equiv \{\, x \in \mathbb{Z}^d \colon X_t = x \text{ for some } t \in [0,T] \,\}$$

denote the set of points visited by $\{X_t\}_{t\geq 0}$ up to time $T > 0$ and $N_T \equiv |D_T|$ the cardinality of D_T. Note that N_T coincides with the cardinality of the support of L_T. We consider the transformed probability measures $\{\widehat{\mathbb{P}}_T\}_{T>0}$ arising via (1.1) with $H_T \equiv -N_T$. We do not introduce a positive coupling constant here, because the model does not exhibit a phase transition. Under the measure $\widehat{\mathbb{P}}_T$ it is preferable for the process $\{X_t\}_{t\in[0,T]}$ to visit only a small number of points without choosing a path which is too improbable with respect to $\mathbb{P}$. A strategy to do this is to stay longer at places where the process has already been and to return to these points more often than to venture out to unknown territory. It turns out that for $d = 2$ and in the limit $T \to \infty$ the process $\{X_t\}_{t\in[0,T]}$ is localized in a disk with an appropriate radius, where the center of the disk is random.

The Donsker–Varadhan large deviation theory plays a crucial rôle for the solution of the problems in the following two sections. Since the transition probability functions of our random walk are symmetric with respect to the counting measure on $\mathbb{Z}^d$, we can use the generator $Q \equiv (q_{x,y})_{x,y\in\mathbb{Z}^d}$ of our random walk, given by

$$q_{x,y} \equiv \begin{cases} -d, & \text{if } x = y, \\ 1/2, & \text{if } \|x-y\|_1 = 1, \\ 0, & \text{if } \|x-y\|_1 > 1, \end{cases}$$

and the corresponding Dirichlet form to obtain via [7, (4.2.49)] or [8, Theorem 5] the rate function $J \colon \mathcal{M}_1(\mathbb{Z}^d) \to [0,\infty)$, namely

$$J(\mu) \equiv d - \sum_{\substack{\{x,y\}\subset\mathbb{Z}^d \\ \|x-y\|_1=1}} \sqrt{\mu(x)\mu(y)} = \frac{1}{2} \sum_{\substack{\{x,y\}\subset\mathbb{Z}^d \\ \|x-y\|_1=1}} \bigl(\sqrt{\mu(x)} - \sqrt{\mu(y)}\bigr)^2 \tag{1.10}$$

for all $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$. Donsker and Varadhan have shown in [10] that the family of measures $\{\mathbb{P}L_T^{-1}\}_{T>0}$ satisfies a weak large deviation principle with rate function J, i. e.,

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{P}(L_T \in C) \le - \inf_{\mu\in C} J(\mu) \tag{1.11}$$

for every compact subset C of $\mathcal{M}_1(\mathbb{Z}^d)$ and

$$\liminf_{T\to\infty} \frac{1}{T} \log \mathbb{P}(L_T \in U) \ge - \inf_{\mu\in U} J(\mu) \tag{1.12}$$

for every open subset U of $\mathcal{M}_1(\mathbb{Z}^d)$.

2. Self-attraction caused by a potential

In this section we treat the transformed probability measures $\{\widehat{\mathbb{P}}_T\}_{T>0}$, which arise via (1.1) with $H_T \equiv TH(L_T)$, where H is defined by (1.6). Since our random walk is not even recurrent for $d \ge 3$, the measures $\{\mathbb{P}L_T^{-1}\}_{T>0}$ cannot satisfy a full large deviation principle with rate function J. Therefore, we cannot use Varadhan's theorem [7, Theorem 2.1.10] directly to prove the following proposition, which is nevertheless correct and describes the value b defined in (1.9).

Proposition 2.1. *The limit b in* (1.9) *exists and satisfies*

$$b = \sup_{\mu\in\mathcal{M}_1(\mathbb{Z}^d)} \Lambda(\mu), \quad \textit{where} \quad \Lambda \equiv H - J.$$

As explained in the introduction, we need a sufficiently strong self-attraction such that the paths under $\widehat{\mathbb{P}}_T$ tend to clump together much more than under the measure $\mathbb{P}$ of the free walk. Therefore, we assume in the following:

Condition 2.2. *Let the potential V be chosen such that $b > 0$.*

Let $K \equiv \{\, \mu \in \mathcal{M}_1(\mathbb{Z}^d) : \Lambda(\mu) = b \,\}$ be the set of optimal measures. We can show that K is a nonempty closed subset of $\mathcal{M}_1(\mathbb{Z}^d)$, when $\mathcal{M}_1(\mathbb{Z}^d)$ is equipped with the total variation distance $\|\mu-\nu\| \equiv \frac{1}{2}\sum_{x\in\mathbb{Z}^d} |\mu(x)-\nu(x)|$. For every $x \in \mathbb{Z}^d$ we define the shift transformation θ_x by $\theta_x(\mu)(y) \equiv \mu(y-x)$ for all $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$ and $y \in \mathbb{Z}^d$. Note that H and J are shift-invariant, hence $\Lambda \circ \theta_x = \Lambda$ for all $x \in \mathbb{Z}^d$. Therefore, if $\mu \in K$, then the whole shift-equivalence class $[\mu] \equiv \{\, \theta_x(\mu) : x \in \mathbb{Z}^d \,\}$ is contained in K and K cannot be compact. Let $\widetilde{\mathcal{M}}_1(\mathbb{Z}^d) \equiv \{\, [\mu] : \mu \in \mathcal{M}_1(\mathbb{Z}^d) \,\}$ denote the set of shift-equivalence classes and let $\widetilde{K} \equiv \{\, [\mu] : \mu \in K \,\}$ denote the subset of optimal ones.

Using a perturbation argument, we can show that $\mu(x) > 0$ for every $x \in \mathbb{Z}^d$ and $\mu \in K$. Hence, for every $\mu \in K$, there exists a unique Markovian path measure $\mathbb{Q}^\mu$ on $(\Omega, \mathcal{A})$ with $\mathbb{Q}^\mu(X_0 = 0) = 1$, whose conservative generator $Q^\mu \equiv (q^\mu_{x,y})_{x,y\in\mathbb{Z}^d}$ is determined by

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

With this notation we can state our main theorem for the special case where $\widetilde{K}$ contains just one shift-equivalence class.

Theorem 2.3. *If $\widetilde{K} = \{\varrho\}$ for some $\varrho \in \widetilde{\mathcal{M}}_1(\mathbb{Z}^d)$, then*

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T L_T^{-1} = \frac{1}{\zeta_\varrho} \sum_{\mu\in\varrho} \sqrt{\mu(0)}\, \delta_\mu, \tag{2.4}$$

where $\zeta_\varrho \equiv \sum_{\mu\in\varrho} \sqrt{\mu(0)} < \infty$,

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T = \frac{1}{\zeta_\varrho} \sum_{\mu\in\varrho} \sqrt{\mu(0)}\, \mathbb{Q}^\mu, \tag{2.5}$$

and

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T X_T^{-1} = \frac{1}{\zeta_\varrho^2} \sum_{\mu\in\varrho} \sqrt{\mu(0)}\sqrt{\mu}. \tag{2.6}$$

The limit in (2.6) denotes convergence with respect to the weak topology on $\mathcal{M}_1(\mathbb{Z}^d)$. Note that the total variation distance is a metric for the weak topology on $\mathcal{M}_1(\mathbb{Z}^d)$. Since $(\mathcal{M}_1(\mathbb{Z}^d), \|\cdot\|)$ is a Polish space, we can consider the weak topology on $\mathcal{M}_1(\mathcal{M}_1(\mathbb{Z}))$, which is used in (2.4). The space $\Omega = D([0,\infty), \mathbb{Z}^d)$ is equipped with the standard Skorohod metric [12, Chapter 3, (5.2)], which turns Ω into a Polish space with Borel σ-algebra $\mathcal{A}$ [12, Chapter 3, Theorem 5.6 and Proposition 7.1]. The corresponding weak topology on $\mathcal{M}_1(\Omega)$ is used in (2.5).

For every $\mu \in K$ we can show the existence of some $x \in \mathbb{Z}^d$ with $\mu(x) \geq b/\|V\|_1$ and the existence of a constant $c > 0$ such that

$$\theta_x(\mu)(y) \leq 4\exp(-c\|y\|_1) \tag{2.7}$$

for all $y \in \mathbb{Z}^d$, where c only depends on b, R_V, $\|V\|_1$, and $\|V\|_\infty$. Hence, the maximizing measures have a uniform exponential decay. Since $\zeta_\varrho = \sum_{x\in\mathbb{Z}^d} \sqrt{\mu(x)}$ for every $\mu \in \varrho$, the relation $\zeta_\varrho < \infty$ in Theorem 2.3 follows easily from (2.7).

Let us comment on the measures appearing as limits in Theorem 2.3. First note that, because $\zeta_\varrho = \sum_{x\in\mathbb{Z}^d} \sqrt{\mu(x)}$ for every $\mu \in \varrho$, the measures $\{\sqrt{\mu}/\zeta_\varrho\}_{\mu\in\varrho}$ are indeed in $\mathcal{M}_1(\mathbb{Z}^d)$. The series over $\mu \in \varrho$ with the weights $\sqrt{\mu(0)}/\zeta_\varrho$ represents in all three cases a convex combination (or mixture) of the probability measures $\{\delta_\mu\}_{\mu\in\varrho}$, $\{\mathbb{Q}^\mu\}_{\mu\in\varrho}$, or $\{\sqrt{\mu}/\zeta_\varrho\}_{\mu\in\varrho}$, respectively. For $\mu \in K$ define $DH(\mu) : \mathbb{Z}^d \to [0,\infty)$, which can be interpreted as the derivative of H at μ, by

$$DH(\mu)(x) = 2\sum_{y\in\mathbb{Z}^d} V(x-y)\mu(y), \quad x \in \mathbb{Z}^d. \tag{2.8}$$

The path measure $\mathbb{Q}^\mu$ on $(\Omega, \mathcal{A})$ then arises from $\mathbb{P}$ via the formula

$$\mathbb{Q}^\mu(A) = \frac{\mathrm{e}^{-\lambda^\mu t}}{\sqrt{\mu(0)}} \mathbb{E}\left[1_A \exp\left(\int_0^t DH(\mu)(X_s)\,\mathrm{d}s\right)\sqrt{\mu(X_t)}\right],$$

where $t \in [0,\infty)$, $A \in \sigma(X_s : s \in [0,t])$, and $\lambda^\mu \equiv H(\mu) + b$. This representation of $\mathbb{Q}^\mu$ indicates how the square roots in Theorem 2.3 appear.

We want to discuss now Condition 2.2 and the hypothesis of Theorem 2.3. If there exists $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$ with $H(\mu) > d$, then Condition 2.2 is satisfied, because the first representation in (1.10) shows that J is bounded above by d and, therefore, $b \geq H(\mu) - J(\mu) > 0$. It is easy to prove that Condition 2.2 always holds in one dimension:

Lemma 2.9. *If $d = 1$, then $b > 0$.*

Proof. There exists $k \in \mathbb{N}_0$ with $V(k) > 0$. Choose an even number $n \in \mathbb{N}$ such that $(n-k)V(k) \geq 96$. Define $\mu \in \mathcal{M}_1(\mathbb{Z})$ by $\mu(i) = N^{-1}(\max\{0, 1 - |i|/n\})^2$ for all $i \in \mathbb{Z}$, where $N \equiv 1 + 2\sum_{i=1}^{n-1} i^2/n^2 \leq 2n$. Then $J(\mu) = 2/(nN)$ and

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

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

In general it is a quite delicate problem to decide whether $\widetilde{K}$ contains just one element or not. For a sufficiently strong Dirac-type interaction, we can solve this problem:

Theorem 2.10. *For $\beta > 0$ define V by (1.4). If $\beta \geq 2d$, then $|\widetilde{K}| = 1$.*

Remark 2.11. The corresponding variational problem for the Brownian motion on the real line is

$$\sup\Big\{\beta \int_{\mathbb{R}} f^4(x)\,\mathrm{d}x - \frac{1}{2}\int_{\mathbb{R}} |\nabla f|^2\,\mathrm{d}x \colon f \in H^1(\mathbb{R}),\, \|f\|_{L_2} = 1\Big\}.$$

For every $\beta > 0$ it has a solution which can easily be determined explicitely, see [16]. 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. Furthermore, an additional argument would be needed to exclude the possibility of several symmetric, monotonely decreasing solutions. We expect that the uniqueness in Theorem 2.10 holds for all $\beta > 0$ in the one-dimensional case, but our method does not allow us to reach zero.

Let us state our general result which includes the case $|\widetilde{K}| \geq 2$. Let

$$K(0) \equiv \{\, \mu \in K \colon \mu(0) > b/\|V\|_1 \,\} \tag{2.12}$$

denote the set of optimal measures which have substantial mass at the origin. It follows from (2.7) that $K(0)$ is compact. One can show that $K(0)$ contains at least one representative of every equivalence class of $\widetilde{K}$, i. e.,

$$K = \{\, \theta_x(\mu) \colon \mu \in K(0),\, x \in \mathbb{Z}^d \,\}.$$

If we equip $\widetilde{\mathcal{M}}_1(\mathbb{Z}^d)$ with the metric

$$\|[\mu] - [\nu]\| = \inf_{x \in \mathbb{Z}^d} \|\mu - \theta_x(\nu)\|, \quad \mu, \nu \in \mathcal{M}_1(\mathbb{Z}^d),$$

then the projection from $\mathcal{M}_1(\mathbb{Z}^d)$ to $\widetilde{\mathcal{M}}_1(\mathbb{Z}^d)$ is continuous and $\widetilde{K}$ as the image of $K(0)$ is compact. In the general case, every accumulation point is characterized by a probability measure Σ on this compact set $\widetilde{K}$, where Σ describes the weights in the convex combination of the kind of measures appearing on the right-hand sides of (2.4), (2.5), and (2.6).

Theorem 2.13. *The set* $\{\widehat{\mathbb{P}}_T[L_T]^{-1}\}_{T\geq 1}$ *is relatively compact in* $\mathcal{M}_1(\widetilde{\mathcal{M}}_1(\mathbb{Z}^d))$. *Every weak accumulation point* Σ *of* $\{\widehat{\mathbb{P}}_T[L_T]^{-1}\}_{T\geq 1}$ *as* $T\to\infty$ *is concentrated on* $\widetilde{K}$. *If*

$$\lim_{n\to\infty}\widehat{\mathbb{P}}_{T_n}[L_{T_n}]^{-1}=\Sigma\in\mathcal{M}_1(\widetilde{K}) \tag{2.14}$$

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

$$\lim_{n\to\infty}\widehat{\mathbb{P}}_{T_n}L_{T_n}^{-1}=\int_{\tilde{K}}\frac{1}{\zeta_\varrho}\sum_{\mu\in\varrho}\sqrt{\mu(0)}\,\delta_\mu\,\Sigma(\mathrm{d}\varrho), \tag{2.15}$$

$$\lim_{n\to\infty}\widehat{\mathbb{P}}_{T_n}=\int_{\tilde{K}}\frac{1}{\zeta_\varrho}\sum_{\mu\in\varrho}\sqrt{\mu(0)}\,\mathbb{Q}^\mu\,\Sigma(\mathrm{d}\varrho), \tag{2.16}$$

and

$$\lim_{n\to\infty}\widehat{\mathbb{P}}_{T_n}X_{T_n}^{-1}=\int_{\tilde{K}}\frac{1}{\zeta_\varrho^2}\sum_{\mu\in\varrho}\sqrt{\mu(0)}\sqrt{\mu}\,\Sigma(\mathrm{d}\varrho), \tag{2.17}$$

where ζ_ϱ *is defined as in Theorem 2.3.*

If there is only one accumulation point Σ of $\{\widehat{\mathbb{P}}_T[L_T]^{-1}\}_{T\geq 1}$ as $T\to\infty$, then we actually have convergence of the whole sequences. Besides the case $|\widetilde{K}|=1$, which was treated in Theorem 2.3, we do not have any criterion for this to occur.

In the following subsections we will give the main ideas and arguments, which lead to a proof of Theorem 2.13. For the actual proof of the convergence in (2.15), (2.16), and (2.17), we refer the reader to [4, Section 2] or to the corresponding survey in [3, Section 3]. Both papers do not directly apply to the present situation because $\mathbb{Z}^d$ is not compact, but the ideas apply when the tube problem is solved and tightness is shown.

The tube problem. In the case $d=1$ we could visualize K as an unbounded cylinder over the compact base set $\widetilde{K}$. If $|\widetilde{K}|=1$, then K is just a (discrete) line. If $\varepsilon>0$, then the ε-neighbourhood $U_\varepsilon(K)$, taken with respect to the total variation distance, is a kind of a tube. The tube problem is to show that L_T stays in $\widehat{\mathbb{P}}_T$-law inside $U_\varepsilon(K)$ as $T\to\infty$, i. e., to prove

Proposition 2.18. *For every* $\varepsilon>0$,

$$\lim_{T\to\infty}\frac{1}{T}\log\widehat{\mathbb{P}}_T(L_T\notin U_\varepsilon(K))<0.$$

The difficulty in proving this proposition is coming from the fact that there is only a weak large deviation principle for $\{\mathbb{P}L_T^{-1}\}_{T>0}$ at our disposal. This means that the upper bound in (1.11) does not hold for all closed subsets of $\mathcal{M}_1(\mathbb{Z}^d)$. The usual way to circumvent this obstacle is to roll up the state space of the random walk $\{X_t\}_{t\geq 0}$ to create a compact one.

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 $\pi_l\colon\mathbb{Z}^d\to\mathbb{Z}_l^d$ 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 $\mathcal{M}_1(\mathbb{Z}_l^d)$ with the total variation distance. It follows from [7, Theorem 4.2.58] that the measures $\{\mathbb{P}(L_T^l)^{-1}\}_{T>0}$ with

$$L_T^l\equiv\frac{1}{T}\int_{[0,T)}\delta_{X_t^l}\,\mathrm{d}t,\quad T>0,$$

satisfy a full large deviation principle with the good rate function

$$J_l(\mu) \equiv \frac{1}{2} \sum_{x \in \mathbb{Z}_l^d} \sum_{i=1}^{d} \bigl(\sqrt{\mu(x)} - \sqrt{\mu(x+e_i)}\bigr)^2, \quad \mu \in \mathcal{M}_1(\mathbb{Z}_l^d),$$

where $e_i \equiv (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{Z}_l^d$ with the 1 at position i. We can define the potential function V_l on the torus by $V_l(x) \equiv \sum_{y \in \pi_l^{-1}(x)} V(y)$ for all $x \in \mathbb{Z}_l^d$ and the corresponding Hamiltonian by

$$H_l(\mu) \equiv \sum_{x,y \in \mathbb{Z}_l^d} V_l(x-y)\mu(x)\mu(y), \quad \mu \in \mathcal{M}_1(\mathbb{Z}_l^d).$$

The full large deviation principle implies via Varadhan's theorem that the limit

$$b_l \equiv \lim_{T \to \infty} \frac{1}{T} \log \mathbb{E}[\exp(T H_l(L_T^l))]$$

exists and is given by $b_l = \sup_{\mu \in \mathcal{M}_1(\mathbb{Z}_l^d)} \Lambda_l(\mu)$, where $\Lambda_l \equiv H_l - J_l$. Since Λ_l is continuous and $\mathcal{M}_1(\mathbb{Z}_l^d)$ is compact, the set

$$K_l \equiv \{\, \mu \in \mathcal{M}_1(\mathbb{Z}_l^d) \colon \Lambda_l(\mu) = b_l \,\}$$

of optimal measures on the discrete torus is nonvoid and compact.

If $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$, then $\mu^l \equiv \mu \pi_l^{-1}$ is in $\mathcal{M}_1(\mathbb{Z}_l^d)$. It is easy to show the inequalities $H_l(\mu^l) \geq H(\mu)$ and $J_l(\mu^l) \leq J(\mu)$ for all $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$. Therefore, $b_l \geq b$. If μ is the uniform distribution on $\mathbb{Z}_l^d$, then $J_l(\mu) = 0$ and $H_l(\mu) = \|V\|_1 / l^d$. Hence, for $l > (\|V\|_1/b)^{1/d}$, the uniform distribution on $\mathbb{Z}_l^d$ is not optimal. This already indicates that, for large l, the optimal measures are essentially concentrated on small regions of $\mathbb{Z}_l^d$; this can in fact be proved.

For every $\mu \in \mathcal{M}_1(\mathbb{Z}_l^d)$ there exists a minimal $z \in \{1, 2, \ldots, l\}^d \subset \mathbb{Z}^d$ with respect to the lexicographic order such that, if z is taken as the center of a chart of the torus $\mathbb{Z}_l^d$ and $\bar{\mu} \in \mathcal{M}_1(\mathbb{Z}^d)$ is defined to coincide with μ on this chart and to be zero otherwise, then $\Lambda(\bar{\mu})$ is maximal (with respect to all other possible choices of z). If $\mu \in \mathcal{M}_1(\mathbb{Z}_l^d)$ is optimal and therefore concentrated in a small region, then $\pi_l(z)$ will be in this region and $\Lambda(\bar{\mu})$ will be close to $\Lambda_l(\mu)$. Concerning the relation between the optimal measures on $\mathbb{Z}^d$ and $\mathbb{Z}_l^d$, we can prove for small $\varepsilon > 0$ and all sufficiently large $l \in \mathbb{N} \setminus \{1\}$:

(1) If $\mu \in \mathcal{M}_1(\mathbb{Z}_l^d)$ is nearly optimal in the sense that $\Lambda_l(\mu) \geq (1-\varepsilon) b_l$, then the cutting procedure leads to a nearly optimal measure $\bar{\mu} \in \mathcal{M}_1(\mathbb{Z}^d)$ in the sense that $\Lambda(\bar{\mu}) \geq \Lambda_l(\mu) - 4\varepsilon$. This implies the estimate $b_l \leq b + 4\varepsilon$ and, therefore, gives the missing upper bound for Proposition 2.1.

(2) Project the optimal measures in K from $\mathbb{Z}^d$ to the torus $\mathbb{Z}_l^d$. Then the optimal measures K_l on the torus are close to the projected ones in the sense that $K_l \subset U_{\varepsilon,l} \equiv U_\varepsilon(\{\mu^l\}_{\mu \in K})$. Furthermore,

$$\sup\{\, \Lambda_l(\mu) \colon \mu \in \mathcal{M}_1(\mathbb{Z}_l^d) \setminus U_{\varepsilon,l} \,\} < b. \tag{2.19}$$

Using (2.19), the following lemma is easy to prove:

Lemma 2.20. *For every $\varepsilon > 0$ there exists $l_0 \in \mathbb{N}$ such that*

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

Lemma 2.20 already seems to be very close to Proposition 2.18, except for one very annoying point which causes the main part of the work necessary to solve the tube problem. We already know that the elements in K_l are essentially concentrated on small sets, even uniformly for large l. Therefore, Lemma 2.20 says that L_T is essentially concentrated on the union of 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.

If we are in the one-dimensional case or if the potential V has all symmetries of the form

$$V(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_d) = V(x_1, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_d) \qquad (2.21)$$

with $i \in \{1, \ldots, d\}$ and $x = (x_1, \ldots, x_d) \in \mathbb{Z}^d$, then we can use a reflection argument, which we want to describe now. For every $\varepsilon > 0$ and sufficiently small $\delta > 0$ there exists $l_0 \in \mathbb{N}$ such that for all $l \geq l_0$ the following works: Suppose that $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$ is not close to an optimal measure, which means that $\mu \notin U_\varepsilon(K)$, but that its projection μ^l to the torus nearly looks like the projection of an optimal one, which means that $\mu^l \in U_{\delta,l}$. Then μ must have substantial mass in at least two l-translates of a small region, in which $\mathbb{Z}^d$-optimal measures have their mass. We can then find a coordinate direction $\kappa \in \{1, \ldots, d\}$ and an integer $i \in \mathbb{Z}$ such that the corresponding hyperplane $h_{i,\kappa} \equiv \{ (x_1, \ldots, x_d) \in \mathbb{Z}^d \colon x_\kappa = i \}$ does not intersect the l-translates of the small region but separates the mass of μ. If the hyperplane is properly chosen, then we can reflect one half of $\mathbb{Z}^d$ on this hyperplane into the other half of $\mathbb{Z}^d$ such that the small regions remain properly separated. This is one place where we use that R_V, given by (1.5), is finite. Let $\widehat{\varphi}_{i,\kappa}(\mu)$ denote the measure arising from μ via the reflection of the mass at the hyperplane $h_{i,\kappa}$. When we project $\widehat{\varphi}_{i,\kappa}(\mu)$ to the torus, the translates of the small region do not come one upon another, hence $(\widehat{\varphi}_{i,\kappa}(\mu))^l$ must be far away from the projections of the optimal measures in K, which means that $(\widehat{\varphi}_{i,\kappa}(\mu))^l \notin U_{c(d)\varepsilon,l}$, where $c(d)$ is a constant depending on the dimension. We need (2.21) to prove that $|H(\widehat{\varphi}_{i,\kappa}(\mu)) - H(\mu)|$ is small. Furthermore, we have to estimate the probabilistic "cost" of the constraint, that the random walk when visiting the hyperplane can only leave it to the side the walk came from.

If we do not have (2.21), then we have to find an appropriate slab $s_{i,\kappa,3w}$ of width $3w$, which we can fold up to a slab $s_{i,\kappa,w}$ of width w by using two of the above reflections. If μ is the above mentioned measure, then its mass outside the slab $s_{i,\kappa,3w}$ is at most shifted and the corresponding terms in (1.6) are not affected. We only have to estimate the terms $V(x-y)\mu(x)\mu(y)$ with $x \in s_{i,\kappa,3w}$ or $y \in s_{i,\kappa,3w}$.

Tightness. The key step for proving the tightness of $\{\widehat{\mathbb{P}}_T L_T^{-1}\}_{T \geq 1}$ is the next lemma which says the following. Given an optimal measure $\mu \in K$ with substantial mass at the origin, the $\widehat{\mathbb{P}}_T$-probability of the event, that the empirical measure L_T looks like a shifted copy $\theta_x(\mu)$ of μ and the process $\{X_s\}_{s \in [0,T]}$ is on a far-reaching excursion from the main bulk of $\theta_x(\mu)$ at time t, is negligible.

Lemma 2.22. *There exists $\varepsilon_0 > 0$ such that for every $\eta > 0$ and $\mu \in K(0)$ there exist $n \in \mathbb{N}$ and $T_0 > 0$ satisfying*

$$\sup_{T>T_0} \sup_{t\in[0,T]} \sum_{x\in\mathbb{Z}^d} \widehat{\mathbb{P}}_T(L_T \in U_{\varepsilon_0}(\theta_x(\mu)), \|X_t - x\|_\infty \geq n) \leq \eta,$$

where $K(0)$ is given by (2.12).

To prove tightness, it suffices to use this lemma for the case $t = 0$, but its full strength is needed for the proof of the convergence in (2.15), (2.16), and (2.17).

Let us explain the main line of the proof of Lemma 2.22. If L_T is in an ε_0-neighbourhood of $\theta_x(\mu)$, then 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, 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 want to discretize time for technical reasons in the proof, we need that the random walk rests for a short moment at the edge of the main bulk of $\theta_x(\mu)$ when it starts and ends the excursion. The first time after t, at which the process is at the edge of the main bulk of $\theta_x(\mu)$, is a stopping time, but the last one before t is not. This problem can be handled by reversing the time in the interval $[0, T]$. Similar as in (1.8) define

$$L_{u,v,T} = \frac{1}{u + T - v} \int_{[0,u)\cup[v,T)} \delta_{X_s} \,\mathrm{d}s.$$

Splitting L_T as

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

we can use (2.8) to decompose the Hamiltonian in the following way

$$\begin{aligned} TH(L_T) = \frac{(v-u)^2}{T} H(L_{u,v}) + T\Big(1 - \frac{v-u}{T}\Big)^2 H(L_{u,v,T}) \\ + (v-u)\Big(1 - \frac{v-u}{T}\Big)\langle DH(L_{u,v,T}), L_{u,v}\rangle. \end{aligned}$$

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 $\langle DH(L_{u,v,T}), L_{u,v}\rangle$ is small. If $v - u \leq t_0$ for an appropriate $t_0 > 0$ and if n is sufficiently large, then the process must have jumped very often during $[u, v]$; we can show that the probability of such a short but far-reaching excursion is negligible by using a large deviation argument concerning the sum of the holding times. If $v - u > t_0$, then we consider a partially exchanged path, which is identically to the original one during $[0, u] \cup [v, T]$ but hangs around the main bulk of $\theta_x(\mu)$ during $[u, v]$. For this modified path, $\langle DH(L_{u,v,T}), L_{u,v}\rangle$ is not small any more and $H(L_T)$ is considerably increased without paying too much "entropy". Therefore, the partially exchanged path has a substantially higher probability with respect to $\widehat{\mathbb{P}}_T$, i. e., the $\widehat{\mathbb{P}}_T$-probability of the original path was small.

Using Lemma 2.22 it is not difficult to prove

Proposition 2.23.

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

$$\limsup_{T\to\infty} \widehat{\mathbb{P}}_T(L_T \notin U_\varepsilon(K(n))) \le \eta,$$

where $K(n) \equiv \{\, \theta_x(\mu) \colon \mu \in K(0),\, x \in [-n,n]^d \cap \mathbb{Z}^d \,\}$.

(2) *The set $\{\widehat{\mathbb{P}}_T L_T^{-1}\}_{T\ge 1}$ is tight.*

Convergence for strong Dirac-type interactions. Let the potential V be given by (1.4) and let us show how to prove that for any coupling constant $\beta \ge 2d$ there is only one kind of optimal measures on $\mathbb{Z}^d$, as asserted by Theorem 2.10. The following algebraic lemma is easy to prove:

Lemma 2.24. *If $a \ge \frac{1}{2}$ and $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$ satisfy $\mu(x) \le a$ for all $x \in \mathbb{Z}^d$, then*

$$\sum_{x\in\mathbb{Z}^d} \mu^2(x) \le a^2 + (1-a)^2.$$

Define $K' \equiv \{\, \mu \in K \colon \mu(0) = \max_{x\in\mathbb{Z}^d} \mu(x) \,\}$. By comparing an optimal measure $\mu \in K'$ with the Dirac measure $\delta_0 \in \mathcal{M}_1(\mathbb{Z}^d)$, we can prove that a substantial amount (at least the half) of the mass of μ is concentrated at the origin.

Lemma 2.25. *If $\beta \ge 2d$ and $\mu \in K'$, then $\mu(0) \ge \frac{1}{2}\bigl(1 + \sqrt{1-(2d/\beta)^2}\bigr)$.*

Proof. Note that $H(\delta_0) = \beta$ and $J(\delta_0) = d$. Define

$$a_n = \frac{1}{2}\Bigl(1 + \sqrt{1-(2d/\beta)^{2(1-2^{-n})}}\Bigr)$$

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 2.24. Since $J(\mu) > 0$, it follows that $\Lambda(\mu) < \beta/2$. This contradicts $\Lambda(\mu) \ge \Lambda(\delta_0) = \beta - d \ge \beta/2$.

Assume that $\mu(0) \le a_{n+1}$. Then, by Lemma 2.24,

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

According to the induction hypotheses, $\mu(0) > a_n$, hence $\mu(y) < 1 - a_n$ for all $y \in \mathbb{Z}^d$ with $\|y\|_1 = 1$. Restricting the sum in (1.10) to all $\{0, y\} \subset \mathbb{Z}^d$ with $\|y\|_1 = 1$, it follows that

$$J(\mu) > d\bigl(\sqrt{a_n} - \sqrt{1-a_n}\bigr)^2 = d - 2d\sqrt{a_n(1-a_n)} = d - d\Bigl(\frac{2d}{\beta}\Bigr)^{1-2^{-n}},$$

hence $\Lambda(\mu) < \beta - d$. Again, this contradicts $\Lambda(\mu) \ge \Lambda(\delta_0) = \beta - d$. □

The result of Lemma 2.25 would be sufficient to prove Theorem 2.10 for every $\beta \ge 3.1766d$. To prove it for all $\beta \ge 2d$, we need a refinement of Lemma 2.25, which is obtained by comparison with the measure $\nu \in \mathcal{M}_1(\mathbb{Z}^d)$ given by

$$\nu(x) = \begin{cases} \frac{1}{2}\bigl(1 + \sqrt{1 - d/(2\beta^2)}\bigr), & \text{if } x = 0, \\ \frac{1}{2d}(1 - \nu(0)), & \text{if } \|x\|_1 = 1, \\ 0, & \text{if } \|x\|_1 > 1. \end{cases}$$

At least for us it was useful to have a small electronic device which did the computing in the course of the proof of the following refinement.

Lemma 2.26. *If $\beta \geq 2d$ and $\mu \in K'$, then $\mu(0) \geq \frac{1}{2}\big(1 + \sqrt{1-(1.19d/\beta)^2}\big)$.*

Proof of Theorem 2.10. Assume that there exist $\mu, \widetilde{\mu} \in K'$ with $\mu \neq \widetilde{\mu}$. Define $\varphi, \widetilde{\varphi} \in l_2(\mathbb{Z}^d)$ by $\varphi(x) = \sqrt{\mu(x)}$ and $\widetilde{\varphi}(x) = \sqrt{\widetilde{\mu}(x)}$ for all $x \in \mathbb{Z}^d$. Let

$$\chi \equiv \frac{\widetilde{\varphi} - \langle \widetilde{\varphi}, \varphi\rangle \varphi}{\|\widetilde{\varphi} - \langle \widetilde{\varphi}, \varphi\rangle \varphi\|_{l_2}} \quad \text{and} \quad \varepsilon_0 \equiv \arcsin \|\widetilde{\varphi} - \langle \widetilde{\varphi}, \varphi\rangle \varphi\|_{l_2}.$$

Then χ is a unit vector in $l_2(\mathbb{Z}^d)$ with $\chi \perp \varphi$. Define $\psi\colon [0,\varepsilon_0] \to l_2(\mathbb{Z}^d)$ by $\psi_\varepsilon \equiv \varphi \cos\varepsilon + \chi \sin\varepsilon$ and $\nu\colon [0,\varepsilon_0] \to \mathcal{M}_1(\mathbb{Z}^d)$ by $\nu_\varepsilon = \psi_\varepsilon^2$. Note that ψ turns the unit vector φ on the surface of the unit ball in $l_2(\mathbb{Z}^d)$ into the vector $\widetilde{\varphi} = \psi_{\varepsilon_0}$. Define $\lambda(\varepsilon) = \Lambda(\nu_\varepsilon)$ for all $\varepsilon \in [0,\varepsilon_0]$. If we can prove that $\lambda''(\varepsilon) < 0$ for all $\varepsilon \in [0,\varepsilon_0]$, then we have the contradiction we need.

Using $\psi'' = -\psi$, it follows that

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

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

$$|\psi'(x)\psi(x)| = \Big| \sum_{y\in\mathbb{Z}^d\setminus\{x\}} \psi'(y)\psi(y)\Big| \leq \sqrt{1-(\psi'(x))^2}\sqrt{1-\psi^2(x)}.$$

Squaring and solving for $(\psi'(x))^2$ yields $(\psi'(x))^2 \leq 1 - \psi^2(x)$. Since $J \leq d$ on $\mathcal{M}_1(\mathbb{Z}^d)$, it follows that $\lambda'' < 2d + 12\beta - 16H \circ \nu$. Since the estimate in Lemma 2.26 holds for $\mu(0)$ and $\widetilde{\mu}(0)$, it is also valid for $\nu_\varepsilon(0)$ with $\varepsilon \in [0,\varepsilon_0]$. Therefore,

$$\lambda'' < 2d + 12\beta - 16\beta\nu^2(0) \leq 2d + 4\frac{1.19^2 d^2}{\beta} - 4\beta\big(2\sqrt{1-(1.19d/\beta)^2} - 1\big).$$

For $\beta \geq 2.38d/\sqrt{3} \approx 1.374d$ this upper bound is obviously decreasing in β and it is negative for $\beta = 2d$. Hence $\lambda'' < 0$ for every $\beta \geq 2d$. □

3. Self-attraction caused by the number of visited points

Let Z_T be given by (1.2) with the Hamiltonian $H_T \equiv -N_T$, where $N_T \equiv |D_T|$ denotes the number of points visited by $\{X_t\}_{t\in[0,T]}$. A celebrated result of Donsker and Varadhan [11] states that

$$\lim_{T\to\infty} T^{-d/(d+2)} \log Z_T = \chi_d \equiv -\frac{d+2}{2}\left(\frac{2\lambda_d}{d}\right)^{d/(d+2)} \omega_d^{2/(d+2)}, \tag{3.1}$$

where $\omega_d = \pi^{d/2}\Gamma(d/2+1)$ is the volume of the unit ball and λ_d is the principal Dirichlet eigenvalue of $-\frac{1}{2}\Delta$ in the unit ball. For $x \in \mathbb{Z}^d$ and $r > 0$ let

$$B(x,r) \equiv \{\, y \in \mathbb{Z}^d\colon \|x-y\|_2 < r \,\}.$$

The main result of [2] is the following:

Theorem 3.2. *Let $\varepsilon > 0$ and $\varrho_d \equiv (2\lambda_d/d\omega_d)^{1/(d+2)}$. If $d = 2$, then*

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T\Big(\bigcup_{x\in B_T(0)} \{\, (1-\varepsilon)B_T(x) \subset D_T \subset (1+\varepsilon)B_T(x)\,\}\Big) = 1, \tag{3.3}$$

where $B_T(x) \equiv B(x, \varrho_d T^{d/(d+2)})$ for all $T > 0$ and $x \in \mathbb{Z}^d$.

Remark 3.4.

(1) A similar model with N_T replaced by the volume of the Wiener sausage has been discussed by Sznitman [18] for $d = 2$. The one-dimensional case was treated in [17].
(2) The paper [2] gives a discussion of the probabilistic aspects in arbitrary dimensions. We have no doubts that the result is true in any dimension. For a complete proof, there is however still lacking a delicate sharpening of the isoperimetric inequality. This will be discussed below.
(3) The ε in the statement of Theorem 3.2 can be replaced by $T^{-\delta}$ for some (small) δ. This follows by an inspection of the proof in [2]. It is however questionable if the method could be used to give a precise description of the boundary fluctuations.

We present here a sketch of the main arguments, ideas, and difficulties to prove the localization property formulated in Theorem 3.2. The starting point is the Donsker–Varadhan analysis of (3.1). The reader may find a detailed presentation of (3.1) for the Wiener sausage case in [7, Chapter 4.3]. Since the analysis is quite delicate, we give a short outline here. The main difficulties for proving (3.1) also show up in the proof of our Theorem 3.2 in aggravated form.

Sketch of the proof of Equation (3.1). The *lower bound* is quite easy. For any $r > 0$ we have

$$Z_T \geq \exp(-|B(0,r)|)\, \mathbb{P}(X_t \in B(0,r) \text{ for all } t \in [0,T]). \tag{3.5}$$

If we choose $r = \varrho T^{1/(d+2)}$ with $\varrho > 0$, then

$$|B(0,r)| = \varrho^d T^{d/(d+2)}\omega_d + O(T^{(d-1)/(d+2)}).$$

To get an approximation of the second factor on the right-hand side of (3.5), it is convenient to perform a Brownian rescaling. Let $\xi_t \equiv T^{-1/(d+2)} X_{tT^{2/(d+2)}}$ for all $t \leq \tau \equiv T^{d/(d+2)}$. Then $\{\xi_t\}_{t\in[0,\tau]}$ is a process on the rescaled lattice $T^{-1/(d+2)}\mathbb{Z}^d$, which more and more looks like a Brownian motion $\{\beta_t\}_{t\in[0,\tau]}$ on $\mathbb{R}^d$ as $T \to \infty$. It is not difficult to prove that

$$\begin{aligned}\mathbb{P}(X_t \in B(0, \varrho T^{1/(d+2)}) \text{ for all } t \in [0,T]) \;&=\; \mathbb{P}(\xi_t \in B(0,\varrho) \text{ for all } t \in [0,\tau]) \\ &\overset{\log}{\sim} \mathbb{P}(\|\beta_t\|_2 < \varrho \text{ for all } t \in [0,\tau]),\end{aligned}$$

where the relation $\overset{\log}{\sim}$ means that the quotient of the logarithms is converging to 1 as $\tau \to \infty$. It is well known that

$$\mathbb{P}(\|\beta_t\|_2 < \varrho \text{ for all } t \in [0,\tau]) \overset{\log}{\sim} \exp(-\tau\varrho^{-2}\lambda_d). \tag{3.6}$$

Substituting all this into (3.5) and choosing the optimal ϱ, namely $\varrho = \varrho_d$, gives

$$\log Z_T \geq \chi_d \tau + o(\tau) = \chi_d T^{d/(d+2)} + o(T^{d/(d+2)}),$$

where $o(T^{d/(d+2)})$ comes from the relation $\overset{\log}{\sim}$ in (3.6). A closer inspection reveals that one can replace $o(T^{d/(d+2)})$ by $O(T^{(d-1)/(d+2)})$, see [2, Section 2].

Much more delicate is the *upper bound,* which is obtained by a sophisticated counting argument. Donsker and Varadhan had the idea to use large deviation techniques for the empirical distribution process $\{L_T\}_{T>0}$ given by (1.8). The (weak) large deviation principle given by (1.11) and (1.12) is, however, of no direct use, because there is no T in front of N_T in the definition $Z_T = \mathbb{E}[\exp(-N_T)]$. (To discuss the path measures with the Hamiltonian $H_T \equiv -TN_T$ would in fact be very simple.) To cope with this problem, one uses the same Brownian rescaling as for the lower bound. To fix notations, we define the local times by

$$l_\tau(x) \equiv \int_0^\tau 1_{\{x\}}(\xi_t)\,\mathrm{d}t, \quad x \in T^{-1/(d+2)}\mathbb{Z}^d.$$

Remember that $\tau \equiv T^{d/(d+2)}$. We don't put the factor $1/\tau$ in front, because we want to view l_τ also as an empirical density: In fact, if we simply extend l_τ to $\mathbb{R}^d$ by defining $l_\tau(y) = l_\tau(x)$ for every y in the cube $\prod_{i=1}^d [x_i, x_i + T^{-1/(d+2)})$, then

$$\int_{\mathbb{R}^d} l_\tau(y)\,\mathrm{d}y = 1.$$

Of course, one can also look at the empirical measure, say $\widetilde{L}_\tau$, defined by this empirical density, and it is not difficult to prove that $\{\mathbb{P}\widetilde{L}_\tau^{-1}\}_{\tau>0}$ satisfies a (weak) large deviation principle in the weak topology of $\mathcal{M}_1(\mathbb{R}^d)$. Note that

$$N_T = \tau |\operatorname{supp}(\widetilde{L}_\tau)|, \tag{3.7}$$

where $|\cdot|$ here denotes the Lebesgue measure on $\mathbb{R}^d$. Therefore, one might hope that the (weak) large deviation principle can be applied, but there is one crucial difficulty, which becomes even more prominent when discussing the path measures $\{\widehat{\mathbb{P}}_T\}_{T>0}$. The map $\mu \mapsto |\operatorname{supp}(\mu)|$ has no continuity property at all in the weak topology of $\mathcal{M}_1(\mathbb{R}^d)$. An important observation of Donsker and Varadhan was that $f \mapsto |\{\, x \in \mathbb{R}^d : f(x) > 0 \,\}|$ is lower semicontinuous in the L_1-topology on the set $\mathcal{D}$ of all probability densities with respect to the Lebesgue measure on $\mathbb{R}^d$. Therefore, one may ask whether $\{\mathbb{P}l_\tau^{-1}\}_{\tau>0}$ satisfies a large deviation principle in the L_1-topology.

Before proceeding further, one has to cope with a technical nuisance. There is no full large deviation principle in any topology for our random walk, which means that, in general, the large deviation upper bound in (1.11) does not hold for all closed sets, but only for compact ones. This difficulty is remedied by a suitable compactification. In our case, we replace the random walk $\{\xi_t\}_{t\in[0,\tau]}$ on the rescaled lattice $T^{-1/(d+2)}\mathbb{Z}^d$ by a periodized random walk on a discrete torus

$$\mathcal{T}_T^R \equiv T^{-1/(d+2)}\mathbb{Z}^d \cap [-R, R)^d,$$

adjusted, if necessary, at the seams. We then have our random walk $\{\xi_t^R\}_{t\in[0,\tau]}$ and the empirical densities $\{l_\tau^R\}_{\tau>0}$ on this torus. We will switch freely between the lattice torus $\mathcal{T}_T^R$ and the continuous torus $[-R,R)^d$ without much comments, hoping that this will not confuse the reader. Of course, finally one must get rid of the torus and be able to compare the unrestricted situation with the one on the torus. This is fairly easy for the discussion of Z_T, but has some delicacies when looking at $\widehat{\mathbb{P}}_T$. Now $\{\mathbb{P}(\widetilde{L}_\tau^R)^{-1}\}_{\tau>0}$ is easily seen to satisfy a good large deviation principle in $\mathcal{M}_1([-R,R)^d)$, equipped with the weak topology, but the important result is the following one:

Theorem 3.8 (Donsker and Varadhan)**.** *If $R>0$, then the set $\{\mathbb{P}(l_\tau^R)^{-1}\}_{\tau>0}$ of measures satisfies a full large deviation principle in the L_1-topology on the space $\mathcal{D}_R$ of probability densities on the continuous torus $[-R,R)^d$ with the rate function*

$$I_R(f) \equiv \begin{cases} \frac{1}{2}\int_{[-R,R)^d} |\nabla\sqrt{f}|^2 \,\mathrm{d}x, & \text{if } \sqrt{f}\in H^1([-R,R)^d), \\ \infty, & \text{otherwise.} \end{cases} \tag{3.9}$$

Remark 3.10.

(1) The result does not appear in the above form in the papers by Donsker and Varadhan. A version of it is proved for smoothed empirical densities of Brownian motion in [9]. They also discuss random walks in discrete time (which are more difficult than in continuous time) in [11] by suitable approximations using local central limit theorems. For continuous-time random walks, the result follows by an easy adaptation of their techniques.

(2) The exponent of the rescaling is of crucial importance for the validity of Theorem 3.8. Instead of taking $\xi_t \equiv T^{-1/(d+2)} X_{tT^{2/(d+2)}}$ for $t\le\tau$, we could also, more generally, consider $\xi_{a,t} \equiv T^{-a/2} X_{tT^a}$ for $t\le\tau_a\equiv T^{1-a}$, where a is in $(0,1)$. Theorem 3.8 remains valid for $0<a<2/d$, but is false for $a>2/d$ in three and more dimensions. This has been proved in [1] for the Wiener sausage. The random walk case has not been worked out in details, but should follow by an adaptation of the method. To relish the delicacy of Theorem 3.8, the reader should convince himself that the corresponding large deviation result for $\{\mathbb{P}(\widetilde{L}_{\tau_a}^R)^{-1}\}_{\tau_a>0}$ in the weak topology is valid for any $a\in(0,1)$.

With Theorem 3.8, a proof of the upper bound in (3.1) is now easy: Since the map $\mathcal{D}_R \ni f \mapsto |\{f>0\}|$ is lower semicontinuous in the L_1-topology, it follows with (3.7) and Varadhan's theorem that, for every $R>0$,

$$\begin{aligned} \limsup_{T\to\infty} T^{-d/(d+2)} \log Z_T &= \limsup_{\tau\to\infty} \tau^{-1} \log \mathbb{E}[\exp(-\tau|\{l_\tau>0\}|)] \\ &\le \limsup_{\tau\to\infty} \tau^{-1} \log \mathbb{E}[\exp(-\tau|\{l_\tau^R>0\}|)] \\ &\le -\inf_{f\in\mathcal{D}_R} \{|\{f>0\}| + I_R(f)\}. \end{aligned} \tag{3.11}$$

There is now a slightly nasty technical point. One has to replace on the right-hand side R by infinity. In fact, it is true, that

$$\lim_{R\to\infty} \inf_{f\in\mathcal{D}_R} \{|\{f>0\}| + I_R(f)\} = \inf_{f\in\mathcal{D}} \{|\{f>0\}| + I(f)\}, \tag{3.12}$$

where $\mathcal{D}$ denotes the set of probability densities on $\mathbb{R}^d$, and $I(f)$ is defined by (3.9) but with $[-R,R)^d$ replaced by $\mathbb{R}^d$. For a detailed proof of this, see [7, Chapter 4.3]. There can be no doubt that

$$\inf_{f\in\mathcal{D}_R}\{|\{f>0\}|+I_R(f)\}=\inf_{f\in\mathcal{D}}\{|\{f>0\}|+I(f)\} \tag{3.13}$$

for all sufficiently large R. This is easy for $d=2$, but not yet proved for $d\geq 3$. The delicacy to prove (3.13) for $d\geq 3$ is in fact connected with the difficulties to prove Theorem 3.2 in higher dimensions. We will come to this a bit later on. Anyway, (3.12) is good enough to prove (3.1). Since (3.11) holds for all R, we get

$$\begin{aligned}\limsup_{T\to\infty}T^{-d/(d+2)}\log Z_T &\leq -\inf_{f\in\mathcal{D}}\{|\{f>0\}|+I(f)\}\\ &=-\inf_{v>0}\inf_{\substack{G\subset\mathbb{R}^d\text{ open}\\|G|=v}}\Big\{v+\inf_{\substack{f\in\mathcal{D}\\ \mathrm{supp}(f)\subset G}}I(f)\Big\}\\ &=-\inf_{v>0}\Big\{v+\inf_{\substack{G\subset\mathbb{R}^d\text{ open}\\|G|=v}}\lambda(G)\Big\},\end{aligned} \tag{3.14}$$

where $\lambda(G)$ is the principle Dirichlet eigenvalue of $-\frac{1}{2}\Delta$ in G. If G is open and connected and if g_G is the L_2-normalized principal Dirichlet eigenfunction of $-\frac{1}{2}\Delta$ in G, then $f_G\equiv g_G^2$ is the unique probability density satisfying

$$I(f_G)=\inf\{\,I(f)\colon f\in\mathcal{D}\text{ with }\{f>0\}\subset G\,\}=\lambda(G).$$

The celebrated Faber-Krahn theorem states that $\inf\{\,\lambda(G)\colon G\text{ open},\,|G|=v\,\}$ is attained exactly when G is a ball of volume v. Using $\lambda(G)=(\omega_d/v)^{2/d}\lambda_d$ and taking the minimum over v on the right-hand side of (3.14) then gives the upper bound in (3.1). The minimizing balls have radius ϱ_d.

For use in the next subsection, we let $\mathcal{B}$ denote the set of all balls of radius ϱ_d in $\mathbb{R}^d$ and

$$\mathcal{F}\equiv\{\,f_B\colon B\in\mathcal{B}\,\} \tag{3.15}$$

the set of all squares of the corresponding L_2-normalized principal Dirichlet eigenfunctions of $-\frac{1}{2}\Delta$.

Outline of the proof of Theorem 3.2. From the derivation of (3.1) in the previous subsection, one sees what is behind Theorem 3.2. The measure $\widehat{\mathbb{P}}_T$ is, in a sense, the law of the typical paths contributing to Z_T. The large deviation theory suggests that the main contribution is coming from those paths whose empirical density is close to the square of the principal Dirichlet eigenfunction in a ball of radius ϱ_d. Therefore, the support of the empirical density should be close to one of these balls. However, the essential difficulty can be expressed by the fact that for a closed subset A of $\mathbb{R}^d$ the Hausdorff distance function

$$\mathcal{D}\ni f\mapsto \mathrm{dist}_{\mathrm{H}}(A,\mathrm{supp}(\mu_f))$$

where $\mu_f\in\mathcal{M}_1(\mathbb{R}^d)$ denotes the measure with density f, has no continuity property in the L_1-topology. (The L_∞-topology would have some advantages, but it is a

simple fact that Theorem 3.8 is false in the L_∞-topology, except for $d = 1$.) There is also another problem which is somewhat connected with that, namely to make the link between the "compactified" situation on the torus, and the original one on $\mathbb{R}^d$. To solve this problem, we need that (3.13) is true for sufficiently large R. As remarked above, this is easy for $d = 2$ and is probably true in any dimension. Let us assume (3.13) and see what we can obtain from this and the discussion of the partition function Z_T. If R is large enough, in particular larger than ϱ_d, then we can consider the functions in (3.15) as being functions defined on the torus $[-R, R)^d$. For clarity, we then write $\mathcal{F}^R$ for this class of functions, and $\mathcal{B}^R$ for the set of balls of radius ϱ_d on this torus. As a consequence of (3.1), (3.13), and Theorem 3.8, we easily get, for sufficiently large R and all $\varepsilon > 0$,

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T\Big(\inf_{f\in\mathcal{F}^R} \|l_\tau^R - f\|_1 \geq \varepsilon\Big) = 0,$$

and we can conclude that

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T\Big(\inf_{B\in\mathcal{B}^R} |\{l_\tau^R > 0\} \triangle B| \geq \varepsilon\Big) = 0. \tag{3.16}$$

The reader may wonder if one can conclude from (3.16) the nontorus version

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T\Big(\inf_{B\in\mathcal{B}} |\{l_\tau > 0\} \triangle B| \geq \varepsilon\Big) = 0, \tag{3.17}$$

which, although not implying Theorem 3.2, would be some step in that direction and at least a partial result. However, here the essential difficulties start. Although (3.3) is clearly a consequence of a torus version of (3.3), (3.17) is not a consequence of (3.16) except for $d = 1$. The problem is that to be able to forget about the torus, one must know that there are no windings around it, which is implied by the torus version of (3.3) (if R is large enough) but not by (3.16). Therefore, one would have to prove a torus version of (3.3) even if one would modestly be satisfied with (3.17).

Let us look at the kind of sharpenings of (3.16) needed to get the result. We should remember that our lattice is becoming finer and finer as T grows. This means that even if

$$|\{l_\tau^R > 0\} \triangle B|$$

is very small for some $B \in \mathcal{B}^R$, this does not imply that there are no long excursions from the ball. For $d \geq 2$, an excursion of length of order 1 (on the rescaled grid) gives a minimal contribution of order $T^{-(d-1)/(d+2)}$, which is not felt by the analysis done so far. The aim is therefore to sharpen (3.16) by replacing ε by something which is going to zero. For our aim to "detorize" the situation, we can as well blow up the balls in $\mathcal{B}^R$ a little bit. If $B = B(x, \varrho_d) \in \mathcal{B}^R$, then we define $B^a = B(x, a\varrho_d)$ for $a > 0$. For $\gamma, \delta > 0$ let $A^R(\gamma, \delta)$ be the statement

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T(|\{l_\tau^R > 0\} \smallsetminus B^{(1+\delta)}| \leq T^{-\gamma} \text{ for some } B \in \mathcal{B}^R) = 1.$$

If $A^R(\gamma, \delta)$ is true for some $\gamma > (d-1)/(d+2)$ and all $\delta > 0$, then we have proved that with $\widehat{\mathbb{P}}_T$-probability converging to 1, the whole path is staying inside a chart of

the torus, and from this one easily gets the corresponding nontorus statement. This is one half of our Theorem 3.2. The other half is to prove that, for all $\delta \in (0,1)$,

$$\lim_{T\to\infty} \widehat{\mathbb{P}}_T(B^{(1-\delta)} \subset \{l_\tau > 0\} \text{ for some } B \in \mathcal{B}) = 1. \tag{3.18}$$

The derivation of the torus version of (3.18) is quite easy for $d = 2$, and the main task is to prove the $A^R(\gamma,\delta)$-statements. Equation (3.18) is actually used in the proof of these as a technical means.

The discussion given here is actually oversimplified. The events which have to be discussed in [2] are more involved. We only want to convey here the main line of the argument.

The statement $A^R(\gamma,\delta)$ is proved by increasing γ inductively, but the foundation has to be a proof of $A^R(\gamma,\delta)$ for some $\gamma > 0$. This turned out to be very involved. Essentially, this can be proved by a sharpening of the upper bound for Z_T and similar but more complicated expectations. The problem to get this is stemming from the fact that despite of the lower semicontinuity of the map $f \mapsto |\{f > 0\}|$, it is still a very badly behaved functional. The way, chosen in [2], is to smooth it somewhat and replace N_T by

$$\sum_{x\in\mathbb{Z}^d} \varphi_T(L_T(x)),$$

where $\varphi_T : [0,\infty) \to [0,1]$ is continuous, but approaches $1_{(0,\infty)}$ as $T \to \infty$. With the help of some other techniques, we are then able to prove a slight improvement for the upper bound of Z_T: There exists some $\varepsilon > 0$ such that for all sufficiently large T

$$\log Z_T \le \chi_d T^{d/(d+2)} + T^{d/(d+2)-\varepsilon}. \tag{3.19}$$

A somewhat weaker upper bound in the two-dimensional Wiener sausage case has also been obtained by Sznitman [18, (7)]. A proof of (3.19) is based on the following analytical property. For sufficiently large R and some $\delta > 0$

$$\inf\{\, |\{f > 0\}| + I_R(f) \colon f \in \mathcal{D}_R,\, \mathrm{dist}_{L_1}(f, \mathcal{F}^R) \ge t \,\} \ge -\chi_d + t^\delta, \tag{3.20}$$

for $t > 0$, where $\mathrm{dist}_{L_1}(f,\mathcal{F}^R) \equiv \inf\{\, \|f - g\|_{L_1} : g \in \mathcal{F}^R \,\}$.

Estimate (3.20) is proved in [2] for $d = 2$, but a proof is lacking in higher dimensions. There is no reasonable doubt that (3.20) is always true. It actually would follow from a sharpening of the isoperimetric inequality. We would need a result stating that if a (nice) set of volume ω_d has a surface which is in size close to the one of the unit ball, then there is a unit ball whose symmetric difference with the set has small Lebesgue measure. This is easy for $d = 2$, but not known for $d \ge 3$.

The paper [2] contains a complete proof that (3.20) implies (3.19) in any dimension. Actually, some sharpenings of the argument give the following result:

Proposition 3.21. *If* (3.20) *is true, then there is a* $\gamma > 0$ *such that for all* $\delta > 0$ *(and large* R*) the statement* $A^R(\gamma,\delta)$ *is also true.*

The second main probabilistic argument developed in [2] is an induction scheme to increase γ. The following statement is not exactly what has been proved, but it catches the spirit:

"Proposition 3.22". *For all $\gamma_0 > 0$, there exists $\kappa(\gamma_0) > 0$ such that for all $\delta > 0$ and $\gamma \geq \gamma_0$, there exists $\delta' > 0$ such that $A^R(\gamma, \delta')$ implies $A^R(\gamma + \kappa, \delta)$.*

From Propositions 3.21 and 3.22 we clearly get $A^R(\gamma, \delta)$ for $\gamma > (d-1)/(d+2)$ and all $\delta > 0$. Together with (3.18), this proves the Theorem 3.2.

The main idea to prove Proposition 3.22 is a conditioning argument. If $A^R(\gamma, \delta')$ is true, then we already know that there are only relatively few excursions from the "good" ball. We then take the conditional law of these excursions given the situation inside the good ball. The law of these excursions is quite messy, but assume for the moment that they would behave as if they form an ordinary random walk patched together of these excursion pieces. The total length of this "excursion walk" would be substantially smaller than τ (everything always in the rescaled situation). Strictly speaking, this does not quite follow from $A^R(\gamma, \delta')$, because the latter only restricts the number of the visited points and not the length of the excursions, but we want to neglect such accessory things here. Anyway, we could apply a rescaled $A^R(\gamma, \delta')$-statement to the "excursion walk" proving that under $\widehat{\mathbb{P}}_T$ (strictly speaking under the conditional law given what happens inside the good ball) the path would clump to a considerable extent having only a substantially shorter "really bad" excursion. This gives the increase of γ. (The argument really needs that one starts with a positive γ_0. The $\kappa(\gamma_0)$ given in [2] goes to zero as $\gamma_0 \downarrow 0$.) In [2] we give a complete proof of this bootstrapping procedure in arbitrary dimensions.

There are in fact two gaps for proving Theorem 3.2 in arbitrary dimensions. One is the above mentioned isoperimetric problem. The other one is a proof of (3.18), which is easy for $d = 2$. The latter can probably be proved by a variant of the method of excluding the "outer" excursions, which has just been sketched.

References

1. Bolthausen, E., *On the volume of the Wiener sausage*, Ann. Probab. **18** (1990), 1576–1582.
2. ———, *Localization of a two-dimensional random walk with an attractive path interaction*, Ann. Probab. (1994).
3. ———, *Stochastic processes with long range interactions of the paths*, Doeblin and Modern Probability: Proceedings of the Doeblin Conference (H. Cohn, ed.), Contemporary Mathematics, 149, Amer. Math. Soc., Providence, Rhode Island, 1993, pp. 297–319.
4. Bolthausen, E., J.-D. Deuschel, and U. Schmock, *Convergence of path measures arising from a mean field or polaron type interaction*, Probab. Theory Relat. Fields **95** (1993), 283–310.
5. Bolthausen, E., and U. Schmock, *On self-attracting d-dimensional random walks*, preprint (1994).
6. Brydges, D., and G. Slade, *The diffusive phase of a model of self-interacting walks*, preprint (1994).
7. Deuschel, J.-D., and D. W. Stroock, *Large Deviations*, Academic Press, San Diego, 1989.
8. Donsker, M., and S. R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time, I*, Comm. Pure Appl. Math. **28** (1975), 1–47.
9. ———, *Asymptotic evaluation of certain Markov process expectations for large time, II*, Comm. Pure Appl. Math. **28** (1975), 279–301.
10. ———, *Asymptotic evaluation of certain Markov process expectations for large time, III*, Comm. Pure Appl. Math. **29** (1976), 389–461.
11. ———, *On the number of distinct sites visited by a random walk*, Comm. Pure Appl. Math. **32** (1979), 721–747.
12. Ethier, S. N., and T. G. Kurtz, *Markov Processes, Characterization and Convergence*, John Wiley & Sons, New York, 1986.

13. König, W., *The drift of a one-dimensional self-avoiding random walk*, Probab. Theory Relat. Fields (1993), 521–543.
14. ———, *The drift of one-dimensional self-repellent random walk with bounded increments*, Probab. Theory Relat. Fields (1994).
15. Madras, N., and G. Slade, *The Self-Avoiding Walk*, Birkhäuser, Boston, 1993.
16. Mansmann, U., *The free energy of the Dirac polaron, an explicit solution*, Stochastics **34** (1991), 93–125.
17. Schmock, U., *Convergence of one-dimensional Wiener sausage path measures to a mixture of Brownian taboo processes*, Stochastics **29** (1989), 203–220.
18. Sznitman, A.-S., *On the confinement property of two-dimensional Brownian motion among Poissonian obstacles*, Comm. Pure Appl. Math. **44** (1991), 1137–1170.

Institut für Angewandte Mathematik, Universität Zürich, Winterthurer Str. 190, CH-8057 Zürich, Switzerland

E-mail addresses: eb@amath.unizh.ch (Erwin Bolthausen)
schmock@amath.unizh.ch (Uwe Schmock)