Weierstraß-Institut

für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Self-intersection local times of random walks: Exponential moments in subcritical dimension

Mathias Becker[1], Wolfgang König[2]

submitted: August 24, 2010

[1] Weierstraß–Institut
für Angewandte Analysis und Stochastik
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: mathias.becker@wias-berlin.de

[2] Weierstraß–Institut
für Angewandte Analysis und Stochastik
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: wolfgang.koenig@wias-berlin.de
und
Technische Universität Berlin
Institut für Mathematik
Str. des 17. Juni 136
10623 Berlin
Germany
E-Mail: koenig@math.tu-berlin.de

No. 1532
Berlin 2010

2010 *Mathematics Subject Classification.* 60K37, 60F10, 60J55.

Key words and phrases. Self-intersection local time, upper tail, Donsker-Varadhan large deviations, variational formula, Gagliardo-Nirenberg inequality.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract: Fix $p > 1$, not necessarily integer, with $p(d-2) < d$. We study the p-fold self-intersection local time of a simple random walk on the lattice $\mathbb{Z}^d$ up to time t. This is the p-norm of the vector of the walker's local times, ℓ_t. We derive precise logarithmic asymptotics of the expectation of $\exp\{\theta_t \|\ell_t\|_p\}$ for scales $\theta_t > 0$ that are bounded from above, possibly tending to zero. The speed is identified in terms of mixed powers of t and θ_t, and the precise rate is characterized in terms of a variational formula, which is in close connection to the *Gagliardo-Nirenberg inequality*. As a corollary, we obtain a large-deviation principle for $\|\ell_t\|_p/(tr_t)$ for deviation functions r_t satisfying $tr_t \gg \mathbb{E}[\|\ell_t\|_p]$.

Informally, it turns out that the random walk homogeneously squeezes in a t-dependent box with diameter of order $\ll t^{1/d}$ to produce the required amount of self-intersections. Our main tool is an upper bound for the joint density of the local times of the walk.

1. Introduction

In this paper, we give precise logarithmic asymptotics for the exponential moments of self-intersection local times of random walks on $\mathbb{Z}^d$ on various scales. This topic has been studied a lot in the last decade, since it is a natural question, and a rich phenemonology of critical behaviours of the random walk arises, depending on the dimension, the intersection parameter, the scale, and the type of the random process. Furthermore, the question is technically very difficult to handle, due to bad continuity and boundedness properties of the self-intersection local time. A couple of different techniques for studying self-intersections have been introduced yet, wich turned out to be more or less fruitful in various situations.

In this paper, we introduce a recently developed method to the study of self-intersections, which enables us to derive limits in terms of an explicit variational formula describing the asymptotics; this formula explains the optimal behaviour of the random walk to produce many self-intersections. We are working in sub-critical dimensions, where this behaviour consists of a homogeneous squeezing of the path over the whole time interval in a box of a certain time-dependent diameter.

Our method is strongly influenced by the celebrated Donsker-Varadhan large-deviations theory. The main obstacle that has to be overcome to make these ideas work is the lack of continuity, and this is serious. To overcome this, we use an explicit upper bound for the joint density of the walker's local times, which has been derived recently by Brydges, van der Hofstad and König [BHK07]. The main task left after applying this bound is to identify the scaling limit of the arising formula, and this is the main novelty of the present paper.

1.1 Self-intersection local time

Let $(S_t)_{t\in[0,\infty)}$ be a simple random walk in $\mathbb{Z}^d$ started from the origin. We denote by $\mathbb{P}$ the underlying probability measure and by $\mathbb{E}$ the corresponding expectation. The main object of this paper is the *self-intersection local time* of the random walk. In order to introduce this object, we need the local times of the random walk at time $t > 0$,

$$\ell_t(z) = \int_0^t \mathbb{1}_{\{S_t = z\}} \,\mathrm{d}t, \qquad \text{for } z \in \mathbb{Z}^d. \tag{1.1}$$

Fix $p \in (1,\infty)$ and consider the p-norm of the local times:

$$\|\ell_t\|_p = \Big(\sum_{z\in\mathbb{Z}^d} \ell_t(z)^p \Big)^{1/p}, \qquad \text{for } t > 0. \tag{1.2}$$

If p is an integer, then, clearly

$$\|\ell_t\|_p^p = \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_p \, \mathbb{1}_{\{S_{t_1} = \cdots = S_{t_p}\}} \tag{1.3}$$

is equal to the p-fold self-intersection local time of the walk, i.e., the amount of time tuples that it spends in p-fold self-intersection sites. For $p = 2$, this is usually called the *self-intersection local time*. For $p = 1$, $\|\ell_t\|_p^p$ is just the number t, and for $p = 0$, it is equal to $\#\{S_r \colon r \in [0,t]\}$, the *range* of the walk. It is certainly also of interest to study $\|\ell_t\|_p^p$ for non-integer values of $p > 1$, see for example [HKM06], where this received technical importance.

The typical behaviour of $\|\ell_t\|_p^p$ for continuous-time random walks cannot be found in the literature, to the best of our knowledge, but we have no doubt that it is, up to the value of the prefactor, equal to the behaviour of the self-intersection local time, $\|\ell_n\|_p^p$, of a centred random walk in discrete time. This has been identified as

$$\mathbb{E}[\|\ell_n\|_p^p] \sim C a_{d,p}(n), \qquad \text{where} \qquad a_{d,p}(n) = \begin{cases} n^{(p+1)/2} & \text{if } d = 1, \\ n(\log n)^{p-1} & \text{if } d = 2, \\ n & \text{if } d \geq 3, \end{cases} \tag{1.4}$$

where

$$C = \begin{cases} \frac{\Gamma(p+1)}{(2\pi\sqrt{\det \Sigma})^{p-1}} & \text{if } d = 2, \\ \gamma^2 \sum_{j=1}^{\infty} j^p (1-\gamma)^{1-j} & \text{if } d \geq 3, \end{cases} \tag{1.5}$$

where $\gamma = \mathbb{P}(S_n \neq 0 \text{ for any } n \in \mathbb{N})$ denotes the escape probability and Σ the covariance matrix of the random walk. See [Ce07] for $d = 2$ and [BK09] for $d \geq 3$, but we could find no reference for $d = 1$.

1.2 Main results

In this paper, we study the behaviour of the random walk when the walker produces extremely many self-intersections. We restrict to the *subcritical* dimensions, where $d(p-1) < 2p$. Before we formulate our results, let us informally describe the optimal behaviour to produce many self-intersections in these dimensions. It is a homogeneous self-squeezing strategy: the walker does not leave a box with radius on a particular scale $\alpha_t \ll \sqrt{t}$ (we write $b_t \ll c_t$ if $\lim_{t\to\infty} c_t/b_t = \infty$), and the sizes of all the local times are on the same scale t/α_t^d within this box. Furthermore, their rescaled shape approximates a certain deterministic profile, which is given in terms of a characteristic variational formula.

In our result, we do not prove this path picture, but we derive precise logarithmic asymptotics, as $t \to \infty$, for the exponential moments of $\theta_t \|\ell_t\|_p$ for various choices of weight functions $\theta_t \in (0,\infty)$ that are bounded from above. As a direct consequence, this leads to asymptotics of the probability of the event $\{\|\ell_t\|_p > t r_t\}$ for various choices of scale functions $r_t \in [0,1]$ satisfying $a_{d,p}(t) \ll t r_t$.

In order to formulate the result, we have to introduce more notation. By $L^q = L^q(\mathbb{R}^d)$ we denote the usual Lebesgue space, which is equipped with the norm $\|\cdot\|_q$ if $q \geq 1$. The space $H^1 = H^1(\mathbb{R}^d)$ is the usual Sobolev space. By $\mathcal{M}_1(X)$ we denote the set of probability measures on a metric space X equipped with the Borel sigma field.

Now we formulate our main result.

Theorem 1.1 (Exponential Moments)**.** *Assume that $p > 1$ and $d(p-1) < 2p$.*

(i) *For any $\theta > 0$,*

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\Big(\mathrm{e}^{\theta \|\ell_t\|_p}\Big) = \rho_{d,p}^{(\mathrm{d})}(\theta), \tag{1.6}$$

where

$$\rho_{d,p}^{(\mathrm{d})}(\theta) = \sup\Big\{\theta\|\mu\|_p - J(\mu) \colon \mu \in \mathcal{M}_1(\mathbb{Z}^d)\Big\} \in (0,\theta], \tag{1.7}$$

and $J(\mu) = \frac{1}{2}\sum_{x\sim y}\big(\sqrt{\mu(x)} - \sqrt{\mu(y)}\big)^2$ denotes the Donsker-Varadhan rate functional.

(ii) *Define* $\lambda = \frac{2p+d-dp}{2p} \in (0,1)$ *and let* $(\theta_t)_{t>0}$ *be a function in* $(0,\infty)$ *such that*

$$\left(\frac{\log t}{t}\right)^{\frac{2\lambda}{d+2}} \ll \theta_t \ll 1. \tag{1.8}$$

Furthermore, put

$$\rho^{(\mathrm{c})}_{p,d}(\theta) = \sup\Big\{\theta\|g^2\|_p - \frac{1}{2}\|\nabla g\|_2^2 \colon g \in H^1, \|g\|_2 = 1\Big\}, \qquad \theta > 0. \tag{1.9}$$

Then $\rho^{(\mathrm{c})}_{p,d}(\theta) \in (0,\infty)$, *and*

(a)

$$\frac{1}{t}\log \mathbb{E}\Big(\mathrm{e}^{\theta_t\|\ell_t\|_p}\Big) \geq \theta_t^{1/\lambda}\big(\rho^{(\mathrm{c})}_{d,p}(1) + o(1)\big), \qquad t \to \infty. \tag{1.10}$$

(b) *If, additionally to* $d(p-1) < 2p$, *the stronger assumption* $d(p-1) < 2$ *is fulfilled, then*

$$\frac{1}{t}\log \mathbb{E}\left(e^{\theta_t\|\ell_t\|_p}\right) \leq \theta_t^{1/\lambda}\big(\rho^{(\mathrm{c})}_{d,p}(1) + o(1)\big), \qquad t \to \infty. \tag{1.11}$$

A heuristic derivation of Theorem 1.1 is given in Section 1.5. The proof is given in Section 2. Some comments on the related literature are given in Section 1.4. We proceed with a couple of remarks.

1.3 Remarks

Remark 1.2 (Connection between $\rho^{(\mathrm{c})}_{d,p}$ and $\rho^{(\mathrm{d})}_{d,p}(\theta)$)**.** Note that the Donsker-Varadhan functional is equal to the walk's Dirichlet form at $\sqrt{\mu}$, i.e.,

$$J(\mu) = \frac{1}{2}\|\nabla\sqrt{\mu}\|_2^2, \qquad \mu \in \mathcal{M}_1(\mathbb{Z}^d),$$

where ∇ here denotes the discrete gradient. Hence, we see that $\rho^{(\mathrm{c})}_{p,d}(\theta)$ is the continuous version of $\rho^{(\mathrm{d})}_{d,p}(\theta)$. An important step in our proof of Theorem 1.1(ii) is to show that the continuous version of this formula describes the small-θ asymptotics of the discrete one, i.e.,

$$\rho^{(\mathrm{d})}_{d,p}(\theta) \sim \theta^{1/\lambda}\rho^{(\mathrm{c})}_{d,p}(1), \qquad \theta \downarrow 0. \tag{1.12}$$

(Actually, we only prove a version of this statement on large boxes, see Lemma 2.1.) In the light of this, we can heuristically explain the transition between the two cases in Theorem 1.1. Indeed, if we use (i) for θ replaced by $\theta_t \to 0$, then we formally obtain

$$\frac{1}{t}\log \mathbb{E}\left(\mathrm{e}^{\theta_t\|\ell_t\|_p}\right) \sim \rho^{(\mathrm{d})}_{p,d}(\theta_t) \sim \theta_t^{1/\lambda}\rho^{(\mathrm{c})}_{d,p}(1). \tag{1.13}$$

Hence, (1.12) shows that Theorem 1.1(i) and (ii) are consistent. ◇

Remark 1.3 (On the constant $\rho^{(\mathrm{c})}_{d,p}(\theta)$)**.** We will show in the following that

$$\rho^{(\mathrm{c})}_{p,d}(\theta) = \theta^{1/\lambda}\lambda\left(\frac{2p}{d(p-1)}\chi_{d,p}\right)^{\frac{\lambda-1}{\lambda}} \qquad \theta \in (0,\infty), \tag{1.14}$$

where

$$\chi_{d,p} = \inf\Big\{\frac{1}{2}\|\nabla g\|_2^2 \colon g \in L^{2p} \cap L^2 \cap H^1 \colon \|g\|_2 = 1 = \|g\|_{2p}\Big\}. \tag{1.15}$$

It turned out in [GKS07, Lemma 2.1] that $\chi_{d,p}$ is positive if and only if $d(p-1) \leq 2p$, i.e., in particular in the cases considered in the present paper. This implies in particular, that $\rho^{(\mathrm{c})}_{p,d}(\theta)$ is finite and positive for any $\theta > 0$. Because of (1.12), also $\rho^{(\mathrm{d})}_{p,d}(\theta)$ is finite and positive, for any sufficiently small $\theta \in (0,\infty)$. By monotonicity, it is positive for any $\theta \in (0,\infty)$. It is also finite (even not larger than θ), since $J(\mu) \geq 0$ and $\|\mu\|_p \leq 1$ for any $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$.

Let us now prove (1.14). In the definition (1.9) of $\rho^{(c)}_{p,d}(\theta)$, we replace g, for any $\beta \in (0,\infty)$, with $\beta^{d/2} g(\beta\,\cdot)$, which is also L^2-normalized. This gives, for any $\beta > 0$,

$$\rho^{(c)}_{p,d}(\theta) = \sup_{g\in H^1\colon \|g\|_2=1} \Big\{\theta\beta^{\frac{d(p-1)}{p}} \|g^2\|_p - \frac{1}{2}\beta^2 \|\nabla g\|_2^2\Big\}.$$

Picking the optimal value

$$\beta^* = \Big(\theta \frac{d(p-1)}{p} \frac{\|g^2\|_p}{\|\nabla g\|_2^2}\Big)^{1/(2\lambda)},$$

we get

$$\rho^{(c)}_{p,d}(\theta) = \theta^{1/\lambda}\lambda \sup_{g\in H^1\colon \|g\|_2=1} \Big(\frac{2p}{d(p-1)}\Big[\frac{\frac{1}{2}\|\nabla g\|_2^2}{\|g\|_{2p}^{2/(1-\lambda)}}\Big]\Big)^{\frac{\lambda-1}{\lambda}}.$$

Note that the term in square brackets remains invariant under the transformation $g \mapsto \beta^{d/2} g(\beta\,\cdot)$, which keeps the L^2-norm fixed. Thus we may freely add the condition that $\|g\|_{2p} = 1$. Recall (1.15) to see that the proof of (1.14) is finished. ◇

Remark 1.4 (Relation to the Gagliardo-Nirenberg constant)**.** In dimensions $d \geq 2$, the constant $\chi_{d,p}$ in (1.15) can be identified in terms of the *Gagliardo-Nirenberg* constant, $K_{d,p}$, as follows. Assume that $d \geq 2$ and $p < \frac{d}{d-2}$. Then $K_{d,p}$ is defined as the smallest constant $C > 0$ in the *Gagliardo-Nirenberg inequality*

$$\|\psi\|_{2p} \leq C \|\nabla\psi\|_2^{\frac{d(p-1)}{2p}} \|\psi\|_2^{1-\frac{d(p-1)}{2p}}, \qquad \text{for } \psi \in H^1(\mathbb{R}^d). \tag{1.16}$$

This inequality received a lot of interest from physicists and analysts, and it has deep connections to Nash's inequality and logarithmic Sobolev inequalities. Furthermore, it also plays an important role of work of Chen [Ch04], [BC04] on intersections of random walks and self-intersections of Brownian motion. See [Ch04, Sect. 2] for more on the Gagliardo-Nirenberg inequality.

It is clear that

$$K_{d,p} = \sup_{\substack{\psi\in H^1(\mathbb{R}^d)\\ \psi\neq 0}} \frac{\|\psi\|_{2p}}{\|\nabla\psi\|_2^{\frac{d(p-1)}{2p}} \|\psi\|_2^{1-\frac{d(p-1)}{2p}}} = \Big(\inf_{\substack{\psi\in H^1(\mathbb{R}^d)\\ \|\psi\|_2=1}} \|\psi\|_{2p}^{-\frac{4q}{d}} \|\nabla\psi\|_2^2\Big)^{-\frac{d}{4q}}. \tag{1.17}$$

Clearly, the term over which the infimum is taken remains unchanged if ψ is replaced by $\psi_\beta(\cdot) = \beta^{\frac{d}{2}}\psi(\cdot\,\beta)$ for any $\beta > 0$. Hence, we can freely add the condition $\|\psi\|_{2p} = 1$ and obtain that

$$K_{d,p} = \chi_{d,p}^{-\frac{d}{4q}}.$$

In particular, the variational formulas for $K_{d,p}$ in (1.17) and for $\chi_{d,p}$ in (1.15) have the same maximizer(s) respectively minimizer(s). It is known that (1.17) has a maximizer, and this is a smooth, positive and rotationally symmetric function (see [We83]). Some uniqueness results are in [MS81]. ◇

Remark 1.5 (Large deviations)**.** In the spirit of the Gärtner-Ellis theorem (see [DZ98, Sect. 4.5]), from Theorem 1.1(ii) large-deviation principles for $\|\ell_t\|_p$ on various scales follow. Indeed, fix some function $(r_t)_{t>0}$ satisfying

$$\left(\frac{\log t}{t}\right)^{\frac{d(p-1)}{p(d+2)}} \ll r_t \ll 1 \qquad \text{as} \qquad t \to \infty. \tag{1.18}$$

Then, as $t \to \infty$, under the conditions $p > 1$ and $d(p-1) < 2$, we have

$$\frac{1}{t}\log \mathbb{P}\big(\big\|\ell_t/t\big\|_p \geq r_t\big) \sim -\chi_{d,p}\frac{d(p-1)}{2p} r_t^{\frac{2p}{d(p-1)}}. \tag{1.19}$$

Applying this to ur_t instead of r_t, one obtains that $\|\ell_t\|_p/(tr_t)$ satisfies a large-deviation principle on the scale $tr_t^{2p/(d(p-1))}$ with strictly convex and continuous rate function $(0,\infty) \ni u \mapsto \chi_{d,p}\frac{d(p-1)}{2p} u^{2p/(d(p-1))}$.

In order to prove the upper bound in (1.19), put $\theta_t = (r_t\lambda/\rho^{(c)}_{d,p}(1))^{\lambda/(1-\lambda)}$ and note that the assumption in (1.8) is satisfied. Now use the exponential Chebyshev inequality to see that

$$\frac{1}{t}\log\mathbb{P}\big(\big\|\ell_t/t\big\|_p \geq r_t\big) \leq \mathbb{E}\left(\mathrm{e}^{\theta_t\|\ell_t\|_p}\right)\mathrm{e}^{-tr_t\theta_t}.$$

Finally use (1.11) and summarize to see that the upper bound in (1.19) is true. The lower bound is derived in a standard way using an exponential change of measure, like in the proof of the Gärtner-Ellis theorem. The point is that the limiting logarithmic moment generating function of $\theta_t\|\ell_t\|_p$ is differentiable throughout $(0,\infty)$, as is seen from (1.10) and (1.11).

However, it is not clear to us from Theorem 1.1(i) whether or not $\|\ell_t\|_p/t$ satisfies a large-deviation principle. Indeed, it is unclear if the map $\rho^{(\mathrm{d})}_{d,p}$ is differentiable since the map $\mu \mapsto \theta\|\mu\|_p - J(\mu)$ is a difference of convex functions and therefore not necessarily strictly convex. As a result, we do not know if the maximiser is uniquely attained. ◇

Remark 1.6. One might wonder what (1.19) might look like in the critical case $p = \frac{d}{d-2}$. Note that the right-hand side is then equal to $-\chi_{d,d/(d-2)}r_t$, which is nontrivial according to [GKS07, Lemma 2.1], recall Remark 1.3. However, in $d \geq 3$, [Ca10, Theorem 2] shows that (1.19) holds, for any $t^{\frac{1}{p}-1} \ll r_t \ll 1$, with $\chi_{d,d/(d-2)}$ replaced by $\sup\{\|\nabla f\|_2^2\colon f \in \ell^{2p}(\mathbb{Z}^d), \|f\|_{2p} = 1\}$, which is a discrete version of $\chi_{d,d/(d-2)}$. This interestingly shows that the critical dimension $d = \frac{2p}{p-1}$ seems to exhibit a different regime and is not the boundary regime of the cases considered here. ◇

Remark 1.7 (Restrictions for r_t)**.** We believe that the large-deviations principle in (1.19) should hold for more functions $(r_t)_{t>0}$ than those that satisfy (1.18), more precisely for all r_t satisfying $a_{d,p}(t) \ll (tr_t)^p \ll t^p$, recall (1.4). In $d \geq 3$, this would mean $t^{(1-p)/p} \ll r_t \ll 1$. Likewise, we believe that also Theorem 1.1(ii) should be true for the corresponding θ_t, i.e., for θ_t of order $r_t^{\lambda/(1-\lambda)}$.

This restriction in our Theorem 1.1(ii) is necessary for a technical reason that comes from the error terms in [BHK07, Theorem 2.1, Prop. 3.6], which is an important ingredient of our proof of the upper bound, see (2.7). Our proof of the lower bound in (1.10) does not use this and is indeed true in greater generality. ◇

Remark 1.8 (Restriction in the dimension)**.** Certainly, we expect also the upper bound in (1.11) to be true for any $p > 1$ satisfying $d(p-1) < 2p$. However, for some technical reason, the method of our proof does not seem to give this. The point is that in the proof of the statement in (1.12), we have to approximate a certain step function with its interpolating polygon line in L^{2p}-sense, and the difference is essentially equal to the gradient of the polygon line. A control in L^2-sense is possible by comparison to the energy term, but the required L^{2p}-control represents a problem that we did not overcome to full extent, see (2.23). ◇

Remark 1.9. A partial result in direction of the statement in (1.10) has been derived by Xia Chen [Ch04]. Let $\psi\colon \mathbb{R}^d \to [0,\infty)$ be an L^1-normalized, smooth function, and use $\psi_\varepsilon(\cdot) = \varepsilon^d\psi(\,\cdot\,\varepsilon)$ as an approximation of the Dirac measure at zero as $\varepsilon \downarrow 0$. Then [Ch04, Theorem 3.1, (3.3)] states that, for $p \in \mathbb{N}$ with $p > 1$ and $d(p-1) < 2p$, and for any $\varepsilon, \theta > 0$,

$$\lim_{t\to\infty}\frac{\alpha_t^2}{t}\log\mathbb{E}\Big[\exp\Big\{\theta k_t\big\|L_t * \psi_\varepsilon\big\|_p\Big\}\Big] = \sup\Big\{\theta\big\|g^2 * \psi_\varepsilon\big\|_p - \frac{1}{2}\|\nabla g\|_2^2\colon g \in H^1, \|g\|_2 = 1\Big\}, \tag{1.20}$$

where $\alpha_t = \theta_t^{-1/(2\lambda)}$ and $L_t(x) = \frac{\alpha_t^d}{t}\ell_t\big(\lfloor x\alpha_t\rfloor\big)$ for $x \in \mathbb{R}^d$. It is elementary to derive that the right hand side of (1.20) for $\varepsilon = 0$ (interpreting $\|g^2 * \psi_0\|_p$ as $\|g^2\|_p$) is equal to $\rho^{(c)}_{d,p}(\theta)$ on the right hand side of (1.10). In Section 1.5 we see that the whole statement in (1.20) for $\varepsilon = 0$ and $\theta = 1$ is equivalent to our statement in (1.10) and (1.11). Hence, (1.20) can be seen as a smoothed version of Theorem 1.1(ii).

The proof of (1.20) in [Ch04] uses a sophisticated, technically rather involved functional analytic method, which reduces the problem to its Brownian analogue and makes it possible to use the well-known large-deviation principle for the occupation times measures of Brownian motion.

However, we did not find any way to apply the result in (1.20) for the proof of Theorem 1.1. Indeed, we use an alternate route. ◇

Remark 1.10. In [HKM06, Prop. 2.1], for any $p > 1$ satisfying $d(p-1) < 2p$, it is shown that

$$\limsup_{\theta \downarrow 0} \limsup_{t \to \infty} \frac{1}{t\theta_t^{1/\lambda}} \log \mathbb{E}\left(\mathrm{e}^{\theta\theta_t \|\ell_t\|_p}\right) \leq 0,$$

under the assumption

$$t^{-\frac{2\lambda}{d+1}-\varepsilon} \ll \theta_t \ll 1 \qquad \text{for some } \varepsilon > 0,$$

which is slightly less restrictive than our assumption in (1.8). The method of the proof is based on combinatorics; high polynomial moments of $\|\ell_t\|_p$ where asymptotically evaluated. ◇

1.4 Literature remarks

For decades, and particularly in this millenium, there is an active interest in the deviations of self-intersection local times and their connections with the theory of large deviations. See the recent monograph [Ch09] for a host of related results and concepts. This subject is a rich source of various phenomena that arise, depending on the dimension d, the intersection parameter p and the scale of the deviation, r_t. In spite of this interest, there are not many results that identify the precise logarithmic asymptotics of exponential expectations or of the decay of the probability of a large value of this local time. The reason is that it is difficult to get a precise control on the p-norm of the local times, which is a highly discontinuous object, and furthermore unbounded. Most of the available results identify the logarithmic rate only, but not the precise prefactors. In Theorem 1.1, we identify this prefactor, on the cost of some loss of generality in the scale and in the dimension. In many other investigations, the parameter p is assumed to be an integer, which we do not do.

Various methods have been employed in this field. Le Gall [Le86] introduced a technique of successive division of the path into equally long pieces and controlling the mutual interaction. This method has been further developed by Asselah in a series of papers, out of which we want to mention [A08], [A09], and [AC07]. Another strategy was developed by Xia Chen, who made the application of Donsker-Varadhan's large-deviation technique possible by a sophisticated compactification procedure, which uses a lot of abstract functional analysis and goes back to de Acosta. See [Ch09] for a thorough and self-contained presentation of the field and of his method and results. As mentioned in Remark 1.10 above, a combinatorial method was applied in [HKM06, Prop. 2.1]. Recently, Castell [Ca10] used Dynkin's isomorphism for deriving precise logarithmic asymptotics for the deviations of the intersection local times in $d \geq 3$ for the critical parameter $p = \frac{d}{d-2}$, which is the boundary of our restriction $d(p-1) < 2p$. See the introduction of [Ca10] for an extensive but concise summary of related results.

The present paper uses a new strategy that goes back to a formula for the joint density of the local times of any continuous-time finite-state space Markov chain. The kernel is an explicit upper bound for this density, which basically implies the upper bound in Donsker-Varadhan's large-deviation principle for these local times without using any topology. In this way, one obtains a discrete, t-dependent variational formula, and the main task is to find its large-t asymptotics. This is done via techniques in the spirit of Γ-convergence. An example of this technique was carried out in [HKM06, Section 5].

1.5 Heuristic derivation of Theorem 1.1

We now give a heuristic derivation of Theorem 1.1(ii), which is based on large-deviation theory.

Recall that $\lambda = (2p - dp + d)/2p \in (0,1)$. For some scale function $\alpha_t \to \infty$, to be specified later, define the random step function $L_t\colon \mathbb{R}^d \to [0,\infty)$ as the scaled normalized version of the local times ℓ_t, i.e.,

$$L_t(x) = \frac{\alpha_t^d}{t}\ell_t\big(\lfloor x\alpha_t\rfloor\big), \qquad \text{for } x \in \mathbb{R}^d. \tag{1.21}$$

Then L_t is a random element of the set

$$\mathcal{F} = \Big\{f \in L^1\colon f \geq 0, \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = 1\Big\} \tag{1.22}$$

of all probability densities on $\mathbb{R}^d$. In the spirit of the celebrated large-deviation theorem of Donsker and Varadhan, if α_t satisfies $1 \ll \alpha_t^d \ll a_{d,0}(t)$ (see (1.4)), then the distributions of L_t satisfy a weak large-deviation principle in the weak L^1-topology on $\mathcal{F}$ with speed $t\alpha_t^{-2}$ and rate function $\mathcal{I}\colon \mathcal{F} \to [0,\infty]$ given by

$$\mathcal{I}(f) = \begin{cases} \frac{1}{2}\big\|\nabla\sqrt{f}\big\|_2^2 & \text{if } \sqrt{f} \in H^1(\mathbb{R}^d), \\ \infty & \text{otherwise.} \end{cases} \tag{1.23}$$

Roughly, this large-deviation principle says that,

$$\mathbb{P}(L_t \in \,\cdot\,) = \exp\Big\{-\frac{t}{\alpha_t^2}\Big[\inf_{f\in\,\cdot}\mathcal{I}(f) + o(1)\Big]\Big\}, \tag{1.24}$$

and the convergence takes place in the weak topology. This principle has been partially proved in a special case in [DV79], a proof in the general case was given in [HKM06, Prop. 3.4].

In order to heuristically recover Theorem 1.1(ii) in terms of the statement in (1.24), note that

$$\theta_t\|\ell_t\|_p = \theta_t\Big(\sum_{z\in\mathbb{Z}^d}\ell_t(z)^p\Big)^{1/p} = t\theta_t\alpha_t^{-d}\Big(\sum_{z\in\mathbb{Z}^d}L_t\big(\tfrac{z}{\alpha_t}\big)^p\Big)^{1/p} = t\theta_t\alpha_t^{\frac{d(1-p)}{p}}\|L_t\|_p.$$

Now we choose $\alpha_t = \theta_t^{-1/(2\lambda)}$ and have therefore that

$$\theta_t\|\ell_t\|_p = \frac{t}{\alpha_t^2}\|L_t\|_p \qquad \text{and} \qquad t\theta_t^{1/\lambda} = \frac{t}{\alpha_t^2}. \tag{1.25}$$

Therefore, the scale t/α_t^2 of the large-deviation principle coincides with the logarithmic scale $t\theta_t^{1/\lambda}$ of the expectation under interest in Theorem 1.1(ii). A formal application of Varadhan's lemma yields

$$\mathbb{E}\Big(\mathrm{e}^{\theta_t\|\ell_t\|_p}\Big) = \mathbb{E}\Big(\exp\Big\{\frac{t}{\alpha_t^2}\|L_t\|_p\Big\}\Big) = \exp\Big\{\frac{t}{\alpha_t^2}(\widetilde{\rho} + o(1))\Big\},$$

where

$$\begin{aligned}\widetilde{\rho} &= \sup\Big\{\|f\|_p - \mathcal{I}(f)\colon f \in \mathcal{F}\Big\} \\ &= \sup\big\{\|g^2\|_p - \frac{1}{2}\|\nabla g\|_2^2 \,:\, g \in L^2 \cap L^{2p} \cap H^1,\ \|g\|_2 = 1\big\} \\ &= \rho_{d,p}^{(\mathrm{c})}(1).\end{aligned} \tag{1.26}$$

This ends the heuristic derivation of Theorem 1.1(ii). In the same way, one can derive also Theorem 1.1(i); this is similar to the line of argument used in [GM98].

Hence, we see informally that, in Theorem 1.1(ii), the main contribution to the exponential moments should come from those random walk realisations that make the rescaled local times, L_t, look like the minimiser(s) of the variational formula $\rho_{d,p}^{(\mathrm{c})}(1)$. In particular, the random walk should stay within a region with diameter $\alpha_t \ll t^{1/d}$, and each local time should be of order $t/\alpha_t^d \gg 1$. That is, there is a time-homogeneous sqeezing strategy. In Theorem 1.1(i), the interpretation is analogous, but the diameter of the preferred region is now of finite order in t. This is why a discrete picture arises in the variational formula $\rho_{d,p}^{(\mathrm{d})}(1)$.

There are several serious obstacles to be removed when trying to turn the above heuristics into an honest proof: (1) the large-deviation principle only holds on compact subsets of $\mathbb{R}^d$, (2) the functional $L_t \mapsto \|L_t\|_p$ is not bounded, and (3) this functional is not continuous. Removing the obstacle (1) is easy and standard, but it is in general notoriously difficult to overcome the obstacles (2) and (3) for related problems.

2. Proof of Theorem 1.1

We prove Theorem 1.1(i) (that is, (1.6)) in Section 2.1, the lower-bound part (1.10) of Theorem 1.1(ii) in Section 2.2 and the upper-bound part (1.11) in Section 2.3.

2.1 Proof of (1.6)

This is analogous to the proof of [GM98, Theorem 1.2]; we will sketch the argument. First we explain the lower bound. Let Q_R denote the box $[-R,R]^d \cap \mathbb{Z}^d$ and insert an indicator on the event $\{\mathrm{supp}(\ell_t) \subset Q_R\}$ in the expectation, to get, for any $\theta > 0$,

$$\mathbb{E}\left(\mathrm{e}^{\theta\|\ell_t\|_p}\right) \geq \mathbb{E}\left(\mathrm{e}^{\theta t\|\ell_t/t\|_p}\mathbb{1}_{\{\mathrm{supp}(\ell_t)\subset Q_R\}}\right).$$

Now observe that the functional $\mu \mapsto \|\mu\|_p$ is continuous and bounded on the set $\mathcal{M}_1(Q_R)$ of all probability measures on Q_R. Furthermore, under the sub-probability measure $\mathbb{P}(\cdot, \mathrm{supp}(\ell_t) \subset Q_R)$, the distributions of ℓ_t/t satisfy a large-deviations principle with scale t and rate function equal to the restriction of J defined in Theorem 1.1(i) to $\mathcal{M}_1(Q_R)$. Hence, Varadhan's lemma [DZ98, Lemma 4.3.4] yields that

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left(\mathrm{e}^{\theta\|\ell_t\|_p}\right) \geq \sup_{\mu\in\mathcal{M}_1(Q_R)} \big(\theta\|\mu\|_p - J(\mu)\big). \tag{2.1}$$

Letting $R \to \infty$ and using an elementary approximation argument, we see that the right-hand side converges towards $\rho^{(\mathrm{d})}_{d,p}(\theta)$. This ends the proof of the lower bound in (1.6).

Now we explain the upper bound. Introduce the periodized version of the local times in Q_R,

$$\ell^{(R)}_t(z) = \sum_{x\in\mathbb{Z}^d} \ell_t(z + Rx), \qquad t \in (0,\infty), R \in \mathbb{N}, z \in \mathbb{Z}^d. \tag{2.2}$$

Then it is easy to see that $\|\ell_t\|_p \leq \|\ell^{(R)}_t\|_p$. Hence, for any $\theta \in (0,\infty)$,

$$\mathbb{E}\big(\mathrm{e}^{\theta\|\ell_t\|_p}\big) \leq \mathbb{E}\big(\mathrm{e}^{\theta\|\ell^{(R)}_t\|_p}\big) = \mathbb{E}\big(\mathrm{e}^{\theta t\|\ell^{(R)}_t/t\|_p}\big).$$

It is well-known that $(\ell^{(R)}_t/t)_{t>0}$ satisfies a large-deviations principle on the set of probability measures on Q_R with rate function $\mu \mapsto J_{R,\mathrm{per}}(\mu)$ equal to the Dirichlet form at $\sqrt{\mu}$ of $-\frac{1}{2}\Delta$ in Q_R with periodic boundary condition. By continuity and boundedness of the map $\mu \mapsto \|\mu\|_p$, it is clear from Varadhan's lemma that

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left(\mathrm{e}^{\theta\|\ell_t\|_p}\right) \leq \sup_{\mu\in\mathcal{M}_1(Q_R)} \big(\theta\|\mu\|_p - J_{R,\mathrm{per}}(\mu)\big). \tag{2.3}$$

In the same way as in the proof of [GM98, Lemma 1.10], one shows that the difference between the variational formulas on the right-hand sides of (2.3) and (2.1) vanish in the limit as $R \to \infty$. This ends the proof of (1.6).

2.2 Proof of the lower bound (1.10)

Fix $q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, and consider a continuous bounded function $f\colon \mathbb{R}^d \to \mathbb{R}$ such that $\|f\|_q = 1$. According to Hölder's inequality, we have

$$\|L_t\|_p \geq \langle f, L_t\rangle.$$

Recall from Section 1.5 that $\alpha_t = \theta_t^{-1/(2\lambda)}$. Using (1.25), we obtain, for any $R > 0$, the lower bound

$$\mathbb{E}\left(\mathrm{e}^{\theta_t \|\ell_t\|_p}\right) = \mathbb{E}\left(\mathrm{e}^{t\alpha_t^{-2}\|L_t\|_p}\right) \geq \mathbb{E}\left(\mathrm{e}^{t\alpha_t^{-2}\langle f, L_t\rangle} \mathbb{1}_{\{\mathrm{supp}(L_t)\subset B_R\}}\right), \tag{2.4}$$

where we denote $B_R = [-R, R]^d$. According to [GKS07, Lemma 3.2], the distributions of L_t under the sub-probability measure $\mathbb{P}(\cdot\,, \mathrm{supp}(L_t) \subset B_R)$ satisfy, as $t \to \infty$, a large-deviation principle on the set of probability densities on $\mathbb{R}^d$ with support in B_R. The rate function is

$$g^2 \mapsto \frac{1}{2}\|\nabla g\|_2^2, \qquad g \in H^1, \|g\|_2 = 1, \mathrm{supp}(g) \subset B_R,$$

and we put the value of the rate function equal to $+\infty$ if g is not in H^1 or not normalized or if its support is not contained in B_R. The speed is $t\alpha_t^{-2}$, which is identical to the logarithmic scale in (1.10), $t\theta_t^{1/\lambda}$. Since the map $L_t \mapsto \langle f, L_t\rangle$ is continuous, we obtain, by Varadhan's lemma, from that large-deviation principle the following estimate for the exponential moments:

$$\liminf_{t\to\infty} \frac{\alpha_t^2}{t} \log \mathbb{E}\left(\mathrm{e}^{\theta_t \|\ell_t\|_p}\right) \geq \sup_{\substack{g\in H^1\colon \|g\|_2=1 \\ \mathrm{supp}(g)\subset B_R}} \left(\langle f, g^2\rangle - \frac{1}{2}\|\nabla g\|_2^2\right).$$

Certainly we can restrict the supremum over g to $g \in L^{2p}$. Since the left-hand side does not depend on f, we can take on the right-hand side the supremum over all continuous bounded $f\colon B_R \to \mathbb{R}$ satisfying $\|f\|_q = 1$. Using an elementary approximation argument and the duality between L^q and L^p, we see that

$$\sup_{f\in\mathcal{C}(B_R), \|f\|_q=1} \sup_{\substack{g\in H^1\colon \|g\|_2=1 \\ \mathrm{supp}(g)\subset B_R}} \left(\langle f, g^2\rangle - \frac{1}{2}\|\nabla g\|_2^2\right) \geq \sup_{\substack{g\in H^1\cap L^{2p}\colon \|g\|_2=1 \\ \mathrm{supp}(g)\subset B_R}} \left(\|g^2\|_p - \frac{1}{2}\|\nabla g\|_2^2\right).$$

Letting $R \to \infty$ and using another elementary approximation argument, we see that the right-hand side converges to $\rho_{d,p}^{(\mathrm{c})}(1)$. This ends the proof of the lower bound in (1.10).

2.3 Proof of the upper bound (1.11)

Fix $\theta_t \in (0, \infty)$ satisfying (1.8). Recall from Section 1.5 that $\lambda = (2p - dp + d)/2p \in (0, 1)$ and that $\alpha_t = \theta_t^{-1/(2\lambda)}$. As in the proof of (1.6), we estimate from above against a periodized version of the walk, but now in the t-dependent box $Q_{R\alpha_t} = [-R\alpha_t, R\alpha_t]^d \cap \mathbb{Z}^d$. Recall that $\ell_t^{(R\alpha_t)}$ denotes the periodized version of the local times, see (2.2). We estimate

$$\mathbb{E}\left(\mathrm{e}^{\theta_t \|\ell_t\|_p}\right) \leq \mathbb{E}\left(\mathrm{e}^{\theta_t \|\ell_t^{(R\alpha_t)}\|_p}\right) = \mathbb{E}\left(\exp\left\{t\alpha_t^{-2\lambda} \left\|\tfrac{1}{t}\ell_t^{(R\alpha_t)}\right\|_p\right\}\right). \tag{2.5}$$

Note that $\ell_t^{(R\alpha_t)}$ is the local time vector of the continuous-time random walk on $Q_{R\alpha_t}$ with generator $A_{R\alpha_t}$, which is $\frac{1}{2}$ times the Laplace operator in $Q_{R\alpha_t}$ with periodic boundary condition.

Now we employ a recently developed method for effectively deriving large-deviation upper bounds without continuity and boundedness assumptions. The base of this method has been laid in [BHK07] and has been applied first in [BHK07, Theorem 3.7] and [HKM06, Section 5]. The main point is the identification of a joint density of the local time vector $\ell_t^{(R\alpha_t)}$ and of an explicit upper bound for this density. In this way, no continuity or boundedness is required, which is a great improvement over classical large-deviations arguments. The upper bound is in terms of a discrete-space variational formula and additional error terms involving the box size. Let us remark that these error terms give us the lower restriction for θ_t in (1.8). The main work after the application of the upper bound is to derive the large-t asymptotics of the discrete variational formula, which requires Gamma-convergence techniques.

We apply [BHK07, Theorem 3.6] to get, for any $t \geq 1$,

$$\begin{aligned}\log \mathbb{E}\left(\exp\left\{t\alpha_t^{-2\lambda}\big\|\tfrac{1}{t}\ell_t^{(R\alpha_t)}\big\|_p\right\}\right) \leq t \sup_{\mu\in\mathcal{M}_1(Q_{R\alpha_t})}\Big[\alpha_t^{-2\lambda}\|\mu\|_p - \big\|(-A_{R\alpha_t})^{1/2}\sqrt{\mu}\big\|_2^2\Big] \\ + |Q_{R\alpha_t}|\log\left(2d\sqrt{8\mathrm{e}}\,t\right) + \log|Q_{R\alpha_t}| + \frac{|Q_{R\alpha_t}|}{4t}\end{aligned} \tag{2.6}$$

Here we have used that $\eta_{Q_{R\alpha_t}}$, defined in [BHK07, (3.2)], is equal to $2d$.

Let us first show that the terms in the second line on the right-hand side are asymptotically negligible on the scale $t/\alpha_t^2 = t\theta_t^{1/\lambda}$. Indeed, these terms are of order $\alpha_t^d \log t$, and we see that

$$\alpha_t^d \log t = \frac{t}{\alpha_t^2}\alpha_t^{d+2}\frac{\log t}{t} = \frac{t}{\alpha_t^2}\theta_t^{-\frac{d+2}{2\lambda}}\frac{\log t}{t} \ll \frac{t}{\alpha_t^2}, \tag{2.7}$$

where we used (1.8).

Hence, substituting this in (2.5) and (2.6), it is clear that, for the proof of the upper bound in (1.11), it is sufficient to prove the following.

Lemma 2.1.

$$\limsup_{R\to\infty}\limsup_{t\to\infty}\alpha_t^2 \sup_{\mu\in\mathcal{M}_1(Q_{R\alpha_t})}\Big[\alpha_t^{-2\lambda}\|\mu\|_p - \big\|(-A_{R\alpha_t})^{1/2}\sqrt{\mu}\big\|_2^2\Big] \leq \rho_{p,d}^{(\mathrm{c})}(1). \tag{2.8}$$

Proof. We will adapt the method described in [HKM06, Prop. 5.1]. First, we pick sequences $R_n \to \infty$, $t_n \to \infty$ and $\mu_n \in \mathcal{M}_1(B_{R_n\alpha_{t_n}})$ such that

$$\text{L.h.s. of (2.8)} \leq \alpha_{t_n}^{d(p-1)/p}\|\mu_n\|_p - \alpha_{t_n}^2\big\|\big(-A_{R_n\alpha_{t_n}}\big)^{1/2}\sqrt{\mu_n}\big\|_2^2 + \tfrac{1}{n}, \qquad n\in\mathbb{N}. \tag{2.9}$$

We may assume that μ_n is a probability measure on $\mathbb{Z}^d$ with support in $B_{R_n\alpha_{t_n}}$.

In the following, we will construct a sequence $(h_n)_n$ in H^1 such that (1) h_n is L^2-normalized, (2) the term $\alpha_{t_n}^2\|(-A_{R_n\alpha_{t_n}})^{1/2}\sqrt{\mu_n}\|_2^2$ is approximately equal to its energy, $\frac{1}{2}\|\nabla h_n\|_2^2$, and (3) the term $\alpha_{t_n}^{d(p-1)/p}\|\mu_n\|_p$ is approximately equal to $\|h_n^2\|_p$. Having constructed such a series, the proof is quickly finished.

We are using finite-element methods to construct such function h_n, see [B07] for the general theory. We split $\mathbb{R}^d$ along the integer grid into half-open unit cubes $C(k) = \times_{i=1}^d(k_i, k_i+1]$ with $k = (k_1,\dots,k_d) \in \mathbb{Z}^d$. Each such cube is split into $d!$ 'tetrahedra' as follows. For $\sigma \in \mathfrak{S}_d$, the set of permutations of $1,\dots,d$, we denote by $T_\sigma(k)$ the intersection of $C(k)$ with the convex hull generated by $k, k+\mathrm{e}_{\sigma(1)},\dots,k+\mathrm{e}_{\sigma(1)}+\ldots+\mathrm{e}_{\sigma(d)}$, where e_i denotes the i-th unit vector in $\mathbb{R}^d$. Up to the boundary, the tetrahedra $T_\sigma(k)$ with $\sigma\in\mathfrak{S}_d$ are pairwise disjoint. One can easily see that, for $x \in C(k)$,

$$x \in T_\sigma(k) \qquad\Longleftrightarrow\qquad x_{\sigma(1)} - \lfloor x_{\sigma(1)}\rfloor \geq \ldots \geq x_{\sigma(d)} - \lfloor x_{\sigma(d)}\rfloor > 0.$$

The interior of $T_\sigma(k)$ is characterised by strict inequalities. A site x belongs to two different of these tetrahedra if and only if at least one of the inequalities is an equality.

Now we introduce the following functions. For $n\in\mathbb{N}$, $i\in\{1,\dots,d\}$ and $y\in\mathbb{R}^d$, put

$$f_{n,\sigma,i}(y) = \Big[\sqrt{\mu_n\big(\lfloor y\rfloor + \mathrm{e}_{\sigma(1)}+\cdots+\mathrm{e}_{\sigma(i)}\big)} - \sqrt{\mu_n\big(\lfloor y\rfloor + \mathrm{e}_{\sigma(1)}+\cdots+\mathrm{e}_{\sigma(i-1)}\big)}\Big]\,(y_{\sigma(i)} - \lfloor y_{\sigma(i)}\rfloor). \tag{2.10}$$

Furthermore, given $k\in\mathbb{Z}^d$ and $y\in C(k)$, we pick some $\sigma(y)\in\mathfrak{S}_d$ such that $y\in T_{\sigma(y)}(k)$ and define

$$g_n(x) = \alpha_{t_n}^{d/2}\mu_n(\lfloor\alpha_{t_n}x\rfloor)^{1/2} + \alpha_{t_n}^{d/2}\sum_{i=1}^d f_{n,\sigma(\alpha_{t_n}x),i}(\alpha_{t_n}x). \tag{2.11}$$

Now we show that g_n is well-defined, i.e., that if y lies in $T_{\sigma_1}(k)$ and in $T_{\sigma_2}(k)$, then

$$\sum_{i=1}^{d} f_{n,\sigma_1,i}(y) = \sum_{i=1}^{d} f_{n,\sigma_2,i}(y). \tag{2.12}$$

For sake of simplicity, we do this only for the special case

$$\sigma_1(1) = 1 = \sigma_2(2) \qquad \text{and} \qquad \sigma_1(2) = 2 = \sigma_2(1), \qquad \text{and } \sigma_1(i) = \sigma_2(i) \text{ for } i \geq 3.$$

That is, $y_1 - \lfloor y_1 \rfloor = y_2 - \lfloor y_2 \rfloor > y_i - \lfloor y_i \rfloor$ for any $i \in \{3, \dots, d\}$. We calculate

$$\begin{aligned}
\sum_{i=1}^{d} f_{n,\sigma_1,i}(y) - \sum_{i=1}^{d} f_{n,\sigma_2,i}(y) = {} & \Big[\sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_1\big)} - \sqrt{\mu_n\big(\lfloor y \rfloor\big)}\Big]\,(y_1 - \lfloor y_1 \rfloor) \\
& + \Big[\sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_1 + \mathrm{e}_2\big)} - \sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_1\big)}\Big]\,(y_2 - \lfloor y_2 \rfloor) \\
& - \Big[\sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_2\big)} - \sqrt{\mu_n\big(\lfloor y \rfloor\big)}\Big]\,(y_2 - \lfloor y_2 \rfloor) \\
& - \Big[\sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_2 + \mathrm{e}_1\big)} - \sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_2\big)}\Big]\,(y_1 - \lfloor y_1 \rfloor).
\end{aligned}$$

This is equal to zero, since $y_1 - \lfloor y_1 \rfloor = y_2 - \lfloor y_2 \rfloor$. Hence, we know that g_n is well-defined. Furthermore, this also shows that g_n is continuous within each $C(k)$. From now on, we abbreviate $f_{n,\sigma(y),i}(y)$ by $f_{n,\sigma,i}(y)$.

Similarly, we see that g_n is also continuous at the boundary of each of the cubes. Indeed, a site $y \in \mathbb{R}^d$ belongs to this boundary if and only if it has at least one integer coordinate. For the sake of simplicity assume that only for $i = 1$ it holds that $y_i - \lfloor y_i \rfloor = 1$. It is clear that $y \in T_{\sigma_1}(k) \cap \overline{T_{\sigma_2}(k + \mathrm{e}_1)}$ where $\sigma_1, \sigma_2 \in \mathfrak{S}_d$ are given by

$$\sigma_1(1) = 1, \sigma_2(d) = 1 \text{ and } \sigma_1(i+1) = \sigma_2(i) \qquad \forall i \in \{1, \dots, d-1\}$$

Choose an arbitrary sequence $(y^{(m)})_m \in T_{\sigma_2}(k + \mathrm{e}_1)$ that converges to y. For sufficiently large m it holds that $\lfloor y \rfloor + \mathrm{e}_1 = \lfloor y^{(m)} \rfloor$. We see now that:

$$\begin{aligned}
\mu_n(\lfloor y^{(m)} \rfloor)^{1/2} + \sum_{i=1}^{d} f_{n,\sigma_2,i}(y^{(m)}) = {} & \mu_n(\lfloor y \rfloor + \mathrm{e}_1)^{1/2} + \sum_{i=1}^{d} (y^{(m)}_{\sigma_2(i)} - \lfloor y^{(m)}_{\sigma_2(i)} \rfloor) \\
& \times \Big[\sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_1 + \mathrm{e}_{\sigma_2(1)} + \cdots + \mathrm{e}_{\sigma_2(i)}\big)} - \sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_1 + \mathrm{e}_{\sigma_2(1)} + \cdots + \mathrm{e}_{\sigma_2(i-1)}\big)}\Big].
\end{aligned}$$

Note that the summand for $i = d$ converges to zero, since $\lim_{m\to\infty}(y^{(m)}_{\sigma_2(d)} - \lfloor y^{(m)}_{\sigma_2(d)} \rfloor) = 0$. In the remaining sum on $i = 1, \dots, d-1$, we shift the sum by substituting $i = j - 1$ and replace $\sigma_2(j-1)$ by $\sigma_1(j)$, to get, as $m \to \infty$,

$$\begin{aligned}
& \mu_n(\lfloor y^{(m)} \rfloor)^{1/2} + \sum_{i=1}^{d} f_{n,\sigma_2,i}(y^{(m)}) \\
& \quad = \mu_n(\lfloor y \rfloor)^{1/2} + \Big[\mu_n(\sqrt{\lfloor y \rfloor + \mathrm{e}_1}) - \sqrt{\mu(\lfloor x \rfloor)}\Big](y^{(m)}_{\sigma_1(1)} - \lfloor y_{\sigma_1(1)} \rfloor) \\
& \qquad + \sum_{j=2}^{d} \Big[\sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_{\sigma_1(1)} + \mathrm{e}_{\sigma_1(2)} + \cdots + \mathrm{e}_{\sigma_1(j)}\big)} - \sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_{\sigma_1(1)} + \mathrm{e}_{\sigma_1(2)} + \cdots + \mathrm{e}_{\sigma_1(j-1)}\big)}\Big] \\
& \qquad\qquad \times (y^{(m)}_{\sigma_1(j)} - \lfloor y_{\sigma_1(j)} \rfloor) + o(1) \\
& \quad = \mu_n(\lfloor y \rfloor)^{1/2} + \sum_{j=1}^{d} \Big[\sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_{\sigma_1(1)} + \cdots + \mathrm{e}_{\sigma_1(j)}\big)} - \sqrt{\mu_n\big(\lfloor y \rfloor + \mathrm{e}_{\sigma_1(1)} + \cdots + \mathrm{e}_{\sigma_1(j-1)}\big)}\Big] \\
& \qquad\qquad \times (y^{(m)}_{\sigma_1(j)} - \lfloor y_{\sigma_1(j)} \rfloor) + o(1).
\end{aligned}$$

As $y_j^{(m)} \to y_j$, we see that the last term converges towards $\mu_n(\lfloor y \rfloor)^{1/2} + \sum_{j=1}^d f_{n,\sigma_2,j}(y)$. Hence, we proved the continuity of g_n at the border of each $C(k)$ and thus the continuity of g_n on the entire $\mathbb{R}^d$. In addition, as g_n is clearly differentiable in the interior of each tetrahedron $T_\sigma(k)$, we get that g_n lies in H^1. An elementary calculation shows that

$$\|\nabla g_n\|_2^2 = \alpha_{t_n}^2 \big\| \big(-A_{R_n \alpha_{t_n}}\big)^{1/2} \sqrt{\mu_n} \big\|_2^2, \qquad n \in \mathbb{N}. \tag{2.13}$$

Furthermore, we will prove at the end of the proof of this lemma that, for any $n \in \mathbb{N}$,

$$\alpha_{t_n}^{d(p-1)/p} \|\mu_n\|_p \le \|g_n^2\|_p \big[1 + C\alpha_{t_n}^{[d(p-1)-2]/2p}\big] + \big[\|\nabla g_n\|_2^2 + 1\big] C \big[\alpha_{t_n}^{[d(p-1)-2]/2p} + \alpha_{t_n}^{[d(p-1)-2]/p}\big], \tag{2.14}$$

where C depends on d and p only. Our assumption $d(p-1) < 2$ implies that the exponents at α_{t_n} on the right-hand side are negative. Recalling that α_{t_n} tends to infinity as $t_n \to \infty$ we have (at the cost of choosing a subsequence of $(R_n, t_n, \mu_n)_n$):

$$\alpha_{t_n}^{d(p-1)/p} \|\mu_n\|_p \le \|g_n^2\|_p (1 + \tfrac{1}{n}) + \tfrac{1}{n} \big(\|\nabla g_n\|_2^2 + 1\big). \tag{2.15}$$

Note that g_n asymptotically satisfies periodic boundary condition in the box $[-R, R]^d$. Now we compare it to some version that satisfies zero boundary condition. To this end, introduce $\Psi_{R_n} = \bigotimes_{i=1}^d \psi_{R_n} \colon \mathbb{R}^d \to [0,1]$, where $\psi_{R_n} \colon \mathbb{R} \to [0,1]$ is zero outside $[-R_n, R_n]$, one in $[-R_n + R_n^\varepsilon, R_n - R_n^\varepsilon]$ and linearly interpolates between $-R_n$ and $-R_n + R_n^\varepsilon$ and between $R_n - R_n^\varepsilon$ and R_n. Here $\varepsilon > 0$ is picked so small that $p < \frac{d}{d-1+\varepsilon}$, where we point out that our assumption $d(p-1) < 2$ also implies $p < \frac{d}{d-1}$. We are going to estimate the changes of the functionals when going from g_n to $g_n \Psi_{R_n}$. We first use the triangle inequality in (2.11) and note that the L^2-norm of $x \mapsto \alpha_{t_n}^{d/2} \sum_{i=1}^d f_{n,\sigma,i}(\alpha_{t_n} x)$ is not larger than $C\alpha_{t_n}^{-1} \|\nabla g_n\|_2$, where $C \in (0, \infty)$ depends on d only, to see that

$$\|g_n\|_2 \le \Big\| \sqrt{\alpha_{t_n} \mu_n\big(\lfloor \alpha_{t_n} \cdot \rfloor\big)} \Big\|_2 + \big\| \alpha_{t_n}^{d/2} \sum_{i=1}^d f_{n,\sigma,i}(\alpha_{t_n} \cdot) \big\|_2 \le 1 + C\alpha_{t_n}^{-1} \|\nabla g_n\|_2.$$

Analogously we derive $\|g_n\|_2 \ge 1 - C\alpha_{t_n}^{-1} \|\nabla g_n\|_2$ and thus we get

$$\big| \|g_n\|_2 - 1 \big| \le C\alpha_{t_n}^{-1} \|\nabla g_n\|_2, \qquad n \in \mathbb{N}. \tag{2.16}$$

Using Schwarz's inequality and using n so large that $dR_n^{-\varepsilon} \le 1$, this allows us to show that

$$\begin{aligned}
\|\nabla(g_n \Psi_{R_n})\|_2^2 &\le \int_{\mathbb{R}^d} \sum_{i=1}^d \big| \partial_i g_n(x) \Psi_{R_n}(x) + g_n(x) \partial_i \Psi_{R_n}(x) \big|^2 \mathrm{d}x \\
&\le \|\nabla g_n\|_2^2 + 2R_n^{-\varepsilon} \|g_n\|_2 \|\nabla g_n\|_2 + dR_n^{-2\varepsilon} \|g_n\|_2^2 \\
&\le \|\nabla g_n\|_2^2 + 2R_n^{-\varepsilon} \tfrac{1}{2} \big(\|g_n\|_2^2 + \|\nabla g_n\|_2^2\big) + dR_n^{-2\varepsilon} \|g_n\|_2^2 \\
&\le \|\nabla g_n\|_2^2 \big[1 + R_n^{-\varepsilon}\big] + 2R_n^{-\varepsilon} \|g_n\|_2^2 \\
&\le \|\nabla g_n\|_2^2 \big[1 + R_n^{-\varepsilon}\big] + 2R_n^{-\varepsilon} \big(1 + 2C\alpha_{t_n}^{-1} \|\nabla g_n\|_2 + C^2 \alpha_{t_n}^{-2} \|\nabla g_n\|_2^2\big).
\end{aligned}$$

Hence, we have (probably by choosing again a subsequence of $(R_n, t_n, \mu_n)_n$):

$$\|\nabla g_n\|_2^2 \ge \|\nabla(g_n \Psi_{R_n})\|_2^2 (1 - \tfrac{1}{n}) - \tfrac{1}{n}, \qquad n \in \mathbb{N}. \tag{2.17}$$

Furthermore, we note that, without loss of generality, we may assume that

$$\int_{B_{R_n} \setminus B_{R_n - R_n^\varepsilon}} g_n^{2p}(x) \, \mathrm{d}x \le \frac{|B_{R_n} \setminus B_{R_n - R_n^\varepsilon}|}{|B_{R_n}|} \int_{B_{R_n}} g_n^{2p}(x) \, \mathrm{d}x, \qquad n \in \mathbb{N}. \tag{2.18}$$

This can be easily derived using the shift invariance of the second integral due to periodic boundary conditions. To see this, assume that for every shift $\theta_z(x) = x + z$ modulo B_{R_n} with $z \in B_{R_n}$ it holds

that:

$$\int_{B_{R_n}\setminus B_{R_n-R_n^\varepsilon}} g_n^{2p}(\theta_z(x))\,\mathrm{d}x > \frac{|B_{R_n}\setminus B_{R_n-R_n^\varepsilon}|}{|B_{R_n}|}\int_{B_{R_n}} g_n^{2p}(x)\,\mathrm{d}x. \tag{2.19}$$

Now, integrate both sides over all $z\in B_{R_n}$, to get a contradiction by changing the order of the integration. Hence, for some $z\in B_{R_n}$, the opposite of (2.19) holds, and we continue to work with $g_n\circ\theta_z$ instead of g_n. All properties considered so far are preserved by periodicity.

Note that the quotient on the right-hand side of (2.18) can be estimated against $CR_n^{\varepsilon-1}$ where C does not depend on n. Thus, we have:

$$\|g_n^2\|_p^p - \|(g_n\Psi_{R_n})^2\|_p^p \le CR_n^{\varepsilon-1}\|g_n^2\|_p^p$$

which leads (after probably choosing again a subsequence of $(R_n,t_n,\mu_n)_n$) to

$$\|g_n^2\|_p \le \|(g_n\Psi_{R_n})^2\|_p(1+\tfrac{1}{n}),\qquad n\in\mathbb{N}. \tag{2.20}$$

Summarizing, substituting (2.13) and (2.15), and using (2.17) and (2.20), for any n we have

$$\begin{aligned}\text{R.h.s. of (2.9)} &\le \|g_n^2\|_p(1+\tfrac{1}{n}) - \|\nabla g_n\|_2^2(1-\tfrac{2}{n}) - \tfrac{1}{n}\|\nabla g_n\|_2^2 + \tfrac{2}{n}\\ &\le \|(g_n\Psi_{R_n})^2\|_p(1+\tfrac{3}{n}) - \|\nabla(g_n\Psi_{R_n})\|_2^2(1-\tfrac{3}{n}) - \tfrac{1}{n}\|\nabla g_n\|_2^2 + \tfrac{3}{n}\\ &= \Big[\frac{\|(g_n\Psi_{R_n})^2\|_p}{\|g_n\Psi_{R_n}\|_2^2}\,\frac{1+\frac{3}{n}}{1-\frac{3}{n}} - \frac{\|\nabla(g_n\Psi_{R_n})\|_2^2}{\|g_n\Psi_{R_n}\|_2^2}\Big](1-\tfrac{3}{n})\|g_n\Psi_{R_n}\|_2^2 - \tfrac{1}{n}\|\nabla g_n\|_2^2 + \tfrac{3}{n}.\end{aligned} \tag{2.21}$$

Now observe that $h_n = g_n\Psi_{R_n}/\|g_n\Psi_{R_n}\|_2$ is an L^2-normalized element of H^1 and of L^{2p}. Hence, we may estimate the term in the brackets against the supremum over all such functions, which is equal to $\rho_{d,p}^{(\mathrm{c})}(\frac{1+3/n}{1-3/n})$, see (1.9). Since $\rho_{d,p}^{(\mathrm{c})}(\frac{1+3/n}{1-3/n})>0$ by (1.14) and, obviously, $\|g_n\Psi_{R_n}\|_2^2\le\|g_n\|_2^2$, we can proceed with

$$\text{R.h.s. of (2.9)} \le \rho_{d,p}^{(\mathrm{c})}\big(\tfrac{1+3/n}{1-3/n}\big)(1-\tfrac{3}{n})\|g_n\|_2^2 - \tfrac{1}{n}\|\nabla g_n\|_2^2 + \tfrac{3}{n} \le \rho_{d,p}^{(\mathrm{c})}\big(\tfrac{1+3/n}{1-3/n}\big)\|g_n\|_2^2 - \tfrac{1}{n}\|\nabla g_n\|_2^2 + \tfrac{3}{n}.$$

By (2.16) and at the cost of chosing again a subsequence of $(R_n,t_n,\mu_n)_n$, we have that $\|g_n\|_2^2 \le 1+\frac{1}{n}\|\nabla g_n\|_2^2/\rho_{d,p}^{(\mathrm{c})}(\frac{1+3/n}{1-3/n})$. Using this in the last display, we arrive for all n, at

$$\text{R.h.s. of (2.9)} \le \rho_{d,p}^{(\mathrm{c})}\big(\tfrac{1+3/n}{1-3/n}\big) + \tfrac{3}{n}.$$

Recalling (1.14), we see that the right-hand side converges to $\rho_{d,p}^{(\mathrm{c})}(1)$ as $n\uparrow\infty$. This ends the proof of the lemma.

Now we give the proof of (2.14). Recall (2.10) and (2.11) and that we write $f_{n,\sigma,i}(y)$ instead of $f_{n,\sigma(y),i}(y)$. The triangle inequality gives that

$$\begin{aligned}\|g_n\|_{2p} &\ge \alpha_{t_n}^{d/2}\|\mu_n(\lfloor\alpha_{t_n}\cdot\rfloor)^{1/2}\|_{2p} - \alpha_{t_n}^{d/2}\Big\|\sum_{i=1}^d f_{n,\sigma,i}(\alpha_{t_n}\cdot)\Big\|_{2p}\\ &= \alpha_{t_n}^{d(p-1)/2p}\|\mu_n\|_p^{\frac12} - \alpha_{t_n}^{d(p-1)/2p}\Big\|\sum_{i=1}^d f_{n,\sigma,i}\Big\|_{2p}.\end{aligned} \tag{2.22}$$

Now we estimate, for any $y\in\mathbb{R}^d$,

$$\begin{aligned}\Big|\sum_{i=1}^d f_{n,\sigma,i}(y)\Big|^{2p} &\le d^{2p}\sum_{i=1}^d\Big|\sqrt{\mu_n\big(\lfloor y\rfloor+\mathrm{e}_{\sigma(1)}+\cdots+\mathrm{e}_{\sigma(i)}\big)} - \sqrt{\mu_n\big(\lfloor y\rfloor+\mathrm{e}_{\sigma(1)}+\cdots+\mathrm{e}_{\sigma(i-1)}\big)}\Big|^{2p}\\ &\le d^{2p}\sum_{i=1}^d\Big|\sqrt{\mu_n\big(\lfloor y\rfloor+\mathrm{e}_{\sigma(1)}+\cdots+\mathrm{e}_{\sigma(i)}\big)} - \sqrt{\mu_n\big(\lfloor y\rfloor+\mathrm{e}_{\sigma(1)}+\cdots+\mathrm{e}_{\sigma(i-1)}\big)}\Big|^{2}\\ &= d^{2p}\alpha_{t_n}^{-d-2}|\nabla g_n(y/\alpha_{t_n})|^2,\end{aligned} \tag{2.23}$$

since the term in brackets is not larger than one (recall that μ_n is a probability measure on a finite set). Using this in (2.22), we obtain

$$\alpha_{t_n}^{d(p-1)/2p}\|\mu_n\|_p^{\frac{1}{2}} \leq \|g_n\|_{2p} + d\alpha_{t_n}^{[d(p-1)-2]/2p}\|\nabla g_n\|_2^{\frac{1}{p}}.$$

Now square both sides and use the estimates $\|\nabla g_n\|_2^{\frac{1}{p}} \leq 1 + \|\nabla g_n\|_2$ and $\|g_n\|_{2p}\|\nabla g_n\|_2 \leq \frac{1}{2}(\|g_n\|_{2p}^2 + \|\nabla g_n\|_2^2)$ and summarize to arrive at (2.14).

□

References

[A08] A. Asselah, Large deviations estimates for self-intersection local times for simple random walk in $\mathbb{Z}^3$, *Probab. Theory Relat. Fields* **141**, 19-45 (2008).

[A09] A. Asselah, Large deviations principle for self-intersection local times for simple random walk in dimension $d > 4$, preprint (2009).

[AC07] A. Asselah and F. Castell, Random walk in random scenery and self-intersection local times in dimensions $d \geq 5$. *Probab. Theory Relat. Fields* **138**, 1-32 (2007).

[B07] D. Braess, *Finite Elements. Theory, Fast Solvers and Applications in Elasticity Theory* (Finite Elemente. Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie), 4th revised and extended ed. (German), Springer, Berlin (2007).

[BC04] R. Bass and X. Chen, Self-intersection local time: Critical exponent, large deviations, and laws of the iterated logarithm. *Ann. Probab.* **32**, 3221-3247 (2004).

[BK09] M. Becker and W. König, Moments and distribution of the local times of a transient random walk on $\mathbb{Z}^d$, *Jour. Theor. Prob.* **22:2**, 365 - 374 (2009).

[BHK07] D. Brydges, R. van der Hofstad and W. König, Joint density for the local times of continuous-time Markov chains, *Ann. Probab.* **35:4**, 1307-1332 (2007).

[Ca10] F. Castell, Large deviations for intersection local time in critical dimension, *Ann. Probab.* **38:2**, 927-953 (2010).

[Ch04] X. Chen. Exponential asymptotics and law of the iterated logarithm for intersection local times of random walks. *Ann. Probab.* **32**, 3248-3300 (2004).

[Ch09] X. Chen, *Random Walk Intersections: Large Deviations and Related Topics.* Mathematical Surveys and Monographs, AMS. (2010) Vol. 157, Providence, RI.

[Ce07] J. Cerny, Moments and distribution of the local times of a two-dimensional random walk, *Stoch. Proc. Appl.* **117**, 262-270 (2007).

[DZ98] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd Edition. Springer, New York (1998).

[DV79] M. Donsker and S.R.S. Varadhan, On the number of distinct sites visited by a random walk, *Comm. Pure Appl. Math.* **32**, 721–747 (1979).

[GKS07] N. Gantert, W. König and Z. Shi, Annealed deviations for random walk in random scenery, *Ann. Inst. Henri Poincaré (B) Prob. Stat.* **43:1**, 47-76 (2007).

[GM98] J. Gärtner and S. Molchanov, Parabolic problems for the Anderson model. II. Second-order asymptotics and structure of high peaks. *Probab. Theory Relat. Fields* **111**, 17–55 (1998).

[HKM06] R. van der Hofstad, W. König and P. Mörters, The universality classes in the parabolic Anderson model, *Commun. Math. Phys.* **267:2**, 307-353 (2006).

[Le86] J.-F. Le Gall, Propriétés d'intersection des marches aléatoires, I. Convergence vers le temps local d'intersection, *Com. Math. Phys.* **104**, 471-507 (1986).

[MS81] K. McLeod and J. Serrin, Uniqueness of solutions of semilinear Poisson equations, *Proc. Natl. Acad. Sci. USA*, **78:11**, 6592–6595 (1981).

[We83] M.I. Weinstein, Nonlinear Schrödinger equations and sharp interpolation estimates, *Commun. Math. Phys.* **87**, 567–576 (1983).