

## ON THE NUMBER OF TURNS IN REDUCED RANDOM LATTICE PATHS

YUNJIANG JIANG,\* *Stanford University*

WEIJUN XU,\*\* *University of Oxford*

### Abstract

We consider the tree-reduced path of a symmetric random walk on  $\mathbb{Z}^d$ . It is interesting to ask about the number of turns  $T_n$  in the reduced path after  $n$  steps. This question arises from inverting the signatures of lattice paths:  $T_n$  gives an upper bound of the number of terms in the signature needed to reconstruct a 'random' lattice path with  $n$  steps. We show that, when  $n$  is large, the mean and variance of  $T_n$  in the asymptotic expansion have the same order as  $n$ , while the lower-order terms are  $O(1)$ . We also obtain limit theorems for  $T_n$ , including the large deviations principle, central limit theorem, and invariance principle. Similar techniques apply to other finite patterns in a lattice path.

*Keywords:* Signature of a path; reduced word; number of turns

2010 Mathematics Subject Classification: Primary 60F10; 60F17; 60G50

### 1. Introduction

Let  $G$  be the free group with  $d$  generators  $e_1, \dots, e_d$ . Start with the empty word at time 0. At each time  $k$ , choose one from the  $2d$  elements ( $d$  generators and their inverses) uniformly randomly to multiply the current word on the right. For example, the first six choices  $e_2, e_3, e_3^{-1}, e_2, e_1^{-1}, e_4$  will produce the reduced word  $e_2e_2e_1^{-1}e_4$  at time 6. Every word at time  $n$  has a unique reduced word with length at most  $n$ . It is then interesting to ask about the length and number of turns in the reduced word.

**Definition 1.1.** Let  $w$  be a word at time  $n$ , and let  $\hat{w} = a_{i_1} \cdots a_{i_k}$  be its reduced word, where  $a_j$  is either  $e_j$  or  $e_j^{-1}$ . Define the number of turns of  $w$  to be

$$T(w) = \#\{a_{i_j}a_{i_{j+1}} : i_j \neq i_{j+1}, 1 \leq j \leq k-1\}.$$

There is a canonical bijection between words of length  $n$  and random walks in  $\mathbb{Z}^d$  with  $n$  steps. The quantity  $T(w)$  counts the number of times the reduced trajectory of the walk corresponding to  $w$  has changed its direction. In the above example, the number of turns in the reduced word  $e_2e_2e_1^{-1}e_4$  is 2. If two words reduce to the same word then they necessarily have the same number of turns.

The main goal of this paper is to calculate asymptotics for  $T_n$  when  $n$  is large. The question of estimating  $T_n$  arises from inverting the signature for lattice paths, where at most  $T+2$  terms in the signature are needed for inversion if one knows in advance that the reduced lattice path has  $T$  turns.

---

Received 18 September 2011; revision received 10 September 2012.

\* Postal address: Department of Mathematics, Stanford University, Building 380, Stanford, CA 94305, USA.

Email address: [jjj@math.stanford.edu](mailto:jjj@math.stanford.edu)

\*\* Postal address: Mathematical and Oxford-Man Institutes, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK. Email address: [xu@maths.ox.ac.uk](mailto:xu@maths.ox.ac.uk)

**1.1. Motivation from inversion of signature for axis paths**

Besides the purely probabilistic interest of  $T_n$ , one motivation for our study is from inversion of the signatures for lattice paths, and our aim is to provide an upper bound of the number of terms needed in the signature to reconstruct a random lattice path with  $n$  steps for large  $n$ . We will give some background material on the path signature in this subsection. A path  $\gamma : [s, t] \rightarrow \mathbb{R}^d$  is a continuous function mapping a time interval into  $\mathbb{R}^d$ . The length of the path is defined as

$$|\gamma| := \sup_{\mathcal{P}} \sum_i |\gamma(u_{i+1}) - \gamma(u_i)|,$$

where the supremum is taken over all finite partitions of  $[s, t]$ . If  $|\gamma| < +\infty$ , we say that  $\gamma$  has bounded variation. Let  $BV(\mathbb{R}^d)$  denote the space of all paths of bounded variations in  $\mathbb{R}^d$ .

**Definition 1.2.** Let  $\gamma : [s, t] \rightarrow \mathbb{R}^d$  be an element in  $BV(\mathbb{R}^d)$ . The signature of  $\gamma$ ,  $X_{s,t}(\gamma)$ , is defined as

$$X_{s,t}(\gamma) = 1 + X_{s,t}^1(\gamma) + \dots + X_{s,t}^n(\gamma) + \dots,$$

where

$$X_{s,t}^n(\gamma) = \int_{s < u_1 < \dots < u_n < t} d\gamma(u_1) \otimes \dots \otimes d\gamma(u_n)$$

is an element in  $(\mathbb{R}^d)^{\otimes n}$ .

Let  $(e_1, e_2, \dots, e_d)$  be a standard basis of  $\mathbb{R}^d$ . Then  $\gamma$  can be written in terms of its coordinate decomposition as

$$\gamma = (\gamma_1, \dots, \gamma_d).$$

If  $w = e_{i_1} \dots e_{i_n}$  is a word of length  $n$ , we write

$$C_{s,t}(w) = C_{s,t}(e_{i_1} \dots e_{i_n}) = \int_{s < u_1 < \dots < u_n < t} d\gamma_{i_1}(u_1) \dots d\gamma_{i_n}(u_n)$$

as the coefficient of  $w$ . As all words of length  $n$  form a basis of  $V^{\otimes n}$ , we can rewrite  $X_{s,t}^n(\gamma)$  as the linear combination of basis elements, i.e.

$$X_{s,t}^n(\gamma) = \sum_{|w|=n} C_{s,t}(w)w,$$

where the sum is taken over all words of length  $n$ .

Re-parametrizing the path does not change the signature. For any paths  $\alpha : [0, s] \rightarrow \mathbb{R}^d$  and  $\beta : [0, t] \rightarrow \mathbb{R}^d$ , we can form the concatenation  $\alpha * \beta : [0, s + t] \rightarrow \mathbb{R}^d$  as

$$\alpha * \beta(u) := \begin{cases} \alpha(u), & u \in [0, s], \\ \beta(u - s) + \alpha(s) - \alpha(0), & u \in [s, s + t]; \end{cases}$$

the decomposition of one path into two can be carried out in the same fashion.

For any path  $\gamma : [s, t] \rightarrow \mathbb{R}^d$ , the path ‘ $\gamma$  run backwards’,  $\gamma^{-1}$ , is defined as

$$\gamma^{-1}(u) := \gamma(s + t - u), \quad u \in [s, t],$$

and the trajectories of  $\gamma * \gamma^{-1}$  cancel each other out.

Concatenation and ‘backwards’ of paths of bounded variation are still paths of bounded variation. In fact, we have  $|\alpha * \beta| = |\alpha| + |\beta|$ , and  $|\gamma^{-1}| = |\gamma|$ . The following proposition, first proved by Chen [1], asserts that the signature map is a homomorphism from  $BV(\mathbb{R}^d)$  to the tensor algebra.

**Proposition 1.1.** *Let  $\alpha, \beta \in \text{BV}(\mathbb{R}^d)$ . Then  $X(\alpha * \beta) = X(\alpha) \otimes X(\beta)$ .*

The signature of a path is an important object to study as it determines paths of bounded variation up to tree-like equivalence. More precisely, Hambly and Lyons [4] showed that if  $\alpha, \beta \in \text{BV}(\mathbb{R}^d)$  then  $X(\alpha) = X(\beta)$  if and only if  $\alpha * \beta^{-1}$  is tree like, which is a continuous analogue of a null path. This tree-like relation defines an equivalence relation on  $\text{BV}(\mathbb{R}^d)$ . Within every equivalent class, there is a unique path with minimal length, called the tree-reduced path. A natural question then is how one can reconstruct the tree-reduced path from the given signature. This is in general a very hard problem, but, for the case of axis paths, the answer was provided by Lyons and Xu [6].

**Definition 1.3.**  $\gamma: [s, t] \rightarrow \mathbb{R}^d$  is a (finite) axis path if its movements are parallel to the Euclidean coordinate axes, has finitely many turns, and each straight line component has finite length.

Any axis path has a unique reduced axis path; integer lattice paths are special cases of axis paths. An  $\mathbb{R}^d$  axis path can move in  $d$  different directions (up to the sign). At time 0, it starts to move along a direction  $e_{i_1}$  for some distance  $r_1$ ; then it turns a right angle, and moves along  $e_{i_2}$  for a distance  $r_2$ , and so on, and stops after finitely many turns. Thus, up to re-parametrization, an axis path  $\gamma$  can be represented as

$$\gamma = (r_1 e_{i_1}) * \dots * (r_n e_{i_n}),$$

where the  $r_i$ s are real numbers, with the sign denoting the direction (we mean  $-r e_j = r e_j^{-1}$ ).

Using Chen’s identity (Proposition 1.1), the signature of  $\gamma$  can be conveniently expressed as

$$X(\gamma) = \exp(r_1 e_{i_1}) \otimes \dots \otimes \exp(r_n e_{i_n}),$$

which should be understood as the product of  $n$  power series in the letters  $\{e_{i_1}, \dots, e_{i_n}\}$ .

If  $\gamma$  is already in its reduced form then it is clear that  $i_k \neq i_{k+1}$ , and we call the word  $w = (e_{i_1}, \dots, e_{i_n})$  the shape of  $\gamma$ . If a word  $w$  is in its reduced form, we use  $|w|$  to denote the number of letters in  $w$ , or the length of  $w$ . We introduce the notion of square-free words to characterize an axis path.

**Definition 1.4.** Let  $w = e_{i_1} \dots e_{i_n}$  be a word. We call it a square-free word if, for all  $k \leq n - 1$ ,  $i_k \neq i_{k+1}$ .

In other words, a word is square free if no two consecutive letters in it are identical. The coefficients of square-free words in the signature give a precise description of the moving directions of the path. The following theorem, provided by Lyons and Xu [6], gives an inversion procedure for finite axis paths with the aid of square-free words.

**Theorem 1.1.** *For any finite axis path  $\gamma: [s, t] \rightarrow \mathbb{R}^d$ , there exists a unique square-free word  $w$  with the property that  $C_{s,t}(w) \neq 0$ , and if  $w'$  is any other square-free word with  $C_{s,t}(w') \neq 0$  then  $|w'| < |w|$ . Furthermore, suppose that the unique longest square-free word is  $w = e_{i_1} \dots e_{i_n}$ , and let*

$$w_k := e_{i_1} \dots e_{i_{k-1}} e_{i_k}^2 e_{i_{k+1}} \dots e_{i_n},$$

which has length  $n + 1$ . Then we have

$$\gamma = (r_1 e_{i_1}) * \dots * (r_n e_{i_n}),$$

where  $r_k = 2C_{s,t}(w_k)/C_{s,t}(w)$ .

Thus, if an axis path has  $n$  turns then at most  $n + 2$  terms in the signature are needed for inversion. For a lattice path with length  $L$ , it can have at most  $L - 1$  turns, so we only need the first  $L + 1$  terms in the signature to recover it.

In practice, lattice paths in  $\mathbb{R}^d$  are often generated by drawing  $n$  letters and their inverses uniformly randomly from an alphabet, and putting them in the order they are drawn. It is then important to understand the behavior of the number of turns in the reduced word after a large number of steps, as this gives an asymptotic upper bound for the number of terms needed in the signature for inverting a random lattice path, and also the efficiency of the inversion algorithm.

**1.2. Outline of the method and summary of results**

Let  $T_n$  be the number of turns of the reduced trajectory of a simple symmetric random walk on  $\mathbb{Z}^d$ , as defined in Definition 1.1. Then, we can write

$$T_n = \sum_{i=1}^n V_i,$$

where  $V_i$  denotes the number of turns created at step  $i$ . Note that the  $V_i$  can be 1, 0, or  $-1$ , and are correlated. The distribution of  $V_i$  depends on the whole history in the past.

On the other hand, one can condition on the length of the reduced path  $L_n$ . Then  $T_n|L_n$  has a binomial distribution. A detailed study of  $L_n$  yields asymptotic behaviors of  $T_n$ . A natural coupling  $L_n = S_n + D_n$  simplifies the study of  $L_n$ , where  $S_n$  is a sum of  $n$  independent and identically distributed (i.i.d.) random variables and  $D_n$  is dominated by a geometric random variable.

The main results in this paper are as follows.

(i) *Proposition 2.5.*

$$\mathbb{E}T_n - \frac{2(d-1)^2}{d(2d-1)}n \rightarrow -\frac{2d^2-4d+1}{d(2d-1)}$$

and

$$\text{var}T_n - \frac{2(d-1)^2(5d-2)}{d^2(2d-1)^2}n$$

also converges.

(ii) *Theorem 3.1—large deviations principle.* The sequence of the laws for the random variables  $\{T_n/n\}_{n \geq 1}$  satisfies the large deviations principle with rate function

$$I(x) = \sup_{\theta} [\theta x - \log h(\theta)],$$

where

$$h(\theta) = \begin{cases} \frac{1}{2d} \left[ 2(d-1)e^\theta + \frac{2d-1}{1+2(d-1)e^\theta} + 1 \right], & \theta \geq \log \frac{\sqrt{2d-1}-1}{2(d-1)}, \\ \frac{\sqrt{2d-1}}{d}, & \theta < \log \frac{\sqrt{2d-1}-1}{2(d-1)}. \end{cases}$$

(iii) *Theorem 4.4—invariance principle.* For each  $n$ , define a  $C^0([0, 1])$ -valued random variable  $\{W_t^{(n)} : t \in [0, 1]\}$  by

$$W_t^{(n)} = \frac{1}{\sigma\sqrt{n}} \left[ T_{tn} - \frac{2(d-1)^2}{d(2d-1)}tn \right] \quad \text{for } t \in [n]/n,$$

and linearly interpolated for other values of  $t$ . Then the sequence converges in law to the standard one-dimensional Brownian motion on  $[0, 1]$  as  $n \rightarrow \infty$ .

As a generalization, analogous results hold for the number of occurrences of any finite collection of finite length pattern  $\mathcal{P} = \{P_i = (e_{i_1}, \dots, e_{i_{k_i}})\}$  in a lattice path: the key is to establish a central limit theorem for the number of occurrences of elements in  $\mathcal{P}$  conditioned on the length of the path  $L_t$ , which is essentially an i.i.d. sum of  $m$ -dependent random variables (see [7]), where  $m$  is bounded above by  $\max_i k_i$ . The number of turns  $T_t$  corresponds to  $\mathcal{P} = \{P_{ij} = (e_i, e_j) : i \neq j\}$ . For the sake of clarity, we will focus only on the number of turns.

The paper is organized as follows. In Section 2 we show that the lower-order terms in the mean and variance of  $L_n$  are  $O(1)$ ; we then obtain similar results for  $T_n$ . A key ingredient in the derivation is to prove that  $\text{cov}(S_n, D_n)$  is  $O(1)$ . We compare it with  $\text{cov}(S_{n+1}, D_{n+1})$ , and show that their difference decays exponentially with  $n$ , thus proving convergence.

Section 3 is devoted to the proof of the large deviations principle for  $\{T_n/n\}$ . We derive the rate function, and, thus, prove the principle, by comparison of the Laplace transform of  $L_n$  with that of  $S_n$ . It turns out that the rate function deviates from the normal one as predicted by  $S_n$  on the lower side of the real line.

In Section 4 we prove the central limit theorem and invariance principle for  $T_n$ . This result shows that, although the components of  $T_n$  are correlated, the increments are still asymptotically independent under proper scaling.

### 2. Lower-order terms in the mean and variance

Let  $L_n$  denote the length of the reduced path after  $n$  steps. Then,  $T_n|L_n$  has a binomial distribution with parameters  $(L_n - 1, (2d - 2)/(2d - 1))$ . Let  $X_i$  be a sequence of i.i.d. random variables with  $\mathbb{P}(X_i = 1) = (2d - 1)/2d$  and  $\mathbb{P}(X_i = -1) = 1/2d$ . Let  $L_0 = 0$ . Then  $L_n$  can be defined inductively as

$$L_{i+1} = L_i + X_{i+1}, \quad L_i > 0; \quad L_{i+1} = 1, \quad L_i = 0.$$

We want to compare  $L_n - S_n$ , where  $S_n = \sum_{i=1}^n X_i$ . It is well known that  $L_n/n \rightarrow (d - 1)/d$  almost surely. We compute a finer estimate to show that  $\mathbb{E}L_n - n(d - 1)/d = O(1)$ .

Since  $L_i - S_i$  does not change when  $L$  is away from 0, a difference only occurs when  $L$  hits 0. Let  $R_n$  denote the number of times that  $L_n$  hits 0 after the first step up to time  $n$ . Since  $\mathbb{P}(L \text{ ever comes back to } 0) = 1/(2d - 1)$ ,  $R_n$  converges to a geometric distributed random variable  $R$ , with

$$\mathbb{P}(R = k) = \frac{2d - 2}{2d - 1} \left( \frac{1}{2d - 1} \right)^{k-1}$$

(see, e.g. [3]). Here, step 0 is counted as a return, because  $L$  and  $S$  can be different in the first move. Let  $D_n = L_n - S_n$ . Then  $\frac{1}{2}D_n|R_{n-1}$  has a binomial distribution with parameters  $(R_{n-1}, 1/2d)$ . In particular,  $D_n \leq 2R_{n-1}$ . So the mean of the difference is

$$\mathbb{E}D_n = \mathbb{E}\mathbb{E}(D_n | R_{n-1}) = \frac{1}{d}\mathbb{E}R_{n-1} \rightarrow \frac{2d - 1}{2d(d - 1)}.$$

Thus, we get an error term for  $\mathbb{E}L_n$ .

**Lemma 2.1.** *Let  $L_n$  denote the length of the reduced word after  $n$  steps. Then*

$$\lim_{n \rightarrow \infty} \left( \mathbb{E}L_n - \frac{d-1}{d}n \right) = \frac{2d-1}{2d(d-1)}.$$

Now we compute the lower-order terms in  $\text{var}L_n$ . Since  $\text{var}L_n = \text{var}S_n + 2\text{cov}(S_n, D_n) + \text{var}D_n$ , it suffices to show that  $\text{cov}(S_n, D_n) = O(1)$ . We show it in the following lemma.

**Lemma 2.2.** *Under the above coupling  $L_n = S_n + D_n$ , we have*

$$\lim_{n \rightarrow \infty} (\mathbb{E}S_n D_n - \mathbb{E}S_n \mathbb{E}D_n) = -u(d),$$

where  $0 \leq u(d) \leq 2d^2(2d-1)(d^2+2d-1)/(d-1)^5$ .

*Proof.* Let  $U_n = (S_n - n(d-1)/d)D_n$ . We will show that  $\mathbb{E}U_{n+1} - \mathbb{E}U_n$  decays exponentially fast. Note that we have

$$\begin{aligned} U_{n+1} &= \mathbf{1}_{\{L_n=0\}} \left( S_n + X_{n+1} - \frac{d-1}{d}(n+1) \right) (D_n + 1 - X_{n+1}) \\ &\quad + \mathbf{1}_{\{L_n>0\}} \left( S_n + X_{n+1} - \frac{d-1}{d}(n+1) \right) D_n \\ &= U_n + \left( X_{n+1} - \frac{d-1}{d} \right) D_n + \mathbf{1}_{\{L_n=0\}} (1 - X_{n+1}) \left( S_{n+1} - \frac{d-1}{d}(n+1) \right). \end{aligned}$$

Since  $X_{n+1}$  is independent of  $D_n$ , and  $-n \leq S_n \leq 0$  when conditioned on  $L_n = 0$ , taking the expectation of both sides yields

$$-4(n+1)\mathbb{P}(L_n = 0) \leq \mathbb{E}U_{n+1} - \mathbb{E}U_n \leq 0. \tag{2.1}$$

Also,  $\mathbb{P}(L_{2n+1} = 0) = 0$ , and

$$\mathbb{P}(L_{2n} = 0) \leq \mathbb{P}(S_{2n} \leq 0) \leq \sum_{k=0}^n 2^{2k} \left( \frac{1}{2d} \right)^{n+k} \left( \frac{2d-1}{2d} \right)^{n-k} \leq \frac{2d-1}{2(d-1)} \left( \frac{2d-1}{d^2} \right)^n.$$

We can then deduce from (2.1) that  $\mathbb{E}U_n$  is decreasing and bounded from below, and, thus, that it has a finite limit. Adding up all  $(\mathbb{E}U_n - \mathbb{E}U_{n-1})$  gives  $\mathbb{E}U_n \rightarrow -u(d)$ , where  $0 \leq u(d) \leq 2d^2(2d-1)(d^2+2d-1)/(d-1)^5$ .

**Remark 2.1.** The negative correlation agrees with probabilistic intuition: when  $S_n$  is small, the process  $L_n$  tends to visit 0 more times, and, thus,  $D_n$  is likely to be large.

**Proposition 2.1.** *Let  $u(d)$  be a constant as in the previous lemma. Then*

$$\lim_{n \rightarrow \infty} \left( \text{var}L_n - \frac{2d-1}{d^2}n \right) = \beta(d),$$

where  $\beta(d) = -2u(d) + (2d-1)(4d^2-6d+3)/4d^2(d-1)^2$ .

*Proof.* Since

$$\text{var}D_n = \mathbb{E}\text{var}(D_n \mid R_{n-1}) + \text{var}\mathbb{E}(D_n \mid R_{n-1}) \rightarrow \frac{(2d-1)(4d^2-6d+3)}{4d^2(d-1)^2},$$

we have

$$\text{var}L_n - \frac{2d - 1}{d^2}n = 2\text{cov}(S_n, D_n) + \text{var}D_n \rightarrow \beta(d),$$

where  $\beta(d) = -2u(d) + (2d - 1)(4d^2 - 6d + 3)/4d^2(d - 1)^2$ . This completes the proof.

Combining the above estimates for  $L_n$ , we then have similar estimates for  $T_n$ :

$$\begin{aligned} \mathbb{E}T_n &= \mathbb{E}\mathbb{E}(T_n \mid L_n) = \frac{2d - 2}{2d - 1}\mathbb{E}L_n - \frac{2d - 2}{2d - 1}, \\ \text{var}T_n &= \mathbb{E}\text{var}(T_n \mid L_n) + \text{var}\mathbb{E}(T_n \mid L_n) \\ &= \frac{2(d - 1)}{(2d - 1)^2}\mathbb{E}L_n + \frac{4(d - 1)^2}{(2d - 1)^2}\text{var}L_n - \frac{2(d - 1)}{(2d - 1)^2}. \end{aligned}$$

This gives the following proposition.

**Proposition 2.2.** *Let  $\beta(d)$  be the error term in  $\text{var}L_n$  as above. Then*

$$\lim_{n \rightarrow \infty} \left( \mathbb{E}T_n - \frac{2(d - 1)^2}{d(2d - 1)}n \right) = -\frac{2d^2 - 4d + 1}{d(2d - 1)}$$

and

$$\lim_{n \rightarrow \infty} \left( \text{var}T_n - \frac{2(d - 1)^2(5d - 2)}{d^2(2d - 1)^2}n \right) = \frac{4(d - 1)^2}{(2d - 1)^2}\beta(d) - \frac{2d^2 - 4d + 1}{d(2d - 1)^2}.$$

### 3. Large deviations

The goal of this section is to prove the following large deviations theorem for  $T_n/n$ .

**Theorem 3.1.** *The sequence of the laws for the random variables  $\{T_n/n\}_{n \geq 1}$  satisfies the large deviations principle with rate function*

$$I(x) = \sup_{\theta} [\theta x - \log h(\theta)],$$

where

$$h(\theta) = \begin{cases} \frac{1}{2d} \left[ 2(d - 1)e^{\theta} + \frac{2d - 1}{1 + 2(d - 1)e^{\theta}} + 1 \right], & \theta \geq \log \frac{\sqrt{2d - 1} - 1}{2(d - 1)}, \\ \frac{\sqrt{2d - 1}}{d}, & \theta < \log \frac{\sqrt{2d - 1} - 1}{2(d - 1)}. \end{cases}$$

We postpone the proof of this theorem to the end of the section. In light of the Gartner-Ellis theorem (see [2, Section 2.3]), it suffices to show that

$$\lim_{n \rightarrow +\infty} (\mathbb{E}e^{\theta T_n})^{1/n} = h(\theta)$$

for every  $\theta \in \mathbb{R}$ , with the limit  $h$  essentially smooth.

Note that  $T_n|L_n$  has a binomial distribution with parameter  $((2d - 2)/(2d - 1), L_n - 1)$  for  $L_n \geq 1$ , and  $T_n = 0$  if  $L_n = 0$ , so we have

$$\mathbb{E}e^{\theta T_n} = \mathbb{P}(L_n = 0) + \mathbb{P}(L_n > 0)\mathbb{E}w(\theta)^{L_n - 1},$$

where

$$w(\theta) = \frac{2d - 2}{2d - 1} e^\theta + \frac{1}{2d - 1}.$$

The first term is bounded by

$$\mathbb{P}(L_n = 0) \leq \sum_{k=0}^n \mathbb{P}(S_n = -k) \leq C \left( \frac{\sqrt{2d - 1}}{d} \right)^n,$$

and we need to estimate  $\mathbb{E}w(\theta)^{L_n}$  for large  $n$ .

In the context below, we regard  $w$  to be a positive real number independent of  $\theta$ , and study the asymptotics of  $(\mathbb{E}w^{L_n})^{1/n}$  as  $n \rightarrow +\infty$ .

**Proposition 3.1.** *If  $w \geq 1$  then we have*

$$\lim_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} = \frac{2d - 1}{2d} w + \frac{1}{2d} \frac{1}{w}.$$

*Proof.* We compare the difference between  $\mathbb{E}w^{L_{n+1}}$  and  $\mathbb{E}w^{L_n}$ :

$$\begin{aligned} \mathbb{E}w^{L_{n+1}} &= \mathbb{E} \mathbf{1}_{\{L_n=0\}} w^{L_{n+1}} + \mathbb{E} \mathbf{1}_{\{L_n>0\}} w^{L_{n+1}} \\ &= w \mathbb{P}(L_n = 0) + \mathbb{E} \mathbf{1}_{\{L_n>0\}} w^{L_n + X_{n+1}} \\ &= w \mathbb{P}(L_n = 0) + \mathbb{E} w^{X_{n+1}} \mathbb{E} \mathbf{1}_{\{L_n>0\}} w^{L_n} \\ &= w \mathbb{P}(L_n = 0) + \mathbb{E} w^{X_{n+1}} \mathbb{E} w^{L_n} - \mathbb{E} w^{X_{n+1}} \mathbb{P}(L_n = 0) \\ &= \left( \frac{2d - 1}{2d} w + \frac{1}{2d} \frac{1}{w} \right) \mathbb{E} w^{L_n} + \frac{1}{2d} \left( w - \frac{1}{w} \right) \mathbb{P}(L_n = 0). \end{aligned}$$

Let

$$x_n = \mathbb{E}w^{L_n}, \quad a = \frac{2d - 1}{2d} w + \frac{1}{2d} \frac{1}{w}, \quad b = \frac{1}{2d} \left( w - \frac{1}{w} \right), \quad p_n = \mathbb{P}(L_n = 0).$$

Then we have the following recursive relation:

$$x_n = ax_{n-1} + bp_{n-1}.$$

Since  $x_1 = w$ , adding them up yields

$$x_n = a^{n-1}w + b(a^{n-2}p_1 + a^{n-3}p_2 + \dots + ap_{n-2} + p_{n-1}).$$

For  $w \geq 1$ , we have  $b \geq 0$ . In this case, since  $a^{n-1} \leq x_n \leq na^{n-1}$  from the expression above, we obtain

$$\lim_{n \rightarrow \infty} (x_n)^{1/n} = a = \frac{2d - 1}{2d} w + \frac{1}{2d} \frac{1}{w},$$

thus proving the proposition.

The situation for  $w \in (0, 1)$  is more involved. We prove it based on comparison with  $(\mathbb{E}w^{S_n})^{1/n}$ . Note that  $S_n$  is a sum of i.i.d. random variables, so, by Cramer’s theorem, it satisfies the large deviations principle with rate function

$$J(x) = \sup_{\theta} \left[ \theta x - \log \left( \frac{2d - 1}{2d} e^\theta + \frac{1}{2d} e^{-\theta} \right) \right].$$

**Lemma 3.1.** For any  $w > 0$ , the equation

$$w^\alpha e^{-J(\alpha)} = \frac{2d-1}{2d}w + \frac{1}{2d} \frac{1}{w} \tag{3.1}$$

has a unique solution at  $\alpha^* = ((2d-1)w^2 - 1)/((2d-1)w^2 + 1)$ . Furthermore,  $\alpha^*$  is the global maximizer for

$$f_w(\alpha) = w^\alpha e^{-J(\alpha)}.$$

*Proof.* We first obtain an expression for  $J$  in terms of  $\alpha$  only. It is clear that  $J(\alpha) = +\infty$  for  $|\alpha| > 1$ . For  $\alpha \in [-1, 1]$ , the maximizer  $\theta^*$  is

$$\theta^*(\alpha) = \frac{1}{2} \left[ \log \frac{1+\alpha}{1-\alpha} - \log(2d-1) \right], \quad \alpha \in [-1, 1],$$

passing to the limit  $\pm\infty$  for  $\alpha = \pm 1$ . Substituting into  $J$ , we have

$$J(\alpha) = \frac{1}{2}\alpha \left[ \log \frac{1+\alpha}{1-\alpha} - \log(2d-1) \right] - \log \frac{\sqrt{2d-1}}{2d} - \log \left( \sqrt{\frac{1+\alpha}{1-\alpha}} + \sqrt{\frac{1-\alpha}{1+\alpha}} \right)$$

for  $\alpha \in [-1, 1]$ . Differentiating with respect to  $\alpha$ , we obtain

$$J'(\alpha) = \frac{1}{2} \left[ \log \frac{1+\alpha}{1-\alpha} - \log(2d-1) \right] \tag{3.2}$$

for  $\alpha \in (-1, 1)$ . Note that

$$f'_w(\alpha) = \frac{d}{d\alpha} (w^\alpha e^{-J(\alpha)}) = w^\alpha e^{-J(\alpha)} (\log w - J'(\alpha)),$$

and, since  $J$  is convex,  $f_w$  has the global maximizer  $\alpha^*$  satisfying

$$J'(\alpha^*) = \log w.$$

By (3.2), solving the above first-order condition yields

$$\alpha^* = \frac{(2d-1)w^2 - 1}{(2d-1)w^2 + 1},$$

and, thus,

$$f_w(\alpha^*) = \frac{2d-1}{2d}w + \frac{1}{2d} \frac{1}{w}.$$

Since  $\alpha^*$  is the global maximizer of  $f_w$ , we conclude that (3.1) has a unique solution at  $\alpha^*$ .

**Proposition 3.2.** Let  $\alpha^* = \alpha^*(w) = ((2d-1)w^2 - 1)/((2d-1)w^2 + 1)$  as above. Then

$$\lim_{\varepsilon \downarrow 0} \lim_{n \rightarrow +\infty} (\mathbb{E}w^{S_n} \mathbf{1}_{\{S_n/n \in (\alpha^* - \varepsilon, \alpha^* + \varepsilon)\}})^{1/n} = \frac{2d-1}{2d}w + \frac{1}{2d} \frac{1}{w}.$$

This proposition shows that the major contribution to  $\mathbb{E}w^{S_n}$  is from the  $S_n$ s with values near  $\alpha^*n$ .

**Lemma 3.2.** (a) Let  $\alpha > 0$ . For all  $\varepsilon \in (0, \alpha)$  and all small enough  $\delta > 0$ , there exists  $N = N(\alpha, \varepsilon, \delta)$  such that

$$\mathbb{P} \left( L_n - S_n \leq \delta n \mid \frac{S_n}{n} \in (\alpha - \varepsilon, \alpha + \varepsilon) \right) \geq 1 - n \left( \frac{1 - \alpha + \varepsilon}{1 + \alpha - \varepsilon} \right)^{\delta n/4}$$

for all  $n \geq N$ .

(b) For all  $\varepsilon > 0$  and all small enough  $\delta > 0$ , there exists  $N = N(\varepsilon, \delta)$  such that

$$\mathbb{P}(L_n \leq \delta n \mid S_n \leq -\varepsilon n) \geq 1 - n \left( \frac{1 - \varepsilon}{1 + \varepsilon} \right)^{\delta n/4}$$

for all  $n \geq N$ .

*Proof.* We first prove part (a). Observe that  $D_n = L_n - S_n = 2|\min_{0 \leq k \leq n} S_k|$ . We first consider the quantity  $\mathbb{P}(\min_{0 \leq k \leq n} S_k \geq -\delta n \mid S_n = \alpha n)$ , where, without loss of generality, we have assumed that  $(1 + \alpha)n/2$  is an integer, and have replaced  $\delta/2$  by  $\delta$ . Once conditioned on the event  $\{S_n = \alpha n\}$ , all possible paths contain  $\binom{n}{(1+\alpha)n/2}$  positive movements and  $\binom{n}{(1-\alpha)n/2}$  negative movements. Since all these paths have the same (conditional) weight, the quantity  $\mathbb{P}(\min_{0 \leq k \leq n} S_k \geq -\delta n \mid S_n = \alpha n)$  is independent of  $d$ . Thus, we may assume that  $d = 1$ , where all paths are the trajectories of the (conditional) simple symmetric random walk. That is,

$$\mathbb{P}\left(\min_{0 \leq k \leq n} S_k < -\delta n \mid S_n = \alpha n\right) = \mathbb{P}\left(\min_{0 \leq k \leq n} \tilde{S}_k < -\delta n \mid \tilde{S}_n = \alpha n\right),$$

where  $\tilde{S}_n$  is a one-dimensional simple symmetric random walk. By the reflection principle we have

$$\mathbb{P}\left(\min_{0 \leq k \leq n} \tilde{S}_k < -\delta n, \tilde{S}_n = \alpha n\right) = \mathbb{P}(\tilde{S}_n = -\lfloor(\alpha + 2\delta)n\rfloor).$$

Using Stirling’s approximation, we estimate the ratio

$$\begin{aligned} \frac{\mathbb{P}(\tilde{S}_n = -\lfloor(\alpha + 2\delta)n\rfloor)}{\mathbb{P}(\tilde{S}_n = \alpha n)} &= \frac{\binom{n}{(1-\alpha-2\delta)n/2}}{\binom{n}{(1+\alpha)n/2}} \\ &= \frac{\left(\frac{1-\alpha}{2}n\right)! \left(\frac{1+\alpha}{2}n\right)!}{\left[\left(\frac{1-\alpha-2\delta}{2}n\right)! \left(\frac{1+\alpha+2\delta}{2}n\right)!\right]} \\ &\approx \left[ \frac{(1-\alpha)^{1-\alpha} (1+\alpha)^{1+\alpha}}{(1-\alpha-2\delta)^{1-\alpha-2\delta} (1+\alpha+2\delta)^{1+\alpha+2\delta}} \right]^{n/2} \\ &= e^{[g(\alpha)-g(\alpha+2\delta)]n/2}, \end{aligned}$$

where  $g(x) = (1 - x) \log(1 - x) + (1 + x) \log(1 + x)$ . It is straightforward to check that

$$g'(x) = \log \frac{1+x}{1-x} > 0$$

for  $x \in (0, 1)$ , and, thus, for small enough  $\delta > 0$ , we have

$$\begin{aligned} &\mathbb{P}\left(\min_{0 \leq k \leq n} S_k < -\delta n \mid \frac{S_n}{n} \in (\alpha - \varepsilon, \alpha + \varepsilon)\right) \\ &= \sum_{\beta \in (\alpha - \varepsilon, \alpha + \varepsilon)} \mathbb{P}\left(\min_{1 \leq k \leq n} \tilde{S}_k < -\delta n, \frac{\tilde{S}_n}{n} \in (\alpha - \varepsilon, \alpha + \varepsilon)\right) / \mathbb{P}\left(\frac{\tilde{S}_n}{n} \in (\alpha - \varepsilon, \alpha + \varepsilon)\right) \\ &\leq n \sup_{\beta \in (\alpha - \varepsilon, \alpha + \varepsilon)} \left[ \frac{\mathbb{P}(\tilde{S}_n = -(\beta + 2\delta)n)}{\mathbb{P}(\tilde{S}_n = \beta n)} \right] \\ &\leq n \sup_{\beta \in (\alpha - \varepsilon, \alpha + \varepsilon)} e^{[g(\beta)-g(\beta+2\delta)]n/2} \\ &\leq n \left( \frac{1 - \alpha + \varepsilon}{1 + \alpha - \varepsilon} \right)^{\delta n/2}, \end{aligned}$$

and, consequently,

$$\mathbb{P}\left(\min_{0 \leq k \leq n} S_k \geq -\delta n \mid \frac{S_n}{n} \in (\alpha - \varepsilon, \alpha + \varepsilon)\right) \geq 1 - n \left(\frac{1 - \alpha + \varepsilon}{1 + \alpha - \varepsilon}\right)^{\delta n/2}.$$

As  $L_n - S_n = 2|\min_{0 \leq k \leq n} S_k|$ , replacing  $\delta$  by  $\delta/2$  gives (a).

The proof of part (b) is similar, and is thus omitted.

**Proposition 3.3.** *For  $w \in (0, 1)$ , the limit  $(\mathbb{E}w^{L_n})^{1/n}$  exists as  $n \rightarrow +\infty$ , and we have*

$$\lim_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} = \begin{cases} \frac{2d-1}{2d}w + \frac{1}{2d} \frac{1}{w}, & w \in \left(\frac{1}{\sqrt{2d-1}}, 1\right), \\ \frac{\sqrt{2d-1}}{d}, & w \in \left(0, \frac{1}{\sqrt{2d-1}}\right]. \end{cases}$$

*Proof.* We first consider the case in which  $w \in (1/\sqrt{2d-1}, 1)$  and  $\alpha^* = ((2d-1)w^2 - 1)/((2d-1)w^2 + 1) > 0$ . In this case, we have

$$\begin{aligned} (\mathbb{E}w^{L_n})^{1/n} &\geq \left[ \mathbb{E}\left(w^{L_n} \mathbf{1}_{\{L_n - S_n \leq \delta n\}} \mid \frac{S_n}{n} \in (\alpha^* - \varepsilon, \alpha^* + \varepsilon)\right) \right. \\ &\quad \left. \times \mathbb{P}\left(\frac{S_n}{n} \in (\alpha^* - \varepsilon, \alpha^* + \varepsilon)\right) \right]^{1/n} \\ &\geq w^{\alpha^* + \varepsilon + \delta} \mathbb{P}\left(\frac{S_n}{n} \in (\alpha^* - \varepsilon, \alpha^* + \varepsilon)\right)^{1/n} \\ &\quad \times \mathbb{P}\left(L_n - S_n \leq \delta n \mid \frac{S_n}{n} \in (\alpha^* - \varepsilon, \alpha^* + \varepsilon)\right)^{1/n}. \end{aligned}$$

Taking  $n \rightarrow +\infty$  on both sides yields

$$\liminf_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} \geq w^{\alpha^* + \varepsilon + \delta} e^{-J(\alpha^* + \varepsilon)}.$$

Since  $\varepsilon$  and  $\delta$  are arbitrary, we have

$$\liminf_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} \geq w^{\alpha^*} e^{-J(\alpha^*)} = \frac{2d-1}{2d}w + \frac{1}{2d} \frac{1}{w}.$$

On the other hand,  $\mathbb{E}w^{L_n} \leq \mathbb{E}w^{S_n}$  for all  $w < 1$ , so

$$\limsup_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} \leq \frac{2d-1}{2d}w + \frac{1}{2d} \frac{1}{w}.$$

Thus, we conclude that

$$\lim_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} = \frac{2d-1}{2d}w + \frac{1}{2d} \frac{1}{w}$$

for  $w \in (\sqrt{2d-1}/d, 1)$ .

We now consider the case in which  $w \in (0, 1/\sqrt{2d-1})$  and  $\alpha^* < 0$ . In this case, we have

$$\begin{aligned} (\mathbb{E}w^{L_n})^{1/n} &\geq [\mathbb{E}(w^{L_n} \mathbf{1}_{\{L_n \leq \delta n\}} \mid S_n \leq -\varepsilon n) \mathbb{P}(S_n \leq -\varepsilon n)]^{1/n} \\ &\geq w^{\delta} \mathbb{P}(L_n \leq \delta n \mid S_n \leq -\varepsilon n)^{1/n} \mathbb{P}(S_n \leq -\varepsilon n)^{1/n}. \end{aligned}$$

Again, sending  $n \rightarrow +\infty$ , and taking  $\varepsilon$  and  $\delta$  arbitrarily small yields

$$\liminf_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} \geq e^{-I(0)} = \frac{\sqrt{2d-1}}{d}.$$

On the other hand,

$$\lim_{w \downarrow 1/\sqrt{2d-1}} \lim_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} = \frac{\sqrt{2d-1}}{d}.$$

By monotonicity we have

$$\limsup_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} \leq \frac{\sqrt{2d-1}}{d},$$

and, thus,

$$\limsup_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} = \frac{\sqrt{2d-1}}{d}$$

for all  $w \in (0, 1/\sqrt{2d-1})$ .

Finally, we consider the case in which  $w = 1/\sqrt{2d-1}$  and  $\alpha^* = 0$ . Again, by monotonicity in  $w$  we have

$$\lim_{n \rightarrow +\infty} (\mathbb{E}\sqrt{2d-1}^{L_n})^{1/n} = \frac{\sqrt{2d-1}}{d}.$$

This completes the proof.

The following corollary is an immediate consequence of Propositions 3.1 and 3.3.

**Corollary 3.1.** *The laws for the random variables  $\{L_n/n\}_{n \geq 1}$  satisfy the large deviations principle with rate function*

$$I_L(x) = \sup_{\theta} [\theta x - \log h_L(\theta)],$$

where

$$h_L(\theta) = \begin{cases} \frac{2d-1}{2d}e^\theta + \frac{1}{2d}e^{-\theta}, & \theta \geq -\frac{1}{2} \log(2d-1), \\ \frac{\sqrt{2d-1}}{d}, & \theta < \log \frac{\sqrt{2d-1}-1}{2(d-1)}. \end{cases}$$

Now we are in a position to prove Theorem 3.1.

*Proof of Theorem 3.1.* Recall that

$$\mathbb{E}e^{\theta T_n} = \mathbb{P}(L_n = 0) + \mathbb{P}(L_n \geq 1)\mathbb{E}w(\theta)^{L_n-1},$$

where we have

$$\mathbb{P}(L_n = 0) \leq C \left( \frac{\sqrt{2d-1}}{d} \right)^n$$

for all  $n$ . On the other hand, Propositions 3.1 and 3.3 imply that

$$\lim_{n \rightarrow +\infty} (\mathbb{E}w^{L_n})^{1/n} \geq \frac{\sqrt{2d-1}}{d}$$

for all  $w > 0$ . Thus, we see that

$$\lim_{n \rightarrow +\infty} (\mathbb{E}e^{\theta T_n})^{1/n} = \lim_{n \rightarrow +\infty} (\mathbb{E}w(\theta)^{L_n})^{1/n} = h(\theta),$$

where  $h$  is defined in the theorem. It is also straightforward to check that  $h$  is essentially smooth.

Thus, the Gartner-Ellis theorem implies that the laws for  $\{T_n/n\}$  satisfy the large deviations principle with rate function

$$I(x) = \sup_{\theta} [\theta x - \log h(\theta)].$$

This completes the proof.

We end this section with two remarks.

**Remark 3.1.** As an alternative to the Gartner-Ellis theorem, one can compute the rate function for  $T_n/n$  directly, i.e.

$$\begin{aligned} & \mathbb{P}\left(\frac{T_n}{n} \in (\alpha - \varepsilon, \alpha + \varepsilon)\right) \\ &= \sum_{\beta} \mathbb{P}\left(\frac{T_n}{n} \in (\alpha - \varepsilon, \alpha + \varepsilon) \mid \frac{L_n}{n} \in (\beta - \delta, \beta + \delta)\right) \mathbb{P}\left(\frac{L_n}{n} \in (\beta - \delta, \beta + \delta)\right), \end{aligned}$$

where the sum is taken over appropriate  $\beta \in (0, 1)$ . In each product, the first probability is known as  $T_n|L_n$  and has a binomial distribution, while the second probability can be asymptotically computed from the large deviations principle for  $L_n$ . Finally, sending  $\delta \rightarrow 0$ , one can get the rate function for  $\{T_n/n\}$ .

**Remark 3.2.** We give a heuristic explanation why  $\lim_n (\mathbb{E}w^{L_n})^{1/n}$ , if it exists, is a constant for  $w$  below the critical value. Let  $w = e^\eta$ , and consider the Laplace transform

$$(\mathbb{E}w^{L_n})^{1/n} = (\mathbb{E}e^{\eta L_n})^{1/n}.$$

Call the limit of the right-hand side  $\Lambda(\eta)$ , if it exists. For  $w > w^* = 1/\sqrt{2d - 1}$ , the limit exists and is equal to  $(2d - 1)e^\eta/2d + e^{-\eta}/2d$ . Note that  $\eta^*$  is a maximizer for this quantity on  $[\eta^*, +\infty)$ , so we have

$$\Lambda'_+(\eta^*) = 0. \tag{3.3}$$

On the other hand, as a limit of the Laplace transform,  $\Lambda$  must be convex if it exists. We have already seen that

$$\Lambda(\eta) \leq \frac{\sqrt{2d - 1}}{d}$$

for all  $\eta \leq \eta^*$ , so, together with (3.3), the convexity of  $\Lambda$  forces the graph to be flat on  $(-\infty, \eta^*]$ , and is thus equal to  $\sqrt{2d - 1}/d$ .

#### 4. Central limit theorem and the invariance principle

In this section we derive the asymptotic distribution of  $T_n$  for large  $n$ , including the central limit theorem and invariance principle. Under the natural coupling  $L_n = S_n + D_n$ ,  $D_n \leq 2R_n$ , where  $R_n$  is (almost) surely bounded by a geometric random variable  $R$ . As we will be frequently using this property, we state it as a lemma below.

**Lemma 4.1.** *Let  $L_n = S_n + D_n$ , and let  $R_n$  be defined as above. Then,  $D_n \leq wR_n$  and  $R = \sup_n R_n$  has a geometric distribution with*

$$\mathbb{P}(R_n = k) = \frac{2d - 2}{2d - 1} \left(\frac{1}{2d - 1}\right)^{1/n}.$$

**Theorem 4.1.** *Under the above assumptions,  $n^{-1/2}(L_n - (d-1)n/d)$  converges in distribution to  $N(0, (2d-1)/d^2)$ , and  $n^{-1/2}(T_n - 2(d-1)^2n/d(2d-1))$  converges in distribution to  $N(0, 2(d-1)^2(5d-2)/d^2(2d-1)^2)$ .*

*Proof.* For any  $x \in \mathbb{R}$ ,

$$\begin{aligned} F_n(x) &= \mathbb{P}\left(\frac{1}{\sqrt{n}}\left(L_n - \frac{2d-1}{d^2}n\right) \leq x\right) \\ &= \mathbb{P}\left(\frac{1}{\sqrt{n}}\left(S_n - \frac{2d-1}{d^2}n\right) + \frac{D_n}{\sqrt{n}} \leq x\right) \\ &\rightarrow F(x), \end{aligned}$$

where  $F$  is the distribution function for  $N(0, (2d-1)/d^2)$ . The convergence in the last line follows from the fact that  $D_n/\sqrt{n} \rightarrow 0$  almost surely, as  $R = \sup_n R_n$  is almost surely finite.

Now we compute the moment generating function for  $n^{-1/2}(T_n - 2(d-1)^2n/d(2d-1))$ . For simplicity, let  $\mu = 2(d-1)^2/d(2d-1)$  and  $p = (2d-2)/(2d-1)$ . Then

$$\begin{aligned} \mathbb{E}e^{\theta(T_n - \mu n)/\sqrt{n}} &= e^{-\theta\mu\sqrt{n}}\mathbb{E}e^{\theta T_n/\sqrt{n}} \\ &= e^{-\theta\mu\sqrt{n}}\mathbb{E}\mathbb{E}(e^{\theta T_n/\sqrt{n}} \mid L_n) \\ &= e^{-\theta\mu\sqrt{n}}\mathbb{E}(1 - p + pe^{\theta/\sqrt{n}})^{L_n - 1} \\ &= e^{-\theta\mu\sqrt{n}}\mathbb{E}\exp\left((L_n - 1)\log\left(1 + \frac{p\theta}{\sqrt{n}} + \frac{p\theta^2}{2n} + o\left(\frac{1}{n}\right)\right)\right) \\ &= e^{-\theta\mu\sqrt{n}}\mathbb{E}\exp\left((L_n - 1)\left[\frac{p\theta}{\sqrt{n}} + \frac{1}{2n}(p\theta^2 - p^2\theta^2) + o\left(\frac{1}{n}\right)\right]\right) \\ &= \mathbb{E}\exp\left(\frac{p\theta}{\sqrt{n}}\left(L_n - \frac{\mu}{p}n\right) + \frac{1}{2}\theta^2 p(1-p)\frac{L_n}{n} + o(1)\right) \\ &\rightarrow \exp\left(\frac{(d-1)^2(5d-2)}{d^2(2d-1)^2}\theta^2\right). \end{aligned}$$

The convergence in the last line follows from the law of large numbers and central limit theorem for  $L_n$ , and the dominated convergence. This implies that  $n^{-1/2}(T_n - 2(d-1)^2n/d(2d-1))$  converges in distribution to  $N(0, 2(d-1)^2(5d-2)/d^2(2d-1)^2)$ .

Since  $L_n$  behaves very much like  $S_n$ , we expect that, under proper scaling, it converges to the standard Brownian motion.

**Proposition 4.1.** *Let  $(\Omega, \mathbb{P}, \mathcal{F})$  be a probability space affording the discrete-time process  $\{S_i, i \in \mathbb{N}\}$ . For each  $n \in \mathbb{N}$ , define a  $C^0([0, 1])$ -valued random variable  $\{W_t^{(n)} : t \in [0, 1]\}$  by*

$$W_t^{(n)} = \left[ L_{[tn]} - \frac{d-1}{d}tn \right] / \sqrt{n \frac{2d-1}{d^2}}$$

for  $t \in [n]/n$ , and linearly interpolated for other values of  $t$ . Then the sequence converges in law to the standard one-dimensional Brownian motion on  $[0, 1]$  as  $n \rightarrow \infty$ .

*Proof.* This is plain in light of Lemma 4.1, as  $L_{[tn]} - S_{[tn]}$  is bounded by a geometric random variable, uniformly in  $t$ .

More difficult is the invariance principle for  $T_n$ , the number of turns at time  $n$ . We prove it in the next theorem. Let  $\sigma^2 = 2(d - 1)^2(5d - 2)/d^2(2d - 1)^2$ .

**Theorem 4.2.** For each  $n$ , define a  $C^0([0, 1])$ -valued random variable  $\{W_t^{(n)} : t \in [0, 1]\}$  by

$$W_t^{(n)} = \frac{1}{\sigma\sqrt{n}} \left[ T_{tn} - \frac{2(d - 1)^2}{d(2d - 1)} tn \right]$$

for  $t \in [n]/n$ , and linearly interpolated for other values of  $t$ . Then the sequence converges in law to the standard one-dimensional Brownian motion on  $[0, 1]$  as  $n \rightarrow \infty$ .

**4.1. Proof of Theorem 4.4**

Recall that, for a sequence of processes to converge to Brownian motion, it suffices to check that their finite-dimensional joint distributions converge to that of a Brownian motion and Prohorov tightness criterion (see Theorem 16.5 of [5]).

**Lemma 4.2.** Let  $W_t^{(n)}$  be defined as above. Then

$$(W_{t_1}^{(n)}, W_{t_2}^{(n)}, \dots, W_{t_k}^{(n)}) \rightarrow (B_{t_1}, B_{t_2}, \dots, B_{t_k}),$$

where  $B_t$  is a standard one-dimensional Brownian motion starting at 0.

*Proof.* It suffices to show that  $\mathcal{L}(W_t^{(n)} - W_s^{(n)} \mid W_r^{(n)} : r \leq s)$  is asymptotically  $N(0, t - s)$  and independent of  $\{W_r^{(n)} : r \leq s\}$ . This is a generalization of the central limit theorem for  $T_n$ , which corresponds to the case  $s = 0$ , and also suggests that, for general  $s$ , we can compare with the case  $s = 0$ . We couple  $W_t^{(n)} - W_s^{(n)}$  with two extreme cases, as considered below.

Let  $T_{(\cdot)}$  be the standard process of the number of turns, and define

$$U_{[m,n]} := T_{[m,n]},$$

that is, the number of turns in the segment of the walk during the time interval  $[m, n]$ . Then, one immediately sees that

$$T_{[tn]} - T_{[sn]} \leq U_{[sn],[tn]},$$

as the former may cancel turns created before time  $sn$ .

On the other hand, we have the reversed inequality

$$T_{[tn]} - T_{[sn]} \geq U_{[sn],[tn]} - 2 \left| \min_{1 \leq k \leq \lfloor (t-s)n \rfloor} S_k \right|,$$

and it suffices to estimate the behavior of  $|\min_{0 \leq k \leq n} S_k|$ . Since

$$\mathbb{P} \left( \left| \min_{0 \leq k \leq n} S_k \right| = M \right) \leq \sum_{k=0}^{\lfloor (t-s)n \rfloor} \mathbb{P}(S_k = -M) \leq C \left( \frac{\sqrt{2d-1}}{d} \right)^M,$$

we see that it is bounded by a geometric random variable, and, thus,

$$\frac{1}{\sigma n} \left| \min_{0 \leq k \leq (t-s)n} S_k \right| \rightarrow 0 \text{ in probability.}$$

Note that  $U_{[m,n]}$  has the same distribution as  $T_{n-m}$ ; therefore,  $\mathcal{L}(W_t^n - W_s^n \mid W^n(0, s))$  converges to  $N(0, t - s)$ , and is clearly independent of  $W^n(0, s)$ . By induction, we also get the asymptotically independent increment property for the sequence  $W^n$ .

**Lemma 4.3.**  $W_t^{(n)}$  satisfies Prohorov’s tightness condition, i.e.

$$\lim_{h \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{E} \left( \sup_{|t-s| \leq h} |W_t^{(n)} - W_s^{(n)}| \wedge 1 \right) = 0.$$

*Proof.* Let  $\bar{T}_n := T_n - 2(d - 1)^2 n / d(2d - 1)$  be the centered number of turns at step  $n$ . The idea of the proof is similar to Ottaviani’s maximal inequality for the random walk on  $\mathbb{R}$  (see Lemma 14.8 of [5]).

Fix  $\varepsilon > 0$ . Let  $t \in [0, 1)$  and  $h \in (0, 1 - t)$ . Define  $\tau := \min\{k \in (tn, (t + h)n) : |\bar{T}_k - \bar{T}_{\lfloor tn \rfloor}| \geq 2\varepsilon\sigma\sqrt{n}\}$ . Then

$$\begin{aligned} & \mathbb{P}(|\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_{\lfloor tn \rfloor}| > \varepsilon\sigma\sqrt{n}) \\ & \geq \mathbb{P}(\tau \leq (t + h)n, |\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_\tau| \leq \varepsilon\sigma\sqrt{n}) \\ & = \sum_{k=\lfloor tn \rfloor+1}^{\lfloor (t+h)n \rfloor} \mathbb{P}(\tau = k) \mathbb{P}(|\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_k| \leq \varepsilon\sigma\sqrt{n} \mid \tau = k) \\ & \geq \mathbb{P}(\tau \leq n) \min_{k \in (tn, (t+h)n)} \mathbb{P}(|\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_k| \leq \varepsilon\sigma\sqrt{n} \mid \tau = k). \end{aligned}$$

We want to bound  $\mathbb{P}(\tau \leq n) = \mathbb{P}(\max_k |\bar{T}_k - \bar{T}_{\lfloor tn \rfloor}| > 2\varepsilon\sigma\sqrt{n})$ . First note that, when  $hn$  is large enough,  $\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_{\lfloor tn \rfloor}$  behaves like a Gaussian with mean 0 and variance  $h\sigma^2 n$ . So, there exists  $N(\varepsilon)$  such that, for all  $n > N(\varepsilon)/h$ , we have

$$\mathbb{P}(|\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_{\lfloor tn \rfloor}| > \varepsilon\sigma\sqrt{n}) \leq \exp\left(-\frac{\varepsilon^2}{2h}\right). \tag{4.1}$$

On the other hand,  $|\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_k| \leq U_{(k, (t+h)n)} + M_{\lfloor hn \rfloor}$ . According to Proposition 2.2, the variance of the former term on the right-hand side is  $(t + h)n - k$  with an error uniformly bounded in  $t, h, k$ , and  $n$ . The second term is dominated by a geometric random variable, and, thus, has a finite variance. So, by Chebyshev’s inequality we have

$$\mathbb{P}(|\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_k| \leq \varepsilon\sigma\sqrt{n} \mid \tau = k) \geq 1 - \frac{2[(t + h)n - k + C]}{n}.$$

Since  $k$  takes values between  $tn$  and  $(t + h)n$ , the minimum bound is achieved at  $k = \lfloor tn \rfloor + 1$ , so we obtain

$$\min_{k \in (tn, (t+h)n)} \mathbb{P}(|\bar{T}_{\lfloor (t+h)n \rfloor} - \bar{T}_k| \leq \varepsilon\sigma\sqrt{n} \mid \tau = k) \geq 1 - \frac{2h}{\varepsilon^2} - \frac{C}{n},$$

where  $C$  is independent of  $t, h, k$ , and  $n$ . The last inequality is valid as long as the right-hand side is positive, which requires  $h$  to take small values and  $n$  to take large values. These values depend on  $\varepsilon$  only. The last bound, together with the bound (4.1), implies that

$$\mathbb{P}\left(\max_{k \in (tn, (t+h)n)} |\bar{T}_k - \bar{T}_{\lfloor tn \rfloor}| > 2\varepsilon\sigma\sqrt{n}\right) \leq \left(1 - \frac{2h}{\varepsilon^2} - \frac{C}{n}\right)^{-1} \exp\left(-\frac{\varepsilon^2}{2h}\right). \tag{4.2}$$

Since, for  $n > 16/\varepsilon^2\sigma^2$ , we have

$$\mathbb{P}\left(\sup_{\delta \in (0, h)} |W_{t+\delta}^{(n)} - W_t^{(n)}| > \varepsilon\right) \leq \mathbb{P}\left(\sup_{k \in (tn, (t+h)n)} |\bar{T}_k - \bar{T}_{\lfloor tn \rfloor}| > \frac{\varepsilon}{2}\sigma\sqrt{n}\right),$$

and the bound on the right-hand side of (4.2) is independent of  $t$ , we obtain

$$\sup_{t \in (0,1)} \mathbb{P} \left( \sup_{\delta \in (0,h)} |W_{t+\delta}^{(n)} - W_t^{(n)}| > \varepsilon \right) < \left( 1 - \frac{2h}{\varepsilon^2} - \frac{C}{n} \right)^{-1} \exp \left( -\frac{\varepsilon^2}{32h} \right)$$

for all  $n > \max\{N(\varepsilon)/h, 16/\varepsilon^2\sigma^2\}$ . Now divide the interval  $(0, 1)$  into  $\lfloor 1/h \rfloor + 1$  subintervals, each with length at most  $h$ . Then  $|t-s| < h$  implies that either  $s$  and  $t$  are in the same subinterval or they are in two adjacent intervals. This observation gives

$$\mathbb{P} \left( \sup_{|t-s|<h} |W_t^{(n)} - W_s^{(n)}| > \varepsilon \right) < \left( \frac{2}{h} + 2 \right) \left( 1 - \frac{2h}{\varepsilon^2} - \frac{C}{n} \right)^{-1} \exp \left( -\frac{\varepsilon^2}{32h} \right) \tag{4.3}$$

for all  $n > \max\{N(\varepsilon), 16/\varepsilon^2\sigma^2\}$ . Since

$$\mathbb{E} \left( \sup_{|t-s|<h} |W_t^{(n)} - W_s^{(n)}| \wedge 1 \right) \leq \varepsilon + \mathbb{P} \left( \sup_{|t-s|<h} |W_t^{(n)} - W_s^{(n)}| > \varepsilon \right),$$

the maximal inequality (4.3) quickly gives

$$\lim_{h \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{E} \left( \sup_{|t-s| \leq h} |W_t^{(n)} - W_s^{(n)}| \wedge 1 \right) \leq \varepsilon,$$

which implies Prohorov’s tightness condition since  $\varepsilon$  is arbitrary.

Combining the above two lemmas completes the proof of the invariance principle.

### Acknowledgement

We wish to thank the referee for his/her careful reading of the manuscript and helpful suggestions to improve its presentation.

### References

- [1] CHEN, K.-T. (1957). Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula. *Ann. Math.* **65**, 163–178.
- [2] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications* (Appl. Math. **38**). Springer, New York.
- [3] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd edn. John Wiley, New York.
- [4] HAMBLY, B. AND LYONS, T. (2010). Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. Math.* **171**, 109–167.
- [5] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.
- [6] LYONS, T. J. AND XU, W. (2011). Inversion of signature for paths of bounded variation. In preparation.
- [7] OREY, S. (1958). A central limit theorem for  $m$ -dependent random variables. *Duke Math. J.* **25**, 543–546.