

THREE FAVORITE SITES OCCURS INFINITELY OFTEN FOR ONE-DIMENSIONAL SIMPLE RANDOM WALK¹

BY JIAN DING AND JIANFEI SHEN

University of Chicago

For a one-dimensional simple random walk (S_t) , for each time t we say a site x is a favorite site if it has the maximal local time. In this paper, we show that with probability 1 three favorite sites occurs infinitely often. Our work is inspired by Tóth [*Ann. Probab.* **29** (2001) 484–503], and disproves a conjecture of Erdős and Révész [In *Mathematical Structure—Computational Mathematics—Mathematical Modelling* **2** (1984) 152–157] and of Tóth [*Ann. Probab.* **29** (2001) 484–503].

1. Introduction. Let S_t , $t \in \mathbb{N}$ be a one-dimensional simple random walk with $S_0 = 0$. We define the local time at x by time t to be $L(t, x) = \#\{0 < k \leq t : S_k = x\}$. At time t , we say x is a favorite site if it has the maximal local time, that is, $L(t, x) = \max_y L(t, y)$, and we say that *three favorite sites* occurs if there are exactly three sites which achieve the maximal local time. Our main result states the following.

THEOREM 1.1. *For a one-dimensional simple random walk, with probability 1 three favorite sites occur infinitely often.*

Theorem 1.1 complements the result in [24] which showed that there are no more than three favorite sites eventually, and disproves a conjecture of Erdős and Révész [14–16] and of [24]. Previous to [24], it was shown in [25] that eventually there are no more than three favorite edges.

Besides the number of favorite sites, the asymptotic behavior of favorite sites have been much studied (see [23] for an overview): at time n as $n \rightarrow \infty$, it was shown in [3, 20] that the distance between the favorite sites and the origin in the infimum limit sense is about $\sqrt{n}/\text{poly}(\log n)$ while in the supremum limit sense is about $\sqrt{2n \log \log n}$; it was proved in [8] that the distance between the edge of the range of random walk and the set of favorites increases as fast as $\sqrt{n}/(\log \log n)^{3/2}$; in [7] the jump size for the position of favorite site was studied and shown to be as large as $\sqrt{2n \log \log n}$; a number of other papers [2, 6, 12, 13, 17, 18, 21] studied similar questions in broader contexts including symmetric stable processes, random walks on random environments and so on.

Received March 2017; revised September 2017.

¹Supported in part by NSF Grant DMS-1455049 and an Alfred Sloan fellowship.
MSC2010 subject classifications. 60J15, 60J55.

Key words and phrases. Random walk, favorite sites.

In two dimensions and higher, favorite sites for simple random walks have been intensively studied where some intriguing fractal structure arise; see, for example, [1, 9, 10, 22]. Such fractal structure also plays a central role in the study of cover times for random walks; see, for example, [4, 5, 11]. We refrain from an extensive discussion on the literature on this topic as the mathematical connection to the concrete problem considered in the present article is limited. That being said, we remark that analogous questions on the number of favorite sites in two dimensions and higher are of interest for future research, which we expect to be more closely related to the literature mentioned in this paragraph as well as references therein.

Our proof is inspired by [24], which in turn was inspired by [25]. Following [24], we define the number of upcrossings and downcrossings at x by the time t to be

$$U(t, x) = \#\{0 < k \leq t : S_k = x, S_{k-1} = x - 1\},$$

$$D(t, x) = \#\{0 < k \leq t : S_k = x, S_{k-1} = x + 1\}.$$

It is elementary to check that (see, e.g., [24], equation (1.6))

$$(1.1) \quad \begin{aligned} L(t, x) &= D(t, x) + D(t, x - 1) + \mathbf{1}_{\{0 < x \leq S(t)\}} - \mathbf{1}_{\{S(t) < x \leq 0\}} \\ &= U(t, x) + U(t, x + 1) + \mathbf{1}_{\{S(t) \leq x < 0\}} - \mathbf{1}_{\{0 \leq x < S(t)\}}. \end{aligned}$$

The set of favorite (or most visited) sites $\mathcal{H}(t)$ of the random walk at time $t \in \mathbb{N}$ consists of those sites where the local time attains its maximum value, that is,

$$\mathcal{H}(t) = \left\{ y \in \mathbb{Z} : L(t, y) = \max_{z \in \mathbb{Z}} L(t, z) \right\}.$$

For $r \geq 1$, let $f(r)$ be the (possibly infinite) number of times when the currently occupied site is one of the r favorites:

$$f(r) = \#\{t \geq 1 : S_t \in \mathcal{H}(t), \#\mathcal{H}(t) = r\}.$$

We remark that one of the main conceptual contributions in [24, 25] is the introduction of this function $f(r)$. Effectively, $f(r)$ counts the clusters of instances for r favorite sites; it is plausible that after the random walk leaves one of the favorite sites, within a nonnegligible (random) number of steps those r favorite sites will remain favorite sites. Therefore, the expectation of $f(r)$ is significantly smaller than the expected number of t at which r favorite sites occurs, and in fact it was shown in [24] that $\mathbb{E}f(r) < \infty$ for all $r \geq 4$. It was then conjectured in [24] that $f(3) < \infty$ with probability 1, even though from the computations in [24] it was clear that $\mathbb{E}f(3) = \infty$. In the current article, we will show, using the idea of counting clusters in [24], that the correlation becomes so small that the first moment dictates the behavior. That is to say, we will show that

$$(1.2) \quad f(3) = \infty \quad \text{with probability 1,}$$

which then yields Theorem 1.1.

The rest of the paper is organized as follows: in Section 2, we will set up the framework of our proof following [24]; in Section 3, we first show that $f(3) = \infty$ with positive probability and then prove (1.2) by demonstrating a 0–1 law. We emphasize that the first moment computation in Section 3.1 follows from arguments in [24], and the main novelty of our work is on the second moment computation in Section 3.2.

2. Preliminaries. In this section, we recall the framework of [24] with suitable adaption to our setup, and collect a number of useful and well-understood facts. We claim no originality in this section, and the existence of the current section is mainly for the completeness of notation and definition.

2.1. Three consecutive favorite sites. It turns out that in order to show $f(3) = \infty$ it suffices to consider instances of three favorite sites which are consecutive. To this end, we define the inverse edge local times by

$$T_U(k, x) \triangleq \inf\{t \geq 1 : U(t, x) = k\} \quad \text{and} \quad T_D(k, x) \triangleq \inf\{t \geq 1 : D(t, x) = k\}.$$

We consider the events of three consecutive favorite sites, that is,

$$A_{x,h}^{(k)} \triangleq \{\mathcal{K}(T_U(k+1, x)) = \{x, x+1, x+2\}, L(T_U(k+1, x), x) = h\}.$$

We write the events in $T_U(k+1, x)$ rather than $T_U(k, x)$ as it matches the form of the Ray–Knight representation which we will discuss later. We then let $I_h = (\frac{1}{2}(h + \sqrt{h}), \frac{1}{2}(h + 2\sqrt{h}))$ and define

$$N_H = \sum_{h=1}^H \sum_{k \in I_h} \sum_{x=1}^{\infty} \mathbf{1}_{A_{x,h}^{(k)}} \quad \text{and} \quad N = \lim_{H \rightarrow \infty} N_H = \sum_{h=1}^{\infty} \sum_{k \in I_h} \sum_{x=1}^{\infty} \mathbf{1}_{A_{x,h}^{(k)}}.$$

We observe that for each h , the events $A_{x,h}^{(k)}$ are mutually disjoint. In addition, we have that $f(3) \geq u(x)$ where

$$\begin{aligned} u(x) &= \sum_{t=1}^{\infty} \mathbf{1}_{\{S(t-1)=x-1, S(t)=x, x \in \mathcal{K}(t), \#\mathcal{K}(t)=3\}} \\ &= \sum_{k=1}^{\infty} \mathbf{1}_{\{x \in \mathcal{K}(T_U(k, x)), \#\mathcal{K}(T_U(k, x))=3\}} \\ &= \sum_{k=0}^{\infty} \sum_{h=1}^{\infty} \mathbf{1}_{\{x \in \mathcal{K}(T_U(k+1, x)), \#\mathcal{K}(T_U(k+1, x))=3, L(T_U(k+1, x), x)=h\}}. \end{aligned}$$

Therefore, we have that $f(3) \geq N$, and thus it suffices to show that $N = \infty$. We remark that the preceding discussions are extracted from decompositions in [24], (2.3), (2.4), (2.5), and they are the starting point for all computations in [24] as well as the present article.

2.2. *Additive processes and the Ray–Knight representation.* Throughout this paper, we denote by Y_t a critical Galton–Watson branching process with geometric offspring distribution and by Z_t, R_t critical geometric branching processes with one immigrant in each generation (in different ways). More precisely, we let $X_{t,i}$ ’s be i.i.d. geometric variables with mean 1 and recursively define

$$(2.1) \quad Z_{t+1} = \sum_{i=1}^{Z_t+1} X_{t,i} \quad \text{and} \quad R_{t+1} = 1 + \sum_{i=1}^{R_t} X_{t,i}.$$

One can verify that Y_t, Z_t and R_t are Markov chains with state space \mathbb{Z}_+ and transition probabilities:

$$(2.2) \quad \mathbb{P}(Y_{t+1} = j | Y_t = i) = \pi(i, j) \triangleq \begin{cases} \delta_0(j) & \text{if } i = 0, \\ 2^{-i-j} \frac{(i+j-1)!}{(i-1)!j!} & \text{if } i > 0, \end{cases}$$

$$\mathbb{P}(Z_{t+1} = j | Z_t = i) = \rho(i, j) \triangleq \pi(i+1, j)$$

and $\mathbb{P}(R_{t+1} = j | R_t = i) = \rho^*(i, j) \triangleq \pi(i, j-1).$

Let $k \geq 0$ and x be fixed integers. When $x \geq 1$, define the following three processes:

1. $(Z_t^{(k)})_{t \geq 0}$, is a Markov chain with transition probability $\rho(i, j)$ and initial state $Z_0 = k$.
2. $(Y_t^{(k)})_{t \geq -1}$, is a Markov chain with transition probabilities $\pi(i, j)$ and initial state $Y_{-1} = k$.
3. $(Y_t^{\prime(k)})_{t \geq 0}$, is a Markov chain with transition probabilities $\pi(i, j)$ and initial state $Y_0^{\prime(k)} = Z_{x-1}^{(k)}$.

The three processes are independent, except for the fact that $Y_t^{\prime(k)}$ starts from the terminal state of $Z_t^{(k)}$. We patch the three processes together to a single process:

$$\Delta_x^{(k)}(y) \triangleq \begin{cases} Z_{x-1-y}^{(k)} & \text{if } 0 \leq y \leq x-1, \\ Y_{y-x}^{(k)} & \text{if } x-1 \leq y \leq \infty, \\ Y_{-y}^{\prime(k)} & \text{if } -\infty < y \leq 0. \end{cases}$$

We also define

$$(2.3) \quad \Lambda_x^{(k)}(y) \triangleq \Delta_x^{(k)}(y) + \Delta_x^{(k)}(y-1) + \mathbf{1}_{\{0 < y \leq x\}}.$$

From the Ray–Knight theorems on local time of simple random walks on \mathbb{Z} (cf. [19], Theorem 1.1), it follows that for any integers $x \geq 1$ and $k \geq 0$,

$$(2.4) \quad (D(T_U(k+1, x), y), y \in \mathbb{Z}) \stackrel{\text{law}}{=} (\Delta_x^{(k)}(y), y \in \mathbb{Z}).$$

Using (1.1), (2.3) and (2.4), we get

$$(2.5) \quad (L(T_U(k + 1, x), y), y \in \mathbb{Z}) \stackrel{\text{law}}{=} (\Lambda_x^{(k)}(y), y \in \mathbb{Z}).$$

Similarly, when $x \leq 0$, we define the processes:

1. $(R_t^{(k)})_{t \geq 0}$, is a Markov chain with transition probability $\rho^*(i, j)$ and initial state $R_{-1} = k$.
2. $(Y_t^{(k)})_{t \geq 0}$, is a Markov chain with transition probabilities $\pi(i, j)$ and initial state $Y_0 = k$.
3. $(Y_t^{\prime(k)})_{t \geq -1}$, is a Markov chain with transition probabilities $\pi(i, j)$ and initial state $Y_{-1}^{\prime(k)} = R_{-1-x}^{(k)}$.

In this case, we patch the three processes together by

$$\Delta_x^{(k)}(y) \triangleq \begin{cases} Y_y^{\prime(k)} & \text{if } -1 \leq y < \infty, \\ R_{y-x} & \text{if } x - 1 \leq y \leq -1, \\ Y_{x-1-y}^{(k)} & \text{if } -\infty < y \leq x - 1. \end{cases}$$

The corresponding $\Lambda_x^{(k)}$ is defined by

$$\Lambda_x^{(k)}(y) \triangleq \Delta_x^{(k)}(y) + \Delta_x^{(k)}(y - 1) - \mathbf{1}_{\{x < y \leq 0\}}.$$

By classical Ray–Knight theorems, we get the couplings for the case $k \geq 0, x \leq 0$:

$$(2.6) \quad (D(T_U(k + 1, x), y), y \in \mathbb{Z}) \stackrel{\text{law}}{=} (\Delta_x^{(k)}(y), y \in \mathbb{Z}),$$

$$(2.7) \quad (L(T_U(k + 1, x), y), y \in \mathbb{Z}) \stackrel{\text{law}}{=} (\Lambda_x^{(k)}(y), y \in \mathbb{Z}).$$

In this paper, we will mainly use the Ray–Knight representation (2.4) and (2.5), while (2.6) and (2.7) will be used in the calculation of $\mathbb{E}N_H^2$. In the following, we default $x > 0$ unless mentioned otherwise.

2.3. *Three favorite sites under Ray–Knight representation.* To utilize (2.5), given the additive processes $Y_t^{(k)}, Z_t^{(k)}$ and $Y_t^{\prime(k)}$, we define

$$\tilde{Z}_t^{(k)} \triangleq Z_t^{(k)} + Z_{t-1}^{(k)} + 1, \quad \tilde{Y}_t^{(k)} \triangleq Y_t^{(k)} + Y_{t-1}^{(k)}, \quad \tilde{Y}_t^{\prime(k)} \triangleq Y_t^{\prime(k)} + Y_{t-1}^{\prime(k)}.$$

For $h \in \mathbb{Z}_+$, define the first hitting time of $[h, \infty)$ for $Y_t^{(k)}$ and $Z_t^{(k)}$ to be $\sigma_h^{(k)}$ and $\tau_h^{(k)}$, respectively, and the extinction time of $Y_t^{(k)}$ to be $\omega^{(k)}$. That is,

$$(2.8) \quad \begin{aligned} \sigma_h^{(k)} &\triangleq \inf\{t \geq 0 : Y_t^{(k)} \geq h\}, & \tau_h^{(k)} &\triangleq \inf\{t \geq 0 : Z_t^{(k)} \geq h\}, \\ \text{and } \omega^{(k)} &= \inf\{t \geq 0 : Y_t^{(k)} = 0\}. \end{aligned}$$

Correspondingly, we define the first hitting time of $[h, \infty)$ for the process $\tilde{Y}_t^{(k)}$ and $\tilde{Z}_t^{(k)}$ to be $\tilde{\sigma}_h^{(k)}$ and $\tilde{\tau}_h^{(k)}$, respectively. Namely,

$$\tilde{\sigma}_h^{(k)} \triangleq \inf\{t \geq 0 : \tilde{Y}_t^{(k)} \geq h\}, \quad \tilde{\tau}_h^{(k)} \triangleq \inf\{t \geq 0 : \tilde{Z}_t^{(k)} \geq h\}.$$

Using the notation above, we can write $\mathbb{P}(A_{h,x}^{(k)})$ in its Ray–Knight representation form. That is, $\mathbb{P}(A_{h,x}^{(k)})$ is equal to

$$\mathbb{P}(Y_0^{(k)} = h - k - 1, Y_1^{(k)} = k + 1, Y_2^{(k)} = h - k - 1, \{\tilde{Y}_t^{(k)} < h, \text{ for } t \geq 3\}, \\ \{\tilde{Z}_t^{(k)} < h, \text{ for } 1 \leq t \leq x - 1\}, \{\tilde{Y}'_t^{(k)} < h, \text{ for } t \geq 1\}).$$

For all the notation above, when the initial state of a process is obvious, we omit the superscript “ (k) ” to avoid cumbersome notation. We will also use conditional probability $\mathbb{P}(\cdot | Y_0 = k)$ to indicate the initial state.

2.4. *Standard lemmas.* In this subsection, we record a few well-understood lemmas that will be useful later.

LEMMA 2.1 ([24], (6.14)–(6.15)). *For any $0 \leq k \leq h \leq u$, the following overshoot bounds hold:*

$$\mathbb{P}(Y_{\sigma_h} \geq u | Y_0 = k, \sigma_h < \infty) \leq \mathbb{P}(Y_1 \geq u | Y_0 = h, Y_1 \geq h), \\ \mathbb{P}(Z_{\tau_h} \geq u | Z_0 = k) \leq \mathbb{P}(Z_1 \geq u | Z_0 = h, Z_1 \geq h).$$

LEMMA 2.2. *We have that:*

- (i) *For $i, j \in (\frac{1}{2}(h - 10\sqrt{h}), \frac{1}{2}(h + 10\sqrt{h}))$, there exist positive constants c and C such that $ch^{-\frac{1}{2}} \leq \pi(i, j) \leq Ch^{-\frac{1}{2}}$ for all $h \geq 1$.*
- (ii) *For $i + j = h$, $\pi(i, j) \leq O(1)h^{-\frac{1}{2}}$.*
- (iii) *For $j < i_1 < i_2$, $\pi(i_1, j) > \pi(i_2, j)$.*

PROOF. Properties (i) and (ii) follow from straightforward computation using Stirling’s formula and (2.2). For Property (iii), we see that $\frac{\pi(i+1,j)}{\pi(i,j)} = \frac{i+j}{2i} < 1$ for $j < i$, and (iii) follows from induction. \square

LEMMA 2.3. *We have that $\mathbb{E}\tau_h = \mathbb{E}Z_{\tau_h} - Z_0$. In particular, we have that $\mathbb{E}[\tau_h | Z_0 = k] \geq h - k$.*

PROOF. Applying the optional stopping theorem to the martingale $Z_t - t$ at time τ_h , we get $\mathbb{E}\tau_h = \mathbb{E}Z_{\tau_h} - Z_0 \geq h - k$, as desired. \square

3. Proof of Theorem 1.1. The current section contains three parts: in Section 3.1, we adapt the arguments in [24] and provide a lower bound on the first moment for the number of instances for the consecutive three favorite sites; in Section 3.2 (which contains the main novelty of the present paper), we show that the second moment is of the same order as the square of the first moment, thereby proving that three favorite sites occurs with nonvanishing probability; in Section 3.3, we prove a 0–1 law for three favorite sites, and thus complete the proof of Theorem 1.1.

3.1. *Lower bound on the first moment.* For $x > 0$ and $h \in \mathbb{N}$, in order to bound the probability for three consecutive favorite sites with local time h at vertices $x, x + 1$ and $x + 2$, the main part is to control the probability for the local times below h everywhere except at $x, x + 1$ and $x + 2$. To this end, it suffices to consider the edge local times (i.e., number of downcrossings) in the Ray–Knight representation with appropriate conditioning in the region of $(x, x + 2)$. Then in the region outside of $(0, x + 2)$, these edge local times evolve as martingales [when looking forward spatially in $(x + 2, \infty)$ and backward spatially in $(-\infty, 0)$] and it is fairly standard to control the probability of staying below the level h ; in the region $(0, x)$, the edge local times are not exactly a martingale [when looking backward spatially; see (2.1)] and the analysis is slightly more complicated. In the next lemma, we prove a lower bound on the first moment of $\sum_{t=1}^{\tau_h} \frac{h-Z_t}{h}$. Combined with standard martingale analysis in the region outside of $(0, x + 2)$ and a change of summation when summing over x [see (3.5)], this will then give a lower bound on the first moment of N_H (see Proposition 3.2).

LEMMA 3.1. *Suppose that $Z_0 = k \in [h - 2\sqrt{h}, h - \sqrt{h}]$. Then there exists a constant $c > 0$ such that $\mathbb{E}(\sum_{t=1}^{\tau_h} \frac{h-Z_t}{h}) \geq c\sqrt{h}$.*

PROOF. Let $M_t = \sum_{s=1}^t (Z_s - s) - t(Z_t - t)$, and let $\mathcal{F}_t = \sigma(Z_0, Z_1, \dots, Z_t)$. We see that

$$\mathbb{E}(M_{t+1} | \mathcal{F}_t) = \left[\sum_{s=1}^t (Z_s - s) + (Z_t - t) \right] - (t + 1)(Z_t - t) = M_t.$$

Thus (M_t) is a martingale. By the optional stopping theorem, we see that $\mathbb{E}(\sum_{t=1}^{\tau_h} (Z_t - t)) = \mathbb{E}\tau_h(Z_{\tau_h} - \tau_h)$, and hence

$$(3.1) \quad \mathbb{E}\left(\sum_{t=1}^{\tau_h} \frac{h - Z_t}{h}\right) = \left(1 + \frac{1}{2h}\right)\mathbb{E}\tau_h - \frac{1}{h}\mathbb{E}\left[\tau_h Z_{\tau_h} - \frac{1}{2}\tau_h^2\right].$$

Now consider the process $M'_t = -\frac{1}{4}Z_t^2 + tZ_t - \frac{1}{2}t^2 + \frac{1}{4}t$. By (2.1), we see that

$$\mathbb{E}(M'_{t+1} | \mathcal{F}_t) = -\frac{1}{4}(Z_t^2 + 4Z_t + 3) + (tZ_t + Z_t + t + 1) - \frac{1}{2}(t^2 + 2t + 1) + \frac{1}{4}(t + 1),$$

where equal to M'_t . So (M'_t) is a martingale. Using the optional stopping theorem to (M'_t) at τ_h , we have

$$\begin{aligned}
 \mathbb{E}\left[\tau_h Z_{\tau_h} - \frac{1}{2}\tau_h^2\right] &= \mathbb{E}\left[\frac{1}{4}Z_{\tau_h}^2 - \frac{1}{4}\tau_h\right] - \frac{1}{4}Z_0^2 \\
 (3.2) \qquad \qquad \qquad &= \frac{1}{4}\mathbb{E}(Z_{\tau_h}^2 - Z_0^2) - \frac{1}{4}\mathbb{E}\tau_h.
 \end{aligned}$$

Combining (3.1), (3.2) and Lemma 2.3, we get

$$\begin{aligned}
 \mathbb{E}\left[\sum_{t=1}^{\tau_h} \frac{h - Z_t}{h}\right] &= \left(1 + \frac{1}{4h}\right)\mathbb{E}\tau_h - \frac{1}{4h}\mathbb{E}[Z_{\tau_h}^2 - Z_0^2] \\
 &= \left(1 + \frac{1}{4h}\right)\mathbb{E}(Z_{\tau_h} - Z_0) - \frac{1}{4h}\mathbb{E}[(Z_{\tau_h} - Z_0)(Z_{\tau_h} + Z_0)] \\
 &\geq \frac{1}{4h}\mathbb{E}[(Z_{\tau_h} - Z_0)(4h - (Z_{\tau_h} + Z_0))].
 \end{aligned}$$

Obviously, $Z_{\tau_h} - Z_0 \geq h - k \geq \sqrt{h}$ and by Lemma 2.1 we have that $\mathbb{E}(Z_{\tau_h} - Z_0)(Z_{\tau_h} + Z_0 - 2h) = O(h)$. Therefore, there is a constant c such that $\mathbb{E}[\sum_{t=1}^{\tau_h} \frac{h - Z_t}{h}] \geq c\sqrt{h}$ for sufficiently large h . \square

PROPOSITION 3.2. *For a constant $c > 0$, we have $\mathbb{E}N_H \geq c \log H$.*

PROOF. In what follows, c_i for $i \geq 1$ and c are all constants. By the Ray-Knight representation, $\mathbb{E}N_H$ is equal to the following product:

$$\begin{aligned}
 &\sum_{h=1}^H \sum_{k \in I_h} \mathbb{P}(Y_0^{(k)} = h - k - 1, Y_1^{(k)} = k + 1, Y_2^{(k)} = h - k - 1, \\
 &\quad \{\tilde{Y}_t^{(k)} < h, \text{ for } t \geq 3\}) \\
 &\quad \times \sum_{x=1}^{\infty} \mathbb{P}(\{\tilde{Z}_t^{(k)} < h, 1 \leq t \leq x - 1\}, \{\tilde{Y}'_t^{(k)} < h, \text{ for } t \geq 1\}).
 \end{aligned}$$

Thus, we get that

$$\begin{aligned}
 \mathbb{E}N_H &\geq \sum_{h=1}^H \sum_{k \in I_h} \pi(l, h - k - 1)\pi(h - k - 1, k + 1)\pi(k + 1, h - k - 1) \\
 &\quad \times \mathbb{P}\left(Y_t^{(h-k-1)} < \frac{1}{2}h \text{ for } t \geq 0\right) \cdot \sum_{x=1}^{\infty} \mathbb{P}(\tilde{\tau}_h \geq x, \{\tilde{Y}'_t^{(k)} < h, \text{ for } t \geq 1\}).
 \end{aligned}$$

By Lemma 2.2(i), all $\pi(\cdot, \cdot)$ in the above equation are at the scale $h^{-\frac{1}{2}}$. Since Y_t is a martingale, by using the optional stopping theorem at $\sigma_{\frac{h}{2}} \wedge \omega$ where $\sigma_{\frac{h}{2}}$ and ω

are defined in (2.8), we have

$$\begin{aligned} \mathbb{P}\left(Y_t^{(h-k-1)} < \frac{h}{2} \text{ for } t \geq 0\right) &= \mathbb{P}\left(Y_t^{(h-k-1)} \text{ hits } 0 \text{ before } \frac{h}{2}\right) \\ &\geq \frac{h/2 - (h-k-1)}{h/2} \geq c_1 h^{-\frac{1}{2}}. \end{aligned}$$

So we get

$$(3.3) \quad \mathbb{E}N_H \geq c_2 \sum_{h=1}^H \sum_{k \in I_h} \sum_{x=1}^{\infty} h^{-2} \mathbb{P}(\tilde{\tau}_h \geq x, \{\tilde{Y}_t^{(k)} < h, t \geq 1\}).$$

Let $k_1 = \frac{1}{2}(h - 2\sqrt{h})$. By independence in the Ray–Knight representation,

$$\begin{aligned} &\sum_{x=1}^{\infty} \mathbb{P}(\tilde{\tau}_h \geq x, \{\tilde{Y}_t^{(k)} < h, \text{ for } t \geq 1\}) \\ &\geq \sum_{x=1}^{\infty} \mathbb{P}\left(Z_1^{(k)} \leq k_1, Z_t^{(k)} < \frac{h}{2} \text{ for } 2 \leq t \leq x-1, \left\{Y_t^{(k)} < \frac{h}{2}, \text{ for } t \geq 1\right\}\right) \\ &\geq \sum_{x=2}^{\infty} \sum_{l=0}^{\lfloor \frac{h}{2}-1 \rfloor} \left(\mathbb{P}(Z_1^{(k)} \leq k_1) \cdot \mathbb{P}\left(Z_t^{(k_1)} < \frac{h}{2} \text{ for } 1 \leq t \leq x-2, Z_{x-2}^{(k_1)} = l\right)\right. \\ &\quad \left. \times \mathbb{P}\left(Y_t^{(l)} \text{ hits } 0 \text{ before } \frac{h}{2}\right)\right). \end{aligned}$$

By Lemma 2.2(i), $\mathbb{P}(Z_1^{(k)} \leq k_1) \geq c_3$. Using the optional stopping theorem again, we have $\mathbb{P}(Y_t^{(l)} \text{ hits } 0 \text{ before } \frac{h}{2}) \geq \frac{h/2-l}{h/2}$. So

$$(3.4) \quad \begin{aligned} &\sum_{x=1}^{\infty} \mathbb{P}(\tilde{\tau}_h \geq x, \{\tilde{Y}_t^{(k)} < h, t \geq 1\}) \\ &\geq c_3 \cdot \sum_{x=1}^{\infty} \sum_{l=0}^{\lfloor \frac{h}{2}-1 \rfloor} \mathbb{P}(\tau_{h/2}^{(k_1)} \geq x, Z_{x-1}^{(k_1)} = l) \cdot \frac{h/2-l}{h/2}. \end{aligned}$$

By interchange of the summation and the expectation (which is valid by the monotone convergence theorem) and Lemma 3.1, we have that the right- hand side of (3.4) is equal to

$$(3.5) \quad \begin{aligned} c_3 \cdot \mathbb{E}\left[\sum_{l=0}^{\lfloor \frac{h}{2}-1 \rfloor} \sum_{x=1}^{\tau_{h/2}^{(k_1)}} \frac{h/2-l}{h/2} \cdot \mathbf{1}_{\{Z_{x-1}^{(k_1)}=l\}}\right] &= c_3 \mathbb{E}\left(\sum_{t=0}^{\tau_{h/2}^{(k_1)}-1} \frac{h/2-Z_t^{(k_1)}}{h/2}\right) \\ &\geq c_4 \sqrt{h}, \end{aligned}$$

where in the second inequality we did change of variable $t = x - 1$. Thus by (3.3) and (3.5),

$$\mathbb{E}N_H \geq \sum_{h=1}^H \sum_{k \in I_h} c_5 h^{-\frac{3}{2}} \geq c_6 \cdot \sum_{h=1}^H \frac{1}{h} \geq c \log H,$$

completing the proof of the proposition. \square

3.2. Upper bound on the second moment. The calculation of second moment involves the two three favorite sites that happen in chronological order. The key insight is that two instances of three favorite sites with no spatial overlap are almost independent. Before giving the bound for the second moment, we discuss some useful concepts and tools that characterize the independence of different three favorite sites.

Let $D(t) = (D(t, x), x \in \mathbb{Z}) \in \mathbb{N}^{\mathbb{Z}}$ be the random vector that records the number of downcrossings of each site by the time t . For $\ell \in \mathbb{N}^{\mathbb{Z}}$, we use $\ell(i), i \in \mathbb{Z}$ to denote the i th component of ℓ . For $\ell \in \mathbb{N}^{\mathbb{Z}}$, define $B_x(\ell) = \{\exists t < \infty : D(t) = \ell, S(t - 1) = x - 1, S(t) = x\}$. Note that if $B_x(\ell)$ happens, there exists a unique $t \in \mathbb{N}$ such that $D(t) = \ell, S(t - 1) = x - 1$ and $S(t) = x$. Sometimes we abuse the terminology “after $B_x(\ell)$ happens” by meaning “after the unique t with $D(t) = \ell, S(t - 1) = x - 1, S(t) = x$.” We also say “ $B_x(\ell)$ happens before $B_{x'}(\ell')$ ” by meaning the unique t [corresponding to $B_x(\ell)$] is less than the unique t' [corresponding to $B_{x'}(\ell')$].

Let $\mathcal{P} = \{\ell : \mathbb{P}(B_x(\ell)) > 0 \text{ for some } x\}$. Clearly, for any $\ell \in \mathcal{P}$, ℓ has compact support. For $\mathcal{Q} \subset \mathcal{P}$, denote $B_x(\mathcal{Q}) = \bigcup_{\ell \in \mathcal{Q}} B_x(\ell)$. Then we have $A_{x,h}^{(k)} = B_x(\mathcal{P}_{x,h}^{(k)})$ where $\mathcal{P}_{x,h}^{(k)}$ is the collection of $\ell \in \mathcal{P}$ such that

$$\begin{aligned} \ell(x - 1) &= k, & \ell(x) &= h - k - 1, \\ \ell(x + 1) &= k + 1, & \ell(x + 2) &= h - k - 1; \\ \ell(i - 1) + \ell(i) &< h & \text{for all } i &\neq x, x + 1, x + 2. \end{aligned}$$

Our main intuition on bounding the correlation between two instances of three favorite sites is the following: Suppose at some time (say T_1) we have an instance of three favorite points at $x, x + 1, x + 2$ with edge local time (i.e., downcrossings) given by ℓ . Our crucial observation is that conditioning on $B_x(\ell)$ does not increase much of the probability for producing an instance of three favorite sites in a future time (say T_2) which are spatially different from those of ℓ . To this end, we let ℓ' be one of many local perturbations of ℓ (which are obtained from ℓ by decreasing the values at $x + 1$ and $x + 2$). We note that (see Figure 1 for an illustration)

- The event $B_x(\ell)$ [resp., $B_x(\ell')$] corresponds to that the edge local time is ℓ (resp., ℓ') when the random walk cross the directed edge $(x - 1, x)$ for the $(\ell(x - 1) + 1)$ 'th time [note that $\ell(x - 1) = \ell'(x - 1)$; and note that this corresponds

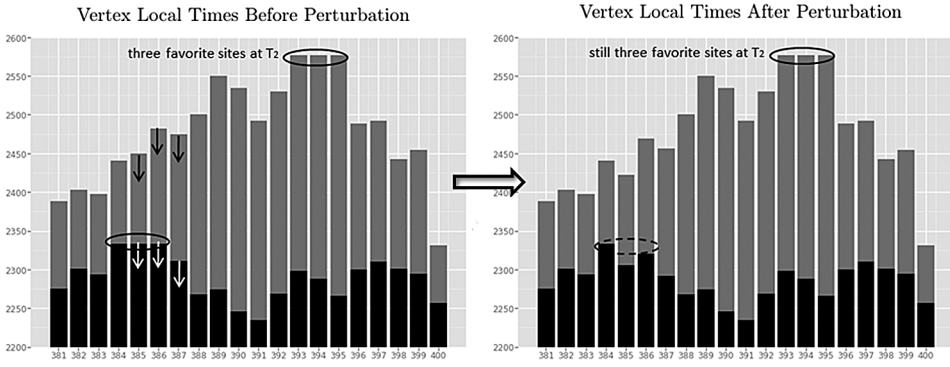


FIG. 1. The black bars represent vertex local times at T_1 and the grey bars represent ones at T_2 . When we decrease the edge local times at $x + 1$ and $x + 2$, descent of vertex local times happens at $x + 1$, $x + 2$ and $x + 3$. After the local time perturbation at time T_1 , we will still get “three favorite sites” at T_2 .

to time T_1 in Figure 1]. Conditioned on $B_x(\ell)$ [resp., $B_x(\ell')$], the edge local time at a later time (which corresponds to T_2 in Figure 1) is ℓ (respectively, ℓ') superposed with an independent edge local time field which we denote by $\tilde{\ell}$. By the strong Markov property for random walks, the law of $\tilde{\ell}$ is the same regardless of conditioning on $B_x(\ell)$ or $B_x(\ell')$.

- If the field $(\ell + \tilde{\ell})$ produces three favorite sites which are spatially different from those of ℓ , then the field $(\ell' + \tilde{\ell})$ also produces three favorite sites.

In summary, we see that the conditional probability of producing an instance of three favorite sites which are spatially different from those of ℓ given $B_x(\ell)$ is the same as the conditional probability given $B_x(\ell')$. But the probability for the union of $B_x(\ell')$'s when ℓ' ranging over all legitimate perturbations is much larger than that of $B_x(\ell)$ —in fact larger by a factor of order $h = \ell(x - 1) + \ell(x) + 1$ (see Lemma 3.4 below). This is a (quantitative) manifestation that the event $B_x(\ell)$ is uncorrelated with a spatially different instance of three favorite sites in the future.

Our formal proof does not exactly follow the discussion above on controlling the conditional probability, as it turns out slightly simpler to directly compute the joint probability for two instances of three favorite sites (but the intuition is the same). For the precise implementation, we let \mathcal{A} be the set of all subsets of \mathcal{P} and define a map $\varphi_x : \mathcal{P} \mapsto \mathcal{A}$ mapping an $\ell \in \mathcal{P}$ to a collection of vectors where we locally push down the values at locations $x + 1$ and $x + 2$. More precisely, we define $\varphi_x(\ell)$ to be

$$\{\ell^* \in \mathcal{P} : \ell^*(i) < \ell(i) \text{ for } i = x + 1, x + 2, \ell^*(i) = \ell(i) \text{ for } i \neq x + 1, x + 2\}.$$

LEMMA 3.3. For $i = 1, 2$ and $\ell_i^* \in \varphi_{x_i}(\ell_i)$ with $\ell_i \in \mathcal{P}_{x_i, h}^{(k_i)}$, we have that $B_{x_1}(\ell_1^*) \cap B_{x_2}(\ell_2^*) = \emptyset$ if $(x_1, \ell_1) \neq (x_2, \ell_2)$. Further, we have $B_{x_1}(\ell_1^*) \cap B_{x_2}(\ell_2^*) = \emptyset$ if $(x_1, \ell_1) = (x_2, \ell_2)$ but $\ell_1^* \neq \ell_2^*$.

PROOF. Case (i): Suppose $x_1 \neq x_2$. Since clearly $B_{x_1}(\ell_1^*)$ and $B_{x_2}(\ell_2^*)$ cannot happen at the same time t , we can then assume without loss of generality that $B_{x_1}(\ell_1^*)$ happens first. Then when $B_{x_2}(\ell_2^*)$ happens the vertex local time at x_1 is at least h , arriving at a contradiction.

Case (ii): Suppose that $x_1 = x_2$ but $\ell_1 \neq \ell_2$. In this case, we have $\ell_1^* \neq \ell_2^*$. Since clearly $B_{x_1}(\ell_1^*)$ and $B_{x_2}(\ell_2^*)$ cannot happen at the same time t , we can then assume without loss of generality that $B_{x_1}(\ell_1^*)$ happens first. In order for $B_{x_2}(\ell_2^*)$ to happen, the random walk has to leave $x_1 (= x_2)$ and revisit x_1 . As a result, the vertex local time at x_1 will be strictly larger than h , arriving at a contradiction.

Case (iii): Suppose that $x_1 = x_2$, $\ell_1 = \ell_2$ but $\ell_1^* \neq \ell_2^*$. This follows from the same reasoning as in Case (ii). \square

LEMMA 3.4. *There exists a constant $c > 0$ such that for any $\ell \in P_{x,h}^{(k)}$ with $k \in I_h$,*

$$\mathbb{P}(B_x(\varphi_x(\ell))) \geq ch\mathbb{P}(B_x(\ell)).$$

PROOF. We consider $\ell^* \in \varphi_x(\ell)$ such that $\ell^*(x + 1) \in [k + 1 - \sqrt{h}, k + 1)$ and $\ell^*(x + 2) \in [h - k - 1 - \sqrt{h}, h - k - 1)$. According to Lemma 2.2(i) and (iii), there is a constant $c > 0$ such that

$$\begin{aligned} \frac{\mathbb{P}(B_x(\ell^*))}{\mathbb{P}(B_x(\ell))} &= \frac{\pi(\ell^*(x), \ell^*(x + 1))\pi(\ell^*(x + 1), \ell^*(x + 2))\pi(\ell^*(x + 2), \ell(x + 3))}{\pi(h - k - 1, k + 1)\pi(k + 1, h - k - 1)\pi(h - k - 1, \ell(x + 3))} \\ &\geq c. \end{aligned}$$

Note that there are about h of such $\ell^* \in \varphi_x(\ell)$ that satisfy the inequality. By Lemma 3.3, we get that $\mathbb{P}(B_x(\varphi_x(\ell))) \geq ch\mathbb{P}(B_x(\ell))$. \square

PROPOSITION 3.5. *We have that $\mathbb{E}N_H^2 = O(\log H) \cdot \mathbb{E}N_H$.*

PROOF. We decompose the second moment into the following three parts:

$$\begin{aligned} \mathbb{E}N_H^2 &= 2 \sum_{1 \leq h < h' \leq H} \sum_{k \in I_h} \sum_{k' \in I_{h'}} \sum_{x=1}^{\infty} \sum_{x'=1}^{\infty} \mathbb{P}(A_{x,h}^{(k)}, A_{x',h'}^{(k')}) + \mathbb{E}N_H \\ (3.6) \quad &\leq O(1) \cdot (\text{I} + \text{II} + \mathbb{E}N_H), \end{aligned}$$

where

$$\begin{aligned} \text{I} &= \sum_{1 \leq h < h' \leq H} \sum_{k' \in I_{h'}} \sum_{k \in I_h} \sum_{|x' - x| > 3} \mathbb{P}(A_{x,h}^{(k)}, A_{x',h'}^{(k')}), \\ \text{II} &= \sum_{1 \leq h < h' \leq H} \sum_{k \in I_h} \sum_{k' \in I_{h'}} \sum_{|x' - x| \leq 3} \mathbb{P}(A_{x,h}^{(k)}, A_{x',h'}^{(k')}). \end{aligned}$$

First, we estimate I. By the strong Markov property,

$$\begin{aligned} \mathbb{P}(A_{x,h}^{(k)}, A_{x',h'}^{(k')}) &= \sum_{\ell \in \mathcal{P}_{x,h}^{(k)}} \sum_{\ell' \in \mathcal{P}_{x',h'}^{(k')}} \mathbb{P}(B_x(\ell), B_{x'}(\ell')) \\ &= \sum_{\ell \in \mathcal{P}_{x,h}^{(k)}} \sum_{\tilde{\ell}: \ell + \tilde{\ell} \in \mathcal{P}_{x',h'}^{(k')}} \mathbb{P}^0(B_x(\ell)) \cdot \mathbb{P}^x(B_{x'}(\tilde{\ell})), \end{aligned}$$

where the x in \mathbb{P}^x indicates the starting point of the random walk. For any $x' \in \mathbb{Z}_+$ and $k' \in I_{h'}$, using Lemma 3.4, we get

$$\begin{aligned} &\sum_{k \in I_h} \sum_{x: |x-x'| > 3} \mathbb{P}(A_{x,h}^{(k)}, A_{x',h'}^{(k')}) \\ &= \sum_{k \in I_h} \sum_{x: |x-x'| > 3} \sum_{\ell \in \mathcal{P}_{x,h}^{(k)}} \sum_{\tilde{\ell}: \ell + \tilde{\ell} \in \mathcal{P}_{x',h'}^{(k')}} \mathbb{P}^0(B_x(\ell)) \cdot \mathbb{P}^x(B_{x'}(\tilde{\ell})) \\ &\leq \sum_{k \in I_h} \sum_{x: |x-x'| > 3} \sum_{\ell \in \mathcal{P}_{x,h}^{(k)}} \sum_{\tilde{\ell}: \ell + \tilde{\ell} \in \mathcal{P}_{x',h'}^{(k')}} O(1)h^{-1} \mathbb{P}^0(B_x(\varphi_x(\ell))) \cdot \mathbb{P}^x(B_{x'}(\tilde{\ell})) \\ &\leq O(1)h^{-1} \sum_{k \in I_h} \sum_{x: |x-x'| > 3} \sum_{\ell \in \mathcal{P}_{x,h}^{(k)}} \sum_{\ell^* \in \varphi_x(\ell)} \sum_{\tilde{\ell}: \ell + \tilde{\ell} \in \mathcal{P}_{x',h'}^{(k')}} \mathbb{P}(B_x(\ell^*), B_{x'}(\ell^* + \tilde{\ell})). \end{aligned}$$

The last inequality follows from Lemma 3.3 and strong Markov property. By Lemma 3.3, all events $B_x(\ell^*)$ for $x \in \mathbb{N}$, $\ell^* \in \varphi_x(\ell)$, $k \in I_h$ and $\ell \in \mathcal{P}_{x,h}^{(k)}$ are disjoint. Note that $|x - x'| > 3$ and φ_x only reduces the downcrossing number at $x + 1$, $x + 2$. So $\ell^* + \tilde{\ell} \in \mathcal{P}_{x',h'}^{(k')}$. Hence we have

$$\sum_{k \in I_h} \sum_{x: |x-x'| > 3} \mathbb{P}(A_{x,h}^{(k)}, A_{x',h'}^{(k')}) \leq O(1)h^{-1} \sum_{\ell' \in \mathcal{P}_{x',h'}^{(k')}} \mathbb{P}(B_{x'}(\ell')).$$

As a result, we obtain that

$$\begin{aligned} \text{(3.7)} \quad \text{I} &\leq O(1) \sum_{1 \leq h < h' \leq H} h^{-1} \sum_{k' \in I_{h'}} \sum_{x'=1}^{\infty} \mathbb{P}(A_{x',h'}^{(k')}) \\ &\leq O(1) \left(\sum_{h=1}^H h^{-1} \right) \left(\sum_{h'=1}^H \sum_{k' \in I_{h'}} \sum_{x'=1}^{\infty} \mathbb{P}(A_{x',h'}^{(k')}) \right) = O(1) \log H \cdot \mathbb{E}N_H. \end{aligned}$$

It remains to estimate II. In the case where the locations for favorite sites have overlap, we do have strong correlation between the two events. However, due to the overlap of locations for favorite sites the enumeration is hugely reduced. As a result, the contribution to the second moment in this case can also be controlled, as we show in what follows.

Since $A_{x',h'}^{(k_1)} \cap A_{x',h'}^{(k_2)} = \emptyset$ for $k_1 \neq k_2$, we have

$$\Pi \leq \sum_{x=1}^{\infty} \sum_{h=1}^H \sum_{k \in I_h} \sum_{h'=h+1}^H 7 \cdot \sup_{x':|x'-x| \leq 3} \mathbb{P}(A_{x,h}^{(k)}, \{\exists k' : A_{x',h'}^{(k')}\}).$$

Note $\mathbb{P}(A_{x,h}^{(k)}, \{\exists k' : A_{x',h'}^{(k')}\}) = \sum_{\ell \in \mathcal{P}_{x,h}^{(k)}} \mathbb{P}(B_x(\ell)) \cdot \mathbb{P}(\exists k' : A_{x',h'}^{(k')} | B_x(\ell))$. Conditioned on $B_x(\ell)$, in order for the event $\{\exists k' : A_{x',h'}^{(k')}\}$ to occur, we must have:

(1) There exists a $k' \geq \ell(x')$ such that at some time t , $S(t - 1) = x' - 1$, $S(t) = x'$ and $D(t, x' - 1) = k'$, $D(t, x') = h' - k' - 1$ (if such k' exists, it is unique).

(2) Once (1) happens, both t and k' are determined. The additional process after $B_x(\ell)$ need to satisfy: $D(t, x' + 1) - \ell(x' + 1) = h' - k' - 1 - \ell(x' + 1)$ and $D(t, x' + 2) - \ell(x' + 2) = k' + 1 - \ell(x' + 2)$.

(3) $L(t, y) < h'$ for all $y \neq x', x' + 1, x' + 2$.

We omit the probability loss for (1) and (3) and only consider the probability for (2). Formally, define T to be the time t such that $S(t - 1) = x' - 1$, $S(t) = x'$, $D(t, x' - 1) + D(t, x') = h' - 1$. Then we have $\mathbb{P}(\exists k' : A_{x',h'}^{(k')} | B_x(\ell))$ is less equal to

$$\sum_{k'=\ell(x')}^{h'} \mathbb{P}(T = T_U(k' + 1, x'), D(T, x') = h' - k' - 1, \\ D(T, x' + 1) = k' + 1, D(T, x' + 2) = h' - k' - 1).$$

Using the Ray–Knight representation for the random walk started at x after $B_x(\ell)$, we have $\mathbb{P}(\exists k' : A_{x',h'}^{(k')} | B_x(\ell))$ is less equal to

$$\sum_{k'=\ell(x')}^{h'} \mathbb{P}(T = T_U(k' + 1, x'), D(T, x') = h' - k' - 1) \\ \times \pi^*(h' - k' - 1 - \ell(x'), k' + 1 - \ell(x' + 1)) \\ \times \pi^*(k' + 1 - \ell(x' + 1), h' - k' - 1 - \ell(x' + 2)),$$

where $\pi^*(\cdot, \cdot)$ is either $\pi(\cdot, \cdot)$ or $\rho^*(\cdot, \cdot)$ depending on the relative position of x and x' [see (2.4) and (2.6)]. Since both $(h' - k' - 1 - \ell_x(x')) + (k' + 1 - \ell(x' + 1))$ and $(k' + 1 - \ell(x' + 1)) + (h' - k' - 1 - \ell(x' + 2))$ are greater than or equal to $h' - h$, by Lemma 2.2(ii) and the relation $\rho^*(i, j) = \pi(i, j - 1)$, we see that

$$\pi^*(h' - k' - 1 - \ell(x'), k' + 1 - \ell(x' + 1)) \cdot \pi^*(k' + 1 - \ell(x' + 1), \\ h' - k' - 1 - \ell(x' + 2))$$

is at most $\frac{O(1)}{h'-h}$ for any $\ell(x') \leq k' \leq h'$. Therefore,

$$\begin{aligned} &\mathbb{P}(\exists k' : A_{x',h'}^{(k')} | B_x(\ell)) \\ &\leq \sum_{k'=\ell(x')}^{h'} \mathbb{P}(T = T_U(k' + 1, x'), D(T, x') = h' - k' - 1) \cdot \frac{O(1)}{h' - h} \\ &= \mathbb{P}(\exists k' : T = T_U(k' + 1, x'), D(T, x') = h' - k' - 1) \cdot \frac{O(1)}{h' - h}, \end{aligned}$$

which is bounded by $\frac{O(1)}{h'-h}$. As a consequence, we get that

$$\begin{aligned} \mathbb{P}(A_{x,h}^{(k)}, \{\exists k' : A_{x',h'}^{(k')}\}) &\leq \sum_{\ell \in \mathcal{P}_{x,h}^{(k)}} \mathbb{P}(B_x(\ell)) \cdot \frac{O(1)}{h' - h} \\ &= \frac{O(1)}{h' - h} \cdot \mathbb{P}(A_{x,h}^{(k)}) \end{aligned}$$

and thus

$$(3.8) \quad \mathbb{II} \leq \sum_{x=1}^{\infty} \sum_{h=1}^H \sum_{k \in I_h} \sum_{h'=h+1}^H \frac{O(1)}{h' - h} \cdot \mathbb{P}(A_{x,h}^{(k)})$$

$$(3.9) \quad \leq O(\log H) \sum_{h=1}^H \sum_{k \in I_h} \sum_{x=1}^{\infty} \mathbb{P}(A_{x,h}^{(k)}) = O(\log H) \mathbb{E}N_H.$$

Combining (3.6), (3.7) and (3.8), we get that $\mathbb{E}N_H^2 = O(\log H) \mathbb{E}N_H$. \square

We are now ready to show that $N = \infty$ with positive probability.

PROPOSITION 3.6. *There exists a constant $\delta > 0$ such that $\mathbb{P}(N = \infty) \geq \delta$ where $N = \lim_{H \rightarrow \infty} N_H$.*

PROOF. By the Cauchy–Schwarz inequality, we get that

$$\begin{aligned} \mathbb{E}N_H &= \mathbb{E}N_H \mathbf{1}_{\{N_H > \log \log H\}} + \mathbb{E}N_H \mathbf{1}_{\{N_H \leq \log \log H\}} \\ &\leq \sqrt{\mathbb{E}N_H^2 \cdot \mathbb{P}(N_H > \log \log H)} + \log \log H. \end{aligned}$$

By Propositions 3.2 and 3.5, there exist constants $c, \delta > 0$ such that

$$\mathbb{P}(N_H > \log \log H) \geq \frac{(\mathbb{E}N_H - \log \log H)^2}{\mathbb{E}N_H^2} \geq c \frac{\mathbb{E}N_H \log H}{\mathbb{E}N_H^2} \geq \delta,$$

for all sufficiently large H . Sending $H \rightarrow \infty$, we get that $\mathbb{P}(N = \infty) \geq \delta$. \square

3.3. *0–1 law.* In this section, building on Proposition 3.6 we show that $N = \infty$ occurs with probability 1. There are a few possible approaches, and here we choose to prove a 0–1 law taking advantage of the result on the transience of favorite sites. Let $V(t)$ be an arbitrary element in $\mathcal{K}(t)$. It was shown in [3] that uniformly in all $V(t) \in \mathcal{K}(t)$ we have with probability 1

$$(3.10) \quad \liminf_{t \rightarrow \infty} \frac{|V(t)|}{t^{\frac{1}{2}}(\log t)^{-11}} = \infty.$$

Denote $\psi(t) = t^{\frac{1}{2}}(\log t)^{-11}$ and $E = \{\liminf_{t \rightarrow \infty} |V(t)| \geq \psi(t)\}$. By (3.10), we have $\mathbb{P}(E) = 1$, and thus without loss of generality we can assume that E occurs in what follows. Our goal is to show that the event $\{f(3) = \infty\}$ is a tail event and it suffices to show that the event $\{f(3) = \infty\}$ is independent of any σ -field F_m (which is the σ -field generated by the first m steps of the random walk) for all $m \in \mathbb{N}$. To this end, for each $m \in \mathbb{N}$ we let M be the first time such that for all $t \geq M$ favorite sites occurs outside of $[-2m, 2m]$. We see that M is not necessarily a stopping time but $M < \infty$ with probability 1. Therefore, the event $\{f(3) = \infty\}$ depends only on whether after M three favorite sites occurs infinitely often. Now consider the event $\{f_m(3) = \infty\}$ where $f_m(3)$ is defined analogously to $f(3)$ but for the random walk started at time m . We claim that the symmetric difference between $\{f(3) = \infty\}$ and $\{f_m(3) = \infty\}$ has probability zero since in the symmetric difference one must have a favorite site (for the original random walk) in the interval $[-2m, 2m]$ after M . Therefore, the event $\{f(3) = \infty\}$ is independent of F_m for all $m \in \mathbb{N}$, and thus is a tail event. By Kolmogorov's 0–1 law, $\mathbb{P}(f(3) = \infty) \in \{0, 1\}$. Combined with Proposition 3.6, it completes the proof of (1.2).

Acknowledgments. We thank Yueyun Hu and Zhan Shi for introducing the problem on favorite sites and for interesting discussions, and we thank Steve Lalley and Bálint Tóth for many helpful discussions and useful comments on an early version of the manuscript.

REFERENCES

- [1] ABE, Y. (2015). Maximum and minimum of local times for two-dimensional random walk. *Electron. Commun. Probab.* **20** 22, 14 pp. [MR3320410](#)
- [2] BASS, R. F., EISENBAUM, N. and SHI, Z. (2000). The most visited sites of symmetric stable processes. *Probab. Theory Related Fields* **116** 391–404. [MR1749281](#)
- [3] BASS, R. F. and GRIFFIN, P. S. (1985). The most visited site of Brownian motion and simple random walk. *Z. Wahrsch. Verw. Gebiete* **70** 417–436. [MR0803682](#)
- [4] BELIUS, D. (2013). Gumbel fluctuations for cover times in the discrete torus. *Probab. Theory Related Fields* **157** 635–689.
- [5] BELIUS, D. and KISTLER, N. (2017). The subleading order of two dimensional cover times. *Probab. Theory Related Fields* **167** 461–552. [MR3602852](#)
- [6] CHEN, D., DE RAPHÉLIS, L. and HU, Y. Favorite sites of randomly biased walks on a supercritical Galton–Watson tree. Available at [arXiv:1611.04497](#).

- [7] CSÁKI, E., RÉVÉSZ, P. and SHI, Z. (2000). Favourite sites, favourite values and jump sizes for random walk and Brownian motion. *Bernoulli* **6** 951–975.
- [8] CSÁKI, E. and SHI, Z. (1998). Large favourite sites of simple random walk and the Wiener process. *Electron. J. Probab.* **3** 14, 31 pp. (electronic).
- [9] DEMBO, A. (2005). Favorite points, cover times and fractals. In *Lectures on Probability Theory and Statistics. Lecture Notes in Math.* **1869** 1–101. Springer, Berlin.
- [10] DEMBO, A., PERES, Y., ROSEN, J. and ZEITOUNI, O. (2001). Thick points for planar Brownian motion and the Erdős–Taylor conjecture on random walk. *Acta Math.* **186** 239–270.
- [11] DEMBO, A., PERES, Y., ROSEN, J. and ZEITOUNI, O. (2004). Cover times for Brownian motion and random walks in two dimensions. *Ann. of Math. (2)* **160** 433–464.
- [12] EISENBAUM, N. (1997). On the most visited sites by a symmetric stable process. *Probab. Theory Related Fields* **107** 527–535.
- [13] EISENBAUM, N. and KHOSHNEVISAN, D. (2002). On the most visited sites of symmetric Markov processes. *Stochastic Process. Appl.* **101** 241–256. [MR1931268](#)
- [14] ERDŐS, P. and RÉVÉSZ, P. (1987). Problems and results on random walks. In *Mathematical Statistics and Probability Theory, Vol. B (Bad Tatzmannsdorf, 1986)* 59–65. Reidel, Dordrecht. [MR0922711](#)
- [15] ERDŐS, P. and RÉVÉSZ, P. (1984). On the favourite points of a random walk. In *Mathematical Structure—Computational Mathematics—Mathematical Modelling* **2** 152–157.
- [16] ERDŐS, P. and RÉVÉSZ, P. (1991). Three problems on the random walk in \mathbf{Z}^d . *Studia Sci. Math. Hungar.* **26** 309–320.
- [17] HU, Y. and SHI, Z. (2000). The problem of the most visited site in random environment. *Probab. Theory Related Fields* **116** 273–302. [MR1743773](#)
- [18] HU, Y. and SHI, Z. (2015). The most visited sites of biased random walks on trees. *Electron. J. Probab.* **20** 62, 14 pp.
- [19] KNIGHT, F. B. (1963). Random walks and a sojourn density process of Brownian motion. *Trans. Amer. Math. Soc.* **109** 56–86.
- [20] LIFSHITS, M. A. and SHI, Z. (2004). The escape rate of favorite sites of simple random walk and Brownian motion. *Ann. Probab.* **32** 129–152. [MR2040778](#)
- [21] MARCUS, M. B. (2001). The most visited sites of certain Lévy processes. *J. Theoret. Probab.* **14** 867–885. [MR1860527](#)
- [22] OKADA, I. Topics and problems on favorite sites of random walks. Available at [arXiv:1606.03787](#).
- [23] SHI, Z. and TÓTH, B. (2000). Favourite sites of simple random walk. *Period. Math. Hungar.* **41** 237–249. Endre Csáki 65.
- [24] TÓTH, B. (2001). No more than three favorite sites for simple random walk. *Ann. Probab.* **29** 484–503. [MR1825161](#)
- [25] TÓTH, B. and WERNER, W. (1997). Tied favourite edges for simple random walk. *Combin. Probab. Comput.* **6** 359–369.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
CHICAGO, ILLINOIS 60637
USA
E-MAIL: jjangding@galton.uchicago.edu
jfshen@galton.uchicago.edu