The asymptotic behavior of rarely visited edges of the simple random walk

Ze-Chun Hu[1], Xue Peng[1,*], Renming Song[2] and Yuan Tan[1]

[1]College of Mathematics, Sichuan University, Chengdu 610065, China

[2]Department of Mathematics, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

Abstract

In this paper, we study the asymptotic behavior of the number of rarely visited edges (i.e., edges that visited only once) of a simple symmetric random walk on $\mathbb{Z}$. Let $\alpha(n)$ be the number of rarely visited edges up to time n. First we evaluate $\mathbb{E}(\alpha(n))$, show that $n \to \mathbb{E}(\alpha(n))$ is non-decreasing in n and that $\lim_{n\to+\infty} \mathbb{E}(\alpha(n)) = 2$. Then we study the asymptotic behavior of $\mathbb{P}(\alpha(n) > a(\log n)^2)$ for any $a > 0$ and use it to show that there exists a constant $C \in (0, +\infty)$ such that $\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} = C$ almost surely.

Keywords and phrases: Random walk, local time, rarely visited edges

Mathematics Subject Classification (2020) 60F15, 60J55.

1 Introduction and the main results

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, $\{S_n\}_{n\geq 0}$ be a simple symmetric random walk on $\mathbb{Z}$ with $S_0 = 0$. Let $X_n := S_n - S_{n-1}, n \geq 1$. Then $\{X_n, n \geq 1\}$ are i.i.d. with $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}$.

For $y \in \mathbb{Z}$, we use $\xi(y, n) := \#\{0 \leq k \leq n : S_k = y\}$ to denote the time spent at y by $\{S_m\}_{m\geq 0}$ up to time n. Here and throughout this paper, $\#D$ denotes the cardinality of the set D. A site $x \in \mathbb{Z}$ is called a favorite (most visited) site of $\{S_m\}_{m\geq 0}$ up to time n if

$$\xi(x, n) = \max_{y\in\mathbb{Z}} \xi(y, n).$$

*Corresponding author.

E-mail addresses: zchu@scu.edu.cn (Z.C. Hu), pengxuemath@scu.edu.cn (X. Peng), rsong@illinois.edu (R. Song), 630769091@qq.com (Y. Tan).

For $y \in \mathbb{Z}$, we use $\langle y, y+1\rangle$ to denote the edge between the sites y and $y+1$. The numbers of upcrossings and downcrossings of $y \in \mathbb{Z}$ by $\{S_m\}_{m\geq 0}$ up to time $n \geq 1$ are defined by

$$L^U(y,n) := \#\{0 < k \leq n : S_k = y, S_{k-1} = y-1\},$$
$$L^D(y,n) := \#\{0 < k \leq n : S_k = y, S_{k-1} = y+1\}.$$

We set

$$L(y,n) := L^U(y+1,n) + L^D(y,n).$$

Then $L(y,n)$ is the number of times that $\{S_m\}_{m\geq 0}$ visits the edge $\langle y, y+1\rangle$ up to time n. An edge $\langle x, x+1\rangle$ is called a favorite edge of $\{S_m\}_{m\geq 0}$ up to time n if

$$L(x,n) = \sup_{y\in\mathbb{Z}} L(y,n).$$

The study of favorite sites of random walks was initiated by Erdös and Révész [5]. Since then, this topic has been intensively studied, see Bass [1], Bass and Griffin [2], Ding and Shen [3], Erdös and Révész [6, 7], Hao [8], Hao et al. [9, 10], Shi and Tóth [14], Toth [16], Toth and Werner [17] and the references therein.

A site $x \in \mathbb{Z}$ is called a rarely visited site of $\{S_m\}_{m\geq 0}$ up to time n if $\xi(x,n) = 1$. Compared to favorite sites, there are only a few papers on rarely visited sites, see Major [11], Newman [12] and Tóth [15]. Following Révész [13], we use $f_1(n)$ to denote the number of rarely visited sites up to time n, i.e.,

$$f_1(n) := \#\{x \in \mathbb{Z} : \xi(x,n) = 1\}.$$

Newman [12] proved that $\mathbb{E}(f_1(n)) = 2$, for all $n \geq 1$. Major [11] proved that there exists a constant $C \in (0,\infty)$ such that $\limsup_{n\to+\infty} \frac{f_1(n)}{(\log n)^2} = C$ almost surely.

An edge $\langle x, x+1\rangle$ is called a rarely visited edge of $\{S_m\}_{m\geq 0}$ up to time n if $L(x,n) = 1$. So far it seems that no one has studied rarely visited edges. The purpose of this paper is to study the asymptotic behavior of the number of rarely visited edges. Define

$$\mathcal{A}_n := \{\langle x, x+1\rangle | L(x,n) = 1\}, \quad \alpha(n) := \#\mathcal{A}_n,\ n \geq 1. \tag{1.1}$$

Then $\mathcal{A}_n$ is the collection of all the rarely visited edges of $\{S_m\}_{m\geq 0}$ up to time n, and $\alpha(n)$ is the number of rarely visited edges of $\{S_m\}_{m\geq 0}$ up to time n. The main results in our paper are as follows:

Theorem 1.1. *(i)* $\mathbb{E}(\alpha(1)) = 1$ *and for all* $n \geq 1$,

$$\begin{cases} \mathbb{E}(\alpha(n+1)) &= \mathbb{E}(\alpha(n)), & \textit{if } n \textit{ is odd}, \\ \mathbb{E}(\alpha(n+1)) &= \mathbb{E}(\alpha(n)) + 2\cdot\frac{(n-1)!!}{(n+2)!!}, & \textit{if } n \textit{ is even}. \end{cases} \tag{1.2}$$

(ii) $\lim_{n\to+\infty} \mathbb{E}(\alpha(n)) = 2$.

Theorem 1.2. *For all* $a > 0$ *and* $\varepsilon > 0$, *there exists an* $N_0 = N_0(a,\varepsilon)$ *such that for all* $n > N_0$,

$$n^{-2a-\varepsilon} < \mathbb{P}\left(\alpha(n) > a(\log n)^2\right) < n^{-2a+\varepsilon}.$$

Theorem 1.3. *There exists a constant $C \in (0, \infty)$ such that*

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} = C\right) = 1.$$

From Theorem 1.1, we can see that, unlike the result that the expected number $\mathbb{E}(f_1(n))$ of rarely visited sites is equal to 2 for all $n \geq 1$, the expected number of rarely visited edges $\mathbb{E}(\alpha(n))$ increases with n and $\lim_{n\to+\infty} \mathbb{E}(\alpha(n)) = 2$. Theorem 1.2 and Theorem 1.3 imply that the asymptotic behavior of rarely visited edges is similar to that of rarely visited sites.

Remark 1.4. *Related to the results above, we think the following open problems are worth studying:*

(1) What is the exact value of the constant C in Theorem 1.3?

(2) If $\{S_n\}_{n\geq 0}$ is an asymmetric simple random walk, i.e., $P(X_1 = 1) \neq P(X_1 = -1)$, what is the asymptotic behavior of rarely visited edges?

The rest of the paper is organized as follows. In Section 2, we give the proof of Theorem 1.1. In Section 3, the proofs of Theorems 1.2 and 1.3 will be given.

2 Proof of Theorem 1.1

Without loss of generality, for the proof of Theorem 1.1, we can assume that

$$\Omega := \{\omega = (\omega_0, \omega_1, \omega_2, \dots) : \omega_n \in \mathbb{Z},\ \forall n \geq 0,\ \text{and } |\omega_n - \omega_{n-1}| = 1,\ \forall n \geq 1\}.$$

Let $\mathscr{F}$ be the σ-field on Ω generated by cylinder sets. For $n \geq 0$, $x \in \mathbb{Z}$ and $x_0, x_1, \dots, x_n \in \mathbb{Z}$ satisfying $|x_k - x_{k-1}| = 1$ for all $k = 1, \dots, n$, we define a probability measure $\mathbb{P}_x$ on $(\Omega, \mathscr{F})$ by

$$\mathbb{P}_x(\omega : \omega_0 = x_0, \omega_1 = x_1, , \dots, \omega_n = x_n) := \frac{1}{2^n}\delta_x(x_0).$$

Let

$$S_n(\omega) := \omega_n,\ \forall n \geq 0, \quad X_0 := S_0,\ X_n = S_n - S_{n-1},\ \forall n \geq 1.$$

Then under $\mathbb{P}_x$, $\{S_n\}_{n\geq 0}$ is a simple symmetric random walk on $\mathbb{Z}$ with $S_0 = x$, and $\{X_n\}_{n\geq 1}$ are i.i.d. random variables with

$$\mathbb{P}_x(X_1 = 1) = \mathbb{P}_x(X_1 = -1) = \frac{1}{2}.$$

$\mathbb{P}_0$ is the probability measure $\mathbb{P}$ of Section 1. We will use $\mathbb{E}_x$ to denote the expectation with respect to $\mathbb{P}_x$.

Proof of Theorem 1.1. (i) Obviously, we have $\mathbb{E}_0(\alpha(1)) = 1$.

Let $\tilde{\alpha}(n)$ be the number of rarely visited edges of the random walk $\{S_k, 1 \le k \le n+1\}$. Since $X_1, X_2, \ldots, X_{n+1}$ are i.i.d., we have

$$\begin{aligned}\mathbb{E}_0(\tilde{\alpha}(n)) &= \mathbb{P}_0(X_1 = 1)\mathbb{E}_0(\tilde{\alpha}_1(n)|X_1 = 1) + \mathbb{P}_0(X_1 = -1)\mathbb{E}_0(\tilde{\alpha}(n)|X_1 = -1) \\ &= \frac{1}{2}\mathbb{E}_1(\alpha(n)) + \frac{1}{2}\mathbb{E}_{-1}(\alpha(n)) = \frac{1}{2}\mathbb{E}_0(\alpha(n)) + \frac{1}{2}\mathbb{E}_0(\alpha(n)) = \mathbb{E}_0(\alpha(n)).\end{aligned}$$

Thus,

$$\begin{aligned}&\mathbb{E}_0(\alpha(n+1)) - \mathbb{E}_0(\alpha(n)) \\ =&[\mathbb{E}_0(\alpha(n+1); X_1 = 1) + \mathbb{E}_0(\alpha(n+1); X_1 = -1)] - \mathbb{E}_0(\alpha(n)) \\ =&[\mathbb{E}_0(\alpha(n+1); L(0, n+1) = 1, X_1 = 1) + \mathbb{E}_0(\alpha(n+1); L(0, n+1) = 2, X_1 = 1) \\ &+\mathbb{E}_0(\alpha(n+1); L(0, n+1) \ge 3, X_1 = 1)] \\ &+ [\mathbb{E}_0(\alpha(n+1); L(-1, n+1) = 1, X_1 = -1) + \mathbb{E}_0(\alpha(n+1); L(-1, n+1) = 2, X_1 = -1) \\ &+\mathbb{E}_0(\alpha(n+1); L(-1, n+1) \ge 3, X_1 = -1)] - \mathbb{E}_0(\alpha(n)) \\ =&[\mathbb{E}_0(\tilde{\alpha}(n) + 1; L(0, n+1) = 1, X_1 = 1) + \mathbb{E}_0(\tilde{\alpha}(n) - 1; L(0, n+1) = 2, X_1 = 1) \\ &+\mathbb{E}_0(\tilde{\alpha}(n); L(0, n+1) \ge 3, X_1 = 1)] \\ &+ [\mathbb{E}_0(\tilde{\alpha}(n) + 1; L(-1, n+1) = 1, X_1 = -1) + \mathbb{E}_0(\tilde{\alpha}(n) - 1; L(-1, n+1) = 2, X_1 = -1) \\ &+\mathbb{E}_0(\tilde{\alpha}(n); L(-1, n+1) \ge 3, X_1 = -1)] - \mathbb{E}_0(\alpha(n)) \\ =&\mathbb{E}_0(\tilde{\alpha}(n)) + [P_0(L(0, n+1) = 1, X_1 = 1) - \mathbb{P}_0(L(0, n+1) = 2, X_1 = 1)] \\ &+ [\mathbb{P}_0(L(-1, n+1) = 1, X_1 = -1) - \mathbb{P}_0(L(-1, n+1) = 2, X_1 = -1)] - \mathbb{E}_0(\alpha(n)) \\ =&[\mathbb{P}_0(L(0, n+1) = 1, X_1 = 1) - \mathbb{P}_0(L(0, n+1) = 2, X_1 = 1)] \\ &+ [\mathbb{P}_0(L(-1, n+1) = 1, X_1 = -1) - \mathbb{P}_0(L(-1, n+1) = 2, X_1 = -1)]. \qquad (2.1)\end{aligned}$$

For $\omega \in \Omega$, define

$$\sigma(\omega) := \sup\{0 < k \le n+1, S_k(\omega) = 1\}. \qquad (2.2)$$

Then

$$\begin{aligned}\mathbb{P}_0(L(0, n+1) = 1, X_1 = 1) =&\mathbb{P}_0(L(0, n+1) = 1, X_1 = 1, \sigma < n+1) \\ &+ \mathbb{P}_0(L(0, n+1) = 1, X_1 = 1, \sigma = n+1). \qquad (2.3)\end{aligned}$$

We deal with $\mathbb{P}_0(L(0, n+1) = 1, X_1 = 1, \sigma < n+1)$ first. For any $\omega \in \{L(0, n+1) = 1, X_1 = 1, \sigma \le n+1\}$, we define $\omega' \in \Omega$ by

$$\omega'_k := \begin{cases} \omega_k, & \text{if} \quad 0 \le k \le \sigma(\omega), \\ 2 - \omega_k, & \text{if} \quad k > \sigma(\omega). \end{cases} \qquad (2.4)$$

For any $\omega \in \{L(0, n+1) = 1, X_1 = 1, S_0 = 0, \sigma < n+1\}$, we have $1 \le \sigma(\omega) < n+1$, $S_{\sigma(\omega)+1}(\omega) = 2$ and $S_k(\omega) \ge 1$ for all $1 \le k \le \sigma(\omega)$. So $S_{\sigma(\omega)+1}(\omega') = 0$. Thus $L(0, n+1, \omega') \ge 2$. Suppose that $L(0, n+1, \omega') \ge 3$. Since $S_k(\omega') = S_k(\omega) \ge 1$ for all $1 \le k \le \sigma(\omega)$, there exists m, $\sigma(\omega) + 2 \le m \le n+1$, such that $S_{m-1}(\omega') = 0, S_m(\omega') = 1$. Then by (2.4) we get $S_{m-1}(\omega) = 2, S_m(\omega) = 1$, which contradicts the definition of $\sigma(\omega)$ in (2.2). Hence, $L(0, n+1, \omega') = 2$. Thus,

$\omega' \in \{L(0,n+1)=2, X_1=1, S_0=0\}$. It is easy to see that the map defined in (2.4) is an injection on the set $\{L(0,n+1)=1, X_1=1, S_0=0, \sigma<n+1\}$, so it is an injection from the set $\{L(0,n+1)=1, X_1=1, S_0=0, \sigma<n+1\}$ to the set $\{L(0,n+1)=2, X_1=1, S_0=0\}$.

For any $\omega \in \{L(0,n+1)=2, S_0=0, X_1=1\}$, we have $1 \le \sigma(\omega) \le n+1$ and there exists a unique $\tilde\sigma(\omega)$ such that $n+1 \ge \tilde\sigma(\omega) \ge 2$ and

$$S_{\tilde\sigma(\omega)-1}(\omega)=1,\ \ S_{\tilde\sigma(\omega)}(\omega)=0. \tag{2.5}$$

Obviously, $\sigma(\omega) \ne \tilde\sigma(\omega)$. If $\tilde\sigma(\omega) < \sigma(\omega)$, then using $S_0(\omega)=0, S_1(\omega)=1, S_{\tilde\sigma(\omega)-1}(\omega)=1, S_{\tilde\sigma(\omega)}(\omega)=0$ and $S_{\sigma(\omega)}(\omega)=1$, we get $L(0,n+1,\omega)\ge 3$, which contradicts $L(0,n+1)=2$. Hence, $n+1 \ge \tilde\sigma(\omega) > \sigma(\omega)$. If $S_{\sigma(\omega)+1}(\omega)=2$, then, since $S_{\tilde\sigma(\omega)}(\omega)=0$, there exists m such that $\sigma(\omega)+2 \le m < \tilde\sigma(\omega)$ and

$$S_{m-1}(\omega)=2,\ \ S_m(\omega)=1,$$

which contradicts the definition of $\sigma(\omega)$ in (2.2). Thus $S_{\sigma(\omega)+1}(\omega)=0$ and $\tilde\sigma(\omega)=\sigma(\omega)+1$. Therefore $S_k(\omega)\ge 1$ for all $1\le k\le\sigma(\omega)$, and $S_k(\omega)\le 0$ for all $\sigma(\omega)+1\le k\le n+1$. Note that $S_0(\omega')=0$, $S_k(\omega')=S_k(\omega)\ge 1$ for all $1\le k\le\sigma(\omega)$, and $S_k(\omega')=2-S_k(\omega)\ge 2$ for all $\sigma(\omega)+1\le k\le n+1$. Thus $L(0,n+1,\omega')=1$ and $\sigma(\omega')=\sigma(\omega)<n+1$. It easy to see that the map defined in (2.4) is an injection on the set $\{L(0,n+1)=2, S_0=0, X_1=1\}$, so it is an injection from the set $\{L(0,n+1)=2, S_0=0, X_1=1\}$ to the set $\{L(0,n+1)=1, S_0=0, X_1=1, \sigma<n+1\}$.

Therefore we have shown that there is a one-to-one correspondence between the sets $\{L(0,n+1)=2, S_0=0, X_1=1\}$ and $\{L(0,n+1)=1, S_0=0, X_1=1, \sigma<n+1\}$. Thus

$$\mathbb{P}_0(L(0,n+1)=1, X_1=1, \sigma<n+1)=\mathbb{P}_0(L(0,n+1)=2, X_1=1). \tag{2.6}$$

Now we deal with $\mathbb{P}_0(L(0,n+1)=1, X_1=1, \sigma=n+1)$. Note that

$$\begin{aligned}
&\mathbb{P}_0(L(0,n+1)=1, X_1=1, \sigma=n+1)\\
&=\mathbb{P}_0(L(0,n+1)=1, X_1=1, S_n=2, S_{n+1}=1)\\
&=\mathbb{P}_0(X_1=1)\mathbb{P}_0(S_j\ge 1,\ 1\le j\le n+1, S_{n+1}=1|X_1=1)\\
&=\frac{1}{2}\mathbb{P}_0(S_j\ge 0,\ 0\le j\le n, S_n=0).
\end{aligned} \tag{2.7}$$

Combining (2.3), (2.6) and (2.7), we get

$$\begin{aligned}
&\mathbb{P}_0(L(0,n+1)=1, X_1=1)\\
=&\mathbb{P}_0(L(0,n+1)=2, X_1=1)+\frac{1}{2}\mathbb{P}_0(S_j\ge 0, 0\le j\le n, S_n=0).
\end{aligned} \tag{2.8}$$

Similarly, by the symmetry of $\{S_m\}_{m\ge 0}$, we have

$$\begin{aligned}
&\mathbb{P}_0(L(-1,n+1)=1, X_1=-1)\\
=&\mathbb{P}_0(L(-1,n+1)=2, X_1=-1)+\frac{1}{2}\mathbb{P}_0(S_j\le 0, 0\le j\le n, S_n=0)
\end{aligned}$$

$$=\mathbb{P}_0(L(-1,n+1)=2, X_1=-1)+\frac{1}{2}\mathbb{P}_0(S_j\geq 0, 0\leq j\leq n, S_n=0). \tag{2.9}$$

Hence, by (2.1), (2.8) and (2.9), we have

$$\mathbb{E}_0(\alpha(n+1))-\mathbb{E}_0(\alpha(n))=\mathbb{P}_0(S_j\geq 0, 0\leq j\leq n, S_n=0). \tag{2.10}$$

When n is odd, we have

$$\mathbb{P}_0(S_j\geq 0,\ 0\leq j\leq n, S_n=0)=0. \tag{2.11}$$

When n is even, we will use the reflection principle to deal with $P_0(S_j\geq 0,\ 0\leq j\leq n, S_n=0)$. For $x\in\mathbb{Z}$, let

$$\sigma_x:=\inf\{n\geq 0, S_n=x\}. \tag{2.12}$$

For $n\geq 2$, we have

$$\mathbb{P}_0(S_n=0)=\mathbb{P}_0(S_n=0,\sigma_{-1}>n)+\mathbb{P}_0(S_n=0,\sigma_{-1}\leq n). \tag{2.13}$$

For $\omega\in\Omega$, we define $\omega''\in\Omega$ by

$$\omega''_k=\begin{cases} -2-\omega_k, & 0\leq k\leq \sigma_{-1}(\omega),\\ \omega_k, & k>\sigma_{-1}(\omega).\end{cases} \tag{2.14}$$

It is easy to see that the map defined in (2.14) gives a one-to-one correspondence between the set $\{S_0=0, S_n=0,\sigma_{-1}\leq n\}$ and $\{S_0=-2, S_n=0\}$. Hence,

$$\mathbb{P}_0(S_n=0,\sigma_{-1}\leq n)=\mathbb{P}_{-2}(S_n=0). \tag{2.15}$$

Thus by (2.13) and (2.15) we have

$$\begin{aligned}&\mathbb{P}_0(S_j\geq 0,\ 0\leq j\leq n, S_n=0)=\mathbb{P}_0(S_n=0,\sigma_{-1}>n)\\ =&\mathbb{P}_0(S_n=0)-\mathbb{P}_0(S_n=0,\sigma_{-1}\leq n)=\mathbb{P}_0(S_n=0)-\mathbb{P}_{-2}(S_n=0)\\ =&\binom{n}{\frac{n}{2}}\frac{1}{2^n}-\binom{n}{\frac{n+2}{2}}\frac{1}{2^n}=2\cdot\frac{(n-1)!!}{(n+2)!!}.\end{aligned} \tag{2.16}$$

Hence, by (2.10), (2.11) and (2.16), we obtain

$$\begin{aligned}&\mathbb{E}(\alpha(n+1))-\mathbb{E}(\alpha(n))=\mathbb{E}_0(\alpha(n+1))-\mathbb{E}_0(\alpha(n))\\ =&\mathbb{P}_0(S_j\geq 0, 0\leq j\leq n, S_n=0)=\begin{cases} 0, & \text{if } n \text{ is odd},\\ 2\cdot\frac{(n-1)!!}{(n+2)!!}, & \text{if } n \text{ is even}.\end{cases}\end{aligned}$$

(ii) By (i) and the Taylor expansion of $\sqrt{1-x}, x\in[-1,1]$, we get

$$\begin{aligned}\lim_{n\to+\infty}\mathbb{E}(\alpha(n+1))&=\mathbb{E}(\alpha(1))+\lim_{n\to+\infty}\sum_{k=1}^{n}[\mathbb{E}(\alpha(k+1))-E(\alpha(k))]\\ &=1+2\sum_{k=1}^{+\infty}\frac{(2k-1)!!}{(2k+2)!!}=2.\end{aligned}$$

□

3 Proofs of Theorems 1.2 and 1.3

Our proofs of Theorems 1.2 and 1.3 are inspired by Major [11]. First we use Kolmogorov's 0-1 law to show that there exists a constant $C \in [0, \infty]$ such that $\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} = C$ almost surely. The proof of this result is routine. However, it is not easy to show that $C \in (0, +\infty)$, which is the assertion of Theorem 1.3.

Proposition 3.1. *If $\{f(n)\}_{n\geq 1}$ satisfies $0 < f(n) < n$ and $\lim_{n\to+\infty} f(n) = +\infty$, then there exists $C \in [0, \infty]$ such that $\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{f(n)} = C\right) = 1$.*

Proof. Let $\alpha'(n)$ be the number of rarely visited edges of $\{S_k, k \in [\sqrt{f(n)}, n]\}$. Then $|\alpha(n) - \alpha'(n)| \leq \sqrt{f(n)}$. Thus

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{f(n)} = \limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)}\right) = 1. \tag{3.1}$$

Noticing that for any $c \in [0, \infty]$, $\left\{\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \geq c\right\}$ is a tail event, by Kolmogorov's 0-1 law, we get that, for any $c \in [0, \infty]$,

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \geq c\right) \in \{0, 1\}.$$

Note that $\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \geq 0\right) = 1$ and that $\left\{\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \geq c\right\}$ decreases as c increases. Define

$$c^* := \sup\left\{c \geq 0 : \mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \geq c\right) = 1\right\}.$$

If $0 \leq c^* < \infty$, we can choose a decreasing sequence $\{c_m\}_{m\geq 1}$ such that $c_m > c^*$ and $\lim_{m\to+\infty} c_m = c^*$. Then we have

$$\begin{aligned}
\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} > c^*\right) &= \mathbb{P}\left(\lim_{m\to+\infty} \left\{\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \geq c_m\right\}\right) \\
&= \lim_{m\to+\infty} \mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \geq c_m\right) = 0,
\end{aligned}$$

which implies that

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} \leq c^*\right) = 1. \tag{3.2}$$

In particular, if $c^* = 0$, we have

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha'(n)}{f(n)} = 0\right) = 1. \tag{3.3}$$

If $0 < c^* \le +\infty$, we can choose an increasing sequence $\{c_m\}_{m\ge1}$ such that $c_m < c^*$ and $\lim_{m\to+\infty} c_m = c^*$. Then

$$\begin{aligned}\mathbb{P}\left(\limsup_{n\to+\infty}\frac{\alpha'(n)}{f(n)}\ge c^*\right)&=\mathbb{P}\left(\lim_{m\to+\infty}\left\{\limsup_{n\to+\infty}\frac{\alpha'(n)}{f(n)}\ge c_m\right\}\right)\\&=\lim_{m\to+\infty}\mathbb{P}\left(\limsup_{n\to+\infty}\frac{\alpha'(n)}{f(n)}\ge c_m\right)=1.\end{aligned}\tag{3.4}$$

In particular, if $c^* = +\infty$, we have

$$\mathbb{P}\left(\limsup_{n\to+\infty}\frac{\alpha'(n)}{f(n)}=+\infty\right)=1.\tag{3.5}$$

By (3.2) and (3.4), we know that if $0 < c^* < \infty$, it holds

$$\mathbb{P}\left(\limsup_{n\to+\infty}\frac{\alpha'(n)}{f(n)}=c^*\right)=1.\tag{3.6}$$

Combining (3.1), (3.3), (3.5) and (3.6), we always have

$$\mathbb{P}\left(\limsup_{n\to+\infty}\frac{\alpha(n)}{f(n)}=c^*\right)=\mathbb{P}\left(\limsup_{n\to+\infty}\frac{\alpha'(n)}{f(n)}=c^*\right)=1,$$

where $0 \le c^* \le +\infty$. □

3.1 Some preparations

It follows from [11, Lemma 3, Remark 6] that

$$\begin{aligned}&\lim_{n\to+\infty} n\mathbb{P}\left(S_j>0 \text{ for all } 0<j\le n \text{ and } S_j<S_n \text{ for all } 0\le j<n\right)\\&=\lim_{n\to+\infty} n\mathbb{P}(0<S_j<S_n, \text{ for all } 0<j<n)=\frac{1}{4}.\end{aligned}\tag{3.7}$$

It is well known (see, for instance, [4, Lemma 4.9.3]) that

$$\mathbb{P}(S_1\ne0,\dots,S_{2n}\ne0)=\mathbb{P}(S_{2n}=0).$$

By symmetry, we have

$$\mathbb{P}(S_1>0,\dots,S_{2n}>0)=\frac{1}{2}\mathbb{P}(S_{2n}=0).\tag{3.8}$$

Since

$$\begin{aligned}\mathbb{P}(S_1\ge0,\dots,S_{2n}\ge0)&=\mathbb{P}(S_1\ge0,\dots,S_{2n-1}\ge0)\\&=2\mathbb{P}(S_1=1,S_2-S_1\ge0,\dots,S_{2n}-S_1\ge0)\end{aligned}$$

$$= 2\mathbb{P}(S_1 > 0, \ldots, S_{2n} > 0),$$

we have

$$\mathbb{P}(S_1 \geq 0, \ldots, S_{2n} \geq 0) = \mathbb{P}(S_{2n} = 0). \tag{3.9}$$

For $k \geq 0$, let $\alpha_k(n)$ denote the number of subsets of $\mathcal{A}_n$ with exactly k elements. Then $\alpha_k(n) = \binom{\alpha(n)}{k}$ for $k \leq \alpha(n)$ and $\alpha_k(n) = 0$ for $k > \alpha(n)$. The following lemma plays a key role in the proof of Theorem 1.2.

Lemma 3.2. *Let $a > 0$. If $k \sim a \log n$ as $n \to +\infty$, then for any $\varepsilon \in (0, 1/2)$, there exists $n_0 = n_0(a, \varepsilon)$ such that for all $n \geq n_0$,*

$$[(\frac{1}{2} - \varepsilon) \log n]^k < \mathbb{E}\, \alpha_k(n) < [(\frac{1}{2} + \varepsilon) \log n]^k.$$

For non-negative integers r and t, we define

$$\begin{aligned}
C_1(t) &:= \{0 < S_l < S_t \text{ for all } 0 < l < t\},\ 0 < t \leq n;\\
C_2(r,t) &:= \{S_r < S_l \leq S_t \text{ for all } r < l \leq t\},\ 0 \leq r < t \leq n;\\
C_2(t) &:= \{0 < S_l \leq S_t \text{ for all } 0 < l \leq t\},\ 0 < t \leq n;\\
D_1(t) &:= \{S_l \leq S_t \text{ for all } 0 \leq l \leq t\},\ 0 \leq t \leq n;\\
D_2(r,t) &:= \{S_r < S_l \text{ for all } r < l \leq t\},\ 0 \leq r < t \leq n.
\end{aligned}$$

Proposition 3.3. *(i)* $\lim\limits_{t \to +\infty} t\mathbb{P}(C_2(t)) = \frac{1}{2}$*;*

(ii) $\lim\limits_{t \to +\infty} \sqrt{t}\mathbb{P}(D_1(t)) = \sqrt{2/\pi}$*;*

(iii) $\lim\limits_{t-r \to +\infty} \sqrt{t-r}\mathbb{P}(D_2(r,t)) = 1/\sqrt{2\pi}$*.*

Proof. (i) Since

$$\begin{aligned}
\{X_{t+1} = 1\} \cap C_2(t) &= \{X_{t+1} = 1, 0 < S_l \leq S_t \text{ for all } 0 < l \leq t\}\\
&= \{0 < S_l \leq S_t < S_{t+1}, \text{for all } 0 < l \leq t\}\\
&= \{0 < S_l < S_{t+1} \text{ for all } 0 < l < t+1\} = C_1(t+1),
\end{aligned}$$

we have

$$\mathbb{P}(C_1(t+1)) = \mathbb{P}(\{X_{t+1} = 1\} \cap C_2(t)) = \mathbb{P}(X_{t+1} = 1)\mathbb{P}(C_2(t)) = \frac{1}{2}\mathbb{P}(C_2(t)).$$

Thus by (3.7), we have

$$\lim_{t \to +\infty} t\mathbb{P}(C_2(t)) = \lim_{t \to +\infty} 2t\mathbb{P}(C_1(t+1)) = \frac{1}{2}.$$

(ii) Let $\check{S}_k^t := S_t - S_{t-k}$, $k = 0, 1, \ldots, t$. Then $\{\check{S}_k^t\}_{0\le k\le t}$ is a simple symmetric random walk with $\check{S}_0^t = 0$. Thus

$$\mathbb{P}(D_1(t)) = \mathbb{P}\left(\check{S}_0^t \ge 0, \check{S}_1^t \ge 0, \ldots, \check{S}_t^t \ge 0\right) = \mathbb{P}(S_{2m} = 0) = \binom{2m}{m}\frac{1}{2^{2m}},$$

where $m = t/2$ or $m = (t+1)/2$ and we used (3.9) in the second equality above. Thus, by Stirling's formula, we have

$$\lim_{t\to+\infty} \sqrt{t}\mathbb{P}(D_1(t)) = \sqrt{2/\pi}.$$

(iii) Let $\widehat{S}_k^r = S_{r+k} - S_r$, $k = 0, 1, 2, \ldots, t-r$. Then $\{\widehat{S}_k^r\}_{0\le k\le t-r}$ is a simple symmetric random walk with $\widehat{S}_0^r = 0$. Thus

$$\mathbb{P}(D_2(r,t)) = \mathbb{P}(\widehat{S}_1 > 0, \widehat{S}_2 > 0, \ldots, \widehat{S}_{t-r} > 0) = \frac{1}{2}\mathbb{P}(S_{2m} = 0) = \frac{1}{2}\cdot\binom{2m}{m}\frac{1}{2^{2m}},$$

where $m = (t-r)/2$ or $m = (t-r-1)/2$ and we used (3.8) in the second equaility above. Hence, by Stirling's formula, we have

$$\lim_{t-r\to+\infty} \sqrt{t-r}\mathbb{P}(D_2(r,t)) = 1/\sqrt{2\pi}.$$

□

We define

$$\mathcal{A}_n^+ := \{z \ge 0 | \langle z, z+1\rangle \in \mathcal{A}_n\}, \quad \mathcal{A}_n^- := \{z \le 0 | \langle z-1, z\rangle \in \mathcal{A}_n\}. \tag{3.10}$$

Then there is a one-to-one correspondence between $\mathcal{A}_n^+$ and the collection of rarely visited edges, on the positive half-axis, of $\{S_m\}_{m\ge 0}$ up to time n. There is also a one-to-one correspondence between $\mathcal{A}_n^-$ and the collection of rarely visited edges, on the negative half-axis, of $\{S_m\}_{m\ge 0}$ up to time n. Let $\alpha^+(n) := \#\mathcal{A}_n^+, \alpha^-(n) := \#\mathcal{A}_n^-$. For $k \ge 0$, let $\alpha_k^+(n)$ be the number of subsets of $\mathcal{A}_n^+$ with exactly k elements. Then $\alpha_k^+(n) = \binom{\alpha^+(n)}{k}$ for $k \le \alpha^+(n)$ and $\alpha_k^+(n) = 0$ for $k > \alpha^+(n)$. We define $\alpha_k^-(n)$ similarly.

Proof of Lemma 3.2. For $k \ge 2$, it holds that

$$\alpha_k^+(n)\mathbf{1}_{\{k\le\alpha^+(n)\}} = \sum_{0\le j_1<\cdots<j_k\le n-1} \mathbf{1}_{D_1(j_1)C_2(j_1,j_2)C_2(j_2,j_3)\cdots C_2(j_{k-1},j_k)D_2(j_k,n)},$$

where $\mathbf{1}_A(\cdot)$ is the indicator function. Hence,

$$\begin{aligned}\mathbb{E}\alpha_k^+(n) &= \sum_{0\le j_1<\cdots<j_k\le n-1} \mathbb{P}(D_1(j_1)C_2(j_1,j_2)C_2(j_2,j_3)\cdots C_2(j_{k-1},j_k)D_2(j_k,n)) \\ &= \sum_{0\le j_1<\cdots<j_k\le n-1} \mathbb{P}(D_1(j_1))\mathbb{P}(C_2(j_1,j_2))\mathbb{P}(C_2(j_2,j_3))\cdots\mathbb{P}(C_2(j_{k-1},j_k))\mathbb{P}(D_2(j_k,n))\end{aligned}$$

$$= \sum_{0\le j_1<\cdots<j_k\le n-1} \mathbb{P}(D_1(j_1))\mathbb{P}(C_2(j_2-j_1))\mathbb{P}(C_2(j_3-j_2))\cdots\mathbb{P}(C_2(j_k-j_{k-1}))\mathbb{P}(D_2(j_k,n)). \tag{3.11}$$

Let $j = j_1, r = j_k - j_1, y_i = j_{i+1} - j_i, 1 \le i \le k-1$. Then we have

$$\mathbb{E}\alpha_k^+(n) = \sum_{r=k-1}^{n-1} \left[\sum_{j=0}^{n-1-r} \mathbb{P}(D_1(j))\mathbb{P}(D_2(j+r,n))\right] \left[\sum_{\substack{0<y_i<r \\ y_1+y_2+\cdots+y_{k-1}=r}} \mathbb{P}(C_2(y_1))\mathbb{P}(C_2(y_2))\cdots\mathbb{P}(C_2(y_{k-1}))\right]. \tag{3.12}$$

It follows from Proposition 3.3 that there exists a positive constant c_1 such that for all integers $n, r \ge 1$ and $j \ge 0$ with $n - j - r \ge 1$,

$$\sqrt{j}\mathbb{P}(D_1(j)) \le c_1, \quad \sqrt{n-j-r}\mathbb{P}(D_2(j+r,n)) \le c_1.$$

Thus

$$\begin{aligned}
&\sum_{j=0}^{n-1-r} \mathbb{P}(D_1(j))\mathbb{P}(D_2(j+r,n)) \\
&= \mathbb{P}(D_2(r,n)) + \sum_{j=1}^{n-1-r} \mathbb{P}(D_1(j))\mathbb{P}(D_2(j+r,n)) \\
&\le c\left(1 + \sum_{j=1}^{n-1-r} \frac{1}{\sqrt{j(n-r-j)}}\right) \\
&= c\left(1 + \sum_{j=1}^{n-1-r} \frac{1}{\sqrt{\frac{j}{n-r}}\sqrt{1-\frac{j}{n-r}}} \cdot \frac{1}{n-r}\right),
\end{aligned} \tag{3.13}$$

where $c = \max\{1, c_1^2\}$. Since

$$\lim_{n-r\to+\infty} \sum_{j=1}^{n-1-r} \frac{1}{\sqrt{\frac{j}{n-r}}\sqrt{1-\frac{j}{n-r}}} \cdot \frac{1}{n-r} = \int_0^1 x^{-1/2}(1-x)^{-1/2}dx = \pi,$$

we know that

$$\sum_{j=1}^{n-1-r} \frac{1}{\sqrt{\frac{j}{n-r}}\sqrt{1-\frac{j}{n-r}}} \cdot \frac{1}{n-r}, \quad n-r \ge 1$$

is bounded. Thus by (3.13), there exists a positive constant C such that for all integers $n, r \ge 1$ with $n - r \ge 1$,

$$\sum_{j=0}^{n-1-r} \mathbb{P}(D_1(j))\mathbb{P}(D_2(j+r,n)) \le C. \tag{3.14}$$

Hence, by (3.12) and (3.14), we have

$$\begin{aligned}\mathbb{E}\alpha_k^+(n) \leq & C \sum_{r=k-1}^{n-1} \sum_{\substack{0<y_i<r \\ y_1+y_2+\cdots+y_{k-1}=r}} \mathbb{P}(C_2(y_1))\mathbb{P}(C_2(y_2))\cdots\mathbb{P}(C_2(y_{k-1})) \\ \leq & C \sum_{r=k-1}^{(k-1)(n-1)} \sum_{\substack{0<y_i<r \\ y_1+y_2+\cdots+y_{k-1}=r}} \mathbb{P}(C_2(y_1))P(C_2(y_2))\cdots\mathbb{P}(C_2(y_{k-1})) \\ = & C \sum_{\substack{0<y_i\leq n-1 \\ i=1,2,\ldots,k-1}} \mathbb{P}(C_2(y_1))\mathbb{P}(C_2(y_2))\cdots\mathbb{P}(C_2(y_{k-1})) \\ = & C\left(\sum_{y=1}^{n-1}\mathbb{P}(C_2(y))\right)^{k-1}.\end{aligned}$$

Combining Proposition 3.3 with Stolz's theorem, we get that

$$\lim_{n\to+\infty}\frac{\sum_{y=1}^{n-1}\mathbb{P}(C_2(y))}{\log n} = \lim_{n\to+\infty}\frac{\sum_{y=1}^{n-1}\mathbb{P}(C_2(y))}{\sum_{y=1}^{n-1}\frac{1}{y}}\cdot\frac{\sum_{y=1}^{n-1}\frac{1}{y}}{\log n} = \frac{1}{2}.$$

It follows that, for any $\varepsilon > 0$, there exists $N_1(\varepsilon)$ such that for all $n > N_1(\varepsilon)$, $\sum_{y=1}^{n-1}\mathbb{P}(C_2(y)) \leq (\frac{1}{2}+\varepsilon)\log n$ and $\frac{C}{(\frac{1}{2}+\varepsilon)\log n} \leq \frac{1}{2}$. Thus for all $n > N_1(\varepsilon)$, it holds that

$$\mathbb{E}\alpha_k^+(n) \leq C[(\frac{1}{2}+\varepsilon)\log n]^{k-1} = \frac{C}{(\frac{1}{2}+\varepsilon)\log n}[(\frac{1}{2}+\varepsilon)\log n]^k \leq \frac{1}{2}[(\frac{1}{2}+\varepsilon)\log n]^k. \tag{3.15}$$

Next, we bound $\mathbb{E}\alpha_k^+(n)$ from below. Since $k \sim a\log n$ as $n \to +\infty$, we know when n is sufficiently large, $\frac{n}{3k} > 1$. Let $j_1 \leq \frac{n}{3}$, $0 < j_l - j_{l-1} \leq \frac{n}{3k}$, $l = 2,3,\ldots,k$. Then $j_k = \sum_{l=2}^{k}(j_l - j_{l-1}) + j_1 < \frac{2n}{3}$. Hence, by (3.11), we have that

$$\begin{aligned}\mathbb{E}\alpha_k^+(n) \geq & \sum_{\substack{0\leq j_1\leq\frac{n}{3} \\ 0<j_l-j_{l-1}\leq\frac{n}{3k}, \\ l=2,3,\ldots,k}} \mathbb{P}(D_1(j_1))\mathbb{P}(C_2(j_2-j_1))\mathbb{P}(C_2(j_3-j_2))\cdots\mathbb{P}(C_2(j_k-j_{k-1}))\mathbb{P}(D_2(j_k,n)) \\ = & \sum_{k-1\leq r\leq(k-1)\frac{n}{3k}} \left[\sum_{0\leq j\leq\frac{n}{3}}\mathbb{P}(D_1(j))\mathbb{P}(D_2(j+r,n))\right] \\ & \cdot\left[\sum_{\substack{0<y_i\leq\frac{n}{3k},1\leq i\leq k-1 \\ y_1+y_2+\cdots+y_{k-1}=r}}\mathbb{P}(C_2(y_1))\mathbb{P}(C_2(y_2))\cdots\mathbb{P}(C_2(y_{k-1}))\right].\end{aligned} \tag{3.16}$$

It follows from Proposition 3.3 that there exists a positive constant c_2 such that for all integers $n, r \geq 1$ and $j \geq 0$ with $n - j - r \geq 1$,

$$\sqrt{j}\mathbb{P}(D_1(j)) \geq c_2,\ \sqrt{n-j-r}\mathbb{P}(D_2(j+r,n)) \geq c_2.$$

Thus

$$\sum_{0\le j\le \frac{n}{3}} \mathbb{P}(D_1(j))\mathbb{P}(D_2(j+r,n)) \ge c_2^2 \sum_{0\le j\le \frac{n}{3}} \frac{1}{\sqrt{nj}} = \frac{c_2^2}{\sqrt{3}} \sum_{0\le j\le \frac{n}{3}} \frac{1}{\sqrt{\frac{3j}{n}}}\cdot\frac{3}{n},$$

which together with $\lim_{n\to+\infty} \sum_{0\le j\le \frac{n}{3}} \frac{1}{\sqrt{\frac{3j}{n}}}\cdot\frac{3}{n} = \int_0^1 x^{-1/2}\mathrm{d}x = 2$ implies that there exists a positive constant $\tilde{C}$ (independent of $r\ge 1$) such that

$$\sum_{0\le j\le \frac{n}{3}} \mathbb{P}(D_1(j))\mathbb{P}(D_2(j+r,n)) \ge \tilde{C}. \tag{3.17}$$

Hence, by (3.16) and (3.17), we have

$$\begin{aligned}\mathbb{E}\alpha_k^+(n) \ge& \tilde{C} \sum_{k-1\le r\le (k-1)\frac{n}{3k}} \left[\sum_{\substack{0<y_i\le \frac{n}{3k},1\le i\le k-1\\ y_1+y_2+\cdots+y_{k-1}=r}} \mathbb{P}(C_2(y_1))\mathbb{P}(C_2(y_2))\cdots\mathbb{P}(C_2(y_{k-1}))\right]\\ =&\tilde{C}\left[\sum_{0<y\le\frac{n}{3k}} \mathbb{P}(C_2(y))\right]^{k-1}.\end{aligned} \tag{3.18}$$

Combining the fact that $k\sim a\log n$ as $n\to+\infty$ with Proposition 3.3, we have

$$\lim_{n\to\infty} \frac{\sum_{0<y\le\frac{n}{3k}}\mathbb{P}(C_2(y))}{\log n} = \lim_{n\to\infty} \frac{\sum_{0<y\le\frac{n}{3k}}\mathbb{P}(C_2(y))}{\sum_{0<y\le\frac{n}{3k}}\frac{1}{y}}\cdot\frac{\sum_{0<y\le\frac{n}{3k}}\frac{1}{y}}{\log\frac{n}{3k}}\cdot\frac{\log\frac{n}{3k}}{\log n} = \frac{1}{2}.$$

It follows that, for any $\varepsilon\in(0,\frac{1}{2})$, there exists $N_2(a,\varepsilon) > N_1(\varepsilon)$ such that for all $n > N_2(a,\varepsilon)$, $\sum_{0<y\le\frac{n}{3k}}\mathbb{P}(C_2(y)) \ge (\frac{1}{2}-\frac{\varepsilon}{2})\log n$, which together with (3.18) implies that for all $n > N_2(a,\varepsilon)$,

$$\mathbb{E}\alpha_k^+(n) \ge \tilde{C}\left[(\frac{1}{2}-\frac{\varepsilon}{2})\log n\right]^{k-1} = \left[(\frac{1}{2}-\varepsilon)\log n\right]^k\cdot\frac{\tilde{C}}{\frac{1}{2}-\varepsilon}\left[\frac{\frac{1}{2}-\frac{\varepsilon}{2}}{\frac{1}{2}-\varepsilon}\right]^{k-1}\frac{1}{\log n}.$$

Since by $k\sim a\log n$ as $n\to+\infty$, we have

$$\lim_{n\to+\infty}\left[\frac{\frac{1}{2}-\frac{\varepsilon}{2}}{\frac{1}{2}-\varepsilon}\right]^{k-1}\frac{1}{\log n} = +\infty.$$

Hence, there exists $N_3(a,\varepsilon)\ge N_2(a,\varepsilon)$ such that for all $n > N_3(a,\varepsilon)$,

$$\mathbb{E}\alpha_k^+(n) \ge \frac{1}{2}\left[(\frac{1}{2}-\varepsilon)\log n\right]^k. \tag{3.19}$$

By the symmetry of $\{S_n\}_{n\ge 0}$, (3.10), (3.15) and (3.19), we obtain that for all $n > N_3(a,\varepsilon)$,

$$\frac{1}{2}\left[(\frac{1}{2}-\varepsilon)\log n\right]^k \le \mathbb{E}\alpha_k^-(n) = \mathbb{E}\alpha_k^+(n) \le \frac{1}{2}[(\frac{1}{2}+\varepsilon)\log n]^k. \tag{3.20}$$

Notice that $\alpha^+(n)\alpha^-(n) = 0$. Hence

$$\alpha_k(n) = \alpha_k^+(n) + \alpha_k^-(n).$$

Thus, by (3.15), (3.19) and (3.20), for any $\varepsilon \in (0, \frac{1}{2})$, there exists $n_0(a, \varepsilon) = N_3(a, \varepsilon)$ such that for all $n > n_0(a, \varepsilon)$,

$$\left[(\frac{1}{2} - \varepsilon)\log n\right]^k \leq \mathbb{E}\alpha_k(n) = \mathbb{E}\alpha_k^+(n) + \mathbb{E}\alpha_k^-(n) \leq [(\frac{1}{2} + \varepsilon)\log n]^k.$$

□

3.2 Proof of Theorem 1.2

In this proof, C stands for a positive constant whose value may change from one appearance to another. We prove the theorem in three steps.

Step 1: In this step, we will prove that, for all $a > 0$ and $\varepsilon > 0$, there exists $N_1(a, \varepsilon)$ such that for all $n > N_1(a, \varepsilon)$,

$$\mathbb{P}(\alpha(n) > a(\log n)^2) < n^{-2a+\varepsilon}.$$

Let $k = \lfloor 2a \log n \rfloor$ and $0 < \delta < 1$. By Markov's inequality, Lemma 3.2 and properties of the Gamma function, there exists $n_1(\delta)$ such that for all $n \geq n_1(\delta)$, we have

$$\begin{aligned}
&\mathbb{P}\left(\alpha(n) > a(\log n)^2\right) \leq \mathbb{P}\left(\alpha(n) > \lfloor a(\log n)^2 \rfloor\right) \\
&= \mathbb{P}\left(\alpha_k(n) > \binom{\lfloor a(\log n)^2 \rfloor}{k}\right) \leq \frac{\mathbb{E}\alpha_k(n)}{\binom{\lfloor a(\log n)^2 \rfloor}{k}} \\
&\leq [\frac{1}{2}(1+\delta)\log n]^k \cdot \frac{\Gamma(k+1)\Gamma(\lfloor a(\log n)^2 \rfloor - k + 1)}{\Gamma(\lfloor a(\log n)^2 \rfloor + 1)} \\
&\leq [\frac{1}{2}(1+\delta)\log n]^{2a\log n} \cdot \frac{\Gamma(2a\log n + 1)\Gamma(a(\log n)^2 - 2a\log n + 2)}{\Gamma(a(\log n)^2)} \\
&= [\frac{1}{2}(1+\delta)\log n]^{2a\log n} \cdot a(\log n)^2 \left(a(\log n)^2 - 2a\log n + 1\right) \\
&\quad \cdot \frac{\Gamma(2a\log n + 1)\Gamma(a(\log n)^2 - 2a\log n + 1)}{\Gamma(a(\log n)^2 + 1)}.
\end{aligned} \tag{3.21}$$

Then by Stirling's formula, we have

$$\begin{aligned}
&\mathbb{P}\left(\alpha(n) > a(\log n)^2\right) \\
\leq &C[\frac{1}{2}(1+\delta)\log n]^{2a\log n} \cdot (\log n)^4 \\
&\cdot \frac{(2a\log n)^{2a\log n + 1/2}[a(\log n)^2 - 2a\log n]^{a(\log n)^2 - 2a\log n + 1/2}}{[a(\log n)^2]^{a(\log n)^2 + 1/2}} \\
\leq &C(1+\delta)^{2a\log n} \cdot (\log n)^{9/2}\left[1 - \frac{2}{\log n}\right]^{a(\log n)^2 - 2a\log n + 1/2}.
\end{aligned} \tag{3.22}$$

By Taylor's expansion, we have

$$\begin{aligned}
&\left[1-\frac{2}{\log n}\right]^{a(\log n)^2-2a\log n+1/2} \\
=&\exp\left\{\left[a(\log n)^2-2a\log n+1/2\right]\log\left(1-\frac{2}{\log n}\right)\right\} \\
=&\exp\left\{\left[a(\log n)^2-2a\log n+1/2\right]\left[-\frac{2}{\log n}+O((\log n)^{-2})\right]\right\} \\
=&\exp\left\{-2a(\log n)+O(1)\right\},
\end{aligned}$$

which together with (3.22) implies that for all $\varepsilon>0$,

$$\begin{aligned}
&\mathbb{P}(\alpha(n)>a(\log n)^2) \\
&\le C(1+\delta)^{2a\log n}(\log n)^{\frac{9}{2}}\cdot\exp\left\{-2a(\log n)\right\} \\
&= n^{-2a+\varepsilon}\cdot\exp\left\{[2a\log(1+\delta)-\varepsilon]\log n+\frac{9}{2}\log(\log n)+C\right\}.
\end{aligned}$$

Hence, for any fixed $\varepsilon>0$, there exists $0<\delta_1(\varepsilon)<\frac{1}{2}$ such that for all $\delta<\delta_1(\varepsilon)$, $2a\log(1+\delta)-\varepsilon<0$. Thus we have $\lim\limits_{n\to+\infty}\exp\{[2a\log(1+\delta)-\varepsilon]\log n+\frac{9}{2}\log(\log n)+C\}=0$. Therefore, for any $\delta\in(0,\delta_1(\varepsilon))$, there exists $N_1(a,\varepsilon)>n_1(\delta)$ such that for all $n>N_1(a,\varepsilon)$,

$$\mathbb{P}(\alpha(n)>a(\log n)^2)<n^{-2a+\varepsilon}. \tag{3.23}$$

Step 2: Let $\delta\in(0,\frac{1}{2})$, $\bar{a}=a(1+2\delta), d=2a$, $k=\lfloor 2a\log n\rfloor=\lfloor d\log n\rfloor$, $k'=\lfloor d(1+\delta)\log n\rfloor$, and $k''=\lfloor d(1+2\delta)\log n\rfloor$. In this step, we will prove that when n is sufficiently large,

$$\sum_{m\ge\bar{a}(\log n)^2}\mathbb{P}(\alpha(n)=m)\binom{m}{k'}\le\frac{1}{3}\mathbb{E}\alpha_{k'}(n) \tag{3.24}$$

and

$$\sum_{m\le a(\log n)^2}\mathbb{P}(\alpha(n)=m)\binom{m}{k'}\le\frac{1}{3}\mathbb{E}\alpha_{k'}(n). \tag{3.25}$$

We will only give the proof of (3.24). The proof of (3.25) is similar.

For any $m\ge k''$, $\frac{\binom{m}{k'}}{\binom{m}{k''}}$ decreases as m increases. Thus

$$\begin{aligned}
&\sum_{m\ge\bar{a}(\log n)^2}\mathbb{P}(\alpha(n)=m)\binom{m}{k'} \\
=&\sum_{m\ge\bar{a}(\log n)^2}\mathbb{P}(\alpha(n)=m)\binom{m}{k''}\cdot\frac{\binom{m}{k'}}{\binom{m}{k''}} \\
\le&\left[\sum_{m\ge\bar{a}(\log n)^2}\mathbb{P}(\alpha(n)=m)\binom{m}{k''}\right]\cdot\frac{\binom{\lfloor\bar{a}(\log n)^2\rfloor}{k'}}{\binom{\lfloor\bar{a}(\log n)^2\rfloor}{k''}}
\end{aligned}$$

$$\leq \mathbb{E}\alpha_{k''}(n)\frac{\binom{\lfloor \bar{a}(\log n)^2\rfloor}{k'}}{\binom{\lfloor \bar{a}(\log n)^2\rfloor}{k''}}. \tag{3.26}$$

By Markov's inequality, Lemma 3.2, properties of the Gamma function, there exists $n_2(\delta)$ such that for all $n > n_2(\delta)$,

$$\begin{aligned}
&\frac{\mathbb{E}\alpha_{k''}(n)}{\mathbb{E}\alpha_{k'}(n)}\cdot\frac{\binom{\lfloor \bar{a}(\log n)^2\rfloor}{k'}}{\binom{\lfloor \bar{a}(\log n)^2\rfloor}{k''}}\\
&\leq \frac{[\frac{1}{2}(1+\delta^3)\log n]^{k''}}{[\frac{1}{2}(1-\delta^3)\log n]^{k'}}\cdot\frac{\Gamma(k''+1)\Gamma(\lfloor \bar{a}(\log n)^2\rfloor - k''+1)}{\Gamma(k'+1)\Gamma(\lfloor \bar{a}(\log n)^2\rfloor - k'+1)}\\
&\leq \frac{[\frac{1}{2}(1+\delta^3)\log n]^{d(1+2\delta)\log n}}{[\frac{1}{2}(1-\delta^3)\log n]^{d(1+\delta)\log n-1}}\cdot\frac{\Gamma(d(1+2\delta)\log n+1)\Gamma(\bar{a}(\log n)^2-d(1+2\delta)\log n+2)}{\Gamma(d(1+\delta)\log n)\Gamma(\bar{a}(\log n)^2-d(1+\delta)\log n)}\\
&= \frac{[\frac{1}{2}(1+\delta^3)\log n]^{d(1+2\delta)\log n}}{[\frac{1}{2}(1-\delta^3)\log n]^{d(1+\delta)\log n-1}}\\
&\quad\cdot[d(1+\delta)\log n][\bar{a}(\log n)^2-d(1+\delta)\log n][\bar{a}(\log n)^2-d(1+2\delta)\log n+1]\\
&\quad\cdot\frac{\Gamma(d(1+2\delta)\log n+1)\Gamma(\bar{a}(\log n)^2-d(1+\delta)\log n+1)}{\Gamma(d(1+\delta)\log n+1)\Gamma(\bar{a}(\log n)^2-d(1+\delta)\log n+1)}\\
&\leq C\frac{[\frac{1}{2}(1+\delta^3)\log n]^{d(1+2\delta)\log n}}{[\frac{1}{2}(1-\delta^3)\log n]^{d(1+\delta)\log n-1}}(\log n)^5\\
&\quad\cdot\frac{(d(1+2\delta)\log n)^{d(1+2\delta)\log n+1/2}(\bar{a}(\log n)^2-d(1+2\delta)\log n)^{\bar{a}(\log n)^2-d(1+2\delta)\log n+1/2}}{(d(1+\delta)\log n)^{d(1+\delta)\log n+1/2}(\bar{a}(\log n)^2-d(1+\delta)\log n)^{\bar{a}(\log n)^2-d(1+\delta)\log n+1/2}}\\
&\leq C(\log n)^6\frac{(1+\delta^3)^{d(1+2\delta)\log n}}{(1-\delta^3)^{d(1+\delta)\log n}}\left(\frac{1+2\delta}{1+\delta}\right)^{d(1+\delta)\log n}\frac{(1-\frac{d(1+2\delta)}{\bar{a}\log n})^{\bar{a}(\log n)^2-d(1+2\delta)\log n+1/2}}{(1-\frac{d(1+\delta)}{\bar{a}\log n})^{\bar{a}(\log n)^2-d(1+\delta)\log n+1/2}}.
\end{aligned} \tag{3.27}$$

By Taylor's expansion, we have

$$\begin{aligned}
&\left(1-\frac{d(1+2\delta)}{\bar{a}\log n}\right)^{\bar{a}(\log n)^2-d(1+2\delta)\log n+1/2}\\
&=\exp\left\{\left[\bar{a}(\log n)^2-d(1+2\delta)\log n+1/2\right]\log\left(1-\frac{d(1+2\delta)}{\bar{a}\log n}\right)\right\}\\
&=\exp\left\{\left[\bar{a}(\log n)^2-d(1+2\delta)\log n+1/2\right]\left[-\frac{d(1+2\delta)}{\bar{a}\log n}+O((\log n)^{-2})\right]\right\}\\
&=\exp\{-d(1+2\delta)\log n+O(1)\}.
\end{aligned}$$

Similarly, we have

$$\left(1-\frac{d(1+\delta)}{\bar{a}\log n}\right)^{\bar{a}(\log n)^2-d(1+\delta)\log n+1/2} = \exp\{-d(1+\delta)\log n+O(1)\},$$

$$\left(\frac{1+2\delta}{1+\delta}\right)^{d(1+\delta)\log n} = \exp\left\{d(1+\delta)[\log(1+2\delta)-\log(1+\delta)]\log n\right\}$$

$$
\begin{aligned}
&= \exp\left\{d(1+\delta)(\delta - \frac{3}{2}\delta^2 + O(\delta^3))\log n\right\} \\
&= \exp\left\{\left(d\delta - \frac{d\delta^2}{2} + O(\delta^3)\right)\log n\right\},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{(1+\delta^3)^{d(1+2\delta)\log n}}{(1-\delta^3)^{d(1+\delta)\log n}} &= \exp\left\{d(1+2\delta)\log(1+\delta^3)\log n - d(1+\delta)\log(1-\delta^3)\log n\right\} \\
&= \exp\{O(\delta^3)\log n\}.
\end{aligned}
$$

Combining the four displays above with (3.27), we get

$$
\begin{aligned}
&\frac{\mathbb{E}\alpha_{k''}(n)}{\mathbb{E}\alpha_{k'}(n)} \cdot \frac{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k'}}{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k''}} \\
\le& C(\log n)^6 \cdot \exp\{O(\delta^3)\log n\} \cdot \exp\left\{\left(d\delta - \frac{d\delta^2}{2} + O(\delta^3)\right)\log n\right\} \cdot \frac{\exp\{-d(1+2\delta)\log n\}}{\exp\{-d(1+\delta)\log n\}} \\
\le& \exp\left\{-\left(\frac{d}{2} - O(\delta)\right)\delta^2 \log n + 6\log\log n + C\right\}.
\end{aligned}
$$

So there exists $\delta_2(d) > 0$ such that for all $\delta < \delta_2(d)$, we have $\frac{d}{2} - O(\delta) = a - O(\delta) > 0$. Thus we have

$$
\lim_{n\to+\infty} \frac{\mathbb{E}\alpha_{k''}(n)}{\mathbb{E}\alpha_{k'}(n)} \cdot \frac{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k'}}{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k''}} = 0. \tag{3.28}
$$

Combining this with (3.26), we get that, for any fixed $\delta \in (0, \delta_2(d))$, there exists $N_2(\delta) > n_2(\delta)$ such that for all $n > N_2(\delta)$,

$$
\sum_{m \ge \bar{a}(\log n)^2} \mathbb{P}(\alpha(n) = m)\binom{m}{k'} \le \mathbb{E}\alpha_{k'}(n) \cdot \frac{\mathbb{E}\alpha_{k''}(n)}{\mathbb{E}\alpha_{k'}(n)} \cdot \frac{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k'}}{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k''}} \le \frac{1}{3}\mathbb{E}\alpha_{k'}(n).
$$

Step 3: In this step, we will prove that when n is sufficiently large,

$$
\mathbb{P}(\alpha(n) > a(\log n)^2) > n^{-2a-\varepsilon}.
$$

By Markov's inequality, Lemma 3.2 and properties of the Gamma function, there exists $n_3(\delta)$ such that for all $n > n_3(\delta)$,

$$
\begin{aligned}
&\frac{\mathbb{E}\alpha_{k'}(n)}{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k'}} \\
\ge& [\frac{1}{2}(1-\delta)\log n]^{k'} \frac{\Gamma(k'+1)\Gamma(\lfloor \bar{a}(\log n)^2 \rfloor - k' + 1)}{\Gamma(\lfloor \bar{a}(\log n)^2 \rfloor + 1)} \\
\ge& [\frac{1}{2}(1-\delta)\log n]^{d(1+\delta)\log n - 1} \frac{\Gamma(d(1+\delta)\log n)\Gamma(\bar{a}(\log n)^2 - d(1+\delta)\log n)}{\Gamma(\bar{a}(\log n)^2 + 1)}
\end{aligned}
$$

$$
\begin{aligned}
&\geq [\frac{1}{2}(1-\delta)\log n]^{d(1+\delta)\log n-1}\,[d(1+\delta)\log n]^{-1}\,\left[\bar{a}(\log n)^2-d(1+\delta)\log n\right]^{-1}\\
&\quad\cdot\frac{\Gamma(d(1+\delta)\log n+1)\Gamma(\bar{a}(\log n)^2-d(1+\delta)\log n+1)}{\Gamma(\bar{a}(\log n)^2+1)}\\
&\geq C[\frac{1}{2}(1-\delta)\log n]^{d(1+\delta)\log n}(\log n)^{-4}\\
&\quad\cdot\frac{(d(1+\delta)\log n)^{d(1+\delta)\log+1/2}(\bar{a}(\log n)^2-d(1+\delta)\log n)^{\bar{a}(\log n)^2-d(1+\delta)\log n+1/2}}{(\bar{a}(\log n)^2)^{\bar{a}(\log n)^2+1/2}}\\
&\geq C(\log n)^{-7/2}\left(\frac{1-\delta}{1+2\delta}\right)^{d(1+\delta)\log n}\left(1-\frac{d(1+\delta)}{\bar{a}\log n}\right)^{\bar{a}(\log n)^2-d(1+\delta)\log n+1/2}. \qquad (3.29)
\end{aligned}
$$

By Taylor's expansion, we have

$$
\begin{aligned}
&\left(1-\frac{d(1+\delta)}{\bar{a}\log n}\right)^{\bar{a}(\log n)^2-d(1+\delta)\log n+1/2}\\
&=\exp\left\{\left[\bar{a}(\log n)^2-d(1+\delta)\log n+1/2\right]\log\left(1-\frac{d(1+\delta)}{\bar{a}\log n}\right)\right\}\\
&=\exp\left\{\left[\bar{a}(\log n)^2-d(1+\delta)\log n+1/2\right]\left[-\frac{d(1+\delta)}{\bar{a}\log n}+O((\log n)^{-2})\right]\right\}\\
&=\exp\left\{-d(1+\delta)\log n+O(1)\right\}
\end{aligned}
$$

and

$$
\begin{aligned}
\left(\frac{1-\delta}{1+2\delta}\right)^{d(1+\delta)\log n} &=\exp\left\{d(1+\delta)[\log(1-\delta)-\log(1+2\delta)]\log n\right\}\\
&=\exp\left\{d(1+\delta)(-3\delta+O(\delta^2))\log n\right\}\\
&=\exp\left\{d(-3\delta+O(\delta^2))\log n\right\}.
\end{aligned}
$$

Combining the two displays above with (3.29), we get that for any $\varepsilon>0$,

$$
\begin{aligned}
\frac{\mathbb{E}\alpha_{k'}(n)}{\binom{\lfloor\bar{a}(\log n)^2\rfloor}{k'}} &\geq C(\log n)^{-7/2}\cdot\exp\left\{d(-3\delta+O(\delta^2))\log n\right\}\cdot\exp\left\{-d(1+\delta)\log n\right\}\\
&=n^{-(d+\varepsilon)}\cdot\exp\left\{C-\frac{7}{2}\log(\log n)+(\varepsilon-4d\delta+O(\delta^2))\log n\right\}.
\end{aligned}
$$

For any $\varepsilon>0$, there exists $\delta_3(\varepsilon)>0$ such that $\varepsilon-4d\delta+O(\delta^2)>0$ for all $\delta<\delta_3(\varepsilon)$. Thus $\lim_{n\to+\infty}\exp\left\{C-\frac{7}{2}\log\log n+(\varepsilon-4d\delta+O(\delta^2))\log n\right\}=+\infty$. Hence, for any $\delta\in(0,\delta_3(\varepsilon))$, there exists $N_3(\delta)>n_3(\delta)$ such that for all $n>N_3(\delta)$,

$$
\frac{\mathbb{E}\alpha_{k'}(n)}{\binom{\lfloor\bar{a}(\log n)^2\rfloor}{k'}}>3n^{-(d+\varepsilon)}=3n^{-2a-\varepsilon}. \qquad (3.30)
$$

Let $\delta=\frac{1}{2}\min\{\delta_1(\varepsilon),\delta_2(d),\delta_3(\varepsilon)\}$. By the analysis above, we know that there exists $N(a,\varepsilon)$ such that for all $n>N(a,\varepsilon)$, (3.23), (3.24), (3.25) and (3.30) hold and $\bar{a}(\log n)^2-a(\log n)^2=2a\delta(\log n)^2>1$. Hence,

$$
\mathbb{P}(\alpha(n)>a(\log n)^2)
$$

$$
\begin{aligned}
&\geq \sum_{a(\log n)^2 < m < \bar{a}(\log n)^2} \mathbb{P}(\alpha(n) = m) \\
&\geq \frac{1}{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k'}} \sum_{a(\log n)^2 < m < \bar{a}(\log n)^2} \mathbb{P}(\alpha(n) = m) \binom{m}{k'} \\
&= \frac{1}{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k'}} \left[\mathbb{E}\alpha_{k'}(n) - \sum_{m \geq \bar{a}(\log n)^2} \mathbb{P}(\alpha(n) = m) \binom{m}{k'} - \sum_{m \leq a(\log n)^2} \mathbb{P}(\alpha(n) = m) \binom{m}{k'} \right] \\
&\geq \frac{1}{3} \frac{\mathbb{E}\alpha_{k'}(n)}{\binom{\lfloor \bar{a}(\log n)^2 \rfloor}{k'}} > n^{-2a-\varepsilon}. \qquad (3.31)
\end{aligned}
$$

Hence, by (3.23) and (3.31), we obtain that for all $n > N(a, \varepsilon)$,

$$
n^{-2a-\varepsilon} < \mathbb{P}\left(\alpha(n) > a(\log n)^2\right) < n^{-2a+\varepsilon}.
$$

□

Remark 3.4. *From the proof of Theorem 1.2, we know that the key is that Lemma 3.2 holds for $\mathbb{E}\alpha_k(n)$. By (3.20) we know that $\mathbb{E}\alpha_k^+(n)$ satisfies a similar inequality as $\mathbb{E}\alpha_k(n)$. So by following the proof of Theorem 1.2, we can get the same conclusion for $\mathbb{E}\alpha^+(n)$, i.e. for all $a > 0$ and $\varepsilon > 0$, there exists an $N_0 = N_0(a, \varepsilon)$ such that for all $n > N_0$*

$$
n^{-2a-\varepsilon} < \mathbb{P}\left(\alpha^+(n) > a(\log n)^2\right) < n^{-2a+\varepsilon}.
$$

3.3 Proof of Theorem 1.3

Step 1. First we deal with the upper bound of $\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2}$.

By Theorem 1.2, for all $\varepsilon > 0$, there exists n_0 such that for all $n > n_0$,

$$
\mathbb{P}\left(\alpha(n) > (\frac{1}{2} + \varepsilon)(\log n)^2\right) < n^{-(1+2\varepsilon)+\varepsilon} = n^{-1-\varepsilon}.
$$

It follows that

$$
\sum_{n=1}^{+\infty} \mathbb{P}\left(\alpha(n) > (\frac{1}{2} + \varepsilon)(\log n)^2\right) \leq \sum_{n=1}^{n_0} \mathbb{P}\left(\alpha(n) > (\frac{1}{2} + \varepsilon)(\log n)^2\right) + \sum_{n=n_0+1}^{+\infty} n^{-1-\varepsilon} < \infty.
$$

Thus by the Borel-Cantelli lemma, we have

$$
\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} \leq \frac{1}{2} + \varepsilon\right) \geq \mathbb{P}\left(\bigcup_{k=1}^{+\infty} \bigcap_{n=k}^{+\infty} \left\{\alpha(n) \leq (\frac{1}{2} + \varepsilon)(\log n)^2\right\}\right) = 1.
$$

So

$$
\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} \leq \frac{1}{2}\right) = \lim_{\varepsilon\to 0^+} \mathbb{P}\left(\limsup_{n\to\infty} \frac{\alpha(n)}{(\log n)^2} \leq \frac{1}{2} + \varepsilon\right) = 1. \qquad (3.32)
$$

Step 2. In this step, we deal with lower bound of $\limsup_{n\to\infty}\frac{\alpha(n)}{(\log n)^2}$.

For $k\geq 1$, define

$$\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k):=\{z\geq k^2,\ z\in\mathcal{A}^+_{\sigma_{k^2}+k}\},$$

where σ_{k^2} is defined in (2.12). Then we have

$$\begin{aligned}\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k)&=\{z\geq k^2:\exists! t\in[\sigma_{k^2},\sigma_{k^2}+k)\ s.t.\ S_t=z,S_{t+1}=z+1\}\\&\in\sigma(X_{\sigma_{k^2}+1},X_{\sigma_{k^2}+2},\ldots,X_{\sigma_{k^2}+k}).\end{aligned}$$

Since $\sigma_{(k+1)^2}-\sigma_{k^2}\geq 2k+1$, we get that $\{\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k),k\geq 1\}$ are independent.

We define $\tilde{S}_t^{\sigma_{k^2}}:=S_{\sigma_{k^2}+t}-S_{\sigma_{k^2}},0\leq t\leq k$. Then $\{\tilde{S}_t^{\sigma_{k^2}}\}_{0\leq t\leq k}$ is a simple symmtric random walk with $\tilde{S}_0^{\sigma_{k^2}}=0$. We denote $\tilde{\mathcal{A}}^{\sigma_{k^2},+}(k)$ the counterpart of $\mathcal{A}_k^+$ in (3.10) for the random walk $\{\tilde{S}_t^{\sigma_{k^2}}\}_{0\leq t\leq k}$. Then we know that $\#\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k)$ and $\#\tilde{\mathcal{A}}^{\sigma_{k^2},+}(k)$ have the same distribution. Remark 3.4 tells us that Theorem 1.2 also holds for $\#\tilde{\mathcal{A}}^{\sigma_{k^2},+}(k)$. Hence, for all $\varepsilon\in(0,\frac{1}{2})$, we have

$$\begin{aligned}&\sum_{k=1}^{+\infty}\mathbb{P}\left(\#\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k)>(\frac{1}{2}-\varepsilon)(\log k)^2\right)\\&=\sum_{k=1}^{+\infty}\mathbb{P}\left(\#\tilde{\mathcal{A}}^{\sigma_{k^2},+}(k)>(\frac{1}{2}-\varepsilon)(\log k)^2\right)\\&\geq\sum_{k=1}^{k_0}\mathbb{P}\left(\#\tilde{\mathcal{A}}^{\sigma_{k^2},+}(k)>(\frac{1}{2}-\varepsilon)(\log k)^2\right)+\sum_{k=k_0+1}^{+\infty}k^{-2(\frac{1}{2}-\varepsilon)-\varepsilon}=+\infty.\end{aligned}$$

Then, by the Borel-Cantelli lemma again, we get

$$\mathbb{P}\left(\#\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k)>(\frac{1}{2}-\varepsilon)(\log k)^2,i.o.\right)=1,$$

which together with the fact that $\#\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k)\leq\alpha^+(\sigma_{k^2}+k)$ implies that

$$\mathbb{P}\left(\alpha^+(\sigma_{k^2}+k)>(\frac{1}{2}-\varepsilon)(\log k)^2,i.o.\right)\geq\mathbb{P}(\#\mathcal{A}^+(\sigma_{k^2},\sigma_{k^2}+k)>(\frac{1}{2}-\varepsilon)(\log k)^2,i.o.)=1.$$

Since $\mathbb{P}(\lim\limits_{n\to+\infty}\frac{\sigma_n}{n^4}=0)=1$, we have

$$\mathbb{P}\left(\left\{\alpha^+(\sigma_{k^2}+k)>(\frac{1}{2}-\varepsilon)(\log k)^2,i.o.\right\}\cap\left\{\lim_{n\to+\infty}\frac{\sigma_n}{n^4}=0\right\}\right)=1.\tag{3.33}$$

For any $\omega\in\{\alpha^+(\sigma_{k^2}+k)>(\frac{1}{2}-\varepsilon)(\log k)^2,i.o.\}\cap\{\lim\limits_{n\to+\infty}\frac{\sigma_n}{n^4}=0\}$, there exists $k_j(\omega)\to+\infty$, as $j\to+\infty$ such that, for all $j\geq 1$, $\frac{\alpha^+(\sigma_{k_j^2}+k_j)}{(\log k_j)^2}>\frac{1}{2}-\varepsilon$, $\sigma_{k_j^2}<\frac{1}{2}k_j^8$ and $k_j<\frac{1}{2}k_j^8$. Thus

$$\frac{\alpha(\sigma_{k_j^2}+k_j)}{[\log(\sigma_{k_j^2}+k_j)]^2}\geq\frac{\alpha^+(\sigma_{k_j^2}+k_j)}{[\log(\sigma_{k_j^2}+k_j)]^2}=\frac{\alpha^+(\sigma_{k_j^2}+k_j)}{(\log k_j)^2}\cdot\frac{(\log k_j)^2}{[\log(\sigma_{k_j^2}+k_j)]^2}$$

$$> (\frac{1}{2} - \varepsilon)\frac{(\log k_j)^2}{(\log k_j^8)^2} = \frac{1-2\varepsilon}{128}.$$

Hence, we have

$$\limsup_{n\to+\infty} \frac{\alpha(n)(\omega)}{(\log n)^2} \geq \limsup_{j\to+\infty} \frac{\alpha(\sigma_{k_j^2} + k_j)}{[\log(\sigma_{k_j^2} + k_j)]^2}(\omega) \geq \frac{1-2\varepsilon}{128}.$$

So by (3.33), we have

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} \geq \frac{1-2\varepsilon}{128}\right) = 1.$$

Thus we have

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} \geq \frac{1}{128}\right) = \lim_{\varepsilon\to 0^+} \mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} \geq \frac{1-2\varepsilon}{128}\right) = 1, \tag{3.34}$$

which together with (3.32) implies that

$$\mathbb{P}\left(\frac{1}{128} \leq \limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} \leq \frac{1}{2}\right) = 1. \tag{3.35}$$

Hence, by Proposition 3.1 and (3.35), we know that there exists a constant $C \in \left[\frac{1}{128}, \frac{1}{2}\right]$ such that

$$\mathbb{P}\left(\limsup_{n\to+\infty} \frac{\alpha(n)}{(\log n)^2} = C\right) = 1.$$

□

Acknowledgments We thank Yinshan Chang and Longmin Wang for their comments and discussions. This work was supported by the National Natural Science Foundation of China (12171335, 12001389, 11931004, 12071011), the Simons Foundation (#960480) and the Science Development Project of Sichuan University (2020SCUNL201).

References

[1] Bass, R. F. (2022). The rate of escape of the most visited site of Brownian motion, arXiv: 1303.2040v4.

[2] 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.

[3] Ding, J. and Shen, J. (2018). Three favorite sites occurs infinitely often for one-dimensional simple random walk. *Ann. Probab.* **46**, 2545–2561.

[4] Durrett, R. (2019). Probability: theory and examples (5th edition). Cambridge University Press, Cambridge.

[5] Erdős, P. and Révész, P. (1984). On the favourite points of a random walk. In *Mathematical Structure–Computational Mathematics–Mathematical Modelling* **2**, 152–157.

[6] 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.

[7] Erdős, P. and Révész, P. (1991). Three problems on the random walk in $\mathbf{Z}^d$. *Studia Sci. Math. Hungar.* **26**, 309–320.

[8] Hao, C.-X. (2023). The escape rate of favorite edges of simple random walk, arXiv: 2303.13210v1.

[9] Hao, C.-X., Hu, Z.-C., Ma, T. and Song, R. (2023). Three favorite edges occurs infinitely often for one-dimensional simple random walk. *Commun. Math. Stat.* (accepted).

[10] Hao, C.-X., Hu, Z.-C., Ma, T. and Song, R. (2023). Favorite downcrossing sites for one-dimensional simple random walk, *J. Sichuan Univ.: Nat. Sci. Ed. (accepted).*

[11] Major, P. (1988). On the set visited once by a random wlk. *Probab. Th. Rel. Fields* **77**, 117–128.

[12] Newman, D. (1984). In a random walk the number of "unique experiences" is two on the average, *SIAM Review* **26**, 573–574.

[13] Révész, P. (1990). Random Walk in Random and Non-Random Environment. World Scientific, Singapore.

[14] Shi, Z. and Tóth, B. (2000). Favourite sites of simple random walk. *Period. Math. Hungar.* **41**, 237–249.

[15] Tóth, B. (1996). Multiple covering of the range of a random walk on $\mathbf{Z}$ (on a question of P. Erdös and P. Révész), *Studia Sci. Math. Hungar.* **31**, 355–359.

[16] Tóth, B. (2001). No more than three favorite sites for simple random walk. *Ann. Probab.* **29**, 484–503.

[17] Tóth, B. and Werner, W. (1997). Tied favourite edges for simple random walk. *Combin. Probab. Comput.* **6**, 359–369.