Long-range one-dimensional internal diffusion-limited aggregation

Conrado da Costa*† Debleena Thacker*‡ Andrew Wade*§

18th November 2024

Abstract

We study internal diffusion limited aggregation on $\mathbb{Z}$, where a cluster is grown by sequentially adding the first site outside the cluster visited by each random walk dispatched from the origin. We assume that the increment distribution X of the driving random walks has $\mathbb{E}\, X = 0$, but may be neither simple nor symmetric, and can have $\mathbb{E}(X^2) = \infty$, for example. For the case where $\mathbb{E}(X^2) < \infty$, we prove that after m walks have been dispatched, all but $o(m)$ sites in the cluster form an approximately symmetric contiguous block around the origin. This extends known results for simple random walk. On the other hand, if X is in the domain of attraction of a symmetric α-stable law, $1 < \alpha < 2$, we prove that the cluster contains a contiguous block of $\delta m + o(m)$ sites, where $0 < \delta < 1$, but, unlike the finite-variance case, one may not take $\delta = 1$.

Key words: Aggregation; random walk; growth process; renewal theory; overshoots.

AMS Subject Classification: 60K35 (Primary) 60F15, 60G50, 60K05, 82C24 (Secondary).

Contents

*Department of Mathematical Sciences, Durham University, Durham DH1 3LE, UK.
†conrado.da-costa@durham.ac.uk
‡debleena.thacker@durham.ac.uk
§andrew.wade@durham.ac.uk

1 Introduction and main results

1.1 Diffusion-generated growth

Internal diffusion-limited aggregation (IDLA) is a discrete-time stochastic growth model on an infinite graph driven by a sequence of random walks, dispatched one after another from a common site (the germ of the aggregate), with each walker expanding the aggregate by the addition of the first site the walker visits outside the current aggregate. The model was introduced by Diaconis & Fulton [16] in the case where the driving random walk is simple symmetric random walk (SSRW) on $\mathbb{Z}^d$.

Lawler, Bramson & Griffeath [34] proved that the long-time shape of the aggregate generated by SSRW on $\mathbb{Z}^d$ converges to a sphere. More recently, a rich picture concerning fluctuations around the limit shape for $d \geq 2$ has been revealed [6,7,26–28] (it was already observed in [16, pp. 107–8] and [34, p. 2118] that in $d = 1$ the fluctuations of the process can be described via Friedman's urn).

The terminology "internal" refers to the fact that the successive walkers are dispatched from inside the current cluster, in contrast to classical DLA [44], in which they are dispatched "from infinity". We refer the readers to [42] for a survey and comparison of the two different DLA models. One major difference between these two models is that the IDLA models exhibit regularity in their asymptotic behaviour, for example in $\mathbb{Z}^d$, $d \geq 2$ the limit shape is a Euclidean ball; whereas the shapes generated by classical DLA models appear to be highly irregular and display fractal structure, although mathematical results are scarce (see [32,42,44]). *Long-range* classical DLA models have been studied in [2–4,8], driven by random walks that are allowed jumps beyond nearest-neighbours.

The vast majority of work on IDLA has been concerned with SSRW as the driving random walk; drifted simple random walk was studied by [38], while the case of walkers starting uniformly in the present cluster, rather than from a fixed origin, was studied in [10]. Similar models on regular trees have been studied in [8] and are related to digital search trees that have received considerable attention in computer science, see [19]. A continuous-time version of IDLA was introduced in [11], driven by an oriented simple random walk on the upper half plane of $\mathbb{Z}^2$, and admitting a coupling with some first-passage percolation models. Further connections of IDLA driven by simple random walks on $\mathbb{Z}^2$ and random forests have been investigated in [15].

Our interest here is *long-range IDLA*, in which the underlying walker is not SSRW, and to investigate the degree to which the regularity exhibited by classical (SSRW-driven)

IDLA is preserved. We will consider random walks whose increments have mean zero, and either finite variance, or are in the domain of normal attraction of a symmetric α-stable distribution with characteristic function $\mathrm{e}^{\beta|t|^\alpha}$ for $1 < \alpha < 2$. Our main results are in two parts, and are presented in detail in Section 1.3 below, after we have introduced necessary notation in Section 1.2). First, we show that the long time behavior observed for SSRW-driven IDLA extends to the mean-zero, finite-variance case (Theorem 1.3). Second, we show that the regular behaviour begins to degrade in the infinite-variance case (Theorem 1.6), demonstrating a phase transition in the model.

The complementary work [9] introduced an IDLA model on $\mathbb{Z}$ where the initial semi-infinite aggregate consists of sites $\{0, -1, -2, \ldots\}$ (say) and each walker comes from $-\infty$ (as opposed to our model where the walker start from the origin), and steps on the renewal points of a bi-infinite renewal chain (therefore, long-range walk) and settles at the first vacant site it encounters. The main object of interest in [9] is the recurrence/transience properties of the occupied sites, and also to show that in case of geometric jumps, there is a connection to ASEP blocking measures.

1.2 Model and notation

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which supports an array of i.i.d. $\mathbb{Z}$-valued random variables X and $X_i^{(m)}$, $m \in \mathbb{N}$, $i \in \mathbb{N}$. Consider $S^{(1)}, S^{(2)}, \ldots$ a sequence of independent random walks, started from the origin, defined via $S_0^{(m)} := 0$ and

$$S_n^{(m)} := \sum_{i=1}^{n} X_i^{(m)}, \text{ for } n \in \mathbb{N}.$$

For convenience, we also consider a random walk whose increment distribution is the same as that of X above, but with an arbitrary starting point; to this end we introduce (for each $x \in \mathbb{Z}$) a probability space $(\Omega', \mathcal{F}', \mathbf{P}_x)$ which supports a random walk $S = (S_0, S_1, S_2, \ldots)$ such that $\mathbf{P}_x(S_0 = x) = 1$, and $S_{n+1} - S_n$ are i.i.d. with $\mathbf{P}_x(S_{n+1} - S_n = y) = \mathbb{P}(X = y)$ for all $x, y \in \mathbb{Z}$. We write $\mathbb{E}, \mathbf{E}_x$ for expectation corresponding to $\mathbb{P}, \mathbf{P}_x$, respectively.

Throughout this paper we assume the following *irreducibility* hypothesis, which ensures that the random walk can visit all of $\mathbb{Z}$:

(I) Suppose that for every $x, y \in \mathbb{Z}$, there is $n \in \mathbb{N}$ such that

$$\mathbf{P}_x(S_n = y) > 0. \tag{1.1}$$

Remark 1.1. An elementary sufficient condition for (I), which holds for simple symmetric random walk and many other examples, is that (i) X is not constant and has $\mathbb{E}\, X = 0$ (*zero mean drift*), and (ii) for no $h > 1$ does it hold that $\mathbb{P}(X \in h\mathbb{Z}) = 1$; this last condition is no real loss of generality, as if $\mathbb{P}(X \in h\mathbb{Z}) = 1$ for $h > 1$, then one can work instead with increment X/h. To see the sufficiency, note that if (i) holds, then the support of X contains at least one positive value $x_+ \in \mathbb{N}$ and at least one negative value $-x_-$ for $x_- \in \mathbb{N}$, while if (ii) holds, x_+, x_- can be chosen to have $\gcd(x_+, x_-) = 1$.

Define $\mathfrak{C}_0 := \{0\}$ (the germ) and then, recursively, for $m \in \mathbb{N}$ let

$$\tau_m := \inf\bigl\{n \in \mathbb{Z}_+ : S_n^{(m)} \notin \mathfrak{C}_{m-1}\bigr\}, \text{ and } \mathfrak{C}_m := \mathfrak{C}_{m-1} \cup \{S_{\tau_m}^{(m)}\}. \tag{1.2}$$

Hypothesis (I) implies that $\limsup_{n\to\infty} |S_n^{(m)}| = \infty$, $\mathbb{P}$-a.s., and hence $\tau_m < \infty$, $\mathbb{P}$-a.s., for every $m \in \mathbb{N}$. Consequently, the number of sites in $\mathfrak{C}_m$ is $\#\mathfrak{C}_m = m + 1$.

We make two comments on the notation. First, we use n for the internal clock for each walker, and keep m for indexing the walkers. Second, while $S^{(m)}$ denotes the mth random walker to be released in our IDLA process, we use S and $\mathbf{P}_x$ to make statements generically about the random walk with increments distributed as X.

We call $\mathfrak{C}_m$ the *IDLA cluster* generated by the first m walkers. By construction, $\mathfrak{C}_m \subset \mathfrak{C}_{m+1}$ are (strictly) increasing; denote the limit by $\mathfrak{C}_\infty := \cup_{m\in\mathbb{Z}_+}\mathfrak{C}_m$, the collection of sites eventually contained in the cluster. The limit set $\mathfrak{C}_\infty$ is an infinite subset of $\mathbb{Z}$; the following elementary result, whose proof is in Appendix A, says that it is, in fact, the whole of $\mathbb{Z}$, with no gaps.

Proposition 1.2. *Suppose that* (I) *holds. Then* $\mathbb{P}(\mathfrak{C}_\infty = \mathbb{Z}) = 1$.

To describe our main results define, for $m \in \mathbb{Z}_+$,

$$r_m := \max\{r \in \mathbb{Z}_+ : \mathbb{Z} \cap [-r, r] \subseteq \mathfrak{C}_m\}, \tag{1.3}$$

the radius of the maximal centred interval contained in $\mathfrak{C}_m$. An easy consequence of the fact that $\#\mathfrak{C}_m = m + 1$ is that

$$r_m \leq m/2, \text{ for all } m \in \mathbb{Z}_+. \tag{1.4}$$

1.3 Main results

It follows from Proposition 1.2 that $\lim_{m\to\infty} r_m = \infty$, a.s. We are interested in quantifying the growth rate of r_m. Loosely speaking, our first result, Theorem 1.3, says that when $\mathbb{E}(X^2) < \infty$ and $\mathbb{E}\, X = 0$, the inner radius r_m grows at the maximal rate permitted by (1.4). Here is the precise statement.

Theorem 1.3. *Suppose that* (I) *holds,* $\mathbb{E}(X^2) < \infty$*, and* $\mathbb{E}\, X = 0$*. Then, a.s.,*

$$\lim_{m\to\infty} \frac{r_m}{m} = \frac{1}{2}. \tag{1.5}$$

Remark 1.4. For SSRW, when $\mathbb{P}(X = +1) = \mathbb{P}(X = -1) = 1/2$, it is well known that $\lim_{m\to\infty} r_m/m = 1/2$, a.s. Indeed, in this case $\mathfrak{C}_m$ is always a contiguous interval, and the equivalence of the model to (Bernard) Friedman's urn, as described at [16, pp. 107–8] and [34, p. 2118], yields $\lim_{m\to\infty} r_m/m = 1/2$, a.s., via (David) Freedman's strong law [22]. As far as the authors are aware, the other cases of Theorem 1.3 are new.

We are also interested in the case where $\mathbb{E}(X^2) = \infty$. We will explore this case under the following stable-domain hypothesis for $1 < \alpha < 2$.

($\mathbf{S}_\alpha$) Suppose that $X \in \mathbb{Z}$ has a symmetric distribution (i.e., $X \overset{d}{=} -X$), and its characteristic function $\phi(t) := \mathbb{E}(\mathrm{e}^{itX})$ satisfies

$$\lim_{t\to 0} \left(|t|^{-\alpha}(1 - \phi(t))\right) = \beta \in (0, \infty). \tag{1.6}$$

Remarks 1.5. (i) The hypothesis (1.6) is equivalent to the assumption that X is in the domain of normal attraction of ζ_α, the symmetric α-stable distribution with characteristic function $\mathrm{e}^{\beta|t|^\alpha}$: see Theorem 2.6.7 [24, pp. 92–3]. Recall that being in the domain of normal attraction ζ_α means that $n^{-1/\alpha}S_n$ converges in distribution to ζ_α, i.e., the slowly-varying component of the scaling sequence is constant.

(ii) For $\alpha \in (1,2)$, assumption (S_α) implies that $\mathbb{E}\,|X| < \infty$ (in fact $\mathbb{E}(|X|^\gamma) < \infty$ for every $\gamma < \alpha$), with $\mathbb{E}\,X = 0$ (due to symmetry), but $\mathbb{E}(X^2) = \infty$ [24, p. 93].

Loosely speaking, our second main result says that, under hypothesis (S_α) with $\alpha \in (1,2)$, it still holds that r_m grows at linear rate with m (in lim inf and lim sup sense), but this rate is now strictly less than the maximal rate $1/2$ permitted by (1.4).

Theorem 1.6. *Suppose that* (I) *holds, and that* (S_α) *holds with* $\alpha \in (1,2)$. *Then, there exist constants* c_α, c'_α *with* $0 < c_\alpha \leq c'_\alpha < 1/2$, *such that,*

$$c_\alpha \leq \liminf_{m\to\infty} \frac{r_m}{m} \leq \limsup_{m\to\infty} \frac{r_m}{m} \leq c'_\alpha, \ \text{a.s.} \tag{1.7}$$

Moreover, one may take for the constant c_α *in* (1.7) *the expression*

$$c_\alpha = \frac{(\alpha-1)(2-\alpha)^{2-\alpha}}{(4-\alpha)(3-\alpha)^3}. \tag{1.8}$$

Remark 1.7. The value of c_α in (1.8) is not the best that can be extracted from our method (see Section 3.5), but was chosen for its relatively simple formula, together with its property that $c_\alpha \uparrow 1/2$ as $\alpha \uparrow 2$, which demonstrates some continuity between Theorem 1.6 and Theorem 1.3. We do not give here an explicit expression for c'_α; though an explicit but not very informative bound could be extracted using our methods.

1.4 Overview and discussion

The bulk of the rest of the paper provides the proofs of Theorems 1.3 and 1.6. Section 2 presents preparatory results concerning properties of the underlying random walk; these bring together known results from random walk and renewal theory (Section 2.1), with some important hitting and exit estimates for integer-valued random walks (Section 2.2), due to Kesten [30, 31]. The proofs of the main results are separated into proofs of lower bounds on the growth rate of r_m/m (Section 3) and upper bounds on the growth rate of r_m/m in the infinite variance case (Section 3.5). The strategy of Section 3 adapts, in part, the approach of [34], combined with the results of Kesten mentioned above; an outline of the argument is presented in Section 3.1. Section 3.5 uses further results of Kesten, in the neighbourhood of the Dynkin–Lamperti renewal theorem (see Section 2.1). The proof of Theorem 1.3 is accomplished in Section 3.3, while the proof of Theorem 1.6 combines a lower bound from Section 3.4 with upper bound from Section 3.5, and is concluded in the latter section. Finally, to avoid disrupting the flow of the paper, we defer to the appendix the proofs of some auxiliary results. In particular, Appendix A gives the short proof of the eventual filling statement in Proposition 1.2, and Appendix B some technical elements about certain families of probability functions that are introduced in Section 2.2.

We finish this section with some remarks and open problems. Theorem 1.6 poses some obvious questions. Firstly:

Problem 1.8. Suppose that (I) holds, and that (S_α) holds with $\alpha \in (1,2)$. Does it hold that $\liminf_{m\to\infty} r_m/m = \ell_\alpha$, a.s., for some constant ℓ_α, a.s.? If so, what is its value?

Of course, if it exists, ℓ_α must satisfy $c_\alpha \leq \ell_\alpha \leq c'_\alpha$, by (1.7), and (see Remark 1.7) it must hold that $\ell_\alpha \to 1/2$ as $\alpha \uparrow 2$. One might hope to be able to apply a zero–one law to obtain existence of ℓ_α, but we have not been able to do so. Two further questions in this regime are the following.

Problem 1.9. Suppose that (I) holds, and that (S_α) holds with $\alpha \in (1, 2)$. Does $\lim_{m\to\infty} r_m/m$ exist in this case, as it does in Theorem 1.3?

Further questions arise in the case with $\mathbb{E}\,|X| = \infty$.

Problem 1.10. Suppose that (S_α) holds with $\alpha \in (0, 1)$, or $\alpha = 1$. What is the behaviour of r_m now?

The underlying random walks are of very different character in the case $\alpha \in (0, 1]$: for example, when $\alpha \in (0, 1)$ the walks are (oscillatory) transient; this does not obviously indicate a more disperse aggregate, however, as the walks will typically aggregate after fewer steps. The boundary case $\alpha = 1$ is likely to be delicate, and there are some technical obstructions: for example, the results of Kesten [30, 31] that we use below often omit the case $\alpha = 1$, although comparable results for stable diffusions are known [14]. Lastly:

Problem 1.11. Consider random walks in $\mathbb{Z}^d$, $d \geq 2$.

For simple symmetric random walk, the seminal work [34] deals with general d, so one aspect of Problem 1.11 would be to try to extend that work to walks with sufficiently many moments. For stable walks of index $\alpha \in (1, 2)$, we know of no multidimensional analogues of Kesten's results, although for symmetric stable processes some Green's function estimates are known [14].

2 Ingredients from random walks and renewal theory

2.1 Ladder processes and overshoots

Recall that on probability space $(\Omega', \mathcal{F}', \mathbf{P}_x)$, $x \in \mathbb{Z}$, we have a random walk $S = (S_n)_{n\in\mathbb{Z}_+}$ started from $S_0 = x$ whose increment distribution is that of the $\mathbb{Z}$-valued random variable X underlying our IDLA model. We introduce some additional notation to enable us to discuss classical fluctuation theory for this random walk. Define the strict ascending ladder times $\lambda_0 := 0$ and

$$\lambda_k := \inf\{n \geq \lambda_{k-1} : S_n > S_{\lambda_{k-1}}\}, \text{ for } k \in \mathbb{N}; \tag{2.1}$$

as usual, $\inf \emptyset := \infty$.

We suppose that $\mathbb{E}\,|X| < \infty$ and $\mathbb{E}\,X = 0$. By a result of Chung and Fuchs, this implies that the random walk S_n is recurrent [21, p. 615], so, for every $x \in \mathbb{Z}$,

$$\mathbf{P}_x\Big(\limsup_{n\to\infty} S_n = \infty\Big) = \mathbf{P}_x\Big(\liminf_{n\to\infty} S_n = -\infty\Big) = 1. \tag{2.2}$$

In particular, $\mathbf{P}_x(\lambda_k < \infty) = 1$ for every $k \in \mathbb{Z}_+$, and $S_{\lambda_k} = \max_{0\leq n\leq\lambda_k} S_n$ are strictly increasing in k.

Define $L_n := S_{\lambda_n}$ for $n \in \mathbb{Z}_+$; then $x = L_0 < L_1 < L_2 < \cdots$ is the (strict, ascending) *ladder height* process associated with S. By the strong Markov property and spatial homogeneity, the increments $Y_n := L_n - L_{n-1}$, $n \in \mathbb{N}$, are i.i.d., $\mathbb{N}$-valued, and $L_n = x + \sum_{i=1}^n Y_i$ represents the ladder height process as a renewal process with increment distribution $Y := Y_1$. By translation invariance, the distribution of the ladder times λ_k and ladder increments Y_n are the same for every $\mathbf{P}_x$, regardless of the starting point $x \in \mathbb{Z}$

of the random walk. In such cases where the starting point is unimportant, we will abuse notation slightly and write simply $\mathbf{P}$, $\mathbf{E}$ on occasion.

Suppose next that $S_0 = x \equiv 0$. For $y \in \mathbb{Z}_+$, define the renewal counting process

$$N_y := \inf\{n \in \mathbb{N} : L_n > y\} = \#\{n \in \mathbb{Z}_+ : L_n \le y\}.$$

Note $L_{N_0} = L_1 = Y_1$, $N_y \in \mathbb{N}$, and $L_{N_y - 1} \le y < L_{N_y}$, $\mathbf{P}_0$-a.s. for every $y \in \mathbb{Z}_+$. Define the *residual life-time* process associated with the ladder-height renewal process by

$$Z_y := L_{N_y} - y, \text{ for all } y \in \mathbb{Z}_+, \tag{2.3}$$

which satisfies $Z_y \in \mathbb{N}$; see [5, pp. 140–1] or [21, §XI.4] for background on the terminology and renewal-theoretic context. A fundamental observation [5, p. 9] is that $Z_0, Z_1, Z_2, \ldots$ forms an irreducible Markov chain on a subset of $\mathbb{N}$ with transitions given by

$$Z_{y+1} = \begin{cases} Z_y - 1 & \text{if } Z_y \ge 2, \\ Y_{N_{y+1}} & \text{if } Z_y = 1. \end{cases} \tag{2.4}$$

Furthermore, the Markov chain $(Z_n)_{n \in \mathbb{Z}_+}$ is aperiodic, i.e. $\gcd\{n \in \mathbb{Z}_+ : Z_n = 1\} = 1$. Indeed (see Remark 1.1) in this case the support of X contains some x_+ and $-x_-$, with $x_+, x_- \in \mathbb{N}$ and $\gcd(x_+, x_-) = 1$. Hence we can find $k, \ell \in \mathbb{N}$ with $kx_+ - \ell x_- = 1$, and with positive probability the random walk S can take ℓ steps of value x_- followed by k steps of value x_+, meaning that the ladder variable has $\mathbb{P}(Y = 1) > 0$.

Returning to the random walk, we denote the first passage time of S above level $y \in \mathbb{Z}_+$ by

$$\rho_y := \inf\{n \in \mathbb{Z}_+ : S_n > y\}, \tag{2.5}$$

and we call $S_{\rho_y} - y$ the first (right) *overshoot* of level y by the random walk S. Since ρ_y is necessarily a ladder time, it holds that

$$S_{\rho_y} - y = L_{N_y} - y = Z_y, \ \mathbf{P}_0\text{-a.s., for every } y \in \mathbb{Z}_+; \tag{2.6}$$

thus overshoots of the random walk are equivalent to residual life-times of the associated ladder-height renewal process.

Remark 2.1. We state all the results in this section for *right* overshoots and *increasing* ladder variables, but, evidently, by working with the increment distribution $-X$, we can translate everything to left overshoots and decreasing ladder variables.

For most of the rest of this section, we assume additionally that the increments have finite variance:

(M) Suppose that $\sigma^2 := \mathbb{E}(X^2) \in (0, \infty)$ and $\mathbb{E}\, X = 0$.

The following result presents some key properties of overshoots.

Proposition 2.2. *Suppose that* (I) *and* (M) *hold. Then* $\mu := \mathbf{E}\, Y$ *satisfies* $1 \le \mu < \infty$, *and*

$$\mu = \exp\left\{ \sum_{n=1}^{\infty} \frac{1}{n} \left[\frac{1}{2} - \mathbf{P}(S_n > 0) \right] \right\}. \tag{2.7}$$

Moreover, define the probability distributions $(\pi_k)_{k \in \mathbb{N}}$ *and* $(\psi_k)_{k \in \mathbb{N}}$ *by*

$$\pi_k := \frac{\mathbf{P}(Y \ge k)}{\mu}, \quad \psi_k := \frac{k\, \mathbf{P}(Y = k)}{\mu}, \ \text{for } k \in \mathbb{N}. \tag{2.8}$$

(i) Let Z_∞, L_∞ denote independent random variables with distributions given by π, ψ from (2.8)*, respectively, and set $U_\infty := \max(Z_\infty, L_\infty)$. Then*

$$\sup_{y\in\mathbb{Z}_+} \mathbf{P}_0(Z_y \geq k) \leq \mathbf{P}_0(U_\infty \geq k), \text{ for all } k \in \mathbb{Z}_+.$$

(ii) It holds that, for every $x \in \mathbb{Z}$ and every $k \in \mathbb{N}$,

$$\lim_{y\to\infty} \mathbf{P}_x(Z_y = k) = \pi_k. \tag{2.9}$$

(iii) Suppose, additionally, that $\mathbb{E}(|X|^p) < \infty$, for some $p > 2$. Then,

$$\sup_{y\in\mathbb{Z}_+} \mathbf{E}_x\big[Z_y^{p-2}\big] < \infty. \tag{2.10}$$

In particular, Proposition 2.2(ii) says that if $\mathbb{E}(X^2) < \infty$, the overshoots are tight. This is in abrupt contrast to the case where $\mathbb{E}(X^2) = \infty$, where, under appropriate conditions, the following consequence of the *Dynkin–Lamperti theorem* (see Proposition 2.3) says that the overshoot over level y lives on scale y, asymptotically. Note that the result is valid for all $\alpha \in (0, 2)$, although we will later use only the case $\alpha \in (1, 2)$.

Proposition 2.3. *Suppose that* (I) *holds, and that* (S_α) *holds with $\alpha \in (0, 2)$. Then, for every $x \in \mathbb{Z}$ and every $u \geq 0$,*

$$\lim_{y\to\infty} \mathbf{P}_x\left(\frac{S_{\rho_y} - y}{y} > u\right) = \int_u^\infty f_\alpha(v)\mathrm{d}v,$$

where

$$f_\alpha(v) := \frac{\sin(\pi\alpha/2)}{\pi}\frac{1}{v^{\alpha/2}(1+v)}, \quad \text{for } v > 0.$$

Propositions 2.2 and 2.3 are well known: Kesten's Lemma 6 [30, p. 255] provides Proposition 2.3 explicitly, and gives Proposition 2.2(ii) under an additional symmetry assumption. Another route to Proposition 2.3 is to combine the Dynkin–Lamperti renewal theorem [12, p. 361] with the result that under hypothesis (S_α), the ladder variable Y is in the domain of attraction of a positive $\alpha/2$-stable law (see Theorem 9 of Rogozin [41, p. 592]); a corresponding local limit theorem is given in [18]. We give below a proof of Proposition 2.2 without Kesten's additional hypothesis, but the proof involves little more than indicating appropriate results in the literature. First, we give some intuition behind the important "loss of moments" phenomenon which the above results exhibit. For example, Proposition 2.2 says that in order for the overshoot to have a uniformly bounded mean, we need to assume $\mathbb{E}(|X|^3) < \infty$.

Theorem 3.4 of Spitzer [43, p. 158] shows that the hypothesis (M) (finite variance) implies integrability of the ladder height, $1 \leq \mathbf{E}\,Y < \infty$, and gives the formula (2.7). We "lose moments" in passing from the walk to its ladder heights (this cannot be avoided, as explained in Remark 2.4 below). The fact that we "lose another moment" in passing from the ladder heights to the (stationary) overshoots is due to the observation that, if Z is a random variable distributed as $\mathbf{P}(Z = k) = \pi_k$ from (2.8), then

$$\mathbf{E}(Z^q) = \sum_{k\in\mathbb{N}} k^q \pi_k = \frac{1}{\mathbf{E}\,Y}\sum_{k\in\mathbb{N}} k^q\, \mathbf{P}(Y \geq k), \tag{2.11}$$

which is finite if and only if $\mathbf{E}(Y^{q+1}) < \infty$. This is a "size-biasing" effect; in the stationary renewal process associated with Y, the intervals that straddle a particular value are more likely to be long.

Remark 2.4. Suppose that (I) and (M) hold. Since $Y \geq 1$ we have $L_n \geq n$ and hence $N_y \leq y+1$, a.s., and $L_{N_y} - y \leq L_{y+1} - y$. Hence, for every $y \in \mathbb{Z}_+$, it holds that

$$\mathbf{E}(Z_y^q) < \infty, \text{ whenever } \mathbf{E}(Y^q) < \infty. \tag{2.12}$$

The purpose of this remark is to explain that we cannot claim that (2.12) holds uniformly in y. Let Z denote a random variable whose distribution is given by $\mathbf{P}(Z = k) = \pi_k$ as given by (2.8); Proposition 2.2(ii) shows that Z_y converges to Z in distribution as $y \to \infty$. Suppose that $\sup_y \mathbf{E}(Z_y^q) < \infty$. Then uniform integrability shows that,

$$\text{for every } q' \in (0,q), \ \mathbf{E}(Z^{q'}) = \lim_{y \to \infty} \mathbf{E}(Z_y^{q'}) < \infty. \tag{2.13}$$

In the special case where X is symmetric, Corollary 2 of Doney [17, p. 250] states that,

$$\text{for } p > 2, \ \mathbb{E}(|X|^p) < \infty \text{ if and only if } \mathbf{E}(Y^{p-1}) < \infty. \tag{2.14}$$

In particular, if for some $q > 1$ one has $\mathbb{E}(|X|^{q+1}) < \infty$ but $\mathbb{E}(|X|^{q+(3/2)}) = \infty$, say, then Doney's result (2.14) says $\mathbf{E}(Y^q) < \infty$ but $\mathbf{E}(Y^{q+(1/2)}) = \infty$, and so $\mathbf{E}(Z^{q-(1/2)}) = \infty$, by the discussion around (2.11). This is a contradiction with (2.13). Hence we cannot, in general, insert a supremum over y into (2.12).

Proof of Proposition 2.2. Suppose, without loss of generality, that $S_0 = 0$, and consider $Z_y = S_{\rho_y} - y$ for some $y \in \mathbb{Z}_+$. As mentioned above, the fact that $\mu < \infty$ satisfies (2.7) is due to Spitzer [43]. Part (i) is an inequality in the vein of Lorden [36], obtained by Chang [13], using a coupling argument.

The most elegant (and probabilistic) argument for part (ii) proceeds from the Markov chain representation (2.4). Indeed, $Z_0 = 1$ and $Z_1 = Y_{N_1} = Y_1$. If we set $\tau := \inf\{n \in \mathbb{N} : Z_n = 1\}$, then $\tau = Y_1$, a.s., and the usual excursion-occupation construction shows that an invariant measure $(\mu(y), y \in \mathbb{N})$ for the Markov chain is given by

$$\mu(y) = \frac{1}{\mathbf{E}\,\tau}\,\mathbf{E} \sum_{n=1}^{\tau} \mathbb{1}\{Z_n = y\} = \frac{1}{\mathbf{E}\,Y_1}\,\mathbf{E} \sum_{n=1}^{Y_1} \mathbb{1}\{Z_n = y\} = \frac{1}{\mathbf{E}\,Y}\,\mathbf{E}\,\mathbb{1}\{Y \geq y\},$$

which is exactly π given by (2.8). As remarked after (2.4), the Markov chain is irreducible and aperiodic under the hypotheses of the proposition, and so the convergence in (2.9) follows from the Markov chain convergence theorem. Part (ii) can be found as Theorem 6.10.3 of [23, pp. 103–4](ii), and may also be derived from the classical renewal theorem for aperiodic lattice random variables [21, p. 363].

Part (iii) follows from part (i), and indeed from Theorem 3 of Lorden [36] (see also Theorem 3.1 of [25]), once one knows that $\mathbf{E}(Y^{p-1}) < \infty$ whenever $\mathbf{E}(|X|^p) < \infty$. This fact (the $p > 2$ analogue of Spitzer's result for $p = 2$ that we already used) is provided by results of Doney [18] and Lai [33]. □

2.2 Estimates from Kesten on hitting before exit

In this section we present some estimates on hitting and exit of $\mathbb{Z}$-valued random walks, derived more-or-less directly from fine results of Kesten [30, 31]. To state the results, for $t \in \mathbb{Z}$ and $A \subset \mathbb{Z}$, define

$$T_t := \inf\{n \in \mathbb{Z}_+ : S_n = t\}, \text{ and } \eta_A := \inf\{n \in \mathbb{Z}_+ : S_n \notin A\}, \tag{2.15}$$

respectively the first hitting time of t and the first exit time from A for the random walk S. First, we describe informally the results that we will use. Throughout this discussion, we assume that (I) holds and that $\mathbb{E}\, X = 0$.

- An easy, but important, consequence of recurrence and irreducibility is a *local hitting* estimate saying that there is high probability of visiting a nearby site before going far away: see Lemma 2.5 below.
- Kesten provides general *gambler's ruin* estimates on the probability of exiting a large interval on one side rather than the other: see Lemma 2.6 below.
- The local hitting and gambler's ruin estimates show no essential distinction between the finite- and infinite-variance cases. Where the distinction arises is in what the walk does when it exits an interval, i.e., the behaviour of *overshoots*, as we have seen with the contrast between Propositions 2.2 (tight overshoots) and 2.3 (large-scale overshoots).
- Kesten combines ingeniously the above elements to obtain precise asymptotics for the probability of hitting a particular point before exit from a large interval. Roughly speaking, in the finite-variance case, the tight overshoots and local hitting estimates show that the gambler's ruin probabilities capture the essential behaviour, while in the infinite-variance case there may be many overshoots of the target point before it is successfully hit. Lemmas 2.8 and 2.11 below present the main estimates we will need of this type.

We now give precise statements of the results described loosely above. We start with the following local hitting property, which is a consequence of recurrence and irreducibility: see equation (2.4) of [30, p. 247].

Lemma 2.5. *Suppose that* (I) *holds and that* $\mathbb{E}\, X = 0$. *Then* $\lim_{N\to\infty} \mathbf{P}_0(T_k < \eta_{[-N,N]}) = 1$ *for every fixed* $k \in \mathbb{Z}$.

Next we present Kesten's general gambler's ruin estimates. The finite-variance part, Lemma 2.6(i), is contained in Theorem 2 of [30, p. 256], while the infinite-variance part, Lemma 2.6(ii), is contained in Corollary 1 of [31, p. 273]. We remark that the strength of part (i) is that no more than finite second moments is assumed, but asymptotic lower and upper bounds match; compare e.g. Theorem 5.1.7 of [35, p. 127], which does not provide matching bounds, or Theorem 5 of [37], which requires the hypothesis $\mathbb{E}(|X|^3) < \infty$.

Lemma 2.6 (Kesten 1961 [30, 31])**.** *Suppose that* (I) *holds.*

(i) Suppose that $\mathbb{E}(X^2) < \infty$ *and* $\mathbb{E}\, X = 0$. *Then, for every* $c \in \mathbb{R}_+$,

$$\lim_{N\to\infty} \mathbf{P}_0\big(S_{\eta_{[-cN,N]}} > N\big) = \frac{c}{1+c}. \tag{2.16}$$

(ii) Suppose that (S_α) *holds with* $\alpha \in (1,2)$. *Then, for every* $c \in \mathbb{R}_+$,

$$\lim_{N\to\infty} \mathbf{P}_0\big(S_{\eta_{[-cN,N]}} > N\big) = \frac{\Gamma(\alpha)}{\Gamma(\alpha/2)^2} \int_{(1+c)^{-1}}^1 u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1} \mathrm{d}u. \tag{2.17}$$

Remarks 2.7. (i) A change of variable followed by the symmetric beta-integral formula [1, p. 258] shows that, for every $\alpha > 0$,

$$2\int_{1/2}^{1} u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1}\mathrm{d}u = \int_{0}^{1} u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1}\mathrm{d}u = \frac{\Gamma(\alpha/2)^2}{\Gamma(\alpha)}. \tag{2.18}$$

Consequently, for exit from a symmetric interval ($c = 1$), both (2.16) and (2.17) yield the asymptotically-fair ruin estimate $\lim_{N\to\infty} \mathbf{P}_0\big(S_{\eta_{[-N,N]}} > N\big) = 1/2$.

(ii) Corollary 1 in Kesten [31, p. 273] is stated slightly differently from (2.17), in terms of $\lim_{N\to\infty} \mathbf{P}_0\big(S_{\eta_{[-cN,N]}} < -cN\big)$. However, one easily deduces (2.17) using (2.18), since

$$\begin{aligned}
\lim_{N\to\infty} \mathbf{P}_0\big(S_{\eta_{[-cN,N]}} > N\big) &= 1 - \lim_{N\to\infty} \mathbf{P}_0\big(S_{\eta_{[-cN,N]}} < -cN\big) \\
&= \frac{\Gamma(\alpha)}{\Gamma(\alpha/2)^2}\left[\int_0^1 u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1}\mathrm{d}u - \int_0^{(1+c)^{-1}} u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1}\mathrm{d}u\right] \\
&= \frac{\Gamma(\alpha)}{\Gamma(\alpha/2)^2}\int_{(1+c)^{-1}}^{1} u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1}\mathrm{d}u.
\end{aligned}$$

We turn to the most delicate results, which are estimates for the probability that a particular site in an interval is visited before exit from the interval. Borrowing notations from Kesten [31], let us define for $N \in \mathbb{Z}_+$, $c > 0$, and $k \in \{0, 1, \ldots, N\}$,

$$q_{\alpha,N}\Big(\frac{k}{N}; c\Big) := \mathbf{P}_k\left(T_0 < \eta_{[-cN,N]}\right), \tag{2.19}$$

where $\alpha = 2$ if $\mathbb{E}(X^2) < \infty$ and $\alpha \in (1, 2)$ if (S_α) is satisfied. Also define $q_{\alpha,N}(y; c)$ over all $y \in [0, 1]$ by linear interpolation, i.e.,

$$q_{\alpha,N}(y;c) := (k+1-Ny)q_{\alpha,N}\Big(\frac{k}{N};c\Big) + (Ny-k)q_{\alpha,N}\Big(\frac{k+1}{N};c\Big), \text{ if } \frac{k}{N} < y < \frac{k+1}{N}.$$

For $c' \geq c > 0$, we have $\eta_{[-cN,N]} \leq \eta_{[-c'N,N]}$, a.s.; this shows the following monotonicity property

$$q_{\alpha,N}(y; c) \leq q_{\alpha,N}(y; c'), \text{ whenever } 0 < c \leq c'. \tag{2.20}$$

The main result that of this subsection is the following. The result is essentially due to Kesten [30, 31], but part (i) is not given explicitly by Kesten, so we give a proof later in this subsection.

Lemma 2.8 (Kesten 1961 [30, 31]). *Suppose that* (I) *holds. Fix* $0 \leq y < 1$ *and* $c > 0$.

(i) Suppose that $\mathbb{E}(X^2) < \infty$ *and* $\mathbb{E}\, X = 0$. *Then there exists the limit*

$$\lim_{N\to\infty} q_{2,N}(y; c) = q_2(y; c) := 1 - y. \tag{2.21}$$

(ii) Suppose that (S_α) *holds with* $\alpha \in (1, 2)$. *Then there exists the limit*

$$\lim_{N\to\infty} q_{\alpha,N}(y; c) = q_\alpha(y; c), \tag{2.22}$$

where

$$q_\alpha(y;c) := (\alpha-1)\, c^{1-\frac{\alpha}{2}}\, (1+c)^{\alpha-1}\, (y+c)^{\frac{\alpha}{2}}\, y^{\alpha-1} \int_y^1 (y+cv)^{-\alpha}\, (1-v)^{\frac{\alpha}{2}-1}\, \mathrm{d}v. \tag{2.23}$$

Remark 2.9. Formula (2.23) defines $q_\alpha(y;c)$ when $y > 0$, and when $y = 0$ the definition is to be understood as the limit $q_\alpha(0;c) := \lim_{y\to 0} q_\alpha(y;c) = 1$ (as can be verified by calculus). Moreover, there is continuity as $\alpha \uparrow 2$ in the sense that $\lim_{\alpha\to 2} q_\alpha(y,c) = 1 - y$ to match with (2.21). See Lemma 2.11 below for proofs of these properties.

We defer the proof of Lemma 2.8 until the end of this section. First, we need some technical results on the equicontinuity of the $q_{\alpha,N}(y;c)$ appearing in Lemma 2.8, as well as the corresponding quantities that appear in Lemma 2.6. In the latter case, we need a little more notation. Similarly to $q_{\alpha,N}(y;c)$, we can define for every $y \in [0,1]$, the exit probabilities

$$p_{\alpha,N}\Big(\frac{k}{N}\Big) := \mathbf{P}_k\big(S_{\eta_{[0,N]}} < 0\big), \text{ for } k \in \{1,2,\dots,N\}.$$

For general $\frac{k}{N} < y < \frac{k+1}{N}$, we define $p_{\alpha,N}(y)$ via linear interpolation, similarly to $q_{\alpha,N}(y;c)$. The equicontinuity results that we need are as follows.

Lemma 2.10 (Kesten 1961 [31])**.** *Suppose that* (I) *holds. For* $1 < \alpha \le 2$*, suppose in addition that (if* $\alpha = 2$*)* $\mathbb{E}(X^2) < \infty$ *and* $\mathbb{E}\, X = 0$*, or (if* $1 < \alpha < 2$*) that* (S_α) *holds. Fix* $0 \le \lambda < 1$ *and* $c > 0$*. The following hold.*

(i) The family of functions $(p_{\alpha,N}(y))_{N\in\mathbb{N}}$ *of* $y \in [0,\lambda]$ *is uniformly equicontinuous.*

(ii) The family of functions $(q_{\alpha,N}(y;c))_{N\in\mathbb{N}}$ *of* $y \in [0,\lambda]$ *is uniformly equicontinuous.*

Part (ii) is available explicitly in Kesten [30, 31]; we give a proof of part (i), using similar ideas, in Appendix B. To exemplify the usefulness of Lemma 2.10, we state two of its consequences which extend the convergence stated in Lemma 2.8.

First, suppose that $y_N \in [0,1)$ is a sequence such that $\lim_{N\to\infty} y_N = y \in [0,1)$. Take $\lambda \in (y,1)$. Then uniform equicontinuity means that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|q_{\alpha,N}(y;c) - q_{\alpha,N}(y';c)| \le \varepsilon$ whenever $y, y' \in [0,\lambda]$ and $|y'-y| \le \delta$. In particular, for all N large enough, we have $|y - y_N| \le \delta$. Consequently, equicontinuity extends the convergence in (2.21) and (2.22) to

$$\lim_{N\to\infty} q_{\alpha,N}(y_N;c) = q_\alpha(y;c), \tag{2.24}$$

whenever $y_N \to y \in [0,1)$ and the relevant hypotheses from Lemma 2.8 hold.

Here is a second consequence. For fixed α, c, the family $(q_{\alpha,N}(y;c))_{N\in\mathbb{N}}$ is uniformly equicontinuous, as functions of $y \in A$ for any compact $A \subset [0,1)$. Hence, by (2.22), $q_{\alpha,N}(y;c)$ converges uniformly as $N \to \infty$ to $q_\alpha(y;c)$, as functions of $y \in A$. In particular, for every compact $A \subset [0,1)$,

$$\lim_{N\to\infty} \inf_{y\in A} q_{\alpha,N}(y;c) = \inf_{y\in A} q_\alpha(y;c). \tag{2.25}$$

We state (2.24) and (2.25) for $q_{\alpha,N}$; analogous statements for $p_{\alpha,N}$ are deduced in the same way.

We will need the following bounds on $q_\alpha(y;c)$; in the proof we make a first use of the equicontinuity from Lemma 2.10.

Lemma 2.11. *Suppose that* (I) *holds, and that* (S_α) *holds with* $\alpha \in (1,2)$*. Let* $q_\alpha(y;c)$ *be as defined in* (2.23)*. Then it holds that, for every* $c > 0$ *and all* $0 \le y < 1$*,*

$$q_\alpha(y;c) \ge (\alpha-1)c^{1-\frac{\alpha}{2}}(c+y)^{\frac{\alpha}{2}-1}(1-y), \tag{2.26}$$

and, moreover, for every $c > 0$,

$$\lim_{y \to 0} q_\alpha(y; c) = 1. \tag{2.27}$$

On the other hand, for every $\delta > 0$ it holds that

$$\sup_{\delta \le y \le 1} q_\alpha(y; c) < 1. \tag{2.28}$$

Proof. Similarly to (2.23), define

$$u_\alpha(y; c) := (\alpha - 1)\, c^{1-\frac{\alpha}{2}}\, (1 + c)^{\alpha - 1}\, (y + c)^{\frac{\alpha}{2}}\, y^{\alpha - 1} \int_y^1 (y + cv)^{-\alpha}\, \mathrm{d}v. \tag{2.29}$$

Some calculus shows that

$$\int_y^1 (y + cv)^{-\alpha}\, \mathrm{d}v = \frac{y^{1-\alpha}}{c(\alpha - 1)} \left[(1 + c)^{1-\alpha} - \left(1 + \frac{c}{y}\right)^{1-\alpha} \right].$$

Hence, by (2.29), we obtain

$$u_\alpha(y; c) = \left(\frac{c + y}{c}\right)^{\frac{\alpha}{2}} \left[1 - \left(\frac{1 + c}{1 + (c/y)} \right)^{\alpha - 1} \right]. \tag{2.30}$$

Since $0 < \alpha/2 < 1$, for all $v \in [0, 1)$ it holds that $(1 - v)^{(\alpha/2)-1} \ge 1$. Comparison of (2.23) and (2.29) then shows that $q_\alpha(y; c) \ge u_\alpha(y; c)$ for all $y \in [0, 1]$. Using the fact that $(1 - x)^{\alpha - 1} \le 1 - (\alpha - 1)x$ for $\alpha \in [1, 2]$ and $x \ge 0$, we have that

$$\left(\frac{1 + c}{1 + (c/y)} \right)^{\alpha - 1} = \left(1 - \frac{c(1 - y)}{c + y} \right)^{\alpha - 1} \le 1 - (\alpha - 1) \left(\frac{c(1 - y)}{c + y} \right).$$

Combining the last bound with (2.30) yields the bound in (2.26). Moreover, since for $0 < y \le 1$ the first factor in the expression on the right-hand side of (2.30) is at least 1, and the second factor (in square brackets) is non-negative, it also follows that

$$\liminf_{y \to 0} q_\alpha(y; c) \ge \liminf_{y \to 0} u_\alpha(y; c) \ge 1 - \limsup_{y \to 0} \left(\frac{1 + c}{1 + (c/y)} \right)^{\alpha - 1} = 1,$$

for every $c > 0$ and $\alpha \in (1, 2)$. This yields (2.27), since $0 \le q_\alpha(y; c) \le 1$.

Finally, we obtain the bound (2.28) by an application of Kesten's gambler's ruin estimate (2.17). Observe that, for $c \ge 0$ and $k \in \mathbb{N}$,

$$\mathbf{P}_k\big(T_0 < \eta_{[-cN,N]}\big) \le \mathbf{P}_k\big(S_{\eta_{[1,N]}} < 1\big) = 1 - \mathbf{P}_k\big(S_{\eta_{[1,N]}} > N\big),$$

since in order to visit 0 before exiting the interval $[-cN, N]$, the walk must exit the (smaller) interval $[1, N]$ on the left. In particular, for $y \in [0, 1)$ and a sequence $k_N \in \mathbb{N}$ such that $\lim_{N\to\infty} k_N/N = y$, then using equicontinuity via (2.24),

$$q_\alpha(y; c) = \lim_{N \to \infty} \mathbf{P}_{k_N} \left(T_0 < \eta_{[-cN,N]} \right).$$

Suppose $y \in [0, 1)$ is rational; then we can choose the sequence k_N such that $k_N = yN$ for a subsequence of N. Hence

$$q_\alpha(y; c) \le 1 - \liminf_{N \to \infty} \mathbf{P}_{yN} \left(S_{\eta_{[1,N]}} > N \right)$$

$$
\begin{aligned}
&= 1 - \liminf_{N\to\infty} \mathbf{P}_0 \left(S_{\eta_{[1-yN,(1-y)N]}} > (1-y)N \right) \\
&\le 1 - \liminf_{N\to\infty} \mathbf{P}_0 \left(S_{\eta_{[-c'N,N]}} > N \right),
\end{aligned}
$$

for every $c' < c_y$ where $c_y := \frac{y}{1-y} \in (0,\infty)$. Then $(1+c_y)^{-1} = 1-y$ for $y \in [0,1]$, and hence, from (2.17), for every rational $y \in [\delta, 1]$,

$$
\begin{aligned}
q_\alpha(y;c) &\le 1 - \inf_{\delta\le y\le 1} \frac{\Gamma(\alpha)}{\Gamma(\alpha/2)^2} \int_{1-y}^1 u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1}\mathrm{d}u \\
&= 1 - \frac{\Gamma(\alpha)}{\Gamma(\alpha/2)^2} \int_{1-\delta}^1 u^{\frac{\alpha}{2}-1}(1-u)^{\frac{\alpha}{2}-1}\mathrm{d}u,
\end{aligned}
$$

which yields (2.28), using continuity of $y \mapsto q_\alpha(y;c)$ over $[\delta,1]$. ☐

We conclude this section with a proof of Lemma 2.8. Here and subsequently, we write $\mathcal{F}_n := \sigma(S_0, S_1, \ldots, S_n)$ for the σ-algebra ($\mathcal{F}_n \subseteq \mathcal{F}'$) generated by the first n steps of the random walk.

Proof of Lemma 2.8. Part (ii) is Theorem 2 in [31, p. 277]. Part (i) is not explicitly stated in [30,31], but can be deduced from results therein, as we now demonstrate.

Suppose that $0 < y < 1$, and set $\lambda := \inf\{n \in \mathbb{Z}_+ : S_n < 0\}$, which is a stopping time with respect to filtration $\mathcal{F}_n$. Under either of the hypotheses of the lemma, we have $\mathbb{E}\, X = 0$ and hence $\mathbf{P}_k(\lambda < \infty) = 1$ for every $k \in \mathbb{Z}$, by recurrence. Suppose that $k_N \in \mathbb{N}$ is a sequence such that $k_N/N \to y$ as $N \to \infty$, and consider events

$$
\begin{aligned}
E_1 &:= E_1(N) := \{S_{\eta_{[0,N]}} < 0\}, \\
E_2 &:= E_2(A) := \{S_\lambda \in [-A, 0]\}, \\
E_3 &:= E_3(N,c) := \{T_0 < \eta_{[-cN,N]}\},
\end{aligned}
$$

and note that both E_1 and E_2 are $\mathcal{F}_\lambda$-measurable. By equicontinuity, similarly to (2.24) but applied to p, we have that

$$
\lim_{N\to\infty} \mathbf{P}_{k_N}(E_1) = 1 - \lim_{N\to\infty} \mathbf{P}_{k_N}\left(S_{\eta_{[0,N]}} > N\right) = \lim_{N\to\infty} p_{2,N}(y).
$$

For rational y, there is a sequence of N_m, $m \in \mathbb{N}$, for which $yN_m \in \mathbb{N}$, and then

$$
\begin{aligned}
\lim_{N\to\infty} p_{2,N}(y) = \lim_{m\to\infty} \mathbf{P}_{yN_m}(S_{\eta_{[0,N_m]}} < 0) &= 1 - \lim_{m\to\infty} \mathbf{P}_0(S_{\eta_{[-yN_m,(1-y)N_m]}} > (1-y)N_m) \\
&= 1 - \lim_{N\to\infty} \mathbf{P}_0(S_{\eta_{[-c_yN,N]}} > N) = 1 - y,
\end{aligned}
$$

where $c_y := \frac{y}{1-y}$ and we have used Lemma 2.6(i) for the convergence. In other words, $\lim_{N\to\infty} p_{2,N}(y) = 1 - y$ for all rational y. Moreover, we have from Proposition 2.2 that, since $y > 0$, for every $\varepsilon > 0$, there exists $A < \infty$ such that $\lim_{N\to\infty} \mathbf{P}_{k_N}(E_2) \ge 1 - \varepsilon$. Started from a site in $[-A, 0]$, the probability that the walk visits 0 before exit from $[-cN, N]$ tends to 1, by Lemma 2.5, i.e., on event E_2,

$$
\mathbf{P}(E_3 \mid \mathcal{F}_\lambda) \ge \inf_{z\in[-A,0]} \mathbf{P}_z(T_0 < \eta_{[-cN,N]}) \to 1,
$$

as $N \to \infty$. Hence, given $\varepsilon > 0$ we can choose A large enough and then N large enough so that $\mathbf{P}(E_3 \mid \mathcal{F}_\lambda) \ge 1 - \varepsilon$ on E_2, that $\mathbf{P}_{k_N}(E_1) \ge 1 - y - \varepsilon$, and that $\mathbf{P}_{k_N}(E_2) \ge 1 - \varepsilon$. Then, by the strong Markov property at time λ,

$$
\mathbf{P}_{k_N}(T_0 < \eta_{[-cN,N]}) \ge \mathbf{E}_{k_N}\left[\mathbf{P}(E_3 \mid \mathcal{F}_\lambda) \mathbb{1}_{E_1 \cap E_2}\right]
$$

$$\geq (1-\varepsilon)\,\mathbf{P}_{k_N}(E_1) - \mathbf{P}_{k_N}(E_2^{\mathrm{c}})$$
$$\geq (1-\varepsilon)(1-y-\varepsilon) - \varepsilon \geq 1 - y - 3\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we deduce that $\liminf_{N\to\infty} \mathbf{P}_{k_N}(T_0 < \eta_{[-cN,N]}) \geq 1-y$. Combined with a similar argument in the other direction, we verify (2.21), with the restriction that $y > 0$, but this restriction is easily removed by a continuity argument, as in Kesten [31, p. 275]. □

3 Proofs of the bounds on cluster growth

3.1 Lower bounds: Overview and some heuristics

Recall that r_m defined at (1.3) is the maximal $r \in \mathbb{Z}_+$ such that $[-r, r] \cap \mathbb{Z}$ is contained in $\mathfrak{C}_m$. The purpose of this section is to study the asymptotics of r_m. In Section 3.3, we prove (1.5) of Theorem 1.3, which covers the case where increments have finite variance. To prove (1.7) of Theorem 1.6, the infinite variance case, we prove the lower bound in Section 3.4 and the upper bound in Section 3.5.

In view of the deterministic bound $\limsup_{m\to\infty} r_m/m \leq 1/2$, a.s., from (1.4), to prove (1.5) it is sufficient to prove that $\liminf_{m\to\infty} r_m/m \geq 1/2$, a.s. It turns out to be more convenient at this point to work with the coverage times

$$\sigma_x := \inf\{m \in \mathbb{Z}_+ : r_m \geq x\}. \tag{3.1}$$

The sequences r_m and σ_x are related by the inversion:

$$\text{It holds that } r_m \geq x \text{ if and only if } \sigma_x \leq m. \tag{3.2}$$

In particular, for $c \in (0, \infty)$, equivalent to the statement $\liminf_{m\to\infty}(r_m/m) \geq c$, a.s., is the statement $\limsup_{x\to\infty}(\sigma_x/x) \leq 1/c$, a.s. Hence to prove (1.5) in Theorem 1.3, it is enough to prove the following.

Proposition 3.1. *Suppose that* (I) *holds,* $\mathbb{E}(X^2) < \infty$*, and* $\mathbb{E}\, X = 0$*. Then, a.s.,*

$$\limsup_{x\to\infty} \frac{\sigma_x}{x} \leq 2. \tag{3.3}$$

In the finite-variance case, the trivial bound $\sigma_x/x \geq 2$ that follows from (1.4) means that to prove Theorem 1.3, the upper bound (3.3) is sufficient. In the infinite-variance case, for Theorem 1.6 we need not only an upper bound, presented in Proposition 3.2 that follows, but also non-trivial lower bounds, which are the subject of Section 3.5. Define the constant

$$C'_\alpha := (\alpha-1)^{-1}(4-\alpha)(3-\alpha)^3(2-\alpha)^{\alpha-2}. \tag{3.4}$$

Proposition 3.2. *Suppose that* (S_α) *holds, and that* (S_α) *holds with* $\alpha \in (1,2)$*. Then, for* $C'_\alpha \in (2,\infty)$ *given by* (3.4)*, it holds that, a.s.,*

$$\limsup_{x\to\infty} \frac{\sigma_x}{x} \leq C'_\alpha. \tag{3.5}$$

The outline of the proofs of the upper bounds in both Propositions 3.1 and 3.2, has similarities to the corresponding argument in [34, §3], with the estimates of Section 2.2 (derived from Kesten [30, 31]) providing the central probabilistic components.

Recall from Section 1.2 that $S^{(j)}$ denotes the jth random walk in the IDLA process, and that τ_j, defined at (1.2), is the internal time index of the walk $S^{(j)}$ when it first exits the prior cluster $\mathfrak{C}_{j-1}$. In this section we develop several arguments using the sequence of random walks $S^{(j)}$. We note that the IDLA process depends only on each $S^{(j)}$ up to its associated time τ_j, but in fact the walk $S^{(j)}$ is defined for all time, and several arguments make use of this, for example, to overcome certain dependence. In arguments that emphasize the finite-walk perspective, we sometimes describe walk $S^{(j)}$ as being *active* up until time τ_j, and then *terminating*, while in arguments that use the full trajectory of the walk, we sometimes refer to the walk $S^{(j)}$ as *indefinitely extended*.

Let us describe, informally, how the results of Section 2.2 enable us to understand the behaviour of σ_x. Consider time σ_x, so that the interval $[-x, x]$ is fully occupied by the cluster $\mathfrak{C}_{\sigma_x}$. We fix $u > 1$ and study the IDLA process up to the time at which the interval $[-ux, ux]$ is fully occupied, in order to bound from above $\sigma_{ux} - \sigma_x$. We must argue that unoccupied sites are filled rather rapidly by subsequent walkers. The outline of the argument is as follows.

1. The gambler's ruin estimates of Lemma 2.6 show that any subsequent walker will exit $[-x, x]$ on the right or left each with probability almost $1/2$, and, since $[-x, x]$ is fully occupied, the walker will still be active when it does so. This is the case in both the settings of Propositions 3.1 and 3.2.

2a. Consider some site $t \in [x, ux] \cap \mathbb{Z}$. In the finite-variance case (Proposition 3.1) by tightness of overshoots (Proposition 2.2(ii)) and recurrence (Lemma 2.5), a random walk that exits $[-x, x]$ on the right, will, with probability close to 1 (if $u \approx 1$) visit t before exiting from the interval $[-x, sx]$, where $s > u$. Hence (using point 1 above) the probability that the walk exits $[-x, x]$ on the right *and* visits t before exit from $[-x, sx]$ is approximately $q \approx 1/2$.

2b. Consider some site $t \in [x, ux] \cap \mathbb{Z}$. In the infinite-variance case (Proposition 3.2) a random walk that exits $[-x, x]$ on the right will visit t before exiting from the interval $[-x, sx]$, where $s > u$, with a probability bounded below by some strictly positive $q = q_\alpha(u, s) \in (0, 1/2)$, by Lemma 2.8 and the bounds in Lemma 2.11.

3. To obtain statements that hold with very high probability, from the walk-by-walk probability statements in 2a and 2b, we consider the $k = Ax$ walkers directly after time σ_x and exploit binomial concentration. The number of these walkers that exit $[-x, x]$ on the right and then visit $t \in [x, ux] \cap \mathbb{Z}$ before exit from the interval $[-x, sx]$ is binomial, and, by binomial concentration, we can, by balancing small constants, ensure is, with very high probability, at least about Aqx.

4. There are at most $(s-1)x$ sites of $\mathbb{Z}$ in the interval $[x, sx]$, and each can accommodate at most one walker adding to the aggregate, so if $Aq > s - 1$, then at least one walker is active when it visits t. Hence t is contained in the cluster at time about $\sigma_x + Ax$, where the optimal A is $A \approx (s-1)/q$. (See Lemma 3.4 for a precise version of this argument.)

5. The binomial concentration is sufficient to show that this occurs with high probability for all $t \in [-ux, ux] \cap \mathbb{Z}$, i.e., with high probability by time $\approx \sigma_x + (s-1)x/q$ we cover $[-ux, ux]$. This shows $\sigma_{ux} - \sigma_x \leq (s-1)x/q$, with high probability. With

an interpolation argument, this translates to a bound of the form

$$\limsup_{x\to\infty} \frac{\sigma_x}{x} \le \frac{1}{q}\left(\frac{s-1}{u-1}\right) u, \text{ a.s.;}$$

see (3.25) (for the finite variance case) and (3.32) (infinite variance) below. Choosing constants carefully, in the case of Proposition 3.1 we can take $s \approx u \approx 1$ and $q \approx 1/2$, giving the constant 2 in the bound (3.3). In the setting of Proposition 3.2, choosing $s > u > 1$ appropriately leads to the upper bound in (3.5) with the constant C'_α from (3.4).

3.2 Preliminaries

Analogously to the notation in (2.15), for the random walk $S^{(j)}$, $j \in \mathbb{N}$, define the first hitting time $T_t^{(j)}$ of $t \in \mathbb{Z}$ by

$$T_t^{(j)} := \inf\{n \in \mathbb{Z}_+ : S_n^{(j)} = t\}, \tag{3.6}$$

and the first exit time $\eta_A^{(j)}$ from the set $A \subseteq \mathbb{Z}$ by

$$\eta_A^{(j)} := \inf\{n \in \mathbb{Z}_+ : S_n^{(j)} \notin A\}. \tag{3.7}$$

Note that, τ_m as defined at (1.2) has the representation $\tau_m = \eta^{(m)}_{\mathfrak{C}_{m-1}}$ in the notation at (3.7). For $t \in \mathbb{Z}$ and $x \in \mathbb{N}$, define

$$N^+_{m,k}(t;x) := \sum_{j=m+1}^{m+k} \mathbb{1}\{S^{(j)}_{\eta^{(j)}_{[-x,x]}} > x,\, T_t^{(j)} \le \eta^{(j)}_{\mathfrak{C}_{j-1}}\}, \tag{3.8}$$

the number of the next k random walks, released after time m, that (i) exit the interval $[-x, x]$ on the right, and (ii) visit t before they exit the contemporary cluster. Similarly, define

$$N^-_{m,k}(t;x) := \sum_{j=m+1}^{m+k} \mathbb{1}\{S^{(j)}_{\eta^{(j)}_{[-x,x]}} < -x,\, T_t^{(j)} \le \eta^{(j)}_{\mathfrak{C}_{j-1}}\}. \tag{3.9}$$

We will always be assuming (I), so that the random walk S is irreducible, and hence $\eta^{(j)}_{[-x,x]} < \infty$ for every j and every x, almost surely. Note that the sum $N^+_{m,k}(t;x) + N^-_{m,k}(t;x)$ does not depend on x and hence we may define

$$N_{m,k}(t) := \sum_{j=m+1}^{m+k} \mathbb{1}\{T_t^{(j)} \le \eta^{(j)}_{\mathfrak{C}_{j-1}}\} = N^+_{m,k}(t;x) + N^-_{m,k}(t;x). \tag{3.10}$$

Remark 3.3. The definitions of $N_{m,k}(t)$, $N^{\pm}_{m,k}(t;x)$ are motivated from similar quantities defined in [34], where the authors also make use of the property (3.11) below.

Observe that, if there exists a (smallest) j for which $T_t^{(j)} \le \eta^{(j)}_{\mathfrak{C}_{j-1}}$, then $t \in \mathfrak{C}_j$. In particular,

$$\{t \in \mathfrak{C}_{m+k},\, t \notin \mathfrak{C}_m\} = \{N_{m,k}(t) > 0\}, \text{ and } \mathbb{P}(t \notin \mathfrak{C}_{m+k} \setminus \mathfrak{C}_m) = \mathbb{P}(N_{m,k}(t) = 0). \tag{3.11}$$

It turns out to be convenient to work, instead of with the quantities defined in (3.8)–(3.10), with some related quantities, for the indefinitely extended walks, that dispense with the dependence on the contemporary cluster and so enjoy greater independence structure. Define, for $s > 1$,

$$K^{+}_{m,k}(t,x,s) := \sum_{j=m+1}^{m+k} \mathbb{1}\{S^{(j)}_{\eta^{(j)}_{[-x,x]}} > x,\, T^{(j)}_t < \eta^{(j)}_{[-x,sx]}\}, \tag{3.12}$$

the number of the next k random walks, released after time m, that (i) exit the interval $[-x,x]$ on the right and (ii) visit t before they exit $[-x,sx]$. The events in the indicators in (3.12) are i.i.d., so that

$$K^{+}_{m,k}(t,x,s) \sim \mathrm{Bin}\left(k, p^{+}(t,x,s)\right), \tag{3.13}$$

where

$$p^{+}(t,x,s) := \mathbf{P}_0\big(S_{\eta_{[-x,x]}} > x,\, T_t < \eta_{[-x,sx]}\big). \tag{3.14}$$

Similarly, let

$$K^{-}_{m,k}(t,x,s) := \sum_{j=m+1}^{m+k} \mathbb{1}\{S^{(j)}_{\eta^{(j)}_{[-x,x]}} < -x,\, T^{(j)}_t < \eta^{(j)}_{[-sx,x]}\}, \tag{3.15}$$

be the number of the next k random walks, released after time m, that (i) exit the interval $[-x,x]$ on the left and (ii) visit t before they exit $[-sx,x]$. The events in the indicators in (3.15) are i.i.d., so that

$$K^{-}_{m,k}(t,x,s) \sim \mathrm{Bin}\left(k, p^{-}(t,x,s)\right), \tag{3.16}$$

where

$$p^{-}(t,x,s) := \mathbf{P}_0\big(S_{\eta_{[-x,x]}} < -x,\, T_t < \eta_{[-sx,x]}\big). \tag{3.17}$$

The following is our basic tool for obtaining upper bounds on σ_x.

Lemma 3.4. *Fix $s > u > 1$. If, for $x, k \in \mathbb{N}$, it holds that*

$$\min_{t\in\mathbb{Z}\cap(x,ux]} \min\left(K^{+}_{\sigma_x,k}(t,x,s), K^{-}_{\sigma_x,k}(-t,x,s)\right) \geq \lceil (s-1)x \rceil, \tag{3.18}$$

then $\sigma_{ux} \leq \sigma_x + k$.

Proof. Consider $t \in \mathbb{Z} \cap (x, ux]$. If $K^{+}_{\sigma_x,k}(t,x,s) \geq \lceil (s-1)x \rceil$, then at least $\lceil (s-1)x \rceil$ (indefinitely extended) random walks, released after time σ_x, visit site t before exiting $[-x, sx]$. The set $[-x,sx] \setminus [-x,x]$ contains no more than $(s-1)x$ sites of $\mathbb{Z}$, and so not all of the $\lceil (s-1)x \rceil$ particles can terminate before reaching t. Hence

$$K^{+}_{\sigma_x,k}(t,x,s) \geq \lceil (s-1)x \rceil \text{ implies that } t \in \mathfrak{C}_{\sigma_x+k}.$$

Similarly, $K^{-}_{\sigma_x,k}(-t,x,s) \geq \lceil (s-1)x \rceil$ implies that $-t \in \mathfrak{C}_{\sigma_x+k}$. It follows that (3.18) implies that $t \in \mathfrak{C}_{\sigma_x+k}$ for all $t \in [-ux, ux] \cap \mathbb{Z}$, and hence $\sigma_{ux} \leq \sigma_x + k$. $\square$

3.3 Finite variance

To apply Lemma 3.4, we need to choose $k = k_x$, depending on x, in such a way that the event in (3.18) occurs with high probability. To do so, we will use binomial concentration applied to (3.13) and (3.16), and hence we need to quantify the asymptotics of p^+ and p^- defined at (3.14) and (3.17), respectively. We achieve this through results which make the heuristic ideas in Section 3.1 formal, and will enable us to prove Proposition 3.1 and hence Theorem 1.3.

Lemma 3.5. *Suppose that* (I) *holds,* $\mathbb{E}(X^2) < \infty$*, and* $\mathbb{E}\, X = 0$*. Then for every* $\varepsilon > 0$*, there is a* $u_\varepsilon > 1$ *such that the following holds. For every* $u \in (1, u_\varepsilon)$ *and every* $s > u$*,*

$$\lim_{x\to\infty} \sup_{t\in\mathbb{Z}\cap[x,ux]} \left| p^{\pm}(t, x, s) - \frac{1}{2} \right| < \varepsilon. \tag{3.19}$$

Proof. First observe that, by the gambler's ruin asymptotics in Lemma 2.6(i),

$$\lim_{x\to\infty} \mathbf{P}_0\big(S_{\eta_{[-x,x]}} > x\big) = 1/2. \tag{3.20}$$

Consequently, for every $s > 1$, we have the upper bound

$$\limsup_{x\to\infty} \sup_{t\in\mathbb{Z}} \mathbf{P}_0\big(S_{\eta_{[-x,x]}} > x,\, T_t < \eta_{[-x,sx]}\big) \le \frac{1}{2}. \tag{3.21}$$

To obtain a lower bound, consider $t \in \mathbb{Z} \cap [x, ux]$, and write, for $B \in \mathbb{R}_+$,

$$\mathbf{P}_0\big(S_{\eta_{[-x,x]}} > x,\, T_t < \eta_{[-x,sx]}\big) \ge \mathbf{E}_0\Big[\mathbf{P}_0\big(T_t < \eta_{[-x,sx]} \;\big|\; \mathcal{F}_{\eta_{[-x,t]}}\big) \mathbb{1}\{t < S_{\eta_{[-x,t]}} < t + B\}\Big],$$

where, as in Section 2.2, $\mathcal{F}_n = \sigma(S_0, S_1, \ldots, S_n)$, with respect to which $\eta_{[-x,t]}$ is a stopping time. By Lemma 2.2(ii) (tightness of the overshoots) we have that, for every $\varepsilon > 0$, we may choose B large enough so that $\sup_{t\in\mathbb{Z}} \mathbf{P}_0(S_{\rho_t} \ge t + B) \le \varepsilon$; fix $\varepsilon > 0$ and such a B. By the strong Markov property, it holds that, on $\{t < S_{\eta_{[-x,t]}} < t + B\}$,

$$\begin{aligned}
\min_{t\in[x,ux]} \mathbf{P}_0\big(T_t < \eta_{[-x,sx]} \;\big|\; \mathcal{F}_{\eta_{[-x,t]}}\big) &\ge \min_{t\in[x,ux]} \min_{y\in[t,t+B]\cap\mathbb{Z}} \mathbf{P}_y\big(T_t < \eta_{[-x,sx]}\big) \\
&\ge \min_{t\in[x,ux]} \min_{y\in[0,B]\cap\mathbb{Z}} \mathbf{P}_y\big(T_0 < \eta_{[-x-t,sx-t]}\big) \\
&\ge \min_{y\in[0,B]\cap\mathbb{Z}} \mathbf{P}_y\big(T_0 < \eta_{[-x,(s-u)x]}\big),
\end{aligned}$$

since $\eta_{[-x,(s-u)x]} \le \eta_{[-x-t,sx-t]}$ for every $t \in [x, ux]$. This last bound tends to 1 as $x \to \infty$, provided $s > u > 1$, by the local hitting property, Lemma 2.5. Hence for every $\varepsilon > 0$, and every $s > u > 1$, for all x large enough

$$\min_{t\in[x,ux]} \mathbf{P}_0\big(S_{\eta_{[-x,x]}} > x,\, T_t < \eta_{[-x,sx]}\big) \ge \min_{t\in[x,ux]} \mathbf{P}_0\big(t < S_{\eta_{[-x,t]}} < t + B\big) - \varepsilon.$$

Recalling from (2.5) and (2.15) that $\rho_t = \eta_{(-\infty,t]}$, we have

$$\begin{aligned}
\mathbf{P}_0\big(t < S_{\eta_{[-x,t]}} < t + B\big) &= \mathbf{P}_0\big(S_{\eta_{[-x,t]}} > t,\, S_{\rho_t} < t + B\big) \\
&\ge \mathbf{P}_0\big(S_{\eta_{[-x,t]}} > t\big) - \mathbf{P}_0(S_{\rho_t} \ge t + B) \\
&\ge \mathbf{P}_0\big(S_{\eta_{[-x,t]}} > t\big) - \varepsilon,
\end{aligned}$$

by choice of B. Since $\{S_{\eta_{[-x,t+1]}} \geq t+1\} \subseteq \{S_{\eta_{[-x,t]}} \geq t\}$ for every $t \in \mathbb{Z}_+$, another application of the gambler's ruin asymptotics in Lemma 2.6(i) shows that, for every $\varepsilon > 0$ and $u > 1$, for all x large enough,

$$\min_{t\in[x,ux]} \mathbf{P}_0\big(S_{\eta_{[-x,t]}} > t\big) \geq \mathbf{P}_0\big(S_{\eta_{[-x,ux]}} > ux\big) \geq \frac{1}{1+u} - \varepsilon.$$

Thus we conclude that for every $\varepsilon > 0$ and $s > u > 1$, for all x large enough,

$$\min_{t\in[x,ux]} \mathbf{P}_0\big(S_{\eta_{[-x,x]}} > x,\, T_t < \eta_{[-x,sx]}\big) \geq \frac{1}{1+u} - \varepsilon.$$

In particular, we can choose $u > 1$ close enough to 1 so that this last probability is arbitrarily close to $1/2$, which combines with (3.21) to conclude the proof of the statement for p^+ in (3.19). The proof of the statement for p^- is analogous. ☐

The following high-probability statements are obtained from the probability bounds in Lemma 3.5 via binomial concentration.

Lemma 3.6. *Suppose that* (I) *holds,* $\mathbb{E}(X^2) < \infty$*, and* $\mathbb{E}\,X = 0$*. There exists* $\varepsilon_0 \in (0,1)$ *such that the following holds. Take* $\varepsilon \in (0,\varepsilon_0)$*, and let* $u_\varepsilon > 1$ *be as given in Lemma 3.5. Then, for every* $u \in (1,u_\varepsilon)$ *and every* $s > u$*, there exist* $c > 0$ *(depending on* ε*) and* $x_0 > 0$ *such that, for all* $x > x_0$*, with* $k_x := \lceil (2+7\varepsilon)(s-1)x \rceil$*,*

$$\max_{t\in\mathbb{Z}\cap[x,ux]} \mathbb{P}\left(K^+_{\sigma_x,k_x}(t,x,s) < (s-1)x\right) \leq \mathrm{e}^{-c(s-1)x}, \tag{3.22}$$

and

$$\max_{t\in\mathbb{Z}\cap[x,ux]} \mathbb{P}\left(K^-_{\sigma_x,k_x}(-t,x,s) < (s-1)x\right) \leq \mathrm{e}^{-c(s-1)x}. \tag{3.23}$$

Proof. Standard Chernoff bounds for binomial large deviations (see e.g. [40, p. 16]) say that if $X \sim \mathrm{Bin}(n,p)$ and $a \in (0,1)$, then $\mathbb{P}(X < anp) \leq \exp(-c_a np)$, where $c_a > 0$. Recall from (3.13) that that $K^+_{\sigma_x,k}(t,x,s)$ follows a binomial distribution with mean $kp^+(t,x,s)$. Choose $k_x = \lceil (2+7\varepsilon)(s-1)x \rceil$. For $u \in (1,u_\varepsilon)$, $s > u$, and x large enough, Lemma 3.5 shows that, for all $t \in \mathbb{Z}\cap[x,ux]$,

$$k_x p^+(t,x,s) \geq (2+7\varepsilon)(s-1)x \cdot \left(\frac{1}{2} - \varepsilon\right) > (1+\varepsilon)(s-1)x,$$

for all $\varepsilon \in (0,\varepsilon_0)$ sufficiently small. The binomial Chernoff bound stated above then yields (3.22), where $c > 0$ depends only on ε. A similar argument yields (3.23). ☐

Proof of Proposition 3.1. Let $\varepsilon_0 > 0$ be the constant from Lemma 3.6. Take $\varepsilon \in (0,\varepsilon_0)$, and let $u_\varepsilon > 1$ be as given in Lemma 3.5. It follows from Lemma 3.6 that, for every $u \in (1,u_\varepsilon)$ and every $s > u$, there exist $c > 0$ (depending on ε) and $x_0 > 0$ such that, for all $x > x_0$, with $k_x = \lceil (2+7\varepsilon)(s-1)x \rceil$,

$$\begin{aligned}
&\mathbb{P}\left(\min_{t\in\mathbb{Z}\cap(x,ux]} \min\left(K^+_{\sigma_x,k_x}(t,x,s), K^-_{\sigma_x,k_x}(-t,x,s)\right) < \lceil (s-1)x \rceil\right)\\
&\quad\leq 2ux \max_{t\in\mathbb{Z}\cap(x,ux]} \max\left\{\mathbb{P}\left(K^+_{\sigma_x,k_x}(t,x,s) < (s-1)x\right), \mathbb{P}\left(K^-_{\sigma_x,k_x}(-t,x,s) < (s-1)x\right)\right\}\\
&\quad\leq 2ux\mathrm{e}^{-c(s-1)x}.
\end{aligned}$$

In particular, we obtain

$$\sum_{x\in\mathbb{N}} \mathbb{P}\left(\min_{t\in\mathbb{Z}\cap(x,ux]} \min\left(K^{+}_{\sigma_x,k_x}(t,x,s), K^{-}_{\sigma_x,k_x}(-t,x,s)\right) < \lceil (s-1)x \rceil \right) < \infty.$$

It follows from Lemma 3.4 and the Borel–Cantelli lemma that, almost surely, for all but finitely many $x \in \mathbb{N}$, (3.18) holds and thus $\sigma_{ux} \le \sigma_x + k_x$. In particular, considering the subsequence $x = u^m$, it follows that there is a (random, a.s. finite) $m_0 \in \mathbb{N}$ such that

$$\sigma_{u^{m+1}} - \sigma_{u^m} \le k_{u^m} \le 1 + (2+7\varepsilon)(s-1)u^m, \text{ for all } m \ge m_0.$$

Consequently, for all $m \ge m_0$,

$$\begin{aligned}
\sigma_{u^m} &= \sigma_{u^{m_0}} + \sum_{\ell=m_0}^{m-1} (\sigma_{u^{\ell+1}} - \sigma_{u^\ell}) \\
&\le \sigma_{u^{m_0}} + m + (2+7\varepsilon)(s-1)\sum_{\ell=0}^{m-1} u^\ell \\
&\le \sigma_{u^{m_0}} + m + (2+7\varepsilon)\left(\frac{s-1}{u-1}\right) u^m. \qquad (3.24)
\end{aligned}$$

For fixed $u \in (1, u_\varepsilon)$ and every $x \in \mathbb{N}$, there exists $m_x \in \mathbb{Z}_+$ such that $u^{m_x} \le x < u^{m_x+1}$; note that $m_x \to \infty$ as $x \to \infty$. Since $\sigma_x \le \sigma_{u^{m_x+1}}$, it follows from (3.24) that, for each $u \in (1, u_\varepsilon)$ and $s > u$,

$$\begin{aligned}
\limsup_{x\to\infty} \frac{\sigma_x}{x} &\le \limsup_{x\to\infty} \frac{\sigma_{u^{m_x+1}}}{u^{m_x}} \\
&\le (2+7\varepsilon)\left(\frac{s-1}{u-1}\right) u + \limsup_{x\to\infty} \frac{\sigma_{u^{m_0}} + m_x + 1}{u^{m_x}} \\
&= (2+7\varepsilon)\left(\frac{s-1}{u-1}\right) u. \qquad (3.25)
\end{aligned}$$

Since, for fixed $\varepsilon \in (0, \varepsilon_0)$, the choices of s and u were arbitrary subject to $s > u$ and $u \in (1, u_\varepsilon)$, it follows from (3.25) that $\limsup_{x\to\infty}(\sigma_x/x) \le 2 + 7\varepsilon$, a.s. Since $\varepsilon \in (0, \varepsilon_0)$ was arbitrary, this completes the proof. □

Proof of Theorem 1.3. As explained in Section 3.1, Theorem 1.3 follows from the bound (1.4) together with the bound from Proposition 3.1 and the inversion described by (3.2). □

3.4 Infinite variance

In the infinite-variance case, the present section deals with the proof of the upper bound given in Proposition 3.2. For Theorem 1.6 we also need a *lower* bound $\sigma_x/x > c$ with $c > 2$, which we establish in Section 3.5 below. This section is concerned with the upper bound on σ_x/x required for Proposition 3.2.

The structure of this section parallels that of Section 3.3. The strategy is to once more apply Lemma 3.4, but now the asymptotics of p^+ and p^- defined in (3.14) and (3.17), respectively, are different. Define

$$\underline{q}_\alpha(u,s) := \left(\frac{\alpha-1}{1+u}\right)\left(\frac{s-u}{s}\right)^{1-\frac{\alpha}{2}}. \qquad (3.26)$$

Lemma 3.7. *Suppose that* (I) *holds, and that* (S_α) *holds with* $\alpha \in (1,2)$*. For every* $u > 1$ *and every* $s > u$*, with* $\underline{q}_\alpha$ *as defined at* (3.26)*, it holds that*

$$\liminf_{x\to\infty} \min_{t\in\mathbb{Z}\cap[x,ux]} p^{\pm}(t,x,s) \geq \underline{q}_\alpha(u,s). \tag{3.27}$$

In what follows, as in Section 3.3, we are free to choose s and u such that $s > u > 1$, and, as before, for a fixed probability lower bound on $p^{\pm}$, the optimal choice would be to take $u \approx 1$ and $s \approx u$. However, the bound in (3.27) has $\underline{q}_\alpha(u,s) \to 0$ as $s - u \to 0$, and in fact this is inevitable in the infinite-variance case, since the overshoot is not tight. Thus we must keep $s-u$ strictly positive, and then (compare (3.25)) one must also keep $u-1$ strictly positive. The balance is then to choose u, s so that $s-u$ and $u-1$ are positive, but not too large. The optimal choice can be found by some calculus, but provides a somewhat complicated formula. As our aim here is not to obtain the optimal constants, but to provide reasonable bounds that capture important asymptotics (such as $\alpha \uparrow 2$ behaviour) we instead will choose $u = 1+h$, $s = u + h^2 = 1 + h + h^2$. Then

$$\underline{q}_\alpha(1+h, 1+h+h^2) = \frac{(\alpha-1)}{(2+h)} \frac{h^{2-\alpha}}{(1+h+h^2)^{1-\alpha/2}}. \tag{3.28}$$

The bound (3.28) has the property that it goes to $1/2$ as $h \downarrow 0$ and $\alpha \uparrow 2$ appropriately.

Proof of Lemma 3.7. Note that for $t \in \mathbb{Z} \cap [x, ux]$ we have

$$p^+(t,x,s) = \mathbb{P}_0(T_t < \eta_{[-x,sx]}) = \mathbb{P}_{-t}(T_0 < \eta_{[-x-t,sx-t]}) = \mathbb{P}_t(T_0 < \eta_{[t-sx,x+t]}),$$

by a change of sign and the symmetry hypothesis in (S_α). Hence, using the definition of (2.19), we obtain

$$p^+(t,x,s) = q_{\alpha,x+t}\left(\frac{t}{x+t}; \frac{sx-t}{x+t}\right).$$

Note that for $t \in [x, ux] \cap \mathbb{Z}$ we have $\frac{sx-t}{x+t} \geq \frac{s-u}{1+u}$ and thus, by the monotonicity property (2.20) we obtain that

$$q_{\alpha,x+t}\left(\frac{t}{x+t}; \frac{sx-t}{x+t}\right) \geq q_{\alpha,x+t}\left(\frac{t}{x+t}; \frac{s-u}{1+u}\right).$$

It follows from (2.24) and (2.25), which are consequences of Kesten's convergence and equicontinuity results as presented in Lemmas 2.8 and 2.10 that

$$\lim_{x\to\infty} \min_{t\in\mathbb{Z}\cap[x,ux]} q_{\alpha,x+t}\left(\frac{t}{x+t}; \frac{s-u}{1+u}\right) = \inf_{y\in[\frac{1}{2},\frac{u}{1+u}]} q_\alpha\left(y; \frac{s-u}{1+u}\right).$$

Using the lower bound from (2.26) and observing that $\frac{\alpha}{2} - 1 < 0$ we get

$$\begin{aligned}
\inf_{y\in[\frac{1}{2},\frac{u}{1+u}]} q_\alpha\left(y; \frac{s-u}{1+u}\right) &\geq (\alpha-1)\left(\frac{s-u}{1+u}\right)^{1-\frac{\alpha}{2}} \inf_{y\in[\frac{1}{2},\frac{u}{1+u}]} \left[\left(y + \frac{s-u}{1+u}\right)^{\frac{\alpha}{2}-1}(1-y)\right] \\
&= (\alpha-1)\left(\frac{s-u}{1+u}\right)^{1-\frac{\alpha}{2}} \left(\frac{s}{1+u}\right)^{\frac{\alpha}{2}-1} \frac{1}{1+u},
\end{aligned}$$

which is equal to $\underline{q}_\alpha(u,s)$ as defined at (3.26). The proof of the statement for p^- is analogous. ☐

The next result will substitute for Lemma 3.6 in the infinite-variance setting.

Lemma 3.8. *Suppose that* (I) *holds and that* (S_α) *holds with* $\alpha \in (1,2)$. *Take* $u, s \in \mathbb{R}$, *such that* $1 < u < s$ *and let* $q := \underline{q}_\alpha(u,s)$ *as given in definition* (3.26). *Then, there exist* $\varepsilon_0 \in (0,1)$, $K > 0$, *and* $x_0 > 0$ *such that the following holds. For all* $x > x_0$, *every* $\varepsilon \in (0, \varepsilon_0)$, *with* $k_x := \lceil (q^{-1} + K\varepsilon)(s-1)x \rceil$, *there is* $c > 0$ *for which*

$$\max_{t \in \mathbb{Z} \cap [x, ux]} \mathbb{P}\left(K^+_{\sigma_x, k_x}(t, x, s) < (s-1)x \right) \leq \mathrm{e}^{-c(s-1)x}, \tag{3.29}$$

and

$$\max_{t \in \mathbb{Z} \cap [x, ux]} \mathbb{P}\left(K^-_{\sigma_x, k_x}(-t, x, s) < (s-1)x \right) \leq \mathrm{e}^{-c(s-1)x}. \tag{3.30}$$

Proof. Given u, s with $1 < u < s$, take $q := \underline{q}_\alpha(u,s) > 0$ where $\underline{q}_\alpha(u,s)$ is defined at (3.26). Then there exist $\varepsilon_0 > 0$ and $K \in \mathbb{N}$ (both depending on q and hence on u and s) such that

$$\left(q^{-1} + K\varepsilon\right) \cdot (q - \varepsilon) > (1 + \varepsilon), \text{ for all } \varepsilon \in (0, \varepsilon_0). \tag{3.31}$$

Define $k_x = \lceil (q^{-1} + K\varepsilon)(s-1)x \rceil$, as in the lemma, for this choice of K. Take $\varepsilon \in (0, \varepsilon_0)$. By Lemma 3.7, it holds that $p^+(t,x,s) > \underline{q}_\alpha(u,s) - \varepsilon$ for all x large enough. Then it follows from the choice of k_x, and property (3.31), that

$$k_x p^+(t,x,s) \geq \left(q^{-1} + K\varepsilon\right)(s-1)x \cdot (q - \varepsilon) > (1+\varepsilon)(s-1)x.$$

The binomial Chernoff bound stated in the proof of Lemma 3.6 then yields (3.29), where $c > 0$ depends only on ε. A similar argument yields (3.30). ☐

Proof of Proposition 3.2. Let $1 < u < s$ and $\varepsilon_0 > 0$ be the constants from Lemma 3.8. It follows from Lemma 3.8 that for every $\varepsilon \in (0, \varepsilon_0)$ there exist $K > 0$, depending on $q = \underline{q}_\alpha(u,s)$, $c > 0$ and $x_0 > 0$, depending on ε, such that, for all $x > x_0$, with $k_x = \lceil (q^{-1} + K\varepsilon)(s-1)x \rceil$,

$$\begin{aligned}
&\mathbb{P}\left(\min_{t \in \mathbb{Z} \cap (x, ux]} \min\left\{ K^+_{\sigma_x, k_x}(t,x,s), K^-_{\sigma_x, k_x}(-t,x,s) \right\} < \lceil (s-1)x \rceil \right) \\
&\leq ux \max_{t \in \mathbb{Z} \cap [x, ux]} \Big(\mathbb{P}\left(K^+_{\sigma_x, k_x}(t,x,s) < (s-1)x \right) + \mathbb{P}\left(K^-_{\sigma_x, k_x}(-t,x,s) < (s-1)x \right) \Big) \\
&\leq 2ux\mathrm{e}^{-c(s-1)x}.
\end{aligned}$$

In particular, we obtain

$$\sum_{x \in \mathbb{N}} \mathbb{P}\left(\min_{t \in \mathbb{Z} \cap (x, ux]} \min\left(K^+_{\sigma_x, k_x}(t,x,s), K^-_{\sigma_x, k_x}(-t,x,s) \right) < \lceil (s-1)x \rceil \right) < \infty.$$

It follows from Lemma 3.4 and the Borel–Cantelli lemma that, almost surely, for all but finitely many $x \in \mathbb{N}$, (3.18) holds and thus $\sigma_{ux} \leq \sigma_x + k_x$. Proceeding as in the proof of Proposition 3.1, but with k_x given in Lemma 3.6 replaced by that in Lemma 3.8, repeating the steps through (3.24)–(3.25), we obtain

$$\limsup_{x \to \infty} \frac{\sigma_x}{x} \leq (q^{-1} + K\varepsilon) \left(\frac{s-1}{u-1} \right) u, \text{ a.s.}$$

Here $\varepsilon \in (0, \varepsilon_0)$ was arbitrary, and K, u, s do not depend on ε, so it holds that

$$\limsup_{x\to\infty} \frac{\sigma_x}{x} \le \frac{1}{\underline{q}_\alpha(u,s)} \left(\frac{s-1}{u-1}\right) u, \text{ a.s.} \tag{3.32}$$

Then using the expression (3.28) together with (3.26), it follows from (3.32) that

$$\limsup_{x\to\infty} \frac{\sigma_x}{x} \le \frac{1+u}{\alpha-1} \left(\frac{s}{s-u}\right)^{1-\frac{\alpha}{2}} \left(\frac{s-1}{u-1}\right) u, \text{ a.s.} \tag{3.33}$$

Now choose $u = 1 + h$ and $s = 1 + h + h^2$ for some $h \in [0, 1]$. Then using the fact that $\alpha \in (1, 2)$ and $h \in [0, 1]$, in the bound (3.33) we get, a.s.,

$$\begin{aligned}
\limsup_{x\to\infty} \frac{\sigma_x}{x} &\le (\alpha-1)^{-1}(2+h)(1+h)^2(1+h+h^2)^{1-\alpha/2}h^{\alpha-2} \\
&\le (\alpha-1)^{-1}(2+h)(1+h)^2(1+h+h^2)^{1/2}h^{\alpha-2} \\
&\le (\alpha-1)^{-1}(2+h)(1+h)^2(1+2h)^{1/2}h^{\alpha-2} \\
&\le (\alpha-1)^{-1}(2+h)(1+h)^3h^{\alpha-2}.
\end{aligned}$$

Now we take $h = 2 - \alpha$, to give

$$\limsup_{x\to\infty} \frac{\sigma_x}{x} \le (\alpha-1)^{-1}(4-\alpha)(3-\alpha)^3(2-\alpha)^{\alpha-2}, \text{ a.s.}$$

This completes the proof, with the value for C'_α given at (3.4). □

3.5 Upper bounds: Infinite variance

In this section we prove lower bounds $\sigma_x/x > C'_\alpha$ with $C'_\alpha > 2$. In particular, by the inversion argument relating σ_x and r_m presented in (3.2), we will establish the upper bound in Theorem 1.6. Define $u_\alpha : (1, \infty) \to (0, 1)$ by

$$u_\alpha(w) := 2^{\frac{\alpha}{2}-1}(\alpha-1)\frac{\sin(\pi\alpha/2)}{\pi} \int_{w-1}^{\infty} \frac{\mathrm{d}v}{v^{\alpha/2}(2+v)^{\alpha/2}(1+v)}, \text{ for } w > 1. \tag{3.34}$$

Proposition 3.9. *Suppose that* (I) *holds, and that* (S_α) *holds with* $\alpha \in (1, 2)$. *With* C'_α *given by* (3.4), *set*

$$C''_\alpha := 2 + \sup\left\{\frac{(C-2) \wedge u_\alpha((3/2)(C-1))}{C+1} : C > C'_\alpha\right\}.$$

Then it holds that, a.s.,

$$\liminf_{x\to\infty} \frac{\sigma_x}{x} \ge C''_\alpha > 2. \tag{3.35}$$

A key component in the proof of Proposition 3.9 is Lemma 3.10 below. The result is underpinned by an extension due to Kesten [30] of the Dynkin–Lamperti theorem (Proposition 2.3), which shows that when each walk overshoots $[-x, x]$ there is a positive probability it ends up at a distance at least Ax, say, for appropriately large A. The proof of Lemma 3.10 combines the latter result with Kesten's hitting estimates from Lemma 2.8. To state the lemma we need some further notation.

Let $\mathcal{G}_m := \sigma(S^{(1)}, \ldots, S^{(m)})$, the σ-algebra generated by the first m random walks, and note that σ_x defined in (3.1) is a stopping time with respect to the filtration $(\mathcal{G}_m)_{m\in\mathbb{Z}_+}$. For $m \in \mathbb{Z}_+$, $A > 0$ set

$$k_x(m, A) := \#\big(\mathfrak{C}_{\sigma_x+m} \setminus [-Ax, Ax]\big), \tag{3.36}$$

the number of occupied sites outside $[-Ax, Ax]$ at time $\sigma_x + m$. For $A > 1$, we think of $k_x(m, A)$ as "lost particles" that are too far away to contribute in the near future to the filling of the main cluster. The following result is a probability estimate that shows that lost particles are readily generated.

Lemma 3.10. *Let $A > 1$, $B > 0$. Recall the definition of u_α from (3.34). Then for every $u \in (0, u_\alpha(A+B))$ and every $\theta \in (0, B \wedge u)$, there is an $x_0 \in \mathbb{N}$ such that, for all $x \geq x_0$,*

$$\mathbb{P}(k_x(x, A) \leq \theta x \mid \mathcal{G}_{\sigma_x}) \leq \exp(-(\theta - u)^2 x/2), \ \ \textit{on } \{k_x(0, A) < Bx\}. \tag{3.37}$$

Proof. For $A > 0$, $x \in \mathbb{Z}_+$, and $m \in \mathbb{Z}_+$, let $\Delta_x(m, A) := k_x(m+1, A) - k_x(m, A)$ and note that $\Delta_x(m, A) = \mathbb{1}\{|S^{(\sigma_x+m+1)}_{\tau_{\sigma_x+m+1}}| > Ax\}$. The first step in the proof of (3.37) is to obtain a lower bound on

$$\mathbb{P}(\Delta_x(m, A) = 1 \mid \mathcal{G}_{\sigma_x+m}) = \mathbb{P}\big(\big|S^{(\sigma_x+m+1)}_{\tau_{\sigma_x+m+1}}\big| > Ax \bigm| \mathcal{G}_{\sigma_x+m}\big). \tag{3.38}$$

Let $B > 0$. For $y \in \mathbb{Z}$, $x \in \mathbb{Z}_+$, and $j > \sigma_x$, define

$$I_B(x, y, j) := \Big(\big[y - \tfrac{Bx}{2}, y + \tfrac{Bx}{2}\big] \cap \mathbb{Z}\Big) \setminus \mathfrak{C}_j, \tag{3.39}$$

the sites within distance $Bx/2$ of y which are not occupied by the cluster at time j. Recall, from (1.2) and (3.7), that $\tau_j = \eta^{(j)}_{\mathfrak{C}_{j-1}}$ and thus $\mathfrak{C}_j$ grows from $\mathfrak{C}_{j-1}$ by addition of $S^{(j)}_{\tau_j}$, that is, $S^{(j)}_{\tau_j}$ is the first site outside of $\mathfrak{C}_{j-1}$ visited by the jth random walk. Recall the definitions (3.6)–(3.7), and define $\widetilde{\eta}^{(j)}_B(x, y) := \inf\{n > \eta^{(j)}_{[-x,x]} : S^{(j)}_n \notin [y - Bx, y + Bx]\}$. With this notation $\widetilde{\eta}^{(j)}_B(x, y)$ denotes the first time the jth random walk lands outside the interval $[y - Bx, y + Bx]$ after it exits the interval $[-x, x]$. Note that for $j > \sigma_x$, the jth random walk is still active when it exits $[-x, x]$. If $t \in I_B(x, y, j)$ and $T^{(j)}_t \leq \widetilde{\eta}^{(j)}_B(x, y)$, then $S^{(j)}_{\tau_j} \in [y - Bx, y + Bx]$ and $|S^{(j)}_{\tau_j}| > Ax$ for any $|y| > A + B$. Therefore, for any $j > \sigma_x$, we have

$$\bigcup_{y\in\mathbb{Z}\colon |y|>(A+B)x} \Big\{S^{(j)}_{\eta^{(j)}_{[-x,x]}} = y, \ \min_{t\in I_B(x,y,j)} T^{(j)}_t < \widetilde{\eta}^{(j)}_B(x, y)\Big\} \subseteq \Big\{\big|S^{(j)}_{\tau_j}\big| > Ax\Big\}. \tag{3.40}$$

Now, for $I \subset \mathbb{Z}$, set

$$F_B(x, y, I) := \mathbf{P}_0\Big(S_{\eta_{[-x,x]}} = y, \min_{t\in I} T_t < \widetilde{\eta}_B(x, y)\Big).$$

By the strong Markov property applied at time $\eta_{[-x,x]}$, we have

$$\begin{aligned} F_B(x, y, I) &= \mathbf{P}_0\big(S_{\eta_{[-x,x]}} = y\big)\, \mathbf{P}_y\Big(\min_{t\in I} T_t < \eta_{[y-Bx,y+Bx]}\Big) \\ &\geq \mathbf{P}_0\big(S_{\eta_{[-x,x]}} = y\big) \inf_{t\in I} \mathbf{P}_y\big(T_t < \eta_{[y-Bx,y+Bx]}\big). \end{aligned}$$

Combined with (3.40), it follows that, for $j > \sigma_x$, on $\{I_B(x,y,j) \neq \emptyset\}$,

$$\begin{aligned}
&\mathbb{P}\big(\big|S^{(j+1)}_{\tau_{j+1}}\big| \geq Ax \mid \mathcal{G}_j\big) \\
&\quad \geq \sum_{y\in\mathbb{Z}:\, |y|>(A+B)x} F_B(x,y,I_B(x,y,j)) \\
&\quad \geq \sum_{y\in\mathbb{Z}:\, |y|>(A+B)x} \mathbf{P}_0\big(S_{\eta_{[-x,x]}} = y\big) \inf_{t\in[y-\frac{Bx}{2},y+\frac{Bx}{2}]\cap\mathbb{Z}} \mathbf{P}_y\big(T_t < \eta_{[y-Bx,y+Bx]}\}\big). \qquad (3.41)
\end{aligned}$$

Now, by the symmetry assumption of the increments contained in (S_α), we have

$$\inf_{t\in[y-\frac{Bx}{2},y+\frac{Bx}{2}]\cap\mathbb{Z}} \mathbf{P}_y\big(T_t < \eta_{[y-Bx,y+Bx]}\}\big) = \inf_{t\in[y-\frac{Bx}{2},y]\cap\mathbb{Z}} \mathbf{P}_y\big(T_t < \eta_{[y-Bx,y+Bx]}\}\big)$$

By translation invariance,

$$\mathbf{P}_y\big(T_t < \eta_{[y-Bx,y+Bx]}\big) = \mathbf{P}_{y-t}\big(T_0 < \eta_{[y-t-Bx,y-t+Bx]}\}\big),$$

for all y and t. Therefore, we have

$$\inf_{t\in[y-\frac{Bx}{2},y]\cap\mathbb{Z}} \mathbf{P}_y\big(T_t < \eta_{[y-Bx,y+Bx]}\}\big) = \inf_{z\in[0,\frac{Bx}{2}]\cap\mathbb{Z}} \mathbf{P}_z\big(T_0 < \eta_{[z-Bx,z+Bx]}\}\big).$$

Since $\mathbf{P}_z(T_0 < \eta_{[z-Bx,z+Bx]}) \geq \mathbf{P}_z(T_0 < \eta_{[-\frac{Bx}{2},Bx]})$ for $z \in [0,\frac{Bx}{2}]\cap\mathbb{Z}$, it follows that

$$\inf_{z\in[0,\frac{Bx}{2}]\cap\mathbb{Z}} \mathbf{P}_z\big(T_0 < \eta_{[z-Bx,z+Bx]}\big) \geq \inf_{z\in[0,\frac{Bx}{2}]\cap\mathbb{Z}} \mathbf{P}_z\big(T_0 < \eta_{[-\frac{Bx}{2},Bx]}\big).$$

Using the notation in (2.19) we may write

$$\inf_{z\in[0,\frac{Bx}{2}]\cap\mathbb{Z}} \mathbf{P}_z\big(T_0 < \eta_{[-\frac{Bx}{2},Bx]}\big) = \inf_{\frac{z}{Bx}\in[0,\frac{1}{2}],\, z\in\mathbb{Z}} q_{\alpha,Bx}\big(\tfrac{z}{Bx};1/2\big) \quad \geq \inf_{y\in[0,\frac{1}{2}]} q_{\alpha,Bx}\big(y;1/2\big).$$

From the equicontinuity of $q_{\alpha,x}$ as at (2.25) and the lower bound from (2.26) we obtain

$$\lim_{x\to\infty} \inf_{y\in[0,\frac{1}{2}]} q_{\alpha,Bx}\big(y;1/2\big) = \inf_{0\leq y\leq 1/2} q_\alpha(y;1/2) \geq (\alpha-1)(1/2)^{2-\frac{\alpha}{2}} =: r_\alpha. \qquad (3.42)$$

Let $\varepsilon > 0$. Then, using the bound (3.42) in (3.41), we obtain, for all $x \geq x_0$ large enough, on the event $\bigcap_{y:|y|>(A+B)x}\{I(x,y,j) \neq \emptyset\}$,

$$\begin{aligned}
\mathbb{P}\big(\big|S^{(j+1)}_{\tau_{j+1}}\big| > Ax \mid \mathcal{G}_j\big) &\geq (r_\alpha - \varepsilon) \sum_{y\in\mathbb{Z}:\, |y|>(A+B)x} \mathbf{P}_0\big(S_{\eta_{[-x,x]}} = y\big) \\
&= (r_\alpha - \varepsilon)\, \mathbf{P}_0\big(\big|S_{\eta_{[-x,x]}}\big| > (A+B)x\big). \qquad (3.43)
\end{aligned}$$

Kesten [31, p. 270] considers the quantity $G_\alpha(w;x,1) := \mathbf{P}_0(x < S_{\eta_{[-x,x]}} \leq (w+1)x)$. Under hypothesis ($\mathrm{S}_\alpha$), Kesten proved in [30] that $\lim_{x\to\infty} G_\alpha(w;x,1) = G_\alpha(w,1)$ exists, and gives a formula for $G_\alpha(w,1)$ in Theorem 1 of [31, p. 271]. Moreover, the symmetry assumed in (S_α) shows that $\mathbf{P}_0(|S_{\eta_{[-x,x]}}| > (w+1)x) = 2\,\mathbf{P}_0(S_{\eta_{[-x,x]}} > (w+1)x)$. In particular, it follows from Theorem 1 of [31] that, for $w > 0$,

$$\lim_{x\to\infty} \mathbf{P}_0\big(\big|S_{\eta_{[-x,x]}}\big| > (w+1)x\big) = \frac{2\sin(\pi\alpha/2)}{\pi} \int_w^\infty \frac{\mathrm{d}v}{v^{\alpha/2}(2+v)^{\alpha/2}(1+v)} =: s_\alpha(w). \quad (3.44)$$

Combining (3.43) and (3.44), we obtain for $\delta \in (0, 1/2)$ that, for every $\varepsilon > 0$, for all $x \geq x_0$ large enough, on $\{I(x, y, j) \neq \emptyset\}$,

$$\mathbb{P}\big(\big|S^{(j+1)}_{\tau_{j+1}}\big| > Ax \bigm| \mathcal{G}_j\big) \geq u_\alpha(A + B) - \varepsilon, \tag{3.45}$$

where $u_\alpha(w) = r_\alpha s_\alpha(w - 1)$, with r_α and s_α given at (3.42) and (3.44), which gives the expression at (3.34).

Recall from (3.36) that $k_x(m, A)$ counts the number of particles in $\mathfrak{C}_{\sigma_x+m} \setminus [-Ax, Ax]$. Now observe that if $k_x(m, A) < Bx$, then at time $\sigma_x + m$, for every y with $|y| > (A+B)x$ we have at least one unoccupied site in $[y - \frac{Bx}{2}, y + \frac{Bx}{2}]$, since the length of the interval $[y - \frac{Bx}{2}, y + \frac{Bx}{2}]$ is Bx, and thus, with the notation at (3.39), we have $I(x, y, \sigma_x + m) \neq \emptyset$. Hence, from (3.38) with the $j = \sigma_x + m$ case of (3.45),

$$\mathbb{P}(\Delta_x(m, A) = 1 \mid \mathcal{G}_{\sigma_x+m}) \geq u_\alpha(A + B) - \varepsilon, \text{ on } \{k_x(m, A) < Bx\}. \tag{3.46}$$

Now for $\theta \in (0, B)$, let

$$\lambda_x := \inf\{m \in \mathbb{Z}_+ : k_x(m, A) > \theta x\}. \tag{3.47}$$

Note that $\{k_x(x, A) \leq \theta x\} = \{\lambda_x > x\}$. For $0 < u < u_\alpha(A + B)$, let $M_m := k_x(m \wedge \lambda_x, A) - u(m \wedge \lambda_x)$. Recall that $k_x(m, A) = k_x(0, A) + \sum_{j=0}^{m-1} \Delta_x(j, A)$. Since $\theta < B$, we can (and do) assume that x is large enough so that $\theta x + 1 < Bx$. Then observe that, for all x sufficiently large, it follows from (3.46) that

$$\mathbb{E}\left(\Delta_x(m \wedge \lambda_x, A) \mid \mathcal{G}_{\sigma_x+m}\right) \geq u, \text{ a.s.},$$

since for any $m < \lambda_x$, $k_x(m \wedge \lambda_x, A) \leq \theta x + 1 < Bx$ by the definition of λ_x. Thus, for all x sufficiently large, from the Doob decomposition of $k_x(m, A)$ (see e.g. Theorem 5.2.10 of [20]), it follows that $(M_m)_{m\in\mathbb{Z}_+}$ is a submartingale adapted to $(\mathcal{G}_{\sigma_x+m})_{m\in\mathbb{Z}_+}$.

From (3.46) and the fact that $\Delta_x(m, A) \in \{0, 1\}$ it follows that for $0 < \theta < u \wedge B$

$$\begin{aligned}
\mathbb{P}(\lambda_x > x \mid \mathcal{G}_{\sigma_x}) &= \mathbb{P}(k_x(x, A) < \theta x,\ \lambda_x > x \mid \mathcal{G}_{\sigma_x}) \\
&= \mathbb{P}(k_x(x, A) - ux < (\theta - u)x,\ \lambda_x > x \mid \mathcal{G}_{\sigma_x}) \\
&= \mathbb{P}(M_x < (\theta - u)x,\ \lambda_x > x \mid \mathcal{G}_{\sigma_x}) \\
&\leq \mathbb{P}(M_x < (\theta - u)x \mid \mathcal{G}_{\sigma_x}) \leq \exp(-(\theta - u)^2 x/2),
\end{aligned} \tag{3.48}$$

where we use the Azuma–Hoeffding inequality for submartingales, Theorem 2.4.14 of [39]. This yields (3.37). □

Next we prove Proposition 3.9.

Proof of Proposition 3.9. First observe that from (3.5) established in Proposition 3.2 above, for every $C > C'_\alpha > 2$ as given at (3.4), a.s., for all but finitely many $x \in \mathbb{N}$, it holds that $\sigma_x \leq Cx$. Thus, at time Cx, there are $2x$ of the walks that have occupied sites in $[-x, x]$, and at most $C - 2x$ of the walks that are outside $[-x, x]$. It follows from definition of k_x at (3.36) that, for every $A \geq 1$, a.s., for all but finitely many $x \in \mathbb{N}$, $k_x(0, A) \leq (C - 2)x$. In particular, we can choose $A > 1$ and $B = C - 2 > 0$, to conclude that $k_x(0, A) \leq Bx$ for all but finitely many $x \in \mathbb{N}$. Then from (3.37) we have that for every $u \in (0, u_\alpha(A + B))$ and every $\theta \in (0, B \wedge u)$, $\sum_{x\in\mathbb{N}} \mathbb{P}(k_x(x, A) \leq \theta x \mid \mathcal{G}_{\sigma_x}) < \infty$. Hence, by Lévy's conditional Borel–Cantelli lemma [29, p. 131], we conclude that $k_x(x, A) > \theta x$ for all but finitely many $x \in \mathbb{N}$, a.s.

Now choose $A = (C+1)/2$. If $\sigma_y \le Cy$ and $k_y(y, A) > \theta y$, then $\sigma_y + y \le (C+1)y = 2Ay$, meaning that at time $2Ay$ there are more than θy sites outside $[-Ay, Ay]$. Therefore at time $2Ay$ there are at least θy empty sites in $[-Ay, Ay]$. Those empty sites must be filled before time σ_{Ay} and consequently, $\sigma_{Ay} \ge 2Ay + \theta y$. Taking $y = \lfloor x/A \rfloor$, we conclude that, a.s.,

$$\liminf_{x\to\infty} \frac{\sigma_x}{x} \ge 2 + \frac{2\theta}{C+1}, \tag{3.49}$$

where $C > C'_\alpha$. The inequality (3.49) holds true for any positive $\theta \le B \wedge u_\alpha(A+B)$, with $B = C - 2$ and $A > 1$. Therefore, we can choose A such that $A + B < (3/2)(C-1)$ to conclude from $u_\alpha(A+B) \ge u_\alpha((3/2)(C-1))$ that

$$\liminf_{x\to\infty} \frac{\sigma_x}{x} \ge 2 + \frac{(C-2) \wedge u_\alpha((3/2)(C-1))}{C+1}, \text{ a.s.}$$

Since $C > C'_\alpha$ was arbitrary, this completes the proof. □

Finally, we conclude this section with the proof of Theorem 1.6.

Proof of Theorem 1.6. First, with $C'_\alpha \in (2, \infty)$ as given at (3.4), we have from (3.5) in Proposition 3.2 that $\limsup_{x\to\infty} \sigma_x/x \le C'_\alpha$, a.s. Together with the inversion relation (3.1)–(3.2), it follows immediately that $\liminf_{m\to\infty} r_m/m \ge 1/C'_\alpha$, a.s., which yields the lower bound in (1.7) with $c_\alpha := 1/C'_\alpha \in (0, 1/2)$. The formula (1.8) for c_α follows directly from formula (3.4) for C'_α.

In the other direction, we have from (3.35) in Proposition 3.9 that $\liminf_{x\to\infty} \sigma_x/x \ge C''_\alpha$, a.s., where $C'_\alpha \ge C''_\alpha > 2$ is as defined in Proposition 3.9. Again, inversion then yields $\limsup_{m\to\infty} r_m/m \le c'_\alpha$, a.s., where $c'_\alpha := 1/C''_\alpha$ satisfies $0 < c_\alpha \le c'_\alpha < 1/2$, completing the proof of (1.7) and hence of Theorem 1.6. □

A Eventual filling

This short appendix presents the proof of Proposition 1.2. As in Section 3.5, we write $\mathcal{G}_m = \sigma(S^{(1)}, \ldots, S^{(m)})$.

Proof of Proposition 1.2. It suffices to prove that for every $A \subseteq \mathbb{Z}$ finite, we have $A \subseteq \mathfrak{C}_\infty$, a.s. Fix such an A. Let $z \in A$. Then, by the irreducibility property (1.1), there exist $p_z > 0$, $n_z \in \mathbb{Z}_+$, and $x_{z,1}, \ldots, x_{z,n_z-1} \in \mathbb{Z} \setminus \{0, z\}$ such that

$$\mathbb{P}(S_1 = x_{z,1}, \ldots, S_{n_z-1} = x_{z,n_z-1}, S_n = z) = p_z > 0.$$

With probability $p_z^{n_z}$, a sequence of n_z successive random walkers will follow path $x_{z,1}, \ldots, x_{z,n_z-1}, z$ for their first n_z steps; on this event, at least one of the walkers will reach z before terminating, and so $z \in \mathfrak{C}_\infty$. Set $p_A := \prod_{z\in A} p_z^{n_z} > 0$ and $n_A := \sum_{z\in A} n_z < \infty$. Given $\mathfrak{C}_m$, iterating the above argument shows that (regardless of the existing configuration), with probability p_A, the sequence of random walkers $m+1, m+2, \ldots, m+n_A$ executes an event such that $A \subseteq \mathfrak{C}_{m+n_A}$; that is,

$$\mathbb{P}(A \subseteq \mathfrak{C}_\infty \mid \mathcal{G}_m) \ge \mathbb{P}(A \subseteq \mathfrak{C}_{m+n_A} \mid \mathcal{G}_m) \ge p_A, \text{ a.s.}$$

Denoting $\mathcal{G}_\infty := \sigma(\cup_{m\in\mathbb{Z}_+}\mathcal{G}_n)$, it follows from Lévy's zero–one law (see e.g. Theorem 5.5.8 of [20]) that

$$\mathbb{P}(A \subseteq \mathfrak{C}_\infty \mid \mathcal{G}_\infty) = \lim_{m\to\infty} \mathbb{P}(A \subseteq \mathfrak{C}_\infty \mid \mathcal{G}_m) \geq p_A, \text{ a.s.}$$

But since $\{A \subseteq \mathfrak{C}_\infty\} \in \mathcal{G}_\infty$, this means that $\mathbb{1}\{A \subseteq \mathfrak{C}_\infty\} \geq p_A$, a.s., for deterministic $p_A > 0$, from which it must hold that $\mathbb{P}(A \subseteq \mathfrak{C}_\infty) = 1$. ☐

B Equicontinuity

The aim of this section is to prove Lemma 2.10. The first step is the following lemma, which is essentially provided already by Kesten. Recall that the local hitting property, Lemma 2.5, expresses the fact that a recurrent random walk will very likely hit a point at finite distance from its starting point before going far away; Lemma B.1 extends this statement from points at a finite distance to points at a distance allowed to grow slowly with the size of the interval. In this section of the appendix, we use the same notation as Section 2.2; we recall in particular the notation T_t and η_A from (2.15).

Lemma B.1 (Kesten 1961 [31])**.** *Suppose that* (I) *holds. For* $1 < \alpha \leq 2$*, suppose in addition that (if* $\alpha = 2$*)* $\mathbb{E}(X^2) < \infty$ *and* $\mathbb{E}\, X = 0$*, or (if* $1 < \alpha < 2$*) that* (S_α) *holds. Then the following hold.*

(i) For any $\varepsilon > 0$*, there exist* $\delta > 0$ *and* $N_0 s \in \mathbb{N}$*, such that, for all* $N \geq N_0$ *and every* $k \in \mathbb{Z}$ *with* $|k| \leq \delta N$*,*

$$\mathbf{P}_k\left(T_0 < \eta_{(-\infty,N]}\right) \geq 1 - \varepsilon. \tag{B.1}$$

(ii) For any $\varepsilon > 0$ *and* $c > 0$*, there exist* $\delta > 0$ *and* $N_0 \in \mathbb{N}$*, such that, for all* $N \geq N_0$ *and every* $k \in \mathbb{Z}$ *with* $|k| \leq \delta N$*,*

$$\mathbf{P}_k\left(T_0 < \eta_{[-cN,N]}\right) \geq 1 - \varepsilon. \tag{B.2}$$

Proof. Part (i) is already available as Lemma 3 in [31] for $1 < \alpha < 2$. For $\alpha = 2$, it is available as Lemma 3 in [31] under the additional assumption of symmetric increment distribution, i.e., $X \overset{d}{=} -X$. If we remove this assumption, but instead assume that $\mathbb{E}\, X = 0$ and $\mathbb{E}(X^2) < \infty$, then it is easy to show that from Taylor expansion and dominated convergence that

$$\lim_{t\to 0} \frac{1}{t^2}\left(1 - \mathbb{E}\left[\mathrm{e}^{itX}\right]\right) = \mathbb{E}(X^2);$$

observe also that $\mathbb{E}\, X = 0$ and (I) together imply that $\mathbb{E}(X^2) > 0$. This observation is the key ingredient in the proof of Lemma 3 in [31] (see equation (2.3) which leads to (2.31) in [31]), and hence an exact verbatim proof as in Lemma 3 in [31] yields (B.1).

Part (ii) can be deduced from part (i), as is also indicated in the proof of Lemma 4 in [31]. Indeed, observe that for any $c > 0$, we have

$$\{T_0 < \eta_{(-\infty,N]}\} = \bigcup_{m\in\mathbb{N}} \{T_0 < \eta_{[-cm,N]}\},$$

where, for $m \in \mathbb{N}$, the monotonicity property $\{T_0 < \eta_{[-cm,N]}\} \subseteq \{T_0 < \eta_{[-c(m+1),N]}\}$ is satisfied. Hence, by continuity of monotone limits,

$$\lim_{m\to\infty} \mathbf{P}_k\left(T_0 < \eta_{[-cm,N]}\right) = \mathbf{P}_k\left(T_0 < \eta_{(-\infty,N]}\right).$$

This, together with (B.1) completes the proof of (B.2). ☐

We are now ready to present the proof of Lemma 2.10.

Proof of Lemma 2.10. The proof for (ii) is already available in Lemma 4 in [31], so we only prove part (i) here, which is also very similar.

Suppose that $S_0 = k$, and observe that, for $0 < k < N$,

$$\{T_{-1} < \eta_{(-\infty,N]}\} \subseteq \{S_{\eta_{[0,N]}} < 0\}. \tag{B.3}$$

This shows that for any $|y| < \delta(\varepsilon)$,

$$p_{\alpha,N}(y) \geq 1 - \varepsilon. \tag{B.4}$$

Therefore, for any $0 \leq y_1, y_2 \leq \delta(\varepsilon)$, since $0 \leq p_{\alpha,N}(y) \leq 1$

$$p_{\alpha,N}(y_1) \geq (1-\varepsilon)\, p_{\alpha,N}(y_2), \tag{B.5}$$

which proves equicontinuity for $|y| < \delta(\varepsilon)$. Hence, it is enough to show equicontinuity for any $y > \delta(\varepsilon)$, the proof of which is similar to that of part (ii). We provide the details for the purpose of completeness.

Let $k_1, k_2 \geq \delta(\varepsilon)N$, it follows from the Markov property and translation invariance, that

$$p_{\alpha,N}\Big(\frac{k_1}{N}\Big) \geq \mathbf{P}_{k_1}\Big(S_n = k_2, \text{ for some } 1 \leq n < \eta_{[\frac{\delta(\varepsilon)}{2}N,N]}\Big)\, p_{\alpha,N}\Big(\frac{k_2}{N}\Big). \tag{B.6}$$

It follows from translation invariance and (B.2), that for any given $\varepsilon > 0$, there exits $\delta_1(\varepsilon)$, such that

$$\mathbf{P}_{k_1}\Big(S_n = k_2, \text{ for some } 1 \leq n < \eta_{[\frac{\delta(\varepsilon)}{2}N,N]}\Big) \geq 1 - \varepsilon, \tag{B.7}$$

whenever $k_1, k_2 \leq N(1-\varepsilon)$, and $|k_1 - k_2| \leq N\delta_1(\varepsilon)$. Thus we can choose $\delta_2(\varepsilon) > 0$, such that, for $0 \leq y_1, y_2 < 1 - \varepsilon$ and $|y_1 - y_2| < \delta_2(\varepsilon)$, such that,

$$p_{\alpha,N}(y_1) \geq (1-\varepsilon)p_{\alpha,N}(y_2). \tag{B.8}$$

Since, y_1 and y_2 are arbitrary and $0 \leq p_{\alpha,N}(y) \leq 1$, this proves (i). ☐

Acknowledgements

The authors are grateful to M. Balász, E. Crane, and J. Jay for valuable discussions, and for making available the draft version of [9]. CdC and AW were supported by EPSRC grant EP/W00657X/1. Some of this work was carried out during the programme "Stochastic systems for anomalous diffusion" (July–December 2024) hosted by the Isaac Newton Institute, under EPSRC grant EP/Z000580/1.

References

[1] M. Abramowitz and I.A. Stegun (Eds.), *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, **55**. U.S. Government Printing Office, Washington D.C., 1965.

[2] G. Amir, O. Angel, I. Benjamini, and G. Kozma, One-dimensional long-range diffusion-limited aggregation I. *Ann. Probab.* **44** (2016) 3546–3579.

[3] G. Amir, O. Angel, and G. Kozma, One-dimensional long-range diffusion limited aggregation II: The transient case. *Ann. Appl. Probab.* **27** (2017) 1886–1922.

[4] G. Amir, One-dimensional long-range diffusion-limited aggregation III: The limit aggregate. *Ann. Inst. H. Poincaré Probab. Statist.* **53** (2017) 1513–1527.

[5] S. Asmussen, *Applied Probability and Queues.* 2nd ed., Springer, New York, 2003.

[6] A. Asselah and A. Gaudillière, From logarithmic to subdiffusive polynomial fluctuations for internal DLA and related growth models. *Ann. Probab.* **41** (2013) 1115–1159.

[7] A. Asselah and A. Gaudillière, Sub-logarithmic fluctuations for internal DLA. *Ann. Probab.* **41** (2013) 1160–1179.

[8] M.T. Barlow, R. Pemantle, and E.A. Perkins, Diffusion-limited aggregation on a tree. *Probab. Theory Relat. Fields* **107** (1997) 1–60.

[9] M. Balázs, E. Crane, J. Jay, and D. Thacker, Internal diffusion-limited aggregation with totally asymmetric long-range jumps, and its connection to ASEP blocking measures and dynamical Mallows permutations of the integers (in preparation).

[10] I. Benjamini, H. Duminil-Copin, G. Kozma, and C. Lucas, Internal diffusion-limited aggregation with uniform starting points. *Ann. de l'Institut H. Poincaré - Probab. Stat.* **56** (2020) 391–404.

[11] N. Berger, J.J. Kagan, and E.B. Procaccia, Stretched IDLA. *ALEA.* **11** (2014) 471–481.

[12] N.H. Bingham, C.M. Goldie and J.L. Teugels, *Regular Variation.* Cambridge University Press, Cambridge, 1989.

[13] J.T. Chang, Inequalities for the overshoot. *Ann. Appl. Probab.* **4** (1994) 1223–1133.

[14] Z.-Q. Chen, P. Kim and R. Song, Heat kernel estimates for the Dirichlet fractional Laplacian. *J. Eur. Math. Soc.* **12** (2010) 1307–1329.

[15] N. Chenavier, D. Coupier and A. Rousselle, The bi-dimensional directed IDLA forest, *Ann. Appl. Probab.* **33** (2023) 2247–2290.

[16] P. Diaconis and W. Fulton, A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Semin. Mat. Univ. Politec. Torino* **49** (1991) 95–119.

[17] R.A. Doney, Moments of ladder heights in random walks. *J. Appl. Probab.* **17** (1980) 248–252.

[18] R.A. Doney, One-sided local large deviation and renewal theorems in the case of infinite mean. *Probab. Theory Relat. Fields* **107** (1997) 451–465.

[19] M. DrMota, M. Fuchs, K.-H. Hwang and R. Neininger, Node profiles of symmetric digital search trees: Concentration properties. *Random Struct. Algorithms* **58** (2020) 430–467.

[20] R. Durrett, *Probability: Theory and Examples.* 4th ed., Cambridge University Press, Cambridge, 2010.

[21] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. II.* 2nd ed., John Wiley & Sons, Inc., New York, 1971.

[22] D.A. Freedman, Bernard Friedman's urn. *Ann. Math. Statist.* **36** (1965) 956–970.

[23] A. Gut, *Stopped Random Walks.* 2nd ed., Springer, New York, 2009.

[24] I.A. Ibragimov and Ju.V. Linnik, *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen, 1971.

[25] S. Janson, Renewal theory for m-dependent variables. *Ann. Probab.* **11** (1983) 558–568.

[26] D. Jerison, L. Levine, and S. Sheffield, Logarithmic fluctuations for internal DLA. *J. Amer. Math. Soc.* **25** (2012) 271–301.

[27] D. Jerison, L. Levine, and S. Sheffield, Internal DLA in higher dimensions. *Electron. J. Probab.* **18** (2013) no. 14.

[28] D. Jerison, L. Levine, and S. Sheffield, Internal DLA and the Gaussian free field. *Duke Math. J.* **163** (2014) 267–308.

[29] O. Kallenberg, *Foundations of Modern Probability.* 2nd edition, Springer, New York, 2002.

[30] H. Kesten, On a theorem of Spitzer and Stone and random walks with absorbing barriers. *Illinois J. Math.* **5** (1961) 246–266.

[31] H. Kesten, Random walks with absorbing barriers and Toeplitz forms. *Illinois J. Math.* **5** (1961) 267–290.

[32] H. Kesten, How long are the arms in DLA? *J. Phys. A: Math. Gen.* **20** (1987) L29–L33.

[33] T.L. Lai, Asymptotic moments of random walks with applications to ladder variables and renewal theory. *Ann. Probab.* **4** (1976) 51–66.

[34] G.F. Lawler, M. Bramson, and D. Griffeath, Internal diffusion limited aggregation. *Ann. Probab.* **20** (1992) 2117–2140.

[35] G.F. Lawler and V. Limic, *Random Walk: A Modern Introduction.* Cambridge University Press, Cambridge, 2010.

[36] G. Lorden, On the excess over the boundary. *Ann. Math. Statist.* **41** (1970) 520–527.

[37] V.I. Lotov, Bounds for the probability to leave the interval. *Statist. Probab. Letters* **145** (2019) 141–146.

[38] C. Lucas, The limiting shape for drifted internal diffusion limited aggregation is a true heat ball. *Probab. Theory Relat. Fields* **159** (2014) 197–235.

[39] M. Menshikov, S. Popov, and A. Wade, *Non-homogeneous Random Walks.* Cambridge University Press, Cambridge, 2016.

[40] M. Penrose, *Random Geometric Graphs.* Oxford University Press, Oxford, 2003.

[41] B.A. Rogozin, On the distribution of the first ladder moment and height and fluctuations of a random walk. *Theory Probab. Appl.* **16** (1971) 575–595.

[42] E. Sava-Huss, From fractals in external DLA to internal DLA on fractals. pp. 273–298 in *Fractal Geometry and Stochastics VI*, Progress in Probability vol. 76, Springer, 2021.

[43] F. Spitzer, A Tauberian theorem and its probability interpretation. *Trans. Amer. Math. Soc.* **94** (1960) 150–169.

[44] T.A. Witten Jr. and L.M. Sander, Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Letters* **47** (1981) 1400.