

## How long is the convex minorant of a one-dimensional random walk?

Gerold Alsmeyer\*    Zakhar Kabluchko†    Alexander Marynych‡  
 Vladislav Vysotsky§

### Abstract

We prove distributional limit theorems for the length of the largest convex minorant of a one-dimensional random walk with independent identically distributed increments. Depending on the increment law, there are several regimes with different limit distributions for this length. Among other tools, a representation of the convex minorant of a random walk in terms of uniform random permutations is utilized.

**Keywords:** convex minorant; random permutation; random walk.

**MSC2020 subject classifications:** Primary 60F05; 60G55, Secondary 60J10.

Submitted to EJP on September 26, 2019, final version accepted on July 23, 2020.

Supersedes arXiv:1909.12322v1.

## 1 Introduction and main results

Given a sequence  $(\xi_k)_{k \in \mathbb{N}}$  of independent and identically distributed (i.i.d.) real-valued random variables with a generic copy  $\xi$ , consider the associated random walk  $(S_n)_{n \in \mathbb{N}_0}$ ,  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ , defined by

$$S_0 := 0 \quad \text{and} \quad S_n := \xi_1 + \xi_2 + \cdots + \xi_n \quad \text{for } n \in \mathbb{N},$$

and the random piecewise linear function  $t \mapsto S(t)$ ,  $t \geq 0$ , obtained by linear interpolation between the values  $S(n) := S_n$  for  $n \in \mathbb{N}_0$ . For any fixed  $T > 0$ , let  $t \mapsto S_T^-(t)$  and  $t \mapsto S_T^+(t)$  be, respectively, the convex minorant and the concave majorant of the

---

\*Institute of Mathematical Stochastics, Department of Mathematics and Computer Science, University of Münster, Orleans-Ring 10, D-48149 Münster, Germany. E-mail: [gerolda@uni-muenster.de](mailto:gerolda@uni-muenster.de)

†Institute of Mathematical Stochastics, Department of Mathematics and Computer Science, University of Münster, Orleans-Ring 10, D-48149 Münster, Germany. E-mail: [zakhar.kabluchko@uni-muenster.de](mailto:zakhar.kabluchko@uni-muenster.de)

‡Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, 01601 Kyiv, Ukraine. E-mail: [marynych@unicyb.kiev.ua](mailto:marynych@unicyb.kiev.ua)

§University of Sussex, Pevensey 2 Building, Falmer Campus, Brighton BN1 9QH, United Kingdom and St. Petersburg Department of Steklov Mathematical Institute, Fontanka 27, 191011 St. Petersburg, Russia. E-mail: [v.vysotskiy@sussex.ac.uk](mailto:v.vysotskiy@sussex.ac.uk)

function  $t \mapsto S(t)$  on the interval  $[0, T]$ . Let us recall that the convex minorant (concave majorant) of a function  $f$  on an interval  $[a, b]$  is the largest convex (least concave) function  $f^\sim$  ( $f^\wedge$ ) such that  $f^\sim(x) \leq f(x)$  ( $f(x) \leq f^\wedge(x)$ ) for all  $x \in [a, b]$ . Clearly, both  $t \mapsto S_T^\sim(t)$  and  $t \mapsto S_T^\wedge(t)$  are piecewise linear continuous functions and have therefore well-defined finite lengths, here denoted by  $L_T^\sim$  and  $L_T^\wedge$ , respectively.

In this paper, we provide distributional limit theorems for  $L_n^\sim$  and  $L_n^\wedge$  as  $n \rightarrow \infty$  in the following three regimes:

- (A)  $\mathbb{E}\xi^2 < \infty$  and  $\mathbb{E}\xi = 0$ ;
- (B) the law of  $\xi$  lies in the domain of attraction of an  $\alpha$ -stable law with  $\alpha \in (1, 2)$  and  $\mathbb{E}\xi = 0$ ;
- (C) the law of  $\xi$  lies in the domain of attraction of an  $\alpha$ -stable law with  $\alpha \in (0, 1)$ .

The case  $\mathbb{E}\xi \neq 0$  in parts (A) and (B) turns out to be less intriguing and will be discussed in Section 4.

Let us point out at the outset that it suffices to consider the length of the convex minorant  $L_n^\sim$  because it has the same law as  $L_n^\wedge$ , so

$$L_n^\sim \stackrel{d}{=} L_n^\wedge \quad \text{for all } n \in \mathbb{N}. \tag{1.1}$$

Although nonintuitive, this follows fairly easily from the observation that the concave majorant of  $(S_0, \dots, S_n)$  for any  $n$  coincides with the negative of the convex minorant of the reflected vector  $(-S_0, \dots, -S_n)$  in combination with a distributional representation of  $L_n^\sim$ , stated as (2.2) and owing to Abramson et al. [1, 2], which only involves the squares of the  $S_k$ .

Before putting our work into some context by pointing out connections with earlier work on convex minorants and the convex hulls of random walks, we present our main results, stated as Theorems 1.1, 1.2 and 1.3.

It is well-known that in each of the cases (A), (B) and (C), there exists a sequence  $(a_n)_{n \in \mathbb{N}}$  of positive constants such that, with  $\mathcal{S}_\alpha = (\mathcal{S}_\alpha(t))_{t \in [0,1]}$  denoting an  $\alpha$ -stable Lévy process,

$$\left( \frac{S(nt)}{a_n} \right)_{t \in [0,1]} \xrightarrow{n \rightarrow \infty} (\mathcal{S}_\alpha(t))_{t \in [0,1]} \tag{1.2}$$

in the Skorokhod space  $D[0, 1]$  endowed with the standard  $J_1$ -topology. Note that  $\mathcal{S}_2$  is just a centered Brownian motion. Throughout the paper, we always use  $a_n$  and  $\mathcal{S}_\alpha$  for the normalization and the  $\alpha$ -stable Lévy process such that (1.2) holds. Also, we let  $\mathcal{N}(0, s^2)$  denote the normal distribution with mean zero and positive variance  $s^2$ . The notation  $\simeq$  stands for asymptotic equivalence, that is,  $f(x) \simeq g(x)$ , as  $x \rightarrow x_0$ , holds if and only if  $\lim_{x \rightarrow x_0} f(x)/g(x) = 1$ .

In order to state our result for case (A), put

$$\sigma_n^2 := \text{Var}(\xi \mathbb{1}_{\{|\xi| \leq \sqrt{n}\}}), \quad n \in \mathbb{N}, \tag{1.3}$$

and  $\log^+ x := \max(\log x, 0)$  for  $x > 0$ .

**Theorem 1.1.** *Suppose that  $\mathbb{E}\xi = 0$  and  $\mathbb{E}\xi^2 =: \sigma^2 = \lim_{n \rightarrow \infty} \sigma_n^2 < \infty$ . Then*

$$\frac{1}{\sqrt{\log n}} \left( L_n^\sim - n - \sum_{j=1}^n \frac{\sigma_j^2}{2j} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}\left(0, \frac{3\sigma^4}{4}\right), \tag{1.4}$$

which may be simplified to

$$\frac{1}{\sqrt{\log n}} \left( L_n^\sim - n - \frac{\sigma^2}{2} \log n \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N} \left( 0, \frac{3\sigma^4}{4} \right) \tag{1.5}$$

under the additional assumption  $\mathbb{E}\xi^2 \log^+ |\xi| < \infty$ .

In view of the previous result, one could expect that in case (B) a suitable normalization of  $L_n^\sim$  converges in law to some stable law. It may therefore be surprising that the true answer, stated in the next theorem, looks more complicated.

**Theorem 1.2.** *Suppose that the following assumptions hold:*

(B1) *The function  $t \mapsto \mathbb{P}\{|\xi| > t\}$  is regularly varying at infinity with index  $\alpha \in (1, 2)$  and  $\mathbb{E}\xi = 0$ .*

(B2) *For  $p, q \in [0, 1]$  such that  $p + q = 1$ , we have*

$$\lim_{x \rightarrow +\infty} \frac{\mathbb{P}\{\xi > x\}}{\mathbb{P}\{|\xi| > x\}} = p \quad \text{and} \quad \lim_{x \rightarrow +\infty} \frac{\mathbb{P}\{\xi < -x\}}{\mathbb{P}\{|\xi| > x\}} = q.$$

Then

$$\frac{n}{a_n^2} (L_n^\sim - n) \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2} \sum_{k=1}^{\infty} \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{Z_k},$$

where  $(Z_1, Z_2, \dots)$  is a random sequence that has a Poisson-Dirichlet distribution with parameter  $\theta = 1$  and  $(\mathcal{S}_\alpha^{(k)}(t))_{t \in [0,1]}$ ,  $k = 1, 2, \dots$ , are independent copies of the  $\alpha$ -stable process  $(\mathcal{S}_\alpha(t))_{t \in [0,1]}$  appearing in (1.2).

The proofs of both theorems rely on a known representation of the convex minorant in terms of uniform random permutations that will be described in Subsection 2.1, followed by some explanations of the main arguments in Subsection 2.2. The main difference between the cases (A) and (B) is that, roughly speaking, in case (A) the main contributions to the fluctuations of  $L_n^\sim$  are due to a large number of “small” segments of the convex minorant, whereas in case (B) they are rather due to few “large” segments.

Our third result deals with the case when  $\xi$  lies in the domain of attraction of a stable law with index  $\alpha \in (0, 1)$ . This is the simplest case because rather than making use of the connection with random permutations, a simple comparison argument applies; see Subsection 3.4 below.

**Theorem 1.3.** *Suppose that the following assumptions hold:*

(C1) *The function  $t \mapsto \mathbb{P}\{|\xi| > t\}$  is regularly varying at infinity with index  $\alpha \in (0, 1)$ .*

(C2) *For some  $p, q \in [0, 1]$  with  $p + q = 1$ ,*

$$\lim_{x \rightarrow +\infty} \frac{\mathbb{P}\{\xi > x\}}{\mathbb{P}\{|\xi| > x\}} = p \quad \text{and} \quad \lim_{x \rightarrow +\infty} \frac{\mathbb{P}\{\xi < -x\}}{\mathbb{P}\{|\xi| > x\}} = q.$$

Then

$$\left( \frac{L_n^\sim}{a_n}, \frac{L_n^\wedge}{a_n} \right) \xrightarrow[n \rightarrow \infty]{d} \left( \mathcal{S}_\alpha(1) - 2 \inf_{t \in [0,1]} \mathcal{S}_\alpha(t), 2 \sup_{t \in [0,1]} \mathcal{S}_\alpha(t) - \mathcal{S}_\alpha(1) \right),$$

where  $(a_n)_{n \in \mathbb{N}}$  and  $(\mathcal{S}_\alpha(t))_{t \in [0,1]}$  are as in (1.2).

**Remark 1.4.** There is an interesting connection of our results, notably Theorem 1.1, with the work by Wade and coauthors [17, 22, 23] on the convex hulls of planar random walks. Assuming  $\mathbb{E}\xi = 0$ , one can regard the bivariate sequence  $\{(n, S_n)\}_{n \in \mathbb{N}_0}$  as a degenerate walk in the plane whose increments are supported on the line orthogonal to the mean vector  $(1, 0)$  and thus to the  $x$ -axis. Except for this degenerate case, it was shown by Wade and Xu [23, Thms. 1.1 and 1.2] that the perimeter of the convex hull of

the first  $n$  steps of any square-integrable planar random walk satisfies a central limit theorem and has linearly growing variance as  $n \rightarrow \infty$ . But in the degenerate case, their approach only provides that this growth is sublinear. Note that Theorem 1.1 provides no information on the asymptotic behavior of the moments of  $L_n^\smile$ , as it does not claim any type of uniform integrability. The moment asymptotics are specified in Theorem 1.5 below. Remarkably, the variance of  $L_n^\smile$  grows logarithmically if  $\xi$  has a finite third moment but it may grow polynomially when  $\mathbb{E}|\xi|^3 = \infty$ , see (1.8) and (1.9).

Note further that the perimeter  $L_n$  of the convex hull of  $\{(j, S_j) : j = 0, \dots, n\}$  equals  $L_n^\smile + L_n^\frown$ . Under the assumptions of our Theorem 1.3 (Case (C)), we therefore immediately infer a distributional limit result for  $L_n$ , namely

$$\frac{L_n}{2a_n} \xrightarrow[n \rightarrow \infty]{d} \sup_{t \in [0,1]} \mathcal{S}_\alpha(t) - \inf_{t \in [0,1]} \mathcal{S}_\alpha(t).$$

On the other hand and despite relation (1.1), we do not know in the other cases whether joint convergence of  $(L_n^\smile, L_n^\frown)$  holds which would give a limit theorem for the perimeter  $L_n$  in all cases. The connection with random permutations seems to be insufficient for this purpose and we leave this as an open problem.

**Theorem 1.5.** *Suppose that  $\mathbb{E}\xi = 0$  and  $\sigma^2 = \mathbb{E}\xi^2 \in (0, \infty)$ . Then*

$$\mathbb{E}L_n^\smile - n = \frac{\sigma^2}{2} \log n + o(\log n) \quad \text{as } n \rightarrow \infty, \tag{1.6}$$

where the term  $o(\log n)$  can be replaced by  $O(1)$  if and only if  $\mathbb{E}\xi^2 \log^+ |\xi| < \infty$ .

Furthermore, if  $\mathbb{E}|\xi|^p < \infty$  for some  $p \in [2, 3)$ , then

$$\text{Var } L_n^\smile = o(n^{3-p}) \quad \text{as } n \rightarrow \infty \tag{1.7}$$

and also

$$\text{Var } L_n^\smile \geq 0.02n^3 \mathbb{P}(|\xi| \geq 2n) + O(1) \quad \text{as } n \rightarrow \infty. \tag{1.8}$$

Finally, if  $\mathbb{E}|\xi|^3 < \infty$ , then

$$\text{Var } L_n^\smile \simeq \frac{3\sigma^4}{4} \log n \quad \text{as } n \rightarrow \infty. \tag{1.9}$$

**Corollary 1.6.** *If  $\mathbb{E}\xi = 0$  and  $\mathbb{E}|\xi|^p < \infty$ , then  $\text{Var } L_n = o(n^{3-p})$  for  $p \in [2, 3)$  and  $\text{Var } L_n = O(\log n)$  for  $p \geq 3$ , as  $n \rightarrow \infty$ .*

The asymptotics of  $\mathbb{E}L_n^\smile$  in (1.6) have been known before and follow, for example, from Theorem 1.8 in [17] and the fact that  $L_n^\smile \stackrel{d}{=} L_n^\frown$ . The main results here are about the asymptotic behavior of  $\text{Var } L_n^\smile$  and the corollary on the behavior of  $\text{Var } L_n$ . A recent weaker result, Theorem 6.2.6 in [16], asserts that  $\text{Var } L_n$  grows sub-polynomially if  $\xi$  is centered and bounded.

In view of our discussion about the joint law of  $(L_n^\smile, L_n^\frown)$  in Remark 1.4, it is not clear if, under the assumption  $\mathbb{E}\xi^2 < \infty$ , (1.9) could lead to the stronger result  $\text{Var } L_n \simeq c \log n$  for some constant  $c > 0$  as suggested by Conjecture 1.13 in [17]. However, by (1.8),  $\text{Var } L_n^\smile$  has non-logarithmic behavior when  $\mathbb{E}|\xi|^3 = \infty$ , and the same is very plausible for  $\text{Var } L_n$ .

**Remark 1.7.** In view of the functional convergence (1.2), it is natural to ask whether the above theorems can be obtained by applying the continuous mapping theorem to the functional  $F^\smile : D[0, 1] \rightarrow (0, \infty)$  which assigns to each function  $f$  in  $D[0, 1]$  the length of its convex minorant. An approach of this kind has been applied in [15, 17, 22] to various functionals of convex hulls of multidimensional random walks. As we are interested in

the *graph* of a one-dimensional random walk, this functional limit approach cannot be used here because time and space are scaled by different sequences in (1.2). In fact, in the cases (A) and (B), the scaling in time (which is  $n$ ) is stronger than the scaling in space (which is  $a_n$  and thus regularly varying with index  $1/\alpha$ , with  $\alpha = 2$  in case (A)), whereas in case (C) the scaling in space is the stronger one.

There is just one “critical” case where the two scalings coincide and the functional limit approach does work. Assume that the law of  $\xi$  is such that

$$\left( \frac{S(nt)}{a_n} \right)_{t \in [0,1]} \xrightarrow{n \rightarrow \infty} (S_1(t))_{t \in [0,1]} \tag{1.10}$$

in the Skorokhod space  $D[0, 1]$  with the standard  $J_1$ -topology, where  $a_n/n \rightarrow c \in (0, \infty)$  and  $(S_1(t))_{t \in [0,1]}$  is the standard symmetric Cauchy process. Since the functional  $F^\sim$  is continuous on a set of measure 1 with respect to the law of the Cauchy process, the continuous mapping theorem implies that

$$\frac{L_n^\sim}{n} \xrightarrow[n \rightarrow \infty]{d} L_\infty^\sim(c),$$

where  $L_\infty^\sim(c)$  is the length of the convex minorant of the Cauchy process  $(cS_1(t))_{t \in [0,1]}$ . Note that the above argument does not completely cover the domain of attraction of the symmetric Cauchy distribution. Even if we assume that there is no centering, the sequence  $a_n$  in (1.10) is in general of the form  $a_n = n\ell(n)$  with some slowly varying function  $\ell$ .

## 2 Proofs explained

### 2.1 Connection with uniform random permutations

Our approach relies crucially on the following representation of the convex minorant of a random walk, observed already in 1950th by Sparre Andersen [20]. The version presented below is borrowed from [2, Thms. 1 and 2], see also [1, Thm. 1.1], and valid under the assumption that the law of the increment  $\xi$  is continuous.

Set  $[n] := \{1, 2, \dots, n\}$  and let  $\Pi_n$  be a permutation of  $[n]$  picked uniformly at random from the symmetric group  $\mathfrak{S}_n$ , that is

$$\mathbb{P}\{\Pi_n = \pi\} = \frac{1}{n!}, \quad \pi \in \mathfrak{S}_n.$$

Denote by  $Z_{n,1}, Z_{n,2}, \dots, Z_{n,K_n}$  the nonincreasingly ranked cycle lengths of  $\Pi_n$ , with  $K_n$  being the total number of cycles. The convex minorant  $t \mapsto S_n^\sim(t)$  being a piecewise linear function, let  $F_n$  denote the number of intervals where it is linear. Denote by  $C_{n,1}, \dots, C_{n,F_n}$  the nonincreasingly ordered lengths of these intervals (on the horizontal axis). Then the basic result we shall rely on states that

$$(F_n, C_{n,1}, \dots, C_{n,F_n}, 0, 0, \dots) \stackrel{d}{=} (K_n, Z_{n,1}, Z_{n,2}, \dots, Z_{n,K_n}, 0, 0, \dots).$$

Furthermore, given  $(F_n, C_{n,1}, \dots, C_{n,F_n})$ , the increments of the convex minorant over the linearity intervals are conditionally independent and the conditional law of any such increment over an interval of length  $\ell$  equals the law of  $S_\ell$ . In what follows, we formally put  $Z_{n,k} := 0$  for  $k > K_n$ .

For  $j \in [n]$ , let  $K_{n,j}$  be the number of cycles of length  $j$  in  $\Pi_n$ , that is

$$K_{n,j} := \#\{k : Z_{n,k} = j\}, \quad j = 1, \dots, n.$$

Note that

$$\sum_{k=1}^n Z_{n,k} = n \quad K_n = \sum_{j=1}^n K_{n,j}, \quad \text{and} \quad \sum_{j=1}^n jK_{n,j} = n.$$

From the above observations, we immediately derive two equivalent distributional representations for the length of the convex minorant, namely

$$L_n^\sim - n \stackrel{d}{=} \sum_{k=1}^n \left( \sqrt{Z_{n,k}^2 + S_{k,Z_{n,k}}^2} - Z_{n,k} \right), \tag{2.1}$$

and

$$L_n^\sim - n \stackrel{d}{=} \sum_{j=1}^n \sum_{i=1}^{K_{n,j}} \left( \sqrt{j^2 + S_{i,j}^2} - j \right), \tag{2.2}$$

where the  $S_{i,j}$  for  $i \in \mathbb{N}$  and  $j \in \mathbb{N}_0$  are independent random variables that are also independent of  $(Z_{n,k})_{n,k \in \mathbb{N}}$  and satisfy  $S_{i,j} \stackrel{d}{=} S_j$  for all  $i$  and  $j$ . Note that the summand  $-n$  on the left-hand side (matched by the summands  $-Z_{n,k}$  and  $-j$ , respectively, on the right-hand sides) corresponds to the length of the horizontal interval  $[0, n]$  and should be viewed as a very rough first order approximation to the total length  $L_n^\sim$  in the cases (A) and (B).

The following smoothing argument shows that (2.1) and (2.2) do not require that the random walk has continuous increment law. In other words:

*The representations (2.1) and (2.2) remain valid without the continuity assumption.*

Fixing any  $n \in \mathbb{N}$ , consider the random walk  $S_{k;\varepsilon} := \xi_{\varepsilon,1} + \dots + \xi_{\varepsilon,k}$  for  $1 \leq k \leq n$  and any  $\varepsilon > 0$ , where  $\xi_{\varepsilon,k} := \xi_k + \varepsilon N_k$  and  $(N_k)_{k \in \mathbb{N}}$  are i.i.d. standard normal random variables independent of  $(\xi_k)_{k \in \mathbb{N}}$ . Let  $(S_\varepsilon(t))_{t \in [0,n]}$  be its linear interpolation, defined the same way as  $S(t)$  above. The distribution of  $\xi_{\varepsilon,1}$  is continuous, hence the representations (2.1) and (2.2) hold for  $L_{\varepsilon,n}^\sim$ , the length of the convex minorant of  $(S_\varepsilon(t))_{t \in [0,n]}$ . As  $\varepsilon \downarrow 0$ , the process  $(S_\varepsilon(t))_{t \in [0,n]}$  converges to  $(S(t))_{t \in [0,n]}$  weakly in the space  $C[0, n]$ , and since the functional assigning to each continuous function the length of its convex minorant is continuous on  $C[0, n]$  (by the Cauchy–Crofton formula), the continuous mapping theorem implies that  $L_{\varepsilon,n}^\sim$  converges in distribution to  $L_n^\sim$ , as  $\varepsilon \downarrow 0$ . Finally, the claim follows because the right-hand sides of (2.1) and (2.2) for  $(S_\varepsilon(t))_{t \in [0,n]}$  converge, as  $\varepsilon \downarrow 0$ , to the corresponding expressions for  $(S(t))_{t \in [0,n]}$ .

## 2.2 Explanation of the proofs in the cases (A) and (B)

The structure of uniform random permutations is well understood. In particular, see Theorem 1.3. in [3], it is known that

$$(K_{n,1}, K_{n,2}, \dots, K_{n,n}, 0, 0, \dots) \xrightarrow[n \rightarrow \infty]{d} (P_1, P_2, P_3, \dots), \tag{2.3}$$

where the  $P_j$  are mutually independent and the law of  $P_j$  is Poisson with mean  $1/j$ . Moreover, the convergence is fast: there exists a coupling (called the *Feller coupling*) such that

$$\mathbb{E}|K_{n,j} - P_j| \leq \frac{2}{n+1}, \quad j = 1, \dots, n, \tag{2.4}$$

and thus

$$\mathbb{E} \sum_{j=1}^n |K_{n,j} - P_j| < 2, \tag{2.5}$$

see [3, Remark on p. 18 and Eq. (1.26)]. This coupling will be crucial for the proof in case (A). Put

$$Y_j := \sum_{i=1}^{P_j} \left( \sqrt{j^2 + S_{i,j}^2} - j \right)$$

for  $j \in \mathbb{N}$ . Formula (2.5) in conjunction with representation (2.2) strongly suggests that the asymptotic behavior of  $L_n^\sim - n$  should be well approximated by that of the sum

$$V_n := \sum_{j=1}^n Y_j,$$

which is simply a partial sum of independent (but not identically distributed) random variables. The corresponding limit laws are usually called distributions of class  $L$ , but in our case (A) the limit turns out to be normal.

Below we will prove Theorem 1.1 by showing that the distributions of normalized random variables  $L_n^\sim - n$  and  $V_n$  are asymptotically close, and then checking the classical conditions for convergence in distribution of the  $V_n$  after suitable normalization, which are the row sums of triangular arrays whose rows consist of independent random variables.

An interesting observation is that the above argument, based on replacing  $K_{n,j}$  by  $P_j$  in representation (2.2), fails to work in the cases (B) and (C). In order to heuristically explain our approach in case (B), we recall another classical fact from the theory of random permutations, namely (see Vershik and Schmidt [21] or Kingman [14])

$$\left( \frac{Z_{n,1}}{n}, \frac{Z_{n,2}}{n}, \dots \right) \xrightarrow[n \rightarrow \infty]{d} (Z_1, Z_2, \dots), \tag{2.6}$$

where the random sequence  $(Z_1, Z_2, \dots)$  has a Poisson–Dirichlet distribution with parameter  $\theta = 1$ . Furthermore, there exists a coupling such that

$$\mathbb{E} \left( \sum_{k=1}^{\infty} \left| \frac{Z_{n,k}}{n} - Z_k \right| \right) \simeq \frac{\log n}{4n} \quad \text{as } n \rightarrow \infty, \tag{2.7}$$

see [4, Theorem 8.10]. Put  $b_n := a_n^2/n$  and note that regular variation of  $(a_n)_{n \in \mathbb{N}}$  with index  $\frac{1}{\alpha}$  implies regular variation of the sequence  $(b_n)_{n \in \mathbb{N}}$  with index  $\frac{2}{\alpha} - 1 > 0$ . In particular,  $b_n \rightarrow \infty$  if  $\alpha < 2$ . Recalling relation (2.1), we can argue heuristically as follows:

$$\begin{aligned} \frac{L_n^\sim - n}{b_n} &\stackrel{d}{=} \frac{n}{a_n^2} \sum_{k=1}^n \left( \sqrt{Z_{n,k}^2 + S_{k,Z_{n,k}}^2} - Z_{n,k} \right) \\ &= \frac{n}{a_n^2} \sum_{k=1}^n Z_{n,k} \left( \sqrt{1 + \frac{S_{k,Z_{n,k}}^2}{Z_{n,k}^2}} - 1 \right) \\ &\approx \frac{n}{a_n^2} \sum_{k=1}^n \frac{S_{k,Z_{n,k}}^2}{2Z_{n,k}} \\ &\approx \frac{1}{a_n^2} \sum_{k=1}^n \frac{S_{k,[nZ_k]}^2}{2Z_k} \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2} \sum_{k=1}^{\infty} \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{Z_k}, \end{aligned}$$

where the first approximation stems from the one-term Taylor expansion, the second is a consequence of (2.7) and the asserted convergence follows from (1.2), the  $\mathcal{S}_\alpha^{(k)}$  being independent copies of  $\mathcal{S}_\alpha$  which are also independent of  $(Z_1, Z_2, \dots)$ . Proposition 3.2 below will show that the series in the last line is a.s. finite. Moreover, the above heuristic turns out to be correct and this will provide the proof of Theorem 1.2.

### 3 Proofs

#### 3.1 Proof of Theorem 1.1

Let  $\tilde{L}_n^\sim$  denote the random variable on the right-hand side of (2.2) increased by  $n$  so that  $\tilde{L}_n^\sim \stackrel{d}{=} L_n^\sim$ .

As discussed in Section 2.2, the first step of the proof is to show that

$$(\tilde{L}_n^\sim - n - V_n) / \sqrt{\log n} \xrightarrow[n \rightarrow \infty]{P} 0. \tag{3.1}$$

We have

$$\sqrt{j^2 + S_{1,j}^2} - j = \frac{S_{1,j}^2}{j + \sqrt{j^2 + S_{1,j}^2}} \leq \frac{S_{1,j}^2}{2j}, \tag{3.2}$$

in particular

$$\mathbb{E} \left( \sqrt{j^2 + S_{1,j}^2} - j \right) \leq \frac{\mathbb{E} S_{1,j}^2}{2j} = \frac{\sigma^2}{2}. \tag{3.3}$$

Using the definition of  $V_n$ , this entails

$$\begin{aligned} \mathbb{E} \left| \tilde{L}_n^\sim - n - V_n \right| &\leq \sum_{j=1}^n \mathbb{E} \left| \sum_{i=1}^{K_{n,j}} \left( \sqrt{j^2 + S_{i,j}^2} - j \right) - \sum_{i=1}^{P_j} \left( \sqrt{j^2 + S_{i,j}^2} - j \right) \right| \\ &= \sum_{j=1}^n \mathbb{E} |K_{n,j} - P_j| \mathbb{E} \left( \sqrt{j^2 + S_j^2} - j \right) \leq \sigma^2 < \infty, \end{aligned}$$

where the penultimate inequality follows from (2.5) and (3.3). Now (3.1) follows from the last line and the Markov inequality.

The second step is to simplify  $V_n$ . Put

$$W_n := \sum_{j=1}^n \left( \sqrt{j^2 + S_{1,j}^2} - j \right) \mathbb{1}_{\{P_j=1\}}, \quad n \in \mathbb{N}. \tag{3.4}$$

Observe that

$$\sum_{j=1}^{\infty} \mathbb{1}_{\{P_j \geq 2\}} < \infty \quad \text{a.s.}$$

by the Borel–Cantelli lemma. Then

$$\begin{aligned} 0 \leq V_n - W_n &= \sum_{j=1}^n \sum_{i=1}^{P_j} \left( \sqrt{j^2 + S_{i,j}^2} - j \right) \mathbb{1}_{\{P_j \geq 2\}} \\ &\leq \sum_{j=1}^{\infty} \sum_{i=1}^{P_j} \left( \sqrt{j^2 + S_{i,j}^2} - j \right) \mathbb{1}_{\{P_j \geq 2\}} < \infty \quad \text{a.s.}, \end{aligned}$$

which implies that  $(V_n - W_n)_{n \in \mathbb{N}}$  is bounded in probability. Hence, by (3.1), it suffices to prove the theorem for  $W_n$  instead of  $L_n^\sim - n$ .

The third step is to simplify  $W_n$ . Let us rewrite it as

$$\begin{aligned} W_n &= \sum_{j=1}^n j \left( \sqrt{1 + \frac{S_{1,j}^2}{j^2}} - 1 \right) \mathbb{1}_{\{P_j=1\}} \\ &= \sum_{j=1}^n j \left( \frac{S_{1,j}^2}{2j^2} - \frac{S_{1,j}^4}{8j^4} \cdot \theta \left( \frac{S_{1,j}^2}{j^2} \right) \right) \mathbb{1}_{\{P_j=1\}}, \end{aligned} \tag{3.5}$$



where  $\theta : [0, \infty) \rightarrow [0, 1]$  is a continuous bounded function resulting from Lagrange's form of the remainder in the Taylor expansion of  $x \mapsto \sqrt{1+x}$  at  $x = 0$ . We claim that

$$\frac{1}{\sqrt{\log n}} \sum_{j=1}^n \frac{S_{1,j}^4}{j^3} \mathbb{1}_{\{P_j=1\}} \xrightarrow[n \rightarrow \infty]{P} 0, \tag{3.6}$$

for which it obviously suffices to verify that

$$\sum_{j=1}^{\infty} \frac{S_{1,j}^4}{j^3} \mathbb{1}_{\{P_j=1\}} < \infty \text{ a.s.} \tag{3.7}$$

Put  $p_j := \mathbb{P}\{P_j = 1\}$  and note that

$$\frac{1}{j} \geq p_j = e^{-1/j} \frac{1}{j} = \frac{1}{j} + O\left(\frac{1}{j^2}\right) \text{ as } j \rightarrow \infty. \tag{3.8}$$

As  $\mathbb{E}\xi = 0$  and  $\sigma^2 = \mathbb{E}\xi^2 < \infty$ , it follows by [11, Thm. 10.2 on p. 46] (or [8, Thm. 4]) that

$$\sum_{j=1}^{\infty} \frac{1}{j} \mathbb{P}\left\{|S_j| > a\sigma\sqrt{j \log \log j}\right\} < \infty \tag{3.9}$$

for any  $a > \sqrt{2}$ . By combining this and the Borel–Cantelli lemma, only finitely many events  $\{|S_{1,j}| > 2\sigma\sqrt{j \log \log j}, P_j = 1\}$ ,  $j \in \mathbb{N}$ , occur with probability 1. Now the proof of (3.7), which in turn implies (3.6), can be completed by using the inequality

$$\sum_{j=1}^{\infty} \frac{S_{1,j}^4}{j^3} \mathbb{1}_{\{P_j=1\}} \leq \sum_{j=1}^{\infty} \frac{S_{1,j}^4}{j^3} \mathbb{1}_{\{|S_{1,j}| > 2\sigma\sqrt{j \log \log j}, P_j=1\}} + \sum_{j=1}^{\infty} \frac{16\sigma^4}{j} (\log \log j)^2 \mathbb{1}_{\{P_j=1\}}.$$

The first sum on the right-hand side is finite a.s. since it contains a.s. only finitely many non-zero terms, and the second sum is finite a.s. because it has finite expectation.

In view of (3.5) and (3.6), it suffices to prove Theorem 1.1 with  $L_n^\sim - n$  replaced by

$$W'_n := \sum_{j=1}^n \frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1\}}.$$

Therefore, as the fourth step of the proof, we will show that

$$\frac{1}{\sqrt{\log n}} \left( W'_n - \sum_{j=1}^n \frac{\sigma_j^2}{2j} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}\left(0, \frac{3\sigma^4}{4}\right). \tag{3.10}$$

Even though the variables  $W'_n$  are sums of independent random variables of a rather simple structure, it is not easy to obtain a central limit theorem for  $W'_n$  without additional moment assumptions. For example, when trying to verify the Lindeberg condition, the fourth moment of  $S_{1,j}$  appears, which is not assumed to be finite here.

We use a general result [18, Thm. 18 in Chap. IV, §4], which ensures convergence of row sums in triangular arrays of random variables to a normal distribution. According to the three conditions of the theorem, we need to check that

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbb{P}\left\{\frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1\}} \geq \varepsilon\sqrt{\log n}\right\} = 0 \text{ for every } \varepsilon > 0, \tag{3.11}$$

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{j=1}^n \text{Var}\left(\frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1, S_{1,j}^2/(2j) < \sqrt{\log n}\}}\right) = \frac{3\sigma^4}{4}, \tag{3.12}$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{\log n}} \left[ \sum_{j=1}^n \mathbb{E} \left( \frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1, S_{1,j}^2/(2j) < \sqrt{\log n}\}} \right) - \sum_{j=1}^n \frac{\sigma_j^2}{2j} \right] = 0. \quad (3.13)$$

Our idea is to approximate  $S_{1,j}/\sqrt{j}$  by  $\sigma_j \mathcal{N}$ , where  $\mathcal{N}$  is a standard normal variable and  $\sigma_j^2$  is given by (1.3). We claim that, for any  $\gamma \geq 0$  and  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{(\log n)^{\gamma/2}} \sum_{j=1}^n \frac{1}{j} \left[ \mathbb{E} \left( \frac{|S_{1,j}|^{2\gamma}}{j^\gamma} \mathbb{1}_{\{S_{1,j}^2/j < \varepsilon \sqrt{\log n}\}} \right) - \mathbb{E} \left( \sigma_j^{2\gamma} |\mathcal{N}|^{2\gamma} \mathbb{1}_{\{\sigma_j^2 \mathcal{N}^2 < \varepsilon \sqrt{\log n}\}} \right) \right] = 0. \quad (3.14)$$

To prove this, we first note that

$$\mathbb{E}(X^\gamma \mathbb{1}_{\{X < a\}}) = - \int_{[0,a)} x^\gamma d\mathbb{P}\{X \geq x\} = -a^\gamma \mathbb{P}\{X \geq a\} + \gamma a^\gamma \int_0^1 t^{\gamma-1} \mathbb{P}\{X \geq at\} dt \quad (3.15)$$

for any non-negative random variable  $X$  and  $a, \gamma > 0$ . Upon applying this formula twice with  $X = S_{1,j}^2/j$  and  $X = \sigma_j^2 \mathcal{N}^2$ , we obtain

$$\begin{aligned} & \frac{1}{a^\gamma} \left| \mathbb{E} \left( \frac{|S_{1,j}|^{2\gamma}}{j^\gamma} \mathbb{1}_{\{S_{1,j}^2/j < a\}} \right) - \mathbb{E}(\sigma_j^{2\gamma} |\mathcal{N}|^{2\gamma} \mathbb{1}_{\{\sigma_j^2 \mathcal{N}^2 < a\}}) \right| \\ & \leq \left| \mathbb{P} \left\{ \frac{S_{1,j}^2}{j} \geq a \right\} - \mathbb{P} \{ \sigma_j^2 \mathcal{N}^2 \geq a \} \right| + \gamma \int_0^1 t^{\gamma-1} \left| \mathbb{P} \left\{ \frac{S_{1,j}^2}{j} \geq at \right\} - \mathbb{P} \{ \sigma_j^2 \mathcal{N}^2 \geq at \} \right| dt. \end{aligned}$$

Then, taking  $a = \varepsilon \sqrt{\log n}$ , we see that for every  $j \leq n$ ,

$$\begin{aligned} & \frac{1}{j(\log n)^{\gamma/2}} \left[ \mathbb{E} \left( \frac{|S_{1,j}|^{2\gamma}}{j^\gamma} \mathbb{1}_{\{S_{1,j}^2/j < \varepsilon \sqrt{\log n}\}} \right) - \mathbb{E} \left( \sigma_j^{2\gamma} |\mathcal{N}|^{2\gamma} \mathbb{1}_{\{\sigma_j^2 \mathcal{N}^2 < \varepsilon \sqrt{\log n}\}} \right) \right] \\ & \leq 4\varepsilon^\gamma \cdot \frac{1}{j} \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left\{ \frac{S_{1,j}}{\sqrt{j}} < x \right\} - \mathbb{P} \{ \sigma_j \mathcal{N} < x \} \right|. \quad (3.16) \end{aligned}$$

Clearly, this bound is also valid for  $\gamma = 0$ . On the other hand, the dominated convergence theorem ensures

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{j(\log n)^{\gamma/2}} \left( \frac{|S_{1,j}|^{2\gamma}}{j^\gamma} \mathbb{1}_{\{S_{1,j}^2/j < \varepsilon \sqrt{\log n}\}} - \sigma_j^{2\gamma} |\mathcal{N}|^{2\gamma} \mathbb{1}_{\{\sigma_j^2 \mathcal{N}^2 < \varepsilon \sqrt{\log n}\}} \right) \right] = 0,$$

for every  $j \in \mathbb{N}$  because the random variable in square brackets is bounded by  $2\varepsilon^\gamma/j$ . Since, furthermore, the sequence on the right-hand side of (3.16) is summable over  $j \in \mathbb{N}$ , see [8, p. 1480] and [18, p. 130],<sup>1</sup> (3.14) follows by another appeal to the dominated convergence theorem.

We can now finish the proof of (3.10). Recall that  $P_j$  is independent of  $S_{1,j}$ . Using (3.8) and then (3.14) with  $\gamma = 0$ , we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbb{P} \left\{ \frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1\}} \geq \varepsilon \sqrt{\log n} \right\} &= \lim_{n \rightarrow \infty} \sum_{j=1}^n \frac{1}{j} \mathbb{P} \left\{ \frac{S_{1,j}^2}{2j} \geq \varepsilon \sqrt{\log n} \right\} \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^n \frac{1}{j} \mathbb{P} \left\{ \sigma_j^2 \mathcal{N}^2 \geq 2\varepsilon \sqrt{\log n} \right\}. \end{aligned}$$

<sup>1</sup>This result was actually proved in [10] but stated there without uniformity in  $x$ .

The last limit is equal to zero since  $\lim_{j \rightarrow \infty} \sigma_j^2 = \sigma^2 < \infty$  and in view of the following simple Markov-type inequality

$$\mathbb{E} \left( \sigma_j^{2\gamma} \mathcal{N}^{2\gamma} \mathbb{1}_{\{\sigma_j^2 \mathcal{N}^2 \geq a\}} \right) \leq a^{-3} \sigma_j^{2\gamma+6} \mathbb{E} \mathcal{N}^{2\gamma+6} < \infty, \tag{3.17}$$

which holds for all  $j \in \mathbb{N}$ ,  $\gamma \geq 0$ ,  $a > 0$ , and which we applied with  $\gamma = 0$ . This finishes the proof of (3.11).

Furthermore, we have

$$\begin{aligned} & \sum_{j=1}^n \text{Var} \left( \frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1, S_{1,j}^2/(2j) < \sqrt{\log n}\}} \right) \\ &= \sum_{j=1}^n p_j \mathbb{E} \left( \frac{S_{1,j}^4}{4j^2} \mathbb{1}_{\{S_{1,j}^2/j < 2\sqrt{\log n}\}} \right) - \sum_{j=1}^n \left( p_j \mathbb{E} \left( \frac{S_{1,j}^2}{2j} \mathbb{1}_{\{S_{1,j}^2/j < 2\sqrt{\log n}\}} \right) \right)^2. \end{aligned}$$

The second sum on the right-hand side is increasing in  $n$  and bounded by  $\sum_{j=1}^{\infty} \frac{\sigma^4}{4j^2} < \infty$  in view of (3.8). Appealing once again to (3.8) and using (3.14) with  $\gamma = \varepsilon = 2$ , we obtain

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{j=1}^n \text{Var} \left( \frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1, S_{1,j}^2/(2j) < \sqrt{\log n}\}} \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{j=1}^n p_j \mathbb{E} \left( \frac{S_{1,j}^4}{4j^2} \mathbb{1}_{\{S_{1,j}^2/j < 2\sqrt{\log n}\}} \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{j=1}^n \frac{1}{j} \mathbb{E} \left( \frac{S_{1,j}^4}{4j^2} \mathbb{1}_{\{S_{1,j}^2/j < 2\sqrt{\log n}\}} \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{4 \log n} \sum_{j=1}^n \frac{1}{j} \mathbb{E} \left( \sigma_j^4 \mathcal{N}^4 \mathbb{1}_{\{\sigma_j^2 \mathcal{N}^2 < 2\sqrt{\log n}\}} \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{4 \log n} \sum_{j=1}^n \frac{3\sigma_j^4}{j} = \frac{3\sigma^4}{4} \end{aligned}$$

using that the family  $(\sigma_j^4 \mathcal{N}^4)_{j \in \mathbb{N}}$  is uniformly integrable by  $\sigma_j^2 \rightarrow \sigma^2$ . This proves (3.12).

Finally, from (3.14) and (3.17) with  $\gamma = 1$ ,  $\varepsilon = 2$ , and  $a = 2\sqrt{\log n}$ ,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{\sqrt{\log n}} \left[ \sum_{j=1}^n \mathbb{E} \left( \frac{S_{1,j}^2}{2j} \mathbb{1}_{\{P_j=1, S_{1,j}^2/(2j) < \sqrt{\log n}\}} \right) - \sum_{j=1}^n \frac{\sigma_j^2}{2j} \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{2\sqrt{\log n}} \sum_{j=1}^n \frac{1}{j} \left[ \mathbb{E} \left( \sigma_j^2 \mathcal{N}^2 \mathbb{1}_{\{\sigma_j^2 \mathcal{N}^2 < 2\sqrt{\log n}\}} \right) - \sigma_j^2 \right] = 0. \end{aligned}$$

This proves (3.13). The proof of (3.10) is herewith complete. Thus, we established (1.4).

If  $\mathbb{E} \xi^2 \log^+ |\xi| < \infty$ , then  $\sigma_j^2$  can be replaced by  $\sigma^2$  throughout the proof, implying (1.5). The key observation is that the modified sequence on the right-hand side of estimate (3.16) remains summable over  $j \in \mathbb{N}$  by Lemma 1 and the Theorems on p. 1480 in [8]. This finishes the proof of Theorem 1.1.

### 3.2 Proof of Theorem 1.5

We first give explicit formulae for  $\mathbb{E} L_n^\sim$  and  $\text{Var} L_n^\sim$  and put for simplicity

$$\eta_j := \sqrt{j^2 + S_j^2} - j, \quad j \in \mathbb{N}.$$

**Lemma 3.1.** *If  $\mathbb{E}\xi = 0$  and  $\sigma^2 = \mathbb{E}\xi^2 < \infty$ , then*

$$\mathbb{E}L_n^\sim - n = \sum_{j=1}^n \frac{\mathbb{E}\eta_j}{j} \quad \text{and} \quad \text{Var} L_n^\sim = \sum_{j=1}^n \frac{\mathbb{E}\eta_j^2}{j} + O(1) \quad \text{as } n \rightarrow \infty.$$

*Proof.* It will be used that for any integers  $1 \leq j \neq k \leq n$ ,

$$\mathbb{E}K_{n,j} = \frac{1}{j}, \quad \mathbb{E}(K_{n,j}(K_{n,j} - 1)) = \frac{\mathbb{1}_{\{2j \leq n\}}}{j^2} \quad \text{and} \quad \mathbb{E}(K_{n,j}K_{n,k}) = \frac{\mathbb{1}_{\{j+k \leq n\}}}{jk}; \quad (3.18)$$

see [3, Lemma 1.1]. Thus,  $K_{n,j}$  and  $K_{n,k}$  are uncorrelated whenever  $j + k \leq n$ .

The formula for  $\mathbb{E}L_n^\sim$  follows immediately from representation (2.2) and (3.18). For the variance of  $L_n^\sim$ , we obtain with the help of the formula for the variance of a random sum of i.i.d. random variables:

$$\begin{aligned} \text{Var} \left( \sum_{i=1}^{K_{n,j}} \left( \sqrt{j^2 + S_{i,j}^2} - j \right) \right) &= \mathbb{E}K_{n,j} \text{Var} \eta_j + (\mathbb{E}\eta_j)^2 \text{Var} K_{n,j} \\ &= \frac{\mathbb{E}\eta_j^2}{j} - \frac{(\mathbb{E}\eta_j)^2}{j} + (\mathbb{E}\eta_j)^2 \left( \frac{j - \mathbb{1}_{\{2j > n\}}}{j^2} \right) = \frac{\mathbb{E}\eta_j^2}{j} - \frac{(\mathbb{E}\eta_j)^2 \mathbb{1}_{\{2j > n\}}}{j^2}. \end{aligned}$$

Similarly, by conditioning on  $K_{n,j}$  and  $K_{n,k}$  and setting  $\gamma_{jk}^{(n)} := \text{Cov}(K_{n,j}, K_{n,k})$ ,

$$\text{Cov} \left( \sum_{i=1}^{K_{n,j}} \left( \sqrt{j^2 + S_{i,j}^2} - j \right), \sum_{i=1}^{K_{n,k}} \left( \sqrt{k^2 + S_{i,k}^2} - k \right) \right) = \gamma_{jk}^{(n)} \mathbb{E}\eta_j \mathbb{E}\eta_k.$$

Combining these formulas with (2.2) and (3.18), we obtain

$$\begin{aligned} \text{Var}(L_n^\sim) &= \sum_{j=1}^n \text{Var} \left( \sum_{i=1}^{K_{n,j}} \left( \sqrt{j^2 + S_{i,j}^2} - j \right) \right) + \sum_{1 \leq j \neq k \leq n} \gamma_{jk}^{(n)} \mathbb{E}\eta_j \mathbb{E}\eta_k \\ &= \sum_{j=1}^n \frac{\mathbb{E}\eta_j^2}{j} - \sum_{\substack{1 \leq j \leq n: \\ 2j > n}} \frac{(\mathbb{E}\eta_j)^2}{j^2} - \sum_{\substack{1 \leq j \neq k \leq n: \\ j+k > n}} \frac{\mathbb{E}\eta_j \mathbb{E}\eta_k}{jk} \\ &= \sum_{j=1}^n \frac{\mathbb{E}\eta_j^2}{j} - \sum_{\substack{1 \leq j, k \leq n: \\ j+k > n}} \frac{\mathbb{E}\eta_j \mathbb{E}\eta_k}{jk}. \end{aligned}$$

The last term on the right-hand side can be estimated by using inequality (3.3), viz.

$$\begin{aligned} 0 &\leq \sum_{\substack{1 \leq j, k \leq n: \\ j+k > n}} \frac{\mathbb{E}\eta_j \mathbb{E}\eta_k}{jk} \leq \frac{\sigma^4}{4} \sum_{\substack{1 \leq j, k \leq n: \\ j+k > n}} \frac{1}{jk} = \frac{\sigma^4}{4} \sum_{j=1}^n \frac{1}{j} \sum_{k=n-j+1}^n \frac{1}{k} \\ &= \frac{\sigma^4}{4} \sum_{j=1}^{n-1} \frac{1}{j} \sum_{k=n-j+1}^n \frac{1}{k} + O\left(\frac{\log n}{n}\right) \\ &\leq \frac{\sigma^4}{4} \sum_{j=1}^{n-1} \frac{1}{j} \sum_{k=n-j+1}^n \int_{k-1}^k \frac{dx}{x} + O\left(\frac{\log n}{n}\right) \\ &= \frac{\sigma^4}{4} \sum_{j=1}^{n-1} \frac{\log n - \log(n-j)}{j} + O\left(\frac{\log n}{n}\right) \\ &= -\frac{\sigma^4}{4n} \sum_{j=1}^{n-1} \left(\frac{j}{n}\right)^{-1} \log\left(1 - \frac{j}{n}\right) + O\left(\frac{\log n}{n}\right) \\ &= -\frac{\sigma^4}{4} \int_0^1 x^{-1} \log(1-x) dx + O\left(\frac{\log n}{n}\right), \end{aligned}$$

as  $n \rightarrow \infty$ , where the last passage follows from formula (21) in [13] with  $b = k = 1$ . This proves the formula for  $\text{Var } L_n^\sim$  because the last integral is finite.  $\square$

We are ready to prove our claims on the asymptotics of the moments of  $L_n^\sim$ .

*Proof of Theorem 1.5.* By the equality in (3.2), the law of large numbers and the central limit theorem,

$$\eta_n \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2} \mathcal{N}^2(0, \sigma^2). \tag{3.19}$$

Under the assumptions  $\mathbb{E}\xi^2 < \infty$  and  $\mathbb{E}\xi = 0$ , the family  $(S_n^2/n)_{n \in \mathbb{N}}$  is uniformly integrable [11, Thm. 1.6.3] whence, by the inequality in (3.2), the same holds for  $(\eta_n)_{n \in \mathbb{N}}$ . In conjunction with (3.19), this yields

$$\lim_{n \rightarrow \infty} \mathbb{E}\eta_n = \frac{1}{2} \mathbb{E}\mathcal{N}^2(0, \sigma^2) = \frac{\sigma^2}{2}. \tag{3.20}$$

Combined with Lemma 3.1, this gives the asymptotics of  $\mathbb{E}L_n^\sim - n$  stated in (1.6).

For the proof of the remaining claims, let us introduce the function

$$g(x) := x^2/2 - \sqrt{1+x^2} + 1, \quad x \geq 0,$$

which satisfies  $g(x) \simeq x^4/8$  as  $x \rightarrow 0$ , and is non-negative and strictly increasing for  $x > 0$  since  $g'(x) > 0$ . We have

$$\frac{\sigma^2}{2} - \mathbb{E}\eta_j = \mathbb{E} \left( \frac{S_j^2}{2j} - \left( \sqrt{j^2 + S_j^2} - j \right) \right) = j \mathbb{E}g \left( \frac{|S_j|}{j} \right), \tag{3.21}$$

hence by Lemma 3.1,

$$\mathbb{E}L_n^\sim - n = \sum_{j=1}^n \frac{\mathbb{E}\eta_j}{j} = \sum_{j=1}^n \frac{\sigma^2}{2j} - \sum_{j=1}^n \mathbb{E}g \left( \frac{|S_j|}{j} \right).$$

To study convergence of the last sum, we employ [19, Thm. 1] which reads as follows: there exist positive constants  $C_1, C_2$  such that for every  $x > 0$ ,

$$C_1 x^{-2} \mathbb{E}(\xi^2 \mathbb{1}_{\{|\xi| \geq x\}}) \leq F(x) := \sum_{j=1}^{\infty} \mathbb{P}\{|S_j| \geq xj\} \leq C_2 x^{-2} \mathbb{E}(\xi^2 \mathbb{1}_{\{|\xi| \geq x\}}). \tag{3.22}$$

By Fubini's theorem, we have

$$\sum_{j=1}^{\infty} \mathbb{E}g(|S_j|/j) = \int_0^{\infty} g'(x) F(x) dx.$$

Therefore, by (3.22) the right-hand side is finite if and only if the integrals

$$\int_0^{\infty} \frac{g'(x)}{x^2} \mathbb{E}(\xi^2 \mathbb{1}_{\{|\xi| \geq x\}}) dx = \int_0^{\infty} t^2 \left( \int_0^t \frac{g'(x)}{x^2} dx \right) \mathbb{P}\{|\xi| \in dt\} \tag{3.23}$$

are finite. The function  $g'(x)/x^2$  is integrable at  $x = 0$  since  $g'(x) \simeq x^3/2$  as  $x \rightarrow 0$ . On the other hand, since  $g'(x) = x - 1 + o(1)$  as  $x \rightarrow \infty$ , we have  $\int_0^t (g'(x)/x^2) dx = \log t + O(1)$  as  $t \rightarrow \infty$ . Hence, by  $\mathbb{E}\xi^2 < \infty$ , the right-hand side of (3.23) is finite if and only if  $\mathbb{E}\xi^2 \log^+ |\xi| < \infty$ , which is thus equivalent to the asymptotic relation  $\mathbb{E}L_n^\sim = n + \frac{1}{2}\sigma^2 \log n + O(1)$  as  $n \rightarrow \infty$ .

Let us now prove the claims about  $\text{Var } L_n^\smile$ . Note that, by (3.21),

$$\mathbb{E}\eta_j^2 = \mathbb{E}\left(\sqrt{j^2 + S_j^2} - j\right)^2 = \mathbb{E}\left(S_j^2 - 2j\left(\sqrt{j^2 + S_j^2} - j\right)\right) = 2j^2\mathbb{E}g\left(\frac{|S_j|}{j}\right).$$

Thus, by Lemma 3.1,

$$\text{Var}(L_n^\smile) = \sum_{j=1}^n \frac{\mathbb{E}\eta_j^2}{j} + O(1) = 2 \sum_{j=1}^n j\mathbb{E}g\left(\frac{|S_j|}{j}\right) + O(1) \quad \text{as } n \rightarrow \infty. \quad (3.24)$$

In particular, by (3.20) and (3.21), this implies  $\text{Var}(L_n^\smile) = o(n)$ , that is, (1.7) holds for  $p = 2$ .

From now on we assume  $\mathbb{E}|\xi|^p < \infty$  for some  $p \in (2, 3]$ . To prove the remaining claims we first need to estimate truncated moments of  $S_j$ . Let us show that for every  $\varepsilon > 0$ ,

$$\mathbb{E}(S_j^2 \mathbb{1}_{\{|S_j| \geq \varepsilon j\}}) = o(j^{3-p}) \quad \text{as } j \rightarrow \infty, \quad (3.25)$$

and

$$j^{-2}\mathbb{E}(S_j^4 \mathbb{1}_{\{|S_j| < \varepsilon j\}}) = \begin{cases} o(j^{3-p}), & p \in (2, 3) \\ 3\sigma^4 + \varepsilon c(j, \varepsilon), & p = 3 \end{cases} \quad \text{as } j \rightarrow \infty, \quad (3.26)$$

where  $\limsup_{j \rightarrow \infty} |c(j, \varepsilon)|$  is uniformly bounded in  $\varepsilon > 0$ .

To this end, we employ [18, Claim 22 in Chap. V, §5]: for every  $p \in (2, 3)$ ,  $j \in \mathbb{N}$ , and  $x \in \mathbb{R}$ ,

$$\left| \mathbb{P}\left\{\frac{S_j}{\sigma\sqrt{j}} < x\right\} - \mathbb{P}\{\mathcal{N} < x\}\right| \leq \frac{\psi(\sqrt{j}(1+|x|))}{j^{p/2-1}(1+|x|)^p}, \quad (3.27)$$

where  $\psi$  is a bounded decreasing function on  $[1, \infty)$  such that  $\psi(x) \rightarrow 0$  as  $x \rightarrow \infty$ . This inequality is also valid for  $p = 3$  with a constant function  $\psi$ , see [18, Thm. 14 in Chap. V].

Notice that for any non-negative random variable  $X$  such that  $\mathbb{E}X^2 < \infty$  and any  $a > 0$ ,

$$\mathbb{E}(X^2 \mathbb{1}_{\{X \geq a\}}) = - \int_{[a, \infty)} x^2 d\mathbb{P}\{X \geq x\} = a^2\mathbb{P}\{X \geq a\} + 2 \int_a^\infty x\mathbb{P}\{X \geq x\} dx.$$

Using this formula twice with  $X = \frac{|S_j|}{\sigma\sqrt{j}}$  and  $X = |\mathcal{N}|$ , we obtain from (3.27)

$$\begin{aligned} & \left| \mathbb{E}\left(\frac{S_j^2}{\sigma^2 j} \mathbb{1}_{\{|S_j| \geq \sigma a \sqrt{j}\}}\right) - \mathbb{E}(\mathcal{N}^2 \mathbb{1}_{\{|\mathcal{N}| \geq a\}}) \right| \\ & \leq \frac{2a^2\psi(\sqrt{j})}{j^{p/2-1}a^p} + 2 \int_a^\infty \frac{2x\psi(\sqrt{j})}{j^{p/2-1}x^p} dx = \frac{2pa^{2-p}\psi(\sqrt{j})}{(p-2)j^{p/2-1}}. \end{aligned} \quad (3.28)$$

This implies (3.25) for  $p \in (2, 3)$  upon taking  $a = \varepsilon\sqrt{j}/\sigma$  and applying (3.17) with  $\gamma = 1$  and  $\sigma_j$  replaced by  $\sigma$ . However, for  $p = 3$ , this gives only  $\mathbb{E}(S_j^2 \mathbb{1}_{\{|S_j| \geq \varepsilon j\}}) = O(1)$ . In order to replace  $O(1)$  by  $o(1)$ , pick  $R > \varepsilon$  and write

$$\mathbb{E}(S_j^2 \mathbb{1}_{\{|S_j| \geq \varepsilon j\}}) \leq R^2 j^2 \mathbb{P}\{|S_j| \geq \varepsilon j\} + \mathbb{E}(S_j^2 \mathbb{1}_{\{|S_j| \geq Rj\}}).$$

The first summand is  $o(1)$  by a standard estimate of the rate of convergence in the law of large numbers (see [18, Theorem 28 in Chap. IX]). The second summand is bounded by  $6\sigma\psi(1)R^{-1}$  in view of (3.28) and (3.17) applied with  $a = R\sqrt{j}/\sigma$ . This implies (3.25) for  $p = 3$  since  $R$  can be arbitrarily large.

Similarly, from (3.15) and (3.27) it follows that

$$\left| \mathbb{E}\left(\frac{S_j^4}{\sigma^4 j^2} \mathbb{1}_{\{|S_j| < \sigma a \sqrt{j}\}}\right) - \mathbb{E}(\mathcal{N}^4 \mathbb{1}_{\{|\mathcal{N}| < a\}}) \right| \leq \frac{2(8-p)a^{4-p}\psi(\sqrt{j})}{(4-p)j^{p/2-1}}.$$

This implies (3.26) for  $p \in (2, 3]$  upon taking  $a = \varepsilon\sqrt{j}/\sigma$ , where for  $p = 3$ , we put

$$c(j, \varepsilon) := \varepsilon^{-1} \left[ \mathbb{E} \left( \frac{S_j^4}{j^2} \mathbb{1}_{\{|S_j| < \varepsilon j\}} \right) - 3\sigma^4 \right],$$

which satisfies  $|c(j, \varepsilon)| \leq 10\sigma^3\psi(1) + \sigma^4\varepsilon^{-1}\mathbb{E}(\mathcal{N}^4 \mathbb{1}_{\{|\mathcal{N}| \geq \varepsilon\sqrt{j}/\sigma\}})$ , hence  $\limsup_{j \rightarrow \infty} |c(j, \varepsilon)|$  is uniformly bounded in  $\varepsilon > 0$ , as required.

We can now prove an upper bound for  $\text{Var}(L_n^\sim)$ . Using that  $g(x) \leq \min(x^2/2, x^4/8)$ , we have, for every  $\varepsilon > 0$ ,

$$j\mathbb{E}g\left(\frac{|S_j|}{j}\right) \leq \mathbb{E}\left(\frac{S_j^4}{8j^3} \mathbb{1}_{\{|S_j| < \varepsilon j\}}\right) + \mathbb{E}\left(\frac{S_j^2}{2j} \mathbb{1}_{\{|S_j| \geq \varepsilon j\}}\right).$$

Thus, (3.25) and (3.26) yield

$$j\mathbb{E}g\left(\frac{|S_j|}{j}\right) \leq \begin{cases} o(j^{2-p}), & p \in (2, 3) \\ \frac{1}{8j}(3\sigma^4 + \varepsilon c(j, \varepsilon) + o(1)), & p = 3 \end{cases} \quad \text{as } j \rightarrow \infty.$$

Upon summation over  $j$  and using (3.24) we derive  $\text{Var}(L_n^\sim) = o(n^{3-p})$  for  $p \in (2, 3)$ , thereby completing the proof of (1.7). Similarly, for  $p = 3$ , summing over  $j$ , dividing by  $\log n$ , sending  $n \rightarrow \infty$  and then  $\varepsilon \rightarrow 0$ , we obtain  $\limsup_{n \rightarrow \infty} \frac{\text{Var}(L_n^\sim)}{\log n} \leq \frac{3\sigma^4}{4}$ .

To establish the matching lower bound for  $p = 3$  we argue as follows. Since  $g(x) \simeq x^4/8$ , as  $x \rightarrow 0$ , for every fixed  $\varepsilon_1 > 0$  we can find  $\delta_1 = \delta_1(\varepsilon_1) > 0$  such that

$$1 - \varepsilon_1 \leq \frac{8g(x)}{x^4} \quad \text{for } 0 < x \leq \delta_1.$$

Therefore,

$$\begin{aligned} \liminf_{j \rightarrow \infty} \mathbb{E}\left(j^2 g\left(\frac{|S_j|}{j}\right)\right) &\geq \liminf_{j \rightarrow \infty} \mathbb{E}\left(j^2 g\left(\frac{|S_j|}{j}\right) \mathbb{1}_{\{|S_j| \leq \delta_1 j\}}\right) \\ &\geq (1 - \varepsilon_1) \liminf_{j \rightarrow \infty} \mathbb{E}\left(\frac{S_j^4}{8j^2} \mathbb{1}_{\{|S_j| \leq \delta_1 j\}}\right). \end{aligned}$$

By (3.26) and upon sending  $\varepsilon_1$  to zero, this yields

$$\liminf_{j \rightarrow \infty} \mathbb{E}\left(j^2 g\left(\frac{|S_j|}{j}\right)\right) \geq \frac{3\sigma^4}{8}.$$

Thus arriving at  $\liminf_{n \rightarrow \infty} \frac{\text{Var}(L_n^\sim)}{\log n} \geq \frac{3}{4}\sigma^4$ , we have completed the proof of (1.9).

It remains to prove the lower bound (1.8) for  $\text{Var} L_n^\sim$ . We note that for any  $j \in \mathbb{N}$  (see [7, Lemma 2.1] or [9, p. 290]),

$$\mathbb{P}\{|S_j| \geq j\} \geq j\rho_j \mathbb{P}\{|\xi| \geq 2j\},$$

where  $\rho_j := \mathbb{P}\{|S_{j-1}| < j\} - j\mathbb{P}\{|\xi| \geq 2j\}$  and  $S_0 := 0$ . As  $g$  is an increasing function, this yields

$$\mathbb{E}g(|S_j|/j) \geq g(1)\mathbb{P}\{|S_j| \geq j\} \geq g(1)j\rho_j \mathbb{P}\{|\xi| \geq 2j\}.$$

Hence, we obtain by (3.24)

$$\text{Var}(L_n^\sim) = 2 \sum_{j=1}^n j\mathbb{E}g(|S_j|/j) + O(1) \geq 2g(1) \sum_{j=1}^n j^2 \rho_j \mathbb{P}\{|\xi| \geq 2j\} + O(1)$$

as  $n \rightarrow \infty$ . This gives the desired bound  $\text{Var}(L_n^\sim) \geq 0.02n^3\mathbb{P}\{|\xi| \geq 2n\} + O(1)$  since  $2g(1)/3 > 0.02$  and  $\lim_{j \rightarrow \infty} \rho_j = 1$ . The last statement holds by the assumptions  $\mathbb{E}\xi^2 < \infty$  and  $\mathbb{E}\xi = 0$ . This finishes the proof of Theorem 1.5.  $\square$

### 3.3 Proof of Theorem 1.2

We start by showing that the random series representing the limit random variable in Theorem 1.2 is finite a.s.

**Proposition 3.2.** *Let  $(Z_1, Z_2, \dots)$  be a random sequence having a Poisson–Dirichlet distribution with parameter  $\theta = 1$ . Further, let  $\mathcal{S}_\alpha^{(k)} = (\mathcal{S}_\alpha^{(k)}(t))_{t \in [0,1]}$ ,  $k = 1, 2, \dots$ , be independent copies of the strictly stable process  $\mathcal{S}_\alpha$  with index  $\alpha \in (1, 2)$ . Then,*

$$\sum_{k=1}^{\infty} \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{Z_k} < \infty \quad \text{a.s.}$$

*Proof.* Fix an arbitrary  $\delta \in (0, \alpha/2)$ . Using that  $x \mapsto x^\delta$  is subadditive and then the self-similarity of the process  $\mathcal{S}_\alpha$ , we obtain

$$\mathbb{E} \left( \sum_{k=1}^{\infty} \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{Z_k} \right)^\delta \leq \sum_{k=1}^{\infty} \mathbb{E} \left( \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{Z_k} \right)^\delta = \mathbb{E} |\mathcal{S}_\alpha(1)|^{2\delta} \sum_{k=1}^{\infty} \mathbb{E} Z_k^{(2/\alpha-1)\delta}.$$

Since  $\mathbb{E} |\mathcal{S}_\alpha(1)|^{2\delta} < \infty$  (as  $2\delta < \alpha$ ), it remains to check that  $\sum_{k=1}^{\infty} \mathbb{E} Z_k^{(2/\alpha-1)\delta}$  is finite. To this end, formula (2.1) from [4] with  $\phi(x) = x^{(2/\alpha-1)\delta}$  can be used to see that

$$\sum_{k=1}^{\infty} \mathbb{E} Z_k^{(2/\alpha-1)\delta} = \int_0^1 \frac{\phi(x)}{x} dx < \infty.$$

This completes the proof and we note that the same argument applies to any strictly stable process of index  $\alpha \in (0, 2)$ . □

Before passing to the proof of Theorem 1.2, we give an auxiliary lemma.

**Lemma 3.3.** *If  $\mathbb{E}|\xi| < \infty$ , then*

$$\mathbb{E} \left| \left( \sqrt{i^2 + S_i^2} - i \right) - \left( \sqrt{j^2 + S_j^2} - j \right) \right| \leq (2 + \mathbb{E}|\xi|) |i - j|$$

*holds for any  $i, j \in \mathbb{N}$ .*

*Proof.* It is obviously enough to show that

$$\mathbb{E} \left| \sqrt{i^2 + S_i^2} - \sqrt{j^2 + S_j^2} \right| \leq (1 + \mathbb{E}|\xi|) |i - j|,$$

which follows from

$$\begin{aligned} \mathbb{E} \left| \sqrt{i^2 + S_i^2} - \sqrt{j^2 + S_j^2} \right| &= \mathbb{E} \left| \frac{i^2 + S_i^2 - j^2 - S_j^2}{\sqrt{i^2 + S_i^2} + \sqrt{j^2 + S_j^2}} \right| \\ &\leq \mathbb{E} \frac{|i^2 - j^2|}{\sqrt{i^2 + S_i^2} + \sqrt{j^2 + S_j^2}} + \mathbb{E} \frac{|S_i^2 - S_j^2|}{\sqrt{i^2 + S_i^2} + \sqrt{j^2 + S_j^2}} \\ &\leq \frac{|i^2 - j^2|}{\sqrt{i^2} + \sqrt{j^2}} + \mathbb{E} \frac{|S_i^2 - S_j^2|}{\sqrt{S_i^2} + \sqrt{S_j^2}} \\ &= |i - j| + \mathbb{E} ||S_i| - |S_j|| \leq |i - j| + \mathbb{E} |S_i - S_j| \\ &= |i - j| + \mathbb{E} |S_{i-j}| \leq (1 + \mathbb{E}|\xi|) |i - j| \end{aligned}$$

for any  $i, j \in \mathbb{N}$ . □



*Proof of Theorem 1.2.* Fixing a coupling between  $(Z_{n,k})$  and  $(Z_k)$  such that (2.7) holds, we show first that

$$\frac{1}{b_n} \left| \sum_{k=1}^n \left( \sqrt{Z_{n,k}^2 + S_{k,Z_{n,k}}^2} - Z_{n,k} \right) - \sum_{k=1}^n \left( \sqrt{[nZ_k]^2 + S_{k,[nZ_k]}^2} - [nZ_k] \right) \right| \xrightarrow[n \rightarrow \infty]{P} 0. \quad (3.29)$$

Recall that  $b_n = a_n^2/n$  with  $a_n$  given by (1.2). By Lemma 3.3 and use of the triangle inequality, the expectation of the difference in (3.29) is bounded by a constant times

$$\frac{n}{a_n^2} \sum_{k=1}^n \mathbb{E} |Z_{n,k} - [nZ_k]| \leq \frac{n^2}{a_n^2} \sum_{k=1}^{\infty} \mathbb{E} \left| \frac{Z_{n,k}}{n} - Z_k \right| + \frac{n}{a_n^2} \mathbb{E} \sum_{k=1}^{\infty} \{nZ_k\},$$

where  $\{x\}$  denotes the fractional part of  $x \in \mathbb{R}$ . Then use (2.7) to assess that the first summand on the right is of the same order as  $a_n^{-2}n \log n/4$  and thus tending to zero as  $n \rightarrow \infty$  because  $(a_n^2)_{n \in \mathbb{N}}$  is regularly varying with index  $2/\alpha > 1$ . The second summand tends to zero by the same reasoning and the fact that

$$\mathbb{E} \sum_{k=1}^{\infty} \{nZ_k\} \simeq \frac{1}{2} \log n \quad \text{as } n \rightarrow \infty,$$

which can be justified as follows: by formula (2.1) in [4] with  $\phi(x) = \{nx\}$ , we have

$$\begin{aligned} \mathbb{E} \sum_{k=1}^{\infty} \{nZ_k\} &= \int_0^1 \frac{\{nx\} dx}{x} = \int_0^n \frac{\{y\} dy}{y} = \sum_{k=0}^{n-1} \int_k^{k+1} \frac{y-k}{y} dy \\ &= 1 + \sum_{k=1}^{n-1} \left( 1 - k \log \left( \frac{k+1}{k} \right) \right) \simeq 1 + \frac{1}{2} \sum_{k=1}^{n-1} \frac{1}{k} \simeq \frac{1}{2} \log n \end{aligned}$$

as  $n \rightarrow \infty$ . Now (3.29) follows by these estimates and Markov's inequality.

In view of (3.29) and representation (2.1), it remains to prove

$$\frac{\sum_{k=1}^n \left( \sqrt{[nZ_k]^2 + S_{k,[nZ_k]}^2} - [nZ_k] \right)}{b_n} \xrightarrow[n \rightarrow \infty]{d} \sum_{k=1}^{\infty} \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{2Z_k}. \quad (3.30)$$

To this end, we use Theorem 3.2 in [5] for which the following two assertions must be verified: First, for any fixed  $m \in \mathbb{N}$ ,

$$\frac{\sum_{k=1}^m \left( \sqrt{[nZ_k]^2 + S_{k,[nZ_k]}^2} - [nZ_k] \right)}{b_n} \xrightarrow[n \rightarrow \infty]{d} \sum_{k=1}^m \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{2Z_k}, \quad (3.31)$$

and, second, for any  $\varepsilon > 0$ ,

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left\{ \left| \sum_{k=m}^n \left( \sqrt{[nZ_k]^2 + S_{k,[nZ_k]}^2} - [nZ_k] \right) \right| > \varepsilon b_n \right\} = 0. \quad (3.32)$$

Note that (3.31) is obviously equivalent to

$$\sum_{k=1}^m \frac{S_{k,[nZ_k]}^2}{a_n^2 \left( \sqrt{[nZ_k]^2 + S_{k,[nZ_k]}^2} + [nZ_k] \right) / n} \xrightarrow[n \rightarrow \infty]{d} \sum_{k=1}^m \frac{(\mathcal{S}_\alpha^{(k)}(Z_k))^2}{2Z_k}.$$

But this is true by the continuous mapping theorem and the joint convergence

$$\left( \frac{S_{k,[nZ_k]}^2}{a_n^2}, \frac{\sqrt{[nZ_k]^2 + S_{k,[nZ_k]}^2} + [nZ_k]}{n} \right)_{1 \leq k \leq m} \xrightarrow[n \rightarrow \infty]{d} ((\mathcal{S}_\alpha^{(k)}(Z_k))^2, 2Z_k)_{1 \leq k \leq m},$$

which in turn holds because  $a_n = o(n)$  as  $n \rightarrow \infty$  and  $(S_{i,j})_{i,j \in \mathbb{N}}$  and  $(Z_k)_{k \in \mathbb{N}}$  are independent.

For (3.32), we argue as follows. Using formula (3.2), it is enough to check that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left\{ \sum_{k=m}^n \frac{S_{k, \lfloor nZ_k \rfloor}^2}{2 \lfloor nZ_k \rfloor} \mathbb{1}_{\{nZ_k \geq 1\}} > \varepsilon b_n \right\} = 0. \tag{3.33}$$

Fix  $\delta \in (0, \alpha/2)$ . Then, using the subadditivity of  $x \mapsto x^\delta$  and Markov's inequality, we see that (3.33) is a consequence of

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_n^\delta} \sum_{k=m}^n \mathbb{E} \left( \frac{S_{k, \lfloor nZ_k \rfloor}^2}{2 \lfloor nZ_k \rfloor} \mathbb{1}_{\{nZ_k \geq 1\}} \right)^\delta = 0. \tag{3.34}$$

By Lemma 5.2.2 in [12], there exists a constant  $C = C_{\delta, \alpha} > 0$  such that

$$\mathbb{E}|S_l|^{2\delta} \leq C a_l^{2\delta}, \quad l \in \mathbb{N}. \tag{3.35}$$

Therefore,

$$\begin{aligned} \mathbb{E} \left( \frac{S_{k, \lfloor nZ_k \rfloor}^2}{2 \lfloor nZ_k \rfloor} \mathbb{1}_{\{nZ_k \geq 1\}} \right)^\delta &= \sum_{l=1}^n \mathbb{P}\{\lfloor nZ_k \rfloor = l\} \mathbb{E} \left( \frac{S_l^2}{2l} \right)^\delta \\ &\leq C 2^{-\delta} \sum_{l=1}^n \mathbb{P}\{\lfloor nZ_k \rfloor = l\} \frac{a_l^{2\delta}}{l^\delta} = C 2^{-\delta} \mathbb{E} b_{\lfloor nZ_k \rfloor}^\delta \mathbb{1}_{\{\lfloor nZ_k \rfloor \geq 1\}}, \end{aligned}$$

and (3.34) will follow once having shown that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_n^\delta} \sum_{k=m}^n \mathbb{E} \left( b_{\lfloor nZ_k \rfloor}^\delta \mathbb{1}_{\{\lfloor nZ_k \rfloor \geq 1\}} \right) = 0. \tag{3.36}$$

Fix an arbitrary  $\delta' \in (0, (\frac{2}{\alpha} - 1)\delta)$ . Since  $(b_n^\delta)_{n \in \mathbb{N}}$  is regularly varying with index  $(\frac{2}{\alpha} - 1)\delta$ , we can apply Potter's bound to the slowly varying sequence  $(b_n^\delta/n^{(2/\alpha-1)\delta})_{n \in \mathbb{N}}$ , see [6, Thm. 1.5.6(ii)]. In combination with  $\lfloor nZ_k \rfloor \leq n$ , this gives

$$\frac{b_{\lfloor nZ_k \rfloor}^\delta}{b_n^\delta} \leq \text{const} \cdot \left( \frac{\lfloor nZ_k \rfloor}{n} \right)^{(2/\alpha-1)\delta-\delta'} \leq \text{const} \cdot Z_k^{(2/\alpha-1)\delta-\delta'}.$$

Moreover, by formula (2.1) in [4],

$$\sum_{k=1}^\infty \mathbb{E} Z_k^{(2/\alpha-1)\delta-\delta'} = \int_0^1 x^{(2/\alpha-1)\delta-\delta'-1} dx < \infty. \tag{3.37}$$

Hence (3.36) follows and the proof of (3.32) is complete. □

### 3.4 Proof of Theorem 1.3

Put

$$\tau_n := \arg \max_{0 \leq k \leq n} S_k \quad \text{and} \quad \kappa_n := \arg \min_{0 \leq k \leq n} S_k,$$

(where, for the sake of definiteness, the position of the *first* maximum or minimum is taken) and further,

$$M_n = \max_{0 \leq k \leq n} S_k \quad \text{and} \quad m_n = \min_{0 \leq k \leq n} S_k,$$

see Figure 1.

How long is the convex minorant of a one-dimensional random walk?

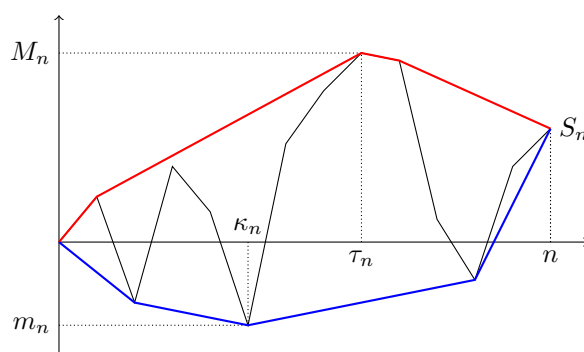


Figure 1: The proof of Theorem 1.3. The thin black piecewise linear line is a linearly interpolated random walk  $(S_k)_{k=0,\dots,n}$ , the thick red line consisting of four line segments is its concave majorant, the thick blue line is its convex minorant.

Our reasoning is geometric and based on a simple comparison argument. First of all note that the concave majorant consists of two piecewise linear subparts: one connecting the origin and the point  $(\tau_n, M_n)$ , and the other connecting  $(\tau_n, M_n)$  and  $(n, S_n)$ . Denote their lengths by  $L_{1,n}^\wedge$  and  $L_{2,n}^\wedge$ , respectively, thus  $L_n^\wedge = L_{1,n}^\wedge + L_{2,n}^\wedge$ . The triangle inequality applied to every segment of the concave majorant provides

$$M_n \leq L_{1,n}^\wedge \leq M_n + \tau_n \leq M_n + n,$$

and also

$$M_n - S_n \leq L_{2,n}^\wedge \leq M_n - S_n + n - \tau_n \leq M_n - S_n + n.$$

Consequently,

$$2M_n - S_n \leq L_n^\wedge \leq 2M_n - S_n + 2n.$$

Similarly,

$$S_n - 2m_n \leq L_n^\vee \leq S_n - 2m_n + 2n.$$

To complete the proof of Theorem 1.3, it remains to note that  $n/a_n \rightarrow 0$  because  $(a_n)_{n \in \mathbb{N}}$  is regularly varying with index  $1/\alpha > 1$  and

$$\left( \frac{S_n}{a_n}, \frac{M_n}{a_n}, \frac{m_n}{a_n} \right) \xrightarrow[n \rightarrow \infty]{d} \left( \mathcal{S}_\alpha(1), \sup_{t \in [0,1]} \mathcal{S}_\alpha(t), \inf_{t \in [0,1]} \mathcal{S}_\alpha(t) \right)$$

as an immediate consequence of (1.2) and the continuous mapping theorem.

#### 4 The case of nonzero mean

Throughout this section, we assume that  $\mu := \mathbb{E}\xi$  exists and is nonzero which rules out case (C). We are left with two possibilities:

(A')  $\mathbb{E}\xi^2 < \infty$  and  $\mu \neq 0$ ;

(B')  $\mathbb{E}\xi^2 = \infty$ ,  $\mu \neq 0$ , and the law of  $\xi$  lies in the domain of attraction of an  $\alpha$ -stable law with  $\alpha \in (1, 2]$ .

We stress that case (B') includes  $\alpha = 2$ , as opposed to case (B).

The essence of the next theorem is that in both cases (A') and (B'), the distributional behavior of  $L_n^\vee$  as  $n \rightarrow \infty$  coincides with that of  $S_n$  up to a linear centering and a scaling as in Theorem 1.7 in [17]. We will see that the same holds for  $L_n$ , the perimeter of the convex hull of  $\{(j, S_j) : j = 0, \dots, n\}$ .

**Theorem 4.1.** *In the cases (A') and (B'), we have*

$$\frac{L_n^\sim - \sqrt{1 + \mu^2} n}{a_n} \xrightarrow[n \rightarrow \infty]{d} \frac{\mu}{\sqrt{1 + \mu^2}} \mathcal{S}_\alpha(1), \tag{4.1}$$

where  $a_n := \sqrt{n}$ ,  $\sigma^2 := \mathbb{E}\xi^2 - \mu^2$ , and  $\mathcal{S}_2(1) := \mathcal{N}(0, \sigma^2)$  in case (A'), and  $(a_n)_{n \in \mathbb{N}}$  is a sequence and  $\mathcal{S}_\alpha(1)$  an  $\alpha$ -stable random variable such that  $(S_n - \mu n)/a_n \xrightarrow{d} \mathcal{S}_\alpha(1)$ .

The convergence (4.1) does also hold with  $L_n^\widehat{}$  and  $L_n/2$  in the place of  $L_n^\sim$ .

The result for  $L_n$  in case (A') is a particular case of Theorem 1.2 in [23] (combined with Theorem 1.8 in [17]), which applies because  $(n, S_n)_{n \in \mathbb{N}_0}$  can be regarded as a random walk in the plane. The approach of [17, 23] is different from the one employed here: we combine the formula  $L_n = L_n^\widehat{ } + L_n^\sim$  with the limit result (4.1) (or, to be more precise, (4.2)) for  $L_n^\sim$  and its version for  $L_n^\widehat{}$ , both obtained from representation (2.2). Recall that this argument does not work when  $\mu = 0$ ; cf. Remark 1.4.

*Proof.* Having  $(S_n - \mu n)/a_n \xrightarrow{d} \mathcal{S}_\alpha(1)$  in both cases (A') and (B') (naturally with  $\alpha = 2$  in (A')), (4.1) follows if we can show that

$$\frac{1}{a_n} \left| L_n^\sim - \sqrt{1 + \mu^2} n - \frac{\mu}{\sqrt{1 + \mu^2}} (S_n - \mu n) \right| \xrightarrow[n \rightarrow \infty]{P} 0, \tag{4.2}$$

which in case (A') is just a particular case of Theorem 1.7 in [17].

We use an extended version of representation (2.2), viz.

$$(L_n^\sim - n, S_n) \stackrel{d}{=} \left( \sum_{j=1}^n \sum_{i=1}^{K_{n,j}} \sqrt{j^2 + S_{i,j}^2} - j, S_n \right). \tag{4.3}$$

Defining

$$\tilde{S}_{i,j} := S_{i,j} - \mu j,$$

for  $i, j \in \mathbb{N}$ , we have

$$\sum_{j=1}^n \sum_{i=1}^{K_{n,j}} \tilde{S}_{i,j} = S_n - \mu n,$$

and

$$\begin{aligned} \Delta_n &:= L_n^\sim - \sqrt{1 + \mu^2} n - \frac{\mu}{\sqrt{1 + \mu^2}} (S_n - \mu n) \\ &\stackrel{d}{=} \sum_{j=1}^n \sum_{i=1}^{K_{n,j}} \left( \sqrt{j^2 + (\mu j + \tilde{S}_{i,j})^2} - \sqrt{1 + \mu^2} j - \frac{\mu}{\sqrt{1 + \mu^2}} \tilde{S}_{i,j} \right) \\ &= \sqrt{1 + \mu^2} \sum_{j=1}^n \sum_{i=1}^{K_{n,j}} j \left( \sqrt{1 + \frac{\tilde{S}_{i,j}^2 + 2\mu j \tilde{S}_{i,j}}{(1 + \mu^2)j^2}} - 1 - \frac{\mu}{1 + \mu^2} \frac{\tilde{S}_{i,j}}{j} \right). \end{aligned}$$

Further, we will use the inequalities

$$1 + \frac{y}{2} \leq \sqrt{1 + x + y} \quad \text{for } x \geq y^2/4, x + y \geq -1, \tag{4.4}$$

and

$$\sqrt{1 + z} \leq 1 + \frac{z}{2} \quad \text{for } z \geq -1. \tag{4.5}$$

## How long is the convex minorant of a one-dimensional random walk?

The first one with  $x := \tilde{S}_{i,j}^2 / ((1 + \mu^2)j^2)$  and  $y := 2\mu j \tilde{S}_{i,j} / ((1 + \mu^2)j^2)$  yields  $\Delta_n \geq 0$  a.s. Then the second one with  $z := \tilde{S}_{i,j}^2 + 2\mu j \tilde{S}_{i,j} / ((1 + \mu^2)j^2)$  implies

$$\mathbb{P}\{|\Delta_n| > \varepsilon a_n\} \leq \mathbb{P}\left\{\sum_{j=1}^n \sum_{i=1}^{K_{n,j}} \frac{\tilde{S}_{i,j}^2}{2\sqrt{1 + \mu^2}j} > \varepsilon a_n\right\}$$

for every  $\varepsilon > 0$ .

Fix  $\delta \in (0, \alpha/2)$ . Then, by subadditivity of  $x \mapsto x^\delta$  and Markov's inequality, it suffices to check that

$$\lim_{n \rightarrow \infty} \frac{1}{a_n^\delta} \mathbb{E}\left(\sum_{j=1}^n \sum_{i=1}^{K_{n,j}} \frac{\tilde{S}_{i,j}^{2\delta}}{j^\delta}\right) = 0. \quad (4.6)$$

We use (3.35), the independence of  $\tilde{S}_{i,j}$  and  $K_{n,j}$  and the formula  $\mathbb{E}K_{n,j} = j^{-1}$  stated in (3.18) to infer

$$\mathbb{E}\left(\sum_{j=1}^n \sum_{i=1}^{K_{n,j}} \frac{\tilde{S}_{i,j}^{2\delta}}{j^\delta}\right) = \sum_{j=1}^n \frac{1}{j} \frac{\mathbb{E}S_j^{2\delta}}{j^\delta} \leq \text{const} \cdot \sum_{j=1}^n \frac{a_j^{2\delta}}{j^{1+\delta}},$$

where the last inequality is a consequence of (3.35). Since  $(a_j)_{j \in \mathbb{N}}$  is regularly varying with index  $1/\alpha$ , the sequence  $(j^{-1-\delta} a_j^{2\delta})_{j \in \mathbb{N}}$  is regularly varying with index  $2\delta/\alpha - 1 - \delta \geq -1$  and the sum is regularly varying with index  $2\delta/\alpha - \delta$ . And since the latter is smaller than  $\delta/\alpha$ , the index of regular variation of  $(a_n^\delta)_{n \in \mathbb{N}}$  for  $\alpha > 1$ , we conclude (4.6) thus completing the proof of (4.1).

Since the representation (4.3) is also valid for  $L_n^\wedge$  instead of  $L_n^\smile$ , the convergence (4.1) and the coupling (4.2) (between  $L_n^\smile$  and a normalization of  $S_n$ ) remain true for  $L_n^\wedge$  in the place of  $L_n^\smile$ . By combining (4.2) for both  $L_n^\smile$ ,  $L_n^\wedge$  finally provides (4.1) for  $L_n/2 = (L_n^\smile + L_n^\wedge)/2$ .  $\square$

## References

- [1] Abramson, J. and Pitman, J. (2011). Concave majorants of random walks and related Poisson processes. *Combinatorics, Probability and Computing* **20(5)**, pp. 651–682. MR-2825583
- [2] Abramson, J., Pitman, J., Ross, N. and Uribe Bravo, G. (2011). Convex minorants of random walks and Lévy processes. *Electron. Comm. Probab.* **16**, pp. 423–434. MR-2831081
- [3] Arratia, R., Barbour, A. and Tavaré, S. (2003). *Logarithmic Combinatorial Structures: A Probabilistic Approach*, European Mathematical Society. MR-2032426
- [4] Arratia, R., Barbour, A. and Tavaré, S. (2006). A tale of three couplings: Poisson–Dirichlet and GEM approximations for random permutations. *Combinatorics, Probability and Computing* **15(1-2)**, pp. 31–62. MR-2195574
- [5] Billingsley, P. (1999). *Convergence of Probability Measures*, Second Ed., John Wiley & Sons. MR-1700749
- [6] Bingham, N., Goldie, C. and Teugels, J. (1987). *Regular Variation*, Cambridge University Press. MR-0898871
- [7] Cai, G.-H. and Zhu, M.-H. (2006). A two-sided estimate in the Hsu–Robbins–Erdős law of large numbers for i.i.d. random variables sequence. *J. Appl. Math. Comput.* **22(1-2)**, pp. 331–338. MR-2248462
- [8] Davis, J. A. (1968). Convergence rates for the law of the iterated logarithm. *Ann. Math. Stat.* **39(5)**, pp. 1479–1485. MR-0253411
- [9] Erdős, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20(2)**, pp. 286–291. MR-0030714

- [10] Friedman, N., Katz, M. and Koopman, L. H. (1966). Convergence rates for the central limit theorem. *Proc. Nat. Acad. Sci. USA* **56(4)**, pp. 1062–1065. MR-0199878
- [11] Gut, A. (2009). *Stopped Random Walks*, Second Ed., Springer. MR-2489436
- [12] Ibragimov, I. and Linnik, Yu. (1971). *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff Publishing, Groningen. MR-0322926
- [13] Iksanov, A., Marynych, A. and Möhle, M. (2009). On the number of collisions in  $\beta(2, b)$ -coalescents. *Bernoulli* **15(3)**, pp. 829–845. MR-2555201
- [14] Kingman, J. F. C. (1977). The population structure associated with the Ewens sampling formula. *Theoretical Population Biology* **11(2)**, pp. 274–283. MR-0682238
- [15] Lo, C. H., McRedmond, J. and Wallace, C. (2018). Functional limit theorems for random walks. arXiv:1810.06275.
- [16] McRedmond, J. (2019). Convex hulls of random walks. Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/13281/>
- [17] McRedmond, J. and Wade, A. R. (2018). The convex hull of a planar random walk: perimeter, diameter, and shape. *Electron. J. Probab.* **23**, paper no. 131. MR-3896868
- [18] Petrov, V. (1975). *Sums of Independent Random Variables*, Springer. MR-0388499
- [19] Pruss, A. R. (1997). A two-sided estimate in the Hsu–Robbins–Erdős law of large numbers. *Stochastic Process. Appl.* **70(2)**, pp. 173–180. MR-1475661
- [20] Sparre Andersen, E. (1954). On the fluctuations of sums of random variables II. *Math. Scand.* **2(2)**, pp. 195–223. MR-0068154
- [21] Vershik, A. M. and Schmidt, A. A. (1977). Limit measures arising in the asymptotic theory of symmetric groups. I. *Theory of Probability and its Applications* **22(1)**, pp. 70–85. MR-0448476
- [22] Wade, A. R. and Xu, C. (2015). Convex hulls of random walks and their scaling limits. *Stochastic Process. Appl.* **125(11)**, pp. 4300–4320. MR-3385604
- [23] Wade, A. R. and Xu, C. (2015). Convex hulls of planar random walks with drift. *Proc. Amer. Math. Soc.* **143(1)**, pp. 433–445. MR-3272767

**Acknowledgments.** G. Alsmeyer and Z. Kabluchko were partially funded by the Deutsche Forschungsgemeinschaft (DFG) under Germany’s Excellence Strategy EXC 2044–390685587, Mathematics Münster: Dynamics–Geometry–Structure. V. Vysotsky was supported in part by the RFBR Grant 19-01-00356.