

Extreme order statistics of random walks

Jim Pitman^{1,a} and Wenpin Tang^{2,b}

¹*Department of Statistics, University of California, Berkeley, Berkeley, USA, apitman@stat.berkeley.edu*

²*Department of Industrial Engineering and Operations Research, Columbia University, New York, NY, USA, wt2319@columbia.edu*

Received 28 July 2020; revised 4 February 2022; accepted 4 February 2022

Abstract. This paper is concerned with the limit theory of the extreme order statistics derived from random walks. We establish the joint convergence of the order statistics near the minimum of a random walk in terms of the Feller chains. Detailed descriptions of the limit process are given in the case of simple symmetric walks and Gaussian walks.

Résumé. Cet article traite de la théorie limite des statistiques d'ordre extrêmes provenant des marches aléatoires. Nous établissons la convergence conjointe des statistiques d'ordre près du minimum d'une marche aléatoire en termes des chaînes de Feller. Des descriptions détaillées du processus limite sont données dans le cas de marches simples symétriques et des marches gaussiennes.

MSC2020 subject classifications: 60G50; 60F17; 60J05

Keywords: Bessel processes; Brownian embedding; Fluctuation theory; Limit theorems; Order statistics; Path decomposition; Random walk

1. Introduction

Extreme value theory and order statistics have received considerable interest in probability theory and statistics. They have a variety of applications including nonparametric tests [37], kernel density estimation [54], physics [39,59], environmental sciences [33], financial risk analysis [22], and reliability engineering [18]. The most studied case is order statistics derived from a sequence of independent and identically distributed (i.i.d.) random variables [3,18,48]. We refer to the books [3, 18,48] for examples and related results.

In this paper, following earlier work of Pollaczek [45], Wendel [62], Sparre Andersen [2] and others, we study the order statistics derived from a random walk $S = (S_k, 0 \leq k \leq n)$, with $S_0 = 0$, and $S_k = \sum_{i=1}^k X_i$ with increments X_1, X_2, \dots, X_n that are either *exchangeable*, or *independent and identically distributed (i.i.d.)*. For $0 \leq k \leq n$, let $M_{k,n} = M_k(S_0, \dots, S_n)$ be the k th order statistics derived from the steps (S_0, \dots, S_n) of the walk S . So

$$(1.1) \quad \{0 = S_0, S_1, \dots, S_n\} = \{M_{k,n}, 0 \leq k \leq n\} \quad \text{with } M_{0,n} \leq M_{1,n} \leq \dots \leq M_{n,n}.$$

We also use the notations

$$(1.2) \quad \underline{S}_n := M_{0,n} := \min_{0 \leq k \leq n} S_k \quad \text{and} \quad \bar{S}_n := M_{n,n} := \max_{0 \leq k \leq n} S_k.$$

The fluctuation theory of random walks, developed in the 1950s and 1960s by Sparre Andersen, Spitzer, Baxter, Feller and Wendel, describes the distributions of these order statistics, especially the minimum, the maximum, and related random variables such as the random times when these values are attained, either for the first or last time. Development of this theory was motivated by diverse applications of random walks, first in queueing theory [4,45] and mathematical statistics [6,26], and more recently in finance [16,32]. The systematic study of large n limit distributions of these functionals of random walks led via Donsker's theorem and its generalizations to the fluctuation theory of Brownian motion and Lévy processes, as reviewed in the recent monographs [8,36].

Beside this mature fluctuation theory of random walks and Lévy processes, there has been a recent resurgence of interest in the order statistics of random walks in the physics community. The papers [52,53] are typical of this literature, starting from particular models of random walk such as symmetric walks with continuous increment distributions, and

making sustained calculations in these models. Many of these calculations recover results that are known in the fluctuation theory of random walks, even if by now somewhat buried in the literature. But some of these calculations have led to new limit distributions and asymptotic formulas, whose place relative to the standard fluctuation theory does not seem obvious. As a case in point, Schehr and Majumdar [52] focused attention on the spacings or gaps between the random walk order statistics,

$$(1.3) \quad D_{k,n} := M_{k,n} - M_{k-1,n}, \quad 1 \leq k \leq n.$$

It was observed by Schehr and Majumdar [52, (9)] that if the distribution of X_1 has a symmetric density with finite variance, then as $n \rightarrow \infty$ the expected spacing $\mathbb{E}D_{k,n}$ has a limit, for which they gave an integral expression involving the Fourier transform of the density of X_1 . Some known results, reviewed in Section 3.2, provide explicit descriptions of both the exact distribution of $M_{k,n}$, and its asymptotic distributions in various limit regimes as $n \rightarrow \infty$. But much less is easily found in the literature regarding the large n limit behavior of the differences $D_{k,n}$, which involves the joint distributions of variables in the sequence $(M_{0,n}, M_{1,n}, \dots)$.

Our first result provides a complete description of the joint limit behavior of the order statistics near the minimum of a random walk, relative to the location of the minimum, for any distribution of the increments of the random walk. The construction relies on the Feller chains S^\uparrow and S^\downarrow , introduced by Feller [24]. The *upward Feller chain* S^\uparrow with $S_0^\uparrow = 0$ is the sequence of partial sums of those increments X_k of the random walk S with $S_k > 0$, and the *downward Feller chain* S^\downarrow with $S_0^\downarrow = 0$ is the sequence of partial sums of those increments X_k of the random walk S with $S_k \leq 0$.

Theorem 1.1. *Let $(S_k, k \geq 0)$ be a real-valued walk with exchangeable increments X_1, X_2, \dots , whose common distribution satisfies $\mathbb{P}(X_1 = 0) < 1$. For $0 \leq k \leq n$, let $M_{k,n}$ be the sequence of order statistics defined by (1.1) of the n -step walk $(S_k, 0 \leq k \leq n)$, and $W_{k,n} := M_{k,n} - M_{0,n}$. Let $0 = W_0 \leq W_1 \leq \dots$ be the order statistics*

$$(1.4) \quad W_k := M_k(\{-S_n^\downarrow, n \geq 0\} \cup \{S_n^\uparrow, n \geq 1\}),$$

derived from the two Feller chains $(S_n^\downarrow, n \geq 0)$ and $(S_n^\uparrow, n \geq 1)$ generated by the walk S . Then we have the following:

(i) *For each finite K , there is the convergence in total variation of finite-dimensional distributions of order statistics*

$$(1.5) \quad (W_{k,n}, 1 \leq k \leq K) \xrightarrow{\text{TV}} (W_k, 1 \leq k \leq K) \quad \text{as } n \rightarrow \infty.$$

(ii) *For each fixed $w > 0$, there is the convergence in total variation of laws of counting processes*

$$(1.6) \quad \left(\sum_{k=1}^{\infty} 1(W_{k,n} \leq v), 0 \leq v \leq w \right) \xrightarrow{\text{TV}} \left(\sum_{k=1}^{\infty} 1(W_k \leq v), 0 \leq v \leq w \right) \quad \text{as } n \rightarrow \infty.$$

As will be seen in the proof of Theorem 1.1, by letting $\alpha_n := \max\{k \leq n : S_k = M_{0,n}\}$, the values W_k derived from $(-S_n^\downarrow, n \geq 0)$ represent contributions to the limiting process from values of S_j with j at or before time α_n , while the values W_k derived from $(S_n^\uparrow, n \geq 1)$ are contributions from strictly after time α_n . See also [61, Theorem 4] for a related result.

Recall from (1.3) that $D_{k,n} := M_{k,n} - M_{k-1,n} = W_{k,n} - W_{k-1,n}$. Observe the reversibility of spacings

$$(1.7) \quad (D_{n,n}, \dots, D_{1,n}) \stackrel{d}{=} (D_{1,n}, \dots, D_{n,n}),$$

which follows from the well known duality relation $(S_n - S_k, 0 \leq k \leq n) \stackrel{d}{=} (S_k, 0 \leq k \leq n)$, implied by the reversibility of increments $(X_n, \dots, X_1) \stackrel{d}{=} (X_1, \dots, X_n)$. Combining Theorem 1.1 with some known results (see Lemma 3.1) regarding the distributions of $M_{k,n}$ yields the following proposition.

Proposition 1.2. *With notations as in Theorem 1.1, and $X := X_1$ the generic increment of the walk with exchangeable increments, we have:*

(i) *For each finite K , there is the convergence in total variation of finite-dimensional distributions of gaps of order statistics*

$$(1.8) \quad (D_{k,n}, 1 \leq k \leq K) \xrightarrow{\text{TV}} (D_k, 1 \leq k \leq K) \quad \text{as } n \rightarrow \infty,$$

where $D_k := W_k - W_{k-1}$. Moreover, these convergences hold together with convergence of p^{th} moments for every $0 < p < \infty$ such that $\mathbb{E}|X|^p < \infty$.

(ii) If $\mathbb{E}|X| < \infty$, then for each fixed $k = 1, 2, \dots$

$$(1.9) \quad \lim_{n \rightarrow \infty} \mathbb{E}D_{k,n} = \mathbb{E}D_k = \frac{\mathbb{E}S_k^+}{k} + \mathbb{E}[\mathbb{E}(X|\mathcal{E})^-],$$

where \mathcal{E} is the exchangeable σ -field, and $x^+ := \max(x, 0)$ and $x^- := -\min(x, 0)$ for all real x . We have $\lim_{k \rightarrow \infty} \mathbb{E}D_k = \mathbb{E}|\mathbb{E}(X|\mathcal{E})|$. Assume further that $\mathbb{E}X^2 < \infty$ and $\mathbb{E}(X|\mathcal{E}) = 0$ almost surely. Then

$$(1.10) \quad \mathbb{E}D_k \sim \frac{\mathbb{E}\sqrt{\mathbb{E}(X^2|\mathcal{E})}}{\sqrt{2\pi}} k^{-1/2} \quad \text{and} \quad \mathbb{E}W_k \sim \mathbb{E}\sqrt{\mathbb{E}(X^2|\mathcal{E})} \sqrt{\frac{2k}{\pi}} \quad \text{as } k \rightarrow \infty.$$

The distribution of $(W_k, k \geq 0)$ is further simplified in the case of a random walk, due to the following description of the Feller chains in that case, by Bertoin [7].

Theorem 1.3 ([7]). *Let $(S_k, k \geq 0)$ be a random walk with i.i.d. increments X_1, X_2, \dots , whose common distribution satisfies $\mathbb{P}(X_1 = 0) < 1$. Then the two Feller chains S^\uparrow and S^\downarrow are independent, and both are Markov chains, with stationary transition probabilities given by*

$$(1.11) \quad p^\uparrow(x, dy) := \mathbb{P}(x + S_1^\uparrow \in dy) = 1(y > 0) \frac{h^\uparrow(y)}{h^\uparrow(x)} \mathbb{P}(x + X_1 \in dy), \quad x \geq 0,$$

$$(1.12) \quad p^\downarrow(x, dy) := \mathbb{P}(x + S_1^\downarrow \in dy) = 1(y \leq 0) \frac{h^\downarrow(y)}{h^\downarrow(x)} \mathbb{P}(x + X_1 \in dy), \quad x \leq 0,$$

where $h^\uparrow(x) := \mathbb{E}(\sum_{k=0}^{\tau^+-1} 1(S_k > -x))$ with $\tau^+ := \inf\{n > 0 : S_n > 0\}$, and $h^\downarrow(x) := \mathbb{E}(\sum_{k=0}^{\tau^--1} 1(S_k \leq -x))$ with $\tau^- := \inf\{n > 0 : S_n \leq 0\}$.

The following corollary is an immediate consequence of Theorem 1.1 and Tanaka's construction of Feller's chains of a random walk (see Proposition 2.3).

Corollary 1.4. *For a random walk with i.i.d. increments whose distribution is symmetric and continuous, let $\tau^+ := \inf\{n > 0 : S_n > 0\}$ and $\tau^- := \inf\{n > 0 : S_n \leq 0\}$, with the conventions that $S_{\tau^+} = \infty$ if $\tau^+ = \infty$ and $-S_{\tau^-} = \infty$ if $\tau^- = \infty$. The common distribution of the lowest gap $D_{1,n} := M_{1,n} - M_{0,n}$ and the topmost gap $D_{n,n} := M_{n,n} - M_{n-1,n}$ converges in total variation norm as $n \rightarrow \infty$ to the distribution of $D_1 = W_1$ defined by*

$$(1.13) \quad \mathbb{P}(W_1 > w) = \mathbb{P}(S_{\tau^+} > w) \mathbb{P}(-S_{\tau^-} > w), \quad w > 0.$$

In the case of a random walk, Proposition 1.2 specializes to the following corollary.

Corollary 1.5. *Let $(S_k, k \geq 0)$ be a random walk with i.i.d. increments X_1, X_2, \dots , whose common distribution satisfies $\mathbb{P}(X_1 = 0) < 1$. Let $X := X_1$ be the generic increment of the random walk. We have:*

(i) Proposition 1.2(i) holds.

(ii) If $\mathbb{E}|X| < \infty$ and $\mathbb{E}X = \mu$, then for each fixed $k = 1, 2, \dots$

$$(1.14) \quad \lim_{n \rightarrow \infty} \mathbb{E}D_{k,n} = \mathbb{E}D_k = \frac{\mathbb{E}S_k^+}{k} + \mu^-.$$

Assume further that $\mu = 0$ and $\mathbb{E}X^2 = \sigma^2 < \infty$. Then

$$(1.15) \quad \mathbb{E}D_k \sim \frac{\sigma}{\sqrt{2\pi}} k^{-1/2} \quad \text{and} \quad \mathbb{E}W_k \sim \sigma \sqrt{\frac{2}{\pi}} k^{1/2} \quad \text{as } k \rightarrow \infty.$$

As mentioned earlier, Schehr and Majumdar [52] showed the existence of the limit (1.14) for symmetric distributions of X with a density, with a more complicated description of the limit by a Fourier integral. Their result reduces to the

simpler formula in (1.14) by application of the formula [11,28,60]

$$\mathbb{E}|X| = \frac{2}{\pi} \int_0^\infty (1 - \mathbb{E}e^{itX}) \frac{dt}{t^2}$$

for any symmetric distribution of X .

Another interesting question is how fast $(W_{k,n}, 1 \leq k \leq K)$ and $(D_{k,n}, 1 \leq k \leq K)$ converges as $n \rightarrow \infty$ to the limiting processes $(W_k, 1 \leq k \leq K)$ and $(D_k, 1 \leq k \leq n)$ respectively. The following proposition quantifies a convergence rate in (1.5) and (1.8) for a random walk with i.i.d. increments which is continuous, and satisfies some moment condition. We leave the question of whether the established rate is optimal as open problem.

Proposition 1.6. *Let $(S_k, k \geq 0)$ be a random walk with i.i.d. increments X_1, X_2, \dots , whose distribution is symmetric and continuous, with $\mathbb{E}e^{\theta|X_1|} < \infty$ for some $\theta > 0$. With notations as in Theorem 1.1 and Proposition 1.2, let $\text{TV}(K, n)$ be the total variation distance between the distributions of the two sequences in (1.5) or (1.8). Then for each finite K , there is a constant $C_K > 0$ such that*

$$(1.16) \quad \text{TV}(K, n) \leq C_K n^{-\frac{1}{4}}.$$

It is well known that in the setting of Corollary 1.4, the distributions F_+ and F_- of the strict ascending and weak descending ladder heights $S_{\tau+}$ and $S_{\tau-}$ with supports in $(0, \infty)$ and $(-\infty, 0]$ respectively, are uniquely determined by the Wiener–Hopf equation $F_+ + F_- - F_+ * F_- = F$, where F is the distribution of increments. The joint limit law of (W_1, W_2) can also be described just as explicitly for a general increment distribution, and in principle so can that of (W_1, W_2, W_3) and so on, but these general descriptions become more complex as the number of variables increases. Three special cases for which much more detailed descriptions are possible are:

- (i) simple random walk, with increments $X_i = \pm 1$ and $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$;
- (ii) symmetric Gaussian random walk, whose increments have the common distribution $\mathbb{P}(X_i \in dx) = (2\pi\sigma^2)^{-\frac{1}{2}} \times \exp(-\frac{x^2}{2\sigma^2}) dx$ for some $\sigma^2 > 0$;
- (iii) symmetric Laplacian random walk, whose increments have the common distribution $\mathbb{P}(X_i \in dx) = (2b)^{-1} \times e^{-b|x|} dx$ for some $b > 0$.

In this paper we will illustrate results for walks with simple ± 1 and symmetric Gaussian increments. The case of symmetric Laplacian increments involves further symmetries and will be treated in a separate paper [44]. The key to the analysis of these random walks is to embed them into a Brownian motion. By the Skorokhod embedding [55], every random walk with mean zero and finite variance σ^2 per step may be embedded in a Brownian motion $(B(t), t \geq 0)$ as $S_k = B(T_k)$ where $0 = T_0 \leq T_1 \leq \dots$ is an increasing sequence of stopping times of B , with the $(T_k - T_{k-1}, k \geq 1)$ i.i.d. as T_1 , and $\mathbb{E}T_k = k\sigma^2$. In the above three examples,

- (i) Simple symmetric walk, with $T_k = \inf\{t > T_{k-1} : |B_t - B_{T_{k-1}}| = 1\}$;
- (ii) symmetric Gaussian random walk, with $T_k = \sigma k$;
- (iii) symmetric Laplacian random walk, with $T_k = 2b^{-2}\gamma_k$, where $\gamma_k = \sum_{i=1}^k \varepsilon_i$ for a sequence of independent standard exponential variables.

The upward Feller chain $(S_k^\uparrow, k \geq 0)$ derived from a simple symmetric walk was studied by Pitman [43] who gave a number of constructions of this Markov chain, including the embedding $S_k^\uparrow = R_3(T_k)$, $k \geq 0$ where R_3 is a BES(3) process, and T_k is the succession of random times at which R_3 hits integer values, defined inductively by $T_0 := 0$ and $T_{k+1} := \inf\{t > T_k : |R_3(t) - R_3(T_k)| = 1\}$. It follows from this embedding, and the easily established almost sure convergence $T_n/n \rightarrow 1$ as $n \rightarrow \infty$, that there is the scaling limit

$$(1.17) \quad (n^{-1/2}S_{un}^\uparrow, u \geq 0) \xrightarrow{d} (R_3(u), u \geq 0)$$

in the sense of functional limit theorems in the path space $C[0, \infty)$, that is the same sense in which $(n^{-1/2}S_{un}, u \geq 0)$ converges in distribution to Brownian motion. See [27] for embedding the Feller chain $(S_k^\uparrow, k \geq 0)$ derived from a random walk with a more general increment distribution in a BES(3) process.

Organization of the paper: In Section 2, we recall basic constructions of the Feller chains. In Section 3, we study the extreme order statistics derived from general random walks. There Theorem 1.1, Proposition 1.2, Corollary 1.4 and Proposition 1.6 are proved. Section 4 provides detailed descriptions of the limit process of order statistics derived from simple symmetric random walks. Section 5 describes the limit process of order statistics derived from Gaussian random walks.

2. Feller chains – old & new

In this section we present some basic constructions in the theory of conditioned random walks. This theory was systematically developed in the 1980's and 1990's by Tanaka [58], Bertoin and Doney [7,9] following earlier contributions by Feller [24], Williams [63,64] and Pitman [43]. Conditioned random walks also appear in the study of localization of random polymers, see e.g. Comets [15, Chapter 7].

Let $S_0 = 0$ and $S_n := X_1 + \dots + X_n$ be a real-valued walk with exchangeable increments X_1, X_2, \dots . Let $(S_k^\uparrow, k \geq 0)$ and $(S_k^\downarrow, k \geq 0)$ be the upward and downward Feller chains derived from the walk $(S_k, k \geq 0)$ defined in the Introduction. These random sequences with a finite time horizon n were introduced by Feller [24, XII.8, Lemma 3] to provide a combinatorial proof of Sparre Andersen's equivalence principle [1]. Formally, the upward Feller chain S^\uparrow is the sequence of partial sums of those increments X_k of the walk S with $S_k > 0$, and the downward Feller chain S^\downarrow is the sequence of partial sums of those increments X_k of the walk S with $S_k \leq 0$. Let $N_n^+ := \#\{k \leq n : S_k > 0\}$ and $N_n^- := n - N_n^+$. The above construction gives partial sum processes $(S_k^\uparrow, 0 \leq k \leq N_n^+)$ and $(S_k^\downarrow, 0 \leq k \leq N_n^-)$ of random lengths N_n^+ and N_n^- respectively. Then there are identities

$$(2.1) \quad N_k^+ = N_{k-1}^+ + 1(S_k > 0), \quad N_k^- = N_{k-1}^- + 1(S_k \leq 0), \quad S_k = S_{N_k^+}^\uparrow + S_{N_k^-}^\downarrow.$$

It follows that for each fixed n , the original sequence $(S_k, 0 \leq k \leq n)$ is encoded as a measurable function of the two Feller chains $(S_k^\uparrow, 1 \leq k \leq N_n^+)$ and $(S_k^\downarrow, 1 \leq k \leq N_n^-)$. The final values of the two chains give the value of $S_n = S_{N_n^+}^\uparrow + S_{N_n^-}^\downarrow$. This also yields the value of $1(S_n > 0)$, hence the values of N_{n-1}^+ and N_{n-1}^- by (2.1) for $k = n$, then the value of S_{n-1} by (2.1) for $k = n - 1$, and so on recursively down to the value of S_1 .

Lemma 2.1 ([24]). *Assume that $(S_k, 0 \leq k \leq n)$ is a walk with exchangeable increments, and let $\alpha_n := \max\{k \leq n : S_k = \underline{S}_n\}$. Whatever the common distribution of exchangeable increments of S , there is the identity in distribution*

$$(2.2) \quad \begin{aligned} & ((S_{\alpha_n+k} - S_{\alpha_n}, 0 \leq k \leq n - \alpha_n), (S_{\alpha_n-k} - S_{\alpha_n}, 0 \leq k \leq \alpha_n)) \\ & \stackrel{d}{=} ((S_k^\uparrow, 0 \leq k \leq N_n^+), (-S_k^\downarrow, 0 \leq k \leq N_n^-)). \end{aligned}$$

Assume further that the increments X_1, X_2, \dots are independent.

(i) *The common distribution of α_n and N_n^- is given by*

$$(2.3) \quad \mathbb{P}(\alpha_n = \ell) = \mathbb{P}(N_n^- = \ell) = \mathbb{P}\left(\bigcap_{k=1}^{\ell} \{S_k \leq 0\}\right) \mathbb{P}\left(\bigcap_{k=1}^{n-\ell} \{S_k > 0\}\right).$$

(ii) *Given $\alpha_n = \ell$, the common distribution of each pair of path fragments displayed in (2.2) is that of two independent path fragments distributed as $(S_k, 0 \leq k \leq n - \ell)$ given $\bigcap_{k=1}^{n-\ell} \{S_k > 0\}$ and $(-S_k, 0 \leq k \leq \ell)$ given $\bigcap_{k=1}^{\ell} \{S_k \leq 0\}$ respectively.*

Here the main results involve Feller's chains with infinite time horizon. Observe that the first n steps of the walk $(S_k, 0 \leq k \leq n)$ define only the first N_n^+ steps of S^\uparrow and the first N_n^- steps of S^\downarrow . As $n \uparrow \infty$, it is obvious that $N_n^+ \uparrow N_\infty^+$ and $N_n^- \uparrow N_\infty^-$ for some limiting random variables N_∞^+ and N_∞^- with values in $\{0, 1, \dots, \infty\}$. When $N_\infty^+ = \infty$ there are infinitely many X_k with $S_k > 0$, and the infinite horizon upward chain S^\uparrow has this infinite list of X -values as increments. Whereas if $N_\infty^+ < \infty$, there is the convention that $S_k^\uparrow = \infty$ for $k > N_\infty^+$. So the random variable N_∞^+ is encoded in S^\uparrow as $N_\infty^+ = \inf\{n : S_{n+1}^\uparrow = \infty\}$ with the convention that $\inf \emptyset = \infty$. Similar remarks apply to the definition of S^\downarrow . The following lemma summarizes some basic facts about Feller's chains.

Lemma 2.2 ([7,24]). *Assume that $(S_k, k \geq 0)$ is a real-valued walk with exchangeable increments X_1, X_2, \dots , whose common distribution satisfies $\mathbb{P}(X_1 = 0) < 1$.*

(i) *There is the convergence in total variation as $n \rightarrow \infty$*

$$(2.4) \quad ((S_{\alpha_n+k} - S_{\alpha_n}, 0 \leq k \leq n - \alpha_n), (S_{\alpha_n-k} - S_{\alpha_n}, 0 \leq k \leq \alpha_n)) \xrightarrow{\text{TV}} (S^\uparrow, -S^\downarrow),$$

where $\alpha_n := \max\{k \leq n : S_k = \underline{S}_n\}$.

(ii) Each Feller chain is upwardly transient:

$$\lim_{n \rightarrow \infty} S_n^\uparrow = \lim_{n \rightarrow \infty} -S_n^\downarrow = \infty \quad \text{almost surely.}$$

Assume further that the increments X_1, X_2, \dots are independent. We have the following:

(iii) The two Feller chains S^\uparrow and S^\downarrow are independent.

(iv) Each of the Feller chains S^\uparrow and $-S^\downarrow$ is a Markov chain with stationary transition probabilities, with state space $[0, \infty]$ and ∞ as an absorbing state.

(v) The chain S^\uparrow is a Doob h -transform of the original random walk S stopped on first reaching $(-\infty, 0]$, for the super-harmonic function $h^\uparrow(x) := \mathbb{E}(\sum_{k=0}^{\tau^+-1} 1(S_k > -x))$ where $\tau^+ := \inf\{n > 0 : S_n > 0\}$; while the chain S^\downarrow is a Doob h -transform of the original random walk S stopped on first reaching $(0, \infty)$, for the super-harmonic function $h^\downarrow(x) := \mathbb{E}(\sum_{k=0}^{\tau^--1} 1(S_k \leq -x))$ where $\tau^- := \inf\{n > 0 : S_n \leq 0\}$.

Note that neither [7] nor [24] mentions the convergence in total variation (2.4). It is only stated in [7] that the left side of (2.4) converges in distribution to its right side. However, the argument in [7], as is clear in the proof of Theorem 1.1, implies the convergence in total variation.

An interesting issue which does not seem to have been addressed in the literature is whether the path of the original walk $(S_k, k \geq 0)$ can be recovered almost surely from the pair of infinite horizon Feller chains $(S^\uparrow, S^\downarrow)$. While this turns out to be the case, it is not a trivial consequence of the fact that the walk $\{S_k, 0 \leq k \leq n\}$ can be recovered from the finite horizon chain segments $(S_k^\uparrow, 1 \leq k \leq N_n^+)$ and $(S_k^\downarrow, 1 \leq k \leq N_n^-)$. As discussed before, the finite segment recovery involves a reverse induction using the values of N_n^+ and N_n^- to progressively determine the values of N_k^+ and N_k^- , and hence $S_k = S_{N_k^+}^\uparrow + S_{N_k^-}^\downarrow$ for smaller values of k . But given the entire paths of the two infinite horizon chains, there is no obvious way to determine for any particular n how the n steps are split into N_n^+ and N_n^- , so no obvious beginning for an inductive recovery. This difficulty is avoided by an alternative expression for the recovery of $(S_k, k \geq 0)$ from $(S_k^\uparrow, k \geq 1)$ and $(S_k^\downarrow, k \geq 1)$ via the ascending and descending chain segments.

To illustrate, we first describe how $(S_k, 0 \leq k \leq n)$ for any fixed n is recovered from $(S_k^\uparrow, 1 \leq k \leq N_n^+)$ and $(S_k^\downarrow, 1 \leq k \leq N_n^-)$ by such an encoding. In terms of the segment $(S_k^\uparrow, 0 \leq k \leq N_n^+)$, define the sequence of *future minimum times* inductively by $G_0^\uparrow = 0$ and for $j \geq 1$,

$$(2.5) \quad G_j^\uparrow := \max \left\{ k > G_{j-1}^\uparrow : S_k^\uparrow = \min_{G_{j-1}^\uparrow < i \leq N_n^+} S_i^\uparrow \right\},$$

with the convention that $\min \emptyset = 0$ and $\max \emptyset = \infty$. Let $J_n^\uparrow := \max\{j : G_j^\uparrow < \infty\}$ be the total number of future minimum times derived from the segment $(S_k^\uparrow, 0 \leq k \leq N_n^+)$. So $G_j^\uparrow < \infty$ (which is well-defined) if and only if $J_n^\uparrow \geq j$. By construction, G_1^\uparrow is the last time that the minimum of S^\uparrow over $\{1, \dots, N_n^+\}$ is attained. If $G_1^\uparrow = N_n^+$ then $J_n^\uparrow = 1$, else G_2^\uparrow is the last time after G_1^\uparrow that the minimum of S^\uparrow over $\{G_1^\uparrow + 1, \dots, N_n^+\}$ is attained. If $G_2^\uparrow = N_n^+$ then $J_n^\uparrow = 2$, else \dots , and so on. Now the idea is to decompose $(S_k^\uparrow, 0 \leq k \leq N_n^+)$ into the sequence of *ascending chain segments*

$$(2.6) \quad (S_{G_{j-1}^\uparrow+k}^\uparrow, 0 \leq k \leq G_j^\uparrow - G_{j-1}^\uparrow), \quad 1 \leq j \leq J_n^\uparrow.$$

Note that in this construction,

$$(i) \quad 0 = G_0^\uparrow < G_1^\uparrow < \dots < G_{J_n^\uparrow}^\uparrow = N_n^+;$$

$$(ii) \quad 0 = S_{G_0^\uparrow}^\uparrow < S_{G_1^\uparrow}^\uparrow < \dots < S_{G_{J_n^\uparrow}^\uparrow}^\uparrow = S_{N_n^+}^\uparrow;$$

(iii) the increments of each segment of the original walk $(S_k, 0 \leq k \leq n)$ over positive excursions, i.e. segments $(S_{U+i}, 0 \leq i \leq V)$ with $S_U \leq 0$ and $S_{U+k} > 0$ for $k \leq V$, and either $S_{U+V+1} \leq 0$ or $U + V = n$ and $S_n > 0$, are encoded in S^\uparrow as the concatenation of increments of some consecutive sequences of future minimum segments.

Now a similar construction with S^\downarrow yields a sequence of *future maximum times* $G_j^\downarrow, 0 \leq j \leq J_n^\downarrow$, and a corresponding sequence of *descending chain segments* derived from $(S_k^\downarrow, 0 \leq k \leq N_n^-)$. The increments of the original walk over negative excursions are encoded in S^\downarrow , with analogs of (i)–(iii) above, except that if $\mathbb{P}(S_k = 0) > 0$ for some k there is the

complication that instead of all strict inequalities in (ii), the first inequality might be an equality, so the analog of (ii) is slightly different:

$$(ii') \quad 0 = S_{G_0^\downarrow}^\downarrow \geq S_{G_1^\downarrow}^\downarrow > \cdots > S_{G_{J_n^\uparrow}^\downarrow}^\downarrow = S_{N_n^-}^\downarrow.$$

The issue of recovering the first n segments of the original walk $(S_k, 0 \leq k \leq n)$ from the two finite chains $(S_k^\uparrow, 1 \leq j \leq N_n^+)$ and $(S_k^\downarrow, 1 \leq j \leq N_n^-)$ is now framed as one of determining the order in which the J_n^\uparrow ascending chain segments and the J_n^\downarrow descending chain segments are riffled together as they appeared in the original path. According to the finite n recovery by reverse induction, there is one and only one possible order to do this riffing. Moreover, it can be seen directly from the ascending and descending chain segments what this order is: it is the only order of arrangement of these segments in which the final values of the segments are weakly increasing in absolute value, with the additional requirement that if two segments have the same absolute final value, the ascending segment is placed first.

The extension of the definition of ascending and descending chain segments to the infinite horizon is obvious. The idea is to recover the infinite walk $(S_k, k \geq 0)$ from $(S_k^\uparrow, 1 \leq k \leq N_\infty^+)$ and $(S_k^\downarrow, 1 \leq k \leq N_\infty^-)$ via ascending and descending chain segments. Thus, it does not require the values of N_n^\pm for any particular n . The result is stated as follows.

Proposition 2.3. *Let $(S_k, k \geq 0)$ be a real-valued walk, and $(S_k^\uparrow, 0 \leq k \leq N_\infty^+)$ and $(S_k^\downarrow, 0 \leq k \leq N_\infty^-)$ be the associated upward and downward Feller chain respectively. Define the future minimum times G^\uparrow of S^\uparrow by $G_0^\uparrow := 0$ and*

$$G_j^\uparrow := \max \left\{ k > G_{j-1}^\uparrow : S_k^\uparrow = \min_{G_{j-1}^\uparrow < i \leq N_\infty^+} S_i^\uparrow \right\},$$

with the convention that $\min \emptyset = 0$ and $\max \emptyset = \infty$, and the sequence of ascending chain segments by

$$(S_{G_{j-1}^\uparrow + k}^\uparrow, 0 \leq k \leq G_j^\uparrow - G_{j-1}^\uparrow), \quad 1 \leq j \leq J_\infty^\uparrow,$$

where $J_\infty^\uparrow := \max\{j : G_j^\uparrow < \infty\}$ is the number of future minimum times derived from $(S_k^\uparrow, 0 \leq k \leq N_\infty^+)$. Similarly, define the future maximum times G^\downarrow of S^\downarrow , and the sequence of descending chain segments. Then the increments of S are recovered from its sequence of ascending and descending chain segments according to the rule:

- (i) Place the increments from these segments in the order of least absolute final values of the segments.
- (ii) If two segments have the same absolute final value, place the ascending segment first.

Assume further that $(S_k, k \geq 0)$ is an oscillating random walk with i.i.d. increments whose distribution is symmetric and continuous. Then $N_\infty^+ = N_\infty^- = \infty$ almost surely, and there are the identities in distribution

$$(2.7) \quad ((G_j^\uparrow, S_{G_j^\uparrow}^\uparrow), j = 1, 2, \dots) \stackrel{d}{=} ((\tau_j^+, S_{\tau_j^+}), j = 1, 2, \dots),$$

$$(2.8) \quad ((G_j^\downarrow, S_{G_j^\downarrow}^\downarrow), j = 1, 2, \dots) \stackrel{d}{=} ((\tau_j^-, S_{\tau_j^-}), j = 1, 2, \dots),$$

where $((\tau_j^+, S_{\tau_j^+}), j = 1, 2, \dots)$ is the sequence of strictly ascending ladder indices and heights of S , and $((\tau_j^-, S_{\tau_j^-}), j = 1, 2, \dots)$ is the sequence of weakly descending ladder indices and heights of S . In particular, $(\tau_j^+, S_{\tau_j^+})$ is the sum of j independent copies of the first strictly ascending ladder index and height (τ^+, S_{τ^+}) with $\tau^+ := \inf\{n > 0 : S_n > 0\}$, and $(\tau_j^-, S_{\tau_j^-})$ is the sum of j independent copies of the first weakly descending ladder index and height (τ^-, S_{τ^-}) with $\tau^- := \inf\{n > 0 : S_n \leq 0\}$.

Proof. The first part of this proposition follows the finite n recovery algorithm described in the preceding paragraph. The second part is a consequence of Tanaka's construction [58] of $(-S)^\uparrow$ from the walk S , or equivalently the Feller's chains $(S^\uparrow, S^\downarrow)$. \square

3. Order statistics of general random walks

In this section we prove the main results, Theorem 1.1, Proposition 1.2, Corollary 1.4 and Proposition 1.6 regarding the order statistics of general random walks.

3.1. *Proof of Theorem 1.1*

Let $0 = \tilde{S}_{0,n}, \tilde{S}_{1,n}, \dots, \tilde{S}_{n,n}$ be the random sequence with increments obtained by concatenation of the first N_n^- increments $(X_1^\downarrow, \dots, X_{N_n^-}^\downarrow)$ of S^\downarrow in reverse order, followed by the first N_n^+ increments $(X_1^\uparrow, \dots, X_{N_n^+}^\uparrow)$ of S^\uparrow in their original order:

$$(\tilde{X}_{1,n}, \dots, \tilde{X}_{n,n}) := (X_{N_n^-}^\downarrow, \dots, X_1^\downarrow, X_1^\uparrow, \dots, X_{N_n^+}^\uparrow).$$

By Lemma 2.1, $(\tilde{S}_{k,n}, 0 \leq k \leq n) \stackrel{d}{=} (S_k, 0 \leq k \leq n)$. Regard the sequence $(W_{k,n}, 0 \leq k \leq n)$ as a measurable function, say Φ_n , of $(S_k, 0 \leq k \leq n)$. Then for each fixed n there is the identity of joint distributions

$$(3.1) \quad (W_{k,n}, 0 \leq k \leq n) \stackrel{d}{=} (\tilde{W}_{k,n}, 0 \leq k \leq n),$$

where the r.h.s. is $\Phi_n(\tilde{S}_{k,n}, 0 \leq k \leq n)$.

Observe that as n varies, the two sequences displayed in (3.1) develop by very different stochastic mechanisms. For instance, on the l.h.s., as n increments to $n + 1$, if X_{n+1} is sufficiently large negative, leading to a sum S_{n+1} with $S_{n+1} = M_{0,n} - a$ with $a > 0$, then $M_{0,n+1} = M_{0,n} - a$, so $W_{1,n} = a$ and $W_{k+1,n+1} = a + W_{k,n}$ for every $1 \leq k \leq n$. In particular, if the process of downward ladder indices of S is recurrent, this will happen infinitely often, and for any fixed K as n varies the values of $(W_{k,n}, 0 \leq k \leq K)$ will keep changing as $n \rightarrow \infty$. In contrast, the right side of (3.1) evolves in a completely different, and much simpler way, which is the key point of this argument. As n increments to $n + 1$, all that happens on the r.h.s. of (3.1) is that one of the two partial Feller chains $(S_k^\uparrow, 0 \leq k \leq N_n^+)$ and $(S_k^\downarrow, 0 \leq k \leq N_n^-)$ has its length incremented by 1, with the consequence that $(\tilde{X}_{1,n+1}, \dots, \tilde{X}_{n+1,n+1})$ is obtained from $(\tilde{X}_{1,n}, \dots, \tilde{X}_{n,n})$ by either prepending or appending X_{n+1} according to whether $S_{n+1} \leq 0$ or $S_{n+1} > 0$. If $S_{n+1} \leq 0$ then $(\tilde{W}_{k,n}, 1 \leq k \leq n + 1)$ is derived by insertion of the value $-S_{N_{n+1}^-}^\downarrow$ into the previous list $(\tilde{W}_{k,n}, 1 \leq k \leq n)$, while if $S_{n+1} > 0$ then $(\tilde{W}_{k,n}, 1 \leq k \leq n + 1)$ is derived by insertion of the value $S_{N_{n+1}^+}^\uparrow$ into the previous list. In either case, the previous list is updated by insertion of a single element, which finds its place somewhere in the previous list. All lower values in the list remain unchanged, while the indexing of all higher values in the list is pushed up by 1. In terms of the associated point processes on $[0, \infty)$, all that happens is that a single point is inserted somewhere in the configuration of n points to make a new configuration of $n + 1$ points. As a consequence, for each fixed k , with probability one,

$$(3.2) \quad \tilde{W}_{k,n} \text{ for } n \geq k \text{ is weakly decreasing in } n \text{ and eventually equal to its limit } W_k.$$

To be more precise, let

$$(3.3) \quad M_{>n} := \min(\{S_j^\uparrow, j \geq N_n^+\} \cup \{-S_j^\downarrow, j \geq N_n^-\}).$$

Let $\text{TV}(K, n)$ be the total variation distance between the distributions of the two sequences displayed in (1.5). Then in view of (3.1) there is the coupling bound [38, (2.6), p. 12]:

$$(3.4) \quad \text{TV}(K, n) \leq \mathbb{P}(\tilde{W}_{n,k} \neq W_k \text{ for some } 1 \leq k \leq K) \leq \mathbb{P}(M_{>n} \leq W_K),$$

since if $M_{>n} > W_K$ then every insertion after time n does not change the value of any $\tilde{W}_{k,n}$ with $1 \leq k \leq K$, and hence $\tilde{W}_{k,m} = W_k$ for all $m \geq n$. Now by Lemma 2.2, the sequences S^\uparrow and $-S^\downarrow$ are upward transient, which implies that $\mathbb{P}(M_{>n} \uparrow \infty) = 1$. So the probability bound in (3.4) decreases to 0 as $n \rightarrow \infty$. This gives part (i). The proof of (ii) is in the same vein.

3.2. *Proof of Proposition 1.2*

The key to the analysis of Proposition 1.2 relies on the following representation of the order statistics $M_{k,n}$, which is due to Pollaczek [45] and Wendel [62, Theorem 2.2(a)].

Lemma 3.1 ([45,62]). *Let $(S_k, 0 \leq k \leq n)$ be a real-valued walk with exchangeable increments, and let $M_{k,n}$ be the order statistics defined by (1.1). For each $0 \leq k \leq n$, the distribution of $M_{k,n}$ is the convolution of the distributions of maximal and minimal values over segments of the walk of length k and $n - k$ respectively:*

$$(3.5) \quad M_{k,n} \stackrel{d}{=} \bar{S}_k + S'_{n-k},$$

where S' is copy of S and independent of S .

Pollaczek proved Lemma 3.1 for a random walk with i.i.d. increments, and attributed this result to Bohnenblust, Spitzer and Welsh. Later Wendel [62] and Port [46, Section V] provided more combinatorial proofs. As pointed out by Dassios [17], Wendel's argument carries over to any walk with exchangeable increments. Continuous analogs of (3.5) on the quantile of stochastic processes were studied in [16,17,23].

We also recall a formula for the expectation of $\max_{0 \leq k \leq n} S_k$ and $\min_{0 \leq k \leq n} S_k$, which is a simple consequence of Spitzer's identity [57, Theorem 3.1] for S a random walk with i.i.d. increments. It is lesser known that this formula also holds for a walk with exchangeable increments [10,56].

Lemma 3.2 ([56,57]). *Let $(S_k, 0 \leq k \leq n)$ be a real-valued walk with exchangeable increments. Then*

$$(3.6) \quad \mathbb{E}\bar{S}_n = \sum_{k=1}^n \frac{\mathbb{E}S_k^+}{k} \quad \text{and} \quad \mathbb{E}\underline{S}_n = - \sum_{k=1}^n \frac{\mathbb{E}S_k^-}{k},$$

where $S_k^+ := \max(S_k, 0)$ and $S_k^- := -\min(S_k, 0)$.

Proof of Proposition 1.2. Part (i) is a consequence of the monotone decreasing nature of the convergence (3.2): the almost sure convergence of $\tilde{W}_{k,n}$ to W_k is dominated by $\tilde{W}_{k,k}$, which is easily seen to have a finite p^{th} moment if so does $|X|$. The argument for $D_{k,n}$ is similar.

For part (ii), by Lemma 3.1 and Lemma 3.2, we have for any distribution of increments X with $\mathbb{E}|X| < \infty$,

$$(3.7) \quad \mathbb{E}M_{k,n} = \sum_{j=1}^k \frac{\mathbb{E}(S_j^+)}{j} - \sum_{j=1}^{n-k} \frac{\mathbb{E}(S_j^-)}{j},$$

and hence by differencing

$$(3.8) \quad \mathbb{E}D_{k,n} = \frac{\mathbb{E}S_k^+}{k} + \frac{\mathbb{E}S_{n-k+1}^-}{n-k+1}.$$

The formula (3.8) implies that $\mathbb{E}D_{n-k+1,n} = \mathbb{E}D_{k,n}$, which is consistent with the reversibility of spacings (1.7). The conclusion of part (ii) follows from (3.8) and the well known fact [20, Section 4.7] that S_n/n is a reversed martingale so $S_{n-k+1}^-/(n-k+1)$ converges as $n \rightarrow \infty$ both almost surely and in L^1 to $\mathbb{E}(X|\mathcal{E})^-$. Moreover, if $\mathbb{E}(X|\mathcal{E}) = 0$ almost surely and $\mathbb{E}X^2 < \infty$, then $\mathbb{E}D_k = \mathbb{E}S_k^+/k$. Note that $\mathbb{E}(S_k^+/\sqrt{k}) = \mathbb{E}[\mathbb{E}(S_k^+/\sqrt{k}|\mathcal{E})]$. The asymptotics (1.10) follows as in the i.i.d. case by de Finetti's theorem and the dominated convergence theorem since $\mathbb{E}(S_k^+/\sqrt{k}|\mathcal{E}) \leq \sqrt{\mathbb{E}(S_k^2/k|\mathcal{E})} = \sqrt{\mathbb{E}(X^2|\mathcal{E})}$. \square

3.3. Proof of Corollary 1.4

Recall from Section 2 that G_1^\uparrow (resp. G_1^\downarrow) is the future minimum time of S^\uparrow (resp. the future maximum time of S^\downarrow). For $w > 0$, we have:

$$(3.9) \quad \begin{aligned} \mathbb{P}(W_1 > w) &= \mathbb{P}(S_{G_1^\uparrow}^\uparrow > w, -S_{G_1^\downarrow}^\downarrow > w) \\ &= \mathbb{P}(S_{G_1^\uparrow}^\uparrow > w)\mathbb{P}(-S_{G_1^\downarrow}^\downarrow > w), \end{aligned}$$

where the first equality is by the construction in Theorem 1.1, and the second equality is due to the independence of the Feller chains $(S^\uparrow, S^\downarrow)$ of a random walk. Further by Proposition 2.3 (Tanaka's construction of Feller's chains of a random walk),

$$(3.10) \quad S_{G_1^\uparrow}^\uparrow \stackrel{d}{=} S_{\tau^+} \quad \text{and} \quad S_{G_1^\downarrow}^\downarrow = S_{\tau^-}.$$

Combining (3.9) and (3.10) yields the equation (1.13).

3.4. Proof of Proposition 1.6

Recall the definition of $M_{>n}$ from (3.3), and it follows from the coupling bound (3.4) that

$$\begin{aligned} \text{TV}(K, n) &\leq \mathbb{P}(M_{>n} \leq W_K) \\ &\leq \mathbb{P}\left(\min_{k \geq N_n^+} S_k^\uparrow \leq W_K\right) + \mathbb{P}\left(\min_{k \geq N_n^-} (-S_k^\downarrow) \leq W_K\right). \end{aligned}$$

For $k \geq 1$, let $W_k^\downarrow := M_k(\{-S_n^\downarrow, n \geq 0\})$ be the k th smallest element in the chain $-S^\downarrow$. By definition, for each finite K we have $W_K \leq W_K^\downarrow$ almost surely. Therefore,

$$(3.11) \quad \mathbb{P}\left(\min_{k \geq N_n^+} S_k^\uparrow \leq W_K\right) \leq \mathbb{P}\left(\min_{k \geq N_n^+} S_k^\uparrow \leq W_K^\downarrow\right).$$

By Theorem 1.3, the two Feller chains S^\uparrow and S^\downarrow are independent, and so are $\min_{k \geq N_n^+} S_k^\uparrow$ and W_K^\downarrow . Thus, the key to the analysis of (3.11) is to bound $\mathbb{P}(\min_{k \geq N_n^+} S_k^\uparrow \leq x)$ for any fixed $x > 0$. Introduce an auxiliary quantity $A(n)$ such that

$$A(n) \rightarrow \infty \quad \text{and} \quad A(n)/\sqrt{n} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

which we will determine later. Note that

$$(3.12) \quad \mathbb{P}\left(\min_{k \geq N_n^+} S_k^\uparrow \leq x\right) \leq \mathbb{P}(S_{N_n^+}^\uparrow \leq A(n)) + \mathbb{P}\left(S_{N_n^+}^\uparrow > A(n), \min_{k \geq N_n^+} S_k^\uparrow \leq x\right).$$

Now we treat the two terms on the r.h.s. of (3.12) separately. First by Lemma 2.1, $S_{N_n^+}^\uparrow \stackrel{d}{=} \max_{0 \leq k \leq n} S_k$. By the KMT embedding [35], there is a coupling between the random walk $(S_k, k \geq 0)$ and Brownian motion $(B_t, t \geq 0)$ such that

$$\mathbb{P}\left(\max_{0 \leq k \leq n} |S_k - B_k| > C_1 \log n\right) \leq n^{-\frac{1}{2}} \quad \text{for some } C_1 > 0.$$

As a consequence,

$$\begin{aligned} \mathbb{P}(S_{N_n^+}^\uparrow \leq A(n)) &= \mathbb{P}\left(\max_{0 \leq k \leq n} S_k \leq A(n)\right) \\ &\leq n^{-\frac{1}{2}} + \mathbb{P}\left(\max_{0 \leq k \leq n} B_k \leq C_1 \log n + A(n)\right) \\ (3.13) \quad &\leq n^{-\frac{1}{2}} + C_2 n^{-\frac{1}{2}} (\log n + A(n)) \leq C_3 n^{-\frac{1}{2}} (\log n + A(n)), \end{aligned}$$

for some $C_2, C_3 > 0$. Now let $\tau_{A(n)}^\uparrow := \inf\{k : S_k^\uparrow > A(n)\}$ the first time at which the upward Feller chain S^\uparrow enters $(A(n), \infty)$. It is easy to see that $\{S_{N_n^+}^\uparrow > A(n), \min_{k \geq N_n^+} S_k^\uparrow \leq x\} \subset \{\tau_{A(n)}^\uparrow < \infty, \min_{k \geq \tau_{A(n)}^\uparrow} S_k^\uparrow \leq x\}$. By conditioning on the value of $S_{\tau_{A(n)}^\uparrow}^\uparrow$, we get

$$\begin{aligned} \mathbb{P}\left(S_{N_n^+}^\uparrow > A(n), \min_{k \geq N_n^+} S_k^\uparrow \leq x\right) &\leq \mathbb{P}\left(\tau_{A(n)}^\uparrow < \infty, \min_{k \geq \tau_{A(n)}^\uparrow} S_k^\uparrow \leq x\right) \\ &= \int_{A(n) \vee x}^{\infty} \mathbb{P}\left(\min_{k \geq \tau_{A(n)}^\uparrow} S_k^\uparrow \leq x \mid S_{\tau_{A(n)}^\uparrow}^\uparrow = y\right) \mathbb{P}(\tau_{A(n)}^\uparrow < \infty, S_{\tau_{A(n)}^\uparrow}^\uparrow \in dy). \end{aligned}$$

According to [12, Lemma 4.1 & A.1],

$$\mathbb{P}\left(\min_{k \geq \tau_{A(n)}^\uparrow} S_k^\uparrow \leq x \mid S_{\tau_{A(n)}^\uparrow}^\uparrow = y\right) = \frac{h^\uparrow(y) - h^\uparrow(y-x)}{h^\uparrow(y)},$$

where $h^\uparrow(\cdot)$ is defined in Theorem 1.3. It is also known (see e.g. [9, p. 2155]) that $h^\uparrow(\cdot)$ is the renewal function of the ladder height S_{τ^+} . By subadditivity of the renewal function, we have $h^\uparrow(y) - h^\uparrow(y-x) \leq h^\uparrow(x)$. Since $\mathbb{E}|X|^2 < \infty$, it follows from the key renewal theorem that $h^\uparrow(y) \sim C_4 y$ as $y \rightarrow \infty$ for some $C_4 > 0$. The above observations imply that

$$(3.14) \quad \mathbb{P}\left(S_{N_n^+}^\uparrow > A(n), \min_{k \geq N_n^+} S_k^\uparrow \leq x\right) \leq C_5 h^\uparrow(x) A(n)^{-1},$$

for some $C_5 > 0$. Combining (3.12), (3.13) and (3.14) yields for $x > 0$,

$$(3.15) \quad \mathbb{P}\left(\min_{k \geq N_n^+} S_k^\uparrow \leq x\right) \leq C_3 n^{-\frac{1}{2}} (\log n + A(n)) + C_5 h^\uparrow(x) A(n)^{-1}.$$

By taking $A(n) = n^{\frac{1}{4}}$ (which optimizes the right side of (3.15) relative to $A(n)$), we get

$$(3.16) \quad \mathbb{P}\left(\min_{k \geq N_n^+} S_k^\uparrow \leq x\right) \leq C_6 (1 + h^\uparrow(x)) n^{-\frac{1}{4}},$$

for some $C_6 > 0$. Now by (3.11) and (3.16), we have

$$\mathbb{P}\left(\min_{k \geq N_n^+} S_k^\uparrow \leq W_K\right) \leq C_6 (1 + \mathbb{E}h^\uparrow(W_K^\downarrow)) n^{-\frac{1}{4}}.$$

Finally, by Proposition 1.2(i) and the fact that $h^\uparrow(y)/y^2 \rightarrow 0$ as $y \rightarrow \infty$, we get $\mathbb{E}h^\uparrow(W_K^\downarrow) < \infty$. A similar bound is derived for $\mathbb{P}(\min_{k \geq N_n^-} (-S_k^\downarrow) \leq W_K)$, which yields the desired result.

4. Order statistics of simple symmetric random walks

This section describes the limit process $(W_k, k \geq 0)$ of shifted order statistics $(M_{k,n} - M_{0,n}, 0 \leq k \leq n)$ derived from the path of a random walk with symmetric ± 1 increments; that is $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$. According to Theorem 1.1, $(W_k, k \geq 0)$ is the increasing rearrangement of the two Feller chains $(S_k^\uparrow, k \geq 1)$ and $(-S_k^\downarrow, k \geq 0)$. The key to the analysis is the occupation counting process $(L_\ell, \ell = 0, 1, \dots)$ defined by

$$(4.1) \quad L_\ell := \sum_{k=0}^{\infty} 1(W_k = \ell) = L_\ell^\uparrow + L_\ell^\downarrow,$$

where

$$(4.2) \quad L_\ell^\uparrow := \sum_{k=1}^{\infty} 1(S_k^\uparrow = \ell) \quad \text{and} \quad L_\ell^\downarrow := \sum_{k=0}^{\infty} 1(-S_k^\downarrow = \ell).$$

Note the convention that the minimal order statistic $W_0 = 0$ is included as a contribution to the count $L_0^\downarrow \geq 1$ of values attaining the minimum, while the 0 value is deliberately excluded as a contribution to $L_0^\uparrow = 0$.

For the simple symmetric random walk the general formula (1.9) for $\mathbb{E}D_k$ simplifies to

$$(4.3) \quad \mathbb{E}D_k = \frac{\mathbb{E}S_k^+}{k} = \frac{1}{2} u_{\lfloor k/2 \rfloor},$$

where $u_m := \mathbb{P}(S_{2m} = 0) = \binom{2m}{m} 2^{-2m} \sim \frac{1}{\sqrt{\pi m}}$ as $m \rightarrow \infty$, is the central term in the Binomial $(2m, 1/2)$ distribution. The second equality in (4.3) follows from Kemperman's formula and the reflection principle. By Kemperman's formula and time reversal,

$$\begin{aligned} j\mathbb{P}(S_k = j) &= k\mathbb{P}(S_1 < j, \dots, S_{k-1} < j, S_k = j) \\ &= k\mathbb{P}(S_1 > 0, \dots, S_{k-1} > 0, S_k = j). \end{aligned}$$

By summing both sides over $j > 0$, we get $\mathbb{E}S_k^+ / k = \mathbb{P}(S_1 > 0, \dots, S_k > 0)$ which equals to $\frac{1}{2}u_{\lfloor k/2 \rfloor}$ by the reflection principle (see e.g. [25, p. 77, (3.2)]). For $\theta > 0$, recall the well known identity:

$$(1 - z)^{-\theta} = 1 + \sum_{k \geq 1} \frac{(\theta)_{k\uparrow}}{k!} z^k,$$

where $(\theta)_{k\uparrow} := \prod_{i=0}^{k-1} (\theta + i)$ is the Pochhammer function. Specializing to $\theta = \frac{1}{2}$ yields the generating function $\sum_{k=0}^{\infty} u_k z^k = (1 - z)^{-1/2}$, which then gives easily the generating function

$$(4.4) \quad \sum_{k=1}^{\infty} \mathbb{E}D_k z^k = \frac{1}{2}((1 + z)(1 - z^2)^{-1/2} - 1).$$

Since the increments X_k of $(S_k, k \geq 0)$ are restricted to ± 1 , the increments D_k of $(W_k, k \geq 0)$ are restricted to be 0 or 1. Thus D_k is an indicator random variable, more precisely

$$(4.5) \quad D_k = \sum_{j=0}^{\infty} 1(\eta_j = k) = 1(\eta_j = k \text{ for some } j \geq 0) \quad \text{for } k = 1, 2, \dots,$$

where

$$(4.6) \quad \eta_j := \sum_{\ell=0}^j L_{\ell} = \sum_{k=1}^{\infty} 1(W_k \leq j),$$

is the number of all order statistics at or below level j . Formula (4.5) looks as if a process of renewal indicators is constructed from the times η_j of renewals generated by spacings L_{ℓ} in a delayed renewal process [4, Chapter VI]. So the formula (4.3) for the indicators D_k defined by (4.5) parallels the well known role of $u_k = \mathbb{P}(S_{2k} = 0)$ in the renewal indicator process $(1(S_{2k} = 0), k \geq 0)$ whose spacings are independent copies of the time of first return to 0 for the original random walk (see e.g. [25, Chapter XIII]). However, the stochastic structure of the $(L_{\ell}, \ell \geq 0)$ is more complex than the independent random spacings in a renewal process. See also [21,40,41,47,49] for the distribution of order statistics and related functionals of simple random walks.

Call a process $(Z_k, k \geq 0)$ a *branching process with immigration distribution F_0 and offspring distribution F_1* if $(Z_k, k \geq 0)$ is a Markov chain with

$$(Z_{k+1} | Z_0, \dots, Z_k) \stackrel{d}{=} Y_0 + \sum_{i=1}^{Z_k} Y_i, \quad k \geq 0,$$

where X_0, X_1, X_2, \dots are independent, with Y_0 distributed according to F_0 and Y_i for $i \geq 1$ distributed according to F_1 . In the present setting, Y_0 will always be a constant, either 0 or 1 or 2, in which case $(Z_k, k \geq 0)$ is said to have either *no immigration* or *single* or *double immigration* as the case may be. We say Y has the $\text{Geo}_0(p)$ distribution on $\{0, 1, \dots\}$ if $\mathbb{P}(Y = k) = p(1 - p)^k$ for $k \geq 0$, and Y has the $\text{Geo}_1(p)$ distribution on $\{1, 2, \dots\}$ if $\mathbb{P}(Y = k) = p(1 - p)^{k-1}$ for $k \geq 1$. The following theorem summarizes some results underlying the construction of the occupation counting process $(L_{\ell}, \ell \geq 0)$.

Theorem 4.1. *Let $(L_{\ell}, \ell \geq 0)$ be the occupation counting process defined by (4.1). Then*

(i) *L may be represented as*

$$(4.7) \quad L_{\ell} = Z_{\ell-1} + Z_{\ell} - 2, \quad \ell \geq 0,$$

where $(Z_{\ell}, \ell = -1, 0, 1, \dots)$ is a branching process with double immigration and $\text{Geo}_0(1/2)$ offspring distribution, with the convention $Z_{-1} = 1$. In particular, $L_0 = Z_0 - 1$ has $\text{Geo}_1(1/2)$ distribution:

$$(4.8) \quad \mathbb{P}(L_0 = k) = \mathbb{P}(Z_0 = k + 1) = 2^{-k} \quad \text{for } k \geq 1.$$

(ii) For each $\ell \geq 1$, the conditional distribution of L_ℓ given $L_0 = k$ is that of

$$(4.9) \quad G_0(\ell) + G_k(\ell) + \sum_{i=1}^{k-1} B_i(\ell) G_i(\ell),$$

where the $(G_i(\ell), 0 \leq i \leq k)$ are independent $\text{Geo}_1(\frac{1}{2\ell})$ variables, and the $(B_i(\ell), 1 \leq i \leq k-1)$ are $k-1$ independent Bernoulli($\frac{1}{\ell}$) variables, independent also of the $(G_i(\ell), 0 \leq i \leq k)$.

(iii) For each $k \geq 1$, η_k defined by (4.6) has the generating function

$$(4.10) \quad \mathbb{E}_z \eta^k = \frac{1}{V_k(1/z) V_{k+1}(1/z)},$$

where V_k is the sequence of Chebyshev polynomials of the third kind, defined by the Chebyshev recurrence

$$(4.11) \quad V_k(x) := 2x V_{k-1}(x) - V_{k-2}(x), \quad k \geq 2,$$

with initial condition $V_0(x) := 1$ and $V_1(x) := 2x - 1$.

To prove Theorem 4.1, we need the following two lemmas regarding the Feller chain S^\uparrow derived from a simple symmetric random walk.

Lemma 4.2 ([34]). Let S^\uparrow and S^\downarrow be the upward and downward Feller chains derived from a simple symmetric random walk S . For $\ell \geq 0$, let

$$(4.12) \quad Z_\ell^\uparrow := \sum_{k=0}^{\infty} 1(S_k^\uparrow = \ell, S_{k+1}^\uparrow = \ell + 1), \quad Z_\ell^\downarrow := \sum_{k=0}^{\infty} 1(-S_k^\downarrow = \ell, -S_{k+1}^\downarrow = \ell + 1),$$

be the total number of upcrossings of S^\uparrow and $-S^\downarrow$ from ℓ to $\ell + 1$. Then we have:

- (i) The sequence $(Z_\ell^\uparrow, \ell \geq 0)$ is a branching process with single immigration and $\text{Geo}_0(1/2)$ offspring distribution.
- (ii) The sequence $(Z_\ell^\downarrow, \ell \geq 0)$ is a branching process with single immigration and $\text{Geo}_0(1/2)$ offspring distribution, which evolves independently of Z^\uparrow , but with a time shift forward by one step. So Z_0^\downarrow starts with the $\text{Geo}_1(1/2)$ distribution.

Lemma 4.3. For $k \geq 1$ let $\tau_k := \inf\{n \geq 1 : S_n = k\}$ be the first passage time to level $k > 0$ for the simple symmetric random walk S . Let $N_{0+}(n) := \sum_{j=1}^n 1(S_j \geq 0)$ be the number of non-negative steps of the walk up to time n , let $\bar{\tau}_k := \inf\{n \geq 1 : S_n - \underline{S}_n = k\}$ be the first passage time to level k for the reflected random walk $(S_n - \underline{S}_n, n \geq 0)$, and let $\eta_k^\uparrow := \sum_{j=1}^{\infty} 1(S_j^\uparrow \leq k)$ be the total occupation time of levels 1 through k for the upward Feller chain. Then we have:

- (i) The three random variables $N_{0+}(\tau_k)$, $\bar{\tau}_k$ and η_k^\uparrow have identical distributions.
- (ii) The common distribution of these random variables has the generating function

$$(4.13) \quad g_k(z) := \mathbb{E}_z N_{0+}(\tau_k) = \mathbb{E}_z \bar{\tau}_k = \mathbb{E}_z \eta_k^\uparrow = \frac{1}{V_k(1/z)},$$

where V_k is the sequence of Chebyshev polynomials of the third kind defined by (4.11).

Proof. (i) It follows from [43] that there is the identity in distribution:

$$(S_j^\uparrow, 0 \leq j \leq \sigma_k) \stackrel{d}{=} (k - S_{\tau_k - j}, 0 \leq j \leq \tau_k),$$

where $\sigma_k := \max\{j, S_j^\uparrow = k\}$ is the last time at which S^\uparrow visits the state k . As a result, $N_{0+}(\tau_k) = \sum_{j=0}^{\tau_k-1} 1(k - S_{\tau_k - j} \leq k) \stackrel{d}{=} \sum_{j=0}^{\tau_k-1} 1(S_j^\uparrow \leq k) = \eta_k^\uparrow$, where the last equality follows from the fact that S^\uparrow does not visit the state k after time σ_k . The further equality in distribution of $N_{0+}(\tau_k)$ and $\bar{\tau}_k$ follows easily from the stronger result that if $\nu_0 := 0$ and ν_k for $k > 0$ is the first time n such that $N_{0+}(n) = k$, then $(S_{\nu_k}, k \geq 0)$, which is just the subsequence of non-negative terms of S , has the same distribution as $(S_n - \underline{S}_n, n \geq 0)$. So each process is just simple symmetric random on the non-negative integers, in which the transition probabilities out of state 0 are $P(0, 0) = P(0, 1) = 1/2$.

(ii) Let $g_k(z) := \mathbb{E}z^{N_{0+}(\tau_k)}$. Let $(\mathcal{F}_k, k \geq 0)$ be the filtration of the random walk S . Note that $S_{\tau_{k-1}+1} = k$ or $k-2$ with equal probabilities $\frac{1}{2}$. By the strong Markov property of S at time τ_{k-1} , we have:

$$(N_{0+}(\tau_k) | \mathcal{F}_{\tau_{k-1}}) = \begin{cases} N_{0+}(\tau_{k-1}) + 1 & \text{with probability } \frac{1}{2}, \\ N_{0+}(\tau_{k-1}) + 1 + N_{0+}(\tau_{k-2:k}) & \text{with probability } \frac{1}{2}, \end{cases}$$

where $N_{0+}(\tau_{k-2:k})$ is the number of non-negative steps of the walk starting at $k-2$ until it first visits k , independent of $N_{0+}(\tau_{k-1})$. Therefore, $g_k(z) = \frac{1}{2}zg_{k-1}(z) + \frac{1}{2}zg_{k-1}(z)\mathbb{E}z^{N_{0+}(\tau_{k-2:k})}$. Further by the strong Markov property of S at time τ_{k-2} , we have $g_{k-2}(z)\mathbb{E}z^{N_{0+}(\tau_{k-2:k})} = g_k(z)$. By eliminating the term $\mathbb{E}z^{N_{0+}(\tau_{k-2:k})}$, we get the recursion

$$(4.14) \quad g_k(z) = g_{k-1}(z)z \left(\frac{1}{2} + \frac{1}{2} \frac{g_k(z)}{g_{k-2}(z)} \right) \quad \text{for } k \geq 2,$$

with the convention that $g_0(z) := 1$ and $g_{-1}(z) := 1$ (so $g_1(z) = \frac{z}{2-z}$). Let $V_k(x) := 1/g_k(1/x)$. Then (4.14) reduces easily to (4.11), with $V_0(x) = 1$ and $V_1(x) = 2x - 1$. \square

Here we do not use explicitly any embedding for the simple symmetric walk. Nevertheless, Lemmas 4.2 and 4.3 are reminiscent of Ray–Knight theorems and Williams’ decomposition, which have been proved via random walk embedding.

Proof of Theorem 4.1. Part (i) follows from Lemma 4.2, and the path construction

$$(4.15) \quad L_\ell^\uparrow = Z_{\ell-1}^\uparrow + Z_\ell^\uparrow - 1 \quad \text{and} \quad L_\ell^\downarrow = Z_{\ell-1}^\downarrow + Z_\ell^\downarrow - 1, \quad \ell \geq 0,$$

with the conventions $Z_{-1}^\uparrow = 0$ and $Z_{-1}^\downarrow = 1$ for $\ell = 0$.

Part (ii) can be read from Révész [50, Lemma 12.5]. For S a simple symmetric walk, let $\tau_0 := \inf\{n \geq 1 : S_n = 0\}$ be the first time at which S returns to 0, and $\xi_\ell := \sum_{k=1}^{\tau_0} 1(S_k = \ell)$ be the number of visits to the state $\ell \geq 0$ up to time τ_0 . It was shown that

$$(4.16) \quad \left(\xi_{n-\ell} | \max_{0 \leq j \leq \tau_0} S_j = n, \xi_n = k \right) \stackrel{d}{=} G_0(\ell, n) + G_k(\ell, n) + \sum_{i=1}^{k-1} B_i(\ell, n) G_i(\ell, n),$$

where the $(G_i(\ell, n), 0 \leq i \leq k)$ are independent $\text{Geo}_1(\frac{n}{2\ell(n-\ell)})$ variables, and the $(B_i(\ell, n), 1 \leq i \leq k-1)$ are $k-1$ independent Bernoulli $(1 - (1 - \frac{1}{\ell})\frac{n}{n-1})$ variables, independent also of the $(G_i(\ell, n), 0 \leq i \leq k)$. Given that $\max_{0 \leq j \leq \tau_0} S_j = n$, the random walk excursion is constructed by joining back to back independently a positive walk running till it first hits n , and a nonnegative walk running till it first hits n . According to [9], as $n \rightarrow \infty$, a walk conditioned to hit n before 0 converges in distribution to S^\uparrow and a walk conditioned to hit n before -1 converges in distribution to $-S^\downarrow$. Thus, near $\max_{0 \leq j \leq \tau_0} S_j = n \rightarrow \infty$, the random walk excursion is decomposed into $(S^\uparrow, -S^\downarrow)$. Consequently, the left side of (4.16) converges in distribution to $(L_\ell | L_0 = k)$. Moreover, $G_i(\ell, n)$ and $B_i(\ell, n)$ converge in distribution to $G_i(\ell)$ and $B_i(\ell)$ respectively, hence the conclusion.

For part (iii), let $\eta_k^\uparrow := \sum_{j=1}^\infty 1(S_j^\uparrow \leq k)$. It is easy to see that $\eta_k \stackrel{d}{=} \eta_k^\uparrow + \tilde{\eta}_{k+1}^\uparrow$, where $\tilde{\eta}_{k+1}^\uparrow$ is a copy of η_{k+1}^\uparrow , independent of η_k^\uparrow . It suffices to apply Lemma 4.3 to conclude. \square

5. Order statistics of Gaussian walks

This section provides further analysis of limiting order statistics of a random walk with i.i.d. Gaussian increments; that is $\mathbb{P}(X_i \in dx) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp(-\frac{x^2}{2\sigma^2}) dx$ for some $\sigma^2 > 0$. As discussed in the Introduction, a random walk $(S_k, 0 \leq k \leq n)$ with Gaussian increments is embedded in a Brownian motion $(B(t), t \geq 0)$ via $S_k = B(\sigma k)$. There is one more variable to be entered into the mix; that is

$$(5.1) \quad M_{-,n} := \min_{0 \leq t \leq \sigma n} B(t).$$

So by definition, $M_{-,n} \leq M_{0,n} \leq \dots \leq M_{n,n}$ are the ranked values of Brownian motion on $[0, \sigma]$ evaluated at a grid of $n+2$ fixed times $0, \sigma, 2\sigma, \dots, \sigma n$ at which the random walk $S_k := B(\sigma k)$ is embedded, and one extra random time

$\alpha(\sigma n) := \inf\{t \leq \sigma n : B(t) = \underline{B}(\sigma n)\}$, the almost surely unique random time at which B attains its minimum on $[0, \sigma n]$. It is a key observation that the structure of limit distributions for differences of order statistics $M_{k,n} - M_{0,n}$ for $k = O(\sqrt{n})$ as $n \rightarrow \infty$ is easily understood by first considering the differences $M_{k,n} - M_{-,n}$, from which the $M_{k,n} - M_{0,n}$ can be recovered using Theorem 5.1 as

$$(5.2) \quad M_{k,n} - M_{0,n} = (M_{k,n} - M_{-,n}) - (M_{0,n} - M_{-,n}).$$

The distribution of $M_{0,n} - M_{-,n}$ was studied in Asmussen, Glynn and Pitman [5] in the Gaussian case, where $M_{0,n} - M_{-,n}$ was interpreted as the discretization error in approximating a reflecting random walk by a natural Euler approximation scheme. The key idea was that when the Gaussian random walk S is embedded in Brownian motion B as $S_k = B(\sigma k)$ for $0 \leq k \leq n$, and n is large, the discretization error $M_{0,n} - M_{-,n}$ is with overwhelming probability determined by the behavior of B in a time window of width $o(n)$ around the minimum time $\alpha(\sigma n)$. According to basic decompositions of a Brownian path at a local minimum, due to Williams [64] and Denisov [19], after centering at $(\alpha, B(\alpha))$ for $\alpha = \alpha(\sigma n)$ the height of the Brownian path above its minimum value becomes locally well approximated by two independent copies of a BES(3) process, say $(R_3(t), t \geq 0)$, and $(\hat{R}_3(t), t \geq 0)$, joined back to back, call it the Brownian valley:

$$(5.3) \quad \check{R}(t) := R_3(t)1(t \geq 0) + \hat{R}_3(t)1(t < 0), \quad \text{for } -\infty < t < \infty.$$

It is easily shown that for each fixed $\varepsilon > 0$, there exists $a(\varepsilon) > 0$ and $n(\varepsilon)$ so that for $n \geq n(\varepsilon)$ the probability that the Gaussian walk attains its minimum over $\{0, 1, \dots, n\}$ at a time k with $|k - \alpha(n)| < a(\varepsilon)$ is at least $1 - \varepsilon$. As shown in [5, Theorem 1], it follows the discretization error for the standard Gaussian random walk S is described by

$$(5.4) \quad M_{0,n} - M_{-,n} = \min_{0 \leq k \leq n} S_k - \min_{0 \leq t \leq n} B_t \xrightarrow{d} \min_{k \in \mathbb{Z}} \check{R}(U + k)$$

with U Uniform $[0, 1]$ and independent of \check{R} . The same argument can establish more, which is stated in the following theorem.

Theorem 5.1. *Let $(S_k, k \geq 0)$ be a Gaussian random walk, for increments with variance σ^2 . For $0 \leq k \leq n$, let $M_{k,n}$ be the sequence of order statistics defined by (1.1) of the n -step walk $(S_k, 0 \leq k \leq n)$. Then for every $m = 1, 2, \dots$, there is the convergence of joint distributions*

$$(5.5) \quad (M_{k,n} - M_{-,n}, 0 \leq k \leq m) \xrightarrow{d} (\sigma M_{k,\infty}, 0 \leq k \leq m),$$

where $M_{k,\infty} := M_k(\{\check{R}(U + j), j \in \mathbb{Z}\})$ is the sequence of order statistics of the values $\{\check{R}(U + j), j \in \mathbb{Z}\}$, and U with uniform distribution on $[0, 1]$ is independent of \check{R} . Moreover, the distribution of $M_{0,\infty} := \min_{k \in \mathbb{Z}} \check{R}(U + k)$ is given by

$$(5.6) \quad \mathbb{P}(M_{0,\infty} > a) = \int_0^1 G^a(u)G^a(1 - u) du,$$

with

$$(5.7) \quad G^a(u) = K^a(u) \prod_{k=0}^{\infty} \frac{H^a(k + u)}{K^a(k + u)},$$

where

$$(5.8) \quad K^a(t) = \sqrt{\frac{2}{\pi t}} a e^{-\frac{a^2}{2t}} + \operatorname{erfc}\left(\frac{a}{\sqrt{2t}}\right),$$

and

$$(5.9) \quad \begin{aligned} H^a(t) = & \frac{a}{\sqrt{2\pi(1+t)}} e^{-\frac{a^2}{2(1+t)}} \left(\operatorname{erfc}\left(\frac{a}{\sqrt{2t(t+1)}}\right) + \operatorname{erfc}\left(\frac{a(2t+1)}{\sqrt{2t(t+1)}}\right) \right) \\ & + \frac{1}{\pi\sqrt{t}} (e^{-\frac{a^2}{2t}} - e^{-\frac{a^2(4t+1)}{2t}}) + \frac{a}{\sqrt{2\pi t}} e^{-\frac{a^2}{2t}} (1 + \operatorname{erfc}(\sqrt{2}a)) + \operatorname{erfc}\left(\frac{a}{\sqrt{2t}}\right) \\ & + \operatorname{erfc}\left(\frac{a}{\sqrt{2(t+1)}}\right) - T\left(\frac{a}{\sqrt{t+1}}, \frac{1}{\sqrt{t}}\right) - T\left(\frac{a}{\sqrt{t+1}}, \frac{2t+1}{\sqrt{t}}\right) - T\left(\frac{a}{\sqrt{t}}, 2\sqrt{t}\right). \end{aligned}$$

Here $\operatorname{erfc}(x) := 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the complementary error function, and $T(h, a)$ is Owen's T function defined by

$$(5.10) \quad T(h, a) := \frac{1}{2\pi} \int_0^a \frac{e^{\frac{1}{2}h^2(1+x^2)}}{1+x^2} dx.$$

Note that in the Gaussian case, the limit in distribution conjectured by Schehr and Majumdar [52] for D_k and W_k can be read from the above theorem, with the process $(D_k, k \geq 1)$ identified as $D_k = M_{k,\infty} - M_{k-1,\infty}$, and the partial sums $(W_k, k \geq 0)$ identified as $W_k = M_{k,\infty} - M_{0,\infty}$. It is a key point of the present approach that the enlargement of the probability space to include an auxiliary Brownian motion enables a simple description of the limit process, involving a single additional variable $M_{0,\infty}$ which places the limiting order statistics in their natural environment, which is a process of random sampling of points in the Brownian valley. See also [14,29] for related works on Lévy processes and random walks with heavy tailed increments.

As discussed in [5], in the Gaussian case, it is a difficult problem to describe the even the limit distribution of the $M_{0,\infty}$ in a simple way, see e.g. formula (5.6)–(5.9). While bounds on the distribution of $M_{0,\infty}$ allow to show that this variable has finite positive moments of all orders, the only exact result found is the mean:

$$(5.11) \quad \mathbb{E}M_{0,\infty} = -\frac{\sigma}{\sqrt{2\pi}} \zeta(1/2) = 0.5826 \cdots \sigma,$$

where $\zeta(\cdot)$ is the Riemann zeta function. In view of (5.6), we get the following non-trivial identity:

$$(5.12) \quad \int_0^\infty \int_0^1 K^a(u) K^a(1-u) \prod_{k=0}^\infty \frac{H^a(k+u) H^a(k+1-u)}{K^a(k+u) K^a(k+1-u)} du da = -\frac{1}{\sqrt{2\pi}} \zeta\left(\frac{1}{2}\right).$$

Further computations of explicit distributions in this case seem very difficult, and it is interesting to see if further moments, or the generating function of $M_{0,\infty}$ can be computed by using the product formula (5.6). See also [13,30,31] for closely related computations of the mean of the ladder height distribution for a Gaussian random walk, which also involve Riemann's zeta function.

Proof of Theorem 5.1. The first part of the theorem is a simple consequence of the proof of [5, Theorem 1]. We focus on the second part of the distribution of $M_{0,\infty}$. The derivation relies on some Bessel process computation, which is of independent interest. Fix $a > 0$, we aim to compute

$$\mathbb{P}(M_{0,\infty} > a) := \mathbb{P}(R_3(k+U) > a \text{ and } \hat{R}_3(k+1-U) > a, \forall k \geq 0),$$

where R_3 and \hat{R}_3 are two independent BES(3) processes. By conditioning on U , we get

$$(5.13) \quad \mathbb{P}(M_{0,\infty} > a) = \int_0^1 \mathbb{P}(R_3(k+u) > a, \forall k \geq 0) \cdot \mathbb{P}(R_3(k+1-u) > a, \forall k \geq 0) du.$$

For $u \in (0, 1)$, by letting $G^a(u) = \mathbb{P}(R_3(u+k) > a, \forall k \in \mathbb{N})$, we get (5.6).

Now we calculate $G^a(u)$. For $t > 0$, let $K^a(t) := \mathbb{P}(R_3(t) > a)$ and $H^a(t) := \mathbb{P}(R_3(t+1) > a, R_3(t) > a)$. By the Markov property of the BES(3) process, we can express $G^a(u)$ by an infinite product:

$$G^a(u) = \mathbb{P}(R_3(u) > a) \prod_{k=1}^\infty \mathbb{P}(R_3(k+u) > a | R_3(k-1+u) > a),$$

which leads to (5.7). Recall from [51, Chapter VI.3] the entrance law $p_t(x) := \mathbb{P}(R_3(t) \in dx)/dx$, and the transition density $q_t(x, y) := \mathbb{P}(R_3(t+t_0) \in dy | R_3(t_0) = x)/dy$ of the BES(3) process:

$$(5.14) \quad p_t(x) = \sqrt{\frac{2}{\pi t^3}} x^2 \exp\left(-\frac{x^2}{2t}\right), \quad x > 0,$$

$$(5.15) \quad q_t(x, y) = \frac{x^{-1}y}{\sqrt{2\pi t}} \left[\exp\left(-\frac{(y-x)^2}{2t}\right) - \exp\left(-\frac{(y+x)^2}{2t}\right) \right], \quad x, y > 0.$$

It is easy to see that for $t > 0$, $K^a(t)$ is given by (5.8). It remains to compute $H^a(t)$ for $t > 0$. By definition,

$$(5.16) \quad H^a(t) = \frac{1}{\pi\sqrt{t^3}} \int_a^\infty x \exp\left(-\frac{x^2}{2t}\right) \int_a^\infty y \left[\exp\left(-\frac{(y-x)^2}{2}\right) - \exp\left(-\frac{(y+x)^2}{2}\right) \right] dy dx.$$

Unfortunately, Mathematica did not provide any simplification for (5.16). Here we simplify (5.16) by hand using some special functions. First,

$$(5.17) \quad \begin{aligned} & \int_a^\infty y \left[\exp\left(-\frac{(y-x)^2}{2}\right) - \exp\left(-\frac{(y+x)^2}{2}\right) \right] dy \\ &= e^{-\frac{1}{2}(a-x)^2} - e^{-\frac{1}{2}(a+x)^2} + \sqrt{\frac{\pi}{2}} x \operatorname{erfc}\left(\frac{a-x}{\sqrt{2}}\right) + \sqrt{\frac{\pi}{2}} x \operatorname{erfc}\left(\frac{a+x}{\sqrt{2}}\right). \end{aligned}$$

So we need to compute the following integrals:

$$\begin{aligned} \text{I} &:= \int_a^\infty x \exp\left(-\frac{1}{2}(a-x)^2 - \frac{x^2}{2t}\right) dx, & \text{II} &:= \int_a^\infty x \exp\left(-\frac{1}{2}(a+x)^2 - \frac{x^2}{2t}\right) dx, \\ \text{III} &:= \int_a^\infty x^2 \exp\left(-\frac{x^2}{2t}\right) \operatorname{erfc}\left(\frac{a-x}{\sqrt{2}}\right) dx, & \text{IV} &:= \int_a^\infty x^2 \exp\left(-\frac{x^2}{2t}\right) \operatorname{erfc}\left(\frac{a+x}{\sqrt{2}}\right) dx. \end{aligned}$$

By direct computations, we get

$$(5.18) \quad \text{I} = a \sqrt{\frac{\pi t^3}{2(t+1)^3}} \exp\left(-\frac{a^2}{2(t+1)}\right) \operatorname{erfc}\left(\frac{a}{\sqrt{2t(t+1)}}\right) + \frac{t}{t+1} \exp\left(-\frac{a^2}{2t}\right),$$

$$(5.19) \quad \text{II} = -a \sqrt{\frac{\pi t^3}{2(t+1)^3}} \exp\left(-\frac{a^2}{2(t+1)}\right) \operatorname{erfc}\left(\frac{a(2t+1)}{\sqrt{2t(t+1)}}\right) + \frac{t}{t+1} \exp\left(-\frac{a^2(4t+1)}{2t}\right).$$

Introduce $F(x) := -\frac{t^{3/2}}{2\sqrt{t+1}} e^{-\frac{a^2}{2(t+1)}} \operatorname{erfc}\left(-\frac{(1+t)x+at/\sqrt{2}}{\sqrt{t(t+1)}}\right) - \frac{t}{2} e^{-\frac{x^2}{t}} \operatorname{erfc}\left(x + \frac{a}{\sqrt{2}}\right)$. By integration by parts, we get $\text{III}/\sqrt{8} = \int_{-\infty}^{-\frac{a}{\sqrt{2}}} x^2 \exp\left(-\frac{x^2}{t}\right) \operatorname{erfc}\left(x + \frac{a}{\sqrt{2}}\right) dx = [xF(x)]_{-\infty}^{-\frac{a}{\sqrt{2}}} - \int_{-\infty}^{-\frac{a}{\sqrt{2}}} F(x) dx$. Note that $[xF(x)]_{-\infty}^{-\frac{a}{\sqrt{2}}} = \frac{at^{3/2}}{\sqrt{8(t+1)}} e^{-\frac{a^2}{2(t+1)}} \times \operatorname{erfc}\left(\frac{a}{\sqrt{2t(t+1)}}\right) + \frac{at}{\sqrt{8}} e^{-\frac{a^2}{2t}}$,

$$\begin{aligned} - \int_{-\infty}^{-\frac{a}{\sqrt{2}}} F(x) dx &= \frac{t^{3/2}}{2\sqrt{t+1}} e^{-\frac{a^2}{2(t+1)}} \int_{-\infty}^{-\frac{a}{\sqrt{2}}} \operatorname{erfc}\left(-\frac{(1+t)x+at/\sqrt{2}}{\sqrt{t(t+1)}}\right) dx \\ &\quad + \frac{t}{2} \int_{-\infty}^{-\frac{a}{\sqrt{2}}} e^{-\frac{x^2}{t}} \operatorname{erfc}\left(x + \frac{a}{\sqrt{2}}\right) dx \end{aligned}$$

and

$$\int_{-\infty}^{-\frac{a}{\sqrt{2}}} \operatorname{erfc}\left(-\frac{(1+t)x+at/\sqrt{2}}{\sqrt{t(t+1)}}\right) dx = \sqrt{\frac{t}{t+1}} \left[-\frac{a}{\sqrt{2t(t+1)}} \operatorname{erfc}\left(\frac{a}{\sqrt{2t(t+1)}}\right) + \frac{1}{\sqrt{\pi}} e^{-\frac{a^2}{2(t+1)}} \right].$$

It remains to evaluate $\int_{-\infty}^{-\frac{a}{\sqrt{2}}} e^{-\frac{x^2}{t}} \operatorname{erfc}\left(x + \frac{a}{\sqrt{2}}\right) dx$. Let (X, Y) be distributed as independent standard normal (i.e. bivariate normal with mean 0, variance 1 and correlation 0). We have

$$\begin{aligned} \int_{-\infty}^{-\frac{a}{\sqrt{2}}} e^{-\frac{x^2}{t}} \operatorname{erfc}\left(x + \frac{a}{\sqrt{2}}\right) dx &= \sqrt{\pi t} \int_{-\infty}^{-a} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \operatorname{erfc}\left(\frac{x+a}{\sqrt{2}}\right) dx \\ &= \sqrt{\pi t} \mathbb{P}(X > a/\sqrt{t}, Y + \sqrt{t}X > a). \end{aligned}$$

It is well known [42, (2.1)] that for (X, Z) a bivariate normal with mean 0, variance 1 and correlation r , $\mathbb{P}(X > h, Z > k) = \frac{1}{2}(\operatorname{erfc}(h/\sqrt{2}) + \operatorname{erfc}(k/\sqrt{2})) - T(h, \frac{k-rh}{h\sqrt{1-r^2}}) - T(k, \frac{h-rk}{k\sqrt{1-r^2}})$, where $T(h, a)$ is defined by (5.10). Specializing to

the case $r = \sqrt{t/(t+1)}$, $h = a/\sqrt{t}$ and $k = a/\sqrt{t+1}$, we get

$$\int_{-\infty}^{-\frac{a}{\sqrt{2}}} e^{-\frac{x^2}{t}} \operatorname{erfc}\left(x + \frac{a}{\sqrt{2}}\right) dx = \frac{\sqrt{\pi t}}{2} \left(\operatorname{erfc}\left(\frac{a}{\sqrt{2t}}\right) + \operatorname{erfc}\left(\frac{a}{\sqrt{2(t+1)}}\right) \right) - \sqrt{\pi t} T\left(\frac{a}{\sqrt{t+1}}, \frac{1}{\sqrt{t}}\right).$$

Combining all the above calculations, we obtain

$$(5.20) \quad \begin{aligned} \text{III} = & a \sqrt{\frac{t^5}{(t+1)^3}} e^{-\frac{a^2}{2(t+1)}} \operatorname{erfc}\left(\frac{a}{\sqrt{2t(t+1)}}\right) + \left(at + \frac{\sqrt{2}t^2}{\sqrt{\pi}(t+1)}\right) e^{-\frac{a^2}{2t}} \\ & + \sqrt{\frac{\pi t^3}{2}} \left(\operatorname{erfc}\left(\frac{a}{\sqrt{2t}}\right) + \operatorname{erfc}\left(\frac{a}{\sqrt{2(t+1)}}\right) \right) - \sqrt{2\pi t^3} T\left(\frac{a}{\sqrt{t+1}}, \frac{1}{\sqrt{t}}\right). \end{aligned}$$

Similarly, we get

$$(5.21) \quad \begin{aligned} \text{IV} = & a \sqrt{\frac{t^5}{(t+1)^3}} e^{-\frac{a^2}{2(t+1)}} \operatorname{erfc}\left(\frac{a(2t+1)}{\sqrt{2t(t+1)}}\right) + at \operatorname{erfc}(\sqrt{2}a) e^{-\frac{a^2}{2t}} - \frac{\sqrt{2}t^2}{\sqrt{\pi}(t+1)} e^{-\frac{a^2(4t+1)}{2t}} \\ & + \sqrt{\frac{\pi t^3}{2}} \left(\operatorname{erfc}\left(\frac{a}{\sqrt{2t}}\right) + \operatorname{erfc}\left(\frac{a}{\sqrt{2(t+1)}}\right) \right) - \sqrt{2\pi t^3} \left(T\left(\frac{a}{\sqrt{t+1}}, \frac{2t+1}{\sqrt{t}}\right) + T\left(\frac{a}{\sqrt{t}}, 2\sqrt{t}\right) \right). \end{aligned}$$

Bringing all the computations together, we obtain $H^a(t) = \frac{1}{\pi t^{3/2}} (\text{I} - \text{II} + \sqrt{\frac{\pi}{2}} \text{III} + \sqrt{\frac{\pi}{2}} \text{IV})$, which yields (5.9). \square

Acknowledgements

We thank G. Schehr for stimulating discussions at the early stage of this work. We also thank two anonymous referees for numerous suggestions which improved the final version of this article.

Funding

Wenpin Tang gratefully acknowledges financial support through an NSF grant DMS-2113779 and through a start-up grant at Columbia University.

References

- [1] E. S. Andersen. On sums of symmetrically dependent random variables. *Skand. Aktuarietidskr.* **36** (1953) 123–138. MR0060173 <https://doi.org/10.1080/03461238.1953.10419466>
- [2] E. S. Andersen. An algebraic treatment of fluctuations of sums of random variables. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*, Vol. II: *Contributions to Probability Theory, Part I* 423–429. Univ. California Press, Berkeley, Calif, 1967. MR0221554
- [3] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja. *A First Course in Order Statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, 1992. MR1178934
- [4] S. Asmussen. *Applied Probability and Queues*, 2nd edition. *Applications of Mathematics* **51**. Springer-Verlag, New York, 2003. MR1978607
- [5] S. Asmussen, P. Glynn and J. Pitman. Discretization error in simulation of one-dimensional reflecting Brownian motion. *Ann. Appl. Probab.* **5** (4) (1995) 875–896. MR1384357
- [6] R. E. Barlow, D. J. Bartholomew, J. M. Bremner and H. D. Brunk. *Statistical Inference Under Order Restrictions. The Theory and Application of Isotonic Regression*. John Wiley & Sons, London-New York-Sydney, 1972. MR0326887
- [7] J. Bertoin. Splitting at the infimum and excursions in half-lines for random walks and Lévy processes. *Stochastic Process. Appl.* **47** (1) (1993) 17–35. MR1232850 [https://doi.org/10.1016/0304-4149\(93\)90092-1](https://doi.org/10.1016/0304-4149(93)90092-1)
- [8] J. Bertoin. *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge University Press, Cambridge, 1996. MR1406564
- [9] J. Bertoin and R. A. Doney. On conditioning a random walk to stay nonnegative. *Ann. Probab.* **22** (4) (1994) 2152–2167. MR1331218
- [10] D. C. Boes and J. D. Salas-La Cruz. On the expected range and expected adjusted range of partial sums of exchangeable random variables. *J. Appl. Probab.* **10** (1973) 671–677. MR0359075 <https://doi.org/10.2307/3212787>
- [11] B. M. Brown. Characteristic functions, moments, and the central limit theorem. *Ann. Math. Stat.* **41** (1970) 658–664. MR0261672 <https://doi.org/10.1214/aoms/1177697109>
- [12] F. Caravenna and L. Chaumont. Invariance principles for random walks conditioned to stay positive. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (1) (2008) 170–190. MR2451576 <https://doi.org/10.1214/07-AIHP119>
- [13] J. T. Chang and Y. Peres. Ladder heights, Gaussian random walks and the Riemann zeta function. *Ann. Probab.* **25** (2) (1997) 787–802. MR1434126 <https://doi.org/10.1214/aop/1024404419>

- [14] A. Chen. *Sampling Error of the Supremum of a Lévy Process*. ProQuest LLC. Thesis (Ph.D.)—University of Illinois at Urbana-Champaign, Ann Arbor, MI, 2011. MR2996014
- [15] F. Comets. *Directed Polymers in Random Environments*. *Lecture Notes in Mathematics* **2175**. Springer, Berlin, 2017. MR3444835 <https://doi.org/10.1007/978-3-319-50487-2>
- [16] A. Dassios. The distribution of the quantile of a Brownian motion with drift and the pricing of related path-dependent options. *Ann. Appl. Probab.* **5** (2) (1995) 389–398. MR1336875
- [17] A. Dassios. Sample quantiles of stochastic processes with stationary and independent increments. *Ann. Appl. Probab.* **6** (3) (1996) 1041–1043. MR1410129 <https://doi.org/10.1214/aoap/1034968241>
- [18] H. A. David and H. N. Nagaraja. *Order Statistics*, 3rd edition. *Wiley Series in Probability and Statistics*. John Wiley & Sons, Hoboken, NJ, 2003. MR1994955 <https://doi.org/10.1002/0471722162>
- [19] I. Denisov. A random walk and a Wiener process near a maximum. *Theory Probab. Appl.* **28** (4) (1984) 821–824. MR0726906
- [20] R. Durrett. *Probability: Theory and Examples*, **49**, 5th edition. Cambridge University Press, Cambridge, 2019. MR3930614 <https://doi.org/10.1017/9781108591034>
- [21] M. Dwass. Simple random walk and rank order statistics. *Ann. Math. Stat.* **38** (1967) 1042–1053. MR0215463 <https://doi.org/10.1214/aoms/1177698773>
- [22] P. Embrechts, C. Klüppelberg and T. Mikosch. *Modelling Extremal Events: For Insurance and Finance*. *Applications of Mathematics (New York)* **33**. Springer-Verlag, Berlin, 1997. MR1458613 <https://doi.org/10.1007/978-3-642-33483-2>
- [23] P. Embrechts, L. C. G. Rogers and M. Yor. A proof of Dassios’ representation of the α -quantile of Brownian motion with drift. *Ann. Appl. Probab.* **5** (3) (1995) 757–767. MR1359828
- [24] W. Feller. *An Introduction to Probability Theory and Its Applications*. Vol. II. John Wiley & Sons, New York-London-Sydney, 1966.
- [25] W. Feller. *An Introduction to Probability Theory and Its Applications*. Vol. I, 3rd edition. John Wiley & Sons, New York-London-Sydney, 1968.
- [26] P. Groeneboom and G. Jongbloed. *Nonparametric Estimation Under Shape Constraints*. *Cambridge Series in Statistical and Probabilistic Mathematics* **38**. Cambridge University Press, Cambridge, 2014. MR3445293 <https://doi.org/10.1017/CBO9781139020893>
- [27] B. M. Hambly, G. Kersting and A. E. Kyprianou. Law of the iterated logarithm for oscillating random walks conditioned to stay non-negative. *Stochastic Process. Appl.* **108** (2) (2003) 327–343. MR2019057 <https://doi.org/10.1016/j.spa.2003.07.002>
- [28] P. L. Hsu. Absolute moments and characteristic functions. *J. Chinese Math. Soc. (N. S.)* **1** (1951) 257–280. MR0072373
- [29] J. Ivanovs. Zooming in on a Lévy process at its supremum. *Ann. Appl. Probab.* **28** (2) (2018) 912–940. MR3784492 <https://doi.org/10.1214/17-AAP1320>
- [30] A. J. E. M. Janssen and J. S. H. van Leeuwen. Cumulants of the maximum of the Gaussian random walk. *Stochastic Process. Appl.* **117** (12) (2007) 1928–1959. MR2437736 <https://doi.org/10.1016/j.spa.2007.03.006>
- [31] A. J. E. M. Janssen and J. S. H. van Leeuwen. On Lerch’s transcendent and the Gaussian random walk. *Ann. Appl. Probab.* **17** (2) (2007) 421–439. MR2308331 <https://doi.org/10.1214/105051606000000781>
- [32] M. Jeanblanc, M. Yor and M. Chesney. *Mathematical Methods for Financial Markets*. *Springer Finance*. Springer-Verlag, London, 2009. MR2568861 <https://doi.org/10.1007/978-1-84628-737-4>
- [33] R. W. Katz, M. B. Parlange and P. Naveau. Statistics of extremes in hydrology. *Adv. Water Resour.* **25** (8–12) (2002) 1287–1304.
- [34] F. B. Knight. Random walks and a sojourn density process of Brownian motion. *Trans. Amer. Math. Soc.* **109** (1963) 56–86. MR0154337 <https://doi.org/10.2307/1993647>
- [35] J. Komlós, P. Major and G. Tusnády. An approximation of partial sums of independent RV’s and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** (1975) 111–131. MR0375412 <https://doi.org/10.1007/BF00533093>
- [36] A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. *Universitext*. Springer-Verlag, Berlin, 2006. MR2250061
- [37] E. L. Lehmann. *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day, Inc., McGraw-Hill International Book Co., New York-Düsseldorf, 1975. MR0395032
- [38] T. Lindvall. *Lectures on the Coupling Method*. John Wiley & Sons, New York, 1992. MR1180522
- [39] N. Moloney, K. Ozogány and Z. Rácz. Order statistics of $1/f^\alpha$ signals. *Phys. Rev. E* **84** (6) (2011), 061101.
- [40] T. V. Narayana. Cyclic permutation of lattice paths and the Chung–Feller theorem. *Skand. Aktuarietidskr.* (1967) 23–30. MR0224486 <https://doi.org/10.1080/03461238.1967.10406204>
- [41] I. Očka. Simple random walk and rank order statistics. *Appl. Math.* **22** (4) (1977) 272–290. MR0438583
- [42] D. B. Owen. Tables for computing bivariate normal probabilities. *Ann. Math. Stat.* **27** (1956) 1075–1090. MR0127562 <https://doi.org/10.1214/aoms/117728074>
- [43] J. W. Pitman. One-dimensional Brownian motion and the three-dimensional Bessel process. *Adv. in Appl. Probab.* **7** (3) (1975) 511–526. MR0375485 <https://doi.org/10.2307/1426125>
- [44] J. W. Pitman and W. Tang. Hidden symmetries and limit laws in the extreme order statistics of the Laplace random walk, 2021. Available at [arXiv:2107.05095](https://arxiv.org/abs/2107.05095).
- [45] F. Pollaczek. Fonctions caractéristiques de certaines répartitions définies au moyen de la notion d’ordre. Application à la théorie des attentes. *C. R. Acad. Sci. Paris Sér. I Math.* **234** (1952) 2334–2336. MR0047945
- [46] S. C. Port. An elementary probability approach to fluctuation theory. *J. Math. Anal. Appl.* **6** (1963) 109–151. MR0145592 [https://doi.org/10.1016/0022-247X\(63\)90097-0](https://doi.org/10.1016/0022-247X(63)90097-0)
- [47] J. Reimann and I. Vincze. On the comparison of two samples with slightly different sizes. *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* **5** (1960) 293–309. MR0132649
- [48] R. D. Reiss. *Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics*. *Springer Series in Statistics*. Springer-Verlag, New York, 1989. MR0988164 <https://doi.org/10.1007/978-1-4613-9620-8>
- [49] M. Renault. Lost (and found) in translation: André’s actual method and its application to the generalized ballot problem. *Amer. Math. Monthly* **115** (4) (2008) 358–363. MR2398417 <https://doi.org/10.1080/00029890.2008.11920537>
- [50] P. Révész. *Random Walk in Random and Non-random Environments*, 3rd edition. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2013. MR3060348 <https://doi.org/10.1142/8678>
- [51] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, 3rd edition. *Grundlehren der Mathematischen Wissenschaften*. **293**. Springer-Verlag, Berlin, 1999. MR1725357 <https://doi.org/10.1007/978-3-662-06400-9>

- [52] G. Schehr and S. N. Majumdar. Universal order statistics of random walks. *Phys. Rev. Lett.* **108** (2012), 040601.
- [53] G. Schehr and S. N. Majumdar. Exact record and order statistics of random walks via first-passage ideas. In *First-Passage Phenomena and Their Applications* 226–251. World Sci. Publ., 2014. MR3363068
- [54] B. W. Silverman. *Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1986. MR0848134 <https://doi.org/10.1007/978-1-4899-3324-9>
- [55] A. V. Skorokhod. *Studies in the Theory of Random Processes*. Addison-Wesley Publishing Co., Reading, MA, 1965. Translated from the Russian by Scripta Technica, Inc.
- [56] M. E. Solari and J. E. A. Dunnage. A note on Spitzer’s lemma and its application to the maximum of partial sums of dependent random variables. *J. Appl. Probab.* **13** (2) (1976) 361–364. MR0400390 <https://doi.org/10.1017/s0021900200094444>
- [57] F. Spitzer. A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82** (1956) 323–339. MR0079851 <https://doi.org/10.2307/1993051>
- [58] H. Tanaka. Time reversal of random walks in one-dimension. *Tokyo J. Math.* **12** (1) (1989) 159–174. MR1001739 <https://doi.org/10.3836/tjm/1270133555>
- [59] S. D. Tremaine and D. O. Richstone. A test of a statistical model for the luminosities of bright cluster galaxies. *Astrophys. J.* **212** (1977) 311–316.
- [60] B. von Bahr. On the convergence of moments in the central limit theorem. *Ann. Math. Stat.* **36** (1965) 808–818. MR0179827 <https://doi.org/10.1214/aoms/1177700055>
- [61] V. Vysotsky. Limit theorems for random walks that avoid bounded sets, with applications to the largest gap problem. *Stochastic Process. Appl.* **125** (5) (2015) 1886–1910. MR3315616 <https://doi.org/10.1016/j.spa.2014.11.017>
- [62] J. G. Wendel. Order statistics of partial sums. *Ann. Math. Stat.* **31** (1960) 1034–1044. MR0119225 <https://doi.org/10.1214/aoms/1177705676>
- [63] D. Williams. Decomposing the Brownian path. *Bull. Amer. Math. Soc.* **76** (1970) 871–873. MR0258130 <https://doi.org/10.1090/S0002-9904-1970-12591-5>
- [64] D. Williams. Path decomposition and continuity of local time for one-dimensional diffusions. I. *Proc. Lond. Math. Soc.* **3** (28) (1974) 738–768. MR0350881 <https://doi.org/10.1112/plms/s3-28.4.738>