Functional limit theorems for the number of occupied boxes in the Bernoulli sieve

Gerold Alsmeyer[a], Alexander Iksanov[b,*], Alexander Marynych[a,b]

[a] *Institute of Mathematical Statistics, Department of Mathematics and Computer Science, University of Münster, Orléans-Ring 10, D-48149 Münster, Germany*
[b] *Faculty of Cybernetics, Taras Shevchenko National University of Kyiv, 01601 Kyiv, Ukraine*

Abstract

The Bernoulli sieve is the infinite Karlin "balls-in-boxes" scheme with random probabilities of stick-breaking type. Assuming that the number of placed balls equals n, we prove several functional limit theorems (FLTs) in the Skorohod space $D[0, 1]$ endowed with the J_1- or M_1-topology for the number $K_n^*(t)$ of boxes containing at most $[n^t]$ balls, $t \in [0, 1]$, and the random distribution function $K_n^*(t)/K_n^*(1)$, as $n \to \infty$. The limit processes for $K_n^*(t)$ are of the form $(X(1) - X((1 - t)-))_{t \in [0,1]}$, where X is either a Brownian motion, a spectrally negative stable Lévy process, or an inverse stable subordinator. The small value probabilities for the stick-breaking factor determine which of the alternatives occurs. If the logarithm of this factor is integrable, the limit process for $K_n^*(t)/K_n^*(1)$ is a Lévy bridge. Our approach relies upon two novel ingredients and particularly enables us to dispense with a Poissonization-de-Poissonization step which has been an essential component in all the previous studies of $K_n^*(1)$. First, for any Karlin occupancy scheme with deterministic probabilities $(p_k)_{k \geq 1}$, we obtain an approximation, uniformly in $t \in [0, 1]$, of the number of boxes with at most $[n^t]$ balls by a counting function defined in terms of $(p_k)_{k \geq 1}$. Second, we prove several FLTs for the number of visits to the interval $[0, nt]$ by a perturbed random walk, as $n \to \infty$. If the stick-breaking factor has a beta distribution with parameters $\theta > 0$ and 1, the process $(K_n^*(t))_{t \in [0,1]}$ has the same distribution as a similar process defined by the number of cycles of length at most $[n^t]$ in a θ-biased random permutation a.k.a. a Ewens permutation with parameter θ. As a consequence, our FLT

* Corresponding author.
 E-mail addresses: gerolda@math.uni-muenster.de (G. Alsmeyer), iksan@univ.kiev.ua (A. Iksanov), marynych@unicyb.kiev.ua (A. Marynych).

with Brownian limit forms a generalization of a FLT obtained earlier in the context of Ewens permutations by DeLaurentis and Pittel (1985), Hansen (1990), Donnelly et al. (1991), and Arratia and Tavaré (1992).

1. Introduction

1.1. The Bernoulli sieve and regenerative compositions

Given a sequence $W_1, W_2, \ldots$ of independent copies of a $(0, 1)$-valued random variable W, consider the random partition of $(0, 1]$ into the subintervals $(V_i, V_{i-1}], i \in \mathbb{N}$, called boxes hereafter, where

$$V_0 := 1, \quad \text{and} \quad V_n := \prod_{i=1}^{n} W_i \quad \text{for } n \in \mathbb{N}.$$

One may also think of the V_i as the successive cut points in a stick-breaking procedure at the end of which a stick with endpoints 0 and 1 is broken into the infinitely many pieces with endpoints V_i and $V_{i-1}, i \in \mathbb{N}$. Next, let $(U_j)_{j \in \mathbb{N}}$ be a sequence of independent uniform $(0, 1)$ random variables which is also independent of $(W_k)_{k \in \mathbb{N}}$. The *Bernoulli sieve* is a random occupancy scheme in which balls, labeled by $1, 2, \ldots$ and of random weights (or random locations) $U_1, U_2, \ldots$, are placed into the boxes in accordance with their weight. Hence, a ball of weight U goes into the box $(V_i, V_{i-1}]$ iff $V_i < U \leq V_{i-1}$. Since its introduction by Gnedin [7], the model has been studied in a series of articles [9–11,13,12,18,19,22]. The term "Bernoulli sieve" can be understood when interpreting the occupancy scheme in terms of a randomized leader-election procedure (see [7] for more details).

Defining the (infinite) vector of occupancy counts

$$Z_{n,i}^* := \#\{1 \leq j \leq n : U_j \in (V_i, V_{i-1}]\}, \quad i \in \mathbb{N}, \tag{1}$$

thus $Z_{n,i}^* \geq 0$ for $i \in \mathbb{N}$ and $\sum_{i \geq 1} Z_{n,i}^* = n$, we see that the $\mathcal{Z}_n^* := (Z_{n,i}^*)_{i \in \mathbb{N}}$ induces a *weak composition* of the integer n. The attribute "weak" is intended to emphasize that zero parts of the composition are allowed. By sequentially allocating the points $U_1, U_2, \ldots$, the sequence $(\mathcal{Z}_n^*)_{n \geq 0}$ is consistently defined for all $n \in \mathbb{N}_0$, and it further has the following two distinguished regenerative properties:

- *Sampling consistency*: If one out of n points is removed uniformly at random from the interval it belongs to, then the resulting weak composition of $n - 1$ has the same law as $\mathcal{Z}_{n-1}^*$.
- *The deletion property*: If the first interval $(V_1, 1]$ contains m points and is removed, then a weak composition of $n - m$ with the same law as $\mathcal{Z}_{n-m}^*$ is obtained.

The class of random compositions generated by a Bernoulli sieve does not cover all regenerative compositions (i.e., those having the two aforementioned properties). In fact, Theorem 5.2 in [14] states that every consistent family of regenerative compositions can be constructed by allocating the points $U_1, U_2, \ldots$ to countably many open intervals forming the complement of the closed range of a multiplicative zero-drift subordinator $(e^{-L_t})_{t \geq 0}$ independent of the uniform

sample. Within this more general framework, weak compositions pertaining to a Bernoulli sieve are those corresponding to a compound Poisson process $(L_t)_{t \geq 0}$.

In the classical occupancy scheme of Karlin [23] balls are placed independently in an infinite array of boxes in accordance with a probability vector $(p_k)_{k \in \mathbb{N}}$, where p_k denotes the probability of choosing box k. The Bernoulli sieve is the Karlin occupancy scheme with random probability

$$p_k^* = V_{k-1} - V_k = W_1 \cdots W_{k-1}(1 - W_k) \tag{2}$$

of choosing box $k \in \mathbb{N}$ and such that, given $(p_k^*)_{k \in \mathbb{N}}$, balls are allocated independently. The Bernoulli sieve can therefore be thought of as an *occupancy scheme in random environment*, the latter being defined by the i.i.d. random variables $W_1, W_2, \ldots$.

For $r = 1, \ldots, n$, denote by $K_{n,r}^* := \sum_{i \geq 1} \mathbf{1}_{\{Z_{n,i}^* = r\}}$ the number of boxes containing exactly r balls, and let $K_n^* := \sum_{r=1}^n K_{n,r}^* = \sum_{i \geq 1} \mathbf{1}_{\{Z_{n,i}^* \geq 1\}}$ denote the number of nonempty boxes. Then define

$$K_n^*(t) := \sum_{r=1}^{[n^t]} K_{n,r}^* = \sum_{i \geq 1} \mathbf{1}_{\{Z_{n,i}^* \in [1, n^t]\}}$$

for $t \in [0, 1]$.

Throughout the rest of the paper, we will use the following notational rule. Quantities related to the Bernoulli sieve are starred, whereas corresponding quantities for the general Karlin scheme are not. The same rule applies, for the most part, to perturbed random walks to be defined in Section 3.

1.2. J_1- and M_1-topology: a brief review

For $T > 0$, let $D[0, T]$ denote the Skorohod space of real-valued functions on $[0, T]$, which are right-continuous with left-hand limits. We will need the J_1- and the M_1-topology on $D[0, T]$ which are commonly used and were introduced in a famous paper by Skorohod [25]. The J_1-topology is generated by the metric

$$d(f, g) := \inf_{\lambda \in \Lambda} \left(\sup_{y \in [0,T]} |f(\lambda(t)) - g(t)| \vee \sup_{t \in [0,T]} |\lambda(t) - t| \right),$$

where Λ denotes the class of strictly increasing, continuous functions $\lambda : [0, T] \rightarrow [0, T]$ with $\lambda(0) = 0$ and $\lambda(T) = T$. Functions f_n which are J_1-convergent to a limit function f are allowed to have a single jump in the vicinity of a jump of f. Furthermore, the positions of the jumps of f_n and their magnitudes should converge to the positions of the jumps of f and their magnitude. This is in contrast to locally uniform convergence which requires the positions of jumps of f_n and f to be the same rather than asymptotically equal.

The M_1-topology is weaker than the J_1-topology and M_1-convergence of f_n to f is equivalent to the convergence of the closed graph of f_n to the closed graph of f. For instance, choosing $f_n(t) := \mathbf{1}_{[1-1/n, 1+1/n)}(t) + 2 \cdot \mathbf{1}_{[1+1/n, 2]}(t)$ and $f(t) := 2 \cdot \mathbf{1}_{[1, 2]}(t)$, the f_n do converge to f in the M_1-topology, but not in the J_1-topology on $D[0, 2]$. Without going into details, we mention that the M_1-topology is typically used in functional limit theorems in which the limit process has jumps unmatched in the convergent sequence of processes. This may happen, for example, if the converging processes are a.s. continuous or have asymptotically vanishing jumps, while the limit process is discontinuous with positive probability. We refer the reader to the monograph [26] for

a comprehensive exposition of the J_1- and the M_1-topologies as well as some other topologies on $D[0, T]$.

Throughout the paper $\overset{J_1}{\Longrightarrow}$ and $\overset{M_1}{\Longrightarrow}$ will mean weak convergence in the Skorohod space when endowed with the J_1-topology and the M_1-topology, respectively. Furthermore, we will use $\overset{\mathbb{P}}{\to}$ and also $\mathbb{P}-\lim$ to denote convergence in probability with respect to $\mathbb{P}$. Finally, $\overset{d}{=}$ will stand for equality in distribution.

1.3. Ewens permutations and Ewens sampling formula

Let $\mathfrak{S}_n$ be the symmetric group of order n. The Ewens family of random permutations is a parametric family $\Pi_n := \Pi_n(\theta), \theta > 0$, of random variables taking values in $\mathfrak{S}_n$ with probabilities

$$\mathbb{P}\{\Pi_n = \sigma\} = \frac{\Gamma(\theta)\theta^{|\sigma|}}{\Gamma(n + \theta)}, \quad \sigma \in \mathfrak{S}_n,$$

where $|\sigma|$ denotes the number of cycles in σ and Γ is the Euler gamma function. Plainly, $\Pi_n(1)$ is a uniform random permutation of $\{1, \ldots, n\}$ for which all $n!$ permutations are equally likely.

For $r = 1, \ldots, n$, denote by $C_{n,r}$ the number of cycles of length r in Π_n. The following is the famous Ewens sampling formula:

$$\mathbb{P}\{C_{n,1} = c_1, \ldots, C_{n,n} = c_n\} = \frac{n!\Gamma(\theta)}{\Gamma(\theta + n)} \prod_{i=1}^{n} \frac{\theta^{c_i}}{i^{c_i} c_i!} \mathbf{1}_{\{\sum_{i=1}^{n} ic_i = n\}}.$$

Define $C_n(t) := \sum_{r=1}^{[n^t]} C_{n,r}$ for $t \in [0, 1]$. A remarkable result, originally due to DeLaurentis and Pittel [5] for the uniform case $\theta = 1$ and to Hansen [17] for the general case $\theta > 0$, asserts that

$$\left(\frac{C_n(t) - \theta t \log n}{\sqrt{\theta \log n}} \right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} (B(t))_{t \in [0,1]}, \quad n \to \infty, \tag{3}$$

where $(B(t))_{t \in [0,1]}$ denotes a standard Brownian motion. Later, much simpler proofs of (3) were found by Donnelly, Kurtz and Tavaré [6] and Arratia and Tavaré [3]. While the first work is based on a Poisson embedding, the second one uses Feller coupling [2, p. 16] as a key tool.

The connection between the Ewens permutations and the Bernoulli sieve emerges when choosing W to have a beta distribution with parameters $\theta > 0$ and 1, i.e. $\mathbb{P}\{W \in \mathrm{d}x\} = \theta x^{\theta-1} \mathbf{1}_{(0,1)}(x)\mathrm{d}x$. In this case (see, for instance, Example 2 in [14] or Section 5.4 in [2])

$$(K_{n,1}^*, \ldots, K_{n,n}^*) \overset{d}{=} (C_{n,1}, \ldots, C_{n,n}), \quad n \in \mathbb{N}$$

which implies that

$$(K_n^*(t))_{t \in [0,1]} \overset{d}{=} (C_n(t))_{t \in [0,1]}. \tag{4}$$

2. Main results and discussion

The asymptotic behavior of the small-parts counts in the Bernoulli sieve is well-understood if $\mathbb{E}(|\log W|) < \infty$. According to Theorem 3.3 in [13], the vector $(K_{n,1}^*, \ldots, K_{n,j}^*)$, with j

fixed, converges in distribution to a similar vector defined in terms of a limiting "balls-in-boxes" scheme in which ball weights are identified with the arrival times of a standard Poisson process on $[0, \infty)$ and boxes are formed by successive points of $\exp(A)$ for a stationary renewal point process A on $\mathbb{R}$ driven by the distribution of $|\log W|$. A criterion for weak convergence of K_n^* is given in Theorem 1.1 and Corollary 1.1 of [10]. One consequence of these results is that, if $\mathbb{E}(|\log W|) < \infty$, the contribution of the small-parts counts to K_n^* becomes asymptotically negligible, as $n \to \infty$, and it raises the natural question which components of the vector $(K_{n,1}^*, \ldots, K_{n,n}^*)$ provide the main contribution to K_n^* for large n. If $\mathbb{E}|\log W| < \infty$, the answer is provided by Proposition 2.1, and for the case $\mathbb{E}|\log W| = \infty$ see (13).

Proposition 2.1. *If* $\mu := \mathbb{E}|\log W| < \infty$, *then*

$$\mathbb{P}-\lim_{n \to \infty} \sup_{t \in [0,1]} \left| \frac{K_n^*(t)}{K_n^*} - t \right| = 0. \tag{5}$$

Thus, if $\mu < \infty$, the random distribution function $t \mapsto K_n^*(t)/K_n^*$ converges uniformly in probability to the uniform distribution function. This provides a definite answer to the question above, namely, for each $t < s$, $t, s \in [0, 1]$, the asymptotic contribution (in probability) of the vector $(K_{n,[nt]}^*, \ldots, K_{n,[ns]}^*)$ to K_n^* is proportional to $s - t$.

Let us compare this observation with some results of a similar flavor from the literature and point out beforehand that $\rho^*(x)$, defined by

$$\rho^*(x) := \#\{k \in \mathbb{N} : p_k^* \geq 1/x\} = \#\{k \in \mathbb{N} : W_1 \cdot \ldots \cdot W_{k-1}(1 - W_k) \geq 1/x\} \tag{6}$$

for $x > 0$, exhibits a logarithmic growth (see Proposition 3.1). Consider now the Karlin occupancy scheme with deterministic or random p_k such that $\rho(x)$, defined by

$$\rho(x) := \#\{k \in \mathbb{N} : p_k \geq 1/x\}, \quad x > 0 \tag{7}$$

is regularly varying at infinity of index α, $\alpha \in (0, 1)$. Let K_n and $K_{n,r}$ denote the number of occupied boxes and the number of boxes containing exactly r balls, respectively, after n balls have been placed. Then $\lim_{n \to \infty} K_{n,r}/K_n = c(r)$ a.s. for explicitly known constants $c(r) > 0$, see Theorems 8 and 9 in [23], Theorem 2.1 in [15] and Corollary 21 in [8] (interesting extensions can be found in [24]). Thus, in sharp contrast to (5), the major contribution to K_n is made by the small-parts counts.

In view of Proposition 2.1, it is natural to ask how fast $K_n^*(t)/K_n^*$ approaches uniformity in $D[0, 1]$. We will answer this question by first proving a FLT for the process $(K_n^*(t))_{t \in [0,1]}$, properly centered and normalized, and then make use of the continuous mapping theorem. Our main results, Theorems 2.2 and 2.5 treat the case of finite and infinite μ, respectively.

Theorem 2.2. *Assume that*

$$\mathbb{E}|\log(1 - W)|^a < \infty \tag{8}$$

for some $a > 0$. *Set*

$$u_n(t) := \mu^{-1} \int_{(1-t)\log n}^{\log n} \mathbb{P}\{|\log(1 - W)| \leq s\} \, \mathrm{d}s$$

and

$$v_n(t) := \mu^{-1} \int_{(1-t)\log n}^{\log n} \mathbb{P}\{|\log(1 - W)| > s\} \, \mathrm{d}s = \mu^{-1} t \log n - u_n(t)$$

for $t \in [0, 1]$, where $\mu = \mathbb{E}|\log W|$.

(A1) If $\sigma^2 := \mathbb{V}\mathrm{ar}|\log W| \in (0, \infty)$, then, as $n \to \infty$,

$$\left(\frac{K_n^*(t) - u_n(t)}{\sqrt{\mu^{-3}\sigma^2 \log n}}\right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} (B(t))_{t \in [0,1]}$$

and

$$\left(\sqrt{\mu\sigma^{-2}\log n}\left(\frac{K_n^*(t)}{K_n^*} - t + \frac{v_n(t) - tv_n(1)}{u_n(1)}\right)\right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} (B(t) - tB(1))_{t \in [0,1]},$$

where $(B(t))_{t \in [0,1]}$ is a standard Brownian motion.

(A2) If $\sigma^2 = \infty$ and

$$\mathbb{E}(\log W)^2 \mathbf{1}_{\{|\log W| \leq x\}} \sim \ell(x), \quad x \to \infty,$$

for some ℓ slowly varying at infinity, then, as $n \to \infty$,

$$\left(\frac{K_n^*(t) - u_n(t)}{\mu^{-3/2}c(\log n)}\right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} (B(t))_{t \in [0,1]}$$

and

$$\left(\frac{\sqrt{\mu}\log n}{c(\log n)}\left(\frac{K_n^*(t)}{K_n^*} - t + \frac{v_n(t) - tv_n(1)}{u_n(1)}\right)\right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} (B(t) - tB(1))_{t \in [0,1]}$$

where c is a positive function satisfying $\lim_{x \to \infty} c(x)^{-2} x\ell(c(x)) = 1$.

(A3) If

$$\mathbb{P}\{|\log W| > x\} \sim x^{-\alpha}\ell(x), \quad x \to \infty, \tag{9}$$

for some $\alpha \in (1, 2)$ and some ℓ slowly varying at infinity, then, as $n \to \infty$,

$$\left(\frac{K_n^*(t) - u_n(t)}{\mu^{-(\alpha+1)/\alpha}c(\log n)}\right)_{t \in [0,1]} \overset{M_1}{\Longrightarrow} (S_\alpha(t))_{t \in [0,1]}$$

and

$$\left(\frac{\mu^{1/\alpha}\log n}{c(\log n)}\left(\frac{K_n^*(t)}{K_n^*} - t + \frac{v_n(t) - tv_n(1)}{u_n(1)}\right)\right)_{t \in [0,1]} \overset{M_1}{\Longrightarrow} (S_\alpha(t) - tS_\alpha(1))_{t \in [0,1]},$$

where c is a positive function satisfying $\lim_{x \to \infty} c(x)^{-\alpha} x\ell(c(x)) = 1$ and $(S_\alpha(t))_{t \in [0,1]}$ is a spectrally negative α-stable Lévy process such that $S_\alpha(1)$ has characteristic function

$$u \mapsto \exp\{-|u|^\alpha \Gamma(1 - \alpha)(\cos(\pi\alpha/2) + i\sin(\pi\alpha/2)\,\mathrm{sgn}(u))\}, \quad u \in \mathbb{R}, \tag{10}$$

with Γ being the gamma function.

Remark 2.3. Along similar lines as in [10] where weak convergence of K_n^* is proved, it can be checked that moment condition (8) is not needed to ensure weak convergence of the *finite-dimensional distributions* of $(K_n^*(t))_{t \in [0,1]}$. Our proof of Theorem 2.2 is based on the decomposition

$$K_n^*(t) - u_n(t) = \left(K_n^*(t) - \left(\rho^*(n) - \rho^*\left(n^{(1-t)-}\right)\right)\right) + \left(\rho^*(n) - \rho^*\left(n^{(1-t)-}\right) - u_n(t)\right) \tag{11}$$

with $\rho^*(x)$ as defined in (6). It will be shown that, irrespective of (8), the first term on the right-hand side of (11), properly normalized, converges to zero uniformly in probability. However, for

dealing with the second term, we need condition (8) (see the proof of Theorem 3.2) but do not know whether it is really necessary.

Remark 2.4. Whenever the second-order term $v_n(t)$ of the centering of $K_n^*(t)$ is killed by the normalizing constants which is the case, for instance, if $\mathbb{E}|\log(1 - W)| < \infty$, the "true" centering for $K_n^*(t)$ is $\mu^{-1}t \log n$. Similarly, the term $\frac{v_n(t)-tv_n(1)}{u_n(1)}$ may then be omitted in the limit theorems for $K_n^*(t)/K_n^*$ because it vanishes asymptotically when multiplied by the corresponding normalization.

Assume now that W has a beta distribution with parameters $\theta > 0$ and 1, giving

$$\mathbb{E}|\log W| = \theta^{-1}, \qquad \mathbb{V}\mathrm{ar}|\log W| = \theta^{-2} \quad \text{and} \quad \mathbb{E}|\log(1 - W)|^a < \infty \quad \text{for all } a > 0.$$

Then part (A1) of Theorem 2.2 together with the preceding remark yields

$$\left(\frac{K_n^*(t) - \theta t \log n}{\sqrt{\theta \log n}} \right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} (B(t))_{t \in [0,1]}, \quad n \to \infty,$$

and in view of (4), this limit relation is equivalent to (3). Thus, we found yet another proof of (3). In fact, our Theorem 2.2 constitutes a generalization of (3).

As shown in [11,13], respectively, similar generalizations exist for the Erdös–Túran law for the order of Ewens permutations [2, Theorem 5.15 on p. 116] and for the weak laws for small cycles in Ewens permutations [2, Theorem 5.1 on p. 96]. On the other hand, the independence-based tools used in the proofs related to these permutations [2, Section 5] are no longer available in the more general framework of the Bernoulli sieve and must therefore be replaced by methods of advanced renewal theory.

Theorem 2.5. *If relation* (9) *holds with* $\alpha \in (0, 1)$, *then, as* $n \to \infty$,

$$\left(\frac{\ell(\log n) K_n^*(t)}{(\log n)^\alpha} \right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} \left(W_\alpha^{\leftarrow}(1) - W_\alpha^{\leftarrow}((1 - t)-) \right)_{t \in [0,1]} \tag{12}$$

and

$$\left(\frac{K_n^*(t)}{K_n^*} \right)_{t \in [0,1]} \overset{J_1}{\Longrightarrow} \left(1 - \frac{W_\alpha^{\leftarrow}((1 - t)-)}{W_\alpha^{\leftarrow}(1)} \right)_{t \in [0,1]}, \tag{13}$$

where $W_\alpha^{\leftarrow}(t) := \inf\{s \geq 0 : W_\alpha(s) > t\}$ *for* $t \geq 0$ *and* $(W_\alpha(s))_{s \geq 0}$ *is an* α-*stable subordinator (nondecreasing Lévy process) with Laplace exponent* $-\log \mathbb{E}(-z W_\alpha(1)) = \Gamma(1 - \alpha)z^\alpha, z \geq 0$.

In the remaining part of this section, we briefly describe our approach and the organization of the paper.

The following heuristic sheds some light on the asymptotic behavior of the number K_n of occupied boxes in the Karlin scheme. Given $(p_j)_{j \in \mathbb{N}}$, call a box k large if $p_k \geq 1/n$. On average, a large box k is occupied because the (conditional) mean number of balls in it is $np_k \geq 1$. One may therefore expect that K_n is asymptotically close to the number of large boxes, which is $\rho(n)$ (see (7) for the definition). For the Bernoulli sieve, this heuristic was justified in [10] by showing that $K_n^* = K_n^*(1)$, properly centered and normalized, converges weakly iff $\rho^*(n)$, defined in (6) and centered and normalized by the same constants, converges weakly to the same law. Arguing similarly, one may expect that $(K_n^* - K_n^*(t))_{t \in [0,1]}$ is well-approximated by $(\rho^*(n^{1-t}))_{t \in [0,1]}$. It will actually be shown in Section 4 that the time-reversal

$$\rho^*(n) - \rho^*(n^{(1-t)-}) = \#\{k \in \mathbb{N} : n^{-1} < p_k^* \leq n^{t-1}\}, \quad t \in [0, 1]$$

provides a uniform approximation for $(K_n^*(t))_{t\in[0,1]}$ which is tight enough to derive that $(K_n^*(t))_{t\in[0,1]}$, properly centered and normalized, converges weakly in the Skorohod space if the same is true for $(\rho^*(n) - \rho^*(n^{(1-t)-}))_{t\in[0,1]}$. This new observation allows us to replace the existing methods based on Poissonization-de-Poissonization used in earlier works on the Bernoulli sieve and constitutes the first principal contribution of the present paper. We stress that our Proposition 4.1 essentially shows that the behavior of $K_n(t) := \sum_{k=1}^{[n^t]} K_{n,r}$ for large n is driven by that of $\rho(n) - \rho(n^{(1-t)-})$ for any Karlin occupancy scheme with deterministic or random $(p_k)_{k\in\mathbb{N}}$ provided that $\rho(x)$ exhibits a logarithmic growth and its increments satisfy an additional condition.

Once a functional limit theorem for $(\rho^*(n^t))_{t\in[0,1]}$ has been proved, the corresponding functional limit theorem for $(\rho^*(n) - \rho^*(n^{(1-t)-}))_{t\in[0,1]}$ follows by an application of the continuous mapping theorem. Put

$$T_n^* := |\log W_1| + \cdots + |\log W_{n-1}| + |\log(1 - W_n)|, \quad n \in \mathbb{N} \tag{14}$$

and note that $\rho^*(n^t) = \#\{k \in \mathbb{N} : T_k^* \leq t \log n\}$ equals the number of visits of the *perturbed random walk* $(T_n^*)_{n\in\mathbb{N}}$ to the interval $[0, t \log n]$. Theorem 3.2 stated in Section 3 provides several functional limit theorems for the number of visits of a general perturbed random walk, not necessarily related to the Bernoulli sieve. Being of independent interest, this result is the second principal contribution of the present paper.

3. Perturbed random walks

Let $(\xi_k, \eta_k)_{k\in\mathbb{N}}$ be a sequence of independent copies of a $\mathbb{R}^2$-valued random vector (ξ, η) with positive components. Set $S_0 := 0$, $S_n := \xi_1 + \cdots + \xi_n$, $n \in \mathbb{N}$ and then

$$T_n := S_{n-1} + \eta_n, \quad n \in \mathbb{N}.$$

The sequence $(T_n)_{n\in\mathbb{N}}$ is called a *perturbed random walk* and has recently attracted some interest in the literature, see [1] and the references therein. Let $N(x)$ denote the number of visits of $(T_n)_{n\in\mathbb{N}}$ to the interval $[0, x]$, i.e.,

$$N(x) := \sum_{k\geq 1} \mathbf{1}_{\{T_k \leq x\}} = \sum_{k\geq 0} \mathbf{1}_{\{S_k + \eta_{k+1} \leq x\}}, \quad x \geq 0.$$

We start with an assertion that will be used in the proof of Proposition 2.1.

Proposition 3.1. *If* $\mathsf{m} := \mathbb{E}\xi < \infty$, *then*

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \left| \frac{\mathsf{m}(N(n) - N(n(1-t)-))}{n} - t \right| = 0 \quad a.s. \tag{15}$$

If $\mathbb{E}\eta < \infty$, the weak convergence of the finite-dimensional distributions of $(N(nt))_{t\geq 0}$, properly centered and normalized, follows from Theorem 2.4 in [21], see also Example 3.2 there. Theorem 3.2 given next is an extension of the aforementioned result to convergence in the Skorohod space. The standard approach to such a strengthening would be to prove tightness. Since this turned out beyond our reach we use an alternative approach.

Theorem 3.2. *Let* $T > 0$ *and* $F(x) = \mathbb{P}\{\eta \leq x\}$, $x \geq 0$. *In* (B1), (B2) *and* (B3) *below, assume further that* $\mathbb{E}\eta^a < \infty$ *for some* $a > 0$.

(B1) *If* $s^2 := \mathbb{V}\mathrm{ar}\,\xi \in (0, \infty)$, *then*

$$\left(\frac{N(nt) - m^{-1} \int_0^{nt} F(u)\, du}{\sqrt{m^{-3}s^2 n}} \right)_{t \in [0,T]} \xRightarrow{J_1} (B(t))_{t \in [0,T]} \tag{16}$$

as $n \to \infty$, *where* $m = \mathbb{E}\xi < \infty$ *and* $(B(t))_{t \in [0,T]}$ *is a standard Brownian motion.*

(B2) *If* $s^2 = \infty$ *and*

$$\mathbb{E}\xi^2 \mathbf{1}_{\{\xi \leq x\}} \sim \ell(x), \quad x \to \infty$$

for some ℓ *slowly varying at infinity, then*

$$\left(\frac{N(nt) - m^{-1} \int_0^{nt} F(u)\, du}{m^{-3/2} c(n)} \right)_{t \in [0,T]} \xRightarrow{J_1} (B(t))_{t \in [0,T]} \tag{17}$$

as $n \to \infty$, *where* c *is a positive function satisfying* $\lim_{x \to \infty} c(x)^{-2} x \ell(c(x)) = 1$.

(B3) *If*

$$\mathbb{P}\{\xi > x\} \sim x^{-\alpha} \ell(x), \quad x \to \infty \tag{18}$$

for some $\alpha \in (1, 2)$ *and some* ℓ *slowly varying at infinity, then*

$$\left(\frac{N(nt) - m^{-1} \int_0^{nt} F(u)\, du}{m^{-(\alpha+1)/\alpha} c(n)} \right)_{t \in [0,T]} \xRightarrow{M_1} (S_\alpha(t))_{t \in [0,T]} \tag{19}$$

as $n \to \infty$, *where* $(S_\alpha(t))_{t \in [0,T]}$ *is an* α-*stable Lévy process,* $S_\alpha(1)$ *has characteristic function* (10), *and* c *is a positive function satisfying* $\lim_{x \to \infty} c(x)^{-\alpha} x \ell(c(x)) = 1$.

(B4) *If* (18) *holds with* $\alpha \in (0, 1)$, *then*

$$\left(\frac{\ell(n) N(nt)}{n^\alpha} \right)_{t \in [0,T]} \xRightarrow{J_1} (W_\alpha^{\leftarrow}(t))_{t \in [0,T]} \tag{20}$$

as $n \to \infty$, *where* $W_\alpha^{\leftarrow}(t)$ *is as defined in* Theorem 2.5.

We close this section with two further results that will be needed in our analysis. The first one tells us that the maximal number of visits of a perturbed random walk to subintervals of $[0, n+b]$ of length b grows stochastically more slowly than any positive power of n.

Proposition 3.3. *For all positive* b *and* c,

$$\mathbb{P}{-}\lim_{n \to \infty} n^{-c} \sup_{t \in [0,1]} \left(N(nt + b) - N(nt) \right) = 0. \tag{21}$$

The second result is a straightforward extension of the fact from renewal theory that the expected number of visits of a random walk with positive increments to intervals of length y is bounded by a linear function in y.

Proposition 3.4. *For all* $x, y \geq 0$ *and appropriate positive* C *and* D,

$$\mathbb{E}\left(N(x + y) - N(x) \right) \leq Cy + D. \tag{22}$$

4. The Karlin occupancy scheme: an approximation result

In this section, we focus on the Karlin occupancy scheme with *deterministic* $(p_k)_{k \in \mathbb{N}}$. For $j = 1, \ldots, n$, let $Z_{n,j}$ denote the number of balls in the jth box, so that

$$K_n(t) := \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} \in [1, n^t]\}}, \quad t \in [0, 1],$$

gives the number of occupied boxes containing at most $[n^t]$ balls. Proposition 4.1 is the first main ingredient to the proof of Theorem 2.2. The connection with the Bernoulli sieve becomes clear when conditioning on (W_k).

Proposition 4.1. *Let $\rho(x)$ be as defined in* (7). *Then*

$$\mathbb{E} \sup_{t \in [0,1]} \left| K_n(t) - \left(\rho(n) - \rho\big(n^{(1-t)-}\big) \right) \right| \leq 6 \left(\rho(n) - \rho\big(x_0^{-1} n (\log n)^{-2}\big) \right)$$

$$+ \frac{3\rho(n)}{\log n} + \int_1^\infty t^{-2} (\rho(nt) - \rho(n)) \, dt + 2 \sup_{t \in [0,1]} \left(\rho(e n^{1-t}) - \rho(e^{-1} n^{1-t}) \right),$$

where $x_0 > 1$ denotes an absolute constant that does not depend on n, nor on $(p_j)_{j \in \mathbb{N}}$.

Proof. Without further notice, all subsequent estimates, including $n^t - n^{3t/4} > 1$ for $t \in [l_n, 1]$ with $l_n := \frac{2 \log \log n}{\log n}$, are meant under the proviso that n be sufficiently large. We start with the basic inequality

$$\left| K_n(t) - \left(\rho(n) - \rho\big(n^{(1-t)-}\big) \right) \right| \leq \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} > n^t, \, np_j \in [1, n^t]\}} + \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} = 0, \, np_j \in [1, n^t]\}}$$

$$+ \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} \in [1, n^t], \, np_j > n^t\}} + \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} \in [1, n^t], \, np_j < 1\}}$$

$$=: S_n^{(1)}(t) + S_n^{(2)}(t) + S_n^{(3)}(t) + S_n^{(4)}(t)$$

and intend to estimate $\mathbb{E} \sup_{t \in [0,1]} S_n^{(i)}(t)$ for $i = 1, 2, 3, 4$.

As for the first summand, we have

$$\sup_{t \in [0,1]} \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} > n^t, \, np_j \in [1, n^t]\}}$$

$$\leq \sup_{t \in [0, l_n)} \sum_{j \geq 1} \mathbf{1}_{\{np_j \in [1, n^t]\}} + \sup_{t \in [l_n, 1]} \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} > n^t, \, np_j \in [1, n^t - n^{3t/4}]\}}$$

$$+ \sup_{t \in [l_n, 1]} \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} > n^t, \, np_j \in (n^t - n^{3t/4}, n^t]\}}$$

$$\leq \sum_{j \geq 1} \mathbf{1}_{\{np_j \in [1, n^{l_n})\}} + \sup_{t \in [l_n, 1]} \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} - np_j > n^{3t/4}, \, np_j \in [1, n^t - n^{3t/4}]\}}$$

$$+ \sup_{t \in [l_n, 1]} \sum_{j \geq 1} \mathbf{1}_{\{np_j \in (n^t - n^{3t/4}, n^t]\}}$$

$$\leq \sum_{j \geq 1} \mathbf{1}_{\{np_j \in [1, \log^2 n)\}} + \sum_{j \geq 1} \mathbf{1}_{\{Z_{n,j} - np_j > (np_j)^{3/4}, \, np_j \in [1, n]\}}$$

$$+ \sup_{t \in [l_n, 1]} \sum_{j \geq 1} \mathbf{1}_{\{np_j \in (n^t - n^{3t/4}, n^t]\}}$$

$$=: S_n^{(11)} + S_n^{(12)} + S_n^{(13)}.$$

By the definition of ρ,

$$S_n^{(11)} = \rho(n) - \rho(n(\log n)^{-2}).$$

The random variable $Z_{n,j}$ has a binomial distribution with parameters n and p_j, so that $\mathbb{E}Z_{n,j} = np_j$ and $\mathbb{V}\mathrm{ar}\, Z_{n,j} = np_j(1 - p_j)$. Use Chebyshev's inequality to obtain

$$\mathbb{E}S_n^{(12)} \leq \sum_{j \geq 1} \frac{np_j(1-p_j)}{(np_j)^{3/2}}\mathbf{1}_{\{np_j \in [1,n]\}} \leq \sum_{j \geq 1}(np_j)^{-1/2}\mathbf{1}_{\{np_j \in [1,n]\}}$$

$$= \int_{[1,n/(\log n)^2]} n^{-1/2}x^{1/2}\,\mathrm{d}\rho(x) + \int_{(n/(\log n)^2,\, n]} n^{-1/2}x^{1/2}\,\mathrm{d}\rho(x)$$

$$\leq \frac{\rho(n)}{\log n} + \left(\rho(n) - \rho\big(n(\log n)^{-2}\big)\right).$$

Finally, the third term $S_n^{(13)}$ can be estimated as follows:

$$S_n^{(13)} = \sup_{t \in [l_n, 1]} \sum_{j \geq 1}\mathbf{1}_{\{p_j \in (n^{t-1} - n^{3t/4-1},\, n^{t-1}]\}}$$

$$\leq \sup_{t \in [l_n, 1]} \sum_{j \geq 1}\mathbf{1}_{\{p_j \in (n^{t-1}(1-n^{-l_n/4}),\, n^{t-1}]\}}$$

$$\leq \sup_{t \in [0,1]} \sum_{j \geq 1}\mathbf{1}_{\{p_j \in (n^{t-1}(1-n^{-l_n/4}),\, n^{t-1}]\}}$$

$$= \sup_{t \in [0,1]} \sum_{j \geq 1}\mathbf{1}_{\{p_j \in (n^{t-1}(1-(\log n)^{-1/2}),\, n^{t-1}]\}}$$

$$\leq \sup_{t \in [0,1]} \left(\rho\big(en^{1-t}\big) - \rho\big(n^{1-t}\big)\right).$$

Summarizing,

$$\mathbb{E}\sup_{t \in [0,1]} S_n^{(1)}(t) \leq 2\left(\rho(n) - \rho\big(n(\log n)^{-2}\big)\right) + \frac{\rho(n)}{\log n} + \sup_{t \in [0,1]}\left(\rho\big(en^{1-t}\big) - \rho\big(n^{1-t}\big)\right).$$

Since $S_n^{(2)}(t)$ and $S_n^{(4)}(t)$ are monotone in t, we further infer more easily that

$$\mathbb{E}\sup_{t \in [0,1]} S_n^{(2)}(t) = \sum_{j \geq 1}\mathbb{P}\{Z_{n,j} = 0\}\mathbf{1}_{\{np_j \in [1,n]\}} = \sum_{j \geq 1}(1 - p_j)^n\mathbf{1}_{\{np_j \in [1,n]\}}$$

$$\leq \sum_{j \geq 1}e^{-p_j n}\mathbf{1}_{\{1/n \leq p_j \leq 1\}}$$

$$= \int_{[1,n/\log n]}e^{-n/x}\,\mathrm{d}\rho(x) + \int_{(n/\log n,\, n]}e^{-n/x}\,\mathrm{d}\rho(x)$$

$$\leq \frac{\rho(n)}{n} + \left(\rho(n) - \rho(n(\log n)^{-1})\right)$$

$$\leq \frac{\rho(n)}{\log n} + \left(\rho(n) - \rho\big(n(\log n)^{-2}\big)\right)$$

and, with the help of Markov's inequality,

$$\mathbb{E}\sup_{t \in [0,1]} S_n^{(4)}(t) = \sum_{j \geq 1}\mathbb{P}\{Z_{n,j} \geq 1\}\mathbf{1}_{\{np_j \leq 1\}} \leq \sum_{j \geq 1}np_j\mathbf{1}_{\{np_j \leq 1\}}$$

$$= \int_{[n, \infty)} nx^{-1}\,\mathrm{d}\rho(x) = \int_{[1, \infty)} x^{-1}\,\mathrm{d}(\rho(nx) - \rho(n)).$$

Upon integration by parts and a use of the fact that $\lim_{x\to\infty} x^{-1}\rho(x) = 0$ (see Lemma 3 in [23]), we arrive at

$$\mathbb{E}\sup_{t\in[0,1]} S_n^{(4)}(t) \leq \int_1^\infty x^{-2}(\rho(nx) - \rho(n))\,\mathrm{d}x + \Big(\rho(n) - \rho(n-)\Big)$$

$$\leq \int_1^\infty x^{-2}(\rho(nx) - \rho(n))\,\mathrm{d}x + \Big(\rho(n) - \rho\big(n(\log n)^{-2}\big)\Big).$$

Left with $S_n^{(3)}(t)$, we write

$$\sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{Z_{n,j}\in[1,n^t],\, np_j > n^t\}}$$

$$\leq \sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{Z_{n,j}\leq n^t,\, np_j - (np_j)^{3/4}\geq n^t\}} + \sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{np_j - (np_j)^{3/4} < n^t \leq np_j\}}$$

$$\leq \sum_{j\geq 1}\mathbf{1}_{\{Z_{n,j}\leq np_j - (np_j)^{3/4},\, np_j\geq 1\}} + \sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{np_j - (np_j)^{3/4} < n^t \leq np_j\}}$$

$$= \sum_{j\geq 1}\mathbf{1}_{\{n - Z_{n,j} - n(1-p_j)\geq (np_j)^{3/4},\, np_j\in[1,n]\}} + \sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{np_j - (np_j)^{3/4} < n^t \leq np_j\}}.$$

Since $n - Z_{n,j}$ has a binomial distribution with parameters n and $1 - p_j$, Chebyshev's inequality along with the estimates used for $\mathbb{E}S_n^{(12)}$ yields

$$\mathbb{E}\sum_{j\geq 1}\mathbf{1}_{\{n - Z_{n,j} - n(1-p_j)\geq (np_j)^{3/4},\, np_j\in[1,n]\}} \leq \frac{\rho(n)}{\log n} + \Big(\rho(n) - \rho\big(n(\log n)^{-2}\big)\Big).$$

To find a proper bound for the second summand, we first verify that

$$x - x^{3/4} < y \leq x \quad\text{and}\quad y \geq 1$$

entail

$$y \leq x < y + (x_0 - 1)y^{3/4}, \tag{23}$$

where $x_0 > 1$ is the unique positive solution to the equation $x - x^{3/4} = 1$. Indeed, if $x < x_0$, then

$$x = (x - x^{3/4}) + x^{3/4} < y + x_0^{3/4} = y + (x_0 - 1) \leq y + (x_0 - 1)y^{3/4},$$

where the last inequality is a consequence of $y \geq 1$. On the other hand, if $x \geq x_0$ we have

$$x = (x - x^{3/4}) + (1 - x^{-1/4})^{-3/4}(x - x^{3/4})^{3/4}$$

$$< y + (1 - x_0^{-1/4})^{-3/4}(x - x^{3/4})^{3/4}$$

$$< y + (1 - x_0^{-1/4})^{-3/4}y^{3/4} = y + (x_0 - 1)y^{3/4},$$

where the first inequality follows from the fact that $x \mapsto (1 - x^{-1/4})^{-3/4}$ is nonincreasing on $(1,\infty)$, giving $(1 - x^{-1/4})^{-3/4} \leq (1 - x_0^{-1/4})^{-3/4}$ for $x \geq x_0$. In view of (23),

$$\sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{np_j - (np_j)^{3/4} < n^t \leq np_j\}} \leq \sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{n^t \leq np_j < n^t + (x_0 - 1)n^{3t/4}\}}$$

$$\leq \sup_{t\in[0,l_n]}\sum_{j\geq 1}\mathbf{1}_{\{n^t \leq np_j < n^t + (x_0 - 1)n^{3t/4}\}} + \sup_{t\in[l_n,1]}\sum_{j\geq 1}\mathbf{1}_{\{n^t \leq np_j < n^t + (x_0 - 1)n^{3t/4}\}}$$

$$\leq \sum_{j\geq 1}\mathbf{1}_{\{1\leq np_j<x_0 n^{l_n}\}} + \sup_{t\in[l_n,1]}\sum_{j\geq 1}\mathbf{1}_{\{p_j\in[n^{t-1},\,n^{t-1}(1+(x_0-1)n^{-t/4}))\}}$$

$$\leq \left(\rho(n) - \rho\big(x_0^{-1}n(\log n)^{-2}\big)\right) + \sup_{t\in[l_n,1]}\sum_{j\geq 1}\mathbf{1}_{\{p_j\in[n^{t-1},\,n^{t-1}(1+(x_0-1)n^{-l_n/4}))\}}$$

$$\leq \left(\rho(n) - \rho\big(x_0^{-1}n(\log n)^{-2}\big)\right) + \sup_{t\in[0,1]}\sum_{j\geq 1}\mathbf{1}_{\{p_j\in[n^{t-1},\,n^{t-1}(1+(x_0-1)(\log n)^{-1/2}))\}}$$

$$\leq \left(\rho(n) - \rho\big(x_0^{-1}n(\log n)^{-2}\big)\right) + \sup_{t\in[0,1]}\left(\rho\big(n^{1-t}\big) - \rho\big(e^{-1}n^{1-t}\big)\right),$$

where $l_n = 2\log\log n/\log n$ should be recalled. A combination of the previous estimates completes the proof of the proposition. $\quad\square$

5. Proofs for Section 3

Proof of Proposition 3.1. If we can prove that

$$\lim_{n\to\infty}\frac{N(ns)}{n} = \lim_{n\to\infty}\frac{N(ns-)}{n} = \frac{s}{\mathsf{m}} \quad\text{a.s.,} \tag{24}$$

for any $s > 0$, then

$$\lim_{n\to\infty}\frac{\mathsf{m}\big(N(n)-N(n(1-t)-)\big)}{n} = t \quad\text{a.s.}$$

for all $t \in [0,1]$, and this yields (15) because, by Dini's theorem, convergence of monotone functions to a continuous limit is uniform on compact sets.

Proof of (24). Since $N(ns)-N(ns-) \leq 1$ for all $n \in \mathbb{N}$ and $s > 0$, it suffices to consider $N(ns)$. Setting

$$\nu(t) := \inf\{k\in\mathbb{N} : S_k > t\} = \sum_{k\geq 0}\mathbf{1}_{\{S_k\leq t\}}, \quad t\in\mathbb{R}, \tag{25}$$

we use the following estimate

$$\frac{\nu(ns-y)}{n} - \frac{1}{n}\sum_{k=1}^{\nu(ns)}\mathbf{1}_{\{\eta_k>y\}} \leq \frac{N(ns)}{n} \leq \frac{\nu(ns)}{n} \quad\text{a.s.} \tag{26}$$

valid for any $y > 0$ and n sufficiently large. The strong law of large numbers provides us with $\lim_{n\to\infty}n^{-1}\sum_{k=1}^{n}\mathbf{1}_{\{\eta_k>y\}} = \mathbb{P}\{\eta > y\}$, while the same law for renewal counting processes [16, Theorem 5.1 on p. 57] gives $\lim_{n\to\infty}n^{-1}\nu(ns) = \mathsf{m}^{-1}s$ a.s. Consequently, $\lim_{n\to\infty}n^{-1}\sum_{k=1}^{\nu(ns)}\mathbf{1}_{\{\eta_k>y\}} = \mathsf{m}^{-1}s\mathbb{P}\{\eta > y\}$ a.s. Finally, (24) follows from (26) by first letting n and then y tend to infinity. $\quad\square$

Proof of Theorem 3.2. (B4) is covered by Corollary 2.6 in [21], see also Example 3.2 there.

Proof of (B1)–(B3). By Theorem 1.1 in [20] applied to $h(t) = F(t) = \mathbb{P}\{\eta \leq t\}$, we know that relations (16), (17) and (19) hold true with $\sum_{k\geq 0}F(nt-S_k)\mathbf{1}_{\{S_k\leq nt\}}$ in place of $N(nt)$. Put

$$X(t) := \sum_{k\geq 0}\big(\mathbf{1}_{\{S_k+\eta_{k+1}\leq t\}} - F(t-S_k)\mathbf{1}_{\{S_k\leq t\}}\big)$$

for $t \geq 0$. In view of the representation

$$N(t) - \frac{1}{m} \int_0^t F(u) \, du = X(t) + \left(\sum_{k \geq 0} F(t - S_k) \mathbf{1}_{\{S_k \leq t\}} - \frac{1}{m} \int_0^t F(u) du \right)$$

and Slutsky's lemma, it suffices to check that

$$\mathbb{P}-\lim_{n \to \infty} n^{-1/2} \sup_{0 \leq t \leq T} |X(nt)| = 0, \tag{27}$$

using also $\lim_{n \to \infty} n^{-1/2} c(n) = \infty$ in the situation of (B2) and (B3) (see Lemma A.1 in the Appendix).

Suppose we can prove that

$$\lim_{t \to \infty} t^{-1/2} X(t) = 0 \quad \text{a.s.} \tag{28}$$

Then, by using

$$n^{-1/2} \sup_{0 \leq t \leq T} |X(nt)| \leq n^{-1/2} \sup_{0 \leq t \leq s} |X(t)| + n^{-1/2} \sup_{s \leq t \leq nT} |X(t)|$$

$$\leq n^{-1/2} \sup_{0 \leq t \leq s} |X(t)| + T^{1/2} \sup_{t \geq s} |t^{-1/2} X(t)|$$

for $0 < s < nT$ and sending first n and then s to infinity, we see that (28) implies (27).

Passing to the proof of (28), we first observe that for each $t \geq 0$, there exists $m \in \mathbb{N}_0$ such that $t \in [m, m+1)$ and

$$t^{-1/2} X(t) \leq m^{-1/2} \sum_{k \geq 0} \left(\mathbf{1}_{\{S_k + \eta_{k+1} \leq m+1\}} - F(m+1 - S_k) \mathbf{1}_{\{S_k \leq m+1\}} \right)$$

$$+ m^{-1/2} \sum_{k \geq 0} \left(F(m+1 - S_k) \mathbf{1}_{\{S_k \leq m+1\}} - F(m - S_k) \mathbf{1}_{\{S_k \leq m\}} \right).$$

Obviously, a lower estimate of similar kind holds as well so that (28) is a consequence of the two limit assertions

$$\lim_{\mathbb{N} \ni m \to \infty} m^{-1/2} X(m) = 0 \quad \text{a.s.} \tag{29}$$

and

$$\lim_{\mathbb{N} \ni m \to \infty} m^{-\delta} \sum_{k \geq 0} \left(F(m+1 - S_k) \mathbf{1}_{\{S_k \leq m+1\}} - F(m - S_k) \mathbf{1}_{\{S_k \leq m\}} \right) = 0 \quad \text{a.s.} \tag{30}$$

for any $\delta > 0$.[1]

Proof of (29). We start by noting that $X(t)$ equals the terminal value of the martingale $(R(k, t), \mathcal{F}_k)_{k \in \mathbb{N}_0}$, where $R(0, t) := 0$, $\mathcal{F}_0 := \{\Omega, \oslash\}$ denotes the trivial σ-algebra,

$$R(k, t) := \sum_{j=0}^{k-1} \left(\mathbf{1}_{\{S_j + \eta_{j+1} \leq t\}} - F(t - S_j) \mathbf{1}_{\{S_j \leq t\}} \right)$$

---

[1] At this point it is enough to prove (30) for $\delta = 1/2$ only. However, we shall need (30) for arbitrary $\delta > 0$ later on.

and $\mathcal{F}_k := \sigma((\xi_j, \eta_j) : 1 \le j \le k)$ for $k \in \mathbb{N}$. For any $l \in \mathbb{N}$, use the Burkholder–Davis–Gundy inequality [4, Theorem 11.3.2] to obtain

$$
\begin{aligned}
\mathbb{E}(X(t))^{2l} \\
&\le C\left(\mathbb{E}\left[\sum_{k \ge 0}\mathbb{E}\big((R(k+1,t) - R(k,t))^2 | \mathcal{F}_k\big)\right]^l + \sum_{k \ge 0}\mathbb{E}\big(R(k+1,t) - R(k,t)\big)^{2l}\right) \\
&= C\left(\mathbb{E}\left[\sum_{k \ge 0} F(t - S_k)(1 - F(t - S_k))\mathbf{1}_{\{S_k \le t\}}\right]^l \right. \\
&\qquad\left. + \sum_{k \ge 0}\mathbb{E}\big(\mathbf{1}_{\{S_k + \eta_{k+1} \le t\}} - F(t - S_k)\mathbf{1}_{\{S_k \le t\}}\big)^{2l}\right) =: C(I_1(t) + I_2(t))
\end{aligned}
$$

for a positive constant C.

Since $\sum_{k \ge 0} F(t - S_k)(1 - F(t - S_k))\mathbf{1}_{\{S_k \le t\}} \le \sum_{k \ge 0}(1 - F(t - S_k))\mathbf{1}_{\{S_k \le t\}}$, we find

$$
I_1(t) = O\left(\left(\int_0^t (1 - F(y))\mathrm{d}y\right)^l\right), \quad t \to \infty
\tag{31}
$$

when applying Lemma A.3 to the nonincreasing function $t \mapsto 1 - F(t)$. Furthermore,

$$
\begin{aligned}
\mathbb{E}&\left(\left(\mathbf{1}_{\{S_k + \eta_{k+1} \le t\}} - F(t - S_k)\mathbf{1}_{\{S_k \le t\}}\right)^{2l}\Big|\mathcal{F}_k\right) \\
&= \big(F(t - S_k)(1 - F(t - S_k))^{2l} + (1 - F(t - S_k))(F(t - S_k))^{2l}\big)\mathbf{1}_{\{S_k \le t\}} \\
&\le (1 - F(t - S_k))\mathbf{1}_{\{S_k \le t\}},
\end{aligned}
$$

giving $I_2(t) \le \mathbb{E}\sum_{k \ge 0}(1 - F(t - S_k))\mathbf{1}_{\{S_k \le t\}}$ and thereupon

$$
I_2(t) = O\left(\int_0^t (1 - F(y))\mathrm{d}y\right), \quad t \to \infty
$$

by Lemma A.3 in the Appendix. We have thus shown that

$$
\mathbb{E}X(t)^{2l} = O\left(\left(\int_0^t (1 - F(y))\,\mathrm{d}y\right)^l\right), \quad t \to \infty,
\tag{32}
$$

and so $\mathbb{E}X(t)^{2l}$ is of order $O(1)$ in the case $\mathbb{E}\eta = \int_0^\infty (1 - F(y))\mathrm{d}y < \infty$. If $\mathbb{E}\eta = \infty$, then our assumption $\mathbb{E}\eta^a < \infty$ for some $a > 0$ entails $a \in (0, 1)$. Using (32) in combination with $\lim_{t \to \infty} t^a(1 - F(t)) = 0$, clearly a consequence of $\mathbb{E}\eta^a < \infty$, yields $\mathbb{E}X(m)^{2l} = O(m^{l(1-a)})$ as $m \to \infty$. Hence, for all $\varepsilon > 0$,

$$
\mathbb{P}\{|X(m)| > \varepsilon m^{1/2}\} \le \frac{\mathbb{E}X(m)^{2l}}{\varepsilon^l m^l} = O(m^{-la}), \quad m \to \infty,
$$

by Markov's inequality. Choosing $l := \min\{j \in \mathbb{N} : ja \ge 2\}$ in the last estimate yields (29) by the Borel–Cantelli lemma.

Proof of (30). Set $J(m) := \int_{[0, m]} \big(F(m + 1 - y) - F(m - y)\big) \, d\nu(y)$ for $m \in \mathbb{N}$, where $\nu(\cdot)$ is defined by (25). We will use the following estimate

$$0 \leq \sum_{k \geq 0} \big(F(m + 1 - S_k) \mathbf{1}_{\{S_k \leq m+1\}} - F(m - S_k) \mathbf{1}_{\{S_k \leq m\}}\big)$$

$$= \int_{[0, m+1]} F(m + 1 - y) \, d\nu(y) - \int_{[0, m]} F(m - y) \, d\nu(y)$$

$$\leq J(m) + \nu(m + 1) - \nu(m).$$

Since $\lim_{m \to \infty} m^{-\delta} (\nu(m + 1) - \nu(m)) = 0$ a.s. by Lemma A.2 in the Appendix, it remains to examine $J(m)$. But

$$J(m) = F(m + 1) - F(m) + \sum_{k=0}^{m-1} \int_{(k, k+1]} \big(F(m + 1 - y) - F(m - y)\big) d\nu(y)$$

$$\leq 1 + \sum_{k=0}^{m-1} \big(F(m + 1 - k) - F(m - 1 - k)\big)(\nu(k + 1) - \nu(k))$$

$$\leq 1 + (F(m) + F(m + 1) - F(1)) \max_{0 \leq k \leq m-1} (\nu(k + 1) - \nu(k)),$$

implies $\lim_{m \to \infty} m^{-\delta} J(m) = 0$ a.s. by Lemma A.2, and this completes the proof of Theorem 3.2. $\quad\square$

Proof of Proposition 3.3. Here it is obviously enough to prove that

$$\lim_{t \to \infty} t^{-c} (N(t + b) - N(t)) = 0 \quad \text{a.s.} \tag{33}$$

for $b > 0$, w.l.o.g. $b = 1$ and $t \to \infty$ along integers only. Put

$$Y(t) := \sum_{j \geq 0} \mathbf{1}_{\{t < S_j + \eta_{j+1} \leq t+1\}} - \sum_{j \geq 0} (F(t + 1 - S_j) \mathbf{1}_{\{S_j \leq t+1\}} - F(t - S_j) \mathbf{1}_{\{S_j \leq t\}}).$$

In view of (30), relation (33) follows if we can show that

$$\lim_{n \to \infty} n^{-c} Y(n) = 0 \quad \text{a.s.}$$

which in turn follows from

$$\mathbb{E} Y(t)^{2l} = O(1), \quad t \to \infty \tag{34}$$

for every $l \in \mathbb{N}$ by a similar argument as in the previous proof using Markov's inequality and the Borel–Cantelli lemma.

Left with (34), the subsequent argument is very similar to the corresponding one for $\mathbb{E} X(t)^{2l}$ in the previous proof. Again, $Y(t)$ is the terminal value of a martingale with respect to the filtration $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$ from there, viz.

$$Y(k, t) := \sum_{j=0}^{k-1} \mathbf{1}_{\{t < S_j + \eta_{j+1} \leq t+1\}} - \sum_{j=0}^{k-1} (F(t + 1 - S_j) \mathbf{1}_{\{S_j \leq t+1\}} - F(t - S_j) \mathbf{1}_{\{S_j \leq t\}}).$$

By another use of the Burkholder–Davis–Gundy inequality, one finds that

$$\mathbb{E}(Y(t))^{2l} \leq C\left(\mathbb{E}\left(\sum_{k\geq 0} p(t-S_k)(1-p(t-S_k))\mathbf{1}_{\{S_k\leq t\}}\right)^l + 1\right)$$

$$\leq C\left(\mathbb{E}\left(\sum_{k\geq 0} p(t-S_k)\mathbf{1}_{\{S_k\leq t\}}\right)^l + 1\right)$$

for a positive constant C and with $p(x) := F(x+1) - F(x)$. By Lemma A.3,

$$\mathbb{E}\left(\sum_{k\geq 0} p(t-S_k)\mathbf{1}_{\{S_k\leq t\}}\right)^l = O\left(\sum_{j=0}^{[t]} \sup_{y\in[j,\,j+1)} (F(y+1)-F(y))\right)^l = O(1)$$

because

$$\left(\sum_{j=0}^{[t]} \sup_{y\in[j,\,j+1)} (F(y+1)-F(y))\right)^l \leq \left(\sum_{j=0}^{[t]} (F(j+2)-F(j))\right)^l \leq 2^l.$$

This establishes (34) thereby finishing the proof of Proposition 3.3. □

Proof of Proposition 3.4. Note that

$$U(x) := \mathbb{E}\nu(x) = \sum_{j\geq 1} \mathbb{P}\{S_{j-1} \leq x\}, \quad x \in \mathbb{R},$$

is the renewal function associated with $(S_n)_{n\in\mathbb{N}}$. As a consequence of the distributional subadditivity of ν (see formula (5.7) on p. 58 in [16]), the monotonicity of ν, and the fact that $\nu(x) = 0$ for $x < 0$, U is subadditive on $\mathbb{R}$, i.e. $U(x+y) \leq U(x)+U(y)$ for all $x, y \in \mathbb{R}$, and so

$$U(x) \leq Cx^+ + C_2$$

for all $x \in \mathbb{R}$ and some positive C and D. By using these facts, we finally obtain

$$\mathbb{E}\big(N(x+y) - N(x)\big) = \sum_{j\geq 1} \mathbb{P}\{x < T_j \leq x+y\}$$

$$= \sum_{j\geq 1} \int_{[0,\infty)} \mathbb{P}\{x-z < S_{j-1} \leq x+y-z\}\,\mathbb{P}\{\eta \in \mathrm{d}z\}$$

$$= \int_{[0,\infty)} \big(U(x+y-z) - U(x-z)\big)\,\mathbb{P}\{\eta \in \mathrm{d}z\}$$

$$\leq U(y) \leq Cy + D.$$

for all $x, y \geq 0$. □

6. Proofs for Section 2

Recall that the Bernoulli sieve is the Karlin occupancy scheme with the random probabilities $(p_k^*)_{k\in\mathbb{N}}$ defined in (2). Condition on $(p_k^*)_{k\in\mathbb{N}}$ and apply Proposition 4.1 to obtain

$$\mathbb{E}\left(\sup_{t\in[0,1]} |K_n^*(t) - (\rho^*(n) - \rho^*(n^{(1-t)-}))|\,\Big|(p_j^*)\right)$$

$$\leq 6\left(\rho^*(n) - \rho^*\big(x_0^{-1}n(\log n)^{-2}\big)\right) + \frac{3\rho^*(n)}{\log n}$$

$$+ \int_1^\infty \frac{\rho^*(nx) - \rho^*(n)}{x^2} \, dx$$

$$+ 2 \sup_{t \in [0,1]} \left(\rho^*(en^{1-t}) - \rho^*(e^{-1}n^{1-t}) \right) =: \varepsilon_n \tag{35}$$

with $\rho^*(x)$ defined in (6). The next lemma shows that ε_n does not grow faster in probability than any power of $\log n$.

Lemma 6.1. *For all $c > 0$,*

$$\mathbb{P}\text{-}\lim_{n \to \infty} \frac{\varepsilon_n}{(\log n)^c} = 0.$$

Proof. Observe that

$$\rho^*(x) = N(\log x) \tag{36}$$

for $N(\cdot)$ corresponding to $\xi = |\log W|$ and $\eta = |\log(1 - W)|$. Using (22), we infer that the expectation of the first three terms of (35) is $O(\log \log n)$. Hence the sum of these terms divided by $(\log n)^c$ converges to zero in probability by Markov's inequality. Finally, the fourth term

$$\sup_{t \in [0,1]} \left(\rho^*(en^{1-t}) - \rho^*(e^{-1}n^{1-t}) \right) = \sup_{t \in [0,1]} \left(N(t \log n + 1) - N(t \log n - 1) \right)$$

is of order $o((\log n)^c)$ in probability for any $c > 0$ by Proposition 3.3.   □

Proof of Proposition 2.1. By Lemma 6.1,

$$\mathbb{P}\text{-}\lim_{n \to \infty} \frac{1}{\log n} \mathbb{E}\left(\sup_{t \in [0,1]} |K_n^*(t) - (\rho^*(n) - \rho^*(n^{(1-t)-}))| \, \Big| \, (p_j^*) \right) = 0$$

which implies

$$\mathbb{P}\text{-}\lim_{n \to \infty} \frac{1}{\log n} \sup_{t \in [0,1]} \left| K_n^*(t) - (\rho^*(n) - \rho^*(n^{(1-t)-})) \right| = 0 \tag{37}$$

by Markov's inequality and the dominated convergence theorem. Furthermore,

$$\left| \frac{K_n^*}{\mu^{-1} \log n} - 1 \right| \leq \frac{|K_n^* - \rho^*(n)|}{\mu^{-1} \log n} + \left| \frac{\rho^*(n)}{\mu^{-1} \log n} - 1 \right| \quad \text{a.s.}$$

The first term on the right-hand side converges to zero in probability by (37) (recall that $K_n^* = K_n^*(1)$), and the second does so by (36) and Proposition 3.1. Hence,[2]

$$\mathbb{P}\text{-}\lim_{n \to \infty} \frac{K_n^*}{\log n} = \frac{1}{\mu}. \tag{38}$$

---

[2] Though not needed here, let us note that the convergence in (38) holds a.s. Just apply the Borel–Cantelli lemma to $K^*_{[\exp(n^2)]}$ and then use the monotonicity of K_n^* in n. Alternatively, this follows from Theorem 1′ and 8 in [23] in combination with $\lim_{n \to \infty} (\log n)^{-1} \rho^*(n) = \mu^{-1}$ a.s. provided by Proposition 3.1.

To complete the proof, we observe that

$$\sup_{t\in[0,1]}\left|\frac{K_n^*(t)}{K_n^*}-t\right|=\frac{\mu^{-1}\log n}{K_n^*}\left(\frac{\sup\limits_{t\in[0,1]}|K_n^*(t)-tK_n^*|}{\mu^{-1}\log n}\right)$$

$$\leq\frac{\mu^{-1}\log n}{K_n^*}\left(\frac{\sup\limits_{t\in[0,1]}|K_n^*(t)-(\rho^*(n)-\rho^*(n^{(1-t)-}))|}{\mu^{-1}\log n}\right.$$

$$\left.+\sup_{t\in[0,1]}\left|\frac{\mu(\rho^*(n)-\rho^*(n^{(1-t)-}))}{\log n}-t\right|+\left|\frac{K_n^*}{\mu^{-1}\log n}-1\right|\right)$$

and note that the right-hand side converges to zero in probability by (37), (38) and Proposition 3.1. □

Proof of Theorem 2.2. *Proof for $K_n^*(t)$.* Denote by a_n the normalization used for $K_n^*(t)$ in the respective parts of Theorem 2.2. Observe that a_n grows faster than some power of the logarithm. Hence, by (35), Lemma 6.1 and Markov's inequality,

$$\mathbb{P}-\lim_{n\to\infty}\mathbb{P}\left\{\sup_{t\in[0,1]}|K_n^*(t)-(\rho^*(n)-\rho^*(n^{(1-t)-}))|>\varepsilon a_n\,\Big|(p_j^*)\right\}=0$$

for all $\epsilon>0$. Using the dominated convergence theorem, this yields

$$\mathbb{P}-\lim_{n\to\infty}a_n^{-1}\sup_{t\in[0,1]}|K_n^*(t)-(\rho^*(n)-\rho^*(n^{(1-t)-}))|=0.$$

In view of representation (11), it remains to prove Theorem 2.2 with $\rho^*(n)-\rho^*(n^{(1-t)-})$ replacing $K_n^*(t)$. Using (36), we see that this is accomplished by an application of Theorem 3.2 to the process $(\rho^*(n^t))_{t\in[0,1]}$ and the subsequent use of the continuous mapping theorem. For the latter, three supporting facts are:

(1) in the J_1- (M_1-) topology on $D[0,1]\times D[0,1]$

$$\left(\frac{\rho^*(n)-u_n(1)}{a_n},\frac{-\rho^*(n^{(1-t)-})+u_n(1)-u_n(t)}{a_n}\right)\Longrightarrow\big(X(1),-X((1-t)-)\big),$$

as $n\to\infty$, where X is a Brownian motion and the convergence is in the J_1-topology in (C1) and (C2), whereas X is an α-stable Lévy process and the convergence is in the M_1-topology in (C3);

(2) if $D[0,1]\times D[0,1]$ is equipped with the J_1 or M_1-topology (which is stronger than the product topology), then the mapping $\psi:D[0,1]\times D[0,1]\to D[0,1]$, defined by $\psi(x,y):=x+y$ is continuous;

(3) $(Y(1)-Y((1-t)-))_{t\in[0,1]}\overset{d}{=}(Y(t))_{t\in[0,1]}$ for any Lévy process Y.

Proof for $K_n^*(t)/K_n^*$. Write

$$\frac{\log n}{\mu a_n}\left(\frac{K_n^*(t)}{K_n^*}-t+\frac{v_n(t)-tv_n(1)}{u_n(1)}\right)=\frac{\log n}{\mu K_n^*}\left(\frac{K_n^*(t)-tK_n^*+v_n(t)-tv_n(1)}{a_n}\right)$$

$$+\frac{\log n}{\mu K_n^*}\frac{v_n(t)-tv_n(1)}{u_n(1)}\frac{K_n^*-u_n(1)}{a_n}\tag{39}$$

with a_n as before. By what has already been proved, we know that

$$\left(\frac{K_n^*(t) - u_n(t)}{a_n}\right)_{t \in [0,1]} \quad \Longrightarrow \quad (Z(t))_{t \in [0,1]}, \quad n \to \infty \tag{40}$$

with some Lévy process Z depending on the respective case. Use the continuous mapping theorem along with continuity of the summation mapping and (38) to obtain

$$\frac{\log n}{\mu K_n^*}\left(\frac{K_n^*(t) - t K_n^* + v_n(t) - t v_n(1)}{a_n}\right) \quad \Longrightarrow \quad (Z(t) - t Z(1))_{t \in [0,1]}, \quad n \to \infty.$$

In view of (38), (40) and

$$\sup_{t \in [0,1]}\left|\frac{v_n(t) - t v_n(1)}{u_n(1)}\right| \le \frac{2 v_n(1)}{u_n(1)} \to 0, \quad n \to \infty,$$

the second term in (39) converges to zero in probability uniformly in $t \in [0, 1]$. The proof is completed by an appeal to Slutsky's lemma. □

Proof of Theorem 2.5. As for $K_n^*(t)$, the argument used in the proof of Theorem 2.2 applies here without changes, and (12) follows. Passing to the proof for $K_n^*(t)/K_n^*$, we immediately conclude that (12) entails

$$\left(\frac{\ell(\log n) K_n^*(t)}{(\log n)^\alpha}, \frac{(\log n)^\alpha}{\ell(\log n) K_n^*}\right) \quad \Longrightarrow \quad \left(W_\alpha^\leftarrow(1) - W_\alpha^\leftarrow((1-t)-), \frac{1}{W_\alpha^\leftarrow(1)}\right), \quad n \to \infty$$

in the J_1-topology on $D[0,1] \times D[0,1]$. To complete the proof, use once again the continuous mapping theorem together with the following fact: if $D[0,1] \times D[0,1]$ is equipped with the J_1-topology, then the mapping $\phi : D[0,1] \times D[0,1] \to D[0,1]$ defined by $\phi(x, y) := xy$ is continuous. □

Acknowledgments

Appendix

We collect three auxiliary results that have been used in the proofs of the main results.

Lemma A.1. *Let $c(x)$ be a function as defined in* (B2) *or* (B3) *of* Theorem 3.2. *Then*

$$\lim_{x \to \infty} x^{-1/2} c(x) = \infty.$$

Proof. For (B3), this is immediate because $\lim_{x \to \infty} c^{-\alpha}(x) x \ell(c(x)) = 1$ implies that c varies regularly with index $1/\alpha > 1/2$. Suppose now the assumptions of part (B2) be valid. Since $c(x)$ is the asymptotic inverse of $x \mapsto x^2/\ell(x)$ and $\lim_{x \to \infty} x^2/\ell(x) = \infty$, we infer that $\lim_{x \to \infty} c(x) = \infty$. Moreover, $\lim_{x \to \infty} \ell(x) = \infty$ in view of $\mathsf{s}^2 = \infty$. Thus, $\lim_{x \to \infty} \ell(c(x)) = \infty$ which in combination with $x^{-1} c^2(x) \sim \ell(c(x)), x \to \infty$, entails $\lim_{x \to \infty} x^{-1/2} c(x) = \infty$. □

Let $(S_n)_{n \in \mathbb{N}_0}$ be a zero-delayed standard random walk with positive steps as in Section 3. Recall the notation $\nu(t) = \inf\{k \in \mathbb{N} : S_k > t\}$ for $t \in \mathbb{R}$. Plainly, $\nu(t) \equiv 0$ for $t \leq 0$.

Lemma A.2. *For all $\delta > 0$,*

$$\lim_{\mathbb{N} \ni n \to \infty} n^{-\delta} \max_{0 \leq j \leq n} (\nu(j+1) - \nu(j)) = 0 \quad a.s. \tag{41}$$

Proof. Fix $a > 0$. Since $\lim_{\gamma \to \infty} \mathbb{E}e^{-\gamma S_1} = \mathbb{P}\{S_1 = 0\} = 0$, we can pick $\gamma > 0$ such that $e^a \mathbb{E}e^{-\gamma S_1} < 1$. Then

$$e^a \frac{\mathbb{E}e^{a\nu(1)} - 1}{e^a - 1} = \sum_{j \geq 1} e^{aj} \mathbb{P}\{\nu(1) \geq j\} = \sum_{j \geq 1} e^{aj} \mathbb{P}\{S_{j-1} \leq 1\}$$

$$\leq e^\gamma \sum_{j \geq 1} e^{aj} \left(\mathbb{E}e^{-\gamma S_1}\right)^{j-1} < \infty \tag{42}$$

by Markov's inequality. Hence, for all $\varepsilon > 0$,

$$\mathbb{P}\{\nu(k+1) - \nu(k) > \varepsilon k^\delta\} \leq \mathbb{P}\{\nu(1) > \varepsilon k^\delta\} \leq e^{-a\varepsilon k^\delta} \mathbb{E}e^{a\nu(1)}$$

having utilized the distributional subadditivity of $\nu(t)$ [16, formula (5.7) on p. 58] for the first inequality and Markov's inequality for the second. By the Borel–Cantelli lemma, we conclude that $\lim_{n \to \infty} n^{-\delta}(\nu(n+1) - \nu(n)) = 0$ a.s. and thereupon (41). $\square$

Lemma A.3. *Let $G : [0, \infty) \to [0, \infty)$ be a locally bounded function, $l \in \mathbb{N}$ be an arbitrary integer, and $[t]$ denote the integer-part of t. Then, as $t \to \infty$,*

$$\mathbb{E}\left(\sum_{k \geq 0} G(t - S_k) \mathbf{1}_{\{S_k \leq t\}}\right)^l = O\left(\left(\sum_{j=0}^{[t]} \sup_{y \in [j, j+1)} G(y)\right)^l\right). \tag{43}$$

If G is nonincreasing, then

$$\mathbb{E}\left(\sum_{k \geq 0} G(t - S_k) \mathbf{1}_{\{S_k \leq t\}}\right)^l = O\left(\left(\int_0^t G(y)\,\mathrm{d}y\right)^l\right) \tag{44}$$

as $t \to \infty$.

Proof. The result is trivial if $G \equiv 0$. Assuming that G is not identically zero, we have

$$G(t) \leq \sum_{n=0}^{[t]} \left(\sup_{y \in [n, n+1)} G(y)\right) \mathbf{1}_{[n, n+1)}(t), \quad t \geq 0.$$

Hence, for t so large that the right-hand side is positive, we infer

$$\left(\sum_{k \geq 0} (G(t - S_k)) \mathbf{1}_{\{S_k \leq t\}}\right)^l \leq \left(\sum_{n=0}^{[t]} \left(\sup_{y \in [n, n+1)} G(y)\right) (\nu(t - n) - \nu(t - (n+1)))\right)^l$$

$$\leq \left(\sum_{k=0}^{[t]} \sup_{y \in [k, k+1)} G(y)\right)^l \sum_{n=0}^{[t]} \frac{\sup_{y \in [n, n+1)} G(y)}{\sum_{k=0}^{[t]} \sup_{y \in [k, k+1)} G(y)} (\nu(t - n) - \nu(t - (n+1)))^l, \tag{45}$$

having utilized $\sum_{j\geq 0} \mathbf{1}_{\{a < S_j \leq b\}} = \nu(b) - \nu(a)$ a.s. for $a < b$ and the convexity of $x \mapsto x^l$. It is clear that

$$\nu(t - [t]) - \nu(t - [t] - 1) = \nu(t - [t]) \leq \nu(1) \quad \text{a.s.}$$

Moreover, $\mathbb{P}\{\nu(t - n) - \nu(t - (n+1)) > x\} \leq \mathbb{P}\{\nu(1) > x\}$ for $n = 0, \ldots, [t] - 1$ and $x \geq 0$ by the distributional subadditivity of $\nu(t)$ mentioned earlier. Consequently,

$$\mathbb{E}(\nu(t - n) - \nu(t - (n+1)))^l \leq \mathbb{E}(\nu(1))^l < \infty \quad \text{for } n = 0, \ldots, [t], \tag{46}$$

where the finiteness follows from (42).

Passing to expectations in (45) and using (46), we see that

$$\mathbb{E}\left(\sum_{j\geq 0} G(t - S_j)\mathbf{1}_{\{S_j \leq t\}}\right)^l \leq \left(\sum_{n=0}^{[t]} \sup_{y \in [n,\, n+1)} G(y)\right)^l \mathbb{E}\nu(1)^l$$

which proves (43). If G is nonincreasing, then

$$\sum_{n=0}^{[t]} \sup_{y \in [n,\, n+1)} G(y) = \sum_{n=0}^{[t]} G(n) = \int_0^t G(y)\mathrm{d}y + O(1), \quad t \to \infty,$$

and (44) follows. $\quad\square$

References

[1] G. Alsmeyer, A. Iksanov, M. Meiners, Power and exponential moments of the number of visits and related quantities for perturbed random walks, J. Theoret. Probab. 28 (2015) 1–40.

[2] R. Arratia, A.D. Barbour, S. Tavaré, Logarithmic Combinatorial Structures: A Probabilistic Approach, in: EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2003.

[3] R. Arratia, S. Tavaré, Limit theorems for combinatorial structures via discrete process approximations, Random Structures Algorithms 3 (1992) 321–345.

[4] Y.S. Chow, H. Teicher, Probability Theory. Independence, Interchangeability, Martingales, third ed., in: Springer Texts in Statistics, Springer-Verlag, New York, 1997.

[5] J.M. DeLaurentis, B.G. Pittel, Random permutations and Brownian motion, Pacific J. Math. 119 (1985) 287–301.

[6] P. Donnelly, T.G. Kurtz, S. Tavaré, On the functional central limit theorem for the Ewens sampling formula, Ann. Appl. Probab. 1 (1991) 539–545.

[7] A. Gnedin, The Bernoulli sieve, Bernoulli 10 (2004) 79–96.

[8] A. Gnedin, B. Hansen, J. Pitman, Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws, Probab. Surv. 4 (2007) 146–171.

[9] A. Gnedin, A. Iksanov, A. Marynych, The Bernoulli sieve: an overview, in: 21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms, AofA'10, in: Discrete Math. Theor. Comput. Sci. Proc., AM, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2010, pp. 329–341.

[10] A. Gnedin, A. Iksanov, A. Marynych, Limit theorems for the number of occupied boxes in the Bernoulli sieve, Theory Stoch. Process. 16 (32) (2010) 44–57.

[11] A. Gnedin, A. Iksanov, A. Marynych, A generalization of the Erdös–Turán law for the order of random permutation, Combin. Probab. Comput. 21 (2012) 715–733.

[12] A. Gnedin, A.M. Iksanov, P. Negadajlov, U. Rösler, The Bernoulli sieve revisited, Ann. Appl. Probab. 19 (2009) 1634–1655.

[13] A. Gnedin, A. Iksanov, U. Roesler, Small parts in the Bernoulli sieve, in: Fifth Colloquium on Mathematics and Computer Science, in: Discrete Math. Theor. Comput. Sci. Proc., AI, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008, pp. 235–242.

[14] A. Gnedin, J. Pitman, Regenerative composition structures, Ann. Probab. 33 (2005) 445–479.

[15] A. Gnedin, J. Pitman, M. Yor, Asymptotic laws for compositions derived from transformed subordinators, Ann. Probab. 34 (2006) 468–492.

[16] A. Gut, Stopped Random Walks. Limit Theorems and Applications, second ed., in: Springer Series in Operations Research and Financial Engineering, Springer, New York, 2009.

[17] J.C. Hansen, A functional central limit theorem for the Ewens sampling formula, J. Appl. Probab. 27 (1990) 28–43.
[18] A. Iksanov, On the number of empty boxes in the Bernoulli sieve II, Stochastic Process. Appl. 122 (2012) 2701–2729.
[19] A. Iksanov, On the number of empty boxes in the Bernoulli sieve I, Stochastics 85 (2013) 946–959.
[20] A. Iksanov, Functional limit theorems for renewal shot noise processes with increasing response functions, Stochastic Process. Appl. 123 (2013) 1987–2010.
[21] A. Iksanov, A. Marynych, M. Meiners, Asymptotics of random processes with immigration I: Scaling limits, Bernoulli (2016) in press.
[22] A. Iksanov, A. Marynych, V.A. Vatutin, Weak convergence of finite-dimensional distributions of the number of empty boxes in the Bernoulli sieve, Theory Probab. Appl. 59 (2015) 87–113.
[23] S. Karlin, Central limit theorems for certain infinite urn schemes, J. Math. Mech. 17 (1967) 373–401.
[24] J. Schweinsberg, The number of small blocks in exchangeable random partitions, ALEA Lat. Am. J. Probab. Math. Stat. 7 (2010) 217–242.
[25] A.V. Skorohod, Limit theorems for stochastic processes, Theory Probab. Appl. 1 (1956) 261–290.
[26] W. Whitt, Stochastic-process Limits: An Introduction to Stochastic-process Limits and Their Application to Queues, in: Springer Series in Operations Research, Springer-Verlag, New York, 2002.