

A law of the iterated logarithm for small counts in Karlin’s occupancy scheme

Alexander Iksanov^{a,*}, Valeriya Kotelnikova^a

^a*Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Ukraine*

iksan@univ.kiev.ua (A. Iksanov), valeria.kotelnikova@unicyb.kiev.ua (V. Kotelnikova)

Received: 19 November 2023, Revised: 18 January 2024, Accepted: 18 January 2024,
Published online: 30 January 2024

Abstract In the Karlin infinite occupancy scheme, balls are thrown independently into an infinite array of boxes $1, 2, \dots$, with probability p_k of hitting the box k . For $j, n \in \mathbb{N}$, denote by $\mathcal{K}_j^*(n)$ the number of boxes containing exactly j balls provided that n balls have been thrown. *Small counts* are the variables $\mathcal{K}_j^*(n)$, with j fixed. The main result is a law of the iterated logarithm (LIL) for the small counts as the number of balls thrown becomes large. Its proof exploits a Poissonization technique and is based on a new LIL for infinite sums of independent indicators $\sum_{k \geq 1} \mathbb{1}_{A_k(t)}$ as $t \rightarrow \infty$, where the family of events $(A_k(t))_{t \geq 0}$ is not necessarily monotone in t . The latter LIL is an extension of a LIL obtained recently by Buraczewski, Iksanov and Kotelnikova (2023+) in the situation when $(A_k(t))_{t \geq 0}$ forms a nondecreasing family of events.

Keywords Independent indicators, infinite occupancy, law of the iterated logarithm, small counts

MSC2020 MSC 60F15, 60G50, 60C05

1 Introduction

1.1 Definition of the model

Let $(p_k)_{k \in \mathbb{N}}$ be a discrete probability distribution with $p_k > 0$ for infinitely many k . The *infinite occupancy scheme* is defined by independent allocation of balls over an

*Corresponding author.

infinite array boxes $1, 2, \dots$, with probability p_k of hitting the box k . The scheme is usually called the *Karlin occupancy scheme* because of Karlin’s remarkable work [12]. We are aware of the two articles [1] and [5] which preceded [12]. A survey of the literature on the infinite occupancy up to 2007 is given in [9]. An incomplete list of very recent contributions includes [3, 4, 6, 7]. Among other things, the authors of [9] discuss applications of the scheme to ecology, database query optimization and literature. Another portion of possible applications can be found in Section 1.1 of [10].

There are deterministic and Poissonized versions of Karlin’s occupancy scheme. In a *deterministic version* the n th ball is thrown at time $n \in \mathbb{N}$. For $j, n \in \mathbb{N}$, denote by $\mathcal{K}_j(n)$ and $\mathcal{K}_j^*(n)$ the number of boxes hit by at least j balls and exactly j balls, respectively, up to and including time n . Observe that $\mathcal{K}_1(n)$ is the number of occupied boxes at time n . Sometimes the variables $\mathcal{K}_j^*(n)$, with j fixed, are referred to as *small counts*.

To define the other version of the scheme we need an additional notation. Let $(S_k)_{k \in \mathbb{N}}$ denote a random walk with independent jumps having an exponential distribution of unit mean. The counting process $\pi := (\pi(t))_{t \geq 0}$ given by $\pi(t) := \#\{k \in \mathbb{N} : S_k \leq t\}$ for $t \geq 0$ is a Poisson process on $[0, \infty)$ of unit intensity.

In a *Poissonized version* of Karlin’s occupancy scheme the n th ball is thrown at time $S_n, n \in \mathbb{N}$, and it is assumed that the allocation process is independent of $(S_k)_{k \in \mathbb{N}}$, hence of π . Thus, in the time interval $[0, t]$ there are $\pi(t)$ balls thrown in the Poissonized version and $\lfloor t \rfloor$ balls thrown in the deterministic version. While the occupancy counts of distinct boxes are dependent in the deterministic version, these are independent in the Poissonized version. The latter fact is a principal advantage of the Poissonized version. It is justified by the thinning property of Poisson processes. For $j \in \mathbb{N}$ and $t \geq 0$, denote by $K_j(t)$ and $K_j^*(t)$ the number of boxes containing at least j balls and exactly j balls, respectively, in the Poissonized scheme at time t . The random variables

$$K_j(t) = \sum_{k \geq 1} \mathbb{1}_{\{\text{the box } k \text{ contains at least } j \text{ balls at time } t\}}$$

and

$$K_j^*(t) = \sum_{k \geq 1} \mathbb{1}_{\{\text{the box } k \text{ contains exactly } j \text{ balls at time } t\}} \tag{1}$$

are the infinite sums of independent indicators. As a consequence, their analysis is much simpler than that of $\mathcal{K}_j(n)$ and $\mathcal{K}_j^*(n)$ which are infinite sums of dependent indicators.

1.2 Main results

Put

$$\rho(t) := \#\{k \in \mathbb{N} : 1/p_k \leq t\}, \quad t > 0,$$

and note that $\rho(t) = 0$ for $t \in (0, 1]$. Following Karlin [12] we assume that ρ varies regularly at ∞ of index $\alpha \in [0, 1]$, that is, $\rho(t) \sim t^\alpha L(t)$ as $t \rightarrow \infty$ for some L slowly varying at ∞ . An encyclopaedic treatment of slowly and regularly varying functions can be found in Section 1 of [2].

The function ρ is said to belong to the *de Haan class* Π if, for all $\lambda > 0$,

$$\lim_{t \rightarrow \infty} \frac{\rho(\lambda t) - \rho(t)}{\ell(t)} = \log \lambda \tag{2}$$

for some ℓ slowly varying at ∞ . The function ℓ is called *auxiliary*. According to Theorem 3.7.4 in [2], the class Π is a subclass of the class of slowly varying functions. Further detailed information regarding the class Π is given in Section 3 of [2] and in [8]. Denote by $\Pi_{\ell, \infty}$ the subclass of the de Haan class Π with the auxiliary functions ℓ satisfying $\lim_{t \rightarrow \infty} \ell(t) = \infty$.

In the case $\alpha \in (0, 1]$, according to Theorems 3, 5 and 5* in [12], both $K_j^*(t)$ and $\mathcal{K}_j^*(n)$, centered by their means and normalized by their standard deviations, converge in distribution to a random variable with the standard normal distribution. In the case $\rho \in \Pi_{\ell, \infty}$, Corollary 1.6 in [11] provides functional central limit theorems for $K_j^*(t)$ and $\mathcal{K}_j^*(n)$, properly scaled. Our purpose is to prove laws of the iterated logarithm (LILs) for $K_j^*(t)$ as $t \rightarrow \infty$ and $\mathcal{K}_j^*(n)$ as $n \rightarrow \infty$. While doing so, we treat the three cases separately: $\alpha = 0$, $\alpha \in (0, 1)$ and $\alpha = 1$. The reason is that the forms of the LILs are slightly or essentially different in these cases. If ρ is slowly varying at ∞ and satisfies an additional assumption, then the actual limit relation is either a law of the single logarithm or a LIL. However, to keep the presentation simple we prefer to call LILs all the limit relations involving upper or lower limits which appear in the paper.

In Theorems 1, 2 and 3 we present LILs for the Poissonized variables $K_j^*(t)$ as $t \rightarrow \infty$. Theorem 1 covers a subcase of the case $\alpha = 0$ in which $\rho \in \Pi_{\ell, \infty}$ with particular ℓ .

Theorem 1. *Assume that (2) holds. If ℓ in (2) satisfies*

$$\ell(t) \sim (\log t)^\beta l(\log t), \quad t \rightarrow \infty, \tag{3}$$

for some $\beta > 0$ and l slowly varying at ∞ , then, for each $j \in \mathbb{N}$,

$$\limsup_{t \rightarrow \infty} \frac{K_j^*(t) - \mathbb{E}K_j^*(t)}{(\text{Var } K_j^*(t) \log \text{Var } K_j^*(t))^{1/2}} = \left(\frac{2}{\beta}\right)^{1/2} \quad \text{a.s.} \tag{4}$$

and

$$\liminf_{t \rightarrow \infty} \frac{K_j^*(t) - \mathbb{E}K_j^*(t)}{(\text{Var } K_j^*(t) \log \text{Var } K_j^*(t))^{1/2}} = -\left(\frac{2}{\beta}\right)^{1/2} \quad \text{a.s.} \tag{5}$$

If ℓ in (2) satisfies

$$\ell(t) \sim \exp(\sigma(\log t)^\lambda), \quad t \rightarrow \infty, \tag{6}$$

for some $\sigma > 0$ and $\lambda \in (0, 1)$, then, for each $j \in \mathbb{N}$,

$$\limsup_{t \rightarrow \infty} \frac{K_j^*(t) - \mathbb{E}K_j^*(t)}{(\text{Var } K_j^*(t) \log \log \text{Var } K_j^*(t))^{1/2}} = \left(\frac{2}{\lambda}\right)^{1/2} \quad \text{a.s.} \tag{7}$$

and

$$\liminf_{t \rightarrow \infty} \frac{K_j^*(t) - \mathbb{E}K_j^*(t)}{(\text{Var } K_j^*(t) \log \log \text{Var } K_j^*(t))^{1/2}} = -\left(\frac{2}{\lambda}\right)^{1/2} \quad \text{a.s.} \tag{8}$$

In both cases

$$\mathbb{E}K_j^*(t) \sim \frac{\ell(t)}{j}, \quad t \rightarrow \infty, \tag{9}$$

and

$$\text{Var } K_j^*(t) \sim \left(\frac{1}{j} - \frac{(2j-1)!}{(j!)^2 2^{2j}} \right) \ell(t), \quad t \rightarrow \infty. \tag{10}$$

Remark 1. Treatment of the situations in which ρ is slowly varying at ∞ , yet $\rho \notin \Pi$, is beyond our reach. To reveal complications arising in this case we only mention that even the large-time asymptotics of $t \mapsto \text{Var } K_j^*(t)$ is not known. To find the asymptotics, a second-order relation for ρ like (2) seems to be indispensable. If $\alpha \in (0, 1]$, then the regular variation of ρ alone ensures that, for all $\lambda > 0$,

$$\lim_{t \rightarrow \infty} \frac{\rho(\lambda t) - \rho(t)}{\rho(t)} = \lambda^\alpha - 1.$$

Thus, no extra conditions are needed in this case.

Remark 2. Our present proof only works provided that, for some $a > 0$, $\rho(t) = O((\ell(t))^a)$ as $t \rightarrow \infty$. In view of this, Theorem 1 does not cover the diverging slowly varying functions ℓ which grow slower than any positive power of the logarithm, for instance, $\ell(t) \sim \log \log t$ as $t \rightarrow \infty$. Indeed, it can be checked that $\lim_{t \rightarrow \infty} \ell(t) = \infty$ entails $\lim_{t \rightarrow \infty} (\rho(t)/\log t) = \infty$, whence trivially, for all $a > 0$, $\lim_{t \rightarrow \infty} (\rho(t)/(\ell(t))^a) = \infty$.

The following results are concerned with the cases $\alpha \in (0, 1)$ and $\alpha = 1$, respectively.

Theorem 2. Assume that, for some $\alpha \in (0, 1)$ and some L slowly varying at $+\infty$,

$$\rho(t) \sim t^\alpha L(t), \quad t \rightarrow \infty.$$

Then, for each $j \in \mathbb{N}$,

$$\limsup_{t \rightarrow \infty} \frac{K_j^*(t) - \mathbb{E}K_j^*(t)}{(\text{Var } K_j^*(t) \log \log \text{Var } K_j^*(t))^{1/2}} = 2^{1/2} \quad a.s. \tag{11}$$

and

$$\liminf_{t \rightarrow \infty} \frac{K_j^*(t) - \mathbb{E}K_j^*(t)}{(\text{Var } K_j^*(t) \log \log \text{Var } K_j^*(t))^{1/2}} = -2^{1/2} \quad a.s., \tag{12}$$

$$\mathbb{E}K_j^*(t) \sim \alpha \frac{\Gamma(j-\alpha)}{j!} t^\alpha L(t) \tag{13}$$

and

$$\text{Var } K_j^*(t) \sim c_{j,\alpha} t^\alpha L(t), \quad t \rightarrow \infty, \tag{14}$$

where Γ is the Euler gamma function and

$$c_{j,\alpha} := \alpha \left(\frac{\Gamma(j-\alpha)}{j!} - \frac{2^\alpha \Gamma(2j-\alpha)}{2^{2j} (j!)^2} \right) > 0. \tag{15}$$

Theorem 3. Assume that, for some L slowly varying at $+\infty$,

$$\rho(t) \sim tL(t), \quad t \rightarrow \infty.$$

Then, for each $j \geq 2$, relation (11) holds,

$$\mathbb{E}K_j^*(t) \sim \frac{1}{j(j-1)}tL(t) \tag{16}$$

and

$$\lim_{t \rightarrow \infty} \frac{\text{Var } K_j^*(t)}{tL(t)} = \frac{1}{j(j-1)} - \frac{(2j-2)!}{2^{2j-1}(j!)^2} = c_{j,1}. \tag{17}$$

Assume that, for each small enough $\gamma > 0$,

$$\lim_{n \rightarrow \infty} \frac{\hat{L}(\exp((n+1)^{1+\gamma}))}{\hat{L}(\exp(n^{1+\gamma}))} = 0, \tag{18}$$

where $\hat{L}(t) := \int_t^\infty y^{-1}L(y)dy$, being well-defined for large t , is a function slowly varying at ∞ and satisfying

$$\lim_{t \rightarrow \infty} \frac{L(t)}{\hat{L}(t)} = 0. \tag{19}$$

Then relation (11) holds with $j = 1$. If (18) does not hold, then

$$\limsup_{t \rightarrow \infty} \frac{K_1^*(t) - \mathbb{E}K_1^*(t)}{(\text{Var } K_1^*(t) \log \log \text{Var } K_1^*(t))^{1/2}} \leq 2^{1/2} \quad a.s. \tag{20}$$

and

$$\liminf_{t \rightarrow \infty} \frac{K_1^*(t) - \mathbb{E}K_1^*(t)}{(\text{Var } K_1^*(t) \log \log \text{Var } K_1^*(t))^{1/2}} \geq -2^{1/2} \quad a.s. \tag{21}$$

In any event

$$\text{Var } K_1^*(t) \sim \mathbb{E}K_1^*(t) \sim t\hat{L}(t), \quad t \rightarrow \infty. \tag{22}$$

Theorems 1, 2 and 3 will be deduced in Section 4 from the LIL for infinite sums of independent indicators given in Theorem 5.

Finally, we present LILs for the variables $\mathcal{K}_j^*(n)$.

Theorem 4. Under the assumptions of Theorems 1, 2 or 3, for $j \in \mathbb{N}$, all the LILs stated there hold true with $\mathcal{K}_j^*(n)$, $\mathbb{E}\mathcal{K}_j^*(n)$ and $\text{Var } \mathcal{K}_j^*(n)$ replacing $K_j^*(t)$, $\mathbb{E}K_j^*(t)$ and $\text{Var } K_j^*(t)$, and $n \rightarrow \infty$ replacing $t \rightarrow \infty$.

A transfer of results available for the Poissonized version to the deterministic version is called *de-Poissonization*. Theorem 4 will be deduced in Section 4 from Theorems 1, 2 and 3 with the help of a de-Poissonization technique.

Remark 3. Following the referee's suggestion, for the sake of comparison, we now provide a verbal description of the LILs obtained in [3]. Under the assumptions of Theorems 1, 2 and 3, the limit relations (4), (5), (7), (8), (11), (12), (20) and (21) hold true with $K_j(t)$, $\mathbb{E}K_j(t)$ and $\text{Var } K_j(t)$ replacing $K_j^*(t)$, $\mathbb{E}K_j^*(t)$ and $\text{Var } K_j^*(t)$. Also, these limit relations hold true with $\mathcal{K}_j(n)$, $\mathbb{E}\mathcal{K}_j(n)$ and $\text{Var } \mathcal{K}_j(n)$ replacing $K_j^*(t)$,

$\mathbb{E}K_j^*(t)$ and $\text{Var} K_j^*(t)$, and $n \rightarrow \infty$ replacing $t \rightarrow \infty$. Typically, the means and the variances of $K_j^*(t)$ and $K_j(t)$, and $\mathcal{K}_j^*(n)$ and $\mathcal{K}_j(n)$ exhibit the same rate of growth, up to a multiplicative constant. The only exception is that, under the assumptions of Theorem 1, for $j \in \mathbb{N}$,

$$\mathbb{E}K_j^*(t) \sim \frac{\ell(t)}{j}, \quad t \rightarrow \infty, \quad \text{and} \quad \mathbb{E}\mathcal{K}_j^*(n) \sim \frac{\ell(n)}{j}, \quad n \rightarrow \infty,$$

whereas

$$\mathbb{E}K_j(t) \sim \rho(t), \quad t \rightarrow \infty, \quad \text{and} \quad \mathbb{E}\mathcal{K}_j(n) \sim \rho(n), \quad n \rightarrow \infty.$$

The latter relations are secured by Lemma 10 and the formula

$$\lim_{n \rightarrow \infty} |\mathbb{E}K_j(n) - \mathbb{E}\mathcal{K}_j(n)| = 0$$

which can be found in Lemma 1 of [9].

2 LIL for infinite sums of independent indicators

Let $(A_1(t))_{t \geq 0}$, $(A_2(t))_{t \geq 0}$, \dots be independent families of events defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that $\sum_{k \geq 1} \mathbb{P}(A_k(t)) < \infty$, for each $t \geq 0$, and then put

$$X(t) := \sum_{k \geq 1} \mathbb{1}_{A_k(t)}, \quad t \geq 0.$$

Since, for $t \geq 0$, $b(t) := \mathbb{E}X(t) = \sum_{k \geq 1} \mathbb{P}(A_k(t)) < \infty$, we infer $X(t) < \infty$ almost surely (a.s.) and further

$$a(t) := \text{Var} X(t) = \sum_{k \geq 1} \mathbb{P}(A_k(t))(1 - \mathbb{P}(A_k(t))) \leq b(t) < \infty.$$

Under the assumption that, for each $k \in \mathbb{N}$ and $0 \leq s < t$, $A_k(s) \subseteq A_k(t)$ a LIL for $X(t)$ can be found in Theorem 1.6 of [3]. As an application, LILs for $K_j(t)$ were proved in that paper, see Theorems 3.1, 3.3 and 3.4 therein. According to (1), the variable $K_j^*(t)$ is a particular instance of $X(t)$. However, for each $k \in \mathbb{N}$, the corresponding events $(A_k(t))_{t \geq 0}$ are not monotone in t , and therefore, a LIL for $K_j^*(t)$ cannot be deduced from Theorem 1.6 of [3]. This serves a motivation for the present section. Here, dropping the monotonicity assumption we provide sufficient conditions under which a LIL for $X(t)$ holds.

We shall prove a LIL for $X(t)$ under the assumptions (A1)–(A5) and (B1)–(B21) or (B22) given below. The lack of monotonicity only affects our proof for the upper bound of $\limsup_{t \rightarrow \infty}$ to be done under (A1)–(A5). In view of this, (A2)–(A5) are modified versions of the corresponding assumptions in [3]. (B1), (B21) and (B22) coincide with the corresponding assumptions in [3] under which the lower bound of $\limsup_{t \rightarrow \infty}$ was found in the cited article.

(A1) $\lim_{t \rightarrow \infty} a(t) = \infty$.

(A2) There exist independent a.s. nondecreasing stochastic processes $(\Phi_1(t))_{t \geq 0}$, $(\Phi_2(t))_{t \geq 0}, \dots$ taking values in $\{0, 1, 2, \dots, M\}$ for some $M \in \mathbb{N}$ and satisfying

- (a) for each $k \in \mathbb{N}$, $0 \leq s < t$, $|\mathbb{1}_{A_k(t)} - \mathbb{1}_{A_k(s)}| \leq \Phi_k(t) - \Phi_k(s)$ a.s.;
- (b) for each $t \geq 0$, $f(t) := \mathbb{E}Y(t) < \infty$, where $Y(t) := \sum_{k \geq 1} \Phi_k(t)$ for $t \geq 0$;
- (c) $b(0) \leq f(0)$.

(A3) Under (A2), there exists $\mu^* \geq 1$ such that $f(t) = O((a(t))^{\mu^*})$ as $t \rightarrow \infty$. In view of (28) and $a(t) \leq b(t)$ for $t \geq 0$, necessarily $\mu^* \geq 1$. Put

$$\mu := \inf\{\mu^* : f(t) = O((a(t))^{\mu^*})\}. \tag{23}$$

If $\mu = 1$, we assume additionally that either f is eventually continuous or

$$\liminf_{t \rightarrow \infty} (\log f(t-1) / \log f(t)) > 0;$$

and that

$$f(t)/a(t) = O(z_q(a(t))), \quad t \rightarrow \infty, \tag{24}$$

where $z_q(t) := (\log t)^q \mathcal{L}(\log t)$ for some $q \geq 0$ and \mathcal{L} is slowly varying at ∞ and, if $q > 0$, $f(t)/a(t) \neq O(z_s(a(t)))$ for $s \in (0, q)$.

Before introducing our next assumption we need some preparation. In view of (A1) and $a(t) \leq f(t)$ for $t \geq 0$, we infer $\lim_{t \rightarrow \infty} f(t) = \infty$. For each $\varrho \in (0, 1)$, put

$$\mu_\varrho := \mu + \varrho \quad \text{if } \mu > 1 \quad \text{and} \quad q_\varrho := q + \varrho \quad \text{if } \mu = 1. \tag{25}$$

In other words, if $\mu > 1$, we work with μ_ϱ , if $\mu = 1$, we work with q_ϱ . Assuming (A3), fix any $\kappa \in (0, 1)$ and $\varrho \in (0, 1)$ and put

$$t_n = t_n(\kappa, \mu) := \inf\{t > 0 : f(t) > v_n(\kappa, \mu)\} \tag{26}$$

for $n \in \mathbb{N}$, where $v_n(\kappa, 1) = v_n(\kappa, 1, q, \varrho) = \exp(n^{(1-\kappa)/(q_\varrho+1)})$ and $v_n(\kappa, \mu) = v_n(\kappa, \mu, \varrho) = n^{\mu_\varrho(1-\kappa)/(\mu_\varrho-1)}$ for $\mu > 1$. Plainly, the sequence $(t_n)_{n \in \mathbb{N}}$ is nondecreasing with $\lim_{n \rightarrow \infty} t_n = +\infty$.

(A4) Fix any $\kappa \in (0, 1)$ and $\varrho \in (0, 1)$. There exists a function a_0 satisfying $a(t) \sim a_0(t)$ as $t \rightarrow \infty$, and, for each n large enough, there exists $s_n = s_n(\kappa, \mu) \in [t_n(\kappa, \mu), t_{n+1}(\kappa, \mu)]$ such that $a_0(t) \geq a_0(s_n)$ for all $t \in [t_n, t_{n+1}]$.

(A5) For each n large enough, there exists $A > 1$ and a partition $t_n = t_{0,n} < t_{1,n} < \dots < t_{j,n} = t_{n+1}$ with $j = j_n$ satisfying

$$1 \leq f(t_{k,n}) - f(t_{k-1,n}) \leq A, \quad 1 \leq k \leq j,$$

and, for all $\varepsilon > 0$, $(j_n \exp(-\varepsilon(a(s_n))^{1/2}))$ is a summable sequence.

Remark 4. (A2a) and (A2b) entail, for $0 \leq s < t$,

$$|b(t) - b(s)| \leq \mathbb{E} \sum_{k \geq 1} |\mathbb{1}_{A_k(t)} - \mathbb{1}_{A_k(s)}| \leq f(t) - f(s). \tag{27}$$

Inequality (27) with $s = 0$ and (A2c) together imply that

$$b(t) \leq f(t) - f(0) + b(0) \leq f(t), \quad t \geq 0. \tag{28}$$

Remark 5. Here is an example of X satisfying (A2) which is motivated by a prospective application of the LIL for $X(t)$ to the variables $K_j^*(t)$. Let $(B_1(t))_{t \geq 0}$, $(B_2(t))_{t \geq 0}$, \dots and $(C_1(t))_{t \geq 0}$, $(C_2(t))_{t \geq 0}$, \dots be two families of independent events satisfying

- (i) for each $k \in \mathbb{N}$ and $t \geq 0$, $C_k(t) \subseteq B_k(t)$;
- (ii) for each $k \in \mathbb{N}$ and $0 \leq s < t$, $B_k(s) \subseteq B_k(t)$ and $C_k(s) \subseteq C_k(t)$;
- (iii) for $t \geq 0$, $\sum_{k \geq 1} \mathbb{P}(B_k(t)) < \infty$.

For each $k \in \mathbb{N}$ and $t \geq 0$, put $A_k(t) := B_k(t) \setminus C_k(t)$ and $\Phi_k(t) := \mathbb{1}_{B_k(t)} + \mathbb{1}_{C_k(t)}$. The so defined Φ_k is a.s. nondecreasing. Since, for $0 \leq s < t$, $\mathbb{1}_{C_k(s)} \leq \mathbb{1}_{C_k(t)}$ and $\mathbb{1}_{B_k(s)} \leq \mathbb{1}_{B_k(t)}$ a.s. we conclude that

$$|\mathbb{1}_{A_k(t)} - \mathbb{1}_{A_k(s)}| = |\mathbb{1}_{B_k(t)} - \mathbb{1}_{C_k(t)} - \mathbb{1}_{B_k(s)} + \mathbb{1}_{C_k(s)}| \leq \Phi_k(t) - \Phi_k(s) \quad \text{a.s.}$$

While (A2b) is a consequence of (iii), (A2c) is justified by $\mathbb{P}(A_k(0)) = \mathbb{P}(B_k(0)) - \mathbb{P}(C_k(0)) \leq \mathbb{E}\Phi_k(0)$.

Putting $C_k(t) := \emptyset$ for all $k \in \mathbb{N}$ and $t \geq 0$ we recover the case of monotone in t families $(A_k(t))_{t \geq 0}$ treated in [3].

Remark 6. The assumption imposed in [3] meaning that, for each $k \in \mathbb{N}$, the family $(A_k(t))_{t \geq 0}$ is nondecreasing in t simplifies significantly the analysis of (43). Indeed, for any $\theta > 0$, we then infer $\sup_{v \in [0, \theta]} |X(t+v) - X(t)| = X(t+\theta) - X(t)$ and $\sup_{v \in [0, \theta]} |b(t+v) - b(t)| = b(t+\theta) - b(t)$. In the absence of the monotonicity assumption, it is necessary to find some monotone majorant for $|X(t+v) - X(t)|$ which is sufficiently close to the true supremum.

One may expect that f behaving like $f(t) = O(b(t))$ as $t \rightarrow \infty$ should do the job. What is not trivial is that f satisfying $\lim_{t \rightarrow \infty} (f(t)/b(t)) = \infty$ may also be suitable. For instance, consider the setting of Theorem 1 and $X(t) := K_j^*(t)$ for $j \in \mathbb{N}$. By (9), $b(t) \sim \text{const } \ell(t)$ as $t \rightarrow \infty$, and by (63) and (48), $f(t) \sim \text{const } \rho(t)$ as $t \rightarrow \infty$. Applying Lemma 9 we conclude that indeed $\lim_{t \rightarrow \infty} (f(t)/b(t)) = \infty$.

Remark 7. A sufficient condition for (A4) is either eventual lower semi-continuity or eventual monotonicity of a_0 . The former means that $\liminf_{y \rightarrow x} a_0(y) \geq a_0(x)$, for all large enough x .

Remark 8. A sufficient condition for (A5) is that f is eventually strictly increasing and eventually continuous. Indeed, one can then choose a partition that satisfies, for large n , $f(t_{k,n}) - f(t_{k-1,n}) = 1$ for $k \in \mathbb{N}, k \leq j-1$ and $f(t_{j,n}) - f(t_{j-1,n}) \in [1, 2)$. As a consequence,

$$j_n = \lfloor v_{n+1}(\kappa, \mu) - v_n(\kappa, \mu) \rfloor = o(a(s_n)), \quad n \rightarrow \infty,$$

by Lemma 2(b) below, so that the sequence $(j_n \exp(-\varepsilon(a(s_n))^{1/2}))$ is indeed summable.

Assuming (A1) and (A3), fix any $\gamma > 0$ and put

$$\tau_n = \tau_n(\gamma, \mu) := \inf\{t > 0 : a(t) > w_n(\gamma, \mu)\} \tag{29}$$

for large $n \in \mathbb{N}$ with μ as given in (23). Here, with q as given in (24), $w_n(\gamma, 1) = w_n(\gamma, 1, q) = \exp(n^{(1+\gamma)/(q+1)})$ if $\mu = 1$ and $w_n(\gamma, \mu) = n^{(1+\gamma)/(\mu-1)}$ if $\mu > 1$.

(B1) The function a is eventually continuous or $\lim_{t \rightarrow \infty} (\log a(t-1) / \log a(t)) = 1$ if $\mu = 1$ and $\lim_{t \rightarrow \infty} (a(t-1) / a(t)) = 1$ if $\mu > 1$.

(B21) For sufficiently large $t > 0$ and each $\zeta > 0$, let $R_\zeta(t)$ denote a set of positive integers satisfying the following two conditions: for each $\zeta > 0$ and each $\gamma > 0$, both close to 0, there exists $n_0 = n_0(\zeta, \gamma) \in \mathbb{N}$ such that the sets $R_\zeta(\tau_{n_0}(\gamma, \mu)), R_\zeta(\tau_{n_0+1}(\gamma, \mu)), \dots$ are disjoint; and

$$\lim_{t \rightarrow \infty} \frac{\text{Var}(\sum_{k \in R_\zeta(t)} \mathbb{1}_{A_k(t)})}{\text{Var } X(t)} = 1 - \zeta.$$

(B22) For sufficiently large $t > 0$, let $R_0(t)$ denote a set of positive integers satisfying the following two conditions: for each $\gamma > 0$ close to 0 there exists $n_0 = n_0(\gamma) \in \mathbb{N}$ such that the sets $R_0(\tau_{n_0}(\gamma, \mu)), R_0(\tau_{n_0+1}(\gamma, \mu)), \dots$ are disjoint; and

$$\lim_{t \rightarrow \infty} \frac{\text{Var}(\sum_{k \in R_0(t)} \mathbb{1}_{A_k(t)})}{\text{Var } X(t)} = 1. \tag{30}$$

Now we are ready to present a LIL for infinite sums of independent indicators.

Theorem 5. *Suppose (A1)–(A5), (B1) and either (B21) or (B22). Then, with $\mu \geq 1$ and $q \geq 0$ as defined in (23) and (24), respectively,*

$$\limsup_{t \rightarrow \infty} \frac{X(t) - \mathbb{E}X(t)}{(2(q+1)\text{Var } X(t) \log \log \text{Var } X(t))^{1/2}} = 1 \quad a.s.$$

and

$$\liminf_{t \rightarrow \infty} \frac{X(t) - \mathbb{E}X(t)}{(2(q+1)\text{Var } X(t) \log \log \text{Var } X(t))^{1/2}} = -1 \quad a.s.$$

if $\mu = 1$ and

$$\limsup_{t \rightarrow \infty} \frac{X(t) - \mathbb{E}X(t)}{(2(\mu-1)\text{Var } X(t) \log \text{Var } X(t))^{1/2}} = 1 \quad a.s.$$

and

$$\liminf_{t \rightarrow \infty} \frac{X(t) - \mathbb{E}X(t)}{(2(\mu-1)\text{Var } X(t) \log \text{Var } X(t))^{1/2}} = -1 \quad a.s.$$

if $\mu > 1$.

Our proof of Theorem 5 given in Section 3.2 is a modified version of the proof of Theorem 1.6 in [3].

3 Proof of Theorem 5

3.1 Auxiliary results

We start with a simple inequality which will be used in the last part of the proof of Proposition 2.

Lemma 1. *Suppose (A2). Then, for $\vartheta \in \mathbb{R}$ and $t > s \geq 0$,*

$$\mathbb{E} \exp(\vartheta(Y(t) - Y(s))) \leq \exp((e^{\vartheta M} - 1)(f(t) - f(s))).$$

Proof. For $k \in \mathbb{N}$ and $t > s \geq 0$, $\mathbb{1}_{\{\Phi_k(t) - \Phi_k(s) > 0\}} = \mathbb{1}_{\{\Phi_k(t) - \Phi_k(s) \geq 1\}} \leq \Phi_k(t) - \Phi_k(s)$ a.s. The equality stems from the fact that Φ_k only takes nonnegative integer values. Hence, for $\vartheta \in \mathbb{R}$ and $0 \leq s < t$,

$$\begin{aligned} \mathbb{E} \exp(\vartheta(Y(t) - Y(s))) &= \prod_{k \geq 1} \mathbb{E} \exp(\vartheta(\Phi_k(t) - \Phi_k(s))) \\ &= \prod_{k \geq 1} (1 + \mathbb{E}(e^{\vartheta(\Phi_k(t) - \Phi_k(s))} - 1) \mathbb{1}_{\{\Phi_k(t) - \Phi_k(s) > 0\}}) \\ &\leq \prod_{k \geq 1} (1 + (e^{\vartheta M} - 1) \mathbb{E} \mathbb{1}_{\{\Phi_k(t) - \Phi_k(s) > 0\}}) \\ &\leq \exp\left((e^{\vartheta M} - 1) \sum_{k \geq 1} \mathbb{E}(\Phi_k(t) - \Phi_k(s))\right) \\ &= \exp((e^{\vartheta M} - 1)(f(t) - f(s))). \quad \square \end{aligned}$$

For each $B \geq 0$ and each $D > 1$, put

$$g_{1,B}(t) := (B+1) \log \log t, \quad t > e, \quad \text{and} \quad g_D(t) := (D-1) \log t, \quad t > 1. \quad (31)$$

Lemma 2 does two things. First, it explains the choice of the sequences (t_n) and (v_n) and the functions g_{1,q_ϱ} and g_{μ_ϱ} (even though (t_n) is not present in Lemma 2 explicitly, it is of crucial importance for defining the sequence (s_n)). Second, it secures a successful application of the Borel–Cantelli lemma in the proof of Proposition 2.

Lemma 2. *Suppose (A1), (A3) and (A4). Fix any $\varrho \in (0, 1)$, any $\kappa \in (0, 1)$ and let q_ϱ and μ_ϱ be as defined in (25).*

(a) *If μ in (23) is equal to 1, then $\exp(-g_{1,q_\varrho}(a(s_n(\kappa, 1)))) = O(n^{-(1-\kappa)})$ as $n \rightarrow \infty$, and if $\mu > 1$, then $\exp(-g_{\mu_\varrho}(a(s_n(\kappa, \mu)))) = O(n^{-(1-\kappa)})$.*

(b) *There exists an integer $r \geq 2$ such that $((v_{n+1}(\kappa, \mu) - v_n(\kappa, \mu))/a(s_n))^r$ is a summable sequence.*

Proof. (a) Using the definition of t_n , the fact that f is nondecreasing and (A3), we conclude that, as $n \rightarrow \infty$,

$$\exp(n^{(1-\kappa)/(q_\varrho+1)}) \leq f(t_n(\kappa, 1)) \leq f(s_n(\kappa, 1)) = O(a(s_n(\kappa, 1))z_q(a(s_n(\kappa, 1)))) \quad (32)$$

and, for $\mu > 1$,

$$n^{\mu_\varrho(1-\kappa)/(\mu_\varrho-1)} \leq f(t_n(\kappa, \mu)) \leq f(s_n(\kappa, \mu)) = O((a(s_n(\kappa, \mu)))^{\mu_\varrho}), \quad n \rightarrow \infty. \quad (33)$$

Since $\lim_{t \rightarrow \infty} (\log z_q(t) / \log t) = 0$, we infer

$$\exp(-g_{1, q_e}(a(s_n(\kappa, 1)))) = (\log a(s_n(\kappa, 1)))^{-(q_e+1)} = O(n^{-(1-\kappa)}), \quad n \rightarrow \infty.$$

Also, for $\mu > 1$,

$$\exp(-g_{\mu_e}(a(s_n(\kappa, \mu)))) = (a(s_n(\kappa, \mu)))^{-(\mu_e-1)} = O(n^{-(1-\kappa)}), \quad n \rightarrow \infty.$$

(b) We start by proving that (A3) with $\mu = 1$ entails

$$\log a(s_n(\kappa, 1)) = O(n^{(1-\kappa)/(q_e+1)}), \quad n \rightarrow \infty. \tag{34}$$

Assume that f is eventually continuous. Then $f(t_n(\kappa, 1)) = v_n(\kappa, 1)$ for large enough n and thereupon $\log a(s_n(\kappa, 1)) \leq \log f(s_n(\kappa, 1)) \leq \log f(t_{n+1}(\kappa, 1)) = (n + 1)^{(1-\kappa)/(q_e+1)}$ for large n . Assuming that $\liminf_{t \rightarrow \infty} (\log f(t - 1) / \log f(t)) > 0$, we obtain (34) as a consequence of $\log f(t_{n+1}(\kappa, 1) - 1) \leq (n + 1)^{(1-\kappa)/(q_e+1)}$ and $\log a(s_n(\kappa, 1)) \leq \log f(s_n(\kappa, 1)) \leq \log f(t_{n+1}(\kappa, 1))$.

We proceed by noting that, as $n \rightarrow \infty$,

$$\begin{aligned} v_{n+1}(\kappa, 1) - v_n(\kappa, 1) &= \exp((n + 1)^{(1-\kappa)/(q_e+1)}) - \exp(n^{(1-\kappa)/(q_e+1)}) \\ &\sim ((1 - \kappa)/(q_e + 1))n^{((1-\kappa)/(q_e+1)-1)} \exp(n^{(1-\kappa)/(q_e+1)}) \end{aligned}$$

and, for $\mu > 1$,

$$\begin{aligned} v_{n+1}(\kappa, \mu) - v_n(\kappa, \mu) &= (n + 1)^{\mu_e(1-\kappa)/(\mu_e-1)} - n^{\mu_e(1-\kappa)/(\mu_e-1)} \\ &\sim (\mu_e(1 - \kappa)/(\mu_e - 1))n^{(1-\mu_e\kappa)/(\mu_e-1)}. \end{aligned}$$

Write

$$\begin{aligned} \frac{1}{a(s_n(\kappa, 1))} &= O((\log a(s_n(\kappa, 1)))^{q_e} \exp(-n^{(1-\kappa)/(q_e+1)})) \\ &= O(n^{q_e(1-\kappa)/(q_e+1)} \exp(-n^{(1-\kappa)/(q_e+1)})), \quad n \rightarrow \infty. \end{aligned}$$

Here, the first equality is implied by $z_q(t) = O((\log t)^{q_e})$ as $t \rightarrow \infty$ and (32), and the second equality is a consequence of (34). In the case $\mu > 1$, invoking (33) we infer

$$\frac{1}{a(s_n(\kappa, \mu))} = O(n^{-(1-\kappa)/(\mu_e-1)}), \quad n \rightarrow \infty.$$

Thus, we have proved that, for $\mu \geq 1$,

$$\frac{v_{n+1}(\kappa, \mu) - v_n(\kappa, \mu)}{a(s_n)} = O(n^{-\kappa}), \quad n \rightarrow \infty.$$

Choosing any integer $r \geq 2$ satisfying $r\kappa > 1$ completes the proof of part (b). \square

For $k \in \mathbb{N}$ and $t \geq 0$, put $X^*(t) := X(t) - \mathbb{E}X(t)$ and $\eta_k(t) := \mathbb{1}_{A_k(t)} - \mathbb{P}(A_k(t))$. Note that

$$X^*(t) = \sum_{k \geq 1} \eta_k(t), \quad t \geq 0,$$

and that $\eta_1(t), \eta_2(t), \dots$ are independent centered random variables.

Lemma 3 provides a uniform bound for higher moments of the increments of X^* . The bound serves a starting point of the chaining argument in the spirit of Lemma 4. A result of an application of Lemma 4 to the present setting is given in Lemma 5.

Lemma 3. *Suppose (A2). Let $r \in \mathbb{N}$ and $t, s \geq 0$. Then*

$$\mathbb{E}(X^*(t) - X^*(s))^{2r} \leq D_r \max(|f(t) - f(s)|^r, |f(t) - f(s)|) \tag{35}$$

for a positive constant D_r which does not depend on t and s .

Proof. In view of the representation

$$X^*(t) - X^*(s) = \sum_{k \geq 1} (\mathbb{1}_{A_k(t)} - \mathbb{P}(A_k(t)) - \mathbb{1}_{A_k(s)} + \mathbb{P}(A_k(s))) =: \sum_{k \geq 1} \eta_k(s, t),$$

the variable $X^*(t) - X^*(s)$ is an infinite sum of independent centered random variables with finite moment of order $2r$.

Invoking Rosenthal’s inequality (Theorem 3 in [14]) in the case $r \geq 2$ we infer

$$\mathbb{E}(X^*(t) - X^*(s))^{2r} \leq C_r \max\left(\left(\sum_{k \geq 1} \mathbb{E}(\eta_k(s, t))^2\right)^r, \sum_{k \geq 1} \mathbb{E}(\eta_k(s, t))^{2r}\right).$$

In the case $r = 1$, the inequality trivially holds with $C_1 = 1$ as is seen from

$$\mathbb{E}(X^*(t) - X^*(s))^2 = \sum_{k \geq 1} \mathbb{E}(\eta_k(s, t))^2.$$

In view of (A2), for $r \in \mathbb{N}$ and $0 \leq s < t$,

$$\begin{aligned} & \sum_{k \geq 1} \mathbb{E}(\eta_k(s, t))^{2r} \\ & \leq 2^{2r-1} \sum_{k \geq 1} (\mathbb{E}(\mathbb{1}_{A_k(t)} - \mathbb{1}_{A_k(s)})^{2r} + (\mathbb{P}(A_k(t)) - \mathbb{P}(A_k(s)))^{2r}) \\ & \leq 2^{2r-1} \sum_{k \in \mathbb{N}} (\mathbb{E}|\mathbb{1}_{A_k(t)} - \mathbb{1}_{A_k(s)}| + |\mathbb{P}(A_k(t)) - \mathbb{P}(A_k(s))|) \\ & \leq 2^{2r} \sum_{k \geq 1} \mathbb{E}(\Phi_k(t) - \Phi_k(s)) = 2^{2r}(f(t) - f(s)). \end{aligned}$$

Here, we have used $(a + b)^{2r} \leq 2^{2r-1}(a^{2r} + b^{2r})$, $a, b \in \mathbb{R}$, for the first inequality, the fact that $|\mathbb{1}_{A_k(t)} - \mathbb{1}_{A_k(s)}| \in \{0, 1\}$ a.s. and $|\mathbb{P}(A_k(t)) - \mathbb{P}(A_k(s))| \in [0, 1]$ for the second and (A2a) for the third. The argument for the case $0 \leq t < s$ is analogous.

Combining fragments together we conclude that (35) holds with $D_r := 12^{2r} C_r$. □

The next result is borrowed from Lemma 2 in [13].

Lemma 4. Let ξ_1, ξ_2, \dots be random variables. Fix any $m \in \mathbb{N}$ and assume that

$$\mathbb{E}|\xi_{i+1} + \dots + \xi_k|^{\lambda_1} \leq (u_{i+1} + \dots + u_k)^{\lambda_2}, \quad 0 \leq i < k \leq m,$$

for some $\lambda_1 > 0$, some $\lambda_2 > 1$ and some nonnegative numbers u_1, \dots, u_m . Then

$$\mathbb{E}\left(\max_{1 \leq k \leq m} |\xi_1 + \dots + \xi_k|\right)^{\lambda_1} \leq A_{\lambda_1, \lambda_2} (u_1 + \dots + u_m)^{\lambda_2}$$

for a positive constant A_{λ_1, λ_2} .

Lemma 5. Suppose (A2) and (A5). Then, for any integer $r \geq 2$, there exists a positive constant A_r such that

$$\mathbb{E}\left(\max_{1 \leq k \leq j} |X^*(t_{k-1, n}) - X^*(t_n)|\right)^{2r} \leq A_r (v_{n+1}(\kappa, \mu) - v_n(\kappa, \mu))^r. \quad (36)$$

Here, j and $(t_{k, n})_{0 \leq k \leq j}$ are as defined in (A5), and $v_n(\kappa, \mu)$ is as defined in (26).

Proof. We first show that the assumption of Lemma 4 holds with $\lambda_1 = 2r$, $\lambda_2 = r$, $m = j - 1$, $\xi_k := X^*(t_{k, n}) - X^*(t_{k-1, n})$ and $u_k := D_r^{1/r} (f(t_{k, n}) - f(t_{k-1, n}))$ for $k \in \mathbb{N}$, where D_r is the constant defined in Lemma 3. Let $0 \leq i < k \leq j - 1$. By (A5), $f(t_{k, n}) - f(t_{i, n}) = \sum_{l=i+1}^k (f(t_{l, n}) - f(t_{l-1, n})) \geq 1$. This in combination with Lemma 3 yields

$$\begin{aligned} \mathbb{E}|\xi_{i+1} + \dots + \xi_k|^{2r} &= \mathbb{E}(X^*(t_{k, n}) - X^*(t_{i, n}))^{2r} \\ &\leq D_r \max((f(t_{k, n}) - f(t_{i, n}))^r, f(t_{k, n}) - f(t_{i, n})) \\ &= D_r (f(t_{k, n}) - f(t_{i, n}))^r = \left(\sum_{l=i+1}^k u_l\right)^r, \end{aligned}$$

thereby proving that the assumption of Lemma 4 does indeed hold. Hence, inequality (36) follows from Lemma 4 and the definition of t_n :

$$\begin{aligned} &\mathbb{E}\left(\max_{1 \leq k \leq j} |X^*(t_{k-1, n}) - X^*(t_n)|\right)^{2r} \\ &= \mathbb{E}\left(\max_{1 \leq k \leq j-1} |\xi_1 + \dots + \xi_k|\right)^{2r} \\ &\leq A_{2r, r} \left(\sum_{l=1}^{j-1} u_l\right)^r = A_{2r, r} D_r (f(t_{j-1, n}) - f(t_n))^r \\ &\leq A_{2r, r} D_r (v_{n+1}(\kappa, \mu) - v_n(\kappa, \mu))^r. \quad \square \end{aligned}$$

3.2 Proof of Theorem 5

We start with a lemma and a proposition which are in essence Lemma 4.13 and Proposition 4.7 in [3]. Although our present assumption (A3) is slightly different from the corresponding assumption in [3], we have checked that the proofs of the aforementioned results in [3] go through.

Lemma 6. *Suppose (A1), (A3), (B1) and either (B21) or (B22), and let $\mu \geq 1$ be as given in (23). For sufficiently small $\delta > 0$, pick $\gamma \in (0, (\sqrt{5} - 1)/2) > 0$ satisfying $(1 + \gamma)(1 - \delta^2/8) < 1$. Then*

$$\limsup_{n \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{1}{(2a(v_n)h_0(a(v_n)))^{1/2}} \sum_{k \geq 1} \eta_k(v_n) \geq 1 - \delta \ (\leq -(1 - \delta)) \quad \text{a.s.},$$

where v_n is either τ_n or $\lfloor \tau_n \rfloor$, and $\tau_n = \tau_n(\gamma, \mu)$, with γ chosen above, is as defined in (29).

Proposition 1. *Suppose (A1), (A3), (B1) and either (B21) or (B22). Then, with $\mu \geq 1$ and $q \geq 0$ as defined in (23) and (24), respectively,*

$$\limsup_{t \rightarrow \infty} \left(\liminf_{t \rightarrow \infty} \right) \frac{X^*(t)}{(2(q + 1)a(t) \log \log a(t))^{1/2}} \geq 1 \ (\leq -1) \quad \text{a.s.}$$

and

$$\limsup_{t \rightarrow \infty} \left(\liminf_{t \rightarrow \infty} \right) \frac{X^*(t)}{(2(\mu - 1)a(t) \log a(t))^{1/2}} \geq 1 \ (\leq -1) \quad \text{a.s.}$$

in the cases $\mu = 1$ and $\mu > 1$, respectively.

Proposition 2 is a counterpart of Proposition 4.6 in [3]. Although it is tempting to believe that the proof of Proposition 4.6 in [3] goes through as well, this is not the case. First, in the proof of (43), instead of dealing with the process X and its mean b (which are not monotone anymore), we work with their nondecreasing majorants Y and f . It is not obvious that Y and f are bounded from above similarly to X and b . Second, unlike in [3] we do not require that the variance a is asymptotically nondecreasing. Hence, putting $a(t_n)$ in the denominator of (37) is not allowed.

Proposition 2. *Suppose (A1)–(A5). Then, with $\mu \geq 1$ and $q \geq 0$ as defined in (23) and (24), respectively,*

$$\limsup_{t \rightarrow \infty} \left(\liminf_{t \rightarrow \infty} \right) \frac{X^*(t)}{(2(q + 1)a(t) \log \log a(t))^{1/2}} \leq 1 \ (\geq -1) \quad \text{a.s.}$$

and

$$\limsup_{t \rightarrow \infty} \left(\liminf_{t \rightarrow \infty} \right) \frac{X^*(t)}{(2(\mu - 1)a(t) \log a(t))^{1/2}} \leq 1 \ (\geq -1) \quad \text{a.s.}$$

in the cases $\mu = 1$ and $\mu > 1$, respectively.

Proof of Proposition 2. In view of (A4), it is enough to show that, for each $\varrho \in (0, 1)$ and each positive κ sufficiently close to 0,

$$\limsup_{n \rightarrow \infty} \frac{\sup_{u \in [t_n, t_{n+1}]} X^*(u)}{(2a(s_n)h_\varrho(a(s_n)))^{1/2}} \leq 1 + \kappa \quad \text{a.s.}, \tag{37}$$

where $t_n = t_n(\kappa, \mu)$ and $s_n = s_n(\kappa, \mu)$ are as defined in (26) and (A4), respectively, $h_\varrho = g_{1, \varrho}$ if μ in (23) is equal to 1 and $h_\varrho = g_{\mu_\varrho}$ if $\mu > 1$ (see (31) for the definitions of $g_{1, \varrho}$ and g_{μ_ϱ}). Indeed, if (37) holds true, then, for large enough n , a.s.

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{X^*(t)}{(2a(t)h_\varrho(a(t)))^{1/2}} &= \limsup_{t \rightarrow \infty} \frac{X^*(t)}{(2a_0(t)h_\varrho(a_0(t)))^{1/2}} \\ &\leq \limsup_{n \rightarrow \infty} \frac{\sup_{u \in [t_n, t_{n+1}]} X^*(u)}{(2a_0(s_n)h_\varrho(a_0(s_n)))^{1/2}} = \limsup_{n \rightarrow \infty} \frac{\sup_{u \in [t_n, t_{n+1}]} X^*(u)}{(2a(s_n)h_\varrho(a(s_n)))^{1/2}} \leq 1 + \kappa. \end{aligned}$$

The relation $\liminf_{t \rightarrow \infty} \frac{X^*(t)}{(2a(t)h_\varrho(a(t)))^{1/2}} \geq -1 - \kappa$ a.s. does not require a separate proof. It follows from the argument for lim sup upon replacing $\eta_k(t)$ with $-\eta_k(t)$.

To obtain (37), we first prove in Lemma 7 that

$$\limsup_{n \rightarrow \infty} \frac{X^*(s_n)}{(2a(s_n)h_\varrho(a(s_n)))^{1/2}} \leq 1 + \kappa \quad \text{a.s.} \tag{38}$$

and then show that

$$\lim_{n \rightarrow \infty} \frac{\sup_{u \in [t_n, t_{n+1}]} |X^*(u) - X^*(s_n)|}{(a(s_n)h_\varrho(a(s_n)))^{1/2}} = 0 \quad \text{a.s.} \tag{39}$$

Lemma 7. *Suppose (A1), (A3) and (A4). Then relation (38) holds for any $\kappa \in (0, (\sqrt{5} - 1)/2)$.*

Proof. Fix any $\kappa \in (0, (\sqrt{5} - 1)/2)$. We first show that there exists $\rho = \rho(\kappa) > 0$ satisfying

$$(1 - \kappa)(1 + \kappa)^2(2 - \exp(2(1 + \kappa)\rho)) > 1. \tag{40}$$

To prove this, note that our choice of κ ensures $(1 - \kappa)(1 + \kappa)^2 > 1$. Observe next that as positive ρ approaches 0, $2 - \exp(2(1 + \kappa)\rho)$ becomes arbitrary close to 1, thereby justifying $2 - \exp(2(1 + \kappa)\rho) > (1 - \kappa)^{-1}(1 + \kappa)^{-2}$.

By Lemma 4.1 in [3], for $\vartheta \in \mathbb{R}$ and $t \geq 0$,

$$\mathbb{E} \exp(\vartheta X^*(t)) \leq \exp(2^{-1}\vartheta^2 \exp(|\vartheta|)a(t)).$$

Fix any $\theta \in \mathbb{R}$ and put $\vartheta = \theta/(2a(s_n)h_\varrho(a(s_n)))^{1/2}$. Observe that, for large enough n , $h_\varrho(a(s_n))/a(s_n) \leq 2\rho^2$ for ρ satisfying (40). An application of Markov's inequality then yields, for large n as above,

$$\begin{aligned} \mathbb{P} \left\{ \frac{X^*(s_n)}{(2a(s_n)h_\varrho(a(s_n)))^{1/2}} > 1 + \kappa \right\} &\leq e^{-(1+\kappa)\theta} \mathbb{E} \exp \left(\theta \frac{X^*(s_n)}{(2a(s_n)h_\varrho(a(s_n)))^{1/2}} \right) \\ &\leq \exp \left(-(1 + \kappa)\theta + \frac{\theta^2}{4h_\varrho(a(s_n))} \exp \left(\frac{\rho|\theta|}{h_\varrho(a(s_n))} \right) \right). \end{aligned}$$

Putting $\theta = 2(1 + \kappa)h_\varrho(a(s_n))$ and then invoking Lemma 2(a), we obtain

$$\begin{aligned} \mathbb{P} \left\{ \frac{X^*(s_n)}{(2a(s_n)h_\varrho(a(s_n)))^{1/2}} > 1 + \kappa \right\} &\leq \exp(-(1 + \kappa)^2(2 - \exp(2(1 + \kappa)\rho))h_\varrho(a(s_n))) \\ &= O \left(\frac{1}{n^{(1-\kappa)(1+\kappa)^2(2 - \exp(2(1+\kappa)\rho))}} \right), \quad n \rightarrow \infty. \end{aligned}$$

According to (40),

$$\sum_{n \geq n_0} \mathbb{P} \left\{ \frac{X^*(t_n)}{(2a(t_n)h_\varrho(a(t_n)))^{1/2}} > 1 + \kappa \right\} < \infty$$

for some $n_0 \in \mathbb{N}$ large enough. An application of the Borel–Cantelli lemma completes the proof of Lemma 7. \square

Next, in order to prove (39) it suffices to show that

$$\lim_{n \rightarrow \infty} \frac{\sup_{u \in [t_n, t_{n+1}]} |X^*(u) - X^*(t_n)|}{(a(s_n))^{1/2}} = 0 \quad \text{a.s.} \tag{41}$$

and

$$\lim_{n \rightarrow \infty} \frac{|X^*(t_n) - X^*(s_n)|}{(a(s_n))^{1/2}} = 0 \quad \text{a.s.} \tag{42}$$

Since (42) is a consequence of (41), we are left with proving (41).

PROOF OF (41). Let $t_n = t_{0,n} < \dots < t_{j,n} = t_{n+1}$ be a partition defined in (A5). With this at hand, write

$$\begin{aligned} & \sup_{u \in [t_n, t_{n+1}]} |X^*(u) - X^*(t_n)| \\ &= \max_{1 \leq k \leq j} \sup_{v \in [0, t_{k,n} - t_{k-1,n}]} \left| (X^*(t_{k-1,n}) - X^*(t_n)) + (X^*(t_{k-1,n} + v) - X^*(t_{k-1,n})) \right| \\ &\leq \max_{1 \leq k \leq j} |X^*(t_{k-1,n}) - X^*(t_n)| \\ &\quad + \max_{1 \leq k \leq j} \sup_{v \in [0, t_{k,n} - t_{k-1,n}]} |X^*(t_{k-1,n} + v) - X^*(t_{k-1,n})| \quad \text{a.s.} \end{aligned}$$

By Markov’s inequality and Lemma 5, for any $r > 0$ and all $\varepsilon > 0$,

$$\begin{aligned} & \mathbb{P} \left\{ \max_{1 \leq k \leq j} |X^*(t_{k-1,n}) - X^*(t_n)| > \varepsilon (a(s_n))^{1/2} \right\} \\ &\leq \frac{\mathbb{E}(\max_{1 \leq k \leq j} |X^*(t_{k-1,n}) - X^*(t_n)|)^{2r}}{\varepsilon^{2r} (a(s_n))^r} \leq \frac{A_r (v_{n+1}(\kappa, \mu) - v_n(\kappa, \mu))^r}{\varepsilon^{2r} (a(s_n))^r}. \end{aligned}$$

By Lemma 2(b), there exists an integer $r \geq 2$ such that the right-hand side forms a sequence which is summable in n . Hence, an application of the Borel–Cantelli lemma yields

$$\lim_{n \rightarrow \infty} \frac{\max_{1 \leq k \leq j} |X^*(t_{k-1,n}) - X^*(t_n)|}{(a(s_n))^{1/2}} = 0 \quad \text{a.s.}$$

Next, we work towards proving that

$$\lim_{n \rightarrow \infty} \frac{\max_{1 \leq k \leq j} \sup_{v \in [0, t_{k,n} - t_{k-1,n}]} |X^*(t_{k-1,n} + v) - X^*(t_{k-1,n})|}{(a(s_n))^{1/2}} = 0 \quad \text{a.s.} \tag{43}$$

According to (A2), for any $0 \leq s < t$, $|X(t) - X(s)| \leq Y(t) - Y(s)$ a.s., where the process Y is a.s. nondecreasing. Taking into account Remark 4 we obtain

$$\sup_{v \in [0, t_{k,n} - t_{k-1,n}]} |X^*(t_{k-1,n} + v) - X^*(t_{k-1,n})|$$

$$\begin{aligned}
 &\leq \sup_{v \in [0, t_k, n - t_{k-1}, n]} |X(t_{k-1, n} + v) - X(t_{k-1, n})| \\
 &\quad + \sup_{v \in [0, t_k, n - t_{k-1}, n]} |b(t_{k-1, n} + v) - b(t_{k-1, n})| \\
 &\leq \sup_{v \in [0, t_k, n - t_{k-1}, n]} (Y(t_{k-1, n} + v) - Y(t_{k-1, n})) \\
 &\quad + \sup_{v \in [0, t_k, n - t_{k-1}, n]} (f(t_{k-1, n} + v) - f(t_{k-1, n})) \\
 &= Y(t_k, n) - Y(t_{k-1, n}) + f(t_k, n) - f(t_{k-1, n}) \quad \text{a.s.}
 \end{aligned}$$

By (A1) and (A5),

$$\frac{\max_{1 \leq k \leq j} (f(t_k, n) - f(t_{k-1, n}))}{(a(s_n))^{1/2}} \leq \frac{A}{(a(s_n))^{1/2}} \rightarrow 0, \quad n \rightarrow \infty. \tag{44}$$

Finally, for all $\varepsilon > 0$,

$$\begin{aligned}
 &\mathbb{P} \left\{ \max_{1 \leq k \leq j} (Y(t_k, n) - Y(t_{k-1, n})) > \varepsilon (a(s_n))^{1/2} \right\} \\
 &\leq \sum_{k=1}^j \mathbb{P} \{ Y(t_k, n) - Y(t_{k-1, n}) > \varepsilon (a(s_n))^{1/2} \} \\
 &\leq e^{-\varepsilon (a(s_n))^{1/2}} \sum_{k=1}^j \mathbb{E} e^{Y(t_k, n) - Y(t_{k-1, n})} \\
 &\leq e^{-\varepsilon (a(s_n))^{1/2}} \sum_{k=1}^j \exp((e^M - 1)(f(t_k, n) - f(t_{k-1, n}))) \\
 &\leq \exp(A(e^M - 1)) j e^{-\varepsilon (a(s_n))^{1/2}}, \tag{45}
 \end{aligned}$$

having utilized Markov’s inequality for the second inequality, Lemma 1 for the third and (A5) for the fourth. Invoking (A5) once again we conclude that the right-hand side is summable in n . Hence, an application of the Borel–Cantelli lemma yields

$$\lim_{n \rightarrow \infty} \frac{\max_{1 \leq k \leq j} (X(t_k, n) - X(t_{k-1, n}))}{(a(s_n))^{1/2}} = 0 \quad \text{a.s.}$$

The proofs of both (41) and Proposition 2 are complete. □

4 Proofs related to Karlin’s occupancy scheme

4.1 Auxiliary results

For ease of reference, we state two known results. The former is an obvious extension of Theorem 1.5.3 in [2]. The latter is Lemma 6.2 in [3].

Lemma 8. *Let f be a function which varies regularly at ∞ of positive index and g a positive nondecreasing function with $\lim_{t \rightarrow \infty} g(t) = \infty$. Then there exists a nondecreasing function h satisfying $f(g(t)) \sim h(t)$ as $t \rightarrow \infty$.*

Lemma 9. (a) Conditions (2) and (3) entail

$$\rho(t) \sim (\beta + 1)^{-1} (\log t)^{\beta+1} l(\log t), \quad t \rightarrow \infty. \tag{46}$$

(b) Conditions (2) and (6) entail

$$\rho(t) \sim (\sigma\lambda)^{-1} \exp(\sigma(\log t)^\lambda) (\log t)^{1-\lambda}, \quad t \rightarrow \infty. \tag{47}$$

4.2 Asymptotic behavior of $\mathbb{E}K_j(t)$ and $\text{Var} K_j(t)$

Given next is a collection of results on the asymptotics of $\mathbb{E}K_j(t)$ and $\text{Var} K_j(t)$ taken from Lemma 6.5 of [3]. Recall that $\Pi_{\ell, \infty}$ denotes the subclass of the de Haan class Π with the auxiliary functions ℓ , see (2), satisfying $\lim_{t \rightarrow \infty} \ell(t) = \infty$.

Lemma 10. Assume that $\rho \in \Pi_{\ell, \infty}$. Then, for each $j \in \mathbb{N}$,

$$\mathbb{E}K_j(t) \sim \rho(t), \quad t \rightarrow \infty, \tag{48}$$

and

$$\text{Var} K_j(t) \sim \left(\log 2 - \sum_{k=1}^{j-1} \frac{(2k-1)!}{(k!)^2 2^{2k}} \right) \ell(t), \quad t \rightarrow \infty.$$

Assume that $\rho(t) \sim t^\alpha L(t)$ as $t \rightarrow \infty$ for some $\alpha \in (0, 1]$ and some L slowly varying at ∞ . If $\alpha \in (0, 1)$ and $j \in \mathbb{N}$ or $\alpha = 1$ and $j \geq 2$, then, as $t \rightarrow \infty$,

$$\mathbb{E}K_j(t) \sim \frac{\Gamma(j-\alpha)}{(j-1)!} \rho(t), \tag{49}$$

and

$$\lim_{t \rightarrow \infty} \frac{\text{Var} K_j(t)}{\rho(t)} = \left(\sum_{i=0}^{j-1} \frac{\Gamma(i+j-\alpha)}{i!(j-1)!2^{i+j-1-\alpha}} - \frac{\Gamma(j-\alpha)}{(j-1)!} \right) > 0. \tag{50}$$

If $\alpha = 1$, then

$$\text{Var} K_1(t) \sim \mathbb{E}K_1(t) \sim t \hat{L}(t), \quad t \rightarrow \infty. \tag{51}$$

4.3 Asymptotic behavior of $\mathbb{E}K_j^*(t)$ and $\text{Var} K_j^*(t)$

For $j \in \mathbb{N}$, the asymptotics of $t \mapsto \mathbb{E}K_j^*(t)$ as stated in Theorems 1, 2 and 3 can be found in Lemma 6.5 of [3]. Next, we show that, for $j \in \mathbb{N}$, the functions $t \mapsto \text{Var} K_j^*(t)$ exhibit the asymptotics given in the aforementioned theorems.

Lemma 11. Assume that $\rho \in \Pi_{\ell, \infty}$. Then, for each $j \in \mathbb{N}$,

$$\text{Var} K_j^*(t) \sim \left(\frac{1}{j} - \frac{(2j-1)!}{(j!)^2 2^{2j}} \right) \ell(t), \quad t \rightarrow \infty. \tag{52}$$

Assume that $\rho(t) \sim t^\alpha L(t)$ as $t \rightarrow \infty$ for some $\alpha \in (0, 1]$ and some L slowly varying at ∞ . If $\alpha \in (0, 1)$ and $j \in \mathbb{N}$ or $\alpha = 1$ and $j \geq 2$, then, as $t \rightarrow \infty$,

$$\lim_{t \rightarrow \infty} \frac{\text{Var} K_j^*(t)}{t^\alpha L(t)} = c_{j, \alpha} > 0 \tag{53}$$

with $c_{j, \alpha}$ as defined in (15) and (17).

If $\alpha = 1$, then

$$\text{Var} K_1^*(t) \sim t \hat{L}(t), \quad t \rightarrow \infty. \tag{54}$$

Proof. Assume that $\rho \in \Pi_{\ell, \infty}$. Putting $u = v = 0$ in formula (11) of [11] we obtain (52).

According to formula (6) in [9],

$$\text{Var } K_j^*(t) = \mathbb{E}K_j^*(t) - 2^{-2j} \binom{2j}{j} \mathbb{E}K_{2j}^*(2t), \quad t \geq 0, j \in \mathbb{N}. \tag{55}$$

Note that (55) does not require even regular variation assumption on ρ .

Assume that ρ is regularly varying at ∞ of index $\alpha = 1$. We first discuss the properties of the function \hat{L} stated in Theorem 3. By Lemma 3 in [12], $\lim_{t \rightarrow \infty} t^{-1} \rho(t) = 0$ and $\int_1^\infty y^{-2} \rho(y) dy \leq 1$. This implies that the function $\hat{L}(t) = \int_t^\infty y^{-1} L(y) dy$ is well-defined for large t and thereupon $\lim_{t \rightarrow \infty} \hat{L}(t) = 0$. According to Proposition 1.5.9b [2], \hat{L} is slowly varying at ∞ and satisfies (19). This in combination with (16), (22) and (55) entails (54).

Assume now $\alpha \in (0, 1)$ and $j \in \mathbb{N}$ or $\alpha = 1$ and $j \geq 2$. Then invoking (55) and either (13) or (16) we obtain

$$\lim_{t \rightarrow \infty} \frac{\text{Var } K_j^*(t)}{t^\alpha L(t)} = \lim_{t \rightarrow \infty} \frac{\mathbb{E}K_j^*(t)}{t^\alpha L(t)} - 2^{-2j} \binom{2j}{j} \lim_{t \rightarrow \infty} \frac{\mathbb{E}K_{2j}^*(2t)}{t^\alpha L(t)} = c_{j, \alpha}.$$

We are left with showing that the constants $c_{j, \alpha}$ are positive for $\alpha \in (0, 1)$ and $j \in \mathbb{N}$ and $\alpha = 1$ and $j \geq 2$ or equivalently

$$\frac{2^\alpha \Gamma(2j - \alpha)}{2^{2j} j! \Gamma(j - \alpha)} < 1.$$

This is a consequence of

$$\frac{2^\alpha \Gamma(2j - \alpha)}{2^{2j} j! \Gamma(j - \alpha)} < \frac{2(2j - 1)!}{2^{2j} j! (j - 1)!} = \frac{(2j - 1)!}{(2j)!! (2j - 2)!!} < 1,$$

where $(2n)!! := 2 \cdot 4 \cdot \dots \cdot (2n)$ for $n \in \mathbb{N}$. Here, the last inequality is justified with the help of mathematical induction. The proof of Lemma 11 is complete. \square

4.4 Proof of Theorems 1, 2 and 3

For $k \in \mathbb{N}$ and $t \geq 0$, denote by $\pi_k(t)$ the number of balls in box k at time t in the Poissonized version. It has already been mentioned in Section 1.1 that the thinning property of Poisson processes implies that the processes $(\pi_1(t))_{t \geq 0}$, $(\pi_2(t))_{t \geq 0}$, \dots are independent. Moreover, for $k \in \mathbb{N}$, $(\pi_k(t))_{t \geq 0}$ is a Poisson process with intensity p_k . As a consequence, both $K_j^*(t)$ and $K_j(t)$ can be represented as the sums of independent indicators

$$K_j^*(t) = \sum_{k=1} \mathbb{1}_{\{\pi_k(t)=j\}} \quad \text{and} \quad K_j(t) = \sum_{k=1} \mathbb{1}_{\{\pi_k(t) \geq j\}}, \quad t \geq 0, j \in \mathbb{N}.$$

Hence, it is reasonable to prove the desired LILs for the small counts by applying Theorem 5.

As a preparation, we start with a lemma which facilitates checking condition (B22) of Theorem 5.

Lemma 12. Assume that either $\rho \in \Pi_{\ell, \infty}$ or ρ is regularly varying at ∞ of index $\alpha \in (0, 1]$. If $\rho \in \Pi_{\ell, \infty}$ and $j \in \mathbb{N}$ or $\alpha \in (0, 1)$ and $j \in \mathbb{N}$ or $\alpha = 1$ and $j \geq 2$, then for any positive functions c and d satisfying $\lim_{t \rightarrow \infty} c(t) = \infty$, $\lim_{t \rightarrow \infty} (c(t)/t) = 0$ and $\lim_{t \rightarrow \infty} (d(t)/t) = \infty$,

$$\text{Var} \left(\sum_{k \geq 1} \mathbb{1}_{\{c(t) < 1/p_k \leq d(t)\}} \mathbb{1}_{\{\pi_k(t) = j\}} \right) \sim \text{Var } K_j^*(t), \quad t \rightarrow \infty.$$

Proof. We start by proving a simple but an important inequality. Since

$$\begin{aligned} \text{Cov}(\mathbb{1}_{\{\pi_k(t) \geq j\}}, \mathbb{1}_{\{\pi_k(t) \geq j+1\}}) \\ = \mathbb{P}\{\pi_k(t) \geq j+1\} - \mathbb{P}\{\pi_k(t) \geq j\} \mathbb{P}\{\pi_k(t) \geq j+1\} \geq 0, \end{aligned}$$

we infer

$$\begin{aligned} \text{Var}(\mathbb{1}_{\{\pi_k(t) = j\}}) \\ = \text{Var}(\mathbb{1}_{\{\pi_k(t) \geq j\}} - \mathbb{1}_{\{\pi_k(t) \geq j+1\}}) \\ = \text{Var}(\mathbb{1}_{\{\pi_k(t) \geq j+1\}}) + \text{Var}(\mathbb{1}_{\{\pi_k(t) \geq j\}}) - 2\text{Cov}(\mathbb{1}_{\{\pi_k(t) \geq j\}}, \mathbb{1}_{\{\pi_k(t) \geq j+1\}}) \\ \leq \text{Var } \mathbb{1}_{\{\pi_k(t) \geq j+1\}} + \text{Var } \mathbb{1}_{\{\pi_k(t) \geq j\}}. \end{aligned}$$

Therefore, it is enough to show that in the setting of the lemma, for all $j \geq 2$ in the case $\alpha = 1$ and for all $j \in \mathbb{N}$ in the other cases,

$$\text{Var} \left(\sum_{k \geq 1} \mathbb{1}_{\{1/p_k > d(t)\}} \mathbb{1}_{\{\pi_k(t) \geq j\}} \right) = o(\text{Var } K_j^*(t)), \quad t \rightarrow \infty, \tag{56}$$

and

$$\text{Var} \left(\sum_{k \geq 1} \mathbb{1}_{\{1/p_k \leq c(t)\}} \mathbb{1}_{\{\pi_k(t) \geq j\}} \right) = o(\text{Var } K_j^*(t)), \quad t \rightarrow \infty. \tag{57}$$

According to formulae (86), (87), (79) and (80) in [3],

$$\text{Var} \left(\sum_{k \geq 1} \mathbb{1}_{\{1/p_k > d(t)\}} \mathbb{1}_{\{\pi_k(t) \geq j\}} \right) = o(\ell(t)), \quad t \rightarrow \infty, \tag{58}$$

$$\text{Var} \left(\sum_{k \geq 1} \mathbb{1}_{\{1/p_k \leq c(t)\}} \mathbb{1}_{\{\pi_k(t) \geq j\}} \right) = o(\ell(t)), \quad t \rightarrow \infty, \tag{59}$$

$$\text{Var} \left(\sum_{k \geq 1} \mathbb{1}_{\{1/p_k > d(t)\}} \mathbb{1}_{\{\pi_k(t) \geq j\}} \right) = o(\rho(t)), \quad t \rightarrow \infty \tag{60}$$

and

$$\text{Var} \left(\sum_{k \geq 1} \mathbb{1}_{\{1/p_k \leq c(t)\}} \mathbb{1}_{\{\pi_k(t) \geq j\}} \right) = o(\rho(t)), \quad t \rightarrow \infty. \tag{61}$$

In view of (52) or (53), depending on the setting, relations (58) or (60), (59) or (61) are equivalent to (56) and (57). □

Proof of Theorems 1, 2 and 3. We first prove Theorem 3 in the case $j = 1$. This setting is much simpler than the others, as the LIL for

$$K_1^*(t) = K_1(t) - K_2(t) \tag{62}$$

can be derived from the already available LILs for $K_1(t)$ and $K_2(t)$.

The statements of Theorem 3 concerning the function \hat{L} has already been justified in the proof of Lemma 11. According to (51) and (54), $\text{Var } K_1^*(t) \sim \text{Var } K_1(t) \sim t\hat{L}(t)$ as $t \rightarrow \infty$. Invoking the latter relation, (50) and (19) we conclude that $\text{Var } K_2(t) \sim 2^{-1}tL(t) = o(\text{Var } K_1(t))$ as $t \rightarrow \infty$. By Theorem 3.4 and Remark 1.7 in [3],

$$\limsup_{t \rightarrow \infty} \left(\liminf_{t \rightarrow \infty} \right) \frac{K_2(t) - \mathbb{E}K_2(t)}{(\text{Var } K_2(t) \log \log \text{Var } K_2(t))^{1/2}} = 2^{1/2} (-2^{1/2}) \quad \text{a.s.}$$

As a consequence, $K_2(t) - \mathbb{E}K_2(t) = o((\text{Var } K_1(t) \log \log \text{Var } K_1(t))^{1/2})$ a.s. as $t \rightarrow \infty$. Now, in view of (62),

$$\begin{aligned} \limsup_{t \rightarrow \infty} \left(\liminf_{t \rightarrow \infty} \right) \frac{K_1^*(t) - \mathbb{E}K_1^*(t)}{(\text{Var } K_1^*(t) \log \log \text{Var } K_1^*(t))^{1/2}} \\ = \limsup_{t \rightarrow \infty} \left(\liminf_{t \rightarrow \infty} \right) \frac{K_1(t) - \mathbb{E}K_1(t)}{(\text{Var } K_1(t) \log \log \text{Var } K_1(t))^{1/2}} \quad \text{a.s.} \end{aligned}$$

Armed with this, the claim of Theorem 3 in the case $j = 1$ is secured by Theorem 3.4 in [3]. Indeed, the theorem states that depending on whether relation (18) holds or not, the right-hand side is either equal to $2^{1/2} (-2^{1/2})$ or is not larger than $2^{1/2}$ (not smaller than $-2^{1/2}$) a.s.

In the remaining part of the proof we treat simultaneously Theorems 1 and 2 and the case $j \geq 2$ of Theorem 3. It has already been announced that our plan is to derive the LILs from Theorem 5. Hence, now we work towards checking the conditions of the aforementioned theorem in the present setting.

CONDITION (A1) holds according to (52) in conjunction with $\lim_{t \rightarrow \infty} \ell(t) = \infty$, (53) and (54).

CONDITION (A2) is justified by a representation $\mathbb{1}_{\{\pi_k(t)=j\}} = \mathbb{1}_{\{\pi_k(t) \geq j\}} - \mathbb{1}_{\{\pi_k(t) \geq j+1\}}$ a.s., for all $k, j \in \mathbb{N}$ and $t \geq 0$, see Remark 5. The corresponding function f is given by

$$\begin{aligned} f_{j, \alpha}(t) &:= \sum_{k \geq 1} (\mathbb{P}\{\pi_k(t) \geq j\} + \mathbb{P}\{\pi_k(t) \geq j + 1\}) = \mathbb{E}K_j(t) + \mathbb{E}K_{j+1}(t) \\ &= \sum_{k \geq 1} \left(1 - \sum_{i=0}^{j-1} e^{-p_k t} \frac{(p_k t)^i}{i!} \right) + \sum_{k \geq 1} \left(1 - \sum_{i=0}^j e^{-p_k t} \frac{(p_k t)^i}{i!} \right). \tag{63} \end{aligned}$$

We bring out the dependence on j and α to distinguish the so defined functions for the different settings. By the same reasoning, we write $a_{j, \alpha}$ instead of a , where $a(t) = \text{Var } K_j^*(t)$ for $t \geq 0$.

CONDITION (A3). Assume first that $\rho \in \Pi_{\ell, \infty}$. According to (48) and (52), for each $j \in \mathbb{N}$, $f_{j,0}(t) \sim 2\rho(t)$ and $a_{j,0}(t) \sim C\ell(t)$ as $t \rightarrow \infty$, respectively. Here and hereafter, C denotes a constant whose value is of no importance and may vary from formula to formula. Under (3), invoking (46) we conclude that (A3) holds with $\mu = 1/\beta + 1$. Under (6), using (47) we infer $\mu = 1$. Thus, we have to check the additional conditions pertaining to the case $\mu = 1$. First, the function $f_{j,\alpha}$ is continuous. Second, $q = 1/\lambda - 1$ and $\mathcal{L}(t) \equiv 1$ for all $t \geq 0$ by another appeal to (47).

Assume now that $\alpha \in (0, 1)$ and $j \in \mathbb{N}$ or $\alpha = 1$ and $j \geq 2$. Then, according to (49) and (53), $f_{j,\alpha}(t) \sim Ca_{j,\alpha}(t)$ as $t \rightarrow \infty$, which entails $\mu = 1$. Further, $f_{j,\alpha}$ is continuous, $q = 0$ and $\mathcal{L}(t) \equiv 1$ for $t \geq 0$.

CONDITION (A4). Denote by $a_{0;j,\alpha}$ a version of a_0 for the different settings. Assume first that $\rho \in \Pi_{\ell, \infty}$. Then, according to (52), $a_{j,0}(t) \sim C\ell(t)$ as $t \rightarrow \infty$. Therefore, under (3), $a_{0;j,0}$ can be chosen as a monotone equivalent of $t \mapsto C(\log t)^\beta l(\log t)$ which exists by Lemma 8. Under (6), $a_{0;j,0}$ can be chosen as $a_{0;j,0}(t) := C \exp(\sigma(\log t)^\lambda)$ for all $t \geq 1$.

Assume now that $\alpha \in (0, 1)$ and $j \in \mathbb{N}$ or $\alpha = 1$ and $j \geq 2$. Then, according to (53), $a_{0;j,\alpha}$ can be chosen as a monotone equivalent of $t \mapsto c_{j,\alpha} t^\alpha L(t)$ which exists by Lemma 8.

Thus, in all settings (A4) holds according to Remark 7.

CONDITION (A5) holds according to Remark 8, for $f_{j,\alpha}$ is continuous and strictly increasing.

CONDITION (B1) holds in view of

$$a_{j,\alpha}(t) = \text{Var } K_j^*(t) = \sum_{k \geq 1} e^{-p_k t} \frac{(p_k t)^j}{j!} \left(1 - e^{-p_k t} \frac{(p_k t)^j}{j!} \right), \quad t \geq 0, \quad (64)$$

which shows that $a_{j,\alpha}$ is a continuous function.

CONDITION (B22). For $t > 1$, put $c(t) := t / \log t$ and $d(t) := t \log t$ and then

$$R_0(t) := \{k \in \mathbb{N} : c(t) < 1/p_k \leq d(t)\}.$$

By Lemma 12, in all settings relation (30), which is the second part of (B22), holds.

Passing to the first part of (B22), we are going to refer to the table below which contains all the necessary information. In the first line, we list the values of μ which have already been found while checking (A3). Recall that the definitions of $w_n(\gamma, \mu)$ and τ_n can be found right after formula (29) and in (29), respectively. We write w_n instead of $w_n(\gamma, \mu)$ and **Set.** is a shorthand for **Setting**. Note that in the case $\alpha = 1$ we only consider $j \geq 2$.

Set.	$\rho \in \Pi_{\ell, \infty}$, (3)	$\rho \in \Pi_{\ell, \infty}$, (6)	$\alpha \in (0, 1]$, $j \in \mathbb{N}$
μ	$1/\beta + 1$	1	1
w_n	$n^{\beta(1+\gamma)}$	$\exp(n^{\lambda(1+\gamma)})$	$\exp(n^{1+\gamma})$
$\tau_n \sim$	$e^{n^{(1+\gamma)}} o(e^{n^{(1+\gamma)}})$	$e^{\sigma^{-1/\lambda} n^{(1+\gamma)}} (1 + o(1))$	$e^{\alpha^{-1} n^{(1+\gamma)}} o(e^{\alpha^{-1} n^{(1+\gamma)}})$

We conclude that in all the settings τ_{n+1}/τ_n diverges to ∞ superexponentially fast, whereas $\log \tau_n$ only grows polynomially fast. Hence, for large enough n , $c(\tau_{n+1}) > d(\tau_n)$, which justifies the first part of (B22).

The proofs of Theorems 1, 2 and 3 are complete. □

4.5 Proofs of Theorem 4

We start with some preparatory work. It is known, see, for instance, Lemma 1 in [9], that for any probability distribution $(p_k)_{k \in \mathbb{N}}$ and $j \in \mathbb{N}$,

$$\lim_{n \rightarrow \infty} |\mathbb{E}K_j^*(n) - \mathbb{E}K_j^*(n)| = 0. \tag{65}$$

However, we are not aware of a counterpart of this relation for variances. Proposition 3 fills up this gap. Recall that $\rho \in \Pi_{\ell, \infty}$ means that $\rho \in \Pi$ and that its auxiliary function ℓ , see (2), satisfies $\lim_{t \rightarrow \infty} \ell(t) = \infty$.

Proposition 3. *Assume that either $\rho \in \Pi_{\ell, \infty}$ or ρ is regularly varying at ∞ of index $\alpha \in (0, 1]$. Then, for $j \in \mathbb{N}$,*

$$\lim_{n \rightarrow \infty} \frac{\text{Var } K_j^*(n)}{\text{Var } K_j^*(n)} = 1.$$

The proof of Proposition 3 is partly based on a lemma which is a slight extension of Lemma 6.9 in [3]. The new aspect of the lemma is that unlike the cited result it covers the case where $j = 1$ and $l \geq 1$ simultaneously.

Lemma 13. *Assume that either $\rho \in \Pi_{\ell, \infty}$ or ρ is regularly varying at ∞ of index $\alpha \in (0, 1]$. Then for $l \geq j$, $l, j \in \mathbb{N}$,*

$$\sum_{k \geq 1} \binom{n}{l} p_k^l (1 - p_k)^n = O(\text{Var } K_j^*(n)), \quad n \rightarrow \infty. \tag{66}$$

Proof. According to the last formula in the proof of Lemma 6.9 in [3],

$$\sum_{k \geq 1} \binom{n}{l} p_k^l (1 - p_k)^n \sim \mathbb{E}K_l^*(n + l), \quad n \rightarrow \infty.$$

According to formulae (9), (13), (16) and (22), the function $t \mapsto \mathbb{E}K_j^*(t)$ is regularly varying at ∞ of index $\alpha \in [0, 1]$. This entails $\mathbb{E}K_l^*(n + l) \sim \mathbb{E}K_l^*(n)$ as $n \rightarrow \infty$.

If $\rho \in \Pi_{\ell, \infty}$ or ρ is regularly varying at ∞ of index $\alpha \in (0, 1)$, then (66) is a consequence of (9) and (10) or (13) and (14). If ρ is regularly varying at ∞ of index $\alpha = 1$ and either $j, l \geq 2$ or $j = l = 1$, then (66) follows from (16) and (17) or (22), respectively. Finally, under the latter regular variation assumption, if $j = 1$ and $l \geq 2$, then (66), with o replacing O , holds true according to (16), (22) and (19). □

Proof of Proposition 3. We start by noting that, in view of (52), (53) or (54), for $j \in \mathbb{N}$,

$$\lim_{t \rightarrow \infty} \frac{\text{Var } K_j^*(t)}{t} = 0. \tag{67}$$

In the case $\alpha = 1$ this is secured by $\lim_{t \rightarrow \infty} \hat{L}(t) = 0$, which follows from the definition of \hat{L} .

For $k, j, n \in \mathbb{N}$, the event {the box k contains exactly j balls out of n } will be denoted by $A_k(j, n)$. Then

$$\begin{aligned} \text{Var } \mathcal{K}_j^*(n) &= \sum_{k \geq 1} \mathbb{P}(A_k(j, n))(1 - \mathbb{P}(A_k(j, n))) \\ &\quad + \sum_{i \neq k} (\mathbb{P}(A_i(j, n) \cap A_k(j, n)) - \mathbb{P}(A_i(j, n))\mathbb{P}(A_k(j, n))). \end{aligned}$$

It is enough to prove that

$$\lim_{n \rightarrow \infty} \frac{\sum_{k \geq 1} \mathbb{P}(A_k(j, n))(1 - \mathbb{P}(A_k(j, n))) - \text{Var } K_j^*(n)}{\text{Var } K_j^*(n)} = 0 \tag{68}$$

and

$$\lim_{n \rightarrow \infty} \frac{\sum_{i \neq k} (\mathbb{P}(A_i(j, n) \cap A_k(j, n)) - \mathbb{P}(A_i(j, n))\mathbb{P}(A_k(j, n)))}{\text{Var } K_j^*(n)} = 0. \tag{69}$$

PROOF OF (68). For $k, j, n \in \mathbb{N}$,

$$\mathbb{P}(A_k(j, n)) = \binom{n}{j} p_k^j (1 - p_k)^{n-j}.$$

In view of this and (64), the numerator in (68) is equal to

$$\begin{aligned} \sum_{k \geq 1} \left(\binom{n}{j} p_k^j (1 - p_k)^{n-j} - e^{-p_k n} \frac{(p_k n)^j}{j!} \right. \\ \left. - \left(\binom{n}{j} p_k^j (1 - p_k)^{n-j} \right)^2 - \left(e^{-p_k n} \frac{(p_k n)^j}{j!} \right)^2 \right). \end{aligned}$$

According to the penultimate inequality in the proof of Lemma 2.13 in [11], for large enough n and any $j \leq n$,

$$-B_j p_k \leq \binom{n}{j} p_k^j (1 - p_k)^{n-j} - e^{-p_k n} \frac{(p_k n)^j}{j!} \leq A_j p_k$$

for some positive constants A_j and B_j . Therefore,

$$\sum_{k \geq 1} \left| \binom{n}{j} p_k^j (1 - p_k)^{n-j} - e^{-p_k n} \frac{(p_k n)^j}{j!} \right| \leq \max(A_j, B_j) = o(\text{Var } K_j^*(n)),$$

as $n \rightarrow \infty$, since under our assumptions $\lim_{n \rightarrow \infty} \text{Var } K_j^*(n) = \infty$. Further, write

$$\sum_{k \geq 1} \left| \left(\binom{n}{j} p_k^j (1 - p_k)^{n-j} \right)^2 - \left(e^{-p_k n} \frac{(p_k n)^j}{j!} \right)^2 \right|$$

$$\begin{aligned}
 &= \sum_{k \geq 1} \left| \binom{n}{j} p_k^j (1-p_k)^{n-j} - e^{-p_k n} \frac{(p_k n)^j}{j!} \right| \left| \binom{n}{j} p_k^j (1-p_k)^{n-j} + e^{-p_k n} \frac{(p_k n)^j}{j!} \right| \\
 &\leq 2 \sum_{k \geq 1} \left| \binom{n}{j} p_k^j (1-p_k)^{n-j} - e^{-p_k n} \frac{(p_k n)^j}{j!} \right| = o(\text{Var } K_j^*(n)), \quad n \rightarrow \infty.
 \end{aligned}$$

The proof of (68) is complete.

PROOF OF (69). For $k, i, j, n \in \mathbb{N}$,

$$\begin{aligned}
 &\mathbb{P}(A_i(j, n) \cap A_k(j, n)) - \mathbb{P}(A_i(j, n))\mathbb{P}(A_k(j, n)) \\
 &= \binom{n}{j} \binom{n-j}{j} p_i^j p_k^j (1-p_i-p_k)^{n-2j} - \binom{n}{j} \binom{n}{j} p_i^j p_k^j (1-p_i)^{n-j} (1-p_k)^{n-j} \\
 &=: C_j(i, k, n).
 \end{aligned}$$

We shall use an appropriate decomposition of C_j

$$\begin{aligned}
 C_j(i, k, n) &= \binom{n}{j} \binom{n-j}{j} p_i^j p_k^j ((1-p_i-p_k)^{n-2j} - (1-p_i)^{n-j} (1-p_k)^{n-j}) \\
 &\quad - \binom{n}{j} \left(\binom{n}{j} - \binom{n-j}{j} \right) p_i^j p_k^j (1-p_i)^{n-j} (1-p_k)^{n-j} \\
 &=: C_j^{(1)}(i, k, n) + C_j^{(2)}(i, k, n).
 \end{aligned}$$

To analyze $C_j^{(1)}$ we argue as in the proof of Lemma 1 on p. 152 in [9]. Invoking an expansion

$$(x-y)^m = x^m + O(mx^{m-1}y), \quad m \rightarrow \infty,$$

which holds for positive x and y , $x > y$, with $x = (1-p_i)(1-p_k)$, $y = p_i p_k$ and $m = n - 2j$, we infer

$$\begin{aligned}
 C_j^{(1)}(i, k, n) &= \binom{n}{j} \binom{n-j}{j} p_i^j p_k^j ((1-p_i)^{n-2j} (1-p_k)^{n-2j} (1 - (1-p_i)^j (1-p_k)^j) \\
 &\quad + O((n-2j)p_i p_k (1-p_i)^{n-2j-1} (1-p_k)^{n-2j-1})) \\
 &=: F_j(i, k, n) + G_j(i, k, n).
 \end{aligned}$$

Next, we intend to show that the contributions of $F_j(i, k, n)$, $G_j(i, k, n)$ and $C_j^{(2)}(i, k, n)$ to the sum are negligible in comparison to $\text{Var } K_j^*(n)$ as $n \rightarrow \infty$.

ANALYSIS OF G_j . With Lemma 13 at hand, we obtain

$$\begin{aligned}
 &\sum_{i \neq k} \binom{n}{j} \binom{n-j}{j} (n-2j) p_i^{j+1} p_k^{j+1} (1-p_i)^{n-2j-1} (1-p_k)^{n-2j-1} \\
 &\leq \sum_{i \geq 1} n \binom{n}{j} p_i^{j+1} (1-p_i)^{n-2j-1} \sum_{k \geq 1} \binom{n-j}{j} p_k^{j+1} (1-p_k)^{n-2j-1}
 \end{aligned}$$

$$= O\left(\text{Var } K_j^*(n) \frac{\text{Var } K_j^*(n)}{n}\right) = o(\text{Var } K_j^*(n)), \quad n \rightarrow \infty,$$

having utilized (67) for the last limit relation.

ANALYSIS OF F_j . For $m \in \mathbb{N}$ and $x \in [0, 1]$, $1 - x^m \leq m(1 - x)$. Using this with $m = j$ and $x = (1 - p_i)(1 - p_k)$ we conclude that $1 - (1 - p_i)^j(1 - p_k)^j \leq j(p_i + p_k - p_i p_k) \leq j(p_i + p_k)$ and thereupon

$$\begin{aligned} F_j(i, k, n) &\leq j \binom{n}{j} \binom{n-j}{j} p_i^j p_k^j (1 - p_i)^{n-2j} (1 - p_k)^{n-2j} (p_i + p_k) \\ &=: F_j^{(1)}(i, k, n) + F_j^{(2)}(i, k, n). \end{aligned}$$

Further, invoking Lemma 13 yields

$$\begin{aligned} 0 \leq \sum_{i \neq k} F_j^{(1)}(i, k, n) &\leq j \sum_{i \geq 1} \binom{n}{j} p_i^{j+1} (1 - p_i)^{n-2j} \sum_{k \geq 1} \binom{n-j}{j} p_k^j (1 - p_k)^{n-2j} \\ &= O\left(\frac{\text{Var } K_j^*(n)}{n} \text{Var } K_j^*(n)\right) = o(\text{Var } K_j^*(n)), \quad n \rightarrow \infty. \end{aligned}$$

Here, the latter asymptotic relation is a consequence of (67).

The argument for $F_j^{(2)}$ is analogous, and we omit details.

ANALYSIS OF $C_j^{(2)}$. Notice that $\binom{n}{j} - \binom{n-j}{j} = O(n^{j-1})$ as $n \rightarrow \infty$. Hence, mimicking the argument used for the analysis of $F_j^{(1)}$ we conclude that

$$\begin{aligned} \sum_{i \neq k} |C_j^{(2)}(i, k, n)| &= O\left(\sum_{i \neq k} n^{2j-1} p_i^j p_k^j (1 - p_i)^{n-j} (1 - p_k)^{n-j}\right) \\ &= O\left(\sum_{i \geq 1} n^j p_i^j (1 - p_i)^{n-j} \sum_{k \geq 1} n^{j-1} p_k^j (1 - p_k)^{n-j}\right) = O\left(\text{Var } K_j^*(n) \frac{\text{Var } K_j^*(n)}{n}\right) \\ &= o(\text{Var } K_j(n)), \quad n \rightarrow \infty. \end{aligned}$$

Combining all the fragments together we arrive at (69). □

With Proposition 3 at hand, we are ready to prove the LIL stated in Theorem 4. We argue along the lines of the proof of Theorem 3.7 in [3].

Proof of Theorem 4. The deterministic and Poissonized schemes discussed in Section 1.1 are not necessarily defined on a common probability space. In other words, we have not assumed so far that the schemes were defined by throwing one and the same collection of balls. Our plan is to deduce LILs for $\mathcal{K}_j^*(n)$ from the corresponding LILs for $K_j^*(t)$. To this end, we need to couple the two schemes. Let X_1, X_2, \dots be independent random variables with distribution $(p_k)_{k \in \mathbb{N}}$, which are independent of a Poisson process π and particularly its arrival sequence $(S_n)_{n \in \mathbb{N}}$. For all $j, n \in \mathbb{N}$

and $t \geq 0$, we define coupled versions of $\mathcal{K}_j(n)$, $\mathcal{K}_j^*(n)$, $K_j(t)$ and $K_j^*(t)$ as follows, keeping the notation for the variables unchanged:

- $\mathcal{K}_j(n)$ = # of distinct values that the variables X_1, X_2, \dots, X_n take at least j times,
- $\mathcal{K}_j^*(n)$ = # of distinct values that the variables X_1, X_2, \dots, X_n take exactly j times,
- $K_j(t)$ = # of distinct values that the variables $X_1, X_2, \dots, X_{\pi(t)}$ take at least j times,
- $K_j^*(t)$ = # of distinct values that the variables $X_1, X_2, \dots, X_{\pi(t)}$ take exactly j times,

To justify the construction, observe that the variable X_i can be thought of as the index of a box hit by the i th ball. The most important conclusion of the preceding discussion is that, for all $j, n \in \mathbb{N}$, $\mathcal{K}_j^*(n) = K_j^*(S_n)$ a.s. (for the coupled variables).

We prove the result in several steps.

STEP 1. According to Step 2 of the proof of Theorem 3.7 in [3],

$$\lim_{n \rightarrow \infty} \frac{K_j(S_n) - K_j(n)}{(\text{Var } K_j(n)m(\text{Var } K_j(n)))^{1/2}} = 0 \quad \text{a.s.},$$

where $m(t) = \log t$ under (2) and (3) and $m(t) = \log \log t$ under the other assumptions of Theorem 4.

By Lemmas 10 and 11, for $j \in \mathbb{N}$, $\text{Var } K_j^*(t)$ and $\text{Var } K_j(t)$ are asymptotically equivalent up to a constant, whence

$$\lim_{n \rightarrow \infty} \frac{K_j(S_n) - K_j(n)}{(\text{Var } K_j^*(n)m(\text{Var } K_j^*(n)))^{1/2}} = 0 \quad \text{a.s.}$$

By Lemma 11, for $j \in \mathbb{N}$, $\text{Var } K_{j+1}^*(t)$ and $\text{Var } K_j^*(t)$ are asymptotically equivalent up to a constant, unless $\alpha = j = 1$. In the latter case, invoking in addition (19) we obtain $\text{Var } K_{j+1}^*(t) = o(\text{Var } K_j^*(t))$ as $t \rightarrow \infty$. This in combination with the last centered limit relation, in which we replace j with $j + 1$, yields

$$\lim_{n \rightarrow \infty} \frac{K_{j+1}(S_n) - K_{j+1}(n)}{(\text{Var } K_j^*(n)m(\text{Var } K_j^*(n)))^{1/2}} = 0 \quad \text{a.s.}$$

Since, for $j \in \mathbb{N}$, $K_j^*(t) = K_j(t) - K_{j+1}(t)$ a.s., subtracting the last two centered limit relations we arrive at

$$\lim_{n \rightarrow \infty} \frac{K_j^*(S_n) - K_j^*(n)}{(\text{Var } K_j^*(n)m(\text{Var } K_j^*(n)))^{1/2}} = 0 \quad \text{a.s.}$$

STEP 2. Halves of LILs (4), (7), (11) and (20) ((5), (8), (12) and (21)) read

$$\limsup_{n \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{K_j^*(n) - \mathbb{E}K_j^*(n)}{(\text{Var } K_j^*(n)m(\text{Var } K_j^*(n)))^{1/2}} \leq C \ (\geq -C) \quad \text{a.s.},$$

where the case-dependent constant C is equal to the right-hand side of (4), (7), (11) or (20) ((5), (8), (12) or (21)), respectively. This taken together with the conclusion of Step 1, formula (65) and Proposition 3 enables us to obtain

$$\limsup_{n \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{\mathcal{K}_j^*(n) - \mathbb{E}\mathcal{K}_j^*(n)}{(\text{Var } \mathcal{K}_j^*(n)m(\text{Var } \mathcal{K}_j^*(n)))^{1/2}} \leq C (\geq -C) \quad \text{a.s.}$$

Here, we have used a decomposition

$$\mathcal{K}_j^*(n) - \mathbb{E}\mathcal{K}_j^*(n) = (K_j^*(S_n) - K_j^*(n)) + (K_j^*(n) - \mathbb{E}K_j^*(n)) + (\mathbb{E}K_j^*(n) - \mathbb{E}\mathcal{K}_j^*(n))$$

a.s. This finishes the proof of

$$\limsup_{t \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{\mathcal{K}_1^*(n) - \mathbb{E}\mathcal{K}_1^*(n)}{(\text{Var } \mathcal{K}_1^*(n) \log \log \text{Var } \mathcal{K}_1^*(n))^{1/2}} \leq 2^{1/2} (\geq -2^{1/2}) \quad \text{a.s.}$$

in the situation that $\alpha = 1$ and relation (18) fails to hold.

According to Lemma 6, for any $\delta > 0$ and the deterministic sequence (τ_n) defined in (29),

$$\limsup_{n \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{K_j^*(\lfloor \tau_n \rfloor) - \mathbb{E}K_j^*(\lfloor \tau_n \rfloor)}{C(\text{Var } K_j^*(\lfloor \tau_n \rfloor)m(\text{Var } K_j^*(\lfloor \tau_n \rfloor)))^{1/2}} \geq 1 - \delta (\leq -(1 - \delta)) \quad \text{a.s.}$$

Combining these inequalities with the conclusion of Step 1, formula (65) and Proposition 3 we arrive at

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{\mathcal{K}_j^*(n) - \mathbb{E}\mathcal{K}_j^*(n)}{C(\text{Var } \mathcal{K}_j^*(n)m(\text{Var } \mathcal{K}_j^*(n)))^{1/2}} \\ \geq (\leq) \limsup_{n \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{K_j^*(\lfloor \tau_n \rfloor) - \mathbb{E}K_j^*(\lfloor \tau_n \rfloor)}{C(\text{Var } K_j^*(\lfloor \tau_n \rfloor)m(\text{Var } K_j^*(\lfloor \tau_n \rfloor)))^{1/2}} \\ \geq 1 - \delta (\leq -(1 - \delta)) \quad \text{a.s.} \end{aligned}$$

Sending $\delta \rightarrow 0+$ yields

$$\limsup_{n \rightarrow \infty} \left(\liminf_{n \rightarrow \infty} \right) \frac{\mathcal{K}_j^*(n) - \mathbb{E}\mathcal{K}_j^*(n)}{(\text{Var } \mathcal{K}_j^*(n)m(\text{Var } \mathcal{K}_j^*(n)))^{1/2}} \geq C (\leq -C) \quad \text{a.s.,}$$

which finishes the proof. □

Acknowledgement

We thank the AE and the anonymous referee for useful comments which have helped us to avoid a couple of unfortunate formulations and made the text more reader-friendly.

Funding

The research was supported by Applied Probability Trust in the framework of a Ukraine Support Scheme.

References

- [1] Bahadur, R.R.: On the number of distinct values in a large sample from an infinite discrete distribution. *Proc. Natl. Inst. Sci. India, A Phys. Sci.* **26**(supplement II), 67–75 (1960). [MR0137256](#)
- [2] Bingham, N.H., Goldie, C.M., Teugels, J.L.: *Regular Variation*. Cambridge University Press (1989). [MR0898871](#). <https://doi.org/10.1017/CBO9780511721434>
- [3] Buraczewski, D., Iksanov, A., Kotelnikova, V.: Limit theorems for sums of independent indicators, with applications to the Ginibre point process and Karlin's occupancy scheme. arXiv preprint [arXiv:2306.15027](#) (2023)
- [4] Chang, J., Grabchak, M.: Necessary and sufficient conditions for the asymptotic normality of higher order Turing estimators. *Bernoulli* **29**(4), 3369–3395 (2023). [MR4632142](#). <https://doi.org/10.3150/23-bej1587>
- [5] Darling, D.A.: Some limit theorems associated with multinomial trials. In: *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, pp. 345–350. Univ. California Press, Berkeley, CA (1967). [MR0216547](#)
- [6] De Blasi, P., Mena, R.H., Prünster, I.: Asymptotic behavior of the number of distinct values in a sample from the geometric stick-breaking process. *Ann. Inst. Stat. Math.* **74**(1), 143–165 (2022). [MR4358167](#). <https://doi.org/10.1007/s10463-021-00791-6>
- [7] Derbazi, Z., Gnedin, A., Marynych, A.: Records in the infinite occupancy scheme. arXiv preprint [arXiv:2308.01739](#) (2023)
- [8] Geluk, J.L., de Haan, L.: *Regular Variation, Extensions and Tauberian Theorems*. CWI Tract, vol. 40, p. 132. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam (1987). [MR0906871](#)
- [9] Gnedin, A., Hansen, B., Pitman, J.: Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. *Probab. Surv.* **4**, 146–171 (2007). [MR2318403](#). <https://doi.org/10.1214/07-PS092>
- [10] Grabchak, M., Kelbert, M., Paris, Q.: On the occupancy problem for a regime-switching model. *J. Appl. Probab.* **57**(1), 53–77 (2020). [MR4094383](#). <https://doi.org/10.1017/jpr.2020.33>
- [11] Iksanov, A., Kotelnikova, V.: Small counts in nested Karlin's occupancy scheme generated by discrete Weibull-like distributions. *Stoch. Model. Appl.* **153**, 283–320 (2022). [MR4481286](#). <https://doi.org/10.1016/j.spa.2022.08.006>
- [12] Karlin, S.: Central limit theorems for certain infinite urn schemes. *J. Math. Mech.* **17**(4), 373–401 (1967). [MR0216548](#). <https://doi.org/10.1512/iumj.1968.17.17020>
- [13] Longnecker, M., Serfling, R.J.: General moment and probability inequalities for the maximum partial sum. *Acta Math. Acad. Sci. Hung.* **30**(1-2), 129–133 (1977). [MR0501261](#). <https://doi.org/10.1007/BF01895656>
- [14] Rosenthal, H.P.: On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables. *Isr. J. Math.* **8**, 273–303 (1970). [MR0271721](#). <https://doi.org/10.1007/BF02771562>