On the number of jumps of random walks with a barrier

Alex Iksanov∗
Faculty of Cybernetics, National T. Shevchenko University,
01033 Kiev, Ukraine

Martin Möhle†
Mathematical Institute, University of Düsseldorf,
40225 Düsseldorf, Germany

January 18, 2008

Abstract

Let \( S_0 := 0 \) and \( S_k := \xi_1 + \cdots + \xi_k \) for \( k \in \mathbb{N} := \{1, 2, \ldots \} \), where \( \{\xi_k : k \in \mathbb{N}\} \) are independent copies of a random variable \( \xi \) with values in \( \mathbb{N} \) and distribution \( p_k := \mathbb{P}\{\xi = k\}, k \in \mathbb{N} \). We interpret the random walk \( \{S_k : k = 0, 1, 2, \ldots\} \) as a particle jumping to the right through integer positions. Fix \( n \in \mathbb{N} \) and modify the process by requiring that the particle is bumped back to its current state each time a jump would bring the particle to a state larger than or equal to \( n \). This constraint defines an increasing Markov chain \( \{R_k^{(n)} : k = 0, 1, 2, \ldots\} \) which never reaches the state \( n \). We call this process a random walk with barrier \( n \).

Let \( M_n \) denote the number of jumps of the random walk with barrier \( n \). This paper focuses on the asymptotics of \( M_n \) as \( n \) tends to infinity. A key observation is that, under \( p_1 > 0 \), \( \{M_n : n \in \mathbb{N}\} \) satisfies the distributional recursion \( M_1 = 0 \) and \( M_n \overset{d}{=} M_{n-1} + 1 \) for \( n = 2, 3, \ldots \), where \( I_n \) is independent of \( M_2, \ldots, M_{n-1} \) with distribution \( \mathbb{P}\{I_n = k\} = p_k/(p_1 + \cdots + p_{n-1}) \), \( k \in \{1, \ldots, n-1\} \).

Depending on the tail behaviour of the distribution of \( \xi \), several scalings for \( M_n \) and corresponding limiting distributions come into play, among them stable distributions and distributions of exponential integrals of subordinators.

∗e-mail address: iksan@unicyb.kiev.ua
†e-mail address: moehle@math.uni-duesseldorf.de
The methods used in this paper are mainly probabilistic. The key tool is to compare (couple) the number of jumps $M_n$ with the first time $N_n$ when the unrestricted random walk $\{S_k : k = 0, 1, \ldots\}$ reaches a state larger than or equal to $n$.

The results are applied to derive the asymptotics of the number of collision events (that take place until there is just a single block) for beta(a,b)-coalescent processes with parameters $0 < a < 2$ and $b = 1$.

Keywords: absorption time; beta coalescent; coupling; exponential integrals; Mittag-Leffler distribution; random recursive equation; random walk; stable limit; subordinator

AMS 2000 Mathematics Subject Classification: Primary 60F05; 60G50 Secondary 05C05; 60E07

1 Introduction and main results

Fix $n \in \mathbb{N} := \{1, 2, \ldots\}$. By a random walk with the barrier $n$ we mean the sequence $\{R^{(n)}_k : k \in \mathbb{N}_0 := \{0, 1, \ldots\}\}$ defined recursively via $R^{(n)}_0 := 0$ and

$$R^{(n)}_k := R^{(n)}_{k-1} + \xi_k 1_{\{R^{(n)}_{k-1} + \xi_k < n\}}, \quad k \in \mathbb{N},$$

where $\{\xi_k : k \in \mathbb{N}\}$ are independent copies of a random variable $\xi$ with some proper and non-degenerate probability distribution

$$p_k := \mathbb{P}\{\xi = k\}, \quad k \in \mathbb{N}, \quad p_1 > 0. \quad (1)$$

Note that the sequence $\{R^{(n)}_k : k \in \mathbb{N}_0\}$ is non-decreasing and that $R^{(n)}_k < n$ for all $k \in \mathbb{N}_0$. Let

$$M_n := \#\{k \in \mathbb{N} : R^{(n)}_{k-1} \neq R^{(n)}_k\} = \sum_{l=0}^{\infty} 1_{\{R^{(n)}_l + \xi_{l+1} < n\}}$$

denote the number of jumps in the process $\{R^{(n)}_k : k \in \mathbb{N}_0\}$. Note that $M_1 = 0$ and that $1 \leq M_n \leq n - 1$ for $n \geq 2$. The aim of the paper is to investigate the asymptotic behaviour of $M_n$ as $n \to \infty$. Hindery and Walk [24] investigate processes more general than random walks with a barrier, but the circle of problems they consider is different from ours.

As $p_1 > 0$, it follows from Lemma 1 in [26] that the marginal distributions of $\{M_n : n \in \mathbb{N}\}$ satisfy the distributional recursion $M_1 = 0$ and

$$M_n \overset{d}{=} M_{n-1} + 1, \quad n \in \{2, 3, \ldots\}, \quad (2)$$
where $I_n$ is a random variable independent of $M_2, \ldots, M_{n-1}$ with distribution
\[
\mathbb{P}(I_n = k) = \frac{p_k}{p_n + \cdots + p_{n-1}}, \quad k, n \in \mathbb{N}, k < n.
\] (3)

Note that $I_n$ is the size of the first jump of $\{R_n^{(k)} : k \in \mathbb{N}_0\}$. It is worth to mention that the asymptotic results presented later apply not only to the number of jumps in a random walk with a barrier, but also to all sequences whose marginal distributions satisfy the distributional recursion (2) with distribution of $I_n$ given via (3). The number of jumps of a random walk with barrier $n$ is just one example for a sequence whose marginal distributions satisfy the recursion (2).

Before we will formulate our asymptotic results for $M_n$ we now briefly discuss closely related and more general models and the corresponding literature. In order to do this assume for a while that the distribution of $I_n$ in (2) does not follow (3) but rather takes the more general form
\[
\mathbb{P}(I_n = k) = \pi_{n,n-k}, \quad k, n \in \mathbb{N}, k < n,
\] (4)

where the $\pi_{ij}, 1 \leq j < i$, are some given non-negative constants satisfying $\sum_{i=1}^{n-1} \pi_{ij} = 1$. Probably the most general description of sequences $\{M_n : n \in \mathbb{N}\}$ satisfying the recursion (2) with distribution of $I_n$ of the form (4) is as follows. Consider a decreasing Markov chain $\{Z_k : k \in \mathbb{N}_0\}$ with state space $\mathbb{N}$ and transition probabilities $\pi_{ij} > 0$ for $i, j \in \mathbb{N}$ with $j < i$ and $\pi_{ij} = 0$ otherwise. For $n \in \mathbb{N}$ let
\[
M_n := \inf\{k \geq 1 : Z_k = 1 \text{ given } Z_0 = n\},
\]
denote the absorption time of the Markov chain conditioned on the event that the chain starts in the initial state $n$. Then the marginal distributions of $\{M_n : n \in \mathbb{N}\}$ satisfy the distributional recursion (2) with distribution of $I_n$ given via (4).

We are only aware of two papers [12] and [37] addressing the asymptotic behaviour of $M_n$ as $n \to \infty$ in the general setting when it is not assumed that $\pi_{ij}$ takes some particular form. The problem is simpler if either the probabilities $\pi_{ij}$ are given explicitly, or if they have some particular functional form. In this latter situation some results on the asymptotic behaviour of the recursion (2) with (4) are available, for example in the context of random composition structures [4, 18, 19, 20], of coalescent theory ([21, 25, 26]; see also Section 7 of this work), and in the context of random trees [11, 15, 26, 30, 31]. We also refer to [3] for a number of interpretations
of the random recursion (2), where $I_n$ satisfies (4) with $\pi_{ij} = (i - 1)^{-1}$, $i, j \in \mathbb{N}$, $j < i$.

Throughout the paper $r(\cdot) \sim s(\cdot)$ means that $r(\cdot)/s(\cdot) \to 1$ as the argument tends to infinity. The symbols $\overset{d}{\to}$, $\Rightarrow$, and $\overset{P}{\to}$ denote convergence in law, weak convergence, and convergence in probability, respectively, and $X_n \overset{d}{\to} (\Rightarrow, \overset{P}{\to}) X$ means that the limiting relation holds when $n \to \infty$. By $L$ we always denote a function slowly varying at infinity.

We now state our main asymptotic results for sequences of random variables $\{M_n : n \in \mathbb{N}\}$ satisfying the distributional recursion (2) with distribution of $I_n$ given via (3). We begin with a weak law of large numbers.

**Theorem 1.1.** If $\sum_{j=1}^{n} \sum_{k=j}^{\infty} p_k \sim L(n)$ for some function $L$ slowly varying at $\infty$, then, as $n \to \infty$,

$$\frac{M_n}{\mathbb{E}M_n} \overset{P}{\to} 1 \quad (5)$$

and $\mathbb{E}M_n \sim n/L(n)$. In particular, if

$$m := \mathbb{E}\xi < \infty, \quad (6)$$

then $\mathbb{E}M_n \sim n/m$. If (6) holds, and if there exists a sequence of positive numbers $\{a_n : n \in \mathbb{N}\}$ such that $M_n/a_n \overset{P}{\to} 1$ as $n \to \infty$, then $a_n \sim n/m$.

To formulate further results we need some more notation. For $C > 0$ and $\alpha \in [1, 2]$ let $\mu_\alpha$ be an $\alpha$-stable distribution with characteristic function $\psi_\alpha(t)$, $t \in \mathbb{R}$ of the form

$$\exp\{-|t|^\alpha CT(1 - \alpha)(\cos(\pi\alpha/2) + i \sin(\pi\alpha/2) \text{sgn}(t))\}, \quad 1 < \alpha < 2;$$

$$\exp\{-|t|C(\pi/2 - i \log |t| \text{sgn}(t))\}, \quad \alpha = 1;$$

$$\exp(-(C/2)t^2), \quad \alpha = 2.$$ 

In the case when (6) holds, the following Theorem 1.2 provides necessary and sufficient conditions ensuring that $M_n$, properly normalized and centered, possesses a weak limit.

**Theorem 1.2.** If $m := \mathbb{E}\xi < \infty$, then the following assertions are equivalent.

(i) There exist sequences of numbers $\{a_n, b_n : n \in \mathbb{N}\}$ with $a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, $(M_n - b_n)/a_n$ converges weakly to a non-degenerate and proper probability law.
(ii) Either \( \sigma^2 := \mathbb{D}\xi < \infty \), or \( \sigma^2 = \infty \) and for some \( \alpha \in [1, 2] \) and some function \( L \) slowly varying at \( \infty \),
\[
\sum_{k=1}^{n} k^2 p_k \sim n^{2-\alpha} L(n), \quad n \to \infty.
\] (7)

If \( \sigma^2 < \infty \), then, with \( b_n := n/m \) and \( a_n := \left( m^{-3}C^{-1}\sigma^2 n \right)^{1/2} \), the limiting law is \( \mu_2 \) (normal with mean zero and variance \( C \)).

If \( \sigma^2 = \infty \) and (7) holds with \( \alpha = 2 \), then, with \( b_n := n/m \) and \( a_n := m^{-3/2}c_n \), where \( c_n \) is any sequence satisfying \( \lim_{n \to \infty} nL(c_n)/c_n^2 = C \), the limiting law is \( \mu_2 \).

If \( \sigma^2 = \infty \) and (7) holds with \( \alpha \in [1, 2) \), then, with \( b_n := n/m \) and \( a_n := m^{-(\alpha+1)/\alpha}c_n \), where \( c_n \) is any sequence satisfying

\[
\lim_{n \to \infty} \frac{nL(c_n)}{c_n^\alpha} = \frac{\alpha}{2 - \alpha} C,
\]

the limiting law is \( \mu_\alpha \).

Remark 1.3. For \( \sigma^2 < \infty \), the same weak convergence result for \( M_n \) was obtained in Theorem 4.1 in [12] in a setting more general than ours. Note that for \( \alpha \in [1, 2) \), (7) is equivalent to \( \mathbb{P}\{\xi \geq n\} \sim (2 - \alpha)n^{-\alpha}L(\alpha)/\alpha \), \( n \to \infty \).

If the mean of \( \xi \) is infinite, the following Theorem 1.4 (respectively Theorem 1.5) points out conditions ensuring that \( M_n \), properly normalized without centering (and centered), possesses a weak limit.

**Theorem 1.4.** Suppose that for some \( \alpha \in (0, 1) \) and some function \( L \) slowly varying at \( \infty \)
\[
\mathbb{P}\{\xi \geq n\} = \sum_{k=n}^{\infty} p_k \sim \frac{L(n)}{n^\alpha}, \quad n \to \infty.
\] (8)

Then, as \( n \to \infty \),
\[
\frac{L(n)}{n^\alpha}M_n \xrightarrow{d} \int_0^\infty e^{-U_t} dt,
\] (9)

where \( \{U_t : t \geq 0\} \) is a drift-free subordinator with Lévy measure
\[
\nu(dt) = \frac{e^{-t/\alpha}}{(1 - e^{-t/\alpha})^{\alpha+1}} dt, \quad t > 0.
\] (10)
Theorem 1.5. Suppose that $E\xi = \infty$ and that for some function $L$ slowly varying at $\infty$

$$\mathbb{P}\{\xi \geq n\} = \sum_{k=n}^{\infty} p_k \sim \frac{L(n)}{n}. \quad (11)$$

Let $c$ be any positive function satisfying $\lim_{x \to \infty} x \mathbb{P}\{\xi \geq c(x)\} = 1$ and set

$$\psi(x) := x \int_{0}^{c(x)} \mathbb{P}\{\xi > y\} \, dy.$$ Let $b$ be any positive function satisfying

$$b(\psi(x)) \sim \psi(b(x)) \sim x,$$

and set $a(x) := x^{-1}b(x)c(b(x))$. Then, $(M_n - b(n))/a(n)$ converges weakly to the stable distribution $\mu_1$ with $C = 1$.

In the literature there exist at least two standard approaches to studying distributional recursions. One approach is purely analytic and based on a singularity analysis of generating functions (see, for example, [15, 31]). Another approach, called contraction method, is more probabilistic (see [29, 35, 36]). It was remarked in [26] that the recursions (2) which satisfy (3) can be successfully investigated by using probabilistic methods alone (completely different from contraction methods). The present work extends ideas laid down in [26] for the particular case

$$\mathbb{P}\{I_n = k\} = \frac{n}{n-1} \frac{1}{k(k+1)}, \quad k \in \{1, \ldots, n-1\}.$$

The basic steps of the technique exploited can be summarized as follows.

Define $S_0 := 0,$

$$S_n := \xi_1 + \cdots + \xi_n \quad \text{and} \quad N_n := \inf\{k \geq 1 : S_k \geq n\}, \quad n \in \mathbb{N}.$$ One may expect that the limiting behaviour of $M_n$ and $N_n$ are similar, or at least that the limiting behaviour of the latter influences that of the former. Similarity in the limiting behaviour of $M_n$ and $N_n$ is well indicated by asymptotic properties of their difference. In particular, we will prove the following.

(a) If $E\xi < \infty$, then $M_n - N_n$ weakly converges. Therefore, $M_n$, properly normalized and centered, possesses a weak limit if and only if the same is true for $N_n$.

(b) Assume now that $E\xi = \infty$. (b1) If $\sum_{k=n}^{\infty} p_k \sim L(n)/n$ and if $(N_n - b_n)/a_n$ weakly converges to some $\mu$, then $(M_n - N_n)/a_n \overset{P}{\to} 0$ which proves that $(M_n - b_n)/a_n$ weakly converges to $\mu$. Thus in cases (a) and (b1) the weak behaviour of $M_n$ and $N_n$ is the same. (b2) If, for some $\alpha \in (0, 1)$, $\sum_{k=n}^{\infty} p_k \sim\)
\( n^{-\alpha}L(n) \) and \( N_n/a_n \) weakly converges to some \( \nu_1 \), then \( (M_n-N_n)/a_n \) weakly converges to some \( \nu_2 \). Even though the argument exploited above does not apply, it will be proved that \( M_n/a_n \) weakly converges to \( \nu_3 \neq \nu_1 \). Thus in this latter case a weak behaviour of \( M_n \) is not completely determined by that of \( N_n \). Now it is influenced by the weak behaviour of both \( N_n \) and \( n-S_{N_n-1} \) to, approximately, the same extent. This observation can be explained as follows. The probability of one big jump of \( S_n \) in comparison to cases (a) and (b1) is higher, and therefore the epoch \( N_n \) comes more ”quickly”. As a consequence, a contribution to \( M_n \) of the number of jumps in the sequence \( \{R_k^{(n)} : k \in \mathbb{N}_0\} \), while \( R_k^{(n)} \) is travelling from \( R_{N_n-1}^{(n)} = S_{N_n-1} \) to \( n-1 \), gets significant.

The referee pointed out the following interpretation of Theorem 1.4 that can be read from (10) in combination with results from [20] on exponential functionals of subordinators. Since \( N_n \) is known to be asymptotic to the local time of an unrestricted Bessel process (which has Mittag-Leffler distribution), then \( M_n \) is asymptotic to the local time of a modified Bessel process, obtained by recursive peeling the meander of the unrestricted Bessel process (the latter has distribution of the right-hand side in (9)).

To close the introduction, it remains to review structural units of the rest of the paper. In Section 2 we investigate both the univariate and the bivariate weak behaviour of \( (N_n, n-S_{N_n-1}) \), and discuss their relation to exponential integrals of subordinators. The proof of Theorem 1.4 along with some comments explaining the appearance of the limiting law in (9) are given in Section 3. Theorems 1.2, 1.1 and 1.5 are proved in Sections 4, 5 and 6 respectively.

In the final Section 7 we apply our results to derive limiting theorems for the number of collision events that take place in certain beta coalescent processes until there is just a single block. It turns out that the results are applicable for \( \beta(a,b) \)-coalescents with \( 0 < a < 2 \) and \( b = 1 \), because for that parameter range the number of collisions satisfy the distributional recursion (2) such that \( I_n \) has a distribution of the form (3).

## 2 Results on \( N_n \) and \( n-S_{N_n-1} \): case \( m = \infty \)

### 2.1 Univariate results

Below necessary and sufficient conditions are given ensuring that the sequence \( \{N_n : n \in \mathbb{N}\} \) (a) properly normalized (without centering), weakly converges to a non-degenerate law (Proposition 2.1) and (b) is relatively stable (Proposition 2.2).
The implication (8) \(\Rightarrow\) (13) (in the case \(\alpha \in (0, 1)\)) in Proposition 2.1 and the equivalence \((a) \Leftrightarrow (b)\) in Proposition 2.2 are well-known (see Theorem 7 in [17] and Corollary 8.1.7 in [9] respectively). Although the whole results may seem classic, we have not been able to locate them in the literature in the present form. Therefore, complete proofs of them are provided in [27] which is a preprint version of this work.

We say that a random variable \(\varsigma_\alpha\) has a scaled Mittag-Leffler distribution with parameter \(\alpha \in (0, 1)\), if

\[
E_{\varsigma_\alpha}^n = \frac{n!}{\Gamma^n(1 - \alpha)\Gamma(1 + n\alpha)}, \quad n \in \mathbb{N}.
\]

Note that the moments \(\{E_{\varsigma_\alpha}^n : n \in \mathbb{N}\}\) uniquely determine the distribution.

**Proposition 2.1.** If (8) holds for some \(\alpha \in [0, 1)\), then

\[
\lim_{n \to \infty} \frac{L_k(n)}{n^{\alpha k}} E N_n^k = \frac{k!}{\Gamma^k(1 - \alpha)\Gamma(1 + k\alpha)}, \quad k \in \mathbb{N},
\]

and therefore

\[
\frac{L(n)}{n^{\alpha}} N_n \Rightarrow \theta_\alpha,
\]

where \(\theta_\alpha\) is the scaled Mittag-Leffler distribution with parameter \(\alpha\).

Conversely, assume that there exist a sequence \(\{a(n) : n \in \mathbb{N}\}\) of positive real numbers such that \(N_n/a(n)\) weakly converges to a non-degenerate and proper law \(\theta\). Then \(a(n) \sim D (\sum_{k=n}^\infty p_k)^{-1} \sim D n^{\alpha}/L(n)\) for some constants \(D > 0, \alpha \in (0, 1)\) and some function \(L\) slowly varying at \(\infty\), and (13) holds.

**Proposition 2.2.** The following conditions are equivalent.

(a) \(\sum_{m=1}^n \sum_{k=m}^\infty p_k \sim L(n)\) for some \(L\) slowly varying at \(\infty\).

(b) \(1 - \sum_{n=1}^\infty e^{-s n} p_n \sim s L(1/s)\) as \(s \downarrow 0\) for some \(L\) slowly varying at \(\infty\).

(c) The sequence \(\{N_n : n \in \mathbb{N}\}\) is relatively stable, i.e. there exist a sequence \(\{a(n) : n \in \mathbb{N}\}\) of positive real numbers, such that \(N_n/a(n) \xrightarrow{P} 1\).

Moreover, if (a) holds, then

\[
\lim_{n \to \infty} \frac{L_k(n)}{n^k} E N_n^k = 1, \quad k \in \mathbb{N},
\]

and \(a(n) \sim E N_n\).
The next result is a corollary of Theorem 1.1 and Proposition 2.2.

**Corollary 2.3.** Assume that (11) holds. Then, $\mathbb{E} N_n \sim \mathbb{E} M_n \sim n/m(n)$, where $m(x) := \int_0^x \mathbb{P}\{\xi > y\} \, dy$, $x > 0$. Moreover,
$$
\frac{m(n)N_n}{n} \overset{P}{\to} 1 \quad \text{and} \quad \frac{m(n)M_n}{n} \overset{P}{\to} 1.
$$

In particular, $M_n/N_n \overset{P}{\to} 1$.

**Proof.** Condition (11) ensures that $m(\cdot)$ belongs to the de Haan class $\Pi$, i.e. 
$$
\lim_{x \to \infty} \frac{m(\lambda x) - m(x)}{L(x)} = \log \lambda.
$$
In particular, $m(\cdot)$ is slowly varying at $\infty$. Since $
\sum_{j=1}^n \sum_{k=j}^\infty p_k \sim m(n)$, Theorem 1.1 and Proposition 2.2 imply the result for $M_n$ and $N_n$ respectively. \hfill $\square$

Proposition 2.4 is a key ingredient for our proof of Theorem 1.5. Define $Y_n := n - S_{N_n-1}$, $n \in \mathbb{N}$.

**Proposition 2.4.** Assume that (11) holds. Then, for fixed $\delta > 0$,
$$
\mathbb{E} Y_n^\delta = O\left(\frac{n^\delta L(n)}{m(n)}\right),
$$
where $m(x) := \int_0^x \mathbb{P}\{\xi > y\} \, dy$, $x > 0$. Furthermore, for functions $a$ and $b$ as used in Theorem 1.5,
$$
\frac{b(n)Y_n}{n a(n)} \overset{P}{\to} 0.
$$

**Proof.** In the same way as in the proof of Proposition 2.6 it follows that
$$
\mathbb{E} Y_n^\delta = \sum_{k=0}^{n-1} (n-k)^\delta \mathbb{P}\{\xi \geq n-k\} u_k, \quad n \in \mathbb{N},
$$
where $u_k := \sum_{i=0}^k \mathbb{P}\{S_i = k\}$, $k \in \mathbb{N}_0$. By Corollary 2.3, $\mathbb{E} N_n \sim n/m(n)$.

On the other hand, $\mathbb{E} N_n \sim \sum_{k=0}^n u_k$, $n \in \mathbb{N}$. Thus, $\sum_{k=0}^n u_k \sim n/m(n)$ and, by Corollary 1.7.3 in [9],
$$
U(s) := \sum_{n=0}^{\infty} s^n u_n \sim \frac{1}{m((1-s)^{-1})(1-s)} \quad \text{as } s \uparrow 1.
$$

By the same Corollary
$$
V(s) := \sum_{n=1}^{\infty} s^n \delta \mathbb{P}\{\xi \geq n\} \sim \frac{\Gamma(\delta) L((1-s)^{-1})}{(1-s)^{\delta}} \quad \text{as } s \uparrow 1.
$$
Therefore,
\[
\sum_{n=1}^{\infty} s^n E Y_n^\delta = U(s) V(s) \sim \frac{\Gamma(\delta)}{(1-s)^{\delta+1}} \frac{L((1-s)^{-1})}{m((1-s)^{-1})} \text{ as } s \uparrow 1.
\]

Therefore, Corollary 1.7.3 in [9] applies and proves (15). Recall that \( \psi(x) = x m(c(x)) \) and \( c(x) \sim x L(c(x)) \), and put \( v(x) := x a(x)/b(x) = c(b(x)) \). Since \( m(x)/L(x) \to \infty \), \( c(x) \to \infty \) and
\[
\frac{\psi(x)}{c(x)} = \frac{x m(c(x))}{c(x)} \sim \frac{m(c(x))}{L(c(x))}
\]
as \( x \to \infty \), we conclude that \( \psi(x)/c(x) \to \infty \) as \( x \to \infty \). Therefore,
\[
\frac{b(x)}{a(x)} = \frac{x}{c(b(x))} \to \infty \quad \text{as } x \to \infty.
\]
The latter relation together with \( m(x)/L(x) \to \infty \) imply that
\[
\frac{L(x) b(x)}{m(x) a(x)} = \frac{L(x) a(x)}{m(x) c(b(x))} \sim \frac{L(x)}{m(x) b(x)L(c(b(x)))} \sim \frac{L(x) m(c(b(x)))}{m(x) L(c(b(x)))}
\]
remains bounded for large \( x \).

For fixed \( \delta \in (0,1) \) and any \( \epsilon > 0 \) we have, by Markov’s inequality and by (15),
\[
P\{Y_n > v(n) \epsilon\} \leq \frac{E Y_n^\delta}{v^\delta(n) \epsilon^\delta} = O\left(\frac{L(n)b(n)}{m(n)a(n)} \left(\frac{b(n)}{a(n)}\right)^{\delta-1}\right) \to 0 \quad \text{as } n \to \infty.
\]
The proof is complete.

2.2 Some results on exponential integrals of subordinators

Let \( \{Z_t : t \geq 0\} \) be a drift-free subordinator which is independent of \( T \), an exponentially distributed random variable with mean one. Set \( Q := \int_0^T e^{-Z_t} dt \), \( M := e^{-Z_T} \), and \( A := \int_T^\infty e^{-Z_t} dt \). As is well-known (see, for example, Lemma 6.2 in [10]) the following equality of distributions holds
\[
A_\infty \overset{d}{=} MA_\infty' + Q,
\]
where \( A_\infty' \) is a copy of \( A_\infty \) which is independent of \( (M,Q) \). The latter means that \( A_\infty \) is a perpetuity (see [2] for the definition and recent results) generated by the random vector \((M,Q)\).

Our next result generalizes Proposition 3.1 in [10] dealing with moments of \( Q \), and a number of results concerning moments of \( \int_0^\infty e^{-Z_t} dt = Q + A \) (see, for example, Proposition 3.3 in [39]).
Proposition 2.5. For $\lambda > 0$ and $\mu \geq 0$

$$EQ^\lambda M^\mu = \frac{\lambda}{1 + \varphi(\lambda + \mu)} EQ^{\lambda-1} M^\mu,$$

where $\varphi(s) := -\log E e^{-sZ}$, $s \geq 0$. In particular,

$$a_{n,m} := EQ^n M^m = \frac{n!}{\prod_{k=0}^{n}(1 + \varphi(m + k))}, \quad m, n \in \mathbb{N}_0,$$  \hfill (18)

$$b_{n,m} := EQ^n A^m = \frac{n!m!}{\prod_{k=0}^{n}(1 + \varphi(m + k))\varphi(1)\cdots\varphi(m)}, \quad m, n \in \mathbb{N}_0.$$  

The moment sequences $\{a_{m,n} : m, n \in \mathbb{N}_0\}$ and $\{b_{m,n} : m, n \in \mathbb{N}_0\}$ uniquely determine the laws of the random vectors $(M, Q)$ and $(A, Q)$ respectively.

Proof. For $t > 0$ define $A_t := \int_0^t e^{-Z_v} dv$. The following is essentially Eq. (3.1) in [10]:

$$A_t^\lambda e^{-\mu Z_t} = \lambda \int_0^t (A_t - A_v)^{\lambda-1} e^{-\mu(Z_t - Z_v)} e^{-(\mu+1)Z_v} dv.$$

Since

$$(A_t - A_v)^{\lambda-1} e^{-\mu(Z_t - Z_v)} = e^{-(\lambda-1)Z_v} \left( \int_0^{t-v} e^{-(Z_s + v - Z_v)} ds \right)^{\lambda-1} e^{-\mu(Z_t - Z_v)},$$

and $\{Z_s + v - Z_v : s \geq 0\}$ is a subordinator which is independent of $\{Z_v : v \leq t\}$ and has the same law as $\{Z_t : t \geq 0\}$, we conclude that $$(\int_0^{t-v} e^{-(Z_s + v - Z_v)} ds)^{\lambda-1} e^{-\mu(Z_t - Z_v)}$$ has the same law as $A_{t-v}^{\lambda-1} e^{-\mu Z_{t-v}}$ and is independent of $e^{-(\lambda-1)Z_v}$. Therefore, using Fubini’s theorem,

$$EA_t^\lambda e^{-\mu Z_t} = \int_0^\infty e^{-t} EA_t^\lambda e^{-\mu Z_t} dt$$

$$= \lambda \int_0^\infty e^{-t} \left( \int_0^t e^{-v\varphi(\lambda+\mu)} EA_{t-v}^{\lambda-1} e^{-\mu Z_{t-v}} dv \right) dt$$

$$= \lambda \int_0^\infty e^{-v\varphi(\lambda+\mu)} \left( \int_v^\infty e^{-t} EA_{t-v}^{\lambda-1} e^{-\mu Z_{t-v}} dt \right) dv$$

$$= \lambda \int_0^\infty e^{-v\varphi(\lambda+\mu)+1} dv \int_0^\infty e^{-u} EA_u^{\lambda-1} e^{-\mu Z_u} du$$

$$= \frac{\lambda}{1 + \varphi(\lambda + \mu)} EA_t^{\lambda-1} e^{-\mu Z_t}.$$
Starting with
\[ E e^{-\mu Z_T} = \int_{0}^{\infty} e^{-t} e^{-\mu Z_t} \, dt = \int_{0}^{\infty} e^{-t(1+\varphi(\mu))} \, dt = \frac{1}{1 + \varphi(\mu)}, \quad (19) \]
the formula for \( a_{n,m} \) follows by induction. To prove that the law of \((M, Q)\) is uniquely determined by \( \{a_{n,m} : n, m \in \mathbb{N}_0\} \), it suffices to check that the marginal laws are uniquely determined by the corresponding moment sequences (see Theorem 3 in [32]). Since \( M \in [0, 1] \) almost surely, the law of \( M \) is trivially moment determinate. From (18) it follows that
\[ EQ^n = \frac{n!}{(1 + \varphi(1)) \cdots (1 + \varphi(n))}, \quad n \in \mathbb{N}. \]
Set \( f_n := EQ^n / n! \). The limit \( f := \lim_{n \to \infty} f_n / f_{n+1} \) exists and is positive (it is finite, if \( Z_t \) is compound Poisson, otherwise it is infinite). By the Cauchy-Hadamard formula, \( f = \sup\{r > 0 : E e^{r Q} < \infty\} \). Therefore, the law of \( Q \) has finite exponential moments of some orders from which we deduce that this law is moment determinate.

According to Proposition 3.3 in [39], \( E A_m^\infty = m! / (\varphi(1) \cdots \varphi(m)) \), \( m \in \mathbb{N}_0 \). In view of (17),
\[ EQ^n A^m = EQ^n M^m E A_m^\infty = \frac{n! m!}{\prod_{k=0}^{n}(1 + \varphi(m+k)) \varphi(1) \cdots \varphi(m)}, \quad m, n \in \mathbb{N}_0. \]
In the same way as above for \((M, Q)\) it can be checked that the law of \((A, Q)\) is determined by the moment sequence. We omit the details. \( \square \)

### 2.3 A bivariate result

Assume that (8) holds, or, equivalently, that
\[ w(n) := \frac{1}{P\{\xi \geq n\}} = \left( \sum_{k=n}^{\infty} p_k \right)^{-1} \sim \frac{n^\alpha}{L(n)} \quad (20) \]
for some \( \alpha \in (0, 1) \). Let \( T \) be an exponentially distributed random variable with mean 1, which is independent of a drift-free subordinator \( \{U_t : t \geq 0\} \) with Lévy measure (10).

From Proposition 2.1 it follows that \( N_n / w(n) \) converges in distribution to a random variable \( \varsigma_\alpha \) with the scaled Mittag-Leffler distribution with parameter \( \alpha \). From (18) or from Proposition 3.1 in [10] we have
\[ E \left( \int_{0}^{T} e^{-U_t} \, dt \right)^n = \frac{n!}{\Gamma^n(1-\alpha) \Gamma(1+n\alpha)}, \quad n \in \mathbb{N}_0, \]
which means that \( \int_0^T e^{-U_t} \, dt \overset{d}{=} \varsigma_{n\alpha} \). Thus, 
\[
\frac{N_n}{w(n)} \overset{d}{\rightarrow} \int_0^T e^{-U_t} \, dt. \tag{21}
\]

Let \( \eta_{n\alpha} \) be a beta-distributed random variable with parameters \( 1 - \alpha \) and \( \alpha \), i.e. with density \( x \mapsto \pi^{-1} \sin(\pi \alpha)x^{-\alpha}(1-x)^{\alpha - 1}, \quad x \in (0,1) \). It is well known (see, for example, Theorem 8.6.3 in [9]) that \( (1 - S_{N_n - 1}/n)^\alpha \overset{d}{\rightarrow} \eta_{n\alpha}^\alpha \). It can be checked that 
\[
E\eta_{n\alpha}^{n\alpha} = \frac{\Gamma(\alpha(n - 1) + 1)}{\Gamma(1 - \alpha)\Gamma(\alpha n + 1)}, \quad n \in \mathbb{N}_0.
\]

From (19) it follows that \( e^{U_T} \) has the same moment sequence. Therefore, since the distribution of \( e^{U_T} \) is concentrated on \([0,1]\), it coincides with the distribution of \( \eta_{n\alpha}^\alpha \). Thus, 
\[
\left( 1 - \frac{S_{N_n - 1}}{n} \right)^\alpha \overset{d}{\rightarrow} e^{-U_T}. \tag{22}
\]

Now we point out a bivariate result generalizing (21) and (22).

**Proposition 2.6.** Suppose (8) holds. Then, 
\[
w^{-1}(n)(w(n - S_{N_n - 1}), N_n) \overset{d}{\rightarrow} (e^{-U_T}, \int_0^T e^{-U_t} \, dt),
\]
where \( \{U_t : t \geq 0\} \) is a drift-free subordinator with Lévy measure (10).

**Remark 2.7.** Corollary 3.3 in [34] states that
\[
\left( \frac{L(n)}{n^n} \left( N_{n+1} - 1, 1 - \frac{S_{N_{n+1} - 1}}{n} \right) \right) \overset{d}{\rightarrow} (X, Y), \tag{23}
\]
where the distribution of a random vector \((X, Y)\) was defined by the moment sequence. Our proof of Proposition 2.6 is different from and simpler than Port’s proof of (23).

**Proof.** According to Proposition 2.5 it suffices to verify that 
\[
\lim_{n \to \infty} \frac{E w^i(n - S_{N_n - 1}) N_n^j}{w^{i+j}(n)} = \frac{j! \Gamma(\alpha(i - 1) + 1)}{\Gamma(j + 1)(1 - \alpha)\Gamma(\alpha(i + j) + 1)}, \quad i, j \in \mathbb{N}_0. \tag{24}
\]
By Proposition 2.1
\[
\lim_{n \to \infty} \frac{L^k(n)}{n^{\alpha k}} \mathbb{E} N^k_n = \frac{k!}{\Gamma^k(1 - \alpha) \Gamma(1 + \alpha k)}, \quad k \in \mathbb{N}. \tag{25}
\]

For \( i = 0 \), Eq. (24) follows from (25). For \( i \in \mathbb{N} \), Eq. (24) is checked as follows.

\[
\mathbb{E} w^i(n - S_{N_n - 1}) N^j_n
\]
\[
= \sum_{k=1}^{n} \sum_{l=0}^{n-1} w^i(n - l) k^j \mathbb{P}\{N_n = k, S_{k-1} = l\}
\]
\[
= w^i(n) \mathbb{P}\{\xi \geq n\} + \sum_{l=1}^{n-1} w^i(n - l) \mathbb{P}\{\xi \geq n - l\} \sum_{k=2}^{l+1} k^j \mathbb{P}\{S_{k-1} = l\}
\]
\[
= w^i(n) \mathbb{P}\{\xi \geq n\} + \sum_{l=1}^{n-1} w^{i-1}(n - l) \sum_{k=2}^{l+1} k^j \mathbb{P}\{S_{k-1} = l\}.
\]

As on p. 26 in [1], define the function \( f(x) := 0 \) on \([0, 1)\) and \( f(x) := (k + 1)^j \) on \([k, k + 1)\) for \( k \in \mathbb{N} \), and set \( F(t) := \int_0^t f(x) \, dx \). Then,
\[
\sum_{l=1}^{n-1} \sum_{k=2}^{l+1} k^j \mathbb{P}\{S_{k-1} = l\} = \sum_{k=1}^{n-1} (k + 1)^j \mathbb{P}\{N_n > k\} = \mathbb{E} F(N_n).
\]

By Karamata’s theorem, \( F(t) \sim (j + 1)^{-1} t^{j+1} \). Since \( \lim_{n \to \infty} N_n = \infty \) almost surely and \( (N_n/w(n))^{j+1} \to \mathbb{E} \alpha^{j+1} \), we have
\[
\frac{F(N_n)}{w^{j+1}(n)} \overset{d}{\to} \frac{\alpha^{j+1}}{j + 1}. \tag{26}
\]

By (25), \( \lim_{n \to \infty} \mathbb{E} (N_n/w(n))^{j+2} = \mathbb{E} \alpha^{j+2} < \infty \). Therefore, the sequence \( \{F(N_n)/w^{j+1}(n) : n \in \mathbb{N}\} \) is uniformly integrable which together with (26) implies
\[
\mathbb{E} F(N_n) \sim \mathbb{E} \frac{\alpha^{j+1}}{j + 1} w^{j+1}(n) \sim \frac{j!}{\Gamma^{j+1}(1 - \alpha) \Gamma(1 + (j + 1)\alpha)} n^{\alpha(j+1)} L^{j+1}(n). \tag{27}
\]

Thus, if \( i = 1 \), we have
\[
\mathbb{E} w(n - S_{N_n - 1}) N^j_n \sim \frac{j!}{\Gamma^{j+1}(1 - \alpha) \Gamma(1 + (j + 1)\alpha)} n^{\alpha(j+1)} L^{j+1}(n),
\]
and (24) follows. Assume now that \( i \geq 2 \). Since \( w^{i-1}(n) \sim n^{\alpha(i-1)/L^{i-1}(n)} \), Corollary 1.7.3 in [9] yields

\[
W(s) := \sum_{n=1}^{\infty} s^n w^{i-1}(n) \sim \frac{\Gamma(1 + \alpha(i - 1))}{(1 - s)^{1+\alpha(i-1)} L^{i-1}((1 - s)^{-1})}, \quad s \uparrow 1.
\]

By the same Corollary, (27) implies

\[
R(s) := \sum_{n=1}^{\infty} s^n \left( \sum_{k=2}^{n+1} k^j \mathbb{P}\{S_{k-1} = l\} \right) \sim \frac{j!}{\Gamma^{j+1}(1 - \alpha)} \frac{1}{(1 - s)^{\alpha(j+1)} L^{j+1}((1 - s)^{-1})}, \quad s \uparrow 1.
\]

Therefore,

\[
W(s)R(s) \sim \frac{\Gamma(1 + \alpha(i - 1))j!}{\Gamma^{j+1}(1 - \alpha)} \frac{1}{(1 - s)^{1+\alpha(i+j)} L^{i+j}((1 - s)^{-1})}, \quad s \uparrow 1.
\]

The sequence \( \{w^{i-1}(n) : n \in \mathbb{N}\} \) is non-decreasing. Hence, the sequence \( \{\sum_{l=1}^{n-1} w^{i-1}(n - l) \sum_{k=2}^{l+1} k^j \mathbb{P}\{S_{k-1} = l\} : n = 2, 3, \ldots\} \) is non-decreasing too. Another appeal to Corollary 1.7.3 in [9] gives, as \( n \to \infty \),

\[
\sum_{l=1}^{n-1} w^{i-1}(n - l) \sum_{k=2}^{l+1} k^j \mathbb{P}\{S_{k-1} = l\} \sim \frac{\Gamma(1 + \alpha(i - 1))j!}{\Gamma^{j+1}(1 - \alpha) \Gamma(1 + \alpha(i + j))} \frac{n^{\alpha(i+j)}}{L^{i+j}(n)}.
\]

From this, (24) follows. \(\square\)

3 Proof of Theorem 1.4 and some comments

Nothing more than (2) and (3) is required for the proof given below.

For \( k, n \in \mathbb{N} \) set \( a_k(n) := \mathbb{E}M^k_n \) and \( b_k(n) := \mathbb{E}N^k_n \). For \( x \geq 0 \) define

\[
\Phi(x) := \frac{\Gamma(1 - \alpha) \Gamma(\alpha x + 1)}{\Gamma(\alpha(x - 1) + 1)} - 1 = \alpha x B(\alpha x, 1 - \alpha) - 1,
\]

where \( B \) denotes the beta function. Note that

\[
B(\alpha x, 1 - \alpha) = \int_0^1 y^{\alpha x - 1} (1 - y)^{-\alpha} \, dy = \alpha^{-1} \int_0^\infty e^{-xy}(1 - e^{-y/\alpha})^{-\alpha} \, dy
\]
and, hence,

$$\Phi(x) = \int_0^\infty xe^{-xy}(1 - e^{-y/\alpha})^{-\alpha} dy - 1$$

$$= \int_0^\infty (1 - e^{-y/\alpha})^{-\alpha} d(1 - e^{-xy}) - 1$$

$$= \int_0^\infty (1 - e^{-xy}) e^{-y/\alpha} \frac{e^{-y/\alpha}}{(1 - e^{-y/\alpha})^{\alpha+1}} dy. \quad (28)$$

Thus, the function $\Phi$ is the Laplace exponent of an infinitely divisible law with zero drift and Lévy measure $\nu$ given in (10).

Remark 3.1. On p. 102 in [7] it was stated that the right-hand side of (28) equals $\Phi(x) + 1$ (in our notation). Thus our formula (28) corrects that oversight.

Assuming that (8) holds we will prove that

$$\lim_{n \to \infty} \frac{L_k(n)}{n^k} a_k(n) = \frac{k!}{\Phi(1) \cdots \Phi(k)} =: a_k, \quad k \in \mathbb{N}. \quad (29)$$

This will imply (see, for example, [7]) that (i) $a_k = \mathbb{E} \eta^k, \ k \in \mathbb{N}$, where $\eta$ is a random variable with distribution of the exponential integral of a drift-free subordinator with Lévy measure $\nu$, and that (ii) the moments $\{a_n : n \in \mathbb{N}\}$ uniquely determine the law of $\eta$. Note that the statement in (i) was first obtained in Example 3.4 in [39]. From (i) and (ii) it will follow that (29) implies (9).

From (2) and (3) it follows that

$$a_1(n) = 1 + r_n \sum_{i=1}^{n-1} a_1(n - i)p_i,$$

and, for $k \in \{2, 3, \ldots\},$

$$a_k(n) = D_k(a_1(n), \ldots, a_{k-2}(n)) + ka_{k-1}(n) + r_n \sum_{i=1}^{n-1} a_k(n - i)p_i$$

$$=: d_k(n) + r_n \sum_{i=1}^{n-1} a_k(n - i)p_i, \quad (30)$$

where $D_k(\cdot)$ denotes the affine function of $k-2$ positive variables of the form

$$D_k(x_1, x_2, \ldots, x_{k-2}) = \gamma_{0,k} + \sum_{i=1}^{k-2} \gamma_{i,k} x_i,$$
with coefficients \( \gamma_{i,k} \in \mathbb{R}, \ i \in \{0,1,\ldots,k-2\} \) (these coefficients can be derived explicitly, but their exact values are of no use here), and \( r_n := 1/(p_1 + \cdots + p_{n-1}) \). Using the equality of distributions

\[
N_n \overset{d}{=} 1 + N'_n \xi \mathbb{1}_{\{\xi < n\}}, \quad n = 2,3,\ldots, \quad N_1 = 1,
\]

where \( \xi \) is independent of \( \{N'_n : n \in \mathbb{N}\} \), a copy of \( \{N_n : n \in \mathbb{N}\} \), we can show that

\[
b_k(n) = c_k(n) + \sum_{i=1}^{n-1} b_k(n-i) p_i, \quad k \in \mathbb{N},
\]

with \( c_1(n) := 1 \) and

\[
c_k(n) := D_k(b_1(n),\ldots,b_{k-2}(n)) + k b_{k-1}(n), \quad k \geq 2.
\]

To prove (29) we will use induction on \( k \). Suppose (29) holds for \( k \in \{1,2,\ldots,j-1\} \). Set

\[
\beta_1 := \frac{1}{1-b_1} \quad \text{and} \quad \beta_l := \frac{1}{b_{l-1} - l^{-1} b_l} \prod_{i=1}^{l-1} \frac{b_{i-1}}{b_{i-1} - i^{-1} b_i}, \quad l \in \{2,3,\ldots\},
\]

where \( b_l := l!/(\Gamma(l-\alpha)\Gamma(1+\alpha l)), \ l \in \mathbb{N} \), and note that

\[
a_{l-1} - \beta_l (b_{l-1} - l^{-1} b_l) = 0, \ l \in \mathbb{N}.
\]

In the following we exploit an idea given in the proof of Proposition 3 in [18]. Suppose there exists an \( \epsilon > 0 \) such that \( a_j(n) > (\beta_j + \epsilon)b_j(n) \) for infinitely many \( n \). It is possible to decrease \( \epsilon \) so that the inequality \( a_j(n) > (\beta_j + \epsilon)b_j(n) + c \) holds infinitely often for any fixed positive \( c \). Thus, we can define \( n_c := \inf\{n \geq 1 : a_j(n) > (\beta_j + \epsilon)b_j(n) + c\} \). Then

\[
a_j(n) \leq (\beta_j + \epsilon)b_j(n) + c \quad \text{for all} \ n \in \{1,2,\ldots,n_c - 1\}.
\]

We have

\[
(\beta_j + \epsilon)b_j(n_c) + c < a_j(n_c) \quad \overset{(30)}{=} \quad d_j(n_c) + r_{n_c-1} \sum_{i=1}^{n_c-1} a_j(n_c-i) p_i
\]

\[
\leq \quad d_j(n_c) + c + (\beta_j + \epsilon) r_{n_c-1} \sum_{i=1}^{n_c-1} b_j(n_c-i) p_i
\]

\[
\overset{(34)}{\leq} \quad D_j(a) + ja_j-1(n_c) + c
\]

\[
+ (\beta_j + \epsilon) (r_{n_c-1} - 1) (b_j(n_c) - D_j(b) - j b_j-1(n_c)) +
\]

\[
+ (\beta_j + \epsilon) b_j(n_c) - (\beta_j + \epsilon) (D_j(b) + j b_j-1(n_c)),
\]

17
or, equivalently,

\[
0 < D_j(a) + ja_{j-1}(n_c) + (\beta_j + \epsilon)(r_{n_c} - 1)(b_j(n_c) - jb_{j-1}(n_c))
- (\beta_j + \epsilon)(D_j(b) + jb_{j-1}(n_c)),
\]

where we have used the abbreviations \(D_j(a) := D_j(a_1(n_c), \ldots, a_{j-2}(n_c))\) and \(D_j(b) := D_j(b_1(n_c), \ldots, b_{j-2}(n_c))\) for convenience. Divide the latter inequality by \(z(c) := n_c^{(j-1)\alpha}/L^{j-1}(n_c)\) and let \(c\) go to \(\infty\) (which implies \(n_c \to \infty\)). Notice that, according to (8), \(r_n - 1 \sim n^{-\alpha}L(n)\) and that, by the induction assumption,

\[
\lim_{c \to \infty} \frac{D_j(a_1(n_c), \ldots, a_{j-2}(n_c))}{z(c)} = 0 \quad \text{and} \quad \lim_{c \to \infty} \frac{a_{j-1}(n_c)}{z(c)} = a_{j-1}.
\]

Using these facts and (25) we obtain

\[
0 \leq ja_{j-1} + (\beta_j + \epsilon)b_j - (\beta_j + \epsilon)jb_{j-1}.
\]

Since the function \(\Phi\) defined at the beginning of the proof is positive for \(x > 0\), and \(jb_{j-1}/b_j - 1 = \Phi(j)\), we conclude that \(jb_{j-1} - b_j > 0\). Therefore,

\[
\epsilon(jb_{j-1} - b_j) \leq j(a_{j-1} - \beta_j(b_{j-1} - j^{-1}b_j)) = 0
\]

by (33). This is the desired contradiction. Thus, we have verified that

\[
\limsup_{n \to \infty} \frac{a_j(n)}{b_j(n)} \leq \beta_j.
\]

A symmetric argument proves the converse inequality for the lower bound. Therefore,

\[
a_j(n) \sim \beta_j b_j(n) \sim \beta_j b_j n^j \alpha = a_j \frac{n^{j\alpha}}{L(n)}.
\]

A similar but simpler reasoning yields the result for \(k = 1\). We omit the details. The proof is complete.

The proof above only exhibits the limiting law but does not give any insight why it is the law of an exponential functional. We intend to explore this issue now in some more detail. Remarkably enough, it seems that we have found a new area where perpetuities appear in a natural way.

Fix \(i, j \in \mathbb{N}\). Define \(\tilde{R}^{(j)}_0(i) := 0\),

\[
\tilde{R}^{(j)}_k(i) := \tilde{R}^{(j)}_{k-1}(i) + \xi_{i+k}1_{\{R^{(j)}_{k-1}(i) + \xi_{i+k} < j\}}, \quad k \in \mathbb{N},
\]
and
\[ \hat{M}_n(i) := \sum_{l=0}^{\infty} \{ \hat{R}_l^{(n)}(i) + \xi_{i+l+1} < n \}, \quad n \in \mathbb{N}. \]

Set also \( Y_n := n - S_{N_n-1}. \)

The subsequent argument relies upon the following decomposition (36).

**Lemma 3.2.** For fixed \( n \in \mathbb{N} \) and any \( i \in \mathbb{N}, \)
\[ \hat{M}_n(i) \overset{d}{=} M_n, \quad (35) \]

and
\[ M_n - N_n + 1 = \hat{M}_{Y_n}(N_n) \overset{d}{=} M'_{Y_n}, \quad (36) \]

where \( \{ M'_n : n \in \mathbb{N} \} \) has the same law as \( \{ M_n : n \in \mathbb{N} \} \) and is independent of \( (N_n, Y_n) \).

**Proof.** We have
\[
M_n = \sum_{l=0}^{\infty} 1_{\{ R_l^{(n)} + \xi_{i+l+1} < n \}} = \sum_{l=0}^{N_n-2} 1 + \sum_{l=N_n}^{\infty} 1_{\{ R_l^{(n)} + \xi_{i+l+1} < n \}} = N_n - 1 + \sum_{l=0}^{\infty} 1_{\{ \hat{R}_l^{(n)}(N_n) + \xi_{N_n+i+l+1} < Y_n \}} = N_n - 1 + \hat{M}_{Y_n}(N_n),
\]
and the first equality in (36) follows. For any fixed \( k \in \mathbb{N}, \)
\[
P\{ \hat{M}_{Y_n}(N_n) = k \}
= \sum_{i=1}^{n} \sum_{j=0}^{n-1} P\{ \hat{M}_{N} = l, N = i, S_{N-1} = j \}
= \sum_{i=1}^{n} \sum_{j=0}^{n-1} P\{ \sum_{l=0}^{\infty} 1_{\{ \hat{R}_l^{(n-j)}(i) + \xi_{i+l+1} < n-j \}} = k, N = i, S_{N-1} = j \}.
\]

The sequence \( \{ \hat{R}_l^{(n-j)}(i) + \xi_{i+l+1} : l \in \mathbb{N}_0 \} \) is independent of \( 1_{\{ N_i = i, S_{N_i-1} = j \}} \)
and has the same law as \( \{ (R_l^{(n-j)})' + \xi_{l+1}' : l \in \mathbb{N}_0 \} \), where \( \{(R_l^{(i)})' : l \in \mathbb{N}_0 \} \)
is constructed in the same way as the sequence without ”prime” by using
\{\xi'_k : k \in \mathbb{N}\}, an independent copy of \{\xi_k : k \in \mathbb{N}\}. This implies (35) and

\[
\mathbb{P}\{\tilde{M}_{Y_n}(N_n) = k\} = \sum_{i=1}^{n} \sum_{j=0}^{n-1} \mathbb{P}\{\sum_{l=0}^{\infty} 1_{\left\{(R_{l}^{(n-j)})'+\xi_{l+1}'<n-j\right\}} = k\} \mathbb{P}\{N_n = i, S_{N_n-1} = j\}
\]

\[
= \mathbb{P}\{\sum_{l=0}^{\infty} 1_{\left\{(R_{l}^{Y_n})'+\xi_{l+1}'<Y_n\right\}} = k\} = \mathbb{P}\{M'_Y = k\},
\]

and the second equality in distribution in (36) follows. \(\square\)

Set \(t(n) := n^\alpha / L(n)\). From the proof above, we already know that \(M_n/t(n)\) converges in law to a random variable \(Z\), say, with a proper law. From \(Y_n \xrightarrow{D} +\infty\) and the result of Lemma 3.2 we conclude that \(\tilde{M}_{Y_n}/t(Y_n)\) converges in law to a random variable \(Z'' \overset{d}{=} Z\). By Proposition 2.6,

\[
\left(\frac{t(Y_n)}{t(n), N_n-1}{t(n)}\right) \overset{d}{\rightarrow} (M, Q) := \left(e^{-U_T}, \int_0^T e^{-U_t} dt\right).
\]

Rewriting (36) in the form

\[
\frac{M_n}{t(n)} = \frac{\tilde{M}_{Y_n} t(Y_n)}{t(Y_n) t(n)} + \frac{N_n - 1}{t(n)}
\]

we conclude that

\[
\left(\frac{\tilde{M}_{Y_n}}{t(Y_n)}, \frac{t(Y_n)}{t(n)}, \frac{N_n - 1}{t(n)}\right) \overset{d}{\rightarrow} (Z', M, Q),
\]

where \(Z' \overset{d}{=} Z\) and using characteristic functions it can be checked that \(Z'\) is independent of \((M, Q)\). Furthermore,

\[
Z \overset{d}{=} MZ' + Q. \tag{37}
\]

From (17) it follows that the distribution of \(\int_0^\infty e^{-U_t} dt\) is a solution of (37). By Theorem 1.5 (i) in [40] this solution is unique. Therefore,

\[
\frac{M_n}{t(n)} \overset{d}{\rightarrow} \int_0^\infty e^{-U_t} dt.
\]

In a similar way, we can prove the following.
Corollary 3.3. Suppose (8) holds. Then,
\[
\left( \frac{M_n - N_n}{t(n - S_{N_n - 1})}, \frac{t(n - S_{N_n - 1})}{t(n)}, \frac{N_n}{t(n)} \right) \n \rightarrow \left( \int_0^\infty e^{-(U_t + T - U_T)} dt, e^{-U_T}, \int_0^T e^{-U_i} dt \right).
\]
Furthermore, \((M_n - N_n) / t(n - S_{N_n - 1})\) and \((t(n - S_{N_n - 1}) / t(n), N_n / t(n))\) are asymptotically independent, and
\[
t^{-1}_n(M_n - N_n, N_n) \n \rightarrow \left( \int_0^\infty e^{-U_i} dt, \int_0^T e^{-U_i} dt \right).
\]

4 Proof of Theorem 1.2

Our proof essentially relies upon the following classical result
\[
\lim_{n \rightarrow \infty} P\{n - S_{N_n - 1} \leq k\} = m^{-1} \sum_{i=1}^k P\{\xi \geq i\} =: P\{W \leq k\}, \ k \in \mathbb{N}. \quad (38)
\]
In order to see why (38) holds, note that
\[
P\{n - S_{N_n - 1} = k\} = \sum_{i=1}^n P\{S_{i-1} = n - k, S_i \geq n\}
\]
\[
= P\{\xi \geq k\} \sum_{i=0}^{n-k} P\{S_i = n - k\}
\]
\[
\rightarrow m^{-1} P\{\xi \geq k\}, \ n \rightarrow \infty,
\]
by the elementary renewal theorem, and (38) follows.

From (36) we conclude that
\[
M_n - N_n \n \rightarrow \left( M'_W - 1 \right), \quad (39)
\]
where \(W\) is a random variable with distribution (38) which is independent of \(\{M'_n : n \in \mathbb{N}\}\). Therefore, for any sequence \(\{d_n : n \in \mathbb{N}\}\) such that \(\lim_{n \rightarrow \infty} d_n = \infty\),
\[
\frac{M_n - N_n}{d_n} \n \rightarrow \left( M'_W - 1 \right), \quad (40)
\]
Assume that the distribution of \(\xi\) does not belong to the domain of attraction of any stable law with index \(\alpha \in [1, 2]\). Then, as is well known, it is not
possible to find sequences $x_n > 0$ and $y_n \in \mathbb{R}$ such that $(S_n - y_n)/x_n$
converges to a proper and non-degenerate law. In view of

$$\mathbb{P}\{N_n > m\} = \mathbb{P}\{S_m \leq n - 1\}, \quad (41)$$

the same is true for $N_n$ (see Theorem 7 in [17] and/or Theorem 2 in [23] for
more details), and according to (40), for $M_n$.

Assume that conditions (ii) of Theorem 1.2 hold. If $\sigma^2 = \infty$ and (7)
holds with $\alpha = 2$, then arguing as in the proof of Theorem 2 in [23] we
conclude that, with $a_n$ and $b_n$ defined in our Theorem 1.2,

$$\frac{N_n - b_n}{a_n} \Rightarrow \mu_2.$$

Theorem 5 in [17] (if $\sigma^2 < \infty$) and Theorem 7 in [17] (if (7) holds for some
$\alpha \in [1, 2)$) leads to the same limiting relation (with corresponding $a_n$ and $b_n$, and with $\mu_2$ replaced by $\mu_\alpha$ in the latter case).

In view of (40) the same limiting relations hold for $M_n$. The proof of
Theorem 1.2 is complete.

5 Proof of Theorem 1.1

Assume first that $m = \infty$. According to (14), $\mathbb{E}N^k_n \sim n^k/L^k(n), k \in \mathbb{N}$. The
same argument as in Section 3 yields

$$\mathbb{E}M^k_n \sim \frac{n^k}{L^k(n)} \sim (\mathbb{E}M_n)^k, \quad k \in \mathbb{N}.$$  

Therefore,

$$\lim_{n \to \infty} \mathbb{E} \left( \frac{M_n}{\mathbb{E}M_n} \right)^k = 1, \quad k \in \mathbb{N},$$

which proves (5). In fact, to arrive at (5), it suffices to know that $\mathbb{E}M_n \sim n/L(n)$ and $\mathbb{E}M^2_n \sim n^2/L^2(n)$ and exploit Chebyshev’s inequality.

Assume now that $m < \infty$. It is well known that

$$\lim_{n \to \infty} \frac{N_n}{n} = \frac{1}{m} \quad (42)$$

In view of (39), $\lim_{n \to \infty}(M_n - N_n)/n = 0$ almost surely, which yields
$\lim_{n \to \infty} M_n/n = 1/m$ almost surely. By the elementary renewal theorem,
$\mathbb{E}N_n \sim n/m$. Using the same approach as in Section 3, it is straightforward
to check that $\mathbb{E}M_n \sim n/m$. Conversely, if $M_n/a_n \overset{P}{\to} 1$, then (40) gives
$(M_n - N_n)/a_n \overset{P}{\to} 0$. Therefore, $N_n/a_n \overset{P}{\to} 1$. An appeal to (42) allows us to
conclude that $a_n \sim n/m$. The proof is complete.
6 Proof of Theorem 1.5

By Theorem 3 (c) and formulae on p. 42 in [8] (see also [22])

\[
\frac{N_n - b(n) - 1}{a(n)} \Rightarrow \mu_1,
\]

where \( \mu_1 \) is the 1-stable law with characteristic function \( \int_{-\infty}^{\infty} e^{iux} \mu_1(dx) = \exp(it \log |t| - |t| \pi/2), \ t \in \mathbb{R} \). By Corollary 2.3,

\[
\frac{M_n}{N_n - 1} \overset{P}{\rightarrow} 1. \tag{43}
\]

Therefore,

\[
\frac{M_n - b(n)}{a(n)} - \frac{M_n - N_n + 1 b(n)}{N_n - 1} \overset{P}{\Rightarrow} \mu_1.
\]

Thus, to prove the theorem it suffices to show that the second summand tends to 0 in probability. Clearly, this can be regarded as a rate of convergence result for (43). Recalling the notation \( Y_n = n - S_{N_n - 1} \) and using (36) gives

\[
\frac{M_n - N_n + 1 b(n)}{N_n - 1} \overset{P}{\Rightarrow} \mu_1.
\]

By Corollary 2.3, \( m(n)M_n/n \overset{P}{\rightarrow} 1 \). Using the equality of distributions (36) and the fact that \( Y_n \overset{P}{\rightarrow} \infty \) allows us to conclude that \( K_1(n) \overset{P}{\rightarrow} 1 \). By Theorem 6 in [16], \( K_2(n) \overset{d}{\rightarrow} 1/R \), where \( R \) is a random variable uniformly distributed on \([0, 1]\). By Proposition 2.4, \( K_3(n) \overset{P}{\rightarrow} 0 \). Finally, by Corollary 2.3, \( K_4(n) \overset{P}{\rightarrow} 1 \). The proof is complete.

7 Number of collisions in beta coalescents

In this section the main results presented in Section 1 are applied to the number of collisions that take place in beta coalescent processes until there is just a single block. Other closely related functionals of coalescent processes such as the total branch length or the number of segregating sites have been studied in [6], [13], [14] and [28] (see also [5]).
Let $\mathcal{E}$ denote the set of all equivalence relations on $\mathbb{N}$. For $n \in \mathbb{N}$ let $\varrho_n : \mathcal{E} \to \mathcal{E}_n$, denote the natural restriction to the set $\mathcal{E}_n$ of all equivalence relations on $\{1, \ldots, n\}$. For $\eta \in \mathcal{E}_n$ let $|\eta|$ denote the number of blocks (equivalence classes) of $\eta$.

Pitman [33] and Sagitov [38] independently introduced coalescent processes with multiple collisions. These Markovian processes with state space $\mathcal{E}$ are characterized by a finite measure $\Lambda$ on $[0, 1]$ and are, hence, also called $\Lambda$-coalescent processes. For a $\Lambda$-coalescent $\{\Pi_t : t \geq 0\}$, it is known that the process $\{|\varrho_{n}\Pi_t| : t \geq 0\}$ has infinitesimal rates

$$g_{nk} := \lim_{t \downarrow 0} \frac{\mathbb{P}\{|\varrho_{n}\Pi_t| = k\}}{t} = \left(\frac{n}{k-1}\right) \int_{[0,1]} x^{n-k-1}(1-x)^{k-1}\Lambda(dx)$$

for all $k, n \in \mathbb{N}$ with $k < n$. Let $g_n := \sum_{k=1}^{n-1} g_{nk}$, $n \in \mathbb{N}$, denote the total rates. We are interested in the number $X_n$ of collisions (jumps) that take place in the restricted coalescent process $\{|\varrho_{n}\Pi_t| : t \geq 0\}$ until there is just a single block. From the structure of the coalescent process it follows that $(X_n)_{n \in \mathbb{N}}$ satisfies the distributional recursion $X_1 = 0$ and $X_n \overset{d}{=} 1 + X_{n-I_n}$, $n \in \{2, 3, \ldots\}$, where $I_n$ is independent of $X_2, \ldots, X_{n-1}$ with distribution $\mathbb{P}\{I_n = k\} = g_{n,n-k}/g_n$, $k \in \{1, \ldots, n-1\}$. The random variable $n-I_n$ is the (random) state of the process $\{|\varrho_{n}\Pi_t| : t \geq 0\}$ after its first jump.

We consider beta coalescents, where, by definition, $\Lambda = \beta(a,b)$ is the beta distribution with density $x \mapsto (B(a,b))^{-1}x^{a-1}(1-x)^{b-1}$ with respect to the Lebesgue measure on $(0, 1)$, and $B(a,b) := \Gamma(a)\Gamma(b)/\Gamma(a+b)$ denotes the beta function, $a, b > 0$. In this case the rates (44) have the form

$$g_{nk} = \left(\frac{n}{k-1}\right) \frac{1}{B(a,b)} \int_{0}^{1} x^{a+n-k-2}(1-x)^{b+k-2} \, dx$$

$$= \left(\frac{n}{k-1}\right) \frac{B(a+n-k-1,b+k-1)}{B(a,b)}, \quad k, n \in \mathbb{N}, k < n. \tag{45}$$

From

$$g_{k+1,k} = \frac{k(k+1)}{2} \frac{B(a,b+k-1)}{B(a,b)}$$

it follows that

$$g_n = \sum_{k=1}^{n-1} (g_{k+1,k} - g_k) = \sum_{k=1}^{n-1} \frac{2}{k+1} g_{k+1,k} = \frac{1}{B(a,b)} \sum_{k=1}^{n-1} kB(a,b+k-1).$$

In the following it is assumed that $b = 1$ such that the rates (45) reduce to

$$g_{nk} = \left(\frac{n}{k-1}\right) \frac{B(a+n-k-1,1)}{B(a,1)} = \frac{n!}{(n-k+1)!} \frac{\Gamma(a+n-k-1)}{\Gamma(a+n-1)},$$

24
and the total rates to
\[ g_n = a \sum_{k=1}^{n-1} kB(a,k) = \begin{cases} 
\frac{a}{a-2} \left(1 - \frac{\Gamma(a)\Gamma(n+1)}{\Gamma(a+n-1)}\right) & \text{for } a > 0, a \neq 2, \\
2(h_n - 1) & \text{for } a = 2.
\end{cases} \]

Here, \( h_n := \sum_{i=1}^{n} 1/i \) denotes the \( n \)-th harmonic number. From the last formula it follows that the parameter \( a = 2 \) plays a special role in this model. Define
\[ p_k := \frac{(2-a)\Gamma(a+k-1)}{\Gamma(a)\Gamma(k+2)}, \quad k \in \mathbb{N}. \]

Assume now that \( 0 < a < 2 \). In this case (and only in this case) we have \( p_k \geq 0 \) for \( k \in \mathbb{N} \), \( \sum_{k=1}^{\infty} p_k = 1 \) and (3) holds. Let \( \xi \) be a random variable with distribution \( \mathbb{P}\{\xi = k\} = p_k, k \in \mathbb{N} \). It follows by induction on \( n \) that
\[ \mathbb{P}\{\xi \geq n\} = \frac{\Gamma(a+n-1)}{\Gamma(a)\Gamma(n+1)}, \quad n \in \mathbb{N}. \]

Using \( \Gamma(n+x) \sim \Gamma(n)x^x \) for \( n \to \infty \), we conclude that
\[ \mathbb{P}\{\xi \geq n\} \sim \frac{n^{a-2}}{\Gamma(a)} = \frac{n^{-\alpha}}{\Gamma(2-\alpha)}, \quad n \to \infty. \]

Thus, if \( 1 < a < 2 \), or, equivalently, \( 0 < \alpha < 1 \), Theorem 1.4 is applicable (with \( L(n) \equiv 1/\Gamma(a) = 1/\Gamma(2-\alpha) \)), and we obtain the following result.

**Theorem 7.1.** For the \( \beta(a,1) \)-coalescent with \( 1 < a < 2 \), i.e., \( 0 < \alpha := 2-a < 1 \), the number \( X_n \) of collision events satisfies
\[ \frac{X_n}{\Gamma(2-\alpha)n^{\alpha}} \xrightarrow{d} \int_{0}^{\infty} e^{-U_t} dt, \]
where \( \{U_t : t \geq 0\} \) is a drift-free subordinator with Lévy measure (10).

Note that, for \( \Lambda = \beta(a,b) \), we have \( \mu_{-1} := \int x^{-1} \Lambda(dx) < \infty \) if and only if \( a > 1 \). Under the condition \( \mu_{-1} < \infty \), limiting results similar to that presented in the above Theorem 7.1 are known for the number of segregating sites (see, for example, Proposition 5.1 in [28]) for general \( \Lambda \)-coalescent processes with mutation.

Assume now that \( 0 < a < 1 \). Then, \( m := \mathbb{E}\xi = 1/(1-a) < \infty \). It is straightforward to verify that
\[ \sum_{k=1}^{n} k^2 p_k \sim \frac{2-a}{\Gamma(a+1)} n^a, \quad n \to \infty. \]
In particular, the variance of ξ is infinite. Thus, Theorem 1.2 is applicable (with $L(n) \equiv (2 - a)/\Gamma(a + 1) = \alpha/\Gamma(3 - \alpha)$, $C := 1/\Gamma(\alpha) = 1/\Gamma(2 - \alpha)$, $b_n := n(1 - a) = n(\alpha - 1)$ and $c_n := n^{1/\alpha}$), and yields the following result.

**Theorem 7.2.** For the $\beta(a,1)$-coalescent with $0 < a < 1$, i.e., $1 < \alpha := 2 - a < 2$, the number $X_n$ of collision events satisfies

$$X_n - n(\alpha - 1)/(\alpha - 1)^{(\alpha+1)/\alpha n^{1/\alpha}} \Rightarrow \mu_\alpha,$$

or, equivalently,

$$X_n - n(\alpha - 1)/(\alpha - 1)n^{1/\alpha} \overset{d}{\to} S_\alpha,$$

where $E \exp(itS_\alpha) = \exp(|t|^\alpha (\cos(\pi\alpha/2) + i \sin(\pi\alpha/2) \text{sgn}(t)))$, $t \in \mathbb{R}$.

Gnedin and Yakubovich [21, Theorem 9] use analytic methods to verify the same convergence result (47) for $\Lambda$-coalescents satisfying $\Lambda([0,x]) = Ax^a + O(x^{a+\zeta})$, $x \to 0$, $0 < a < 1$, $\zeta > \max\{(2 - a)^2/(5 - 5a + a^2), 1 - a\}$.

Theorems 7.1 and 7.2 do not cover the asymptotics of $X_n$ for the Bolthausen-Sznitman coalescent, i.e. the $\beta(a,b)$-coalescent with $a = b = 1$. The limiting behaviour of $X_n$ for the Bolthausen-Sznitman coalescent was studied in [26], and follows also from our Theorem 1.5 with $p_k := 1/(k(k + 1))$, $L(n) \equiv 1$, $c(x) := x$, $b(x) := x/\log x + x \log \log x/(\log x)^2$, and $a(x) := b^2(x)/x \sim x/(\log x)^2$. Therefore, the asymptotics of $X_n$ for all $\beta(a,1)$-coalescent processes with $0 < a < 2$ is clarified. Unfortunately, our method cannot be used to treat the asymptotics of $X_n$ for $\beta(a,1)$-coalescent processes with $a \geq 2$, as in this case the condition (3) is not satisfied. Recently, the limiting behaviour of $X_n$ for $\beta(2,b)$-coalescents with parameter $b > 0$ was obtained in [25] using a completely different approach based on asymptotics of moments.

**Acknowledgement.** The authors thank Alexander Gnedin for fruitful comments and discussions, in particular for pointing out an error in the first version of the manuscript. We furthermore thank the anonymous referee for his profound report leading to a considerable improvement of the style and the presentation of the paper.

**References**


