THE TIME TO ABSORPTION IN Λ-COALESCENTS

By GÖTZ KERSTING and ANTON WAKOLBINGER

Abstract

We present a law of large numbers and a central limit theorem for the time to absorption of Λ-coalescents with dust started from n blocks, as $n \to \infty$. The proofs rely on an approximation of the logarithm of the block-counting process by means of a drifted subordinator.

Keywords: Coalescent; time to absorption; law of large numbers; central limit theorem; subordinator with drift

2010 Mathematics Subject Classification: Primary 60J75
Secondary 60J27; 60F05

1. Introduction and main results

Given a large sample of individuals with a common ancestor, how long are the ancestral lineages back to that ancestor? For a population of constant size, this question concerns the absorption time of a *coalescent* which describes the genealogical tree of n individuals by means of merging partitions. Here we consider coalescents with multiple mergers, also known as Λ-coalescents, as introduced in 1999 by Pitman [6] and Sagitov [7]. If Λ is a finite, nonzero measure on [0, 1], then the Λ-coalescent started with n blocks is a continuous-time Markov chain $(\Pi_n(t),\ t \ge 0)$ taking its values in the set of partitions of $\{1, \dots, n\}$. It has the property that whenever there are b blocks, each possible transition that involves merging $k \ge 2$ of the blocks into a single block happens at rate

$$\lambda_{b,k} = \int_{[0,1]} p^k (1-p)^{b-k} \frac{\Lambda(\mathrm{d}p)}{p^2},$$

and these are the only possible transitions. Let $N_n(t)$ be the number of blocks in the partition $\Pi_n(t),\ t \ge 0$. Then

$$\tau_n := \inf\{t \ge 0 \colon N_n(t) = 1\}$$

is the time of the last merger, also called the *absorption time* of the coalescent started in n blocks. In this paper we study the asymptotic distribution of τ_n as $n \to \infty$.

Our first result is a law of large numbers for the times τ_n. Let

$$\mu := \int_{[0,1]} |\log(1-p)| \frac{\Lambda(\mathrm{d}p)}{p^2},$$

in particular, $\mu = \infty$ when $\Lambda(\{0\}) > 0$ or $\Lambda(\{1\}) > 0$.

Theorem 1. *For any Λ-coalescent, as $n \to \infty$,*

$$\frac{\tau_n}{\log n} \to \frac{1}{\mu} \quad \textit{in probability.} \tag{1}$$

*doi:*10.1017/apr.2018.78

This theorem says that in a Λ-coalescent the number of blocks decays at least at an exponential rate. If $\mu = \infty$ then the right-hand limit is 0, and the coalescent decreases even super-exponentially fast. The $\mu < \infty$ case is equivalently captured by the simultaneous validity of the two conditions

$$\int_{[0,1]} \frac{\Lambda(\mathrm{d}p)}{p} < \infty \quad \text{and} \quad \int_{[0,1]} |\log(1-p)|\Lambda(\mathrm{d}p) < \infty.$$

The first condition is a requirement on Λ in the neighbourhood of 0: it prohibits a swarm of small mergers (these can occur in coalescents coming down from ∞, meaning that the τ_n are bounded in probability uniformly in n). The second is a condition on Λ in the vicinity of 1: it rules out the possibility of mergers which, although appearing only occasionally, are so vast that they make the coalescent collapse. Herriger and Möhle [3] obtained a counterpart to Theorem 1 in which τ_n in (1) is replaced by its expectation.

Our second result is a central limit theorem. Here we confine ourselves to coalescents with $\mu < \infty$. Then the function

$$f(y) := \int_{[0,1]} \frac{1-(1-p)^{\mathrm{e}^y}}{\mathrm{e}^y} \frac{\Lambda(\mathrm{d}p)}{p^2}, \qquad y \in \mathbb{R}, \tag{2}$$

is everywhere finite. Also, f is a positive, monotone decreasing, continuous function with the property $f(y) \to 0$ for $y \to \infty$. Let

$$b_n := \int_{\kappa}^{\log n} \frac{\mathrm{d}y}{\mu - f(y)},$$

where we choose $\kappa \geq 0$ such that

$$f(y) \leq \tfrac{1}{2}\mu \quad \text{for all } y \geq \kappa.$$

Theorem 2. *Assume that $\mu < \infty$ and, moreover,*

$$\sigma^2 := \int_{[0,1]} (\log(1-p))^2 \frac{\Lambda(\mathrm{d}p)}{p^2} < \infty.$$

Then, as $n \to \infty$,

$$\frac{\tau_n - b_n}{\sqrt{\log n}} \xrightarrow{\mathrm{D}} N\left(0, \frac{\sigma^2}{\mu^3}\right). \tag{3}$$

Under the additional condition

$$\int_{[0,1]} |\log p| \frac{\Lambda(\mathrm{d}p)}{p} < \infty, \tag{4}$$

Gnedin *et al.* [1] obtained the CLT (3) with b_n replaced by $(\log n)/\mu$ (condition (9) in [1] is equivalent to the condition at (4); see [4, Remark 13]). Thus, the question arises, whether the simplified centering by $(\log n)/\mu$ is always feasible. The next proposition shows that this can be done under a condition that is weaker than (4), but not in every case.

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

Proposition 1. *Let* $0 \le c < \infty$. *Then*

$$b_n = \frac{\log n}{\mu} + \frac{2c}{\mu^2}\sqrt{\log n} + o(\sqrt{\log n}) \quad \text{as } n \to \infty \tag{5}$$

if and only if

$$\sqrt{|\log r|}\int_{[0,r]} \frac{\Lambda(\mathrm{d}p)}{p} \to c \quad \text{as } r \to 0. \tag{6}$$

Example. Consider for $\gamma \in \mathbb{R}$ the finite measures

$$\Lambda(\mathrm{d}p) = \left(1 + \log\frac{1}{p}\right)^{-\gamma} \mathrm{d}p, \qquad 0 \le p \le 1.$$

For $\gamma = 0$, this gives the Bolthausen–Sznitman coalescent. For $\gamma > 1$, it yields coalescents with $\mu, \sigma^2 < \infty$. Note that (4) is satisfied if and only if $\gamma > 2$, and (6) is fulfilled if and only if $\gamma > \frac{3}{2}$. Thus, within the range $1 < \gamma \le \frac{3}{2}$, we have to come back to the constants b_n in the central limit theorem.

The law of large numbers from Theorem 1 holds for all $\gamma > 1$. For the regime $\gamma \le 1$, Theorem 1 tells us only that $\tau_n = o_P(\log n)$. For $\gamma = 0$, the Bolthausen–Sznitman coalescent, it is known that τ_n is already down to the order $\log\log n$ [2]. For $\gamma < 0$, applying Schweinsberg's criterion [8], it can be shown that the coalescents come down from ∞. There remains the gap $0 < \gamma \le 1$. It is tempting to conjecture that τ_n is of order $(\log n)^\gamma$ for $0 < \gamma < 1$.

When equation (6) does not hold, then the approximation to b_n that follows may be practical. Starting from the identity

$$\frac{1}{\mu - f(y)} = \frac{1}{\mu} + \frac{f(y)}{\mu^2} + \frac{f^2(y)}{\mu^3} + \cdots + \frac{f^k(y)}{\mu^{k+1}} + \frac{f^{k+1}(y)}{\mu^{k+1}(\mu - f(y))},$$

we obtain the expansion

$$b_n = \frac{\log n}{\mu} + \frac{1}{\mu^2}\int_0^{\log n} f(y)\,\mathrm{d}y + \cdots + \frac{1}{\mu^{k+1}}\int_0^{\log n} f^k(y)\,\mathrm{d}y + O\left(\int_0^{\log n} f^{k+1}(y)\,\mathrm{d}y\right).$$

We now explain the method of proving Theorems 1 and 2. We are dealing mainly with Λ-coalescents that have a *dust component*. Briefly speaking, these are the coalescents for which the rate at which a single lineage merges with some others from the sample remains bounded as the sample size tends to ∞. It is well known (see, e.g. [6, Theorem 8]) that this property is characterized by the condition

$$\int_{[0,1]} \frac{\Lambda(\mathrm{d}p)}{p} < \infty. \tag{7}$$

An established tool for the analysis of a Λ-coalescent with dust is the subordinator $S = (S_t)_{t\ge 0}$; this is used to approximate the logarithm of its block-counting process $N_n = (N_n(t))_{t\ge 0}$ (see, e.g. [1], [5], and [6]). We recall this subordinator in Section 3. Indeed, analogues of Theorems 1 and 2 are well known for first-passage times of subordinators with finite first and second moments, respectively, but this approximation neglects the subtlety that a coalescent of b lineages results in a downward jump of size $b - 1$ (and not b) for the process N_n. This effect

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

becomes significant when many small jumps accumulate over time, as happens close to the dustless case (this is readily seen in Proposition 1 and the above example). Then the appropriate approximation is provided by a *drifted* subordinator $Y_n = (Y_n(t))_{t\geq 0}$, given by the stochastic differential equation

$$Y_n(t) = \log n - S_t + \int_0^t f(Y_n(s))\,\mathrm{d}s, \qquad t \geq 0,$$

with initial value $Y_n(0) = \log n$. The drift compensates the difference between b and $b-1$ just mentioned. In Kersting *et al.* [4] it was shown that

$$\sup_{t<\tau_n} |Y_n(t) - \log N_n(t)| = O_P(1) \quad \text{as } n \to \infty,$$

that is, these random variables are bounded in probability. In Section 3 we suitably strengthen this result. In Section 2 we provide the required limit theorems for passage times for a more general class of drifted subordinators. The above results are then proved in Section 4.

It turns out that the regime considered by Gnedin *et al.* [1] is one in which the random variables $\int_0^{\tau_n} f(Y_n(s))\,\mathrm{d}s$ are bounded in probability uniformly in n. This can be seen to be equivalent to the requirement $\int_0^\infty f(y)\,\mathrm{d}y < \infty$, which is likewise equivalent to (4) (see the proof of [4, Corollary 12]). Under this assumption, Gnedin *et al.* [1] proved their central limit theorem also with nonnormal (stable or Mittag-Leffler) limiting distributions for τ_n. A similar generalization of Theorem 2 is feasible in the general dust case, without requirement (4).

2. Limit theorems for a drifted subordinator

Let $S = (S_t)_{t\geq 0}$ be a pure-jump subordinator with Lévy measure λ on $(0, \infty)$. Recall that this requires

$$\int_0^\infty (y \wedge 1)\lambda(\mathrm{d}y) < \infty.$$

More generally than the specific function in (2), let $f\colon \mathbb{R} \to \mathbb{R}$ be an arbitrary positive, nonincreasing, continuous function with

$$\lim_{y\to\infty} f(y) = 0.$$

Let the process $Y^z = (Y_t^z)_{t\geq 0}$ denote the unique solution of the stochastic differential equation

$$Y_t^z = z - S_t + \int_0^t f(Y_s^z)\,\mathrm{d}s \tag{8}$$

with initial value $z > 0$; we investigate the asymptotic behaviour in the limit $z \to \infty$ of its passage times across $x \in \mathbb{R}$, that is, of

$$T_x^z := \inf\{t \geq 0\colon Y_t^z < x\}.$$

Our first result provides a law of large numbers. Denote

$$\mu := \int_{(0,\infty)} y\lambda(\mathrm{d}y). \tag{9}$$

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

Proposition 2. *Assume that $\mu < \infty$. Then, for any $x \in \mathbb{R}$, as $z \to \infty$,*

$$\frac{1}{z} T_x^z \to \frac{1}{\mu} \quad \textit{in probability.}$$

Proof. Let $z > x$. Then

$$\{T_x^z \geq t\} = \{Y_s^z \geq x \text{ for all } s \leq t\} = \left\{S_s \leq z - x + \int_0^s f(Y_u^z)\,\mathrm{d}u \text{ for all } s \leq t\right\}.$$

The positivity of f implies that $\mathbb{P}(T_x^z \geq t) \geq \mathbb{P}(S_t \leq z - x)$, so, for any $\varepsilon > 0$,

$$\mathbb{P}\left(T_x^z \geq (1-\varepsilon)\frac{z}{\mu}\right) \geq \mathbb{P}(S_{(1-\varepsilon)z/\mu} \leq z - x). \tag{10}$$

Now $\mu = \mathbb{E}[S_1]$, so, by the law of large numbers,

$$\lim_{t\to\infty} \frac{S_t}{t} = \mu \quad \text{a.s.;}$$

hence, the right-hand term in (10) converges to 1 as $z \to \infty$ and also

$$\mathbb{P}\left(T_x^z \geq (1-\varepsilon)\frac{z}{\mu}\right) \to 1.$$

On the other hand, $\{T_x^z \geq t\}$ equals

$$\{Y_s^z \geq x \text{ for all } s \leq t\} = \left\{Y_s^z \geq x \text{ for all } s \leq t,\ S_t \leq z - x + \int_0^t f(Y_s^z)\,\mathrm{d}s\right\}.$$

Monotonicity of f implies that $\mathbb{P}(T_x^z \geq t) \leq \mathbb{P}(S_t \leq z - x + tf(x))$. Therefore, since $f(x) \to 0$ as $x \to \infty$,

$$\begin{aligned}\mathbb{P}\left(T_x^z \geq (1+\varepsilon)\frac{z}{\mu}\right) &\leq \mathbb{P}\left(S_{(1+\varepsilon)z/\mu} \leq z - x + (1+\varepsilon)\frac{z}{\mu}f(x)\right)\\ &\leq \mathbb{P}\big(S_{(1+\varepsilon)z/\mu} \leq z\big(1 + \tfrac{1}{2}\varepsilon\big) - x\big)\end{aligned}$$

provided only that x is sufficiently large. Now the right-hand term converges to 0, so it follows that

$$\mathbb{P}\left(T_x^z \geq (1+\varepsilon)\frac{z}{\mu}\right) \to 0.$$

Note that we proved this result only for sufficiently large x, depending on ε. However, this restriction can be omitted, since, for fixed $x_1 < x_2$, the random variables $T_{x_1}^z - T_{x_2}^z$ are bounded in probability uniformly in z. Thus, altogether we have, for any x,

$$\mathbb{P}\left((1-\varepsilon)\frac{z}{\mu} \leq T_x^z < (1+\varepsilon)\frac{z}{\mu}\right) \to 1 \quad \text{as } z \to \infty,$$

which (since $\varepsilon > 0$ is arbitrary) is our assertion. □

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

We turn now to a central limit theorem for passage times of the processes Y^z. Choose κ sufficiently large that $\sup_{y\geq\kappa} f(y) \leq \frac{1}{2}\mu$, and define the function $\beta_z,\ z \geq \kappa$, by

$$\beta_z := \int_\kappa^z \frac{\mathrm{d}y}{\mu - f(y)}. \tag{11}$$

Proposition 3. *Let*

$$\sigma^2 := \int_{(0,\infty)} y^2 \lambda(\mathrm{d}y) < \infty. \tag{12}$$

Then, as $z \to \infty$,

$$\frac{T_x^z - \beta_z}{\sqrt{z}} \xrightarrow{\text{D}} N\bigg(0, \frac{\sigma^2}{\mu^3}\bigg).$$

Proof. (i) Note again that, for $x_1 < x_2$, the random variables $T_{x_1}^z - T_{x_2}^z$ are bounded in probability uniformly in z. Thus, it suffices to prove our theorem for all $x \geq x_0$ for some $x_0 \in \mathbb{R}$. Therefore, we may change $f(x)$ for all $x < x_0$; we do so in such a way that $f(x) \leq \frac{1}{2}\mu$ for all $x \in \mathbb{R}$, without touching the other properties of f. Thus, we assume from now that

$$f(y) \leq \tfrac{1}{2}\mu \quad \text{for all } y \in \mathbb{R}, \tag{13}$$

and set $\kappa = 0$ in (11). Consequently,

$$\frac{z}{\mu} \leq \beta_z \leq \frac{2z}{\mu}, \qquad z > 0. \tag{14}$$

For any $z > 0$, define the function $\rho^z(t) = \rho_t^z,\ 0 \leq t \leq \beta_z$, such that

$$\beta_{\rho^z(t)} = \beta_z - t \quad \text{for } 0 \leq t \leq \beta_z,$$

in particular, $\rho^z(0) = z$ and $\rho^z(\beta_z) = 0$. This means that ρ^z arises by first inverting the function β (restricted to the interval $[0, z]$), and then reversing the time parameter on its domain $[0, \beta_z]$. Differentiation yields $\dot\rho_t^z = f(\rho_t^z) - \mu$, so $\dot\rho_t \leq -\frac{1}{2}\mu$ and

$$\rho_t^z = z - \mu t + \int_0^t f(\rho_s^z)\,\mathrm{d}s.$$

(ii) Inspection of (8) suggests that ρ^z may be a good approximation for the process Y^z; we estimate the difference by observing that

$$Y_t^z - \rho_t^z = -(S_t - \mu t) + \int_0^t (f(Y_s^z) - f(\rho_s^z))\,\mathrm{d}s.$$

For given $t > 0$, define

$$u_t = \begin{cases} \sup\{s < t \colon Y_s^z \leq \rho_s^z\} & \text{on the event } Y_t^z > \rho_t^z, \\ \sup\{s < t \colon Y_s^z \geq \rho_s^z\} & \text{on the event } Y_t^z < \rho_t^z, \end{cases}$$

and $u_t := t$ on the event $Y_t^z = \rho_t^z$. Then $0 \leq u_t \leq t$ since $Y_0^z = z = \rho_0^z$. Because f is a decreasing function, the event $Y_t^z > \rho_t^z$ implies that

$$\begin{aligned} Y_t^z - \rho_t^z &\leq Y_t^z - \rho_t^z - \int_{u_t}^t (f(Y_s^z) - f(\rho_s^z))\,\mathrm{d}s - (Y_{u_t-}^z - \rho_{u_t-}^z) \\ &= -(S_t - \mu t) + (S_{u_t-} - \mu u_t). \end{aligned}$$

On the event $Y_t^z < \rho_t^z$, there is an analogous estimate from below, so taken all together,

$$|Y_t^z - \rho_t^z| \le 2M_t, \quad \text{where} \quad M_t := \sup_{u\le t} |S_u - \mu u|.$$

Consequently, $Y_s^z \ge \rho_s^z - 2M_s \ge \rho_s^z - 2M_t$ for $s \le t$ and from the monotonicity of f,

$$\int_0^t f(Y_s^z)\,\mathrm{d}s - \int_0^t f(\rho_s^z)\,\mathrm{d}s \le \int_0^t f(\rho_s^z - 2M_t)\,\mathrm{d}s - \int_0^t f(\rho_s^z)\,\mathrm{d}s \le 2M_t f(\rho_t^z - 2M_t).$$

An analoguous estimate is valid from below; it yields

$$\left| \int_0^t f(Y_s^z)\,\mathrm{d}s - \int_0^t f(\rho_s^z)\,\mathrm{d}s \right| \le 2M_t f(\rho_t^z - 2M_t). \tag{15}$$

Recall here that, under our assumptions on the subordinator S, Donsker's invariance principle implies that

$$M_t = O_P(\sqrt{t}) \quad \text{as } t \to \infty.$$

(iii) Next we derive some upper estimates of probabilities. Given $a, x \in \mathbb{R}$, for any $c > 0$,

$$\begin{aligned}
&\mathbb{P}(T_x^z \ge \beta_z + a\sqrt{z}) \\
&\quad = \mathbb{P}(Y_t^z \ge x \text{ for all } t \le \beta_z + a\sqrt{z}) \\
&\quad = \mathbb{P}\bigg(S_{\beta_z + a\sqrt{z}} \le z - x + \int_0^{\beta_z + a\sqrt{z}} f(Y_s^z)\,\mathrm{d}s,\ Y_t^z \ge x \text{ for all } t \le \beta_z + a\sqrt{z}\bigg) \\
&\quad \le \mathbb{P}\bigg(S_{\beta_z + a\sqrt{z}} \le z - x + f(x)(c + |a|)\sqrt{z} + \int_0^{\beta_z - c\sqrt{z}} f(Y_s^z)\,\mathrm{d}s\bigg).
\end{aligned}$$

We now use (15). From the definition of ρ^z and writing $\beta(y) = \beta_y$, we have

$$\beta(\rho^z(\beta_z - c\sqrt{z})) = c\sqrt{z},$$

and then because of (14),

$$\rho^z(\beta_z - c\sqrt{z}) \ge \frac{c\sqrt{z}}{2\mu}.$$

So on the event $M_{\beta_z} \le c\sqrt{z}/(8\mu)$,

$$\rho^z(\beta_z - \sqrt{z}) - 2M_{\beta_z - c\sqrt{z}} \ge \frac{c\sqrt{z}}{2\mu} - \frac{c\sqrt{z}}{4\mu} = \frac{c\sqrt{z}}{4\mu}.$$

Consequently, appealing to (15) and since $\beta_z \le 2z/\mu$,

$$\begin{aligned}
&\mathbb{P}(T_x^z \ge \beta_z + a\sqrt{z}) \\
&\quad \le \mathbb{P}\bigg(M_{2z/\mu} > \frac{c\sqrt{z}}{8\mu}\bigg) \\
&\qquad + \mathbb{P}\bigg(S_{\beta_z + a\sqrt{z}} \le z - x + f(x)(c + |a|)\sqrt{z} + \int_0^{\beta_z} f(\rho_s^z)\,\mathrm{d}s + \frac{c\sqrt{z}}{4\mu} f\bigg(\frac{c\sqrt{z}}{4\mu}\bigg)\bigg).
\end{aligned} \tag{16}$$

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

Furthermore, from the definition of ρ^z,

$$z + \int_0^{\beta_z} f(\rho_s^z)\,\mathrm{d}s = \rho^z(\beta_z) + \mu\beta_z = \mu\beta_z.$$

Therefore, if we fix $\varepsilon > 0$, take c so large that the first right-hand probability in (16) is smaller than ε, and then choose z so large that $(c/4\mu) f(c\sqrt{z}/4\mu) \le \varepsilon$ and also choose $x > 0$ and so large that $cf(x)(c + |a|) \le \varepsilon$, then

$$\mathbb{P}(T_x^z \ge \beta_z + a\sqrt{z}) \le \varepsilon + \mathbb{P}(S_{\beta_z + a\sqrt{z}} \le \mu\beta_z + 2\varepsilon\sqrt{z}).$$

Also, by the law of large numbers

$$S_{\beta_z + a\sqrt{z}} - S_{\beta_z} \sim \mu a\sqrt{z} \quad \text{in probability.}$$

Therefore,

$$\mathbb{P}(T_x^z \ge \beta_z + a\sqrt{z}) \le 2\varepsilon + \mathbb{P}(S_{\beta_z} \le \mu\beta_z + (-\mu a + 3\varepsilon)\sqrt{z}).$$

Also, $\mu\beta_z \sim z$, so, for large z,

$$\mathbb{P}(T_x^z \ge \beta_z + a\sqrt{z}) \le 2\varepsilon + \mathbb{P}(S_{\beta_z} \le \mu\beta_z + (-\mu a + 4\varepsilon)\mu^{1/2}\sqrt{\beta_z}).$$

It now follows from (12) and the central limit theorem that

$$\frac{S_t - \mu t}{\sqrt{\sigma^2 t}} \xrightarrow{\mathrm{D}} L,$$

where L denotes a standard normal random variable. Thus,

$$\limsup_{z\to\infty} \mathbb{P}(T_x^z \ge \beta_z + a\sqrt{z}) \le 2\varepsilon + \mathbb{P}(L \le (-\mu a + 4\varepsilon)\mu^{1/2}\sigma^{-1}).$$

Note in our proof that the choice of x depends on ε, but, since the differences $T_{x_1}^z - T_{x_2}^z$ are bounded in probability uniformly in z, this estimate generalizes to all x. Now letting $\varepsilon \to 0$ gives

$$\limsup_{z\to\infty} \mathbb{P}\biggl(\frac{T_x^z - \beta_z}{\sqrt{z}} \ge a\biggr) \le \mathbb{P}\biggl(L \le -\frac{\mu^{3/2}a}{\sigma}\biggr).$$

This is the first part of our claim.

(iv) For the lower estimates, we first introduce the random variable

$$R_{z,x} := \sup\{t \ge 0 \colon Y_t^z \ge x\} - \inf\{t \ge 0 \colon Y_t^z < x\};$$

this is the length of the time interval on which $Y_t^z - x$ changes from a positive sign to ultimately a negative sign (note that the paths of Y^z are *not* monotone). We claim that these random variables are bounded in probability, uniformly in z and x. Indeed, with

$$\eta_{z,x} := \inf\{t \ge 0 \colon Y_t^z < x\},$$

we have, for $t > \eta = \eta_{z,x}$, because of $Y_\eta^z \le x$ and (13),

$$Y_t^z = Y_\eta^z - (S_t - S_\eta) + \int_\eta^t f(Y_s^z)\,\mathrm{d}s \le x - (S_t - S_\eta) + \frac{1}{2}\mu(t - \eta).$$

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

Thus, $R_{z,x}$ is bounded from above by

$$R'_{z,x} := \sup\big\{u \geq 0\colon (S_{\eta_{z,x}+u} - S_{\eta_{z,x}}) - \tfrac{1}{2}\mu u \leq 0\big\}.$$

These random variables are a.s. finite. Moreover, they are identically distributed, since the $\eta_{z,x}$ are stopping times. This proves that the $R_{z,x}$ are uniformly bounded in probability.

For the lower bounds and $a, b \in \mathbb{R}$, we have

$$\begin{aligned}\mathbb{P}(T_x^z \geq \beta_z + a\sqrt{z}) &\geq \mathbb{P}(Y_t^z \geq x \text{ for all } t \leq \beta_z + a\sqrt{z},\ R_{z,x} \leq b)\\ &= \mathbb{P}(Y_t^z \geq x \text{ for all } \beta_z + a\sqrt{z} - b \leq t \leq \beta_z + a\sqrt{z},\ R_{z,x} \leq b).\end{aligned}$$

For these t,

$$Y_t^z = z - S_t + \int_0^t f(Y_s^z)\,\mathrm{d}s \geq z - S_{\beta_z + a\sqrt{z}} + \int_0^{\beta_z + a\sqrt{z} - b} f(Y_s^z)\,\mathrm{d}s;$$

therefore,

$$\begin{aligned}\mathbb{P}(T_x^z \geq \beta_z + a\sqrt{z}) &\geq \mathbb{P}\bigg(S_{\beta_z + a\sqrt{z}} \leq z - x + \int_0^{\beta_z + a\sqrt{z} - b} f(Y_s^z)\,\mathrm{d}s,\ R_{z,x} \leq b\bigg)\\ &\geq \mathbb{P}\bigg(S_{\beta_z + a\sqrt{z}} \leq z - x + \int_0^{\beta_z - c\sqrt{z}} f(Y_s^z)\,\mathrm{d}s\bigg) - \mathbb{P}(R_{z,x} > b)\end{aligned}$$

for sufficiently large c.

We now use (15) as in (iii). Proceeding analogously, instead of estimate (16) we obtain

$$\begin{aligned}\mathbb{P}(T_x^z \geq \beta_z + a\sqrt{z}) \geq{}& -\mathbb{P}(R_{z,x} > b) - \mathbb{P}\bigg(M_{2z/\mu} > \frac{c\sqrt{z}}{8\mu}\bigg)\\ &+ \mathbb{P}\bigg(S_{\beta_z + a\sqrt{z}} \leq z - x + \int_0^{\beta_z - c\sqrt{z}} f(\rho_s^z)\,\mathrm{d}s - \frac{c\sqrt{z}}{4\mu} f\bigg(\frac{c\sqrt{z}}{4\mu}\bigg)\bigg).\end{aligned}$$

Also, since $\rho_{\beta_z}^z = 0$ and $\dot{\rho}_t^z \leq -\frac{1}{2}\mu$,

$$\int_{\beta_z - c\sqrt{z}}^{\beta_z} f(\rho_s^z)\,\mathrm{d}s \leq \int_0^{c\sqrt{z}} f\bigg(\frac{\mu s}{2}\bigg)\,\mathrm{d}s = o(\sqrt{z}).$$

Hence, for given $\varepsilon > 0$ and sufficiently large z,

$$\begin{aligned}\mathbb{P}(T_x^z \geq \beta_z + a\sqrt{z}) \geq{}& -\mathbb{P}(R_{z,x} > b) - \mathbb{P}\bigg(M_{2z/\mu} > \frac{c\sqrt{z}}{8\mu}\bigg)\\ &+ \mathbb{P}\bigg(S_{\beta_z + a\sqrt{z}} \leq z - \varepsilon\sqrt{z} + \int_0^{\beta_z} f(\rho_s^z)\,\mathrm{d}s - \frac{c\sqrt{z}}{4\mu} f\bigg(\frac{c\sqrt{z}}{4\mu}\bigg)\bigg).\end{aligned}$$

Returning to the arguments of part (iii) we choose b, c and then z so large that

$$\mathbb{P}(T_x^z \geq \beta_z + a\sqrt{z}) \geq -2\varepsilon + \mathbb{P}(S_{\beta_z + a\sqrt{z}} \leq \mu\beta_z - 2\varepsilon\sqrt{z})$$

and

$$\liminf_{z\to\infty} \mathbb{P}(T_x^z \geq \beta_z + a\sqrt{z}) \geq -3\varepsilon + \mathbb{P}\bigg(L \leq -\frac{(\mu a + 3\varepsilon)\mu^{1/2}}{\sigma}\bigg).$$

The limit $\varepsilon \to 0$ leads to the desired lower estimate. □

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

3. Approximating the block-counting process

In this section we derive a strengthening of a result of Kersting *et al.* [4] on the approximation to the logarithm of the block-counting processes in the dust case. To this end, let us quickly recall the Poisson point process construction of the Λ-coalescent given in [4]; this is a slight variation of the construction provided by Pitman [6].

This construction requires $\Lambda(\{0\}) = 0$, a condition that is satisfied by coalescents with dust. Consider a Poisson point process Ψ on $(0, \infty) \times (0, 1] \times [0, 1]^n$ with intensity

$$\mathrm{d}t \times p^{-2}\Lambda(\mathrm{d}p) \times \mathrm{d}u_1 \times \cdots \times \mathrm{d}u_n,$$

and let $\Pi_n(0) = \{\{1\}, \ldots, \{n\}\}$ be the partition of the set $\{1, \ldots, n\}$ into singletons. Suppose that $(t, p, u_1, \ldots, u_n)$ is a point of Ψ, and that $\Pi_n(t-)$ consists of the blocks $B_1, \ldots, B_b$, ranked in order by their smallest element. Then $\Pi_n(t)$ is obtained from $\Pi_n(t-)$ by merging together all of the blocks B_i for which $u_i \le p$ into a single block. These are the only times that mergers occur. This construction is well defined because, a.s. for any fixed $t' < \infty$, there are only finitely many points $(t, p, u_1, \ldots, u_n)$ of Ψ for which $t \le t'$ and at least two of $u_1, \ldots, u_n$ are less than or equal to p. The resulting process $\Pi_n = (\Pi_n(t),\ t \ge 0)$ is the Λ-coalescent. When $(t, p, u_1, \ldots, u_n)$ is a point of Ψ, we say that a p-merger occurs at time t.

Condition (7) allows us to approximate the number of blocks in the Λ-coalescent by a subordinator. Let $\phi\colon (0, \infty) \times (0, 1] \times [0, 1]^n \to (0, \infty) \times (0, \infty]$ be the function defined by

$$\phi(t, p, u_1, \ldots, u_n) = (t, -\log(1 - p)).$$

Now $\phi(\Psi)$ is a Poisson point process, and we can define a pure-jump subordinator $(S(t),\ t \ge 0)$ with the property that $S(0) = 0$ and if (t, x) is a point of $\phi(\Psi)$ then $S(t) = S(t-) + x$. With λ the Lévy measure of S, (9) and (12) now read

$$\mu = \int_{[0,1]} |\log(1 - p)| \frac{\Lambda(\mathrm{d}p)}{p^2} \quad \text{and} \quad \sigma^2 = \int_{[0,1]} (\log(1 - p))^2 \frac{\Lambda(\mathrm{d}p)}{p^2}.$$

This subordinator first appeared in [6] and was used to approximate the block-counting process by Gnedin *et al.* [1] and Möhle [5]; the benefits of a refined approximation by a *drifted* subordinator were discovered in [4]. Recall that the drift appears because a merging of b out of $N_n(t)$ lines results in a decrease of $b - 1$ and not of b lines; see [4, Equation (23)] for an explanation of the form of the drift. Our next result provides a refinement of [4, Theorem 10].

Proposition 4. *Suppose that $\int_{[0,1]} p^{-1}\Lambda(\mathrm{d}p) < \infty$. Let f be as in (2), and let Y_n be the solution of (8) with $z := \log n$. Then, for any $\varepsilon > 0$, there exists $\ell < \infty$ such that*

$$\mathbb{P}\Big(\sup_{t<\tau_n} |\log N_n(t) - Y_n(t)| \le \ell,\ Y_n(\tau_n) < \ell\Big) \ge 1 - \varepsilon.$$

Proof. From [4] we know that, for given $\varepsilon > 0$, there exists $r < \infty$ such that

$$\mathbb{P}\Big(\sup_{t<\tau_n} |\log N_n(t) - Y_n(t)| \le r\Big) \ge 1 - \frac{1}{2}\varepsilon.$$

Consider the size Δ_n of the last jump. Letting (u_i, p_i), $i \ge 1$, be the points of the underlying Poisson point process with intensity measure $\mathrm{d}t\,\Lambda(\mathrm{d}p)/p^2$, the associated subordinator S has

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

jumps of size $v_i = -\log(1-p_i)$ at times t_i. So, for any $c > 0$, the event $\{\Delta_n \le \log N_n(\tau_n-)-c\}$ is the same as

$$\begin{aligned}&\{\tau_n = t_i \text{ and } -\log(1-p_i) \le \log N_n(t_i-) - c \text{ for some } i \ge 1\}\\ &\quad = \bigg\{\tau_n = t_i \text{ and } p_i \le 1 - \frac{\mathrm{e}^c}{N_n(t_i-)} \text{ for some } i \ge 1\bigg\}.\end{aligned}$$

Given $N_n(t-)$, this event appears at time t at rate

$$\nu_{n,t} = \int_{[0,1-\mathrm{e}^c/N_n(t-)]} p^{N_n(t-)} \frac{\Lambda(\mathrm{d}p)}{p^2}.$$

Using the inequalities $p^b = (1-(1-p))^b \le \mathrm{e}^{-(1-p)b} \le 1/((1-p)b)$, we obtain

$$\nu_{n,t} \le \int_{[0,1-\mathrm{e}^c/N_n(t-)]} \mathrm{e}^{-(1-p)(N_n(t-)-2)}\Lambda(\mathrm{d}p) \le \int_{[0,1-\mathrm{e}^c/N_n(t-)]} \frac{\mathrm{e}^2}{(1-p)N_n(t-)}\Lambda(\mathrm{d}p).$$

It follows that

$$\mathbb{E}\bigg[\int_0^\infty \nu_{n,t}\,\mathrm{d}t\bigg] \le \mathbb{E}\bigg[\int_{[0,1]}\int_0^\infty \frac{\mathrm{e}^2}{(1-p)N_n(t-)} \mathbf{1}_{\{N_n(t-)\ge\lceil \mathrm{e}^c/(1-p)\rceil\}}\,\mathrm{d}t\,\Lambda(\mathrm{d}p)\bigg].$$

Lemma 14 of [4] yields the estimate

$$\mathbb{E}\bigg[\int_0^\infty \frac{1}{N_n(t-)} \mathbf{1}_{\{N_n(t-)\ge\lceil \mathrm{e}^c/(1-p)\rceil\}}\,\mathrm{d}t\bigg] \le c_1\bigg\lceil \frac{\mathrm{e}^c}{1-p}\bigg\rceil^{-1} \le c_1\frac{1-p}{\mathrm{e}^c}$$

for some $c_1 > 0$; hence,

$$\mathbb{E}\bigg[\int_0^\infty \nu_{n,t}\,\mathrm{d}t\bigg] \le c_1\mathrm{e}^{2-c}\Lambda([0,1]).$$

Therefore, for sufficiently large c,

$$\mathbb{E}\bigg[\int_0^\infty \nu_{n,t}\,\mathrm{d}t\bigg] \le \frac{1}{2}\varepsilon,$$

implying that

$$\mathbb{P}(\Delta_n \le \log N_n(\tau_n-) - c) = 1 - \exp\bigg(-\mathbb{E}\bigg[\int_0^\infty \nu_{n,t}\,\mathrm{d}t\bigg]\bigg) \le \frac{1}{2}\varepsilon.$$

Altogether we obtain

$$\mathbb{P}\Big(\sup_{t<\tau_n}|\log N_n(t) - Y_n(t)| \le r,\ \Delta_n > \log N_n(\tau_n-) - c\Big) \ge 1-\varepsilon.$$

The event in the last relation implies that

$$Y_n(\tau_n) = Y_n(\tau_n-) - \Delta_n < \log N_n(\tau_n-) + r - (\log N_n(\tau_n-) - c) = r + c,$$

and the claim of the proposition follows with $\ell = r + c$. □

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

4. Proof of the main results

Proof of Theorem 1. Assume first that $\mu < \infty$. Then we have a coalescent with dust, and we can apply Proposition 4. Fix $\eta > 0$. Note that, on the event that $Y_n(\tau_n) < \ell$, the event $\tau_n < (1-\eta)\log n/\mu$ implies the inequality $T_\ell^{\log n} < (1-\eta)\log n/\mu$, where T_x^z is defined following (8). Thus, in view of Proposition 4, for any $\varepsilon > 0$, there exists ℓ such that

$$\mathbb{P}\biggl(\tau_n < \frac{(1-\eta)\log n}{\mu}\biggr) \le \mathbb{P}\biggl(T_\ell^{\log n} < \frac{(1-\eta)\log n}{\mu}\biggr) + \varepsilon.$$

Proposition 2 implies that the right-hand probability converges to 0 as $n \to \infty$. Letting $\varepsilon \to 0$ we obtain

$$\lim_{n\to\infty} \mathbb{P}\biggl(\tau_n < \frac{(1-\eta)\log n}{\mu}\biggr) = 0.$$

Also, on the event $\sup_{t<\tau_n} |\log N_n(t) - Y_n(t)| \le \ell$, the event $\tau_n > (1+\eta)\log n/\mu$ implies that $Y_n(t) \ge -\ell$ for all $t \le (1+\eta)\log n/\mu$, and, consequently,

$$\mathbb{P}\biggl(\tau_n > \frac{(1+\eta)\log n}{\mu}\biggr) \le \mathbb{P}\biggl(T_{-\ell}^{\log n} > \frac{(1+\eta)\log n}{\mu}\biggr) + \varepsilon.$$

Again, from Proposition 2, the right-hand probability converges to 0, and we obtain

$$\lim_{n\to\infty} \mathbb{P}\biggl(\tau_n > \frac{(1+\eta)\log n}{\mu}\biggr) = 0.$$

Thus, our claim follows in the $\mu < \infty$ case.

Now assume that $\mu = \infty$. If $\Lambda(\{0\}) > 0$ then the coalescent comes down from ∞ and τ_n remains bounded in probability. The same is true if $\Lambda(\{1\}) > 0$; thus, we may assume that $\Lambda(\{0,1\}) = 0$.

For given $\varepsilon > 0$, define the measure Λ^ε by $\Lambda^\varepsilon(B) := \Lambda(B \cap [\varepsilon, 1-\varepsilon])$. Obviously,

$$\mu^\varepsilon := \int_0^1 |\log(1-p)| \frac{\Lambda^\varepsilon(\mathrm{d}p)}{p^2} < \infty.$$

Thus, for the absorption times τ_n^ε of the Λ^ε-coalescent, we have

$$\frac{\tau_n^\varepsilon}{\log n} \to \frac{1}{\mu^\varepsilon}$$

in probability as $n \to \infty$. Now we may couple the Λ^ε-coalescent in an obvious manner to the Λ-coalescent in such a way that $N_n(t) \le N_n^\varepsilon(t)$ a.s. for all $t \ge 0$, in particular $\tau_n \le \tau_n^\varepsilon$. Hence, it follows that

$$\mathbb{P}\biggl(\frac{\tau_n}{\log n} > \frac{2}{\mu^\varepsilon}\biggr) \to 0.$$

Because $\Lambda(\{0,1\}) = 0$, $\mu^\varepsilon \to \mu = \infty$ as $\varepsilon \to 0$, and, consequently,

$$\mathbb{P}\biggl(\frac{\tau_n}{\log n} > \eta\biggr) \to 0$$

for all $\eta > 0$. This is our claim. □

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

Proof of Theorem 2. Because of the condition $\mu < \infty$ we may again apply Proposition 4; we follow the same line as in the previous proof. For $\varepsilon > 0$, there exists ℓ such that, for all $a \in \mathbb{R}$,

$$\mathbb{P}(\tau_n < b_n + a\sqrt{n}) \le \mathbb{P}(T_\ell^{\log n} < b_n + a\sqrt{n}) + \varepsilon$$

and

$$\mathbb{P}(\tau_n > b_n + a\sqrt{n}) \le \mathbb{P}(T_{-\ell}^{\log n} > b_n + a\sqrt{n}) + \varepsilon.$$

Now apply Proposition 3 and let $\varepsilon \to 0$. □

Proof of Proposition 1. (i) Start by assuming that (6) holds. Because $1 - (1-p)^{1/r} \le \min(p/r, 1)$ for $0 < r < 1$, we have, for $\alpha > 0$,

$$f\Big(\log \frac{1}{r}\Big) \le \int_0^{r^\alpha} \frac{\Lambda(\mathrm{d}p)}{p} + r\int_{r^\alpha}^1 \frac{\Lambda(\mathrm{d}p)}{p^2} \le \int_0^{r^\alpha} \frac{\Lambda(\mathrm{d}p)}{p} + r^{1-\alpha}\int_0^1 \frac{\Lambda(\mathrm{d}p)}{p}. \tag{17}$$

Also, since $1 - (1-p)^{1/r} \ge 1 - \mathrm{e}^{-p/r} \ge \mathrm{e}^{-p/r}p/r$, it follows that, for $\beta > 0$,

$$f\Big(\log \frac{1}{r}\Big) \ge \mathrm{e}^{-r^{\beta-1}} \int_0^{r^\beta} \frac{\Lambda(\mathrm{d}p)}{p}. \tag{18}$$

Together with (6), these two estimates imply that, for $\alpha < 1 < \beta$,

$$c\beta^{-1/2} \le \liminf_{r\to 0} f\Big(\log\frac{1}{r}\Big)\sqrt{\log\frac{1}{r}} \le \limsup_{r\to 0} f\Big(\log\frac{1}{r}\Big)\sqrt{\log\frac{1}{r}} \le c\alpha^{-1/2}.$$

Letting $\alpha, \beta \to 1$ we arrive at $f(y) = (c + o(1))/\sqrt{y}$ as $y \to \infty$ and, consequently,

$$\int_0^{\log n} f(y)\,\mathrm{d}y = (c + o(1))2\sqrt{\log n} \quad \text{as } n \to \infty.$$

Now, since

$$\frac{1}{\mu - f(y)} = \frac{1}{\mu} + \frac{f(y)}{\mu(\mu - f(y))}$$

and $f(y) = o(1)$ as $y \to \infty$,

$$\int_\kappa^z \frac{\mathrm{d}y}{\mu - f(y)} = \frac{z}{\mu} + \frac{1 + o(1)}{\mu^2}\int_0^z f(y)\,\mathrm{d}y + O(1) \quad \text{as } z \to \infty; \tag{19}$$

so, as claimed,

$$b_n = \frac{\log n}{\mu} + \frac{2c + o(1)}{\mu^2}\sqrt{\log n}.$$

(ii) Suppose now that (5) is satisfied. Then, from (19) with $z = \log n$, it follows that

$$\int_0^{\log n} f(y)\,\mathrm{d}y = (2c + o(1))\sqrt{\log n} \quad \text{as } n \to \infty,$$

or, equivalently,

$$\int_0^z f(y)\,\mathrm{d}y = (2c + o(1))\sqrt{z} \quad \text{as } z \to \infty.$$

This implies that $f(z) = (c + o(1))/\sqrt{z}$ as $z \to \infty$. For $c = 0$, this claim follows because f is decreasing; hence,

$$zf(z) \le \int_0^z f(y)\,\mathrm{d}y = o(\sqrt{z}).$$

https://doi.org/10.1017/apr.2018.78 Published online by Cambridge University Press

For $c > 0$, we use the estimate

$$\frac{1}{\eta\sqrt{z}}\int_z^{(1+\eta)z} f(y)\,\mathrm{d}y \le \sqrt{z}f(z) \le \frac{1}{\eta\sqrt{z}}\int_{(1-\eta)z}^{z} f(y)\,\mathrm{d}y$$

with $\eta > 0$. Taking the limit $z \to \infty$ and then $\eta \to 0$ yields $f(z) = (c + o(1))/\sqrt{z}$. Now, similarly to part (i), from (17) and (18), we obtain

$$c\sqrt{\alpha} \le \liminf_{r\to 0} \sqrt{\log\frac{1}{r}}\int_{[0,r]}\frac{\Lambda(\mathrm{d}p)}{p} \le \limsup_{r\to 0}\sqrt{\log\frac{1}{r}}\int_{[0,r]}\frac{\Lambda(\mathrm{d}p)}{p} \le c\sqrt{\beta}.$$

Letting $\alpha, \beta \to 1$ we arrive at (6). □

Acknowledgement

It is our pleasure to dedicate this work to Peter Jagers.

References

[1] Gnedin, A., Iksanov, A., and Marynych, A. (2011). On Λ-coalescents with dust component. *J. Appl. Prob.* **48,** 1133–1151.

[2] Goldschmidt, C. and Martin, J. B. (2005). Random recursive trees and the Bolthausen–Sznitman coalesent. *Electron. J. Prob.* **10,** 718–745.

[3] Herriger, P. and Möhle, M. (2012). Conditions for excheangable coalescents to come down from infinity. *ALEA* **9,** 637–665.

[4] Kersting, G., Schweinsberg, J., and Wakolbinger, A. (2018). The size of the last merger and time reversal in Λ-coalescents. *Ann. Inst. H. Poincaré Prob. Statist.* **54,** 1527–1555.

[5] Möhle, M. (2014). On hitting probabilities of beta coalescents and absorption times of coalescents that come down from infinity. *ALEA* **11,** 141–159.

[6] Pitman, J. (1999). Coalescents with multiple collisions. *Ann. Prob.* **27,** 1870–1902.

[7] Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Prob.* **36,** 1116–1125.

[8] Schweinsberg, J. (2000). A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Commun. Prob.* **5,** 1–11.

GÖTZ KERSTING, *Goethe University Frankfurt*

Institute of Mathematics, Goethe University Frankfurt, Robert Mayer Strasse 10, 60325 Frankfurt am Main, Germany.
Email address: kersting@math.uni-frankfurt.de

ANTON WAKOLBINGER, *Goethe University Frankfurt*

Institute of Mathematics, Goethe University Frankfurt, Robert Mayer Strasse 10, 60325 Frankfurt am Main, Germany.
Email address: wakolbinger@math.uni-frankfurt.de