On the time to absorption in $\Lambda$-coalescents

Götz Kersting* and Anton Wakolbinger*

Abstract

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

AMS 2010 subject classification: 60J75 (primary), 60J27, 60F05 (secondary)

Keywords: coalescents, time to absorption, law of large numbers, central limit theorem, subordinator with drift

1 Introduction and main results

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

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

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

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

is the time of the last merger, also called the absorption time of the coalescent started in $n$ blocks. We will investigate the asymptotic distribution of $\tau_n$ as $n \to \infty$.

Our first result is a law of large numbers for the times $\tau_n$. Let

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

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

*Institut für Mathematik, Goethe Universität, Frankfurt am Main, Germany
kersting@math.uni-frankfurt.de, wakolbinger@math.uni-frankfurt.de

Work partially supported by the DFG Priority Programme SPP 1590 “Probabilistic Structures in Evolution”
Theorem 1. For any \( \Lambda \)-coalescent, \[
{\tau_n \over \log n} \to {1 \over \mu} .
\] (1)
in probability as \( n \to \infty \).

This theorem says that in a \( \Lambda \)-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 case \( \mu < \infty \) is equivalently captured by the simultaneous validity of the conditions
\[
\int_{[0,1]} \frac{\Lambda(dp)}{p} < \infty \quad \text{and} \quad \int_{[0,1]} \log \frac{1}{1-p} \Lambda(dp) < \infty .
\]
The first one is a requirement on \( \Lambda \) in the vicinity of 0, it prohibits a swarm of small mergers (as they occur in coalescents coming down from infinity, meaning that the \( \tau_n \) are bounded in probability uniformly in \( n \)). The second is a condition on \( \Lambda \) in the vicinity of 1. It rules out the possibility of mergers which, although appearing only every now and then, are so vast that they make the coalescent collapse. – A counterpart to Theorem 1 with \( \tau_n \) in (1) replaced by its expectation, was obtained by Herriger and Möhle \[\text{II}\].

Our second result is a central limit theorem. Here we confine ourselves to coalescents with \( \mu < \infty \). Then the function
\[
f(y) := \int_{[0,1]} \frac{1 - (1-p)e^y}{e^y} \frac{\Lambda(dp)}{p^2} , \quad y \in \mathbb{R}
\] (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 \[
b_n := \int_{\kappa}^{\log n} \frac{dy}{\mu - f(y)} ,
\]
where we choose \( \kappa \geq 0 \) such that \( f(y) \leq \frac{\mu}{2} \) for all \( y \geq \kappa \).

Theorem 2. Assume that \( \mu < \infty \) and moreover
\[
\sigma^2 := \int_{[0,1]} \left( \log \frac{1}{1-p} \right)^2 \frac{\Lambda(dp)}{p^2} < \infty .
\]
Then
\[
{\tau_n - b_n \over \sqrt{\log n}} \overset{d}{\to} N(0, \sigma^2/\mu^3)
\] (3)
as \( n \to \infty \).
Under the additional condition
\[
\int_{[0,1]} \log \frac{1}{p} \frac{\Lambda(dp)}{p} < \infty .
\] (4)
the CLT has been obtained by Gnedin, Iksanov and Marynych, with \( b_n \) replaced by \( \log n/\mu \). (Their condition (9) is equivalent to the above condition (4), see Remark 13 in [4]). 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 any case.

**Proposition 3.** Let \( 0 \leq c < \infty \). Then
\[
b_n = \frac{\log n}{\mu} + \frac{2c}{\mu^2} \sqrt{\log n} + o(\sqrt{\log n}) (5)
\]
as \( n \to \infty \), if and only if
\[
\sqrt{\log n} \int_{[0,r]} \frac{\Lambda(dp)}{p} \to c (6)
\]
as \( r \to 0 \).

**Example.** We consider for \( \gamma \in \mathbb{R} \) the finite measures
\[
\Lambda(dp) = (1 + \log \frac{1}{p})^{-\gamma} dp , \ 0 \leq p \leq 1 .
\]
For \( \gamma = 0 \) this gives the Bolthausen-Sznitman coalescent. For \( \gamma > 1 \) it leads to coalescents with \( \mu, \sigma^2 < \infty \). Note that (4) is satisfied iff \( \gamma > 2 \), and (5) is fulfilled iff \( \gamma > 3/2 \). Thus within the range \( 1 < \gamma \leq 3/2 \) one has 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 \( \gamma \leq 1 \), Theorem 1 just tells us that \( \tau_n = o_P(\log n) \). For \( \gamma = 0 \), the Bolthausen-Sznitman coalescent, it is known that \( \tau_n \) is already down to the order \( \log \log n \). For \( \gamma < 0 \), applying Schweinsberg’s criterion, it can be shown that the coalescents come down from infinity. There remains the gap \( 0 < \gamma \leq 1 \). It is tempting to conjecture that \( \tau_n \) is of order \( (\log n)^{\gamma} \) for \( 0 < \gamma < 1 \).

If equation (6) is violated then the subsequent approximation to \( b_n \) 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
\[
b_n = \frac{\log n}{\mu} + \frac{1}{\mu^2} \int_0^{\log n} f(y) dy + \cdots + \frac{1}{\mu^{k+1}} \int_0^{\log n} f^k(y) dy + O\left( \int_0^{\log n} f^{k+1}(y) dy \right) .
\]
Let us now explain the method of proving Theorems 1 and 2. We are mainly dealing with $\Lambda$-coalescents having a dust component. Shortly speaking these are the coalescents for which the rate, at which a single lineage merges with some others from the sample, stays bounded as the sample size tends to infinity. As is well-known this property is characterized by the condition
\[ \int_{[0,1]} \frac{\Lambda(dp)}{p} < \infty . \] (7)
An established tool for the analysis of a $\Lambda$-coalescent with dust is the subordinator $S = (S_t)_{t \geq 0}$, which is used to approximate the logarithm of its block-counting process $N_n = (N_n(t))_{t \geq 0}$ (see e.g. Pitman [6], Möhle [5], and the above mentioned paper by Gnedin et al [2]). We will recall this subordinator in Sec. 3. Indeed, analogues of Theorems 1 and 2 are well-known for first-passage times of subordinators with finite first resp. second moment. However, 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 becomes significant when many small jumps accumulate over time, as it happens close to the dustless case (and as it becomes visible in Proposition 3 and in the above example). Then the appropriate approximation is provided by a drifted subordinator $Y_n = (Y_n(t))_{t \geq 0}$, given by the SDE
\[ Y_n(t) = \log n - S_t + \int_0^t f(Y_n(s)) \, ds , \quad t \geq 0 , \]
with initial value $Y_n(0) = \log n$. The drift compensates the just mentioned difference between $b$ and $b - 1$. In Kersting et al [4] it is shown that
\[ \sup_{t < \tau_n} |Y_n(t) - \log N_n(t)| = O_P(1) \]
as $n \to \infty$, that is, these random variables are bounded in probability. In Sec. 3 we suitably strengthen this result. In Sec. 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 Sec. 4.

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

2 Limit theorems for a drifted subordinator

Let $S = (S_t)_{t \geq 0}$ be a pure jump subordinator with Lévy measure $\lambda$ on $(0, \infty)$. Recall that this requires
\[ \int_0^\infty (y \wedge 1) \lambda(dy) < \infty . \]
With regard to the mentioned properties of the function in [2], let $f : \mathbb{R} \to \mathbb{R}$ be an arbitrary positive, non-increasing, continuous function with
\[ \lim_{y \to \infty} f(y) = 0 . \]
Let the process \( Y^z_t = (Y^z_t)_{t \geq 0} \) denote the unique solution of the SDE
\[
Y^z_t = z - S_t + \int_0^t f(Y^z_s) \, ds
\] (8)
with initial value \( z > 0 \). We will investigate the asymptotic behaviour of its passage times across \( x \in \mathbb{R} \),
\[
T^z_x := \inf\{t \geq 0 : Y^z_t < x\}
\]
in the limit \( z \to \infty \).

The first result provides a law of large numbers. Denote
\[
\mu := \int_{(0,\infty)} y \lambda(dy)
\] (9)

Proposition 4. Assume that \( \mu < \infty \). Then for any \( x \in \mathbb{R} \)
\[
\frac{1}{z} T^z_x \to \frac{1}{\mu}
\]
in probability as \( z \to \infty \).

Proof. Let \( z > x \). Then
\[
\{T^z_x \geq t\} = \{Y^z_s \geq x \text{ for all } s \leq t\} = \left\{ S_s \leq z - x + \int_0^s f(Y^z_u) \, du \text{ for all } s \leq t \right\}.
\]
By positivity of the function \( f \) it follows \( P(T^z_x \geq t) \geq P(S_t \leq z - x) \), thus for any \( \epsilon > 0 \)
\[
P\left(T^z_x \geq (1 - \epsilon) \frac{z}{\mu}\right) \geq P\left(S_{(1-\epsilon)z/\mu} \leq z - x\right).
\] (10)

Now \( \mu = E[S_1] \), thus by the law of large numbers
\[
\frac{S_t}{t} \to \mu
\]
a.s., hence the right-hand term in (10) converges to 1 for \( z \to \infty \) and also
\[
P\left(T^z_x \geq (1 - \epsilon) \frac{z}{\mu}\right) \to 1.
\]

On the other hand,
\[
\{T^z_x \geq t\} = \{Y^z_s \geq x \text{ for all } s \leq t\} = \left\{ Y^z_s \geq x \text{ for all } s \leq t, S_t \leq z - x + \int_0^t f(Y^z_s) \, ds \right\}.
\]
Monotonicity of \( f \) implies \( P(T^z_x \geq t) \leq P\left(S_t \leq z - x + tf(x)\right) \). Therefore, since \( f(x) \to 0 \) as \( x \to \infty \),
\[
P\left(T^z_x \geq (1 + \epsilon) \frac{z}{\mu}\right) \leq P\left(S_{(1+\epsilon)z/\mu} \leq z - x + (1 + \epsilon) \frac{z}{\mu} f(x)\right)
\leq P\left(S_{(1+\epsilon)z/\mu} \leq z(1 + \epsilon/2) - x\right)
\]

5
if only \( x \) is sufficiently large. Now the right-hand term converges to 0, thus it follows that

\[
P\left( T^z_x \geq (1 + \varepsilon) \frac{z}{\mu} \right) \to 0 .
\]

Note that we proved this result only for \( x \) sufficiently large, depending on \( \varepsilon \). However, this restriction may be skipped, since for fixed \( x_1 < x_2 \) the random variables \( T^z_{x_1} - T^z_{x_2} \) are bounded in probability uniformly in \( z \). Thus altogether we have for any \( x \)

\[
P\left( (1 - \varepsilon) \frac{z}{\mu} \leq T^z_x < (1 + \varepsilon) \frac{z}{\mu} \right) \to 1
\]
as \( z \to \infty \), which (since \( \varepsilon > 0 \) was arbitrary) is our assertion. \( \square \)

Now we turn to a central limit theorem for passage times of the processes \( Y^z \). Let the function \( \beta_z, z \geq \kappa \), be given by

\[
\beta_z := \int_{\kappa}^{z} \frac{dy}{\mu - f(y)} ,
\]

where we choose \( \kappa \geq 0 \) so large that

\[
\sup_{y \geq \kappa} f(y) \leq \frac{\mu}{2}.
\]

**Proposition 5.** Suppose that

\[
\sigma^2 := \int_{(0, \infty)} y^2 \lambda(dy) < \infty .
\]

Then

\[
\frac{T^z_x - \beta_z}{\sqrt{\sigma^2 / \mu}} \overset{d}{\to} N(0, \sigma^2 / \mu^3)
\]
as \( z \to \infty \).

**Proof.** (i) Note again that for \( x_1 < x_2 \) the random variables \( T^z_{x_1} - T^z_{x_2} \) 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 it in such a way that \( f(x) \leq \mu/2 \) for all \( x \in \mathbb{R} \), without touching the other properties of \( f \). Thus we assume from now that

\[
f(y) \leq \frac{\mu}{2} \quad \text{for all } y \in \mathbb{R}
\]

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

\[
\frac{z}{\mu} \leq \beta_z \leq \frac{2z}{\mu} , \quad z > 0 .
\]

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

\[
\beta_{\rho^z(t)} = \beta_z - t , \quad 0 \leq t \leq \beta_z,
\]
in particular \( \rho^z(0) = z \) and \( \rho^z(\beta_z) = 0 \). This means that \( \rho^z \) arises by first inverting the function \( \beta \) (restricted to the interval \([0, z]\)) , and then reversing the time parameter on its domain \([0, \beta_z]\) . By differentiation we obtain

\[
\dot{\rho}^z_t = f(\rho^z_t) - \mu ,
\]
consequently \( \dot{\rho}_t \leq -\frac{\mu}{2} \) and

\[
\rho_t^\ast = z - \mu t + \int_0^t f(\rho_s^\ast) \, ds .
\]

(ii) A glimpse on \( \mathbb{S} \) suggests that \( \rho^\ast \) will make a good approximation for the process \( Y^z \). In order to estimate their difference observe that

\[
Y_t^z - \rho_t^\ast = -(S_t - \mu t) + \int_0^t (f(Y_s^z) - f(\rho_s^\ast)) \, ds .
\]

For given \( t > 0 \) define

\[
u_t = \begin{cases}
\sup \{ s < t : Y_s^z \leq \rho_s^\ast \} & \text{on the event } Y_t^z > \rho_t^\ast \\
\sup \{ s < t : Y_s^z \geq \rho_s^\ast \} & \text{on the event } Y_t^z < \rho_t^\ast
\end{cases}
\]

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

\[
Y_t^z - \rho_t^\ast \leq Y_t^z - \rho_t^\ast - \int_{u_t}^t (f(Y_s^z) - f(\rho_s^\ast)) \, ds = (Y_{u_t}^z - \rho_{u_t}^\ast) - (S_t - \mu t) + (S_{u_t} - \mu u_t) .
\]

On the event \( Y_t^z < \rho_t^\ast \) there is an analogous estimate from below, altogether

\[
|Y_t^z - \rho_t^\ast| \leq 2M_t \quad \text{with } M_t := \sup_{u \leq t} |S_u - \mu u| .
\]

Consequently, \( Y_s^z \geq \rho_s^\ast - 2M_s \geq \rho_s^\ast - 2M_t \) for \( s \leq t \) and by means of the monotonicity of \( f \)

\[
\int_0^t f(Y_s^z) \, ds - \int_0^t f(\rho_s^\ast) \, ds \leq \int_0^t f(\rho_s^\ast - 2M_t) \, ds - \int_0^t f(\rho_s^\ast) \, ds \leq 2M_t f(\rho_t^\ast - 2M_t) .
\]

An analogous estimate is valid from below and we obtain

\[
\left| \int_0^t f(Y_s^z) \, ds - \int_0^t f(\rho_s^\ast) \, ds \right| \leq 2M_t f(\rho_t^\ast - 2M_t) . \quad (15)
\]

At this point we recall that under the above assumptions on the subordinator \( S \) by Donsker’s invariance principle we have

\[
M_t = O_P(\sqrt{t})
\]

as \( t \to \infty \).

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

\[
P(T_x^z \geq \beta_x + a\sqrt{z}) = P(Y_t^z \geq x \text{ for all } t \leq \beta_x + a\sqrt{z})
\]

\[
= P \left( S_{\beta_x + a\sqrt{z}} \leq z - x + \int_0^{\beta_x + a\sqrt{z}} f(Y_s^z) \, ds, Y_t^z \geq x \text{ for all } t \leq \beta_x + a\sqrt{z} \right)
\]

\[
\leq P \left( S_{\beta_x + a\sqrt{z}} \leq z - x + f(x)(c + |a|)\sqrt{z} + \int_0^{\beta_x - c\sqrt{z}} f(Y_s^z) \, ds \right)
\]

\[
\leq P \left( S_{\beta_x + a\sqrt{z}} \leq z - x + f(x)(c + |a|)\sqrt{z} + \int_0^{\beta_x - c\sqrt{z}} f(Y_s^z) \, ds \right)
\]

7
We now bring (15) into play. From the definition of $\rho^z$ we have, writing $\beta(y) = \beta_y$, that
\[
\beta(\rho^z(\beta_z - c\sqrt{z})) = c\sqrt{z},
\]
thus because of (14)
\[
\rho^z(\beta_z - c\sqrt{z}) \geq \frac{c\sqrt{z}}{2\mu}.
\]
Then on the event $M_{\beta_z} \leq c\sqrt{z}/(8\mu)$ we have
\[
\rho^z(\beta_z - \sqrt{z}) - 2M_{\beta_z - c\sqrt{z}} \geq \frac{c\sqrt{z}}{2\mu} - \frac{c\sqrt{z}}{4\mu} = \frac{c\sqrt{z}}{4\mu}.
\]
Consequently, by means of (15) and since $\beta_z \leq 2z/\mu$
\[
\begin{align*}
P(T^z_x \geq \beta_z + a\sqrt{z}) &\leq P\left(M_{2z/\mu} > \frac{c\sqrt{z}}{8\mu}\right) \\
&\quad + P\left(S_{\beta_z + a\sqrt{z}} \leq z - x + f(x)(c + |a|)\sqrt{z} + \int_{0}^{\beta_z} f(\rho^z_s) ds + \frac{c\sqrt{z}}{4\mu} f\left(\frac{c\sqrt{z}}{4\mu}\right)\right). \tag{16}
\end{align*}
\]
Moreover, by definition of $\rho^z$,
\[
z + \int_{0}^{\beta_z} f(\rho^z_s) ds = \rho^z(\beta_z) + \mu\beta_z = \mu\beta_z.
\]
Therefore, if we fix $\varepsilon > 0$, let $c$ be so large that the first right-hand probability in (16) is smaller than $\varepsilon$, then choose $z$ so large that $(c/4\mu)f\left(\frac{c\sqrt{z}}{4\mu}\right) \leq \varepsilon$, and also choose $x > 0$ and so large that $c f(x)(c + |a|) \leq \varepsilon$, then we end up with
\[
P(T^z_x \geq \beta_z + a\sqrt{z}) \leq \varepsilon + P\left(S_{\beta_z + a\sqrt{z}} \leq \mu\beta_z + 2\varepsilon\sqrt{z}\right).
\]
Also by the law of large numbers
\[
S_{\beta_z + a\sqrt{z}} - S_{\beta_z} \sim \mu a\sqrt{z}
\]
in probability. Therefore
\[
P(T^z_x \geq \beta_z + a\sqrt{z}) \leq 2\varepsilon + P\left(S_{\beta_z} \leq \mu\beta_z + (-\mu a + 3\varepsilon)\sqrt{z}\right).
\]
Moreover $\mu\beta_z \sim z$, hence
\[
P(T^z_x \geq \beta_z + a\sqrt{z}) \leq 2\varepsilon + P\left(S_{\beta_z} \leq \mu\beta_z + (-\mu a + 4\varepsilon)\mu^{1/2}\sqrt{\beta_z}\right)
\]
for large $z$. Now from assumption (12) and the central limit theorem there follows
\[
\frac{S_t - \mu t}{\sqrt{\sigma^2 t}} \xrightarrow{d} L,
\]
where $L$ denotes a standard normal random variable. Thus
\[
\limsup_{z \to \infty} P(T^z_x \geq \beta_z + a\sqrt{z}) \leq 2\varepsilon + P(L \leq (-\mu a + 4\varepsilon)\mu^{1/2}\sigma^{-1}).
\]
Note that the choice of \( x \) depends on \( \varepsilon \) in our proof. However, since again 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 \( \varepsilon \to 0 \) we obtain
\[
\limsup_{z \to \infty} P \left( \frac{T_z^x - \beta_z}{\sqrt{z}} \geq a \right) \leq P \left( L \leq -\mu^{3/2} \sigma^{-1} a \right).
\]
This is the first part of our claim.

(iv) For the lower estimates we first introduce the random variable
\[
R_{z,x} := \sup \{ t \geq 0 : Y_t^z \geq x \} - \inf \{ t \geq 0 : Y_t^z < x \}
\]
which is the length of the time interval where \( Y_t^z - x \) is changing from positive sign to ultimately 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
\[
\eta_{z,x} := \inf \{ t \geq 0 : Y_t^z < x \}
\]
we have for \( t > \eta = \eta_{z,x} \) because of \( Y_\eta^z \leq x \) and (13)
\[
Y_t^z = Y_\eta^z - (S_t - S_\eta) + \int_\eta^t f(Y_s^z) \, ds \leq x - (S_t - S_\eta) + \frac{\mu}{2} (t - \eta).
\]
Thus \( R_{z,x} \) is bounded from above by
\[
R'_{z,x} := \sup \{ u \geq 0 : (S_{\eta_{z,x} + u} - S_{\eta_{z,x}}) - \mu u / 2 \leq 0 \}
\]
These random variables are a.s. finite. Moreover, they are identically distributed, since \( \eta_{z,x} \) are stopping times. This proves that the \( R_{z,x} \) are uniformly bounded in probability.

Now for the lower bounds we have for \( a, b \in \mathbb{R} \)
\[
P(T_z^x \geq \beta_z + a\sqrt{z}) \geq P(Y_t^z \geq x \text{ for all } t \leq \beta_z + a\sqrt{z}, R_{z,x} \leq b) = P(Y_t^z \geq x \text{ for all } \beta_z + a\sqrt{z} - b \leq t \leq \beta_z + a\sqrt{z}, R_{z,x} \leq b).
\]
For these \( t \) we have
\[
Y_t^z = z - S_t + \int_0^t f(Y_s^z) \, ds \geq z - S_{\beta_z + a\sqrt{z}} + \int_0^{\beta_z + a\sqrt{z} - b} f(Y_s^z) \, ds,
\]
therefore
\[
P(T_z^x \geq \beta_z + a\sqrt{z}) \geq P \left( S_{\beta_z + a\sqrt{z}} \leq z - x + \int_0^{\beta_z + a\sqrt{z} - b} f(Y_s^z) \, ds, R_{z,x} \leq b \right)
\]
\[
\geq P \left( S_{\beta_z + a\sqrt{z}} \leq z - x + \int_0^{\beta_z - c\sqrt{z}} f(Y_s^z) \, ds \right) - P(R_{z,x} > b)
\]
for \( c \) sufficiently large.

We now bring, as in part (iii), (15) into play. Proceeding analogously we obtain instead of (16) the estimate
\[
P(T_z^x \geq \beta_z + a\sqrt{z}) \geq -P(R_{z,x} > b) - P \left( M_{2z/\mu} \geq \frac{c\sqrt{z}}{8\mu} \right)
\]
\[
+ P \left( S_{\beta_z + a\sqrt{z}} \leq z - x + \int_0^{\beta_z - c\sqrt{z}} f(\rho_s^z) \, ds - \frac{c\sqrt{z}}{4\mu} f(\frac{c\sqrt{z}}{4\mu}) \right).
\]
Also, since \( \rho_{\beta_z}^z = 0 \) and \( \dot{\rho}_{t}^z \leq -\mu/2 \),
\[
\int_{\beta_z - c\sqrt{z}}^{\beta_z} f(\rho_{s}^z) \, ds \leq \int_{0}^{\sqrt{z}} f(\mu s/2) \, ds = o(\sqrt{z}) .
\]
Hence, for given \( \varepsilon > 0 \) and \( z \) sufficiently large
\[
P(T_{t}^z \geq \beta_z + a\sqrt{z}) \geq -P(R_{t,x} > b) - P\left(M_{2\sqrt{z}/\mu} > \frac{c\sqrt{z}}{8\mu}\right)
+ P\left(S_{\beta_z + a\sqrt{z}} \leq z - \varepsilon\sqrt{z} + \int_{0}^{\beta_z} f(\rho_{s}^z) \, ds - \frac{c\sqrt{z}}{4\mu} \left(f\left(\frac{c\sqrt{z}}{4\mu}\right)\right)\right).
\]
Returning to the arguments of part (iii) we choose \( b, c \) and then \( z \) so large that we arrive at
\[
P(T_{t}^z \geq \beta_z + a\sqrt{z}) \geq -2\varepsilon + P\left(S_{\beta_z + a\sqrt{z}} \leq \mu\beta_z - 2\varepsilon\sqrt{z}\right)
\]
and further at
\[
\liminf_{z \to \infty} P(T_{t}^z \geq \beta_z + a\sqrt{z}) \geq -3\varepsilon + P(L \leq (-\mu a - 3\varepsilon)\mu^{1/2}\sigma^{-1}) .
\]
The limit \( \varepsilon \to 0 \) leads to the desired lower estimate.

3 Approximating the block counting process

In this section we derive a strengthening of a result in Kersting, Schweinsberg and Wakolbinger \[4\] 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 \( \Lambda \)-coalescent given in \[4\], which is a slight variation of the construction provided by Pitman in \[3\].

This construction requires \( \Lambda(\{0\}) = 0 \), which is fulfilled for coalescents with dust. Consider a Poisson point process \( \Psi \) on \((0, \infty) \times (0, 1] \times [0, 1]^n \) with intensity
\[
dt \times p^{-2}\Lambda(dp) \times du_1 \times \cdots \times du_n ,
\]
and let \( \Pi_n(0) = \{\{1\}, \ldots, \{n\}\} \) be the partition of the integers \( 1, \ldots, n \) into singletons. Suppose \((t, p, u_1, \ldots, u_n)\) is a point of \( \Psi \), and \( \Pi_n(t-) \) consists of the blocks \( B_1, \ldots, B_n \), 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 \( u_i \leq p \) into a single block. These are the only times that mergers occur. This construction is well-defined because almost surely for any fixed \( t' < \infty \), there are only finitely many points \((t, p, u_1, \ldots, u_n)\) of \( \Psi \) for which \( t \leq t' \) and at least two of \( u_1, \ldots, u_n \) are less than or equal to \( p \). The resulting process \( \Pi_n = (\Pi_n(t), t \geq 0) \) is the \( \Lambda \)-coalescent. When \((t, p, u_1, \ldots, u_n)\) is a point of \( \Psi \), we say that a \( p \)-merger occurs at time \( t \).

Condition (7) allows us to approximate the number of blocks in the \( \Lambda \)-coalescent by a subordinator. Let \( \phi : (0, \infty) \times (0, 1] \times [0, 1]^n \to (0, \infty) \times (0, \infty) \) 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 \((S(t), t \geq 0)\) having the property that \( S(0) = 0 \) and, if \((t, x)\) is a point of \( \phi(\Psi) \), then \( S(t) = S(t-) + x \). With \( \lambda \) the Lévy measure of \( S \), the formulas (9) and (12) now read
\[
\mu = \int_{[0,1]} \log \frac{1}{1 - p} \times \frac{\Lambda(dp)}{p^2} \quad \text{and} \quad \sigma^2 = \int_{[0,1]} \left(\log \frac{1}{1 - p}\right)^2 \times \frac{\Lambda(dp)}{p^2} .
\]

10
This subordinator first appeared in the work of Pitman \cite{6} and was used to approximate the block-counting process by Gnedin et al. \cite{2} and Möhle \cite{5}; the benefits of a refined approximation by a drifted subordinator were discovered in \cite{4}. We recall that the drift appears because a merging of $b$ out of $N_n(t)$ lines results in a decrease by $b - 1$ and not by $b$ lines, see equation (23) in \cite{4} for an explanation of the form of the drift. The next result provides a refinement of Theorem 10 in \cite{4}.

**Proposition 6.** Let 
$$
\int_{[0,1]} \frac{\Lambda(dp)}{p} < \infty ,
$$
let $f$ be as in \cite{2}, and let $Y_n$ be the solution of \cite{8} with $z := \log n$. Then for any $\varepsilon > 0$ there is an $\ell < \infty$ such that
$$
P\left( \sup_{t<\tau_n} |\log N_n(t) - Y_n(t)| \leq \ell , \ Y_n(\tau_n) < \ell \right) \geq 1 - \varepsilon .
$$

**Proof.** From \cite{4} we know that for given $\varepsilon > 0$ there is an $r < \infty$ such that
$$
P\left( \sup_{t<\tau_n} |\log N_n(t) - Y_n(t)| \leq r \right) \geq 1 - \varepsilon /2 .
$$

Now we consider the size $\Delta_n$ of the last jump. Letting $(u_i, p_i)$, $i \geq 1$, be the points of the underlying Poisson point process with intensity measure $dt \Lambda(dp)/p^2$, the associated subordinator $S$ has jumps of size $v_i = -\log(1 - p_i)$ at times $t_i$. Thus for any $c > 0$ we have
$$
\{\Delta_n \leq \log N_n(\tau_n -) - c\} = \{\tau_n = t_i \text{ and } -\log(1 - p_i) \leq \log N_n(t_i -) - c \text{ for some } i \geq 1\}
$$
$$
= \{\tau_n = t_i \text{ and } p_i \leq 1 - \frac{e^c}{N_n(t_i -)} \text{ for some } i \geq 1\}
$$

Given $N_n(t-)$ this event appears at time $t$ with rate
$$
\nu_{n,t} = \int_{[0,1-e^c/N_n(t-)]} p^{N_n(t-)} \frac{\Lambda(dp)}{p^2} .
$$

Using the inequalities $p^b = (1 - (1 - p))^b \leq e^{-(1-p)b} \leq 1/((1 - p)b)$ we get
$$
\nu_{n,t} \leq \int_{[0,1-e^c/N_n(t-)]} e^{-(1-p)(N_n(t-)-2)} \Lambda(dp) \leq \int_{[0,1-e^c/N_n(t-)]} \frac{e^2}{(1-p)N_n(t-)} \Lambda(dp) .
$$

It follows
$$
E\left[ \int_0^\infty \nu_{n,t} \, dt \right] \leq E\left[ \int_{[0,1]} \int_0^\infty \frac{e^2}{(1-p)N_n(t-)} I_{\{N_n(t-) \geq [e^c/(1-p)]\}} \, dt \Lambda(dp) \right]
$$

Lemma 14 of \cite{4} yields the estimate
$$
E\left[ \int_0^\infty \frac{1}{N_n(t-)} I_{\{N_n(t-) \geq [e^c/(1-p)]\}} \, dt \right] \leq c_1 \left[ e^c/(1-p) \right]^{-1} \leq c_1 \frac{1 - p}{e^c}
$$

with some $c_1 > 0$, hence
$$
E\left[ \int_0^\infty \nu_{n,t} \, dt \right] \leq c_1 e^{2-c} \Lambda([0,1]) .
$$
Therefore for $c$ sufficiently large

$$E\left[\int_0^\infty \nu_{n,t} \, dt\right] \leq \varepsilon/2,$$

which implies

$$P(\Delta_n \leq \log N_n(\tau_n) - c) = 1 - \exp\left(-E\left[\int_0^\infty \nu_{n,t} \, dt\right]\right) \leq \varepsilon/2.$$

Altogether we obtain

$$P\left(\sup_{t<\tau_n} |\log N_n(t) - Y_n(t)| \leq r, \Delta_n > \log N_n(\tau_n) - c\right) \geq 1 - \varepsilon.$$

The event in the previous formula implies

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

and the claim of the theorem follows with $\ell = r + c$. $\square$

### 4 Proof of the main results

**Proof of Theorem 1.** Let us first assume that $\mu < \infty$. Then we have a coalescent with dust, and we may apply Proposition 6. 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 $T_{\ell}^{\log n} < (1 - \eta) \log n/\mu$. Thus in view of Proposition 6 there exists for any $\varepsilon > 0$ an $\ell$ such that

$$P(\tau_n < (1 - \eta) \log n/\mu) \leq P(T_{\ell}^{\log n} < (1 - \eta) \log n/\mu) + \varepsilon.$$

Proposition 4 implies that the right-hand probability converges to 0 as $n \to \infty$. Letting $\varepsilon \to 0$ we obtain

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

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

$$P(\tau_n > (1 + \eta) \log n/\mu) \leq P(T_{-\ell}^{\log n} > (1 + \eta) \log n/\mu) + \varepsilon.$$

Again the right-hand probability converges to zero in view of Proposition 4 and we obtain

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

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

Now assume $\mu = \infty$. If $\Lambda(\{0\}) > 0$, then the coalescent comes down from infinity and $\tau_n$ stays 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 $\Lambda^\varepsilon$ by $\Lambda^\varepsilon(B) := \Lambda(B \cap [\varepsilon, 1 - \varepsilon])$. Obviously

$$\mu^\varepsilon := \int_0^1 \log \frac{1}{1 - p} \frac{\Lambda^\varepsilon(dp)}{p^2} < \infty.$$
Thus for the absorption times $\tau_n^\varepsilon$ of the $\Lambda^\varepsilon$-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 $\Lambda^\varepsilon$-coalescent in an obvious manner to the $\Lambda$-coalescent in such a way that $N_n(t) \leq N_n^\varepsilon(t)$ a.s. for all $t \geq 0$, in particular $\tau_n \leq \tau_n^\varepsilon$. Hence it follows that
\[
P(\tau_n/\log n > 2/\mu^\varepsilon) \to 0.
\]
Because of $\Lambda(\{0,1\}) = 0$ we have $\mu^\varepsilon \to \mu = \infty$ with $\varepsilon \to 0$, consequently
\[
P(\tau_n/\log n > \eta) \to 0
\]
for all $\eta > 0$. This is our claim.

**Proof of Theorem 3.** Because of the condition $\mu < \infty$ we again may apply Proposition 6. We follow the same line as in the previous proof: For $\varepsilon > 0$ there exists an $\ell$ such that for all $a \in \mathbb{R}$
\[
P(\tau_n < b_n + a\sqrt{n}) \leq P(T^\log n_{\ell} < b_n + a\sqrt{n}) + \varepsilon
\]
and
\[
P(\tau_n > b_n + a\sqrt{n}) \leq P(T^\log n_{-\ell} > b_n + a\sqrt{n}) + \varepsilon
\]
Now apply Proposition 6 and let $\varepsilon \to 0$. 

**Proof of Proposition 3.** (i) Let us first assume (6). Because of $1 - (1 - p)^{1/r} \leq \min(p/r, 1)$ for $0 < r < 1$ we have for $\alpha > 0$
\[
f(\log \frac{1}{r}) \leq \int_0^{e^{\alpha}} \frac{\Lambda(dp)}{p} + r \int_{e^{\alpha}}^1 \frac{\Lambda(dp)}{p^2} \leq \int_0^{e^{\alpha}} \frac{\Lambda(dp)}{p} + r^{1-\alpha} \int_0^1 \frac{\Lambda(dp)}{p} .
\]
(17)
Also, because of $1 - (1 - p)^{1/r} \geq 1 - e^{-p/r} \geq e^{-p/r}p/r$, it follows for $\beta > 0$ that
\[
f(\log \frac{1}{r}) \geq e^{-p^\beta - 1} \int_0^{e^{\beta}} \frac{\Lambda(dp)}{p}.
\]
(18)
Together with (6) these two estimates yield for $\alpha < 1 < \beta$
\[
c\beta^{-1/2} \leq \liminf_{r \to 0} f(\log \frac{1}{r}) \sqrt{\log \frac{1}{r}} \leq \limsup_{r \to 0} f(\log \frac{1}{r}) \sqrt{\log \frac{1}{r}} \leq c\alpha^{-1/2} .
\]
Letting $\alpha, \beta \to 1$ we arrive at $f(y) = (c + o(1))/\sqrt{y}$ as $y \to \infty$ and consequently
\[
\int_0^{\log n} f(y) \, dy = (c + o(1))2\sqrt{\log n}
\]
as $n \to \infty$.

Now, because of
\[
\frac{1}{\mu - f(y)} = \frac{1}{\mu} + \frac{f(y)}{\mu(\mu - f(y))}
\]
and \( f(y) = o(1) \) as \( y \to \infty \), we have
\[
\int_{\nu}^{z} \frac{dy}{\mu - f(y)} = \frac{z}{\mu} + \frac{1 + o(1)}{\mu^2} \int_{0}^{z} f(y) \, dy + O(1)
\]
as \( z \to \infty \), and consequently, as claimed,
\[
b_n = \frac{\log n}{\mu} + \frac{2c + o(1)}{\mu^2} \sqrt{\log n}.
\]

(ii) Now suppose that (5) is satisfied. Then in view of (19) with \( z = \log n \) it follows that
\[
\int_{0}^{\log n} f(y) \, dy = (2c + o(1)) \sqrt{\log n}
\]
as \( n \to \infty \), or equivalently
\[
\int_{0}^{z} f(y) \, dy = (2c + o(1)) \sqrt{z}
\]
for \( z \to \infty \). This implies that \( f(z) = (c + o(1))/\sqrt{z} \) as \( z \to \infty \). For \( c = 0 \) this claim follows because \( f \) is decreasing, which entails
\[
z f(z) \leq \int_{0}^{z} f(y) \, dy = o(\sqrt{z}).
\]

For \( c > 0 \) we use the estimate
\[
\frac{1}{\eta \sqrt{z}} \int_{z}^{(1+\eta)z} f(y) \, dy \leq \sqrt{\eta} f(z) \leq \frac{1}{\eta \sqrt{z}} \int_{(1-\eta)z}^{z} f(y) \, dy
\]
with \( \eta > 0 \). Taking the limit \( z \to \infty \) and then \( \eta \to 0 \) yields \( f(z) = (c + o(1))/\sqrt{z} \). Now, similar as in part (i) we get from (17) and (18)
\[
c \sqrt{\alpha} \leq \liminf_{r \to 0} \sqrt{\log \frac{1}{r}} \int_{[0, r]} \frac{\Lambda(dp)}{p} \leq \limsup_{r \to 0} \sqrt{\log \frac{1}{r}} \int_{[0, r]} \frac{\Lambda(dp)}{p} \leq c \sqrt{\beta}.
\]

With \( \alpha, \beta \to 1 \) we arrive at (6).

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

References

[1] Herriger, Ph. and Möhle, M. (2012). Conditions for exchangeable coalescents to come down from infinity. ALEXA 9, 637–665.

[2] Gnedin, A., Iksanov, A., and Marynych, A. (2011). On \( \Lambda \)-coalescents with dust component, J. Appl. Probab. 48, 1133–1151.

[3] Goldschmidt, Ch. and Martin, J. (2005). Random Recursive Trees and the Bolthausen-Sznitman Coalescent. Electron. J. Probab. 10, 718–745.
[4] Kersting, G., Schweinsberg, J., and Wakolbinger, A. (2017). The size of the last merger and time-reversal in Λ-coalescents. To appear in *AIHP*, electronic version available at [http://imstat.org/aihp/accepted.html](http://imstat.org/aihp/accepted.html)

[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. Probab.* **27**, 1870–1902.

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

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