*J. Appl. Prob.* **48,** 1133–1151 (2011)
*Printed in England*
© *Applied Probability Trust* 2011

# ON Λ-COALESCENTS WITH DUST COMPONENT

ALEXANDER GNEDIN,* *Utrecht University*
ALEXANDER IKSANOV ** *** AND
ALEXANDER MARYNYCH,** **** *National Taras Shevchenko University of Kiev*

## Abstract

We consider the Λ-coalescent processes with a positive frequency of singleton clusters. The class in focus covers, for instance, the beta$(a, b)$-coalescents with $a > 1$. We show that some large-sample properties of these processes can be derived by coupling the coalescent with an increasing Lévy process (subordinator), and by exploiting parallels with the theory of regenerative composition structures. In particular, we discuss the limit distributions of the absorption time and the number of collisions.

*Keywords:* Absorption time; coupling; Λ-coalescent; number of collisions; regenerative composition structure; subordinator

2010 Mathematics Subject Classification: Primary 60C05; 60G09
Secondary 60F05; 60J10; 60K05

## 1. Introduction

The Λ-coalescent with values in partitions of $n$ integers is a Markovian process $\Pi_n = (\Pi_n(t))_{t\geq 0}$ which starts at $t = 0$ with $n$ singletons and evolves according to the rule: for each $t \geq 0$, when the number of clusters is $m$, each $k$ tuple of clusters merges into one cluster at probability rate

$$\lambda_{m,k} = \int_0^1 x^k (1-x)^{m-k} \nu(\mathrm{d}x), \qquad 2 \leq k \leq m, \tag{1.1}$$

where $\nu$ is a measure on the unit interval with finite second moment. The integral representation of rates (1.1) ensures that the processes $\Pi_n$ can be defined consistently for all $n$ as restrictions of a coalescent process $\Pi_\infty$ which starts with infinitely many clusters and assumes values in the set of partitions of $\mathbb{N}$; see [24]. The infinite coalescent $\Pi_\infty$ may be regarded as a limiting form of $\Pi_n$ as $n \to \infty$, and uniquely connected to a process with values in the infinite-dimensional space of partitions of a unit mass. The Λ-coalescents were introduced in the papers by Pitman [24] and Sagitov [26], where parameterization by the finite measure $\Lambda(\mathrm{d}x) = x^2 \nu(\mathrm{d}x)$ was used. The reader is referred to the recent lecture notes [2] and [4] for an accessible introduction to the theory of Λ-coalescents and a survey.

After a number of collisions (merging events), $\Pi_n$ enters the absorbing state with a sole cluster. Two basic characteristics of the speed of the coalescence are the *absorption time* $\tau_n$ and the *number of collisions* $X_n$. The large-$n$ properties of $\tau_n$ and $X_n$ are strongly determined by the concentration of measure $\nu$ on the unit interval near the endpoints of [0, 1].

Received 8 February 2011; revision received 27 May 2011.

* Postal address: Department of Mathematics, Utrecht University, Postbus 80010, 3508 TA Utrecht, The Netherlands. Email address: a.v.gnedin@uu.nl
** Postal address: Faculty of Cybernetics, National Taras Shevchenko University of Kiev, 01033 Kiev, Ukraine.
*** Email address: iksan@univ.kiev.ua
**** Email address: marynych@unicyb.kiev.ua

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

We suppose that $\nu$ has no mass at 1, which excludes forced termination of $\Pi_\infty$ at an independent exponential time. The coalescent is said to come down from infinity if $\Pi_\infty(t)$ has finitely many clusters for each $t > 0$ almost surely; then the $\tau_n$ converge to a finite random variable $\tau_\infty$ which is the absorption time of $\Pi_\infty$. Otherwise, $\Pi_\infty(t)$ almost surely stays with infinitely many clusters for all $t$. There is a delicate criterion in terms of the rates $\lambda_{m,k}$ to distinguish between the two alternatives [27].

In this paper we shall study $\tau_n$ and $X_n$ under the assumption that $\Pi_\infty$ stays infinite because infinitely many of the original clusters do not engage in collisions before any given time $t > 0$. This family of $\Lambda$-coalescents can be characterized by the moment condition

$$\int_0^1 x\nu(\mathrm{d}x) < \infty. \tag{1.2}$$

We call the collection of singleton clusters of $\Pi_\infty(t)$ the *dust component*. The dust component has a positive total frequency, meaning that the number of singletons within $\Pi_n(t)$ grows approximately linearly in $n$ as $n \to \infty$.

The coalescents with dust component do not exhaust all coalescents which stay infinite. One distinguished example is the Bolthausen–Sznitman coalescent with $\nu(\mathrm{d}x) = x^{-2}\,\mathrm{d}x$, which stays infinite although (1.2) fails. Such examples on the border between 'coming down from infinity' and 'possessing the dust component' are more of an exception if one considers, e.g. measures $\nu$ satisfying a condition of regular variation near 0.

Under (1.2), every transition of $\Pi_\infty$ will involve infinitely many singletons. This suggests that most of the collision events of $\Pi_n$ will involve some of the original $n$ clusters for large $n$. Another way to express this idea is to say that in a tree representing the complete merging history of $\Pi_n$, most of the internal nodes are linked directly to one of $n$ leaves. We will show that this intuition is indeed correct, to the extent that the behaviours of $\tau_n$ and $X_n$ can be derived from that of analogous quantities associated with the evolution of the dust component. In turn, the total frequency of the dust component of $\Pi_\infty$ undergoes a relatively simple process, which may be represented as $\exp(-S_t)$, where $S = (S_t)_{t\ge 0}$ is a subordinator. Similarly, for $\Pi_n$, the engagement of original $n$ clusters in their first collisions follows a Markovian process, which has been studied in the context of regenerative composition structures derived from subordinators [10]. A coupling of $\Pi_\infty$ with $S$ will enable us to apply known results about the level passage for subordinators, and about the asymptotics of regenerative composition structures.

The connection between $\Pi_\infty$ and $S$ was first explored in [13] in the special case when $\nu$ is a finite measure; hence, the subordinator $S$ is a compound Poisson process. While in the present paper we are mainly interested in infinite $\nu$, the case of finite $\nu$ is not excluded. Moreover, we will be able to extend the results of [13] by removing a condition on $\nu$ imposed in that paper. In [22] coupling of the dust component with a subordinator was applied to more general coalescents with multiple simultaneous collisions.

In a recent paper by Haas and Miermont [18] results on counting collisions in the coalescent and counting blocks in the regenerative composition were derived separately in the context of absorption times of decreasing Markov chains. Our approach adds some insight to the connection between these two models, and it entails some delicate features, such as differentiating between collisions that involve some original clusters of $\Pi_n$ and collisions that do not.

The possible modes of behaviour of $\tau_n$ and $X_n$ for large $n$ are best illustrated by the family of coalescents driven by a beta measure

$$\nu(\mathrm{d}x) = cx^{a-3}(1-x)^{b-1}\,\mathrm{d}x, \qquad a, b, c > 0. \tag{1.3}$$

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

These coalescents come down from infinity for $a < 1$ and stay infinite for $a \geq 1$. The results of the present paper in combination with the results of previous studies can be summarized as follows.

*Case* $0 < a < 1$. The limit law of $(X_n - (1-a)n)/n^{1/(2-a)}$ is $(2-a)$-stable (see [11], and [20] for the case $b = 1$). The distribution of $\tau_\infty$ is unknown.

*Case* $a = 1$. The instance $b = 1$ is the Bolthausen–Sznitman coalescent, for which the limit distribution of $\tau_n - \log\log n$ is standard Gumbel [9], [17], while $X_n$, suitably scaled and centred, converges weakly to a 1-stable distribution [7], [19]. The case $b \neq 1$ remains open.

*Case* $1 < a < 2$. In the sequel we show that $(\tau_n - c\log n)/(\log n)^{1/2}$ (with suitable $c > 0$) converges weakly to a normal distribution, and that $X_n/n^{2-a}$ converges to the exponential functional of a subordinator. The result about $X_n$ was proved previously in [18], and in [20] for the case $b = 1$.

*Case* $a \geq 2$. In the present paper we prove that normal limits hold for both $\tau_n$ and $X_n$ with explicitly determined scaling and centring. In the case $a > 2$ these asymptotics were previously shown in [13]. In the case $a = 2$ the result for $X_n$ was derived in [21].

## 2. The coalescent and singleton clusters

In the role of the state space of the coalescent $\Pi_n$ with $n$ clusters initially we take the set of partitions of $[n] := \{1, \ldots, n\}$, in which every singleton cluster is classified as either *primary* or *secondary*. We call the collection of primary clusters the dust component of $\Pi_n(t)$. Every nonsingleton cluster of $\Pi_n(t)$ is regarded as secondary. For notational convenience, the clusters are written so that their minimal elements are in increasing order, the elements within the clusters are written in increasing order, and the secondary clusters are written in brackets. For instance, 1 (2) (3 5 6) 4 7, a partition of the set [7], has three primary clusters and two secondary: 1, 4, 7 and (2), (3 5 6), respectively.

Introduce $\lambda_{m,1}$ as in (1.1) with $k = 1$. We have $\lambda_{m,1} < \infty$ by assumption (1.2).

We define the $\Lambda$-coalescent $\Pi_n$ as a càdlàg Markov process with values in such partitions of $[n]$ and the initial state $1\ 2\ \cdots\ n$ with $n$ primary clusters. Each admissible transition either merges some clusters into one cluster, or turns a primary singleton cluster into a secondary singleton cluster. From a partition with $m$ clusters, the transition rate for merging each particular $k$-tuple of $m$ clusters into one cluster is $\lambda_{m,k}$ $(2 \leq k \leq m)$, and the transition rate for turning each particular primary singleton cluster into a secondary singleton cluster is $\lambda_{m,1}$. For instance, the sequence of distinct states visited by $\Pi_7$ could be

$$\begin{aligned} 1\,2\,3\,4\,5\,6\,7 &\to 1\,2\,(3\,5\,6)\,4\,7 \\ &\to 1\,(2)\,(3\,5\,6)\,4\,7 \\ &\to 1\,(2\,4)\,(3\,5\,6)\,7 \\ &\to 1\,(2\,3\,4\,5\,6\,7) \\ &\to (1\,2\,3\,4\,5\,6\,7). \end{aligned}$$

Let $N_n(t)$ be the number of clusters in $\Pi_n(t)$. Then $N_n = (N_n(t))_{t\geq 0}$ is a nonincreasing Markov process, with transition rate

$$\varphi_{m,k} := \binom{m}{k}\lambda_{m,k}$$

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

for jumping from $m$ to $m-k+1$ for $2 \le k \le m$. Turning a primary singleton cluster into a secondary singleton cluster does not cause a jump of $N_n$. The absorption time of $\Pi_n$ can be recast as $\tau_n = \inf\{t \colon N_n(t) = 1\}$, and the number of collisions $X_n$ is equal to the number of jumps the process $N_n$ needs to proceed from $n$ to 1 (which is four in the above example where the second transition does not alter the number of clusters).

Removing element $n$ transforms the partition of $[n]$ into the partition of $[n-1]$. For example, partitions 1 (2 4) (3), 1 (2) 3 4, and 1 (2) 3 (4) all become 1 (2) 3. Restricting in this way $\Pi_n$ to $[n-1]$, pointwise in $t \ge 0$, yields a stochastic copy of $\Pi_{n-1}$. This follows as in [24] since the rates satisfy the recursion $\lambda_{m,k} = \lambda_{m+1,k} + \lambda_{m+1,k+1}$ for all $1 \le k \le m$. Therefore, we may define $\Pi_n$ on the same probability space consistently for all $n$. An explicit realization will appear in the sequel.

The projective limit of the processes $\Pi_n$, $n \in \mathbb{N}$, is a Markov process $\Pi_\infty$ starting at $t = 0$ with the infinite configuration of primary clusters $1\ 2\ \cdots$, and assuming values in the space of partitions of the infinite set $\mathbb{N}$. Each partition $\Pi_\infty(t)$ has only primary singletons, namely those original clusters which do not engage in collisions up to time $t$. For a generic singleton, e.g. labelled 1, the time before its first collision has an exponential distribution with parameter $\lambda_{1,1}$, and when such a collision occurs, infinitely many other clusters are engaged.

The differentiation of singletons of $\Pi_n(t)$ into primary and secondary becomes transparent by considering $\Pi_n$ as the restriction of $\Pi_\infty$ on $[n]$. The secondary singletons of $\Pi_n(t)$ are the unique representatives in $[n]$ of some infinite clusters of $\Pi_\infty(t)$. The primary singletons of $\Pi_n(t)$ are also singletons in the partition $\Pi_\infty(t)$.

There is a construction of $\Pi_\infty$ based on a planar Poisson point process in the strip $[0, 1] \times [0, \infty)$ with intensity measure $\nu(\mathrm{d}x) \times \mathrm{d}t$; see [2], [4], and [24]. With each atom $(t, x)$ we associate a transition of $\Pi_\infty$ performed by tossing a coin with probability $x$ for heads. To pass from $\Pi_\infty(t-)$ to $\Pi_\infty(t)$, the coin is tossed for each cluster of $\Pi_\infty(t-)$, then those clusters marked heads are merged into one, while the clusters marked tails remain unaltered. Although there are infinitely many transitions within any time interval, if $\nu$ is an infinite measure, condition (1.2) ensures that $\Pi_\infty$ does not terminate. In the case of finite $\nu$ transitions of $\Pi_\infty$ occur at the epochs of the Poisson process with rate $\nu([0, 1])$.

Let $N_n^*(t)$ be the number of primary clusters in $\Pi_n(t)$. By homogeneity properties of $\Pi_n$, the process $N_n^* = (N_n^*(t))_{t\ge0}$ is a nonincreasing Markov process, jumping at rate $\varphi_{m,k}$ from $m$ to $m-k$ for $1 \le k \le m$. Let

$$\tau_n^* := \inf\{t \colon N_n^*(t) = 0\}$$

be the random time when the last of $n$ primary clusters disappears. For $1 \le r \le n$, let $K_{n,r}$ be the number of decrements of size $r$ of $(N_n^*)$ on the way from $n$ to 0, let $K_n := \sum_{r=1}^n K_{n,r}$ be the total number of decrements of $(N_n^*)$, and let $X_n^*$ be the number of nonunit decrements of $(N_n^*)$. Obviously,

$$X_n^* = K_n - K_{n,1}. \tag{2.1}$$

We call the clusters of partition $\Pi_n^*(\tau_n^*)$ that remain at time $\tau_n^*$ *residual*, and we denote by $R_n$ the number of residual clusters.

Processes $N_n$ and $N_n^*$ look very similar; thus, at a first glance it might seem surprising that $N_n^*$ is much easier to handle. The simplification comes from the identification of the sequence of decrements of $N_n^*$ with the $n$th level of a regenerative composition structure [10], and further connection to the range of a subordinator. The main new contribution of the present paper is that $N_n^*$ yields a good approximation for $N_n$ for large $n$; thus, $X_n^*$ and $\tau_n^*$ are close to their

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

counterparts $X_n$ and $\tau_n$. In one direction, the connection is quite obvious:

$$X_n^* \le X_n, \qquad N_n^*(t) \le N_n(t), \qquad \tau_n^* \le \tau_n.$$

For instance, the first inequality holds since every collision taking at least two primary clusters contributes to $X_n$, and since, with positive probability, some $R_n \ge 2$ clusters remain at time $\tau_n^*$ when the last primary clusters disappears.

## 3. Coupling with a subordinator

Condition (1.2) implies that there exists a pure-jump subordinator $S = (S_t)_{t \ge 0}$ with the Laplace transform

$$\mathrm{E}(\mathrm{e}^{-zS_t}) = \mathrm{e}^{-t\Phi(z)}, \qquad z \ge 0, \tag{3.1}$$

where the Laplace exponent is given by

$$\Phi(z) := \int_0^1 (1-(1-x)^z)\nu(\mathrm{d}x).$$

The coalescent process will be represented in terms of the passage of $S$ through multiple exponentially distributed levels. We describe first the evolution of the dust component.

Let $\varepsilon_1, \dots, \varepsilon_n$ be independent of $S$ independent and identically distributed standard exponential random variables, and let $\varepsilon_{n:n} < \cdots < \varepsilon_{n:1}$ be their order statistics. It is not difficult to see that $\Phi(n)$ coincides with the probability rate at which the subordinator passes through the level $\varepsilon_{n:n}$ from any state $S_t = s < \varepsilon_{n:n}$. The following lemma extends this observation.

**Lemma 3.1.** *For $t \ge 0$, conditionally given $S_t = s$ with $s \in (\varepsilon_{n:m+1}, \varepsilon_{n:m})$, the subordinator passes through $\varepsilon_{n:m}$ at rate $\Phi(m)$, and hits during this passage each of the intervals $(\varepsilon_{n:m-k+1}, \varepsilon_{n:m-k})$ at rate $\varphi_{m,k}$ for $1 \le k \le m \le n$.*

*Proof.* The proof exploits the Lévy–Khintchine formula (3.1) and the memoryless property of the exponential distribution. See computations around [10, Proof of Theorem 5.2(i)] for details.

Now suppose that each of the primary clusters $1\ 2\ \cdots\ n$ is given an exponential mark $\varepsilon_1, \dots, \varepsilon_n$, and that, for every $t \ge 0$, the marks $\varepsilon_j > S_t$ are associated with primary clusters $j$ existing at time $t$. If $t$ is a jump time of $S$ and the interval $(S_{t-}, S_t]$ covers exactly one mark $\varepsilon_j$, we interpret the event of passage through $\varepsilon_j$ as turning the primary cluster $j$ into a secondary cluster. If $(S_{t-}, S_t]$ covers at least two of the $\varepsilon_j$s, we interpret this event as a collision which involves the corresponding primary clusters. Setting $N_n^*(t) := \#\{j \in [n] \colon \varepsilon_j > S_t\}$ we obtain a process with the desired rates $\varphi_{m,k}$ for transition from $m$ to $m-k$, as it follows from Lemma 3.1. In particular, $\Phi(n) = \sum_{k=1}^n \varphi_{n,k}$ coincides with the total transition rate of the coalescent $\Pi_n$ from the initial state $1\ 2\ \cdots\ n$.

An ordered partition of the set $[n]$ is defined by sending $i, j$ to the same block if and only if $T_{\varepsilon_i} = T_{\varepsilon_j}$. The ordered collection of block sizes of the partition is a composition of integer $n$ (i.e. partition of $n$ taken together with some fixed ordering of the collection of parts). The random composition obtained in this way has the property of regeneration with respect to a part deletion, which is a combinatorial analogue of the regenerative property of the range of a subordinator. See [10] for details. The number of blocks of the partition is equal to the number of jumps of $N_n^*$ prior to the absorption at state 0.

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

These evolutions of primary clusters are consistent in $n$. Assigning the exponential marks $\varepsilon_1, \varepsilon_2, \ldots$ to infinitely many primary clusters $1\ 2\ \cdots$ defines the initial state of the dust component. The frequency of the dust component of $\Pi_\infty$ as time passes is the decaying process $(\exp(-S_t))_{t\geq 0}$.

One straightforward application of the representation by $S$ concerns $\tau_n^*$, the maximal lifetime of primary clusters in $\Pi_n$. Let

$$T_s := \inf\{t \geq 0\colon S_t > s\}$$

be the first passage time through level $s \geq 0$. We can identify $\tau_n^*$ with $T_{\varepsilon_{n:1}}$, and, hence, connect the limit behaviour of $\tau_n^*$ to that of $T_s$ for high levels $s$. Indeed, from the extreme value theory, it is known that $\varepsilon_{n:1} - \log n$ converges in distribution, as $n \to \infty$, to a random variable with the Gumbel distribution. It is also known that the scaled and centred random variables $(T_s - g(s))/f(s)$ can converge in distribution only if the normalizing constant $f(s)$ goes to $\infty$ with $s$. Thus, $T_{\varepsilon_{n:1}}$ and $T_{\log n}$ have the same limit law, if any. Moreover, it can be shown that $(T_s - g(s))/f(s)$ converges weakly to a given proper and nondegenerate probability law if and only if the same holds for $(T_s' - g(s))/f(s)$, where $T_s'$ is the number of points within $[0, s]$ of a random walk which starts at 0 and has the generic step distributed like $S_1$. See [5] (or Proposition 27 of [23]) for a complete list of limit distributions of $T_s'$ and the conditions of convergence. Summarizing the above, we have the following result.

**Proposition 3.1.** *For constants $a_n > 0$ and $b_n \in \mathbb{R}$, if one of the random variables $(\tau_n^* - b_n)/a_n$ and $(T_{\log n} - b_n)/a_n$ converges weakly, as $n \to \infty$, to a nondegenerate proper distribution then the other random variable converges weakly to this distribution too.*

To realize the full dynamics of $\Pi_n$ in terms of the level passage, a mark is assigned to each cluster according to the following rule. At time 0 the marks $\varepsilon_1, \ldots, \varepsilon_n$ represent the primary clusters $1\ 2\ \cdots\ n$. At time $t > 0$ there is some collection of marks on $[S_t, \infty)$ representing the clusters existing at this time. If at time $t > 0$ the subordinator passes through exactly $k$ marks corresponding to some clusters $I_1, \ldots, I_k \subset [n]$, then a new cluster $I_1 \cup \cdots \cup I_k$ is born and assigned a mark $S_t + \varepsilon$, where $\varepsilon$ is a copy of the unit exponential random variable, independent of $S$ and all other marks assigned before $t$. For instance, if at the first passage time $t = T_{\varepsilon_{n:n}}$ the subordinator jumps through exactly $k$ levels $\varepsilon_{j_1}, \ldots, \varepsilon_{j_k}$ out of $\varepsilon_1, \ldots, \varepsilon_n$, then the secondary cluster $J = \{j_1, \ldots, j_k\}$ is born (which is a singleton if $k = 1$) and assigned a mark exponentially distributed on $[S_t, \infty)$.

In particular, when $S$ passes at some time $t$ through only one mark, there is no change in $\Pi_n(t)$, and the mark of the corresponding singleton cluster is just reassigned.

## 4. The absorption time

We wish to exploit the lifetime $\tau_n^*$ of primary clusters as an approximation to the absorption time $\tau_n$. At time $\tau_n^*$ the coalescent process is left with $R_n$ residual clusters, whence the distributional identity

$$\tau_0 := 0, \qquad \tau_n \overset{\mathrm{D}}{=} \tau_n^* + \tilde{\tau}_{R_n}, \quad n \in \mathbb{N}, \tag{4.1}$$

where $\tilde{\tau}_m$ is assumed to be independent of $(\tau_n^*, R_n)$ and distributed like $\tau_m$ for each $m \in \mathbb{N}_0$. To address the quality of approximation, we need to estimate $R_n$.

We begin with some preparatory work. By the first transition, the Markov chain $N_n^*$ goes from $n$ to a state with distribution $p_{n,k} := \varphi_{n,n-k}/\Phi(n)$, $0 \leq k \leq n-1$. Let $g_{n,k}$ be the probability that $N_n^*$ ever visits state $k$, so in terms of the realization via a subordinator,

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

$g_{n,k} = \mathrm{P}(T_{\varepsilon_{n:k+1}} < T_{\varepsilon_{n:k}})$ is the probability that the interval $[\varepsilon_{n:k+1}, \varepsilon_{n:k}]$ intersects the range of $S$. An explicit formula for $g_{n,k}$ in terms of $\Phi$ is available (see [10, Equation (50)]), but it is complicated and inconvenient for computations.

**Lemma 4.1.** *Suppose that $(r_k)_{k\in\mathbb{N}}$ is a nonnegative sequence such that the sequence*

$$\left(\frac{\Phi(k)r_k}{k}\right)_{k\in\mathbb{N}}$$

*is nonincreasing. Then the sequence $(a_n)_{n\in\mathbb{N}_0}$ defined by*

$$a_0 = 0, \qquad a_n := \sum_{k=1}^{n} g_{n,k} r_k, \quad n \ge 1,$$

*satisfies the relation*

$$a_n = O\biggl(\sum_{k=1}^{n} \frac{r_k \Phi(k)}{k}\biggr) \quad \textit{as } n \to \infty.$$

*Proof.* The assertion follows from Lemma A.1 in Appendix A. Indeed, conditioning on the size of the first jump of $N_n^*$, we see that the sequence $(a_n)$ satisfies the recurrence

$$a_0 = 0, \qquad a_n = r_n + \sum_{k=0}^{n-1} p_{n,k} a_k, \quad n \in \mathbb{N}.$$

To apply Lemma A.1, we take $\psi_n = \Phi(n)$. Lemma A.2(ii) holds by the assumptions and Lemma A.2(i) follows from

$$\Phi(n)\sum_{k=0}^{n}\biggl(1-\frac{k}{n}\biggr)p_{n,k} = \frac{1}{n}\sum_{k=0}^{n-1}(n-k)\varphi_{n,n-k} = \frac{1}{n}\sum_{k=1}^{n} k\varphi_{n,k} = \int_0^1 x\theta(\mathrm{d}x) > 0.$$

Note that, since the function $s \mapsto \Phi(s)/s$ is nonincreasing, the sequence $(\Phi(k)r_k/k)$ is nonincreasing whenever $(r_k)$ is itself nonincreasing.

Define $\vec{\nu}(x) := \nu([x, 1]),\ x \in (0, 1)$.

**Lemma 4.2.** *If either of the equivalent conditions*

$$\int_0^1 x^{-1}\,\mathrm{d}x \int_0^x \vec{\nu}(y)\,\mathrm{d}y < \infty \tag{4.2}$$

*or*

$$\sum_{k=1}^{\infty} \frac{\Phi(k)}{k^2} < \infty \tag{4.3}$$

*holds, then*

$$\mathrm{E}R_n = O(1) \quad \textit{as } n \to \infty,$$

*in which case the sequence of distributions of the $R_n$s is tight.*

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

*Proof.* The equivalence of (4.2) and (4.3) is established by repeated integration by parts.

In the genealogical history of each residual cluster there is the last event, collision or switch, involving some primary clusters. If a secondary cluster $b$ is born at some time $t \le \tau_n^*$ of such an event, and if at this time some $j \ge 0$ other primary clusters coexist, then $b$ corresponds to a residual cluster provided that $b$ and its followers do not collide with these $j$ primary clusters or their followers before time $\tau_n^*$. That is to say, $b$ and the $j$ primary clusters belong to distinct branches if the coalescent tree is cut at time $\tau_n^*$. Let $q_j$ be the probability that such a secondary cluster $b$, born at the time of a merger when some $j$ primary clusters remain, becomes a residual cluster. Restricting the coalescent to $j+1$ clusters, it is seen that $q_j$ indeed depends only on $j$. The consistency property of the coalescent with respect to the restrictions entails that $q_j$ is decreasing in $j$. Averaging over the times when some primary clusters get engaged, we find the expected number of residual clusters

$$\mathrm{E}R_n = \sum_{j=0}^{n-1} g_{n,j} q_j, \tag{4.4}$$

where $g_{n,j}$ is the probability that at some instant exactly $j$ primary clusters remain.

Furthermore, given $S_t = s$, we have exactly $j$ exponential marks of the primary clusters larger than $s$. The cluster $b$ is assigned a new exponential mark $u = s + \varepsilon$ which lies within each of the spacings in $(s, \infty)$ generated by $\varepsilon_{n:j}, \dots, \varepsilon_{n:1}$ with the same probability $1/(j+1)$. If this spacing is $(\varepsilon_{n:k+1}, \varepsilon_{n:k})$ then $b$ *may* correspond to a residual cluster only if (i) $T_{\varepsilon_{n:k+1}} < T_u < T_{\varepsilon_{n:k}}$ and (ii) $b$ does not collide further with $k$ primary clusters and their followers before time $\tau_n^*$. If (i) occurs, condition (ii) is not sufficient for the correspondence since possible collisions with some of $j$ primary clusters or their followers are ignored. This leads to the inequality

$$q_j \le \frac{1}{j+1} \sum_{k=0}^{j} g_{j+1,k+1} p_{k+1,k} q_k, \qquad 1 \le j \le n-1$$

(with $q_0 = 1$), where $g_{j+1,k+1}$ is the probability that the number of primary clusters goes down from $j+1$ to $k+1$, $p_{k+1,k}$ is the probability that the next event involves only cluster $b$, and $q_k$ is the probability that $b$ avoids collisions thereafter until no primary cluster remains. Substituting $\varphi_{k,1} = k(\Phi(k) - \Phi(k-1))$, we obtain

$$q_j \le \frac{1}{j+1} \sum_{k=1}^{j+1} g_{j+1,k} \frac{k(\Phi(k) - \Phi(k-1))}{\Phi(k)} q_{k-1} \le \frac{c}{j+1} \sum_{k=1}^{j+1} (\Phi(k) - \Phi(k-1)) q_{k-1},$$

where Lemma 4.1 was applied with

$$r_k = \frac{k(\Phi(k) - \Phi(k-1)) q_{k-1}}{\Phi(k)}.$$

The required monotonicity condition holds since both $q_k$ and $\Phi(k) - \Phi(k-1)$ are decreasing in $k$, the latter by the concavity of $\Phi$. Here and throughout, $c$ will denote a positive constant whose value is not important and may change from line to line.

Setting $a_j = (j+1)q_j$ and $b_j = c(\Phi(j+1) - \Phi(j))/(j+1)$, we obtain, from the above,

$$a_j \le \sum_{k=0}^{j} b_k a_k, \qquad j \in \mathbb{N}_0.$$

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

We want to show that the sequence $(a_j)$ is bounded. To that end, let $M_j := \max_{i=0,\dots,j} a_i$. Then

$$M_j \le \sum_{k=0}^{j} b_k M_k.$$

Since $\Phi(j)/j$ decreases, we have $\Phi(j+1) - \Phi(j) \le \Phi(j+1)/(j+1)$, which taken together with (4.3) implies that the series $\sum_{k=0}^{\infty} b_k$ converges, so we can choose

$$n_0 := \inf\left\{k \ge 0 \colon \sum_{i=k}^{\infty} b_i < \frac{1}{2}\right\}.$$

If $\lim_{n\to\infty} M_n = \infty$ then

$$1 \le \liminf_{n\to\infty} \frac{\sum_{k=0}^{n} b_k M_k}{M_n} = \liminf_{n\to\infty} \frac{\sum_{k=n_0}^{n} b_k M_k}{M_n} \le \sum_{k=n_0}^{\infty} b_k \le \frac{1}{2},$$

which is an obvious contradiction. Therefore, $(a_n)$ is bounded. From this,

$$q_j \le \frac{M_j}{j+1} \le \frac{c}{j}.$$

Substituting this bound into (4.4) and applying Lemma 4.1 leads to the conclusion that $\mathrm{E}R_n$ remains bounded, as $n \to \infty$, by virtue of (4.3).

Recall that the convergence of $T_s$ in distribution always requires a scaling constant going to $\infty$ as $s \to \infty$. Under the conditions of Lemma 4.2, the sequence of laws of $\tau_{R_n}$ is tight. Now, from Proposition 3.1 and the decomposition (4.1), the following main result of this section emerges.

**Theorem 4.1.** *Suppose that (4.2) holds. For some constants $a_n > 0$ and $b_n \in \mathbb{R}$, if one of the variables $(T_{\log n} - b_n)/a_n$ and $(\tau_n - b_n)/a_n$ converges weakly, as $n \to \infty$, to a nondegenerate proper distribution then the other variable converges weakly to the same distribution.*

The value of this result lies in the fact that the limit laws for $T_s$ and the conditions of convergence are immediately translated into the convergence of $\tau_n$. Normalizing and centring constants are known explicitly; see [5] or Proposition 27 of [23]. It follows that only stable laws and the Mittag-Leffler laws can appear as the limit distributions of $\tau_n$.

If the measure $\nu$ is finite, condition (4.2) obviously holds. In this case $S$ is a compound Poisson process. Theorem 4.1 has been proved [13] under the assumptions that $\nu$ is not supported by a geometric sequence $(1 - x^k)_{k>0}$ (meaning that the law of $S_1$ is nonlattice) and that

$$\theta := \int_0^1 |\log x| \nu(\mathrm{d}x) < \infty. \tag{4.5}$$

Theorem 4.1 shows that the result of [13] is still true without requiring (4.5).

Assumption (4.2) is not very restrictive since $\Phi(k) = o(k)$ as $k \to \infty$ always holds. Concretely, suppose that the right tail of $\nu$ has the property of regular variation at 0, that is,

$$\vec{\nu}(x) \sim x^{-\gamma} \ell\left(\frac{1}{x}\right) \quad \text{as } x \downarrow 0 \tag{4.6}$$

for some function $\ell$ of slow variation at $\infty$, and $\gamma \in [0, 1]$. Then condition (4.2) is satisfied for

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

$\gamma \in [0, 1)$. In the edge case $\gamma = 1$ the behaviour of $\ell$ is important, for instance, (4.2) holds for $\ell(y) = (\log y)^{-\delta}$ if $\delta > 2$ and does not hold if $\delta \in (1, 2]$.

We use condition (4.2) to bound $R_n$, although we conjecture that (4.2) can be omitted and the equivalence in Theorem 4.1 holds in full generality for the coalescents with dust component. Note that (4.2) is the local property of $\vec{\nu}$ near 0. More substantially, the limit law is affected by the decay at $\infty$ of the right tail of the distribution of $S_1$, for which the behaviour of $\vec{\nu}$ near 1 is responsible. We illustrate this by two examples.

**Example 4.1.** (*Normal limits.*) Assume in addition to (4.2) that

$$\mathtt{s}^2 := \operatorname{var}(S_1) = \int_0^1 |\log(1-x)|^2 \nu(\mathrm{d}x) < \infty.$$

Then, as $n \to \infty$,

$$\frac{\tau_n - \mathtt{m}^{-1}\log n}{(\mathtt{m}^{-3}\mathtt{s}^2 \log n)^{1/2}} \xrightarrow{\mathrm{D}} \mathcal{N}(0, 1), \tag{4.7}$$

where $\mathtt{m} := \mathrm{E}\, S_1 = \int_0^1 |\log(1-x)|\nu(\mathrm{d}x)$.

This setting applies to the beta coalescents mentioned in the introduction. We choose the constant in (1.3) to be $c = 1/B(a, b)$, where $B(a, b)$ is the beta function. The case $a > 2$ was settled in [13]. We focus on the previously open case $1 < a \leq 2$.

For $a = 2$, we compute the constants as

$$\mathtt{m} = b(b+1)\zeta(2, b), \qquad \mathtt{s}^2 = 2b(b+1)\zeta(3, b),$$

where $\zeta$ is the Hurwitz zeta function.

For $a \in (1, 2)$, we have

$$\mathtt{m} = \frac{a+b-1}{(a-1)(2-a)}(1-(a+b-2)\{\Psi(a+b-1)-\Psi(b)\}),$$

$$\mathtt{s}^2 = \frac{a+b-1}{(a-1)(2-a)}(2\{\Psi(a+b-1)-\Psi(b)\}$$
$$-(a+b-2)\{(\Psi(a+b-1)-\Psi(b))^2 + \Psi'(b) - \Psi'(a+b-1)\}),$$

where $\Psi$ is the logarithmic derivative of the gamma function. Finally, condition (4.2) holds since (4.6) is satisfied with $\gamma = 2-a \in [0, 1)$ and a constant function $\ell$. Therefore, convergence (4.7) holds with the computed $\mathtt{m}$ and $\mathtt{s}$.

**Example 4.2.** (*Stable limits.*) Assume that (4.2) holds and that

$$\vec{\nu}(1-\mathrm{e}^{-y}) \sim y^{-\beta}L(y) \quad \text{as } y \to \infty \tag{4.8}$$

for some function $L$ slowly varying at $\infty$ and $\beta \in (1, 2)$. Then

$$\frac{\tau_n - \mathtt{m}^{-1}\log n}{\mathtt{m}^{-(\beta+1)/\beta} c_{\lfloor \log n \rfloor}} \xrightarrow{\mathrm{D}} \mathcal{S}(\beta) \quad \text{as } n \to \infty, \tag{4.9}$$

where $c_n$ is any sequence satisfying $\lim_{n\to\infty} nL(c_n)/c_n^\beta = 1$, and $\mathcal{S}(\beta)$ is the $\beta$-stable distribution with characteristic function

$$z \mapsto \exp\biggl(-|z|^\beta \Gamma(1-\beta)\biggl(\cos\biggl(\frac{\pi\beta}{2}\biggr) + i\sin\biggl(\frac{\pi\beta}{2}\biggr)\operatorname{sgn}(z)\biggr)\biggr), \qquad z \in \mathbb{R}.$$

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

To illustrate, consider

$$\nu(\mathrm{d}x) = \frac{x^{a-2}\,\mathrm{d}x}{(1-x)|\log(1-x)|^d},$$

where $d \in (2, 3)$ and $a \in (d, d+1)$. Then (1.2) is satisfied, and condition (4.6) holds with $\gamma = d+1-a \in (0, 1)$, which implies (4.2). Condition (4.8) is fulfilled with $\beta = d-1 \in (1, 2)$. Therefore, the absorption time $\tau_n$ of such a coalescent has limit law (4.9).

## 5. The number of collisions

### 5.1. Preliminaries

As an approximation to the number of collisions $X_n$, we shall consider $X_n^*$, the number of jumps of $N_n^*$ of size at least two. We will not be able to derive a complete classification of limit laws, comparable with the implication of Theorem 4.1, because the universal criterion for convergence of $X_n^*$ is not available. The cases when we know the behaviour of $X_n^*$ (from [1], [12], [15], and [16]) are all covered by the assumption that $\nu$ satisfies (4.6). We shall also proceed in this direction but exclude the case $\gamma = 1$ when $K_{n,1}$ is the term of dominating growth in the sum $K_n = \sum_{r=1}^n K_{n,r}$. By Karamata's Tauberian theorem [6], condition (4.6) with $\gamma < 1$ is equivalent to the analogous asymptotics of the Laplace exponent

$$\Phi(z) \sim \Gamma(1-\gamma) z^\gamma \ell(z) \quad \text{as } z \to \infty.$$

The case of finite $\nu$ appears when $\gamma = 0$ and $\Phi$ is an increasing bounded function.

The sequence $(X_n)$ is nondecreasing and satisfies a distributional recurrence

$$X_1 = 0, \qquad X_n \overset{\mathrm{D}}{=} \tilde{X}_{n-J_n+1} + 1(J_n \ge 2), \quad n \ge 2, \tag{5.1}$$

where on the right-hand side $J_n$ is independent of the $\tilde{X}_i$s, $\tilde{X}_i \overset{\mathrm{D}}{=} X_i$, and $J_n$ is distributed like the first decrement of $N_n^*$, that is, $\mathrm{P}(J_n = k) = p_{n,n-k}$ for $1 \le k \le n$. Similarly, the number $X_n^*$ of collisions which involve at least two primary clusters satisfies

$$X_1^* = 0, \qquad X_n^* \overset{\mathrm{D}}{=} \tilde{X}_{n-J_n}^* + 1(J_n \ge 2), \quad n \ge 2, \tag{5.2}$$

with the convention that $X_0^* = 0$. We may decompose $X_n$ as

$$X_n = X_n^* + D_n \overset{(2.1)}{=} K_n - K_{n,1} + D_n \quad \text{almost surely}, \tag{5.3}$$

where $D_n$ is the number of collisions which involve at most one primary cluster. Thus, a collision contributes to $D_n$ if either exactly one primary cluster merges with at least one secondary cluster, or at least two secondary and no primary clusters are merged.

**Lemma 5.1.** *We have*

$$\mathrm{E}D_n \le c \sum_{k=1}^n \left(\frac{\Phi(k)}{k}\right)^2, \qquad n \in \mathbb{N}. \tag{5.4}$$

*In particular, if either of the two equivalent conditions*

$$\int_0^1 x^{-2} \left(\int_0^x \vec{\nu}(y)\,\mathrm{d}y\right)^2 \mathrm{d}x < \infty \tag{5.5}$$

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

*or*

$$\sum_{k=1}^{\infty}\left(\frac{\Phi(k)}{k}\right)^2 < \infty \tag{5.6}$$

*holds, then the sequence of distributions of the $D_n$s is tight.*

*Proof.* The equivalence of (5.5) and (5.6) follows from [3, Proposition 1.4].

Choose some primary cluster $b$ to be definite, let it be the cluster labelled 1, and suppose that $\tilde{X}_{n-1}$ is realised as the number of collisions among $n-1$ primary clusters $[n]\setminus\{b\}$ and their followers. Then $X_n = \tilde{X}_{n-1} + Z_n$, where $Z_n$ is the indicator of the event that the first collision of $b$ involves exactly one other cluster $a$. At the time of the merging of $b$ with $a$ the Markov chain $N_n^*$ decrements by two or one, depending on whether $a$ is primary or secondary. Let $Y_n$ be the indicator of the event that the first involvement of $b$ either turns $b$ into a secondary cluster, or a collision involves at most one other primary cluster and an arbitrary number of secondary clusters. Clearly, $Y_n \geq Z_n$; therefore, from (5.1),

$$X_n \overset{\mathrm{D}}{\leq} \tilde{X}_{n-J_n} + Y_{n-J_n+1} + 1(J_n \geq 2), \tag{5.7}$$

where '$\overset{\mathrm{D}}{\leq}$' stands for 'stochastically smaller'. Passing to expectations in (5.2), (5.3), and (5.7) we see that, for $d_n := \mathrm{E}D_n$ and $y_n := \mathrm{E}Y_n$,

$$d_1 = 0, \qquad d_n \leq \sum_{k=1}^{n-1} p_{n,k}(d_k + y_{k+1}), \quad n = 2, 3, \dots,$$

and iterating this inequality yields

$$d_1 = 0, \qquad d_n \leq \sum_{j=1}^{n-1} g_{n,j}\, y_{j+1}, \quad n = 2, 3, \dots.$$

By exchangeability, we have $y_n = (\mathrm{E}\,K_{n,1} + 2\,\mathrm{E}\,K_{n,2})/n$. Since

$$\mathrm{E}\,K_{n,1} = \sum_{k=1}^{n} g_{n,k}\,p_{k,k-1} = \sum_{k=1}^{n} g_{n,k}\frac{k(\Phi(k)-\Phi(k-1))}{\Phi(k)},$$

using Lemma 4.1 with $r_k = k(\Phi(k)-\Phi(k-1))/\Phi(k)$ yields

$$\mathrm{E}\,K_{n,1} \leq c\Phi(n), \qquad n \in \mathbb{N}. \tag{5.8}$$

Using this, an inequality shown in Appendix A, and the monotonicity of $\Phi$,

$$\mathrm{E}\,K_{n,2} \overset{(A.5)}{\leq} c_1\,\mathrm{E}\,K_{\lceil n/2\rceil,1} \overset{(5.8)}{\leq} c_2\Phi\left(\left\lceil \frac{n}{2}\right\rceil\right) \leq c_2\Phi(n).$$

The variable $K_{n,r}$ can be identified with the number of blocks of size $r$ in the ordered partition of $[n]$ associated with $S$. The derivation of the bound on $\mathrm{E}\,K_{n,2}$ used a related interpretation of $K_{n,r}$ in terms of a random occupancy model, in which $n$ balls are thrown independently into infinitely many boxes. Conditionally given $S$, the probabilities of the boxes are equal to the jump sizes of the process $\exp(-S)$. In this model, $K_{n,r}$ appears as the number of boxes occupied by exactly $r$ out of $n$ balls.

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

It follows that

$$d_n \le c \sum_{k=1}^{n} \frac{g_{n,k}\Phi(k)}{k},$$

and using Lemma 4.1 with $r_k = c\Phi(k)/k$ yields (5.4).

### 5.2. The compound Poisson case

Assume that $\nu$ is a finite measure on (0, 1), not supported by a geometric sequence of the form $(1-x^k)_{k\ge 0}$ for some $x \in (0,1)$. Since a linear time change of the coalescent does not affect the distribution of $X_n$, we will not lose generality by assuming that $\nu$ is a probability measure on (0, 1). Let $(W_k)_{k\in\mathbb{N}}$ be independent copies of a random variable $W$ such that the law of $1-W$ is $\nu$. The subordinator $S$ is then a unit rate compound Poisson process with the generic jump $|\log W|$ having some nonlattice law.

The interpretation of $K_{n,r}$ in terms of the occupancy model involves random probabilities of boxes given by

$$P_k := W_1 \cdots W_{k-1}(1-W_k), \qquad k \in \mathbb{N}$$

(see, e.g. [12]). Introduce

$$\mathtt{m} := \int_0^1 |\log(1-x)|\nu(\mathrm{d}x),$$

and, for $1 \le r \le n$, let $\varkappa_{n,r} := \mathrm{E}\, K_{n,r}$.

**Proposition 5.1.** (a) *If* $\mathtt{m} < \infty$ *then, for every* $r = 1, 2, \ldots,$ *the vector* $(K_{n,1}, K_{n,2}, \ldots, K_{n,r})$ *converges weakly, as* $n \to \infty$, *to a proper multivariate distribution, and* $\varkappa_{n,r} \to (\mathtt{m}r)^{-1}$.

(b) *If* $\mathtt{m} = \infty$ *then* $\varkappa_{n,r} \to 0$, *so* $K_{n,r} \to 0$ *in probability.*

*Proof.* Part (a) was proved in [14, Theorem 3.3].

For (b), consider a random walk $(Q_j)_{j\ge 0}$ with $Q_0 = 0$ and the generic step $|\log W|$. Then $P_j = (1-W_j)\exp(-Q_{j-1})$. Using $1-x \le \mathrm{e}^{-x}$ with $x \in [0,1]$, and substituting $\mathrm{e}^z$ for $n$, we reduce estimating $\varkappa_{n,1} = n\sum_{j\ge 1} P_j(1-P_j)^{n-1}$ to estimating

$$\begin{aligned}\mathrm{E}\biggl(\sum_{j\ge 1} \mathrm{e}^z P_j \mathrm{e}^{-\mathrm{e}^z P_j}\biggr) &= \mathrm{E}\biggl(\sum_{j\ge 1}(1-W_j)\exp(z - Q_{j-1} - \mathrm{e}^{z-Q_{j-1}}(1-W_j))\biggr)\\ &= \int_0^\infty f(z-y)\,\mathrm{d}U(y),\end{aligned}$$

where $f(y) := \mathrm{E}((1-W)\exp(y - \mathrm{e}^y(1-W)))$ and $U(y) := \sum_{j\ge 0}\mathrm{P}\{Q_j \le y\}$ is the renewal function of the random walk. The function $f$ is nonnegative and integrable, since $\int_{-\infty}^{\infty} f(y)\,\mathrm{d}y = 1$. Furthermore, the function $y \to \mathrm{e}^{-y}f(y)$ is nonincreasing. It is known that these properties together ensure that $f$ is directly Riemann integrable (see, for instance, the proof of Corollary 2.17 of [8]). When $\mathtt{m} = \mathrm{E}\,|\log W_j| = \infty$, application of the key renewal theorem yields $\int_0^\infty f(z-y)U(\mathrm{d}y) \to 0$ as $z \to \infty$, whence $\varkappa_{n,1} \to 0$.

For $r > 1$, the argument is similar, or one can use the estimate $\varkappa_{n,r} \le c_r\varkappa_{n,1}$ given in Lemma A.2.

The next theorem improves upon a result from [13] by removing condition (4.5).

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

**Theorem 5.1.** *For constants* $a_n > 0$ *such that* $\lim_{n\to\infty} a_n = \infty$, *and* $b_n \in \mathbb{R}$, *whenever any of the variables*

$$\frac{K_n - b_n}{a_n}, \qquad \frac{X_n^* - b_n}{a_n}, \quad \textit{or} \quad \frac{X_n - b_n}{a_n}$$

*converges weakly, as* $n \to \infty$, *to a nondegenerate proper distribution, then all three variables converge weakly to this distribution.*

*Proof.* Recall representation (5.3). Since $\nu$ is a probability measure, we have $\Phi(k) < 1$; hence, condition (5.6) is satisfied, and the sequence of laws of the $D_n$s is tight by Lemma 5.1. By Proposition 5.1, the sequence of laws of the $K_{n,1}$s is tight as well. By the assumption that $a_n \to \infty$, the result follows.

From [12], it is known that, depending on the behaviour of $\vec{\nu}(x)$ near $x = 1$, there are five different modes of the weak convergence of, suitably normalized and centred, $K_n$. We do not exhibit all these cases here, rather provide an example borrowed from [12] to demonstrate a substantial role of the parameter $\theta = \int_0^1 |\log x|\nu(\mathrm{d}x)$.

**Example 5.1.** Suppose that $\nu$ has the right tail of the form

$$\vec{\nu}(x) = \frac{|\log x|^\rho}{1 + |\log x|^\rho}, \qquad x \in (0, 1],$$

with $\rho > 0$. In the case $\rho \in (0, \frac{1}{2})$ we have $\theta = \infty$, and

$$\frac{X_n - \mathtt{m}^{-1}\log n + (\mathtt{m}(1-\rho))^{-1}\log^{1-\rho} n}{c\log^{1/2} n} \xrightarrow{\mathrm{D}} \mathcal{N}(0, 1) \quad \text{as } n \to \infty,$$

where $\mathtt{m} = \int_0^1 |\log(1-x)|\nu(\mathrm{d}x)$.

In the other case, when $\rho > \frac{1}{2}$ (then $\theta < \infty$ for $\rho > 1$), the centring simplifies, so that

$$\frac{X_n - \mathtt{m}^{-1}\log n}{c\log^{1/2} n} \xrightarrow{\mathrm{D}} \mathcal{N}(0, 1) \quad \text{as } n \to \infty.$$

5.2.1. *Evolution of secondary particles.* In the compound Poisson case the number $V_t$ of secondary clusters of $\Pi_\infty(t)$ is finite for each $t \ge 0$. The process $V = (V_t)_{t\ge 0}$ starts with $V_0 = 0$ and is a Markov chain with transition rate $\varphi_{m,k} = \binom{m}{k}\lambda_{m,k}$ for jumping from $m$ to $m - k + 1$, $0 \le k \le m$, $k \ne 1$. The rate for $k = 0$ is given by the same formula (1.1), and $\varphi_{m,0} < \infty$ because $\nu$ is finite. The $k = 0$ transition, resulting in an increase in the number of secondary clusters by 1, occurs when some (in fact, infinitely many) primary clusters merge without engagement of secondary clusters. The time homogeneity of the transition rates of $V$ is a consequence of the existence of the dust component with infinitely many clusters.

It can be shown that the Markov chain $V$ is positive recurrent and has a unique stationary distribution $(\pi_m)$ found from the balance equation

$$\pi_m = \sum_{k=0}^{\infty} \pi_{m+k-1}\varphi_{m+k-1,k}, \tag{5.9}$$

supplemented by the conditions $\pi_0 = 0$ and $\sum_{m=1}^\infty \pi_m = 1$.

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

Suppose, for example, that $\nu(\mathrm{d}x) = \mathrm{d}x$ is the Lebesgue measure on $[0, 1]$. In this case $\varphi_{m,k} = (m+1)^{-1}$. Equation (5.9) becomes $\pi_m = \sum_{j=m-1}^{\infty} \pi_j/(j+1)$. Differencing yields $\pi_m - \pi_{m+1} = \pi_{m-1}/m$, which is readily solved as

$$\pi_m = \frac{\mathrm{e}^{-1}}{(m-1)!}, \qquad m = 1, 2, \dots,$$

so in this case the stationary distribution is shifted Poisson.

In contrast, the number of secondary clusters in the finite coalescent $\Pi_n$ is not a Markov process, because the transition rates depend on the number of remaining primary particles.

### 5.3. The case of slow variation

Suppose that (4.6) holds with $\gamma = 0$ and slowly varying $\ell(z) \to \infty$, $z \to \infty$. The Laplace exponent then satisfies $\Phi(z) \sim \ell(z)$. Suppose also that the subordinator has finite moments

$$\mathtt{m} = \mathrm{E}\, S_1 = \int_0^1 |\log(1-x)| \nu(\mathrm{d}x), \qquad \mathtt{s}^2 = \operatorname{var} S_1 = \int_0^1 |\log(1-x)|^2 \nu(\mathrm{d}x).$$

Choose the centring/scaling constants as

$$b_n = \frac{1}{\mathtt{m}} \int_0^n \frac{\Phi(z)}{z}\, \mathrm{d}z, \qquad a_n = \sqrt{\frac{\mathtt{s}^2}{\mathtt{m}^3} \int_0^n \frac{\Phi^2(z)}{z}\, \mathrm{d}z}.$$

In [1] it was shown that, for $n \to \infty$,

$$\mathrm{E}\, K_n \sim b_n, \qquad \sqrt{\operatorname{var} K_n} \sim a_n,$$

and that the normal limit $(K_n - b_n)/a_n \overset{\mathrm{D}}{\to} \mathcal{N}(0, 1)$ holds for various classes of functions $\ell$. In particular, this includes functions of slow variation at $\infty$ with asymptotics as diverse as

$$\ell(z) = \log(\log(\cdots(\log(z))\cdots)), \qquad \ell(z) = \log^{\beta} z, \qquad \ell(z) = \exp(\log^{\beta} z),$$

where $\beta > 0$.

Series (5.6) converges for arbitrary $\ell$; hence, by Lemma 5.1, $\mathrm{E}\, D_n = O(1)$. On the other hand, from (5.8) and by the properties of slowly varying functions [6],

$$\mathrm{E}\, K_{n,1} = O(\Phi(n)) = o(a_n).$$

It is immediate now that $(K_n - b_n)/a_n \overset{\mathrm{D}}{\to} \mathcal{N}(0, 1)$ implies that both $(X_n^* - b_n)/a_n \overset{\mathrm{D}}{\to} \mathcal{N}(0, 1)$ and $(X_n - b_n)/a_n \overset{\mathrm{D}}{\to} \mathcal{N}(0, 1)$.

**Example 5.2.** (*Gamma subordinators.*) Consider the classical gamma subordinator with Laplace exponent $\Phi(z) = \alpha \log(1 + z/\beta)$, where $\alpha, \beta > 0$. The corresponding $\nu$ driving the coalescent has density

$$\nu(\mathrm{d}x) = \frac{\alpha(1-x)^{\beta-1}}{|\log(1-x)|}\, \mathrm{d}x.$$

The central limit theorem for $K_n$ was proved by different methods in [1] and [16]. From this we conclude that the number of collisions also satisfies $(X_n - b_n)/a_n \overset{\mathrm{D}}{\to} \mathcal{N}(0, 1)$, where the constants can be chosen as

$$a_n = \sqrt{\frac{\beta \log^3 n}{3}}, \qquad b_n = \frac{\beta \log^2 n}{2}.$$

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

**Example 5.3.** (*beta*$(2,b)$*-coalescents.*) For this family, $\nu(\mathrm{d}x) = x^{-1}(1-x)^{b-1}\,\mathrm{d}x$. The convergence of $X_n$ to the standard normal distribution holds with scaling/centring constants

$$a_n = \sqrt{\frac{\mathtt{s}^2}{3\mathtt{m}^3}\log^3 n}, \qquad b_n = \frac{\log^2 n}{2\mathtt{m}},$$

where $\mathtt{m} = \zeta(2,b)$ and $\mathtt{s}^2 = 2\zeta(3,b)$.

### 5.4. Regular variation with index $0 < \gamma < 1$

A key distribution in this case is the law of the random variable

$$I = \int_0^\infty \exp(-\gamma S_t)\,\mathrm{d}t,$$

known as the exponential functional of the subordinator $\gamma S$. The distribution of $I$ is uniquely determined by the moments

$$\mathrm{E}\, I^k = \frac{k!}{\prod_{i=1}^k \Phi(\gamma i)}.$$

From [15, Theorem 4.1 and Corollary 5.2], $X_n^*/a_n \xrightarrow{\mathrm{D}} I$, where $a_n = \Gamma(2-\gamma)n^\gamma \ell(n)$, and no centring is required. In fact, $K_n/a_n$ and $K_{n,r}/a_n$ $(r \geq 1)$ converge almost surely and in the mean.

To justify the convergence of $X_n$ using (5.3), we need to etimate $\mathrm{E}\, D_n$. For $0 < \gamma < \frac{1}{2}$, we have $\mathrm{E}\, D_n = O(1)$ since $\Phi(z) \sim c\ell(z)z^\gamma$; hence, series (5.6) converges. For $\frac{1}{2} < \gamma < 1$, we have

$$\sum_{k=1}^n \left(\frac{\Phi(k)}{k}\right)^2 \sim cn^{2\gamma-1}\ell^2(n),$$

and, for $\gamma = \frac{1}{2}$, the latter sum, as a function of $n$, has the property of slow variation at $\infty$ (see [6, Proposition 1.5.8]). Thus, in any case $D_n/a_n \to 0$ in probability. It follows that $X_n/a_n \xrightarrow{\mathrm{D}} I$.

**Example 5.4.** (*beta*$(a,b)$*-coalescents with* $1 < a < 2$.) In this case

$$\frac{X_n}{n^{2-a}} \xrightarrow{\mathrm{D}} \frac{\Gamma(2-a)}{2-a}\int_0^\infty \exp(-(2-a)S_t)\,\mathrm{d}t \quad \text{as } n \to \infty,$$

where the Laplace exponent of $S$ is given by

$$\Phi(z) = \int_0^1 (1-(1-x)^z)x^{a-3}(1-x)^{b-1}\,\mathrm{d}x.$$

This result was obtained in [18] by another method, and, with a change of variable, the equivalence with Theorem 7.1 of [20] in the case $b = 1$ can be established.

The subfamily of beta-coalescents with parameters $b = 2-a$ has been intensively studied. In the literature, sometimes $\alpha := 2-a$ is taken as a parameter, so that $\nu$ in this notation becomes

$$\nu(\mathrm{d}x) = x^{-\alpha-1}(1-x)^{\alpha-1}.$$

In this case $N_n^*$ decrements like a random walk conditioned to hit 0 and, moreover, there is an explicit formula (see [10, p. 471]), i.e.

$$g_{n,k} = \frac{(\alpha)_k(\alpha)_{n-k}}{(\alpha)_n}\binom{n}{k},$$

where $(\alpha)_k$ denotes the rising factorial. The variable $K_n$ is then the number of blocks in

Pitman's $(\alpha, \alpha)$-partition (or in the regenerative composition induced by excursions of a Bessel bridge [10]). We refer the reader to [2] and [25] for further multiple connections of these beta-coalescents to various random processes.

# Appendix A

## A.1. A linear recursion

For each $n \in \mathbb{N}$, let $(p_{n,k})_{0\le k\le n}$ be a probability distribution with $p_{n,n} < 1$. Define a sequence $(a_n)_{n\in\mathbb{N}}$ as a (unique) solution to the recursion

$$a_n = r_n + \sum_{k=0}^{n} p_{n,k}a_k, \qquad n \in \mathbb{N}, \tag{A.1}$$

with given $r_n \ge 0$ and the initial value $a_0 = a \ge 0$.

**Lemma A.1.** *Suppose that there exists a sequence* $(\psi_n)_{n\in\mathbb{N}}$ *such that*

(i) $\liminf_{n\to\infty} \psi_n \sum_{k=0}^{n}(1 - k/n)p_{n,k} > 0$,

(ii) *the sequence* $(\psi_k r_k/k)_{k\in\mathbb{N}}$ *is nonincreasing.*

*Then* $(a_n)$ *defined by (A.1) satisfies*

$$a_n = O\biggl(\sum_{k=1}^{n}\frac{r_k\psi_k}{k}\biggr) \quad \textit{as } n \to \infty. \tag{A.2}$$

*In particular,* $(a_n)$ *is bounded if the series* $\sum_{k=1}^{\infty} r_k\psi_k/k$ *converges.*

*Proof.* Write, for simplicity, $p_k$ for $p_{n,k}$, and let $\pi_k = \sum_{j=0}^{k} p_j$. Using (ii), we have

$$\sum_{k=1}^{n}\frac{r_k\psi_k}{k}\pi_{k-1} \ge \frac{r_n\psi_n}{n}\sum_{k=1}^{n}\pi_{k-1} = r_n\psi_n\sum_{j=0}^{n-1}\biggl(1-\frac{j}{n}\biggr)p_j.$$

By (i), there exist $n_0 \in \mathbb{N}$ and $c > 0$ such that

$$c\sum_{k=1}^{n}\frac{r_k\psi_k}{k}\pi_{k-1} \ge r_n, \qquad n \ge n_0. \tag{A.3}$$

We claim that $x_n := c\sum_{k=1}^{n} r_k\psi_k/k$ satisfies

$$x_n \ge r_n + \sum_{k=1}^{n} x_k p_k, \qquad n \ge n_0. \tag{A.4}$$

To check the latter, write

$$\begin{aligned} r_n + \sum_{k=1}^{n} x_k p_k &= r_n + c\sum_{j=1}^{n}\sum_{k=j}^{n}\frac{r_j\psi_j}{j}p_k \\ &= r_n + c\sum_{j=1}^{n}\frac{r_j\psi_j}{j}(1-\pi_{j-1}) \end{aligned}$$

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

$$
\begin{aligned}
&= r_n + c\sum_{j=1}^{n}\frac{r_j\psi_j}{j} - c\sum_{j=1}^{n}\frac{r_j\psi_j}{j}\pi_{j-1}\\
&= x_n + r_n - c\sum_{j=1}^{n}\frac{r_j\psi_j}{j}\pi_{j-1}\\
&\overset{(A.3)}{\le} x_n.
\end{aligned}
$$

Set $x_0 := 0$. Subtracting (A.1) from (A.4) we see that $y_n := x_n + c_0 - a_n$ satisfies $y_n \ge \sum_{k=0}^{n} p_k y_k$ for $n \ge n_0$ and arbitrary $c_0$. By choosing $c_0 \ge \max_{n\le n_0} a_n$, it is easily shown by induction that $y_n \ge 0$ for all $n \in \mathbb{N}$, which implies the desired estimate of $a_n$.

### A.2. Estimates for the occupancy counts

Let $(p_k)_{k\in\mathbb{N}}$ be a probability mass function. Consider the multinomial occupancy scheme in which $n$ balls are thrown independently into boxes, with probability $p_j$ for box $j$. The expected number of boxes occupied by exactly $r$ out of $n$ balls is

$$\varkappa_{n,r} = \binom{n}{r}\sum_{j\ge1} p_j^r(1-p_j)^{n-r}, \qquad 1 \le r \le n.$$

**Lemma A.2.** *For fixed $r < s$, there exists a constant $c$ such that*

$$\varkappa_{n,r} \ge c\varkappa_{2n,s}, \qquad n \in \mathbb{N}. \tag{A.5}$$

*Proof.* Using $(1-x)^{-1} \ge \mathrm{e}^x$ for $x \in (0,1)$,

$$
\begin{aligned}
\frac{\binom{n}{r}x^r(1-x)^{n-r}}{\binom{2n}{s}x^s(1-x)^{2n-s}} &\ge c_1\frac{s!}{2^s r!}(nx)^{r-s}(1-x)^{s-r-n}\\
&\ge c_2(nx)^{r-s}\mathrm{e}^{nx/2}\\
&\ge c_2\min_{y>0} y^{r-s}\mathrm{e}^{y/2}\\
&= c_2\biggl(\frac{\mathrm{e}}{2(s-r)}\biggr)^{s-r}.
\end{aligned}
$$

The result extends immediately to the case of random $(P_k)$. This generalization was used in the proof of Proposition 5.1 with the $P_j$s being the sizes of intervals obtained by splitting $[0, 1]$ at points of the range of the process $\exp(-S)$.

## Acknowledgements

The authors are indebted to an anonymous referee for a number of helpful comments. A. Iksanov gratefully acknowledges the support of a grant from Utrecht University.

## References

[1] Barbour, A. D. and Gnedin, A. V. (2006). Regenerative compositions in the case of slow variation. *Stoch. Process. Appl.* **116**, 1012–1047.

[2] Berestycki, N. (2009). *Recent Progress in Coalescent Theory* (Ensaios Matemáticos **16**). Sociedade Brasileira de Matemática, Rio de Janeiro.

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press

[3] Bertoin, J. (1996). *Subordinators: Examples and Applications* (Lecture Notes Math. **1727**). Springer, Berlin.

[4] Bertoin, J. (2010). Exchangeable coalescents. Lecture Notes, ETH Zürich. Available at http://www.fim.math.ethz.ch/lectures/Lectures_Bertoin.pdf.

[5] Bingham, N. H. (1972). Limit theorems for regenerative phenomena, recurrent events and renewal theory. *Z. Wahrscheinlichkeitsth.* **21,** 20–44.

[6] Bingham N. H., Goldie C. M. and Teugels, J. L. (1989). *Regular Variation*. Cambridge University Press.

[7] Drmota, M., Iksanov, A., Moehle, M. and Roesler, U. (2009). A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures Algorithms* **34,** 319–336.

[8] Durrett, R. and Liggett, T. M. (1983). Fixed points of the smoothing transformation. *Z. Wahrscheinlichkeitsth.* **64,** 275–301.

[9] Freund, F. and Möhle, M. (2009). On the time back to the most recent common ancestor and the external branch length of the Bolthausen–Sznitman coalescent. *Markov Process. Relat. Fields* **15,** 387–416.

[10] Gnedin, A. and Pitman, J. (2005). Regenerative composition structures. *Ann. Prob.* **33,** 445–479.

[11] Gnedin, A. and Yakubovich, Y. (2007). On the number of collisions in Λ-coalescents. *Electron. J. Prob.* **12,** 1547–1567.

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

[13] Gnedin, A., Iksanov, A. and Möhle, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Prob.* **45,** 1186–1195.

[14] Gnedin, A., Iksanov, A. and Roesler, U. (2008). Small parts in the Bernoulli sieve. *Discrete Math. Theoret. Comput. Sci.* **AI,** 235–242.

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

[16] Gnedin, A., Pitman, J. and Yor, M. (2006). Asymptotic laws for regenerative compositions: gamma subordinators and the like. *Prob. Theory Relat. Fields* **135,** 576–602.

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

[18] Haas, B. and Miermont, G. (2011). Self-similar scaling limits of non-increasing Markov chains. To appear in *Bernoulli*.

[19] Iksanov, A. and Möhle, M. (2007). A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Commun. Prob.* **12,** 28–35.

[20] Iksanov, A. and Möhle, M. (2008). On the number of jumps of random walks with a barrier. *Adv. Appl. Prob.* **40,** 206–228.

[21] Iksanov, A., Marynych, A. and Möhle, M. (2009). On the number of collisions in beta(2, $b$)-coalescents. *Bernoulli* **15,** 829–845.

[22] Möhle, M. (2010). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent. *Stoch. Process. Appl.* **120,** 2159–2173

[23] Negadailov, P. (2010). Limit theorems for random recurrences and renewal-type processes. Doctoral Thesis, Utrecht University. Available at http://igitur-archive.library.uu.nl/dissertations/.

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

[25] Pitman, J. (2006). *Combinatorial Stochastic Processes* (Lecture Notes Math. **1875**). Springer, Berlin.

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

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

https://doi.org/10.1239/jap/1324046023 Published online by Cambridge University Press