J. Appl. Prob. Spec. Vol. **51A,** 23–40 (2014)

CELEBRATING 50 YEARS OF THE APPLIED PROBABILITY TRUST

Edited by
S. ASMUSSEN, P. JAGERS, I. MOLCHANOV and L. C. G. ROGERS

Part 2. The 2014 AP lectures

Λ-COALESCENTS: A SURVEY

ALEXANDER GNEDIN, *Queen Mary University of London*
School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK. Email address: a.gnedin@qmul.ac.uk

ALEXANDER IKSANOV, *National Taras Shevchenko University of Kyiv*
Faculty of Cybernetics, National Taras Shevchenko University of Kiev, 01033 Kiev, Ukraine.
Email address: iksan@univ.kiev.ua

ALEXANDER MARYNYCH, *National Taras Shevchenko University of Kyiv*
Faculty of Cybernetics, National Taras Shevchenko University of Kiev, 01033 Kiev, Ukraine.
Email address: marynych@unicyb.kiev.ua

APPLIED PROBABILITY TRUST
DECEMBER 2014

Λ-COALESCENTS: A SURVEY

BY ALEXANDER GNEDIN, ALEXANDER IKSANOV AND ALEXANDER MARYNYCH

Abstract

Λ-coalescents model the evolution of a coalescing system in which any number of components randomly sampled from the whole may merge into larger blocks. This survey focuses on related combinatorial constructions and the large-sample behaviour of the functionals which characterize in some way the speed of coalescence.

Keywords: Coalescent with multiple mergers; exchangeable partitions; large-sample asymptotics

2010 Mathematics Subject Classification: Primary 60J25; 60J27
Secondary 60C05; 60G09; 60G51

1. Introduction

Aldous in his 1999 survey [3] observed that the stochastic models of coalescence (clustering, coagulation, aggregation, gelation), common in many scientific disciplines, had been of only tangential concern in the applied probability literature. That same year Pitman [49] and Sagitov [51] introduced a class of processes, called Λ-coalescents, which have become increasingly popular in the probability community and largely stimulated the research in the area of exchangeable partition-valued processes of coalescence and fragmentation.

The mathematical theory of coalescent processes originated in Kingman's work [40, 41]. The idea comes from the study of genealogical relationships in biology: given a large population of haploid organisms evolving over many generations, a sample of individuals from the current generation is taken and their family history is traced backward in time. The ancestral lineages coalesce at times when two or more individuals in the sample have a common ancestor. In Kingman's coalescent every pair of lineages coalesces at unit probability rate, and every merger is binary.

The same type of process may also be interpreted forward in time as describing the evolution of a system of components (particles, polymers, dust formations, political coalitions, etc.) which accumulate in larger and larger blocks as time passes. There are more complex models which depart from Kingman's process in that they allow the coalescence rate to depend in some way on the masses of the two merging parts [3].

A Λ-coalescent is another kind of generalisation, in which arbitrary multiple mergers are possible. The transition rates depend only on the number of blocks involved in a merger; thus, the exchangeability property inherent in Kingman's coalescent is preserved. Bolthausen and Sznitman [15] studied a remarkable process of this kind in the context of spin glasses. Λ-coalescents and more general exchangeable coalescents with simultaneous multiple mergers [52] appear as limiting forms of genealogy models for populations with large offspring sizes [34, 46, 54].

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

The family exhibits a rich variety of behaviours, especially when seen from the perspective of infinite populations. At one extreme are the coalescents, like Kingman's, in which an infinite myriad of massless dust particles coagulate almost instantaneously into finitely many massive blocks, although a typical merger takes just a few parts. At another edge of the spectrum are the processes where the primary dust persists forever while the mergers, regularly spaced in time, take a good chunk of the dust together with some other massive blocks. Intermediate regimes add colour to the picture, which is further refined by phase transitions at the level of fluctuation theory.

We refer the reader to lecture notes [10, 11] for a graduate-plus level introduction to the theory of Λ-coalescents, their connections to other stochastic models, and many pointers to the literature. The present survey gives a snapshot of results on large-sample asymptotics for functionals that characterize in some way the speed of coalescence. Inevitably, the focus is biased towards the authors' interests in renewal approximations, random recursions, and regenerative combinatorial structures.

2. Fundamentals

2.1. Poisson construction and exchangeability

A Pitman–Sagitov process derives its name, Λ-coalescent, from the infinite-dimensional parameter Λ, which is a positive finite measure on $[0, 1]$. For Kingman's coalescent, Λ is the Dirac mass at 0, and, for the Bolthausen–Sznitman coalescent, Λ is the Lebesgue measure. The dynamics of Λ-coalescents follow the rule: if at some time the process is restricted to a partition with m separate blocks then a transition that involves k particular blocks merging into one block occurs at the rate

$$\lambda_{m,k} = \int_0^1 x^{k-2}(1-x)^{m-k}\,\Lambda(\mathrm{d}x), \qquad 2 \le k \le m. \tag{1}$$

The total merging rate on m blocks is therefore

$$\lambda_m = \sum_{k=2}^{m} \binom{m}{k} \lambda_{m,k} = \int_0^1 [1-(1-x)^m - mx(1-x)^{m-1}]x^{-2}\,\Lambda(\mathrm{d}x).$$

Note that a mass at 0 contributes $\binom{m}{2}\Lambda(\{0\})$ to the cumulative rate of binary mergers.

That these formulae involve a mixture of binomial probabilities is not accidental. Suppose in the first instance that Λ has no mass at 0, and consider the measure

$$\nu(\mathrm{d}x) = x^{-2}\Lambda(\mathrm{d}x), \qquad x \in (0, 1],$$

which is easier to interpret and in some respects more convenient than Λ. Think of a coin with probability of heads x randomly chosen from ν. For every block, toss the coin and if the outcome is heads, include the block in the merger. This naïve picture of transition is made rigorous by considering a random Poisson scatter of points in the strip $[0, 1] \times [0, \infty)$ with intensity measure $\nu(\mathrm{d}x) \times \mathrm{d}t$. The coalescent emerges as the points are scanned in the order of their time coordinates. Given a point at the generic location (x, t), each block of the partition, as defined immediately before time t, joins the merger with probability x, independently of other coexisting blocks. If the measure ν is infinite, the set of times of the Poisson scatter will be dense. However, the coalescence will not occur instantaneously, because only times t when at least two coin tosses land up heads yield a transition, and these appear at a finite rate whenever the number of blocks is finite. If Λ has mass at 0, this Poisson construction needs to be modified to superimpose a binary merging rate.

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

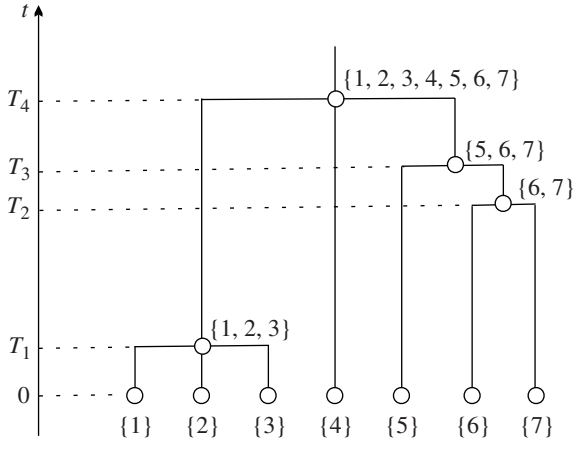

FIGURE 1: A sample path of Π_7 with four mergers.

Any mass at 1 forces a total collapse into one block at a random exponential time. We exclude this possibility throughout the paper.

Up to this point we have intentionally avoided the ritual words 'the process takes values in ...', because the rule is the same for various choices of the state space for Λ-coalescents. The Λ-coalescent on n particles is formally defined as a continuous-time Markov chain $\Pi_n = (\Pi_n(t),\ t \geq 0)$ that takes values in the finite space of partitions of the set $[n] = \{1, 2, \ldots, n\}$, where by 'partition of a set' we mean a representation of the whole as the union of disjoint nonempty subsets (blocks). The process starts with the finest partition of $[n]$ in singleton blocks $\{1\}, \{2\}, \ldots, \{n\}$, to be thought of as primary particles, and moves to coarser partitions until eventually terminating at the trivial partition with one block $[n]$. The process Π_n is exchangeable, i.e. permutations of $[n]$ do not alter the distribution of Π_n. Associated with Π_n is the *arithmetic coalescent* defined on the integer partitions of n and moving from the partition $1 + \cdots + 1$ to n; this less informative process uniquely determines the law of Π_n by exchangeability.

It is customary to graph a sample path of Π_n as a tree with vertical edge lengths representing time elapsed between mergers (see Figure 1). In the genealogical interpretation, the leaves correspond to n individuals from the current generation, and a block of $\Pi_n(t)$ includes the individuals with a common ancestor living at time t backward from the present time. With this in mind, $\Pi_n(t)$ is called the *ancestral partition*.

Given a partition of a set into blocks, we obtain a partition of a subset by removing some elements from the blocks and deleting any blocks that have become empty. It is implicit in the rule of a Λ-coalescent, and obvious from the Poisson construction, that Π_n restricted from $[n]$ to any subset with $m < n$ elements has the same distribution as Π_m, up to relabelling elements of the subset by $[m]$. That a Markov chain projects to a Markov chain is, in fact, a highly nontrivial property which in algebraic terms amounts to the backward-in-n recursion

$$\lambda_{n-1,k} = \lambda_{n,k} + \lambda_{n,k+1}. \tag{2}$$

This recursion is a form of Hausdorff's moment problem, which ensures that every nonnegative solution has a unique integral representation (1) with some finite measure Λ on [0, 1].

By consistency, there is an exchangeable process $\Pi = (\Pi(t),\ t \geq 0)$, called the *infinite Λ-coalescent*, whose restriction for each integer n to $[n]$ is Π_n. The state space for Π is the

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

set of partitions of $\mathbb{N}$, and the initial state is the partition into singleton blocks $\{1\}, \{2\}, \ldots$ corresponding to an infinite set of primary particles labelled by $\mathbb{N}$.

For fixed t, we can view $\{\Pi_n(t) : n \in \mathbb{N}\}$ as a growth process, with $\Pi(t)$ emerging as a limit. One move amounts to either inserting $n+1$ into one of the blocks of $\Pi_n(t)$ or appending $\{n+1\}$ to the set of blocks as a singleton. For instance, given that $\Pi_n(t)$ is the singleton partition of $[n]$, $\Pi_{n+1}(t)$ is the singleton partition of $[n+1]$ with probability $\mathrm{e}^{-(\lambda_{n+1}-\lambda_n)t}$. Unfortunately, there are no simple formulae for more general transition probabilities, nor for the distribution of $\Pi_n(t)$ (see [49, Section 3.8] for the algebra involved). The only known exception is the Bolthausen–Sznitman coalescent, where $\Pi_n(t)$ is a partition from the Ewens–Pitman family with parameters $(\mathrm{e}^{-t}, 0)$ and the growth rule is known under the fancy name Chinese restaurant process [50]. There is also a simple formula for the partition resulting from the kth merger in Kingman's coalescent (see [9] for a generalization).

2.2. Basic classification

The Poisson construction suggests that every block of the ancestral partition $\Pi(t)$ is either a singleton primary particle or infinite with positive frequency, equal to the limit proportion of representatives of the block in $[n]$ as $n \to \infty$. The dichotomy is a general consequence of exchangeability and de Finetti's theorem. Similarly, the collection of particles in singleton blocks of $\Pi(t)$ is either empty or has positive cumulative frequency. We shall say that a block with positive frequency is *massive*, and call the collection of singleton blocks of $\Pi(t)$ *the dust*. For the general exchangeable partition of $\mathbb{N}$, there are four types of realizations which may occur with nonzero probability:

(I) finitely many massive blocks and dust;

(II) infinitely many massive blocks and dust;

(III) infinitely many (massive) blocks and no dust;

(IV) finitely many (massive) blocks.

As follows from [49], the type of the ancestral partition $\Pi(t)$ depends neither on chance nor on $t > 0$. This gives the most basic structural classification of Λ-coalescents.

The type is determined by the concentration of the measure Λ near 0. Note that multiplying Λ by a positive constant amounts to a linear time change in the coalescent process, and, thus, does not affect the type. The first two types are easy to identify in terms of the moments

$$\mathtt{m}_r := \int_0^1 x^r \, \Lambda(\mathrm{d}x),$$

which may be finite or infinite for negative r.

Case (I) appears if and only if $\mathtt{m}_{-2} < \infty$. In this case ν is a finite measure, the transitions of Π occur at discrete times of a Poisson process with rate $\mathtt{m}_{-2}$, and a merger takes each block with probability x, with x chosen from the distribution $\nu(\mathrm{d}x)/\mathtt{m}_{-2}$.

Case (II) appears if and only if $\mathtt{m}_{-2} = \infty$ and $\mathtt{m}_{-1} < \infty$. The transitions in Π occur at a dense set of jump times of a subordinator (increasing Lévy process) with an infinite Lévy measure.

If $\mathtt{m}_{-1} = \infty$, there is no dust, but a finer condition is required to distinguish between (III) and (IV). In case (IV) the coalescent is said to come down from infinity, meaning that the initial singleton partition needs arbitrarily small time to coalesce into a partition of finitely many blocks. Equivalently, Π has finite absorption time $\tau = \inf\{t : \Pi(t) = \{\mathbb{N}\}\}$. It turns out that

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

$\tau < \infty$ almost surely if and only if $\mathbb{E}[\tau] < \infty$, which underlies a criterion for (IV) discovered by Schweinsberg [53], namely,

$$\sum_{m=2}^{\infty} \frac{1}{\gamma_m} < \infty, \tag{3}$$

where

$$\gamma_m = \sum_{k=2}^{m} (k-1) \binom{m}{k} \lambda_{m,k} = \int_0^1 [(1-x)^m - 1 + mx]\, x^{-2}\, \Lambda(\mathrm{d}x). \tag{4}$$

The intuitive link between (3) and $\mathbb{E}[\tau] < \infty$ is evident on interpreting γ_m as a rate at which the number of blocks decreases down from m, hence viewing the reciprocal $1/\gamma_m$ as a kind of mean holding time at m blocks.

We note in passing that in other areas involving the application of exchangeable partitions of $\mathbb{N}$, notably in species sampling problems and Bayesian nonparametrics, partitions with dust or finitely many blocks are considered to be a nuisance and the main focus of researchers is on partitions of type (III). In contrast to that, for Λ-coalescents, a nice ancestral partition of type (III) is more of an exception.

2.3. Coalescents on mass partitions

There is a parallel view of an infinite Λ-coalescent as a process in the space of mass partitions $\Delta = \{\boldsymbol{s} = (s_1, s_2, \dots) \colon s_1 \geq s_2 \geq \cdots \geq 0,\ \sum_j s_j \leq 1\}$. The process appears as a scaled limit form of the arithmetic coalescent. To obtain this, the decreasing sequence of frequencies of massive blocks of $\Pi(t)$ is viewed as a random element of Δ, with the conventions that $s_j = 0$ if there are fewer than j massive blocks, and that $1 - \sum_j s_j$ is the cumulative frequency of dust. When a transition in Π occurs, a subset of the frequencies is removed and replaced by their sum plus a fraction of the dust frequency, and then the resulting terms are notated in decreasing order.

In the other direction the connection relies on a version of Kingman's correspondence between random mass partitions and exchangeable partitions of $\mathbb{N}$ (see, e.g. [10, 11, 50]); this entails the following law of large numbers. For fixed $\boldsymbol{s} \in \Delta$, let $\mu_{\boldsymbol{s}}$ be a probability measure on the reals (say) with positive masses s_j at some locations and a diffuse component of total mass $1 - \sum_j s_j$. Let $\xi_1, \xi_2, \dots$ be independent random variables sampled from the distribution $\mu_{\boldsymbol{s}}$. Define a partition of $\mathbb{N}$ via the classes of the equivalence relation $k \sim \ell$ if and only if $\xi_k = \xi_\ell$; this partition has frequencies $\boldsymbol{s}$.

For the Bolthausen–Sznitman coalescent, the frequencies of $\Pi(t)$ have the Poisson–Dirichlet distribution, notated PD(e^{-t}, 0), over the facet of the simplex with $\sum_j s_j = 1$ (see [49, 50]). This is the only example of a Λ-coalescent for which the law of frequencies is known explicitly.

2.4. Subfamilies

Like every good infinite-dimensional theory, the theory of Λ-coalescents has a numerical parameter a. We prefer this notation, but use it in parallel with the more widely accepted α that is related to a via $a + \alpha = 2$. The parameter controls the smoothness of Λ near 0 and is reflected in asymptotic power laws.

A *beta coalescent* is a Λ-coalescent with beta density on [0, 1], namely,

$$\Lambda(\mathrm{d}x) = A\, x^{a-1} (1-x)^{b-1}\, \mathrm{d}x,$$

where $A, a, b > 0$. The transition rates in this case are computable in terms of Euler's beta function as $\lambda_{m,k} = A\, \mathrm{B}(a + k - 2, b + m - k)$. Usually, the normalization $A = 1/\mathrm{B}(a, b)$

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

is taken to make Λ a probability measure. The beta coalescents belong to type (I) for $a > 2$, type (II) for $1 < a \leq 2$, type (III) for $a = 1$, and type (IV) for $0 < a < 1$. This confirms the view of type (III) as a bridge between coalescents with dust and coalescents coming down from infinity. The Bolthausen–Sznitman coalescent is beta with $a = b = 1$, and Kingman's coalescent arises as the limiting case $a \to 0$.

To some authors, a beta coalescent is understood to be a member of the one-parameter family $(2-\alpha, \alpha)$, for which we reserve the term *symmetric beta*. For $1 \leq \alpha < 2$ (so $0 < a \leq 1$), this beta coalescent has a tractable time reversal (see the paper [14] resulting from a merger of three groups of authors). In this range the beta coalescent is also important as a large-sample limit of the genealogy of those supercritical branching processes for which the offspring distribution has a power law decay $ck^{-\alpha}$ (see [54]).

A step away from beta coalescents are Λ-coalescents with Λ satisfying some version of the regularity condition

$$\Lambda([0, x]) \sim A\, x^a \quad \text{as } x \to 0, \tag{5}$$

where $A > 0$ and the principal range for the shape parameter is $0 < a < 2$. Variations found in the literature may involve adding an estimated error term, or a counterpart condition on the density, or the counterpart regularity condition on the tail

$$\nu([x, 1]) \sim A_1\, x^{a-2} \quad \text{as } x \to 0. \tag{6}$$

Further modifications may also involve a slowly varying factor, which becomes crucial in the critical cases $a = 0, 1, 2$. For $a > 2$, the measure ν is finite and (5) has a minor effect on properties of the coalescent.

3. Related processes and their jump chains

3.1. The block counting process

The counting process $N_n = (N_n(t),\ t \geq 0)$, where $N_n(t)$ is the number of blocks in $\Pi_n(t)$, is itself a Markov process with transition rate $\binom{n}{k}\lambda_{n,k}$ for jumping from n to $n-k+1$, $2 \leq k \leq n$. The coalescent Π_n can be recovered (in distribution) from the path of N_n by iterative use of sampling without replacement. If at time t the process N_n decrements by $k-1$, a merge is constructed by sampling uniformly without replacement k blocks out of $N_n(t-)$ blocks of $\Pi_n(t-)$ and merging them into one block to create $\Pi_n(t)$.

The jump chain has transition probability

$$q(n, k) = \binom{n}{k} \frac{\lambda_{n,k}}{\lambda_n}, \qquad 2 \leq k \leq n, \tag{7}$$

for the move $n \to n-k+1$. A counterpart of (2) is a nonlinear recursion which allows one to calculate $q(n', \cdot)$ from $q(n, \cdot)$ for $n' < n$. Moreover, each of the three objects uniquely determines the other two: the stochastic matrix $q(\cdot, \cdot)$, the sequence of rates $(\lambda_n,\ n \geq 1)$ normalised by $\lambda_2 = 1$, and the probability measure Λ on $[0, 1]$.

3.2. Freezing and the allelic partition

Suppose that in addition to coalescence there is another kind of transition which takes an *active* block and transforms it into a *frozen* block at a given rate $\rho > 0$. The state of the process with this additional feature is a partition into disjoint blocks each of which is in either the active or frozen condition. Frozen blocks neither turn into active nor engage in mergers with other active or frozen blocks. Assuming that at time 0 all primary particles are active, the process

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

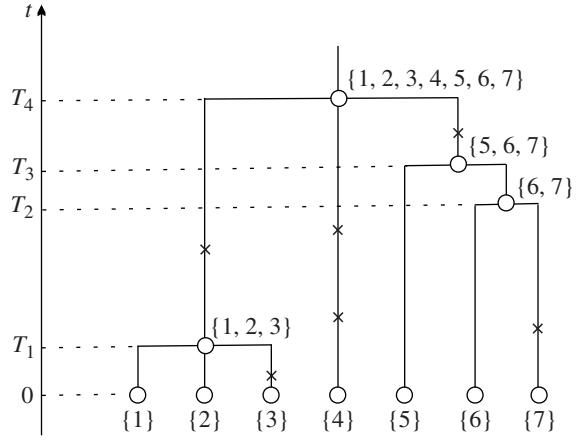

FIGURE 2: A path of coalescent with mutations.

starting with n particles will eventually terminate with only frozen blocks which make up a so-called *allelic partition* of $[n]$. The allelic partitions are exchangeable and consistent under restrictions, and, hence, define a partition of $\mathbb{N}$.

The name 'allelic partition' comes from a biological interpretation. In applications of coalescent theory to genetics, mutations are modelled as sites of a Poisson point process with rate ρ along the coalescent tree. Under the paradigm of the infinite alleles model, every mutation yields a different allelic type; hence, two individuals in the present generation belong to the same allelic type if there is no mutation along their family lines before the individuals coalesce. This is illustrated in Figure 2 where the allelic partition is comprised of blocks $\{1, 2\}, \{3\}, \{4\}, \{5, 6\}, \{7\}$.

The distribution of the allelic partition on $[n]$ is known explicitly only for Kingman's coalescent, where the partition follows the Ewens sampling formula with parameter 2ρ. In general, the distribution of the allelic partition satisfies a recursion due to Möhle [44]. Unfortunately, the recursion is complicated for direct analysis, but can be implemented as the following Markovian algorithm on partitions of $[n]$. Let $\lambda'_n = \lambda_n + n\rho$, and define the transition matrix

$$q'(n,k) = \begin{cases} \binom{n}{k} \dfrac{\lambda_{n,k}}{\lambda'_m} & \text{if } 2 \le k \le n, \\ \dfrac{n\rho}{\lambda'_m} & \text{if } k = 1, \end{cases}$$

for the chain counting the number of active blocks. For $k > 1$, $q'(n,k)$ is the probability that the number of active blocks decreases from n to $n-k+1$ as the result of a k-merger, while $q'(n,1)$ is the probability that the number of active blocks decreases from n to $n-1$ as the result of a freeze. For instance, for Kingman's coalescent, $q(n,1) = 1 - q(n,2) = 2\rho/(n-1+2\rho)$. Given a partition of $[n]$, draw k from the distribution $q(n,\cdot)$. If $k = 1$, remove an element picked uniformly from $[n]$ from its block and append it as a singleton block to the partition. If $k > 1$, choose k elements one-by-one without replacement uniformly from $[n]$, remove the first $k-1$ of them from their blocks, and add them to the block containing the kth element. This defines a recurrent Markov chain on partitions of $[n]$, such that the limit distribution is the allelic partition for the coalescent with freeze (see [17]).

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

3.3. The dust decay subprocess

For coalescents with dust, it is possible to extend definition (1) to the case $k = 1$. It is the condition $\mathrm{m}_{-1} < \infty$ that makes $\lambda_{m,1} < \infty$. Then $\lambda_{m,1}$ is the rate at which a particular block in a system of m blocks merges with some of the infinitely many blocks outside the system. It is convenient to treat such a merger as a unary collision in the context of a subprocess which traces the decay of the dust component. The state of this process at time $t \geq 0$ is the collection of primary particles of $\Pi(t)$. Whenever the process is restricted to m primary particles, any k-tuple of them is taken by a merger in Π at rate $\lambda_{m,k}$, $1 \leq k \leq m$.

The jump chain counting the number of primary particles moves from n to $n - k$ with probability

$$q''(n,k) = \binom{n}{k} \frac{\lambda_{n,k}}{\lambda''_n}, \qquad 1 \leq k \leq n,$$

where $\lambda''_n = \sum_{k=1}^{n} \binom{n}{k} \lambda_{n,k}$. For $k > 1$, this is the probability that the number of primary particles decreases as a result of a merger in Π_n taking k primary particles, whereas $q''(n,1)$ is the probability of a unary merger which eliminates a primary particle.

The sequence of decrements of the jump chain counting the number of primary particles among $[n]$ is a random composition (ordered partition) of n. For instance, on $[n]$ with $n = 7$, this sequence may realise as 3, 1, 2, 2, where the second term results from a *collision* (see Section 4.1 below) which cannot be derived from the path of Π_7 alone. These random compositions for $n = 1, 2, \ldots$ are consistent and constitute a *regenerative composition structure*, as introduced in [22].

4. Functionals and recursions

4.1. Functionals of the coalescent

Intuitively, collision events are spaced over an extended period of time for coalescents with big mergers and accumulate near 0 for coalescents with small mergers. To capture this apparent paradox, it is helpful to consider various functionals associated with the full path of the block counting process N_n such as

- X_n, the number of collisions (merging events) equal to the number of jumps in N_n until termination;
- $\tau_n = \min\{t\colon N_n(t) = 1\}$, the absorption time of Π_n; and
- $L_n = \int_0^{\tau_n} N_n(t)\,\mathrm{d}t$, the total branch length of the coalescent tree, equal to the cumulative lifetime of all blocks until absorption;

together with analogues of these functionals for the partial coalescent up to a given time t. Further related quantities include

- F_n, the number of blocks in the allelic partition for the coalescent with mutation;
- M_n, the number of segregating sites, equal to the number of mutations on the coalescent tree; and
- the external branch length, equal to the cumulative lifetime of primary particles.

More delicate analysis involves the partition $\Pi_n(t)$ itself, in particular the number of blocks $N_n(t)$, the size of the block containing 1 (a size-biased pick from $\Pi_n(t)$), the partition $\Pi_n(\tau_n-)$ before the last collision, etc.

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

Random recursions belong to the toolbox for the analysis of coalescents. One obvious source of them involves decompositions at the time of the first transition. For example, the number of collisions X_n in Π_n satisfies

$$X_1 = 0, \qquad X_n \stackrel{\mathrm{D}}{=} 1 + X'_{n-J_n+1}, \quad n \geq 2,$$

where J_n has the distribution at (7) and, for every fixed k, X'_k is independent of J_n and is an independent copy of X_k.

4.2. Cutting random trees

Representations of coalescents via random tree cutting have been another source of recursions in the symmetric beta case.

A generalized recursive tree [43] has nodes labelled by blocks of a partition of $[n]$, with minimal elements of the blocks increasing along every path from the root to a leaf. In the cutting process an edge of the tree is chosen at unit rate and cut at the node closer to the root, the disconnected subtree is removed and its labels joined to the block at the node. If the process starts with the uniformly random recursive tree with n nodes labelled by $[n]$, the resulting partition-valued process is the Bolthausen–Sznitman coalescent [31]. The representation was used in [31] to study the time reversal of the coalescent, in particular to show that the blocks of $\Pi_n(\tau_n-)$, i.e. the members of the last collision, constitute one giant block and many small blocks with altogether about n^U elements, where U is uniform on $[0, 1]$. The number of collisions X_n of the Bolthausen–Sznitman coalescent has the same law as the number of cuts needed to isolate the root of the random recursive tree. This observation was exploited in [19, 36] to derive a limit distribution for X_n.

In [1] the starting configuration is a Galton–Watson tree conditioned on having n leaves. The leaves are labelled by $[n]$. In one move the tree is cut at a random node, with all labels in the isolated subtree moving to the node which becomes a leaf in the truncated tree. For a suitable choice of the offspring distribution, the dynamics reproduce the symmetric beta coalescent with $\frac{1}{2} \leq \alpha < 1$. The case $\alpha = \frac{1}{2}$ was represented by random binary tree cutting in [2].

The last collision in the symmetric beta case was studied in [33].

5. Coalescents coming down from infinity

5.1. Small-time asymptotics

For Π coming down from infinity, the number of blocks $N(t)$ is finite for all $t > 0$. With improper initial state $N(0) = \infty$, the block counting process $N = (N(t),\ t \geq 0)$ is regarded as an entrance law appearing as a monotonic limit of the N_n as $n \to \infty$.

For Kingman's coalescent, the jump chain associated with N simply goes down by unit jumps, in which the holding time at k blocks is exponential with parameter $\binom{k}{2}$. The holding times at $k = \ldots, 3, 2$ blocks are mutually independent and also independent of the sequence of states passed by Π. For the absorption time of Π, we have a representation

$$\tau = \sum_{k=2}^{\infty} \varepsilon_k \Big/ \binom{k}{2},$$

with independent and identically distributed (i.i.d.) standard exponential ε_k and so $\mathbb{E}[\tau] = 2$. As $t \downarrow 0$, the classical limit theorems can be applied to show that $N(t)$ satisfies the strong law $N(t) \sim 2t^{-1}$ and is approximately normal (see [3, 48]). The function $2t^{-1}$ measures the decay of the number of blocks at small times, and is called the *speed of coming down from infinity*.

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

The analogous speed function exists for all Λ-coalescents of type (IV), as has been shown in different degrees of generality in [6, 9, 13]. Introduce

$$\psi(z) = \int_0^1 [\mathrm{e}^{-zx} - 1 + zx]x^{-2}\,\Lambda(\mathrm{d}x).$$

The behaviour of ψ at infinity is similar to that of (4), but ψ is more natural in the related context of a continuous-state branching process [13, 14]. In particular, Schweinsberg's condition (3) is equivalent to $\int_1^\infty (\psi(z))^{-1}\mathrm{d}z < \infty$.

The speed function $v = v(t)$ is identified as a solution to $\int_{v(t)}^\infty (\psi(z))^{-1}\mathrm{d}z = t$, which is well defined and unique for small enough t. Then (see [6]), as $t \downarrow 0$,

$$N(t) \sim v(t)$$

almost surely and in the pth mean for $p \geq 1$.

When $\Lambda([0, 1]) = 1$ and the regularity condition (5) holds with $0 < a < 1$, the speed of coming down from infinity becomes

$$N(t) \sim c\, t^{-1/(1-a)} \quad \text{as } t \to 0,$$

where $c = [(2-a)/(A\,\Gamma(a))]^{1/(1-a)}$. Similar small-time asymptotics hold for the number of blocks with frequency at most $x\,t^{1/(a-1)}$ [13]. In [42] fluctuations have been studied, in particular, $t^{-1/(2-a)}[N(t)/v(t) - 1]$ is shown to converge in distribution to a $(2-a)$-stable law.

The general bound $\liminf_{t\to 0} N(t)t \geq 2$ almost surely (see [6]) shows that Kingman's coalescent achieves the highest possible speed. This implies that the total tree length for the infinite coalescent is infinite and, since $N_n \uparrow N$ $(n \to \infty)$, also that $L_n \to \infty$ for any Λ.

5.2. Large-sample asymptotics

Fluctuations in the large-n regime have been studied under variants of the regularity condition (5) with $0 < a < 1$. A key observation is the weak convergence of the merger size J_n with distribution (7) to a random variable J with distribution

$$\mathbb{P}\{J = k\} = (2-a)\,\frac{(a)_{k-2}}{k!}, \qquad k \geq 2,$$

where $(x)_k = \Gamma(x+k)/\Gamma(x)$ denotes the Pochhammer factorial. The limit variable has finite mean $\mathbb{E}[J-1] = (1-a)^{-1}$ but the variance is infinite.

The convergence of the step distribution suggests that the jump chain for N_n can be approximated by a renewal sequence $n = R_0 > R_1 > \cdots > 0$ with the generic decrement $J - 1$. Whenever the approximation works, a standard result of renewal theory can be applied to show that the number of collisions X_n satisfies

$$\frac{X_n - (1-a)n}{(1-a)n^{1/(2-a)}} \xrightarrow{\mathrm{D}} \mathcal{S}_{2-a} \quad \text{as } n \to \infty, \tag{8}$$

where $\mathcal{S}_{2-a}$ is a $(2-a)$-stable random variable maximally skewed to the left and with location parameter 0. The weak convergence of J_n is, of course, far from being sufficient to justify the renewal approximation, since rare big mergers could easily affect the asymptotics of X_n.

The heuristic argument leading to (8) was made rigorous by three then (c. 2007) disjoint clusters of authors [16, 24, 35]. In [24], under the assumption that the error in (5) is $O(x^{a+\varsigma})$, where $\varsigma < 1$ but is not too small, tight stochastic bounds on J_n were constructed to squeeze N_n

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

between two renewal processes with the same limit distribution. In [16] a similar assumption on ν was adopted. The setting in [35] concerns decreasing Markov chains with certain transition probabilities and this covers the beta$(a, 1)$ case. This work suggests that the fluctuations might be sensitive to the second term of the asymptotic expansion of Λ near 0.

While X_n depends only on the jump chain, other sources of randomness may cause discontinuities in the limit laws for more complex functionals. Consider the tree length L_n. Replacing the jump chain by the stationary renewal process, L_n should be approximated by the sum

$$\sum_j \frac{R_j\, \varepsilon_j}{\lambda_{R_j}}.$$

Given the R_j, the sum of exponentials has bounded variance, so it may be reasonable to further simplify by replacing the exponentials by their common mean $\mathbb{E}[\varepsilon_j] = 1$. Now $\lambda_n \sim c_1 n^{2-a}$ which suggests the further approximation by a constant multiple of $Y_n = \sum_j R_j^{a-1}$. At this stage renewal theory is applicable, showing the convergence of $(Y_n - c\, n^{2-a})/n^{\delta}$ to a stable law, where $\delta = (2-a)^{-1} + a - 1$ may be positive or negative. If $\delta > 0$, the stable approximation to L_n can indeed be pursued. But, if $\delta < 0$, the fluctuation of Y_n is dominated by the randomness coming from the ε_j; hence, fluctuations of L_n about the mean remain bounded as $n \to \infty$. For positive δ, i.e. for a in the range $2 - \phi < a < 1$, where $\phi = (1 + \sqrt{5})/2$ is the golden ratio, this line of argument was used in [16] to derive a stable limit distribution for the length of a partial tree spanned on the first $(X_n \wedge \lfloor ns \rfloor)$ collisions with $0 < s < 1 - a$, so bounded away from the root. Kersting [38] justified the stable limit for L_n in the symmetric beta case with $2 - \phi \le a < 1$, and showed the weak convergence of centered L_n for the range $0 < a < 2 - \phi$. This result was extended in [39] to a functional limit theorem for the length of a coalescent tree evolving with time.

The number M_n of segregating sites involves a third randomness factor associated with the Poisson sampling, which makes an approximately Gaussian contribution given a path of N_n. For the symmetric beta case, Kersting [38] showed that the limit distribution of M_n is stable for $2 - \sqrt{2} < a < 1$, normal for $0 < a < 2 - \sqrt{2}$, and a mixture of a stable and normal laws in the boundary case $a = 2 - \sqrt{2}$. The analogous phase transition for the number of segregating sites on a partial tree was proved in [16] and independently conjectured at about the same time for F_n [23].

The discontinuities in the limit laws should be regarded as second-order phase transitions, as there is no break in the mean asymptotics.

5.3. The allelic partition

Consider F_n, the number of blocks in the allelic partition for the coalescent with freeze rate ρ. In [7, 8] the small-time asymptotics were translated into the large-n asymptotic behaviour

$$F_n \sim \rho \int_0^n \frac{z\, \mathrm{d}z}{\psi(z)} \quad \text{in probability.}$$

Under a regularity condition with $0 < a < 1$, a stronger result

$$F_n \sim c\rho n^a \quad \text{almost surely} \tag{9}$$

was shown. By Tauberian theory for exchangeable partitions [25], the number of blocks of size k is asymptotic to $c_k \rho\, n^a$ for every $k = 1, 2, \ldots$ (in Pitman's terminology [50], call this property the 'constant a-diversity of allelic partitions').

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

For exchangeable partitions with nonrandom frequencies, the asymptotic behaviour around (9) implies a multivariate normal limit for the small-block counts. In view of the available results on L_n and M_n one can expect that the normal limit is still valid for $a < 2 - \sqrt{2}$ but changes to stable for $a > 2 - \sqrt{2}$.

6. Coalescents with dust

6.1. Construction from subordinator

The dynamics of a Λ-coalescent with dust component are not centered around a single big crunch event as for a coalescent of type (IV). This feature makes types (I) and (II) more appropriate for modelling systems in which clusters emerge from dispersed matter in a discontinuous fashion over an extended time interval. On the analytical side, many features of a coalescent can be understood by following the simpler dust decay subprocess.

Recall the Poisson construction. Every time a merger is driven by (x, t) the cumulative frequency of singletons in $\Pi(t-)$ is diminished by the factor $1 - x$. Hence, the frequency at time t is a product over Poisson atoms in $[0, 1] \times [0, t]$. Passing to logarithms this implies that the frequency can be represented as $\mathrm{e}^{-S(t)}$, where $S = (S(t),\ t \geq 0)$ is an increasing process with independent increments (subordinator), with the Laplace transform

$$\mathbb{E}[\mathrm{e}^{-zS(t)}] = \mathrm{e}^{-t\Phi(z)},$$

where

$$\Phi(z) = \int_0^1 [1 - (1 - x)^z]\, \nu(\mathrm{d}x). \tag{10}$$

Equation (10) is not a standard way of representing the Laplace exponent, but it has a transparent meaning, namely, $\Phi(n)$ is the rate at which at least one of n blocks gets marked by heads, and this is finite by virtue of $\mathrm{m}_{-1} = \int_0^1 x\, \nu(\mathrm{d}x) < \infty$. If (6) holds for $1 < a < 2$, Karamata's theorem implies that $\Phi(z) \sim A_1 \Gamma(a - 1) z^{2-a}$ for $z \to \infty$.

It makes sense to observe when a particular primary particle j engages in its first merger in the infinite coalescent, as this does not occur instantaneously but at an exponential time with parameter $\lambda_1^* := \mathrm{m}_{-1}$. Call the result of the merger a *secondary block*. Restricting the infinite coalescent to $[n]$ (with $n \geq j$) it can happen that the first collision of j is unary, i.e. it does not involve other members of $[n]$, in which case we regard the primary particle as being transformed into a singleton secondary block.

Furthermore, to deal with S, as opposed to e^{-S}, it is convenient to construct the infinite coalescent Π directly from the path of S. Let $\varepsilon_1, \varepsilon_2, \ldots$ be i.i.d. exponential marks assigned initially to the primary particles $1, 2, \ldots$. Secondary blocks evolve and receive exponential marks labelled by subsets of $\mathbb{N}$ as follows. If t is a jump time for S, all labels of exponential marks of primary particles and secondary blocks within the interval $(S(t-), S(t)]$ are merged into one label, which is assigned a new exponential mark located at $S(t) + \varepsilon$, where ε is another independent exponential variable. A snapshot of this process at time t is some collection of marks within $[S(t), \infty)$ with labels making up a partition $\Pi(t)$.

6.2. The absorption time

To illustrate the use of coupling of Π_n with the subordinator S, consider the absorption time τ_n as in [27]. Let τ_n^* be the time of the last (possibly, unary in Π_n) merger taking a primary particle from $[n]$. Clearly, $\tau_n^* \leq \tau_n$ and after τ_n^* some number D_n of secondary blocks remain. It can be shown that the sequence $\{D_n\}$ is tight if $\sum_n \Phi(n)/n^2 < \infty$, and then $\tau_n - \tau_n^*$ is bounded. The condition is not very restrictive since $\Phi(z) = o(z)$ for large z always holds,

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

and (5) with $a > 1$ is sufficient. Letting $T_s = \inf\{t\colon S(t) \geq s\}$ be the first passage time over level s, we have $\tau_n^* = T_{\varepsilon_1 \vee \cdots \vee \varepsilon_n}$, which is the first passage time over the maximal exponential mark. The latter is approximated by $T_{\log n}$, so $(\tau_n - a_n)/b_n$ converges weakly if and only if $(T_{\log n} - a_n)/b_n$ does for any pair (a_n, b_n) of scaling and centering constants. For example, with

$$\mathrm{s}^2 = \operatorname{var}(S(1)) = \int_0^1 |\log(1-x)|^2 \, \nu(\mathrm{d}x) < \infty$$

and

$$\mathrm{m} = \mathbb{E}[S(1)] = \int_0^1 |\log(1-x)| \, \nu(\mathrm{d}x),$$

the standard renewal theorem implies that $(\tau_n - a_n)/b_n$ is asymptotically normal with the choice of constants $a_n = \mathrm{m}^{-1} \log n$ and $b_n = (\mathrm{m}^{-3}\mathrm{s}^2 \log n)^{1/2}$. We stress that for this result the behaviour of ν (or Λ) at the right endpoint of $[0, 1]$ is also important. Possible limit distributions and conditions for the weak convergence of T_s are known, whence we can classify all limit laws for τ_n^*. As another example, the limit is β-stable if the tail of ν satisfies $\nu([1 - \mathrm{e}^{-y}, 1]) \sim y^{-\beta}$ as $y \to \infty$.

6.3. Asymptotics of the number of collisions

Let K_n be the number of mergers in Π_n that take at least one primary particle, and let $K_{n,r}$ be the number of mergers taking r primary particles. Under (6), with $1 < a < 2$, the number of collisions X_n is approximated by $K_n - K_{n,1}$. Introduce the random variable

$$I_{2-a} = \int_0^\infty \mathrm{e}^{-(2-a)S(t)} \, \mathrm{d}t,$$

known as the exponential functional of the subordinator $(2-a)S$. The distribution of I_{2-a} is uniquely determined by the moments

$$\mathbb{E}[I^k] = \frac{k!}{\prod_{i=1}^k \Phi((2-a)i)}.$$

Applying results from [29] on the number of nonunit parts in the regenerative composition structure yields

$$\frac{X_n}{\Gamma(a) n^{2-a}} \xrightarrow{\mathrm{D}} I \quad \text{as } n \to \infty.$$

For $N_n(t)$ with this scaling, the limit law is the partial integral of $\mathrm{e}^{-(2-a)S}$. See [32] for more general results on jump counting which apply to both coalescents and regenerative compositions.

6.4. Dust in coalescents with freeze

Analysing a random recursion, Möhle and Freund [20, Theorem 1.2] proved that the number of blocks in the allelic partition satisfies

$$n^{-1} F_n \xrightarrow{\mathrm{D}} \rho \int_0^\infty \mathrm{e}^{-S(t) - \rho t} \, \mathrm{d}t. \tag{11}$$

Here is another view and a seedling for future research. Superposing the pure-jump dust decay process derived from the coalescent Π with freezing at rate ρ yields a process in which the cumulative frequency of dust decays as $\exp(-S(t) - \rho t)$. By analogy with the regenerative composition structures [22], the terminal partition of $\mathbb{N}$ has singletons with cumulative frequency

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

given by the exponential functional in (11). It follows that (11) can be improved to almost-sure convergence and that the allelic partition itself has a giant dust component. Similarly, it is clear from [22] that, under (6), the number of blocks of size $r = 2, 3, \dots$ in the allelic partition is of the order n^{2-a}. It would be interesting to find explicitly the asymptotics, also for the allelic partition associated with $\Pi(t)$.

6.5. The compound Poisson case

Whenever $\mathrm{m}_{-2} < \infty$, it is natural to scale ν to a probability measure. The coalescent in this case is sometimes called 'simple'. The subordinator S is a compound Poisson process; hence, the dust decay process e^{-S} passes through a stick-breaking sequence and has independent exponential holding times.

In contrast to the case of infinite ν, most of the variability of K_n now comes from the times close to absorption. The typical order of growth of K_n is $\log n$ while, for each fixed r, $K_{n,r}$ remains tight. The variables $(K_n - b_n)/a_n$ and $(X_n - b_n)/a_n$ converge to the same distribution provided the limit exists. From results on K_n in [26], the limit distribution follows by counting renewals on $[0, \log n]$ for a random walk.

7. The boundary cases

Much attention has been devoted to the Bolthausen–Sznitman coalescent [5, 12, 15, 18, 31], which belongs to the critical case $a = 1$. See [30] for limit laws for X_n and L_n in the beta$(1, b)$ case.

A natural setting for the critical cases $a = 0, 1, 2$ is the assumption that $\nu([x, 1]) \sim \ell(x)x^{a-2}$, where ℓ is a function of slow variation as $x \to 0$. Under the condition $\mathrm{m}_{-1} < \infty$, results on regenerative composition structures are still applicable to the dust decay process. If $a = 1$ and $\ell(x) \to \infty$ (but not too fast to guarantee $\mathrm{m}_{-1} < \infty$), the limit law for X_n can be concluded from [29] in much the same way as for the case $1 < a < 2$. If $a = 2$ and $\ell(x) \to \infty$ (hence, $\mathrm{m}_{-2} = \infty$), there are further phase transitions at the level of fluctuations [4, 21]: different modes of behaviour appear when, for instance, ℓ is $\log\log x^{-1}$, $|\log x|^2$, or $\exp(\sqrt{\log x^{-1}})$. The case $a = 2$ and bounded ℓ belongs to the 'simple' type (IV).

We are not aware of any other results on fluctuations for $a = 0, 1$ with $\ell \neq$ constant. The case $a = 1$ must be particularly interesting since it is compatible with any of the types (II), (III), and (IV).

8. Limit distributions for beta coalescents

We summarise in Tables 1–4 known results on the limit laws of X_n, τ_n, L_n, and M_n, respectively, for beta(a, b)-coalescents. The subordinator $(S(t),\ t \geq 0)$ has Laplace exponent

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

The characteristic function of an α-stable law is given by

$$z \mapsto \exp\left\{-|z|^\alpha \left(1 + \mathrm{i}\tan\left(\frac{\pi\alpha}{2}\right)\mathrm{sgn}(z)\right)\right\}, \qquad t \in \mathbb{R}.$$

The constants m and s^2 given in Table 2 are defined by

$$\mathrm{m} = \frac{a + b - 1}{(a - 1)(2 - a)}(1 - (a + b - 2)[\Psi(a + b - 1) - \Psi(b)])$$

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

TABLE 1: Limit distributions for $(X_n - a_n)/b_n$ for beta(a, b)-coalescents. Here $\alpha = 2 - a$, $r_1 = \zeta(2, b)$, $r_2 = 2\zeta(3, b)$, where $\zeta(\cdot, \cdot)$ is the Hurwitz zeta function, $m_1 = \Psi(a - 2 + b) - \Psi(b)$, and $m_2 = \Psi'(b) - \Psi'(a - 2 + b)$, where $\Psi(\cdot)$ is the logarithmic derivative of the gamma function.

a	b	a_n	b_n	Limit law	Source
$a = 0$	$b > 0$	$n - 1$	1	δ_0	Obvious
$0 < a < 1$	$b > 0$	$n(\alpha - 1)$	$(\alpha - 1)n^{1/\alpha}$	α-stable	[16, 24, 30]
$a = 1$	$b > 0$	$n(\log n)^{-1} + n(\log n)^{-2} \log\log n$	$\dfrac{n}{(\log n)^2}$	1-stable	[19, 36] $(b = 1)$, [30]
$1 < a < 2$	$b > 0$	0	$\alpha^{-1}\Gamma(\alpha)n^{\alpha}$	$\int_0^\infty \mathrm{e}^{-\alpha S(t)}\,\mathrm{d}t$	[27, 32]
$a = 2$	$b > 0$	$(2r_1)^{-1}(\log n)^2$	$(3^{-1}r_1^{-3}r_2 \log^3 n)^{1/2}$	Normal	[27, 37]
$a > 2$	$b > 0$	$m_1^{-1} \log n$	$(m_1^{-3}m_2 \log n)^{1/2}$	Normal	[27, 28]

TABLE 2: Limit distributions for $(\tau_n - c_n)/d_n$ for beta(a, b)-coalescents. The ε_k are i.i.d. exponential variables, and $c_1 = b(b + 1)\zeta(2, b)$ and $c_2 = 2b(b + 1)\zeta(3, b)$. The Gumbel distribution function is $x \mapsto \mathrm{e}^{-\mathrm{e}^{-x}}$, $x \in \mathbb{R}$. The constants m_1 and m_2 are the same as in Table 1, and, for $a > 2$, $\gamma = (a - 1 + b)(a - 2 + b)/(a - 1)(a - 2)$. The result for $a > 1$ is a special case of Theorem 4.3 of [27].

a	b	c_n	d_n	Limit law	Source
$a = 0$		0	1	$\sum \varepsilon_k / \binom{k}{2}$	[55]
$0 < a < 1$	$b > 0$	0	1	Unknown	
$a = 1$	$b = 1$	$\log\log n$	1	Gumbel	[31]
$a = 1$	$b \neq 1$			Unknown	
$1 < a < 2$	$b > 0$	$\mathtt{m}^{-1} \log n$	$(\mathtt{m}^{-3}\mathtt{s}^2 \log n)^{1/2}$	Normal	[27]
$a = 2$	$b > 0$	$c_1^{-1} \log n$	$(c_1^{-3}c_2 \log n)^{1/2}$	Normal	[27]
$a > 2$	$b > 0$	$(\gamma m_1)^{-1} \log n$	$\gamma^{-1}(m_1^{-3}(m_2 + m_1^2) \log n)^{1/2}$	Normal	[27, 28]

TABLE 3: Limit distributions for $(L_n - e_n)/f_n$ for beta(a, b)-coalescents. The constants are $c_1 = \Gamma(\alpha + 1)(\alpha - 1)/(2 - \alpha)$, $c_2 = \Gamma(\alpha + 1)(\alpha - 1)^{1+\alpha^{-1}}/\Gamma^{\alpha^{-1}}(2 - \alpha)$, $a_3 = \frac{1}{2}(3 - \sqrt{5})$, $\alpha = 2 - a$, and $\beta = 1 + \alpha - \alpha^2$, and η is some nondegenerate absolutely continuous random variable.

a	b	e_n	f_n	Limit law	Source
$a = 0$		$2 \log n$	2	Gumbel	[18, 55]
$0 < a < a_3$	$b = 2 - a$	$c_1 n^a$	1	η	[38]
$a_3 = \frac{1}{2}(3 - \sqrt{5})$	$b = 2 - a$	$c_1 n^a$	$c_2(\log n)^{\alpha^{-1}}$	α-stable	[38]
$a_3 < a < 1$	$b = 2 - a$	$c_1 n^a$	$c_2(\beta^{-1}n^{\beta})^{\alpha^{-1}}$	α-stable	[38]
$a = 1$	$b > 0$	$n(b \log n)^{-1} + b^{-1}n \log\log n (\log n)^{-2}$	$\dfrac{n}{b(\log n)^2}$	1-stable	[18] $(b = 1)$, [30]
$0 < a < 1$	$b \neq 2 - a$			Unknown	
$a > 1$	$b > 0$	0	n	$\int_0^\infty \mathrm{e}^{-S_t}\,\mathrm{d}t$	[45]

and

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

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

TABLE 4: Limit distributions for $(M_n - g_n)/h_n$ for beta(a, b)-coalescents. Here ρ is the rate of a Poisson process on the coalescent tree, $a_2 = 2 - \sqrt{2}$, the other constants c_1, c_2, α, and β are the same as in Table 3, and ζ is the sum of a centered normal random variable with variance θc_1 and $\rho c_2(\beta)^{-\alpha^{-1}}$ times an independent α-stable random variable.

a	b	g_n	h_n	Limit law	Source
$a = 0$		$2\rho \log n$	$\sqrt{2\rho \log n}$	Normal	ESF
$0 < a < a_2$	$b = 2 - a$	$\rho c_1 n^a$	$\sqrt{\rho c_1} n^{a/2}$	Normal	[38]
$a = a_2$	$b = 2 - a$	$\rho c_1 n^a$	$n^{a/2}$	ζ	[38]
$a_2 < a < 1$	$b = 2 - a$	$\rho c_1 n^a$	$\rho c_2(\beta^{-1} n^\beta)^{\alpha^{-1}}$	α-stable	[38]
$a = 1$	$b > 0$	$n(b \log n)^{-1} +$ $b^{-1} n \log\log n (\log n)^{-2}$	$\dfrac{\rho n}{b(\log n)^2}$	1-stable	[18] ($b = 1$), [30]
$0 < a < 1$	$b \neq 2 - a$			Unknown	
$a > 1$	$b > 0$	0	ρn	$\int_0^\infty \mathrm{e}^{-S_t}\, \mathrm{d}t$	[45]

The limit distribution for X_n in the Bolthausen–Sznitman coalescent was obtained in [19] with the aid of the singularity analysis of generating functions, and later probabilistically via coupling with random walks with barrier in [36] by using a relation to random recursive trees; see Section 4.2. The latter approach proved useful [35] to study collisions in beta$(a, 1)$-coalescents with $0 < a < 2$. Asymptotics of moments of X_n for beta$(a, 1)$-coalescents with $0 < a \leq 1$ appear in [35, 47], and for $a = 2$ and $b > 0$ in [37]. All entries of Table 1 (except for the case $a = 1$) are specializations from more general Λ-coalescents [24, 27, 28, 32]. The value $a = 0$ corresponds to Kingman's coalescent.

References

[1] ABRAHAM, R. AND DELMAS, J.-F. (2013). β-coalescents and stable Galton–Watson trees. Preprint. Available at http://arxiv.org/abs/1303.6882v1.

[2] ABRAHAM, R. AND DELMAS, J.-F. (2013). A construction of a β-coalescent via the pruning of binary trees. *J. Appl. Prob.* **50,** 772–790.

[3] ALDOUS, D. J. (1999). Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli* **5,** 3–48.

[4] BARBOUR, A. D. AND GNEDIN, A. V. (2006). Regenerative compositions in the case of slow variation. *Stoch. Process. Appl.* **116,** 1012–1047.

[5] BASDEVANT, A.-L. AND GOLDSCHMIDT, C. (2008). Asymptotics of the allele frequency spectrum associated with the Bolthausen–Sznitman coalescent. *Electron. J. Prob.* **13,** 486–512.

[6] BERESTYCKI, J., BERESTYCKI, N. AND LIMIC, V. (2010). The Λ-coalescent speed of coming down from infinity. *Ann. Prob.* **38,** 207–233.

[7] BERESTYCKI, J., BERESTYCKI, N. AND LIMIC, V. (2014). Asymptotic sampling formulae for Λ-coalescents. *Ann. Inst. H. Poincaré Prob. Statist.* **50,** 715–731.

[8] BERESTYCKI, J., BERESTYCKI, N. AND SCHWEINSBERG, J. (2007). Beta-coalescents and continuous stable random trees. *Ann. Prob.* **35,** 1835–1887.

[9] BERESTYCKI, J., BERESTYCKI, N. AND SCHWEINSBERG, J. (2008). Small-time behavior of beta coalescents. *Ann. Inst. H. Poincaré Prob. Statist.* **44,** 214–238.

[10] BERESTYCKI, N. (2009). *Recent Progress In Coalescent Theory* (Math. Surveys **16**). Sociedade Brasileira de Matemática, Rio de Janeiro.

[11] BERTOIN, J. (2010). Exchangeable coalescents. Nachdiplom Lectures, ETH Zürich.

[12] BERTOIN, J. AND LE GALL, J.-F. (2000). The Bolthausen–Sznitman coalescent and the genealogy of continuous-state branching processes. *Prob. Theory Relat. Fields* **117,** 249–266.

[13] BERTOIN, J. AND LE GALL, J.-F. (2006). Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math.* **50,** 147–181.

[14] BIRKNER, M. *et al.* (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Prob.* **10,** 303–325.

https://doi.org/10.1239/jap/1417528464 Published online by Cambridge University Press

[15] Bolthausen, E. and Sznitman, A.-S. (1998). On Ruelle's probability cascades and an abstract cavity method. *Commum. Math. Phys.* **197,** 247–276.
[16] Delmas, J.-F., Dhersin, J.-S. and Siri-Jegousse, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Prob.* **18,** 997–1025.
[17] Dong, R., Gnedin, A. and Pitman, J. (2007). Exchangeable partitions derived from Markovian coalescents. *Ann. Appl. Prob.* **17,** 1172–1201.
[18] Drmota, M., Iksanov, A., Moehle, M. and Roesler, U. (2007). Asymptotic results concerning the total branch length of the Bolthausen–Sznitman coalescent. *Stoch. Process. Appl.* **117,** 1404–1421.
[19] 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.
[20] Freund, F. and Möhle, M. (2009). On the number of allelic types for samples taken from exchangeable coalescents with mutation. *Adv. Appl. Prob.* **41,** 1082–1101.
[21] Gnedin, A. and Iksanov, A. (2012). Regenerative compositions in the case of slow variation: a renewal theory approach. *Electron. J. Prob.* **17,** 19pp.
[22] Gnedin, A. and Pitman, J. (2005). Regenerative composition structures. *Ann. Prob.* **33,** 445–479.
[23] Gnedin, A. and Yakubovich, Y. (2007). Collisions and freezing events in coalescents with multiple mergers. Tech. rep. 46/2007, Mathematisches Forschungsinstitut Oberwolfach.
[24] Gnedin, A. and Yakubovich, Y. (2007). On the number of collisions in Λ-coalescents. *Electron. J. Prob.* **12,** 1547–1567.
[25] Gnedin, A., Hansen, B. and Pitman, J. (2007). Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. *Prob. Surveys* **4,** 146–171.
[26] 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.
[27] Gnedin, A., Iksanov, A. and Marynych, A. (2011). On Λ-coalescents with dust component. *J. Appl. Prob.* **48,** 1133–1151.
[28] Gnedin, A., Iksanov, A. and Möhle, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Prob.* **45,** 1186–1195.
[29] Gnedin, A., Pitman, J. and Yor, M. (2006). Asymptotic laws for compositions derived from transformed subordinators. *Ann. Prob.* **34,** 468–492.
[30] Gnedin, A., Iksanov, A., Marynych, A. and Möhle, M. (2014). On asymptotics of the beta coalescents. *Adv. Appl. Prob.* **46,** 496–515.
[31] Goldschmidt, C. and Martin, J. B. (2005). Random recursive trees and the Bolthausen–Sznitman coalescent. *Electron. J. Prob.* **10,** 718–745.
[32] Haas, B. and Miermont, G. (2011). Self-similar scaling limits of non-increasing Markov chains. *Bernoulli* **17,** 1217–1247.
[33] Hénard, O. (2013). The fixation line. Preprint. Available at http://arxiv.org/abs/1307.0784v2.
[34] Huillet, T. E. (2014). Pareto genealogies arising from a Poisson branching evolution model with selection. *J. Math. Biol.* **68,** 727–761.
[35] Iksanov, A. and Möhle, M. (2008). On the number of jumps of random walks with a barrier. *Adv. Appl. Prob.* **40,** 206–228.
[36] 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.
[37] Iksanov, A., Marynych, A. and Möhle, M. (2009). On the number of collisions in beta(2, *b*)-coalescents. *Bernoulli* **15,** 829–845.
[38] Kersting, G. (2012). The asymptotic distribution of the length of beta-coalescent trees. *Ann. Appl. Prob.* **22,** 2086–2107.
[39] Kersting, G., Schweinsberg, J. and Wakolbinger, A. (2014). The evolving beta coalescent. Preprint. Available at http://arxiv.org/abs/1402.4534v1.
[40] Kingman, J. F. C. (1982). The coalescent. *Stoch. Process. Appl.* **13,** 235–248.
[41] Kingman, J. F. C. (1982). On the genealogy of large populations. In *Essays in Statistical Science* (J. Appl. Prob. Spec. Vol. **19A**), eds J. Gani and E. J. Hannan, Applied Probability Trust, Sheffield, pp. 27–43.
[42] Limic, V. and Talarczyk, A. (2014). Second-order asymptotics for the block counting process in a class of regularly varying λ-coalescents. Preprint. Available at http://arxiv.org/abs/1304.5183v2.
[43] Meir, A. and Moon, J. W. (1974). Cutting down recursive trees. *Math. Biosci.* **21,** 173–181.
[44] Möhle, M. (2006). On sampling distributions for coalescent processes with simultaneous multiple collisions. *Bernoulli* **12,** 35–53.
[45] Möhle, M. (2006). On the number of segregating sites for populations with large family sizes. *Adv. Appl. Prob.* **38,** 750–767.
[46] Möhle, M. and Sagitov, S. (2001). A classification of coalescent processes for haploid exchangeable population models. *Ann. Prob.* **29,** 1547–1562.

[47] Panholzer, A. (2003). Non-crossing trees revisited: cutting down and spanning subtrees. In *Discrete random walks* (Paris, 2003; Discrete Math. Theoret. Comput. Sci. Proc. **AC**), Assoc. Discrete Math. Theoret. Comput. Sci., Nancy, pp. 265–276.
[48] Pfaffelhuber, P., Wakolbinger, A. and Weisshaupt, H. (2011). The tree length of an evolving coalescent. *Prob. Theory Relat. Fields* **151,** 529–557.
[49] Pitman, J. (1999). Coalescents with multiple collisions. *Ann. Prob.* **27,** 1870–1902.
[50] Pitman, J. (2006). *Combinatorial Stochastic Processes* (Lecture Notes Math. **1875**). Springer, Berlin.
[51] Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Prob.* **36,** 1116–1125.
[52] Schweinsberg, J. (2000). Coalescents with simultaneous multiple collisions. *Electron. J. Prob.* **5,** 50pp.
[53] Schweinsberg, J. (2000). A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Commun. Prob.* **5,** 11pp.
[54] Schweinsberg, J. (2003). Coalescent processes obtained from supercritical Galton–Watson processes. *Stoch. Process. Appl.* **106,** 107–139.
[55] Tavaré, S. (2004). Ancestral inference in population genetics. In *Lectures on Probability Theory and Statistics* (Lecture Notes Math. **1837**), Springer, Berlin, pp. 1–188.

ALEXANDER GNEDIN, *Queen Mary University of London*

School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK.
Email address: a.gnedin@qmul.ac.uk

ALEXANDER IKSANOV, *National Taras Shevchenko University of Kyiv*

Faculty of Cybernetics, National Taras Shevchenko University of Kiev, 01033 Kiev, Ukraine.
Email address: iksan@univ.kiev.ua

ALEXANDER MARYNYCH, *National Taras Shevchenko University of Kyiv*

Faculty of Cybernetics, National Taras Shevchenko University of Kiev, 01033 Kiev, Ukraine.
Email address: marynych@unicyb.kiev.ua