The Annals of Applied Probability
2018, Vol. 28, No. 6, 3857–3883
https://doi.org/10.1214/18-AAP1409

THE COLLISION SPECTRUM OF Λ-COALESCENTS[1]

BY ALEXANDER GNEDIN*, ALEXANDER IKSANOV†, ALEXANDER MARYNYCH†,[2] AND MARTIN MÖHLE‡

Queen Mary University of London, Taras Shevchenko National University of Kyiv† and Eberhard Karls University of Tübingen‡*

Λ-coalescents model the evolution of a coalescing system in which any number of blocks randomly sampled from the whole may merge into a larger block. For the coalescent restricted to initially n singletons, we study the collision spectrum $(X_{n,k} : 2 \le k \le n)$, where $X_{n,k}$ counts, throughout the history of the process, the number of collisions involving exactly k blocks. Our focus is on the large n asymptotics of the joint distribution of the $X_{n,k}$'s, as well as on functional limits for the bulk of the spectrum for simple coalescents. Similar to the previous studies of the total number of collisions, the asymptotics of the collision spectrum largely depends on the behaviour of the measure Λ in the vicinity of 0. In particular, for beta(a, b)-coalescents different types of limit distributions occur depending on whether $0 < a \le 1$, $1 < a < 2$, $a = 2$ or $a > 2$.

1. Introduction. The Λ-coalescents introduced by Pitman [43] and Sagitov [45] are partition-valued Markov processes which evolve according to the rule: if at some time the process is restricted to a partition with m separate blocks, then any k particular blocks collide and merge in one block at rate

$$\lambda_{m,k} = \int_{[0,1]} x^k (1-x)^{m-k} x^{-2} \Lambda(\mathrm{d}x), \qquad 2 \le k \le m, \tag{1.1}$$

where Λ is a finite measure on $[0, 1]$. For a Λ-coalescent $\Pi_n = (\Pi_n(t))_{t \ge 0}$ which starts with n singleton blocks and terminates with a single block, the possible states are partitions of the set $[n] := \{1, \dots, n\}$. Due to consistency for various n, one can also view Π_n as the restriction to $[n]$ of an infinite Λ-coalescent $\Pi_\infty = (\Pi_\infty(t))_{t \ge 0}$ with values in the set of partitions of the infinite set $\mathbb{N}$.

Among the functionals characterising the speed of coalescence, the total number of collisions X_n (transitions of Π_n until absorption) has attracted the most attention; see [24, 27, 31, 33, 36, 38] and a survey paper [25]. Here, we are interested

Received August 2017; revised April 2018.

[1]The publication contains the results of studies conducted by President of Ukraine grant for competitive projects Φ70 of the State Fund for Fundamental Research.

[2]Supported by the return fellowship of the Alexander von Humboldt Foundation.

MSC2010 subject classifications. Primary 60J25, 60F17; secondary 60C05, 60G09.

Key words and phrases. Collision spectrum, coupling, exchangeable coalescent, functional approximation.

in more delicate properties of the coalescent process by distinguishing mergers of various sizes, that is decomposing the total number of collisions as

$$X_n = \sum_{k=2}^{n} X_{n,k}, \qquad 2 \le k \le n,$$

where $X_{n,k}$ is the number of collisions resulting in a merger of k blocks. We call the collection of counts $(X_{n,k} : 2 \le k \le n)$ the *collision spectrum* of Π_n.

There are different threads motivating our study of the collision spectrum. In the genealogical representation of the path of Π_n by a tree, directed from leaves up to the root, a k-fold collision corresponds to a node with in-degree k and out-degree 1. Thus the collision spectrum carries the same information as the degree sequence, which is a fundamental characteristic previously studied for other types of random trees [12, 16, 39].

The classical Kingman coalescent with only binary collisions (where Λ is an atomic mass at 0) is widely used in genetics to model the ancestry of neutral populations. Other Λ-coalescents, with collisions of arbitrary multiplicity have been found useful to simulate the ancestry affected by various evolutionary factors such as the large offspring size, selective sweeps, population bottlenecks and rapid adaptation [19, 42, 46]. Outside of biology, in the field of machine learning coalescents have been adopted in hierarchical clustering networks [34, 49]. The choice of model is of paramount importance for the applications. The components of the collision spectrum are easy to compute and can be used for statistical inference on the parameter measure Λ. See [9] for related example of computing likelihoods and model fitting.

A tractable parametric family of Λ-coalescents are the beta-coalescents with characteristic measure

$$\Lambda(\mathrm{d}x) = \frac{1}{\mathrm{B}(a,b)} x^{a-1}(1-x)^{b-1}\mathbb{1}_{(0,1)}(x)\,\mathrm{d}x, \qquad a, b > 0, \tag{1.2}$$

which have collision rates expressible in terms of the beta function $\mathrm{B}(\cdot,\cdot)$. The instance $a = b = 1$, where Λ is the uniform distribution on $[0, 1]$, is known as the Bolthausen–Sznitman coalescent. In application to beta(a, b)-coalescents, our results on limit distributions of $(X_{n,k} - a_{n,k})/b_{n,k}$ are summarised in Table 1. Here, $\mathcal{S}_{2-a}$ denotes a random variable with a $(2-a)$-stable distribution, $\mathcal{E}_{2-a}$ a random variable representable as the exponential functional of a subordinator (nondecreasing Lévy process), and $\mathcal{N}$ a standard normal random variable. Explanation for the scaling/centering constants will appear in a due course.

Some beta coalescents can be associated with a continuous time process of cutting a random tree. In [32], it was observed that the Bolthausen–Sznitman coalescent can be realised by pruning a random recursive tree at a uniformly chosen edge. The same procedure applied to a random binary tree yields the beta$(\frac{3}{2}, \frac{1}{2})$-coalescent [1]. In [2], a beta$(2-\alpha, \alpha)$-coalescent, $\alpha \in (0, 1/2]$ is realised by pruning at nodes the stable Galton–Watson tree. The interest so far was focused on X_n,

TABLE 1
Limits of $(X_{n,k} - a_{n,k})/b_{n,k}$ *for* beta(a, b)*-coalescents*, $b > 0$

a	$a_{n,k}$	$b_{n,k}$	Limit	Source
$0 < a < 1$	$p_{k-1}^{(a)}(1-a)n$	$p_{k-1}^{(a)}(1-a)n^{1/(2-a)}$	$\mathcal{S}_{2-a}$	Theorem 5.1(i)
$a = 1$	$p_{k-1}^{(1)}\frac{n\log(n\log n)}{(\log n)^2}$	$p_{k-1}^{(1)}\frac{n}{(\log n)^2}$	$\mathcal{S}_1$	Theorem 5.1(ii)
$1 < a < 2$	0	$\frac{\Gamma(k+a-2)}{k!}n^{2-a}$	$\mathcal{E}_{2-a}$	Theorem 4.2
$a = 2$	$(k\nu_1)^{-1}\log n$	$\sqrt{(\frac{\nu_2}{\nu_1^3}\frac{1}{k^2} + \frac{1}{k\nu_1})\log n}$	$\mathcal{N}$	Theorem 4.3
$a > 2$	0	1	exists	Theorem 2.1
$a = 3$	0	1	Poisson$(\frac{b}{k-1})$	[41], Theorem 3.1

interpreted in this context as the number of cuts needed to isolate the root of the tree. The collision spectrum gives a much more detailed account on the sizes of subtrees removed in the pruning process.

In two cases, the collision spectrum is trivial. If Λ is the unit mass at 0, then Π_∞ is Kingman's coalescent with only binary collisions, hence $X_n = X_{n,2} = n - 1$. If Λ is the unit mass at 1, then Π_n has a sole transition, and $X_n = X_{n,n} = 1$. In the sequel, we shall assume that Λ has no atoms at 0 and 1. Still, the two extremes offer the intuition that concentration of the measure Λ near 0 affects the abundance of collisions involving small number of blocks. A rough quantification of the concentration of the measure Λ near 0 involves the first two moments of negative order

$$\mathtt{m}_{-r} := \int_{[0,1]} x^{-r}\Lambda(\mathrm{d}x), \qquad r = 1, 2 \tag{1.3}$$

which may be finite or infinite. If $\mathtt{m}_{-1} < \infty$, then the infinite coalescent has a nontrivial *dust* component, which is the collection of singleton blocks of $\Pi_\infty(t)$, for every $t \geq 0$. Under the stronger condition $\mathtt{m}_{-2} < \infty$, the coalescent is *simple*, in the sense that transitions of Π_∞ occur at isolated times of a Poisson process. If $\mathtt{m}_{-1} = \infty$, there is a further division in coalescents Π_∞ that terminate in finite time, and coalescents that have infinitely many blocks for every $t \geq 0$.

In many respects the collision spectrum is similar to the collection of counts of component sizes in decomposable combinatorial structures [4, 44]. By this analogy, it is of interest to look at the large-n behaviour of the first few components of the collision spectrum, as well as to identify the counts that make a major contribution to the total number of collisions. Other types of spectra were studied for Λ-coalescents with mutations (or freeze; see, e.g., [15]), where the terminal state of the process is the so-called allelic partition. These are the site frequency spectrum [7, 10, 11, 14, 18, 48] (for the general Λ-coalescents) and the allelic partition frequency spectrum [6] (for the Bolthausen–Sznitman coalescent). Compared to

these, the collision spectrum is simpler and better amenable to analysis for a large class of measures Λ by the methods previously developed for X_n.

The rest of the paper is organised as follows. In Section 2, we consider the case $\mathtt{m}_{-2} < \infty$ of simple coalescents, showing that the $X_{n,k}$'s jointly converge without scaling as n grows. Thus for every fixed k the contribution of $X_{n,k}$ to X_n is asymptotically negligible, and so is for every finite collection of the values of k. In Section 3 for simple coalescents we assess the joint contribution to X_n in the form of a functional limit theorem for the process of cumulative counts

$$X_n(s) := \sum_{k=2}^{\lfloor n^s \rfloor} X_{n,k}, \qquad s \in [0,1], \tag{1.4}$$

where $\lfloor \cdot \rfloor$ is the floor function. In Section 4, we consider the case where $\mathtt{m}_{-2} = \infty$ but $\mathtt{m}_{-1} < \infty$, showing that under an assumption of regular variation each $X_{n,k}$ is asymptotic to a constant multiple of X_n, while for coalescents related to the gamma-type subordinators the spectrum has a nondegenerate multivariate normal distribution. In Section 5, we consider beta(a, b)-coalescents with $a \in (0, 1]$ (for which $\mathtt{m}_{-1} = \infty$), and show that the $X_{n,k}$'s, properly centered and normalised, jointly converge to multiples of the same stable random variable.

2. Simple coalescents: Convergence of the spectrum. Let $N_n(t)$ be the number of blocks of the partition $\Pi_n(t)$. The counting process $N_n := (N_n(t))_{t\geq 0}$ is Markovian, starting at $N_n(0) = n$ and decrementing from m to $m - k$ at rate $\binom{m}{k+1}\lambda_{m,k+1}$, $1 \leq k < m$. The total collision rate on n blocks is

$$\lambda_n := \sum_{k=2}^{n} \binom{n}{k} \lambda_{n,k} = \int_{[0,1]} x^{-2}\big(1 - nx(1-x)^{n-1} - (1-x)^n\big)\Lambda(\mathrm{d}x), \qquad n \geq 2,$$

and the first transition of Π_n is a collision of $k + 1$ blocks with probability

$$p_{n,k} := \binom{n}{k+1} \frac{\lambda_{n,k+1}}{\lambda_n}, \qquad 1 \leq k < n. \tag{2.1}$$

Hence by the first collision event N_n decrements from state n to $n - I_n$, where I_n has distribution $\mathbb{P}\{I_n = k\} = p_{n,k}$.

For the remainder of this and in the next section, we assume that $\mathtt{m}_{-2} < \infty$. For the simple Λ-coalescents $\lambda_n \to \mathtt{m}_{-2}$, which is the rate of the Poisson process of collision times of Π_∞. Let W be a random variable taking values in $(0, 1)$ with distribution function

$$\mathbb{P}\{W \leq x\} = \frac{\int_{[1-x,1]} y^{-2}\Lambda(\mathrm{d}y)}{\mathtt{m}_{-2}}, \qquad x \in [0,1]. \tag{2.2}$$

Using the weak law of large numbers for binomial random variables, we infer

$$\lim_{n\to\infty} \sum_{k=2}^{\lfloor nx \rfloor + 1} \binom{n}{k} y^k (1-y)^{n-k} = \mathbb{1}_{\{y \leq x\}}, \qquad x, y \in [0,1),$$

and thereupon

$$\begin{aligned}\lim_{n\to\infty}\sum_{k=1}^{\lfloor nx\rfloor} p_{n,k} &= \frac{1}{\mathtt{m}_{-2}}\lim_{n\to\infty}\sum_{k=2}^{\lfloor nx\rfloor+1}\binom{n}{k}\lambda_{n,k}\\ &= \frac{1}{\mathtt{m}_{-2}}\int_{[0,1]}\lim_{n\to\infty}\left(\sum_{k=2}^{\lfloor nx\rfloor+1}\binom{n}{k}y^k(1-y)^{n-k}\right)y^{-2}\Lambda(\mathrm{d}y) \qquad (2.3)\\ &= \frac{1}{\mathtt{m}_{-2}}\int_{[0,x]} y^{-2}\Lambda(\mathrm{d}y) = \mathbb{P}\{1-W\le x\}, \qquad x\in[0,1).\end{aligned}$$

Since the expression in the parentheses is bounded by one the second equality is justified by the dominated convergence theorem. Thus, the variable W has the intuitive meaning of the asymptotic proportion $(n-I_n)/n$ of the blocks remaining after the first collision in Π_n.

Introduce the logarithmic moments

$$\mu := \mathbb{E}|\log W|, \qquad \sigma^2 := \mathrm{Var}(|\log W|)$$

which may be finite or infinite and note that

$$\mu = \frac{1}{\mathtt{m}_{-2}}\int_{[0,1]}|\log(1-x)|x^{-2}\Lambda(\mathrm{d}x).$$

Our first result states the joint convergence in distribution of the collision spectrum. We set $X_{n,k} := 0$ for $k > n$.

THEOREM 2.1. *Suppose* $\mathtt{m}_{-2} < \infty$ *and that the distribution of* $|\log W|$ *is nonarithmetic.*

(i) *If* $\mu < \infty$, *then there exists a nondegenerate random vector* $(X_{\infty,k})_{k\ge2}$ *such that*

$$(X_{n,k})_{k\ge2} \Longrightarrow (X_{\infty,k})_{k\ge2}, \qquad n\to\infty,$$

where $\Longrightarrow$ *denotes the weak convergence in the product space* $\mathbb{R}^\infty$.

(ii) *If* $\mu = \infty$, *then*

$$X_{n,k} \xrightarrow{\mathbb{P}} 0, \qquad n\to\infty$$

for every fixed $k \ge 2$.

Note that the limit variable $X_{\infty,k}$ appearing in this theorem is not the number of k-collisions for the infinite coalescent. For simple coalescents, partition $\Pi_\infty(t)$ has infinitely many blocks and every collision takes infinitely many of them.

The proof of Theorem 2.1 will be based on three lemmas.

LEMMA 2.2. *Let ξ be a random variable with values in $[1,\infty)$. There exists a nondecreasing function $\varphi:[1,\infty)\to[0,\infty)$ slowly varying at ∞ such that*

$$\lim_{x\to\infty}\varphi(x)=\infty \quad \text{and} \quad \mathbb{E}\varphi(\xi)<\infty.$$

PROOF. This follows from the two rather obvious facts. First, there exists a positive function ψ diverging to ∞ such that $\mathbb{E}\psi(\xi)<\infty$. Second, there exists a nondecreasing slowly varying function φ diverging to ∞ and satisfying $\lim_{x\to\infty}(\varphi(x)/\psi(x))=0$. One possible construction of such a φ can be found in the proof of Lemma 4.4 in [20]. □

Applying Lemma 2.2 to $\xi=1/(1-W)$ gives the following.

COROLLARY 2.3. *If $\mathtt{m}_{-2}<\infty$ then there exists a nondecreasing function $\varphi:[1,\infty)\to[0,\infty)$ slowly varying at ∞ such that $\lim_{x\to\infty}\varphi(x)=\infty$ and*

$$\mathbb{E}\varphi\big(1/(1-W)\big)=\int_{[0,1]}y^{-2}\varphi\big(y^{-1}\big)\Lambda(\mathrm{d}y)<\infty. \tag{2.4}$$

LEMMA 2.4. *Let $(a_k)_{k\in\mathbb{N}}$ be a sequence defined recursively as follows*:

$$a_1=a_2=\cdots=a_\ell=0, \qquad a_n=p_{n,\ell}\varphi(n)+\sum_{k=1}^{n-1}p_{n,k}a_{n-k}, \qquad n>\ell,$$

where $\ell\in\mathbb{N}$ is fixed, and φ is a positive function slowly varying at ∞ such that (2.4) *holds. Then the sequence (a_n) is bounded.*

PROOF. Choose $\theta\in(0,1)$ such that $\int_{[1-\theta,1]}y^{-2}\Lambda(\mathrm{d}y)>0$. We first prove by induction that there is a constant $C_\ell>0$ such that

$$a_n\le C_\ell\sum_{m=\ell+1}^{n}\frac{p_{m,\ell}\varphi(m)}{m}, \qquad n>\ell.$$

By adjusting C_ℓ if necessary it is enough to show this for $n>n_0$, where n_0 is any fixed integer. We have

$$\begin{aligned}
a_n&=p_{n,\ell}\varphi(n)+\sum_{k=1}^{n-1}p_{n,n-k}a_k=p_{n,\ell}\varphi(n)+\sum_{k=\ell+1}^{n-1}p_{n,n-k}a_k\\
&\le p_{n,\ell}\varphi(n)+C_\ell\sum_{k=\ell+1}^{n-1}p_{n,n-k}\sum_{m=\ell+1}^{k}\frac{p_{m,\ell}\varphi(m)}{m}\\
&=p_{n,\ell}\varphi(n)+C_\ell\sum_{m=\ell+1}^{n-1}\frac{p_{m,\ell}\varphi(m)}{m}\sum_{k=1}^{n-m}p_{n,k}.
\end{aligned}$$

We need to check that for some $C_\ell > 0$ the inequality

$$p_{n,\ell}\varphi(n) + C_\ell \sum_{m=\ell+1}^{n-1} \frac{p_{m,\ell}\varphi(m)}{m} \sum_{k=1}^{n-m} p_{n,k} \le C_\ell \sum_{m=\ell+1}^{n} \frac{p_{m,\ell}\varphi(m)}{m}$$

holds for large enough n. This is equivalent to

$$p_{n,\ell}\varphi(n) \le C_\ell \sum_{m=\ell+1}^{n} \frac{p_{m,\ell}\varphi(m)}{m} \sum_{j=1}^{m-1} p_{n,n-j} \tag{2.5}$$

for large enough n. We have

$$\begin{aligned} &\sum_{m=\ell+1}^{n} \frac{p_{m,\ell}\varphi(m)}{m} \sum_{j=1}^{m-1} p_{n,n-j} \\ &\quad\ge \sum_{m=\lfloor \theta n\rfloor}^{n} \frac{p_{m,\ell}\varphi(m)}{m} \sum_{j=1}^{m-1} p_{n,n-j} \\ &\quad\ge \inf_{\lfloor \theta n\rfloor \le m \le n} \left(\binom{m}{\ell+1} \frac{\varphi(m)}{m\lambda_m} \right) \lambda_{n,\ell+1} \sum_{k=\lfloor \theta n\rfloor}^{n} \sum_{j=1}^{k-1} p_{n,n-j}. \end{aligned} \tag{2.6}$$

Since $\lambda_n \to \mathtt{m}_{-2} \in (0,\infty)$, the sequence $(\binom{m}{\ell+1}\frac{\varphi(m)}{m\lambda_m})$ is regularly varying with index ℓ. Therefore,

$$\inf_{\lfloor \theta n\rfloor \le m \le n} \left(\binom{m}{\ell+1} \frac{\varphi(m)}{m\lambda_m} \right) \ge C'_\ell \binom{n}{\ell+1} \frac{\varphi(n)}{n} \tag{2.7}$$

for some $C'_\ell > 0$ and large enough n, by [8] (Theorem 1.5.3). Finally,

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{k=\lfloor \theta n\rfloor}^{n} \sum_{j=1}^{k-1} p_{n,n-j} > 0 \tag{2.8}$$

is a consequence of

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{k=\lfloor \theta n\rfloor}^{n} \sum_{j=1}^{k-1} p_{n,n-j} \ge (1-\theta) \liminf_{n\to\infty} \sum_{j=1}^{\lfloor \theta n\rfloor - 1} p_{n,n-j} > 0,$$

where the last inequality follows from (2.3) and our choice of θ.

Inequality (2.5) now follows from (2.6), (2.7), (2.8), Definition (2.1) and the convergence of λ_n.

Next, we argue that the series $\sum_{m\le\ell+1}\frac{p_{m,\ell}\varphi(m)}{m}$ converges. Indeed, since $\lambda_n = O(1)$ we can find some $C''_\ell > 0$ such that

$$\sum_{m\ge\ell+1}\frac{p_{m,\ell}\varphi(m)}{m} \le C''_\ell \sum_{m\ge\ell+1} m^\ell\varphi(m)\lambda_{m,\ell+1}$$
$$= C''_\ell\int_{[0,1]} x^{\ell-1}(1-x)^{-\ell-1}\Big(\sum_{m\ge\ell+1} m^\ell\varphi(m)(1-x)^m\Big)\Lambda(\mathrm{d}x).$$

By dominated convergence the integrand is bounded in some left vicinity of 1. By the Abelian theorem for power series [see, e.g., [21] (Chapter XIII.5, Theorem 5)],

$$\sum_{m\ge\ell+1} m^\ell\varphi(m)(1-x)^m \sim \text{const}\cdot x^{-\ell-1}\varphi(1/x), \qquad x\to 0+$$

which in combination with (2.4) shows that the integral converges in some right vicinity of 0. □

Now we are in position to prove Theorem 2.1.

PROOF OF THEOREM 2.1.

CASE $\mu < \infty$. According to the Cramér–Wold device it is enough to show that

$$Z_n^{(N)} := \sum_{k=2}^N \alpha_k X_{n,k} \xrightarrow{\mathrm{d}} \sum_{k=2}^N \alpha_k X_{\infty,k}, \qquad n\to\infty$$

for every choice of $N\ge 2$ and nonnegative reals $\alpha_2,\dots,\alpha_N$. Observe the stochastic recurrence

$$Z_1^{(N)} = 0, \qquad Z_n^{(N)} \stackrel{\mathrm{d}}{=} \sum_{k=2}^N \alpha_k \mathbb{1}_{\{I_n=k-1\}} + \widehat{Z}^{(N)}_{n-I_n}, \qquad n\ge 2,$$

where I_n is the size of the first decrement of N_n, $\widehat{Z}_m \stackrel{\mathrm{d}}{=} Z_m$, and I_n is independent of $\widehat{Z}_m$'s. Let A be the span of $\alpha_2,\dots,\alpha_N$ with nonnegative integer weights. For $z\in A$, we get

$$\mathbb{P}\{Z_n^{(N)} = z\} = \sum_{m=1}^{n-1} p_{n,n-m}\mathbb{P}\Big\{Z_m^{(N)} = z - \sum_{k=2}^N \alpha_k\delta_{n-m,k-1}\Big\}$$
$$= \sum_{m=1}^{n-1} p_{n,n-m}\mathbb{P}\{Z_m^{(N)} = z\}$$
$$+ \sum_{m=1}^{n-1} p_{n,n-m}\Big(\mathbb{P}\Big\{Z_m^{(N)} = z - \sum_{k=2}^N \alpha_k\delta_{n-m,k-1}\Big\} - \mathbb{P}\{Z_m^{(N)} = z\}\Big)$$

$$
\begin{aligned}
&= \sum_{m=1}^{n-1} p_{n,n-m}\mathbb{P}\{Z_m^{(N)} = z\} \\
&\quad + \sum_{m=n-N+1}^{n-1} p_{n,n-m}\big(\mathbb{P}\{Z_m^{(N)} = z - \alpha_{n-m+1}\} - \mathbb{P}\{Z_m^{(N)} = z\}\big) \\
&= \sum_{m=1}^{n-1} p_{n,n-m}\mathbb{P}\{Z_m^{(N)} = z\} \\
&\quad + \sum_{m=2}^{N} p_{n,m-1}\big(\mathbb{P}\{Z_{n-m+1}^{(N)} = z - \alpha_m\} - \mathbb{P}\{Z_{n-m+1}^{(N)} = z\}\big) \\
&=: \sum_{m=1}^{n-1} p_{n,n-m}\mathbb{P}\{Z_m^{(N)} = z\} + s_{n,z}^{(N)},
\end{aligned}
$$

where $\delta_{x,y}$ is the Kronecker delta.

Interpreting, for $1 \le k \le n$ and fixed $z \in A$, $s_{k,z}^{(N)}$ as a "reward" collected by the block-counting process N_n on visiting state k, we conclude that the total expected reward of N_n is $\mathbb{P}\{Z_n^{(N)} = z\}$. For $k \le n$ let $h(n,k)$ be the probability that N_n ever visits state k. Since N_n is a.s. decreasing, $h(n,k)$ is also the expected number of visits of N_n to k, whence

$$
\mathbb{P}\{Z_n^{(N)} = z\} = \delta_{z,0} + \sum_{k=2}^{n} h(n,k)s_{k,z}^{(N)}, \qquad z \in A,
$$

where the initial condition $\mathbb{P}\{Z_1^{(N)} = z\} = \delta_{z,0}$ has to be recalled. Formally, the last formula is verified by induction on n. For L_n the number of blocks involved in the last collision, we have

$$
\mathbb{P}\{L_n = k\} = h(n,k)p_{k,k-1}.
$$

Thus,

$$
\mathbb{P}\{Z_n^{(N)} = z\} = \delta_{z,0} + \mathbb{E}f_z^{(N)}(L_n), \qquad z \in A,
$$

where

$$
f_z^{(N)}(m) := \frac{s_{m,z}^{(N)}}{p_{m,m-1}}, \qquad m \ge 2.
$$

By [40] (Theorem 2), L_n converges in distribution to a random variable L_∞. Hence,

$$
\lim_{n\to\infty} \mathbb{P}\{Z_n^{(N)} = z\} = \delta_{z,0} + \mathbb{E}f_z^{(N)}(L_\infty), \qquad z \in A,
$$

provided that the family $(f_z^{(N)}(L_n))_{n\in\mathbb{N}}$ is uniformly integrable. In view of

$$|f_z^{(N)}(m)| \le \frac{\sum_{k=1}^{N-1} p_{m,k}}{p_{m,m-1}}, \qquad m \ge 2,$$

it suffices to check uniform integrability of $(g_1^{(k)}(L_n))_{n\in\mathbb{N}}$ for every fixed $k \in \mathbb{N}$, where $g_1^{(k)}(m) := p_{m,k}/p_{m,m-1}$. According to the Vallée–Poussin criterion in the form given in [47] (Lemma 3, page 267) the latter is secured by

$$\sup_{n\in\mathbb{N}} \mathbb{E}\big(g_1^{(k)}(L_n)\varphi\big(\log^+ g_1^{(k)}(L_n)\big)\big) < \infty \tag{2.9}$$

with a nondecreasing slowly varying function φ as given in Corollary 2.3. We will show even more, namely that

$$\sup_{n\in\mathbb{N}} \mathbb{E}g_2^{(k)}(L_n) < \infty, \tag{2.10}$$

where $g_2^{(k)}(m) := \frac{p_{m,k}}{p_{m,m-1}}\varphi(|\log p_{m,m-1}|)$. The expectation under the supremum can be written as

$$\mathbb{E}g_2^{(k)}(L_n) = \sum_{m=2}^{n} h(n,m)\, p_{m,k}\varphi\big(|\log p_{m,m-1}|\big).$$

We use Jensen's inequality

$$\begin{aligned} -\log p_{m,m-1} &= -\log \frac{\int_{[0,1]} x^{m-2}\Lambda(\mathrm{d}x)}{\Lambda([0,1])} - \log\Lambda\big([0,1]\big) + \log\lambda_m \\ &\le (m-2)\frac{\int_{[0,1]} |\log x|\Lambda(\mathrm{d}x)}{\Lambda([0,1])} - \log\Lambda\big([0,1]\big) + \log\lambda_m \end{aligned}$$

in combination with $\lim_{m\to\infty}\lambda_m = \mathtt{m}_{-2}$ to conclude that

$$-\log p_{m,m-1} \le \text{const}\cdot m, \qquad m \ge 2$$

and thereupon

$$\mathbb{E}g_2^{(k)}(L_n) \le \text{const}\sum_{m=2}^{n} h(n,m)\, p_{m,k}\varphi(m)$$

by monotonicity of φ. The last sum is uniformly bounded by Lemma 2.4.

CASE $\mu = \infty$. This follows immediately from the observation

$$\lim_{n\to\infty} h(n,m) = 0, \qquad m \in \mathbb{N}$$

which is a consequence of [40] (Theorem 3). □

REMARK. Under the assumptions of Theorem 2.1 we also have convergence of all joint moments: for every $N \geq 2$ and $m_2, m_3, \ldots, m_N \geq 0$,

$$\lim_{n\to\infty} \mathbb{E}(X_{n,2}^{m_2} X_{n,3}^{m_3} \cdots X_{n,N}^{m_N}) = \mathbb{E}(X_{\infty,2}^{m_2} X_{\infty,3}^{m_3} \cdots X_{\infty,N}^{m_N}).$$

In order to see this, it is enough to check that

$$\sup_{n\geq 2} \mathbb{E} X_{n,k}^m < \infty \tag{2.11}$$

for all $k \geq 2$ and all $m \in \mathbb{N}$. Uniform integrability of $(X_{n,2}^{m_2} X_{n,3}^{m_3} \cdots X_{n,N}^{m_N})_{n\geq 2}$ will then follow from Hölder's inequality. While condition (2.11) for $m = 1$ follows from the recursion

$$\mathbb{E} X_{n,k} = p_{n,k-1} + \sum_{j=1}^{n-k} p_{n,n-j} \mathbb{E} X_{j,k}, \qquad n \geq 2,$$

for $m \geq 2$ it is checked by induction. We omit the details.

Theorem 2.1 is a pure existence result. Nevertheless, for beta$(3, b)$-coalescents, it is possible to describe the asymptotic joint distribution of the collision spectrum explicitly. This result is strikingly similar to the classical Poisson limit for the small-block counts of Ewens' partitions [4].

EXAMPLE (Theorem 3.1 in [41]). Suppose Λ is a beta$(3, b)$-distribution. Then

$$(X_{n,k})_{k\geq 2} \Longrightarrow (X_{\infty,k})_{k\geq 2}, \qquad n \to \infty,$$

where $(X_{\infty,k})_{k\geq 2}$ are independent with $X_{\infty,k} \stackrel{\mathrm{d}}{=} \text{Poisson}(b/(k-1))$.

3. Simple coalescents: Functional limits. A consequence of Theorem 2.1 is that the contribution of $X_{n,k}$ to X_n for every fixed k remains bounded as n grows. In this section, we prove functional limit theorems for the process of cumulative counts $(X_n(s))_{s\in[0,1]}$ defined by (1.4). The cases of finite and infinite μ are treated separately (Theorems 3.1 and 3.2).

The process $(X_n(s))_{s\in[0,1]}$ has paths belonging to the Skorohod space $D[0, 1]$ of càdlàg functions. We endow $D[0, 1]$ with either the J_1- or the M_1-topology and denote the associated weak convergence of probability measures by $\stackrel{J_1}{\Longrightarrow}$ and $\stackrel{M_1}{\Longrightarrow}$, respectively.

THEOREM 3.1. *Assume* $\mathtt{m}_{-2} < \infty$ *and that for some* $a > 0$

$$\mathbb{E}|\log(1-W)|^a = \frac{1}{\mathtt{m}_{-2}} \int_{[0,1]} |\log x|^a x^{-2} \Lambda(\mathrm{d}x) < \infty. \tag{3.1}$$

Suppose further that $\mu < \infty$ *and set*

$$u_n(s) := \mu^{-1} \int_{(1-s)\log n}^{\log n} \mathbb{P}\{|\log(1-W)| \le y\}\,\mathrm{d}y,$$

for $s \in [0,1]$.

(A1) *If* $\sigma^2 < \infty$, *then as* $n \to \infty$

$$\left(\frac{X_n(s) - u_n(s)}{\sqrt{\mu^{-3}\sigma^2 \log n}}\right)_{s\in[0,1]} \stackrel{J_1}{\Longrightarrow} \big(B(s)\big)_{s\in[0,1]},$$

where $(B(s))_{s\in[0,1]}$ *is a standard Brownian motion.*

(A2) *If* $\sigma^2 = \infty$ *and*

$$\mathbb{E}(\log W)^2 \mathbb{1}_{\{|\log W| \le x\}} \sim \ell(x), \qquad x \to \infty,$$

for some function ℓ *slowly varying at infinity, then as* $n \to \infty$

$$\left(\frac{X_n(s) - u_n(s)}{\mu^{-3/2} c(\log n)}\right)_{s\in[0,1]} \stackrel{J_1}{\Longrightarrow} \big(B(s)\big)_{s\in[0,1]},$$

where c *is a positive function satisfying* $\lim_{x\to\infty}(c(x))^{-2} x\ell(c(x)) = 1$.

(A3) *If*

$$\mathbb{P}\big\{|\log W| > x\big\} \sim x^{-\alpha}\ell(x), \qquad x \to \infty, \tag{3.2}$$

for some $\alpha \in (1,2)$ *and some* ℓ *slowly varying at infinity, then as* $n \to \infty$

$$\left(\frac{X_n(s) - u_n(s)}{\mu^{-(\alpha+1)/\alpha} c(\log n)}\right)_{s\in[0,1]} \stackrel{M_1}{\Longrightarrow} \big(S_\alpha(s)\big)_{s\in[0,1]},$$

where c *is a positive function satisfying* $\lim_{x\to\infty}(c(x))^{-\alpha} x\ell(c(x)) = 1$ *and* $(S_\alpha(s))_{s\in[0,1]}$ *is a spectrally negative* α*-stable Lévy process such that* $S_\alpha(1)$ *has the characteristic function*

$$z \mapsto \exp\big\{-|z|^\alpha \Gamma(1-\alpha)\big(\cos(\pi\alpha/2) + i\sin(\pi\alpha/2)\operatorname{sgn}(z)\big)\big\}, \qquad z \in \mathbb{R} \tag{3.3}$$

with Γ *being the gamma function.*

Without moment condition (3.1), the conclusions of Theorem 3.1 are still valid in the weaker sense of the convergence of the finite-dimensional distributions of $(X_n(s))_{s\in[0,1]}$; see [3] (Remark 2.3).

THEOREM 3.2. *Assume* $\mathtt{m}_{-2} < \infty$. *If relation* (3.2) *holds with* $\alpha \in (0,1)$, *then as* $n \to \infty$

$$\left(\frac{\ell(\log n) X_n(s)}{(\log n)^\alpha}\right)_{s\in[0,1]} \stackrel{J_1}{\Longrightarrow} \big(W_\alpha^{\leftarrow}(1) - W_\alpha^{\leftarrow}(1-s)\big)_{s\in[0,1]}, \tag{3.4}$$

where $W_\alpha^{\leftarrow}(s) := \inf\{y \ge 0 : W_\alpha(y) > s\}$ *for* $s \ge 0$ *and* $(W_\alpha(y))_{y\ge 0}$ *is an* α*-stable subordinator* (*nondecreasing Lévy process*) *with the Laplace exponent* $-\log \mathbb{E}(-zW_\alpha(1)) = \Gamma(1-\alpha)z^\alpha$, $z \ge 0$.

For simple Λ-coalescents, each partition $\Pi_\infty(t)$ has a dust component. Therefore, a coupling between the infinite coalescent and a subordinator (which is a compound Poisson process in the case $\mathtt{m}_{-2} < \infty$) [24, 27] can be applied to relate N_n with a simpler counting process derived from the dust component. We briefly summarise the combinatorial part of this connection.

Consider an *extended* coalescent, which is a process on partitions of $[n]$ where every initial singleton block $\{1\}, \dots, \{n\}$ is regarded as *primary* and every other block as *secondary*. We stress that a singleton block may become at a later stage a secondary block. Whenever the partition has m blocks (some of which being primary and some secondary), every k-tuple of the blocks is merging in one block at rate $\lambda_{m,k}$ (for $2 \le k \le m$), and every primary block transforms into secondary at rate $\lambda_{m,1}$. The rate $\lambda_{m,1}$ is defined by formula (1.1) with $k = 1$, and we have $0 < \lambda_{m,1} < \infty$ because $\mathtt{m}_{-1} < \infty$. Let $N_n^*(t)$ be the number of primary blocks at time t, and let $K_{n,k}$ be the number of decrements of size $k \in [n]$ of the process $N_n^* := (N_n^*(t))_{t \ge 0}$. With the extended coalescent, we may associate a *partition of* $[n]$ *by the first event*, which has two integers i and $j \ne i$ in the same block if the first event involving $\{i\}$ (which could have been collision or transformation into secondary block) is a collision which is also the first event for $\{j\}$. The number of blocks of size k of this partition is then $K_{n,k}$. With the natural time-ordering of the blocks, the partition by the first event is a regenerative composition as introduced in [28]. Figure 1 represents a realisation of an extended coalescent with seven initial blocks.

Let Z_n be the number of secondary blocks emerging throughout the history of the extended coalescent. Clearly, $Z_n \le K_{n,1} + X_n$, since a secondary block results from either a collision or a transformation of a primary block, while the number of transformations into a secondary block does not exceed $K_{n,1}$. Choosing $c_0 > \mu^{-1}$ we have

$$\mathbb{P}\{Z_n > c_0 \log n\} \to 0, \qquad n \to \infty, \tag{3.5}$$

as follows from [24] (Proposition 5.1 and Theorem 5.1) in conjunction with [23] (Corollary 1.1).

Similarly to [24] [formula (5.3)], the number of collisions involving at most $\lfloor n^s \rfloor$ blocks can be decomposed as

$$X_n(s) = X_n^*(s) + D_n(s), \tag{3.6}$$

where, for $s \in [0, 1]$, $X_n^*(s)$ is the number of collisions involving between 2 and $\lfloor n^s \rfloor$ blocks such that at least two of them are primary, and $D_n(s)$ is the number of collisions involving at most $\lfloor n^s \rfloor$ blocks with at most one block being primary. Trivially, $D_n(s) \le D_n$ where $D_n := D_n(1)$. Furthermore,

$$\sum_{k=2}^{\lfloor n^s \rfloor - Z_n} K_{n,k} \le X_n^*(s) \le \sum_{k=2}^{\lfloor n^s \rfloor} K_{n,k},$$

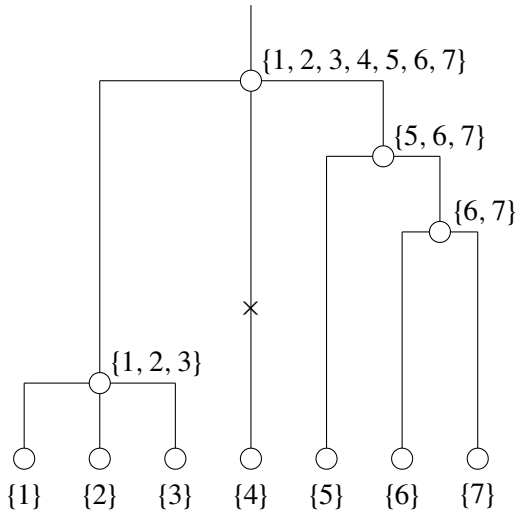

FIG. 1. *A sample path of* Π_7 *with four collisions and one transformation* (*denoted by* $\times$) *of the primary block* 4 *into the secondary block. The process* N_n *has transitions* $7 \to 5 \to 4 \to 3 \to 1$, *thus* $X_7 = 4$, $X_{7,2} = X_{7,3} = 2$, *the process* N_n^* *has transitions* $7 \to 4 \to 3 \to 1 \to 0$, *thus* $K_{7,1} = 2$, $K_{7,2} = 1$, $K_{7,3} = 1$.

where the upper bound is obvious, and the lower bound follows since a decrement of N_n^* of size at least two and at most $\lfloor n^s \rfloor - Z_n$ occurs by a collision of size at most $\lfloor n^s \rfloor$ involving at least two primary particles.

Denoting $K_n(s) := \sum_{k=1}^{\lfloor n^s \rfloor} K_{n,k}$ for $s \in [0,1]$ and combining the aforementioned estimates we arrive at

$$\sup_{s\in[0,1]} \bigl|X_n(s) - K_n(s)\bigr| \le \sup_{s\in[0,1]} \Biggl(\sum_{k=\lfloor n^s - Z_n \rfloor \vee 1}^{\lfloor n^s \rfloor} K_{n,k} \Biggr) + D_n + K_{n,1}. \tag{3.7}$$

The analogues of Theorems 3.1 and 3.2 have been proved for $K_n(s)$; see [3] (Theorems 2.2 and 2.5). Hence, it is sufficient to show that for every fixed $\varepsilon > 0$

$$\frac{\sup_{s\in[0,1]} (\sum_{k=\lfloor (n^s - Z_n)\vee 1 \rfloor}^{\lfloor n^s \rfloor} K_{n,k}) + D_n + K_{n,1}}{(\log n)^{\varepsilon}} \xrightarrow{\mathbb{P}} 0, \qquad n \to \infty. \tag{3.8}$$

According to [24], the sequence $(D_n + K_{n,1})_{n\in\mathbb{N}}$ is tight (see Lemma 5.1, Proposition 5.1 and also the proof of Theorem 5.1 therein). Appealing to (3.5), we see that (3.8) will follow from

$$\frac{\sup_{s\in[0,1]} (\sum_{k=\lfloor (n^s - c_0 \log n)\vee 1 \rfloor + 1}^{\lfloor n^s \rfloor} K_{n,k})}{(\log n)^{\varepsilon}} \xrightarrow{\mathbb{P}} 0, \qquad n \to \infty. \tag{3.9}$$

It remains to prove (3.9). Let $(W_k)_{k\in\mathbb{N}}$ be a sequence of independent copies of the random variable W with distribution (2.2),

$$P_k := W_1 \cdots W_{k-1}(1 - W_k), \qquad k \in \mathbb{N} \quad \text{and}$$
$$\rho(x) := \#\{k \in \mathbb{N} : P_k \geq 1/x\}, \qquad x \geq 0.$$

We will use the following lemma borrowed from [3] [formula (35) and Lemma 6.1].

LEMMA 3.3. *We have, for every fixed* $\varepsilon > 0$,

$$\frac{\sup_{s\in[0,1]} |K_n(s) - (\rho(n) - \rho(n^{(1-s)}))|}{(\log n)^{\varepsilon}} \xrightarrow{\mathbb{P}} 0, \qquad n \to \infty.$$

Set

$$w_n(s) := \frac{\log^+(n^s - c_0 \log n)}{\log n} \quad \text{and} \quad s_0(n) := \frac{\log(1 + c_0 \log n)}{\log n}, \qquad s \in [0, 1]$$

and note that

$$\begin{aligned} &\sup_{s\in[0,1]} \big|(1-s)\log n - \big(1 - w_n(s)\big)\log n\big| \\ &\quad = \sup_{s\in[0,1]} \big|s \log n - \log^+\big(n^s - c_0 \log n\big)\big| \\ &\quad = \max\left\{ s_0(n)\log n, \sup_{s\in[s_0(n),1]} \left|\log\left(1 - \frac{c_0 \log n}{n^s}\right)\right|\right\} \\ &\quad = \log(1 + c_0 \log n). \end{aligned} \tag{3.10}$$

Furthermore, (3.9) is equivalent to

$$\frac{\sup_{s\in[0,1]}(K_n(s) - K_n(w_n(s)))}{(\log n)^{\varepsilon}} \xrightarrow{\mathbb{P}} 0, \qquad n \to \infty, \tag{3.11}$$

and applying Lemma 3.3 we see that (3.11) follows if one can show that

$$\frac{\sup_{s\in[0,1]} |\rho(n^{1-w_n(s)}) - \rho(n^{1-s})|}{(\log n)^{\varepsilon}} \xrightarrow{\mathbb{P}} 0, \qquad n \to \infty.$$

By (3.10), the left-hand side of the last formula satisfies

$$\begin{aligned} &\frac{\sup_{s\in[0,1]} |\rho(e^{(1-w_n(s))\log n}) - \rho(e^{(1-s)\log n}))|}{(\log n)^{\varepsilon}} \\ &\quad \leq \frac{\sup_{s\in[0,1]}(\rho(e^{s\log n + \log(1+c_0\log n)}) - \rho(e^{s\log n}))}{(\log n)^{\varepsilon}}. \end{aligned}$$

Letting $b(t) := \log(1 + c_0 t)$ the right-hand side of the latter inequality is bounded from above by

$$\frac{\sum_{k=0}^{\lfloor b(\log n)\rfloor} \sup_{s\in[0,1]}(\rho(e^{s\log n+k+1}) - \rho(e^{s\log n+k}))}{(\log n)^{\varepsilon}}$$
$$\le \big([b(\log n)] + 1\big)\frac{\sup_{s\in[0,1]}(\rho(e^{(c+1)s\log n+1}) - \rho(e^{(c+1)s\log n}))}{(\log n)^{\varepsilon}},$$

where the last estimate follows from $b(t) \le c_0 t$ and

$$\sup_{s\in[0,1]}\big(\rho\big(e^{s\log n+k+1}\big) - \rho\big(e^{s\log n+k}\big)\big)$$
$$\le \sup_{s\in[0,1]}\big(\rho\big(e^{(c_0+1)s\log n+1}\big) - \rho\big(e^{(c_0+1)s\log n}\big)\big)$$

for $0 \le k \le b(\log n)$. Finally, according to [3] (Proposition 3.3)

$$\frac{\sup_{s\in[0,1]}(\rho(e^{(c_0+1)s\log n+1}) - \rho(e^{(c_0+1)s\log n}))}{((c\log n)^{\varepsilon/2}} \xrightarrow{\mathbb{P}} 0, \qquad n\to\infty,$$

whence (3.9).

4. Other coalescents with dust. We turn to the coalescents with $\mathtt{m}_{-2} = \infty$, yet $\mathtt{m}_{-1} < \infty$. In this case, the collision times of Π_∞ can be identified with the jump times of a subordinator $(S_t)_{t\ge 0}$ with the Laplace exponent

$$\Phi(z) = \log \mathbb{E}e^{-zS_1} = \int_{[0,1]} \big(1 - (1-x)^z\big)x^{-2}\Lambda(\mathrm{d}x).$$

The dust component has frequency $\exp(-S_t)$, that is for large n the partition $\Pi_n(t)$ has about $n\exp(-S_t)$ primary singleton blocks.

We wish to approximate $X_{n,k}$ by $K_{n,k}$, the number of collisions which involve k primary blocks (and possibly some secondary). Let $D_{n,k} := X_{n,k} - K_{n,k}$. Our main tool is the following estimate.

LEMMA 4.1. *For $k = 2, 3, \ldots$*

$$\mathbb{E}|D_{n,k}| \le c_k \sum_{j=1}^{n}\Big(\frac{\Phi(j)}{j}\Big)^2, \tag{4.1}$$

where c_k is a positive constant.

PROOF. We modify the argument for X_n in [24] (Section 5.1). Let $X_{n,k+}$ be the number of collisions involving at least k blocks, and let $K_{n,k+}$ be the number of collisions involving at least k primary blocks. These variables are easier to compare, because $X_{n,k+} \ge K_{n,k+}$ and $d_{n,k+} := \mathbb{E}(X_{n,k+} - K_{n,k+}) \ge 0$.

Decomposing $X_{n,k+}$ at the first collision event of Π_n, we get

$$X_{1,k+} = 0, \qquad X_{n,k+} \stackrel{\mathrm{d}}{=} \widehat{X}_{n-I_n,k+} + \mathbb{1}_{\{I_n \geq k-1\}}, \qquad k \geq 2 \tag{4.2}$$

with the usual convention about the variables in the right-hand side, and I_n having distribution (2.1). In the same way, restricting the coalescent to the set of the primary blocks

$$K_{0,k+} = K_{1,k+} = 0, \qquad K_{n,k+} \stackrel{\mathrm{d}}{=} \widehat{K}_{n-I_n-1,k+} + \mathbb{1}_{\{I_n \geq k-1\}}, \qquad k \geq 2. \tag{4.3}$$

To make the right-hand sides comparable, we need to adjust for -1 in recursion (4.3). To that end, we focus on the singleton $\{n\}$ in the evolution of Π_n, and identify Π_{n-1} with the restriction of Π_n to $[n-1]$. Thus $X_{n-1,k+}$ is realised as the count of mergers of at least k blocks for the restricted process. The first collision involving $\{n\}$ occurs at some random time, say τ, and for $\tau \geq t$ the partitions $\Pi_n(t)$ and $\Pi_{n-1}(t)$ have the same number of blocks, that is, $N_n(t) = N_{n-1}(t)$ for $\tau \geq t$. Now, if $\{n\}$ at time τ is merged together with $k-1$ other blocks (a k-merger), then $X_{n,k+} = X_{n-1,k+} + 1$; otherwise, $X_{n,k+} = X_{n-1,k+}$. We see that $X_{n-1,k}$ and $X_{n,k}$ differ only in the case when the first collision taking $\{n\}$ is a k-merger, hence involving at most k primary blocks. It follows that $X_{n,k+} \leq X_{n-1,k+} + Y_{n,k}$, where $Y_{n,k}$ is the indicator of the event that the first collision with $\{n\}$ takes at most k primary blocks.

From (4.2), we now conclude that

$$X_{n,k+} \stackrel{\mathrm{d}}{\leq} X_{n-I_n-1,k+} + Y_{n-I_n,k} + \mathbb{1}_{\{I_n \geq k-1\}}, \tag{4.4}$$

where the inequality is meant in the sense of stochastic order. With $y_{n,k} := \mathbb{E}Y_{n,k} = \mathbb{P}\{Y_{n,k} = 1\}$, taking expectations in (4.3) and (4.4), and subtracting the first relation from the second we obtain

$$d_{0,k+} = d_{1,k+} = 0, \qquad d_{n,k+} \leq \sum_{j=1}^{n-1} p_{n,n-j}(d_{j-1,k+} + y_{j,k}), \qquad n \geq 2.$$

Consider the exchangeable partition of $[n]$ by the first collision event, where two primary singletons i and j are assigned to the same block if the first collision engaging one of these elements also takes the other. In these terms, $y_{n,k}$ is the probability that in this partition the block containing element n has size at most k. Since there are $\sum_{r=1}^{k} r K_{n,r}$ elements in blocks of size at most k, and since by exchangeability given the block sizes the distribution of partition is uniform, we obtain

$$y_{n,k} = \frac{1}{n} \sum_{r=1}^{k} r \mathbb{E}K_{n,r}.$$

Following the same line as in [24] (Section 5.1) we obtain for some positive constants c_r, c', c_{k+},

$$\mathbb{E}K_{n,r} \leq c_r \Phi(n),$$

whence $y_{n,k} \le c'\Phi(n)$ and, therefore (see [24]),

$$d_{n,k+} \le c_{k+} \sum_{j=1}^{n} \Big(\frac{\Phi(j)}{j}\Big)^2.$$

Using $X_{n,k} = X_{n,k+} - X_{n,(k-1)+}$ and $K_{n,k} = K_{n,k+} - K_{n,(k-1)+}$, we decompose the difference in question as

$$D_{n,k} = (X_{n,k+} - K_{n,k+}) - (X_{n,(k-1)+} - K_{n,(k-1)+}).$$

Since the differences in parantheses are both nonnegative, taking expectations and applying the triangle inequality we obtain

$$\mathbb{E}|D_{n,k}| \le d_{n,k+} + d_{n,(k-1)+}.$$

The desired estimate (4.1) easily follows with constant $c_k = c_{k+} + c_{(k-1)+}$. □

In [24], we used a coupling with a subordinator to derive a limit law for X_n under the condition of regular variation

$$\int_{[x,1]} y^{-2}\Lambda(\mathrm{d}y) \sim x^{-\gamma}\ell(1/x), \qquad x \to 0+ \tag{4.5}$$

for some $\gamma \in (0, 1)$ and a function ℓ slowly varying at infinity. Specifically,

$$\frac{X_n}{n^\gamma \ell(n)} \xrightarrow{\mathrm{d}} \Gamma(2-\gamma)\mathcal{E}_\gamma, \tag{4.6}$$

where the random variable

$$\mathcal{E}_\gamma := \int_0^\infty e^{-\gamma S_t}\,\mathrm{d}t$$

is known as an exponential functional of a subordinator. The following result is the extension for the collision spectrum.

THEOREM 4.2. *If condition* (4.5) *holds, then as* $n \to \infty$,

$$\frac{1}{n^\gamma \ell(n)}(X_{n,k})_{k\ge 2} \quad \Longrightarrow \quad \Big(\frac{\gamma\Gamma(k-\gamma)}{k!}\mathcal{E}_\gamma\Big)_{k\ge 2} \tag{4.7}$$

weakly in the product space $\mathbb{R}^\infty$. *Moreover, the convergence of joint moments holds*: *for all* $m \ge 2$ *and nonnegative integers* $q_2, \dots, q_m$, *as* $n \to \infty$,

$$\begin{aligned} &\mathbb{E}\prod_{k=2}^{m}\Big(\frac{X_{n,k}}{n^\gamma\ell(n)}\Big)^{q_k} \\ &\quad\to \prod_{k=2}^{m}\Big(\frac{\gamma\Gamma(k-\gamma)}{k!}\Big)^{q_k}\mathbb{E}(\mathcal{E}_\gamma)^{q_2+\cdots+q_m} \\ &\quad= \prod_{k=2}^{m}\Big(\frac{\gamma\Gamma(k-\gamma)}{k!}\Big)^{q_k}\prod_{j=1}^{q_2+\cdots+q_m}\frac{j}{\Phi(\gamma j)}. \end{aligned} \tag{4.8}$$

PROOF. Convergence (4.7) is concluded from the counterpart result for $(K_{n,k})_{k\geq 2}$ viewed in the context of Π_∞. Indeed, by [29] (Theorem 4.1),

$$\frac{K_{n,k}}{n^\gamma \ell(n)} \to \frac{\gamma\Gamma(k-\gamma)}{k!}\mathcal{E}_\gamma \qquad \text{a.s.}$$

On the other hand, changing a variable and integrating by parts followed by an application of Karamata's Tauberian theorem [8] (Theorem 1.7.1') give $\Phi(z) \sim \Gamma(1-\gamma)z^\gamma \ell(z)$ as $z \to \infty$. Using this asymptotics along with (4.1) readily yields $\mathbb{E}D_{n,k} = o(n^\gamma \ell(n))$, whence $|K_{n,k} - X_{n,k}|/(n^\gamma \ell(n)) \xrightarrow{\mathbb{P}} 0$.

The assertion about the convergence of moments (4.8) follows by dominated convergence from the analogous fact for X_n proved in [33], and the familiar formula for moments of $\mathcal{E}_\gamma$ (see, e.g., [29] and references therein). □

Formal summation in (4.7) yields (4.6). We see that with the same scaling the variables X_n, $X_{n,k}$'s all converge in distribution to multiples of $\mathcal{E}_\gamma$.

Theorem 4.2 covers beta(a,b)-coalescents with $a \in (1,2)$. In this case, (4.5) holds with $\gamma = 2-a \in (0,1)$ and constant function $\ell(x) \equiv 1/(\gamma \mathrm{B}(a,b))$. The related exponential functional is denoted by $\mathcal{E}_{2-a}$ (Table 1 in Section 1). Extension to the case $\gamma = 1$ is possible for some slowly varying factors (for instance, $\ell(z) = (\log z)^{-\theta}$ with $\theta > 2$). A further extension concerns the spectrum of coalescent $(\Pi_n(t), t \in [0,T])$ with finite time horizon T; the assertions of Theorem 4.2 hold then with $\mathcal{E}_\gamma$ replaced by the incomplete exponential functional $\int_0^T \exp(-\gamma S_t)\,\mathrm{d}t$. Another edge case is that of slow variation, where (4.5) holds with $\gamma = 0$ and some unbounded function ℓ. In that case, the series (4.1) converges which implies by Lemma 4.1 that the sequence $(D_{n,k})_{n\in\mathbb{N}}$ is tight for every k. Moreover, each $\mathbb{E}X_{n,k}$ is then of smaller order of growth than $\mathbb{E}X_n$. However, the limit laws for small block counts $K_{n,k}$ are only available for functions ℓ of logarithmic growth [30], although there are plentiful results on K_n [5, 22] which have their counterparts for X_n [24] (Section 5.3). Thus we confine ourselves with the framework of [30], assuming that the characteristic measure satisfies

$$\int_{[x,1]} y^{-2}\Lambda(\mathrm{d}y) = |\log x| + c + O\big(x^{-\epsilon}\big), \qquad x \to 0+, \tag{4.9}$$

$$\int_{[x,1]} y^{-2}\Lambda(\mathrm{d}y) = O\big((1-x)^{\epsilon}\big), \qquad x \to 1- \tag{4.10}$$

for some constants $\epsilon > 0$ and c. Introduce the logarithmic moments

$$\nu_r := \int_{[0,1]} \big|\log(1-x)\big|^r x^{-2}\Lambda(\mathrm{d}x), \qquad r = 1,2$$

(so $\nu_1 = \mathbb{E}S_1$ and $\nu_2 = \operatorname{Var} S_1$ for the corresponding subordinator).

Our next result follows from [30] (Theorem 15) and the discussion above.

THEOREM 4.3. *If conditions* (4.9) *and* (4.10) *hold, then*

$$\left(\frac{X_{n,k}-(k\nu_1)^{-1}\log n}{\sqrt{\log n}}\right)_{k\geq 2}\Longrightarrow(\mathcal{N}_k)_{k\geq 2}$$

weakly in $\mathbb{R}^\infty$, *where the limit is a zero-mean Gaussian sequence with the covariance matrix*

$$\left(\frac{\nu_2}{\nu_1^3}\frac{1}{ij}+\delta_{i,j}\frac{1}{j\nu_1}\right)_{i,j\geq 2}.$$

Despite the seemingly limited scope, the theorem covers a number of important cases. For instance, $\Lambda(\mathrm{d}x)=\frac{x^2(1-x)^{\theta-1}}{|\log(1-x)|}\mathbb{1}_{(0,1)}(x)\,\mathrm{d}x$ is the case where $(S_t)_{t\geq 0}$ is the classic gamma-subordinator with parameter $\theta>0$ and the Laplace exponent $\Phi(z)=\log(1+z/\theta)$ (see [24] for constants and the normal limit for X_n). Another example is the beta$(2,b)$-coalescent; in that case the logarithmic moments can be evaluated in terms of the Hurwitz zeta function $\zeta(z,b):=\sum_{i\geq 0}(i+b)^{-z}$ as $\nu_1=\zeta(2,b)$ and $\nu_2=2\zeta(3,b)$.

5. Beta-coalescents without dust component. The Λ-coalescents with $\mathtt{m}_{-1}=\infty$ are very different from the coalescents with dust component, and, as such, require other approaches. The most general available result on the total number of collisions X_n states a stable limit distribution for the coalescents with characteristic measure satisfying $\Lambda([0,x])\sim cx^a$ for $x\to 0+$ [where $c>0$, $a\in(0,1)$] and a condition on the remainder of the expansion at 0; see [31] (Theorem 7). This covers, in particular, all beta(a,b)-coalescents with $0<a<1$. We also know that a stable limit for X_n holds for beta$(1,b)$-coalescents [26].

In what follows, we shall confine ourselves to the family of beta(a,b)-coalescents with $a\in(0,1]$. The qualitative difference between the beta coalescents with $0<a<1$ and $a=1$ is that in the first case $\Pi_\infty(t)$ has finitely many blocks for all $t>0$ (the coalescent 'comes down from infinity') and terminates in finite time, while in the second case the number of blocks always stays infinite.

THEOREM 5.1. *Suppose* Λ *is a* beta(a,b)*-distribution with parameters* $a\in(0,1]$ *and* $b>0$. *Let*

$$p_k^{(a)}:=\frac{(2-a)\Gamma(k+a-1)}{\Gamma(a)(k+1)!},\qquad k\in\mathbb{N},$$

in particular, $p_k^{(1)}=(k^2+k)^{-1}$.

(i) *If* $0<a<1$, *then as* $n\to\infty$,

$$\left(\frac{X_{n,k}-p_{k-1}^{(a)}(1-a)n}{(1-a)n^{1/(2-a)}}\right)_{k\geq 2}\Longrightarrow(p_{k-1}^{(a)}\mathcal{S}_{2-a})_{k\geq 2}$$

weakly in $\mathbb{R}^\infty$, *where* $\mathcal{S}_\alpha$, *for* $\alpha \in (1,2)$, *is an* α*-stable random variable with characteristic function*

$$z \mapsto \exp\{|z|^\alpha(\cos(\pi\alpha/2) + \mathrm{i}\sin(\pi\alpha/2)\operatorname{sgn}(z))\}, \qquad z \in \mathbb{R}.$$

(ii) *If* $a = 1$, *then as* $n \to \infty$,

$$(n^{-1}(\log n)^2 X_{n,k} - p_{k-1}^{(1)}(\log n + \log\log n))_{k\geq 2} \Longrightarrow (p_{k-1}^{(1)}\mathcal{S}_1)_{k\geq 2}$$

weakly in $\mathbb{R}^\infty$, *where* $\mathcal{S}_1$ *is a* 1*-stable random variable with characteristic function*

$$z \mapsto \exp\left(\mathrm{i}z\log|z| - \frac{\pi}{2}|z|\right) = (\mathrm{i}z)^{\mathrm{i}z}, \qquad z \in \mathbb{R}.$$

PROOF. For $m \geq 2$ and $\beta_2, \ldots, \beta_m \in \mathbb{R}$, we consider the linear combinations $q^{(a)} := \sum_{k=2}^m \beta_k p_{k-1}^{(a)}$ and $Z_n := \sum_{k=2}^m \beta_k X_{n,k}$. To apply the Cramér–Wold device, we need to prove that

$$\frac{Z_n - q^{(a)}(1-a)n}{(1-a)n^{1/(2-a)}} \xrightarrow{\mathrm{d}} q^{(a)}\mathcal{S}_{2-a}, \qquad n \to \infty \tag{5.1}$$

in case (i), and that

$$n^{-1}(\log n)^2 Z_n - q^{(1)}(\log n + \log\log n) \xrightarrow{\mathrm{d}} q^{(1)}\mathcal{S}_1, \qquad n \to \infty \tag{5.2}$$

in case (ii). As before, we denote by I_n the number of blocks involved in the first collision of Π_n minus one. Decomposing at the first collision, we obtain a stochastic recurrence

$$Z_1 = 0, \qquad Z_n \stackrel{\mathrm{d}}{=} \sum_{k=2}^m \beta_k \mathbb{1}_{\{I_n = k-1\}} + \widetilde{Z}_{n-I_n}, \qquad n \geq 2, \tag{5.3}$$

where $\widetilde{Z}_k \stackrel{\mathrm{d}}{=} Z_k$ for every $k \in \mathbb{N}$ and $(\widetilde{Z}_k)_{k\in\mathbb{N}}$ is independent of I_n.

It is known (see [13] (Lemma 2.1) for $a \in (0,1)$ and [17] (page 1409) for $a = 1$) that under the assumptions of Theorem 5.1 there exists a distributional limit

$$I_n \xrightarrow{\mathrm{d}} \xi, \qquad n \to \infty, \tag{5.4}$$

where ξ is a random variable with distribution $\mathbb{P}\{\xi = k\} = p_k^{(a)}$, $k \in \mathbb{N}$ and $\mathbb{E}\xi = (1-a)^{-1}$.

Consider an ordinary random walk $(S_j)_{j\in\mathbb{N}_0}$ defined by $S_0 := 0$ and $S_j := \xi_1 + \cdots + \xi_j$ for $j \in \mathbb{N}$, where $\xi_1, \xi_2, \ldots$ are independent copies of ξ, and denote by $T_n := \inf\{j \in \mathbb{N}_0 : S_j \geq n\}$, $n \in \mathbb{N}_0$ the level n first-passage time. For $n \in \mathbb{N}_0$ and $k \in \mathbb{N}$, let

$$J_{n,k} := \sum_{j=1}^{T_n} \mathbb{1}_{\{\xi_j = k\}} = \sum_{j\geq 1} \mathbb{1}_{\{\xi_j = k, S_{j-1} \leq n-1\}}$$

be the number of jumps of size k before the random walk passes level n, and set $Y_n := \sum_{k=2}^m \beta_k J_{n,k-1}$. A standard conditioning argument yields the recursion

$$Y_0 = 0, \qquad Y_n \stackrel{\mathrm{d}}{=} \sum_{k=2}^m \beta_k \mathbb{1}_{\{\xi=k-1\}} + \widetilde{Y}_{n-\xi\wedge n}, \qquad n \in \mathbb{N}, \tag{5.5}$$

where $\widetilde{Y}_k \stackrel{\mathrm{d}}{=} Y_k$ for every $k \in \mathbb{N}_0$ and $(\widetilde{Y}_k)_{k\in\mathbb{N}_0}$ is independent of ξ. Comparing (5.3) and (5.5) and keeping in mind (5.4), we may anticipate that if Y_n, properly centered and normalised, converges in distribution, then the same holds for Z_n. The subsequent proof of this intuition is split in two steps. The first step shows that relations (5.1) and (5.2) hold with Y_n replacing Z_n. The second step derives from this the convergence for Z_n.

STEP 1. The representation

$$\begin{aligned} Y_{n+1} &= \sum_{k=2}^m \beta_k \sum_{j=1}^{T_{n+1}} \mathbb{1}_{\{\xi_j=k-1\}} \\ &= \sum_{k=2}^m \beta_k \sum_{j\ge1} \mathbb{1}_{\{\xi_j=k-1, S_{j-1}\le n\}} \\ &= \sum_{j\ge0} \mathbb{1}_{\{S_j\le n\}} \sum_{k=2}^m \beta_k \mathbb{1}_{\{\xi_{j+1}=k-1\}} \\ &= \sum_{j\ge0} \eta_{j+1} \mathbb{1}_{\{S_j\le n\}}, \qquad n \in \mathbb{N}_0 \end{aligned}$$

with $\eta_j := \sum_{k=2}^m \beta_k \mathbb{1}_{\{\xi_j=k-1\}}$, $j \in \mathbb{N}$, shows that $(Y_n)_{n\in\mathbb{N}_0}$ has the same distribution as random process with immigration in the sense of [35] [formula (1)] with random but constant response process $\eta := \sum_{k=2}^m \beta_k \mathbb{1}_{\{\xi=k-1\}}$. Thus, (5.1) with Y_n replacing Z_n follows from [35] (Theorem 2.4 applied with $u = 1$, $\alpha = 2 - a$, $h(t) = \mathbb{E}\eta = q_a$, $\rho = 0$ and $p = 1$). In order to see that (5.2) holds with Y_n replacing by Z_n decompose

$$Y_{n+1} = \sum_{j\ge0} \eta_{j+1} \mathbb{1}_{\{S_j\le n\}} = (\mathbb{E}\eta) \sum_{j\ge0} \mathbb{1}_{\{S_j\le n\}} + \sum_{j\ge0} (\eta_{j+1} - \mathbb{E}\eta) \mathbb{1}_{\{S_j\le n\}}.$$

By [37] (Proposition 2), $n^{-1}(\log n)^2 \sum_{j\ge0} \mathbb{1}_{\{S_j\le n\}} - \log n - \log\log n \stackrel{\mathrm{d}}{\to} \mathcal{S}_1$ as $n \to \infty$. In view of Slutsky's lemma, it would be enough to prove that $n^{-1}(\log n)^2 \sum_{j\ge0} (\eta_{j+1} - \mathbb{E}\eta) \mathbb{1}_{\{S_j\le n\}} \stackrel{\mathbb{P}}{\to} 0$ as $n \to \infty$. The latter convergence even holds in the mean-square sense, since

$$\mathbb{E}\Big(\sum_{j\ge0} (\eta_{j+1} - \mathbb{E}\eta) \mathbb{1}_{\{S_j\le n\}} \Big)^2 = (\operatorname{Var}\eta) \sum_{j\ge0} \mathbb{P}\{S_j \le n\} = o(n)$$

as $n \to \infty$, where the last estimate is a consequence of the elementary renewal theorem.

STEP 2. As in [26], we use probability distances. For $q \in (0, 1]$ and random variables X and Y having finite qth moment, the Wasserstein distance d_q is defined via $d_q(X, Y) = \inf \mathbb{E}|\widehat{X} - \widehat{Y}|^q$, where the infimum is taken over all couplings $(\widehat{X}, \widehat{Y})$ such that $\widehat{X} \stackrel{\mathrm{d}}{=} X$ and $\widehat{Y} \stackrel{\mathrm{d}}{=} Y$. For properties of d_q, we refer the reader to [26] (Proposition 4.1).

To prove (5.1), it suffices to show that

$$d_q(Y_n, Z_n) = o\big(n^{q/(2-a)}\big), \qquad n \to \infty \tag{5.6}$$

for some $q \in (0, 1]$. Likewise, (5.2) will follow from

$$d_q(Y_n, Z_n) = o\big(n^q (\log n)^{-2q}\big), \qquad n \to \infty \tag{5.7}$$

for some $q \in (0, 1]$; see [26] [formulae (5.1) and (5.2)].

From $|J_{i,k} - J_{j,k}| \le |i - j|$ for $i, j \in \mathbb{N}_0$ and $k \in \mathbb{N}$ (the number of k-jumps while the random walks stays in $[i \wedge j - 1, i \vee j - 1)$ is dominated by the number of 1-jumps which is $|i - j|$), it follows that

$$|Y_i - Y_j| \le \|\beta\| |i - j|, \tag{5.8}$$

where $\|\beta\| := \sum_{k=2}^m |\beta_k|$.

Let $(\hat{I}_n, \hat{\xi})$ be a coupling of I_n and ξ such that $d_q(I_n, \xi \wedge n) = \mathbb{E}|\hat{I}_n - \hat{\xi} \wedge n|^q$. Further, let $(\hat{Y}_j)_{j\in\mathbb{N}_0}$ [resp., $(\hat{Z}_j)_{j\in\mathbb{N}}$] be an arbitrary copy of $(Y_j)_{j\in\mathbb{N}_0}$ [resp., $(Z_j)_{j\in\mathbb{N}}$] independent of $(\hat{I}_n, \hat{\xi})$. Using recurrences (5.3) and (5.5) in combination with the inequality $|x + y|^q \le |x|^q + |y|^q$ for $x, y \in \mathbb{R}$, we infer

$$\begin{aligned}
t_n &:= d_q(Y_n, Z_n) \\
&\le \mathbb{E}\left|\sum_{k=2}^m \beta_k \mathbb{1}_{\{\hat{\xi}=k-1\}} + \hat{Y}_{n-\hat{\xi}\wedge n} - \sum_{k=2}^m \beta_k \mathbb{1}_{\{\hat{I}_n=k-1\}} - \hat{Z}_{n-\hat{I}_n}\right|^q \\
&\le \mathbb{E}\left|\sum_{k=2}^m \beta_k (\mathbb{1}_{\{\hat{\xi}=k-1\}} - \mathbb{1}_{\{\hat{I}_n=k-1\}})\right|^q + \mathbb{E}|\hat{Y}_{n-\hat{\xi}\wedge n} - \hat{Z}_{n-\hat{I}_n}|^q \\
&\le \mathbb{E}\left|\sum_{k=2}^m \beta_k (\mathbb{1}_{\{\hat{\xi}=k-1\}} - \mathbb{1}_{\{\hat{I}_n=k-1\}})\right|^q \\
&\quad + \mathbb{E}|\hat{Y}_{n-\hat{\xi}\wedge n} - \hat{Y}_{n-\hat{I}_n}|^q + \mathbb{E}|\hat{Y}_{n-\hat{I}_n} - \hat{Z}_{n-\hat{I}_n}|^q.
\end{aligned}$$

Passing to the infimum over all pairs $((\hat{Y}_j, \hat{Z}_j))_{1\le j\le n-1}$ in the last summand leads to

$$t_n \le c_n + \sum_{k=1}^{n-1} \mathbb{P}\{I_n = k\} t_{n-k}, \qquad n \ge 2$$

with $c_n := \mathbb{E}|\sum_{k=2}^{m}\beta_k(\mathbb{1}_{\{\hat{\xi}=k-1\}}-\mathbb{1}_{\{\hat{I}_n=k-1\}})|^q+\mathbb{E}|\hat{Y}_{n-\hat{\xi}\wedge n}-\hat{Y}_{n-\hat{I}_n}|^q$. Applying (5.8), we obtain

$$\begin{aligned}c_n &\le \|\beta\|^q\big(\mathbb{P}\{\hat{I}_n\neq\hat{\xi}\wedge n\}+\mathbb{E}|\hat{I}_n-\hat{\xi}\wedge n|^q\big)\\ &= \|\beta\|^q\big(\mathbb{P}\big\{|\hat{I}_n-\hat{\xi}\wedge n|^q\ge 1\big\}+\mathbb{E}|\hat{I}_n-\hat{\xi}\wedge n|^q\big)\\ &\le 2\|\beta\|^q\mathbb{E}|\hat{I}_n-\hat{\xi}\wedge n|^q = 2\|\beta\|^q d_q(I_n,\xi\wedge n).\end{aligned}$$

The proof of (5.6) and (5.7) is completed along the lines of the argument in [26] (the bottom of page 504). □

The instance $b=1$ in part (ii) of the theorem is the asymptotics of the collision spectrum for the Bolthausen–Sznitman coalescent. Note that neither the limit law nor the scaling/centering constants depend on b. This suggests that the stable limit for the collision spectrum in part (i) also holds in the more general setting of [31].

REMARK. We conjecture that in the situation of Theorem 5.1 the moments of $X_{n,k}$ differ only little from those of $p_{k-1}^{(a)}X_n$ if n is sufficiently large. While this is open for $a\in(0,1)$, for $a=1$ an adaptation of the method of sequential approximation [see [26] (Section 5.3)] shows that, for every $k\ge 2$ and $j\in\mathbb{N}$,

$$\mathbb{E}X_{n,k}^j = \left(p_{k-1}^{(1)}\frac{n}{\log n}\right)^j\left(1+\frac{m_j}{\log n}+O\left(\frac{1}{\log^2 n}\right)\right),\qquad n\to\infty, \tag{5.9}$$

where the sequence $(m_j)_{j\in\mathbb{N}_0}$ is recursively defined via $m_0:=0$ and $m_j:=m_{j-1}+\kappa_j/j$ for $j\in\mathbb{N}$, with $\kappa_j:=(j+b-1)\Psi(j+b)+j-(b-1)\Psi(b)$, $j\in\mathbb{N}$. Here, Ψ denotes the logarithmic derivative of the gamma function. By [26] (Theorem 3.2), (5.9) coincides with the second order expansion of the jth moment of $p_{k-1}^{(1)}X_n$.

Acknowledgements. We thank two anonymous referees for a careful reading and useful comments which helped improving the presentation.

REFERENCES

[1] ABRAHAM, R. and DELMAS, J.-F. (2013). A construction of a β-coalescent via the pruning of binary trees. *J. Appl. Probab.* **50** 772–790. MR3102514

[2] ABRAHAM, R. and DELMAS, J.-F. (2015). β-coalescents and stable Galton-Watson trees. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** 451–476. MR3368966

[3] ALSMEYER, G., IKSANOV, A. and MARYNYCH, A. (2017). Functional limit theorems for the number of occupied boxes in the Bernoulli sieve. *Stochastic Process. Appl.* **127** 995–1017. MR3605718

[4] ARRATIA, R., BARBOUR, A. D. and TAVARÉ, S. (2003). *Logarithmic Combinatorial Structures: A Probabilistic Approach*. European Mathematical Society (EMS), Zürich. MR2032426

[5] BARBOUR, A. D. and GNEDIN, A. V. (2006). Regenerative compositions in the case of slow variation. *Stochastic Process. Appl.* **116** 1012–1047. MR2238612

[6] BASDEVANT, A.-L. and GOLDSCHMIDT, C. (2008). Asymptotics of the allele frequency spectrum associated with the Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **13** 486–512. MR2386740

[7] BERESTYCKI, J., BERESTYCKI, N. and LIMIC, V. (2014). Asymptotic sampling formulae for Λ-coalescents. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 715–731. MR3224287

[8] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1989). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge Univ. Press, Cambridge. MR1015093

[9] BIRKNER, M. and BLATH, J. (2008). Computing likelihoods for coalescents with multiple collisions in the infinitely many sites model. *J. Math. Biol.* **57** 435–465. MR2411228

[10] BIRKNER, M., BLATH, J. and ELDON, B. (2013). Statistical properties of the site-frequency spectrum associated with Λ-coalescents. *Genetics* **195** 1037–1053.

[11] BLATH, J., CRONJÄGER, M. C., ELDON, B. and HAMMER, M. (2016). The site-frequency spectrum associated with Ξ-coalescents. *Theor. Popul. Biol.* **110** 36–50.

[12] COOPER, C. (2006). Distribution of vertex degree in web-graphs. *Combin. Probab. Comput.* **15** 637–661. MR2248318

[13] DELMAS, J.-F., DHERSIN, J.-S. and SIRI-JEGOUSSE, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.* **18** 997–1025. MR2418236

[14] DIEHL, C. and KERSTING, G. (2018). Tree lengths for general Λ-coalescents and the asymptotic site frequency spectrum around the Bolthausen–Sznitman coalescent. Preprint. Available at https://arxiv.org/abs/1804.00961.

[15] DONG, R., GNEDIN, A. and PITMAN, J. (2007). Exchangeable partitions derived from Markovian coalescents. *Ann. Appl. Probab.* **17** 1172–1201. MR2344303

[16] DRMOTA, M. and GITTENBERGER, B. (1999). The distribution of nodes of given degree in random trees. *J. Graph Theory* **31** 227–253. MR1688949

[17] DRMOTA, M., IKSANOV, A., MOEHLE, M. and ROESLER, U. (2007). Asymptotic results concerning the total branch length of the Bolthausen-Sznitman coalescent. *Stochastic Process. Appl.* **117** 1404–1421. MR2353033

[18] ELDON, B., BIRKNER, M., BLATH, J. and FREUND, F. (2014). Can the site-frequency spectrum distinguish exponential population growth from multiple-merger coalescents? *Genetics* **199** 841–856.

[19] ELDON, B. and WAKELEY, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* **172** 2621–2633.

[20] FAŸ, G., GONZÁLEZ-ARÉVALO, B., MIKOSCH, T. and SAMORODNITSKY, G. (2006). Modeling teletraffic arrivals by a Poisson cluster process. *Queueing Syst.* **54** 121–140. MR2268057

[21] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II.* 2nd ed. Wiley, New York. MR0270403

[22] GNEDIN, A. and IKSANOV, A. (2012). Regenerative compositions in the case of slow variation: A renewal theory approach. *Electron. J. Probab.* **17** Paper no. 77. MR2981902

[23] GNEDIN, A., IKSANOV, A. and MARYNYCH, A. (2010). Limit theorems for the number of occupied boxes in the Bernoulli sieve. *Theory Stoch. Process.* **16(32)** 44–57. MR2777900

[24] GNEDIN, A., IKSANOV, A. and MARYNYCH, A. (2011). On Λ-coalescents with dust component. *J. Appl. Probab.* **48** 1133–1151. MR2896672

[25] GNEDIN, A., IKSANOV, A. and MARYNYCH, A. (2014). Λ-coalescents: A survey. *J. Appl. Probab.* **51A** 23–40. MR3317347

[26] GNEDIN, A., IKSANOV, A., MARYNYCH, A. and MÖHLE, M. (2014). On asymptotics of the beta coalescents. *Adv. in Appl. Probab.* **46** 496–515. MR3215543

[27] GNEDIN, A., IKSANOV, A. and MÖHLE, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Probab.* **45** 1186–1195. MR2484170

[28] GNEDIN, A. and PITMAN, J. (2005). Regenerative composition structures. *Ann. Probab.* **33** 445–479. MR2122798

[29] GNEDIN, A., PITMAN, J. and YOR, M. (2006). Asymptotic laws for compositions derived from transformed subordinators. *Ann. Probab.* **34** 468–492. MR2223948

[30] GNEDIN, A., PITMAN, J. and YOR, M. (2006). Asymptotic laws for regenerative compositions: Gamma subordinators and the like. *Probab. Theory Related Fields* **135** 576–602. MR2240701

[31] GNEDIN, A. and YAKUBOVICH, Y. (2007). On the number of collisions in Λ-coalescents. *Electron. J. Probab.* **12** 1547–1567. MR2365877

[32] GOLDSCHMIDT, C. and MARTIN, J. B. (2005). Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **10** 718–745. MR2164028

[33] HAAS, B. and MIERMONT, G. (2011). Self-similar scaling limits of non-increasing Markov chains. *Bernoulli* **17** 1217–1247. MR2854770

[34] HU, Y., BOYD-GRABER, J., DAUMÉ, H. III and YING, Z. I. (2013). Binary to bushy: Bayesian hierarchical clustering with the beta coalescent. *Adv. Neural Inf. Process. Syst.* **26**.

[35] IKSANOV, A., MARYNYCH, A. and MEINERS, M. (2017). Asymptotics of random processes with immigration I: Scaling limits. *Bernoulli* **23** 1233–1278. MR3606765

[36] IKSANOV, A., MARYNYCH, A. and MÖHLE, M. (2009). On the number of collisions in beta$(2, b)$-coalescents. *Bernoulli* **15** 829–845. MR2555201

[37] 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. Probab.* **12** 28–35. MR2407414

[38] IKSANOV, A. and MÖHLE, M. (2008). On the number of jumps of random walks with a barrier. *Adv. in Appl. Probab.* **40** 206–228. MR2411821

[39] JANSON, S. (2005). Asymptotic degree distribution in random recursive trees. *Random Structures Algorithms* **26** 69–83. MR2116576

[40] KERSTING, G., SCHWEINSBERG, J. and WAKOLBINGER, A. (2018). The size of the last merger and time reversal in Λ-coalescents. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 1527–1555. MR3825890

[41] MÖHLE, M. (2018). On strictly monotone Markov chains with constant hitting probabilities and applications to a class of beta coalescents. *Markov Process. Related Fields* **24** 107–130.

[42] NEHER, R. A. and HALLATSCHEK, A. (2013). Genealogies of rapidly adapting populations. *Proc. Natl. Acad. Sci. USA* **110** 437–442.

[43] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. MR1742892

[44] PITMAN, J. (2006). *Combinatorial Stochastic Processes. Lecture Notes in Math.* **1875**. Springer, Berlin. MR2245368

[45] SAGITOV, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36** 1116–1125. MR1742154

[46] SCHWEINSBERG, J. and DURRETT, R. (2005). Random partitions approximating the coalescence of lineages during a selective sweep. *Ann. Appl. Probab.* **15** 1591–1651. MR2152239

[47] SHIRYAEV, A. N. (1989). *Probability, Vol. I*, 4th ed. Nauka, Moscow.

[48] SPENCE, J. P., KAMM, J. A. and SONG, Y. S. (2016). The site frequency spectrum for general coalescents. *Genetics* **202** 1549–1561.

[49] TEH, Y. W., DAUMÉ, H. III and ROY, D. M. (2008). Bayesian agglomerative clustering with coalescents. *Adv. Neural Inf. Process. Syst.* **20** 1473–1480.

A. GNEDIN
SCHOOL OF MATHEMATICAL SCIENCES
QUEEN MARY UNIVERSITY OF LONDON
LONDON E1 4NS
UNITED KINGDOM
E-MAIL: a.gnedin@qmul.ac.uk

A. IKSANOV
A. MARYNYCH
FACULTY OF COMPUTER SCIENCE AND CYBERNETICS
TARAS SHEVCHENKO NATIONAL UNIVERSITY OF KYIV
01601 KYIV
UKRAINE
E-MAIL: iksan@univ.kiev.ua
marynych@unicyb.kiev.ua

M. MÖHLE
MATHEMATICAL INSTITUTE
EBERHARD KARLS UNIVERSITY OF TÜBINGEN
72076 TÜBINGEN
GERMANY
E-MAIL: martin.moehle@uni-tuebingen.de