On asymptotics of the beta-coalescents

Alexander Gnedin,[*]   Alexander Iksanov,[†]
Alexander Marynych[‡]   and   Martin Möhle[§]

December 11, 2012

Abstract

We show that the total number of collisions in the exchangeable coalescent process driven by the beta $(1, b)$ measure converges in distribution to a 1-stable law, as the initial number of particles goes to infinity. The stable limit law is also shown for the total branch length of the coalescent tree. These results were known previously for the instance $b = 1$, which corresponds to the Bolthausen–Sznitman coalescent. The approach we take is based on estimating the quality of a renewal approximation to the coalescent in terms of a suitable Wasserstein distance. Application of the method to beta (a, b)-coalescents with $0 < a < 1$ leads to a simplified derivation of the known $(2 - a)$-stable limit. We furthermore derive asymptotic expansions for the moments of the number of collisions and of the total branch length for the beta $(1, b)$-coalescent by exploiting the method of sequential approximations.

Keywords: Absorption time; asymptotic expansion; beta-coalescent; coupling; number of collisions; total branch length; Wasserstein distance

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

1   Introduction

Pitman [28] and Sagitov [29] introduced exchangeable coalescent processes with multiple collisions, also known as Λ-coalescents. A counting process associated with the Λ-coalescent is a Markov chain $\Pi_n = \big(\Pi_n(t)\big)_{t \geq 0}$ with right-continuous paths, which starts with n particles $\Pi_n(0) = n$ and terminates when a sole particle remains. The particles

---

[*]Queen Mary, University of London, e-mail: a.gnedin@qmul.ac.uk

[†]Taras Shevchenko National University of Kyiv, e-mail: iksan@univ.kiev.ua
Supported by a grant awarded by the President of Ukraine (project Φ47/431-2012)

[‡]Technical University of Eindhoven, e-mail: O.Marynych@tue.nl
Supported by a grant awarded by the President of Ukraine (project Φ47/431-2012)

[§]University of Tübingen, e-mail: martin.moehle@uni-tuebingen.de

merge according to the rule: for each $t \geq 0$ when the number of particles is $\Pi_n(t) = m > 1$, each k tuple of them is merging in one particle at probability rate

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

where Λ is a given finite measure on the unit interval. The event of merging of two or more particles is called *collision*. By every collision Π_n jumps to a smaller value. When Λ is a Dirac mass at 0 the Λ-coalescent is the classical Kingman coalescent [23], in which every pair of particles is merging at the unit rate and only binary mergers are possible. Another eminent instance, known as the Bolthausen-Sznitman coalescent [6], appears when Λ is the Lebesgue measure on $[0,1]$.

The subclass of *beta-coalescents* are the processes driven by some beta measure on $[0,1]$ with density

$$\Lambda(\mathrm{d}x)/\mathrm{d}x = \frac{1}{\mathrm{B}(a,b)}x^{a-1}(1-x)^{b-1}, \qquad a,b > 0, \tag{2}$$

where $\mathrm{B}(\cdot,\cdot)$ denotes Euler's beta function. This class is amenable to analysis due to the fact that the transition rates (1) can be expressed in terms of $\mathrm{B}(\cdot,\cdot)$. For this reason and due to multiple connections with Lévy processes and random trees, the beta coalescents were the subject of intensive research [2, 3, 5, 8, 9, 15, 19, 28]. We refer to [4] for a survey and further references.

In this paper we study beta-coalescents with parameter $0 < a \leq 1$. Specifically, we are interested in the total number of collisions X_n and the total branch length of the coalescent tree L_n. Note that X_n is equal to the total number of particles born by collisions, and L_n is the cumulative lifetime of all particles from the start of the process to its termination. The variable L_n is closely related to the number of segregating sites M_n, the connection being that given L_n the distribution of M_n is Poisson with mean rL_n for some fixed mutation rate $r > 0$.

Our first main new contribution is the proof of a 1-stable limit law for X_n and L_n as $n \to \infty$. As in much of the previous work (see, for instance, [14] and [20]) we use a renewal approximation to Π_n. A novel element in this context is estimating the quality of approximation in terms of a Wasserstein distance.

The second new contribution are asymptotic expansions for the moments of X_n, L_n and M_n for the beta $(1,b)$-coalescent with arbitrary parameter $b > 0$. These expansions are obtained independently from the weak limiting results mentioned before. The proofs are based on the method of sequential approximations similar to those used in [18].

The rest of the paper is organized as follows. Section 2 gives a summary of some results on limit laws related to the beta-coalescents. In Section 3 general properties of the block-counting Markov chain and basic recurrences are discussed and the main results are stated. Section 4 recalls the definition and properties of a Wasserstein distance. In Section 5 we provide proofs of the main results. Some auxiliary lemmas are collected in the appendix.

2 A summary of limit laws for beta-coalescents

The tables in this section summarize the limit laws for X_n, L_n and the absorption time of the coalescent $\tau_n := \min\{t : \Pi_n(t) = 1\}$. The distributions which appear in the tables will be denoted as follows

(i) $\mathcal{N}$, standard normal,

(ii) $\mathcal{S}_\alpha$ with $1 < \alpha < 2$, (spectrally negative) α-stable distribution with characteristic function

$$z \mapsto \exp\big\{|z|^\alpha\big(\cos(\pi\alpha/2) + \mathrm{i}\sin(\pi\alpha/2)\mathrm{sgn}(z)\big)\big\}, \ z \in \mathbb{R}, \tag{3}$$

(iii) $\mathcal{S}_1$ (spectrally negative) 1-stable distribution with characteristic function

$$z \mapsto \exp\big\{-|z|(\pi/2 - \mathrm{i}\log|z|\,\mathrm{sgn}(z))\big\}, \ z \in \mathbb{R}, \tag{4}$$

(iv) $\mathcal{E}_\gamma(a, b)$ with $a, b, \gamma > 0$, distribution of the *exponential functional* $\int_0^\infty \exp(-\gamma S_{a,b}(t))\mathrm{d}t$, where $\big(S_{a,b}(t)\big)_{t \geq 0}$ is a drift-free subordinator with the Laplace exponent

$$\Phi_{a,b}(z) = \int_0^1 \big(1 - (1-x)^z\big)x^{a-3}(1-x)^{b-1}\mathrm{d}x, \qquad z \geq 0,$$

(v) $\mathcal{G}$, Gumbel with distribution function $x \mapsto \exp\big(-e^{-x}\big), x \in \mathbb{R}$,

(vi) ρ, convolution of infinitely many exponential laws with rates $i(i-1)/2$, $i \geq 2$.

Table 1: Limit distributions of $(X_n - a_n)/b_n$ for beta (a, b)-coalescents.

a	b	a_n	b_n	distribution	source
$0 < a < 1$	$b > 0$	$(1-a)n$	$(1-a)n^{1/(2-a)}$	$\mathcal{S}_{2-a}$	[20]($b = 1$), [14], this paper
$a = 1$	$b > 0$	$n(\log n)^{-1}+$ $n\log\log n(\log n)^{-2}$	$\frac{n}{(\log n)^2}$	$\mathcal{S}_1$	[9, 19]($b = 1$), this paper
$1 < a < 2$	$b > 0$	0	$\frac{\Gamma(a)}{2-a}n^{2-a}$	$\mathcal{E}_{2-a}(a, b)$	[12, 16]
$a = 2$	$b > 0$	$(2r_1)^{-1}(\log n)^2$	$(3^{-1}r_1^{-3}r_2\log^3 n)^{1/2}$	$\mathcal{N}$	[12, 18]
$a > 2$	$b > 0$	$m_1^{-1}\log n$	$(m_1^{-3}m_2\log n)^{1/2}$	$\mathcal{N}$	[12, 13]

Notation and comments: $r_1 = \zeta(2, b)$, $r_2 = 2\zeta(3, b)$, where $\zeta(\cdot, \cdot)$ is the Hurwitz zeta function; $m_1 = \Psi(a - 2 + b) - \Psi(b)$, $m_2 = \Psi'(b) - \Psi'(a - 2 + b)$, where $\Psi(\cdot)$ is the logarithmic derivative of the gamma function.

For the Bolthausen-Sznitman coalescent the limit law of X_n was first obtained in [9] using singularity analysis of generating functions. A probabilistic proof of this result appeared in [19], where a coupling with a random walk with barrier was exploited, and

3

the technique was further extended in [20] to study collisions in the beta $(a, 1)$-coalescents with $a \in (0, 2)$. The aforementioned limit laws for $a > 1$ are specializations of results for more general Λ-coalescents with *dust component*, i.e., those driven by the measures Λ such that $\int_0^1 x^{-1} \Lambda(\mathrm{d}x) < \infty$ [12, 13, 14, 16]. For Kingman's coalescent we have $X_n = n - 1$ for all $n \in \mathbb{N}$.

In the next two tables the value $a = 0$ corresponds to Kingman's coalescent.

Table 2: Limit distributions of $(\tau_n - a_n)/b_n$ for beta (a, b)-coalescents.

a	b	a_n	b_n	distribution	source
$a = 0$		0	1	ρ	[31]
$a = 1$	$b = 1$	$\log \log n$	1	$\mathcal{G}$	[15, 10]
$1 < a < 2$	$b > 0$	$\mathtt{m}^{-1} \log n$	$(\mathtt{m}^{-3} \mathtt{s}^2 \log n)^{1/2}$	$\mathcal{N}$	[12]
$a = 2$	$b > 0$	$c_1^{-1} \log n$	$(c_1^{-3} c_2 \log n)^{1/2}$	$\mathcal{N}$	[12]
$a > 2$	$b > 0$	$(\gamma m_1)^{-1} \log n$	$\gamma^{-1}(m_1^{-3}(m_2 + m_1^2) \log n)^{1/2}$	$\mathcal{N}$	[12, 13]

Notation and comments: The constants $\mathtt{m}$ and $\mathtt{s}^2$ are

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

$$\begin{aligned} \mathtt{s}^2 = {} & \frac{a + b - 1}{(a - 1)(2 - a)} \times \bigg(2\big(\Psi(a + b - 1) - \Psi(b)\big) \\ & - (a + b - 2)\big((\Psi(a + b - 1) - \Psi(b))^2 + \Psi'(b) - \Psi'(a + b - 1)\big)\bigg), \end{aligned}$$

$c_1 = b(b + 1)\zeta(2, b)$, $c_2 = 2b(b + 1)\zeta(3, b)$. The constants m_1 and m_2 are the same as in Table 1, and for $a > 2$

$$\gamma = \frac{(a - 1 + b)(a - 2 + b)}{(a - 1)(a - 2)}.$$

In the case $a \in (0, 1)$, $b > 0$ the beta (a, b)-coalescent has the property of coming down from infinity [30], which implies that τ_n weakly converges without any normalization to some limiting law, which is not known explicitly. The result for $a > 1$ is a special case of Theorem 4.3 in [12]. The case $a = 1$ and $b \neq 1$ is open; in this case the coalescent does not come down from infinity.

Table 3: Limit distributions of $(L_n - a_n)/b_n$ for beta (a, b)-coalescents.

a	b	a_n	b_n	distribution	source
$a = 0$		$2 \log n$	2	$\mathcal{G}$	[8, 31]
$0 < a < \frac{3 - \sqrt{5}}{2}$	$b = 2 - a$	$c_1 n^a$	1	exists	[22]
$a = \frac{3 - \sqrt{5}}{2}$	$b = 2 - a$	$c_1 n^a$	$c_2 (\log n)^{\alpha^{-1}}$	$\mathcal{S}_{2-a}$	[22]
$\frac{3 - \sqrt{5}}{2} < a < 1$	$b = 2 - a$	$c_1 n^a$	$c_2 (\beta n^{-\beta})^{\alpha^{-1}}$	$\mathcal{S}_{2-a}$	[22]
$a = 1$	$b > 0$	$n(b \log n)^{-1} +$ $b^{-1} n \log \log n (\log n)^{-2}$	$\frac{n}{b(\log n)^2}$	$\mathcal{S}_1$	[8]$(b = 1)$, this paper
$a > 1$	$b > 0$	0	$\mathrm{B}(a, b) n$	$\mathcal{E}_1(a, b)$	[24, 25]

Notation and comments: The constants are $\alpha = 2 - a$, $\beta = 1 + \alpha - \alpha^2$, $c_1 = \frac{\Gamma(\alpha+1)(\alpha-1)}{2-\alpha}$, $c_2 = \frac{\Gamma(\alpha+1)(\alpha-1)^{1+\alpha^{-1}}}{\cos(\pi\alpha/2)\Gamma^{\alpha^{-1}}(2-\alpha)}$.

In [24] the weak convergence of properly normalized L_n was proved for Λ-coalescents with dust component. In particular, that result covered the beta (a, b)-coalescents with $a > 1$. Although some partial results for $a \in (0, 1)$ and $b > 0$ were obtained in [7], this case with $b \neq 2 - a$ remains open.

3   Main results

For the general Λ-coalescent, the Markov chain Π_n is a pure-death process which jumps from state m to $m - k + 1$ at rate $\binom{m}{k}\lambda_{m,k}$, where $\lambda_{m,k}, 2 \leq k \leq m$, is given by (1). The total transition rate from state $m \geq 2$ is

$$\lambda_m := \sum_{k=2}^{m} \binom{m}{k}\lambda_{m,k} = \int_0^1 \left(1 - mx(1-x)^{m-1} - (1-x)^m\right)x^{-2}\Lambda(\mathrm{d}x). \tag{5}$$

The first decrement I_n of Π_n has distribution

$$\mathbb{P}\{I_n = k\} = \binom{n}{k+1}\frac{\lambda_{n,k+1}}{\lambda_n}, \quad 1 \leq k \leq n - 1.$$

The strong Markov property of the coalescent entails the distributional recurrences

$$X_1 = 0, \quad X_n \overset{d}{=} 1 + X'_{n-I_n}, \quad n \in \mathbb{N}\backslash\{1\}; \tag{6}$$

$$\tau_1 = 0, \quad \tau_n \overset{d}{=} T_n + \tau'_{n-I_n}, \quad n \in \mathbb{N}\backslash\{1\}; \tag{7}$$

$$L_1 = 0, \quad L_n \overset{d}{=} nT_n + L'_{n-I_n}, \quad n \in \mathbb{N}\backslash\{1\}, \tag{8}$$

where T_n denotes the time of the first collision, hence T_n has the exponential law with parameter λ_n; X'_k (respectively, τ'_k, L'_k) is independent of I_n (respectively, (T_n, I_n)) and is distributed like X_k (respectively, τ_k, L_k), for each $k \in \mathbb{N}$.

Letting Λ be defined by (2) with $a \in (0, 1]$ denote by

$$p_{n,k}^{(a)} := \mathbb{P}\{I_n = n - k\}, \quad k = 1, \ldots, n - 1. \tag{9}$$

Using the leading terms of asymptotic relations (27), (28) and (29) we infer

$$\lim_{n \to \infty} p_{n,n-k}^{(a)} = \frac{(2-a)\Gamma(k+a-1)}{\Gamma(a)(k+1)!} =: p_k^{(a)}, \quad k \in \mathbb{N},$$

hence

$$I_n \overset{d}{\to} \xi, \quad n \to \infty, \tag{10}$$

where ξ is a random variable with distribution $(p_k^{(a)})_{k \in \mathbb{N}}$.

Consider a zero-delayed random walk $(S_n)_{n \in \mathbb{N}_0}$ defined by

$$S_0 := 0, \quad S_n := \xi_1 + \ldots + \xi_n, \quad n \in \mathbb{N},$$

5

where (ξ_j) are independent copies of ξ with distribution $(p_k^{(a)})_{k \in \mathbb{N}}$, and let $(N_n)_{n \in \mathbb{N}_0}$ be the associated first-passage time sequence defined by

$$N_n = \inf\{k \geq 0 : S_k \geq n\}, \quad n \in \mathbb{N}_0.$$

It is plain that

$$N_0 = 0, \quad N_n \overset{d}{=} 1 + N'_{n-\xi \wedge n} = 1 + N'_{n-\xi} 1_{\{\xi < n\}}, \quad n \in \mathbb{N}, \tag{11}$$

where N'_k is independent of ξ and distributed like N_k, for each $k \in \mathbb{N}$. Comparing (6) and (11) one can expect that if N_n (properly centered and normalized) converges weakly to some proper and non degenerate probability law then the same is true for X_n (with the same centering and normalization). This is what we mean by a renewal approximation mentioned in the Introduction. This idea was exploited in [14] (for $a \in (0,1)$, $b > 0$) and in [20] (for $a \in (0,1]$, $b = 1$) to derive the limit distribution of X_n from that of N_n. We shall use a method based on probability metrics to show the stable limits for $a \in (0,1]$ and $b > 0$.

Theorem 3.1. *As $n \to \infty$ the number of collisions X_n in the* beta (a,b)-*coalescent satisfies*

(i) *for $0 < a < 1$ and $b > 0$*

$$\frac{X_n - (1-a)n}{(1-a)n^{1/(2-a)}} \overset{d}{\to} \mathcal{S}_{2-a},$$

(ii) *for $a = 1$ and $b > 0$,*

$$\frac{\log^2 n}{n} X_n - \log n - \log \log n \overset{d}{\to} \mathcal{S}_1.$$

As a consequence of our main theorem we also obtain a weak limit for the total branch length L_n and the number of segregating sites M_n (see [24]) of the beta $(1, b)$-coalescent.

Corollary 3.2. *For the total branch length L_n in the* beta $(1, b)$-*coalescent we have as $n \to \infty$*

$$\frac{b \log^2 n}{n} L_n - \log n - \log \log n \overset{d}{\to} \mathcal{S}_1.$$

Corollary 3.3. *For the number of segregating sites M_n in the* beta $(1, b)$-*coalescent we have as $n \to \infty$*

$$\frac{b \log^2 n}{rn} M_n - \log n - \log \log n \overset{d}{\to} \mathcal{S}_1,$$

where $r > 0$ is the rate of the homogeneous Poisson process on branches of the coalescent tree.

We now turn to the moments of X_n, L_n and M_n. An analysis of these moments provides further insight into the structure of these functionals. Our next result concerns the asymptotics of the moments of the number of collisions X_n in the beta $(1, b)$-coalescent.

Theorem 3.4. *Fix $b \in (0, \infty)$ and $j \in \mathbb{N}_0$. The jth moment of the number of collisions X_n in the* beta $(1, b)$-*coalescent has the asymptotic expansion*

$$\mathbb{E}X_n^j = \frac{n^j}{\log^j n}\left(1 + \frac{m_j}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right), \qquad n \to \infty, \tag{12}$$

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 $\kappa_j := (j + b - 1)\Psi(j + b) + j - (b - 1)\Psi(b)$, $j \in \mathbb{N}_0$.

For some more information on the coefficients m_j, $j \in \mathbb{N}$, we refer the reader to Eq. (23) in the proof of the following Corollary 3.5, which provides asymptotic expansions for the central moments of X_n in the beta $(1, b)$-coalescent.

Corollary 3.5. *Fix $b \in (0, \infty)$ and $j \in \mathbb{N} \setminus \{1\}$. The jth central moment of the number of collisions X_n in the* beta $(1, b)$-*coalescent has the asymptotic expansion*

$$\mathbb{E}(X_n - \mathbb{E}X_n)^j = \frac{(-1)^j}{j}\mathrm{B}(b, j - 1)\frac{n^j}{\log^{j+1} n} + O\left(\frac{n^j}{\log^{j+2} n}\right), \qquad n \to \infty. \tag{13}$$

In particular, $\mathrm{Var}(X_n) = (2b)^{-1}n^2/\log^3 n + O(n^2/\log^4 n)$ as $n \to \infty$.

Remark 3.6. For $b = 1$, Eq. (13) reduces to the asymptotic expansion (see Panholzer [27, p. 277 or Theorem 2.1. with $\alpha = 0$])

$$\mathbb{E}(X_n - \mathbb{E}X_n)^j = \frac{(-1)^j}{j(j-1)}\frac{n^j}{\log^{j+1} n} + O\left(\frac{n^j}{\log^{j+2} n}\right), \qquad n \to \infty$$

of the jth central moment of the number of collisions X_n for the Bolthausen–Sznitman n-coalescent.

The last result concerns the moments end central moments of the total branch length L_n of the beta $(1, b)$-coalescent.

Proposition 3.7. *Fix $b \in (0, \infty)$ and $j \in \mathbb{N}_0$. The jth moment of the total branch length L_n of the* beta $(1, b)$-*coalescent has the asymptotic expansion*

$$\mathbb{E}L_n^j = \frac{1}{b^j}\frac{n^j}{\log^j n}\left(1 + \frac{m_j}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right), \qquad n \to \infty, \tag{14}$$

where the sequence $(m_j)_{j \in \mathbb{N}_0}$ is defined as in Theorem 3.4. Moreover, for $j \in \{2, 3, \ldots\}$, the jth central moment of L_n has the asymptotic expansion

$$\mathbb{E}(L_n - \mathbb{E}L_n)^j = \frac{(-1)^j}{jb^j}\mathrm{B}(b, j - 1)\frac{n^j}{\log^{j+1} n} + O\left(\frac{n^j}{\log^{j+2} n}\right), \qquad n \to \infty. \tag{15}$$

In particular, $\mathrm{Var}(L_n) = (2b^3)^{-1}n^2/\log^3 n + O(n^2/\log^4 n)$ as $n \to \infty$.

Proposition 3.7 indicates that bL_n essentially behaves like X_n, in agreement when comparing Theorem 3.1 (ii) with Corollary 3.2. The proof of Proposition 3.7 works essentially the same as the analogous proofs of Theorem 3.4 and Corollary 3.5 for X_n. Instead

of the distributional recurrence (6) for $(X_n)_{n \in \mathbb{N}}$ one has to work with the distributional recurrence (8) for $(L_n)_{n \in \mathbb{N}}$. Since the expansion of $\mathbb{E}T_n = 1/\lambda_n$ is known (see Lemma 6.4), the proofs concerning X_n are readily adapted for L_n. A proof of Proposition 3.7 is therefore omitted. We finally mention that, for the beta $(1, b)$-coalescent with mutation rate $r > 0$, expansions for the moments and central moments of the number of segregating sites M_n can be easily obtained, since (see, for example, [8, p. 1417]) the descending factorial moments of M_n are related to the moments of L_n via $\mathbb{E}(M_n)_j = r^j \mathbb{E}L_n^j$, $j \in \mathbb{N}_0$.

4   Probability distances χ_T and d_q

For real-valued random variables X and Y and $T > 0$ the χ_T-distance between X and Y is defined by

$$\chi_T(X, Y) = \sup_{|t| \leq T} \left| \mathbb{E}e^{itX} - \mathbb{E}e^{itY} \right|. \tag{16}$$

By the continuity theorem for the characteristic functions convergence in distribution $Z_n \xrightarrow{d} Z$ holds if and only if $\lim_{n \to \infty} \chi_T(Z_n, Z) = 0$, for every $T > 0$.

Let $\mathcal{D}_q, q \in (0, 1]$, be the set of probability laws on $\mathbb{R}$ with finite qth absolute moment. Recall that $|x - y|^q$ is a metric on $\mathbb{R}$. The associated Wasserstein distance on $\mathcal{D}_q$ is defined by

$$d_q(X, Y) = \inf \mathbb{E}|\widehat{X} - \widehat{Y}|^q, \tag{17}$$

where the infimum is taken over all couplings $(\widehat{X}, \widehat{Y})$ such that $X \stackrel{d}{=} \widehat{X}$ and $Y \stackrel{d}{=} \widehat{Y}$.

For ease of reference we summarize properties of d_q in the following proposition.

Proposition 4.1. *Let X, Y be random variables with finite qth absolute moments. The Wasserstein distance d_q has the following properties:*

(Dist) $d_q(X, Y)$ *only depends on marginal distributions of X and Y,*

(Inf) *the infimum in* (17) *is attained for some coupling,*

(Rep) *the Kantorovich-Rubinstein representation holds*

$$d_q(X, Y) = \sup_{f \in \mathcal{F}_q} |\mathbb{E}f(X) - \mathbb{E}f(Y)|,$$

where $\mathcal{F}_q := \{f \in C(\mathbb{R}) : |f(x) - f(y)| \leq |x - y|^q, \ x, y \in \mathbb{R}\}$,

(Hom) $d_q(cX, cY) = |c|^q d(X, Y)$ *for $c \in \mathbb{R}$,*

(Reg) *for X, Y, Z defined on the same probability space $d_q(X + Z, Y + Z) \leq d_q(X, Y)$ provided $Z \in \mathcal{D}_q$ is independent of (X, Y),*

(Aff) $d_q(X + a, Y + a) = d_q(X, Y)$ *for $a \in \mathbb{R}$,*

(Conv) *for $X, X_n \in \mathcal{D}_q$ convergence $d_q(X_n, X) \to 0$, $n \to \infty$ implies $X_n \xrightarrow{d} X$ and $\mathbb{E}|X_n|^q \to \mathbb{E}|X|^q$.*

Proof. We refer to [11, 21] for most of these facts. To prove (Reg) choose an independent of Z coupling (X', Y') on which the infimum in the definition of d_q is attained. Then $X + Z \overset{d}{=} X' + Z$, $Y + Z \overset{d}{=} Y' + Z$ and the definition of d_q entails

$$d_q(X + Z, Y + Z) \leq \mathbb{E}|(X' + Z) - (Y' + Z)|^q = \mathbb{E}|X' - Y'|^q = d_q(X, Y).$$

Property (Conv): the convergence of moments is easy; the rest is a consequence of Lemma 4.2 to follow. $\qquad\square$

Lemma 4.2. *For $T > 0$ and $q \in (0, 1]$ there exists constant $C = C_{T,q} > 0$ such that*

$$\sup_{|t| \leq T} |\mathbb{E}e^{itX} - \mathbb{E}e^{itY}| \leq Cd_q(X, Y), \quad n \in \mathbb{N}.$$

Proof. Assume that the infimum in the definition of $d_q(X, Y)$ is attained on $(\hat{X}, \hat{Y})$. It is easy to check that for arbitrary $q \in (0, 1]$

$$|e^{ix} - e^{iy}| = 2\left|\sin\frac{x - y}{2}\right| \leq 2^{1-q}M_q|x - y|^q, \quad x, y \in \mathbb{R}, \tag{18}$$

where $M_q := \sup_{u > 0} |\sin u| u^{-q} < \infty$. Hence

$$
\begin{aligned}
\sup_{|t| \leq T} |\mathbb{E}e^{itX} - \mathbb{E}e^{itY}| &= \sup_{|t| \leq T} |\mathbb{E}e^{it\hat{X}} - \mathbb{E}e^{it\hat{Y}}| \leq \sup_{|t| \leq T} \mathbb{E}|e^{it\hat{X}} - \mathbb{E}e^{it\hat{Y}}| \\
&\overset{(18)}{\leq} 2^{1-q}M_q \sup_{|t| \leq T} |t|^q \mathbb{E}|\hat{X} - \hat{Y}|^q \leq 2^{1-q}M_q T^q d_q(X, Y),
\end{aligned}
$$

as wanted. $\qquad\square$

5 Proofs

5.1 Proof of Theorem 3.1

Suppose $a = 1$. It is enough to show that

$$\lim_{n \to \infty} \chi_T\left(\frac{\log^2 n}{n}X_n - \log n - \log\log n, \mathcal{S}_1\right) = 0,$$

for every $T > 0$.

Using the triangle inequality yields

$$
\begin{aligned}
\chi_T\left(\frac{\log^2 n}{n}X_n - \log n - \log\log n, \mathcal{S}_1\right) &\leq \\
\chi_T\left(\frac{\log^2 n}{n}X_n - \log n - \log\log n, \frac{\log^2 n}{n}N_n - \log n - \log\log n\right) & \\
+ \quad \chi_T\left(\frac{\log^2 n}{n}N_n - \log n - \log\log n, \mathcal{S}_1\right).&
\end{aligned}
$$

The second term converges to zero by Proposition 2 in [19] on stable limit for the number of renewals. In view of Lemma 4.2 to prove convergence to zero of the first term it is sufficient to check that

$$\lim_{n\to\infty} d_q\Big(\frac{\log^2 n}{n}X_n - \log n - \log\log n, \frac{\log^2 n}{n}N_n - \log n - \log\log n\Big) = 0,$$

for some $q \in (0,1]$, which in view of the properties (Hom) and (Aff) in Proposition 4.1 amounts to the estimate

$$d_q(X_n, N_n) = o(n^q \log^{-2q} n), \quad n \to \infty. \tag{19}$$

In the like way, proving Theorem 3.1 in the case $a \in (0,1)$ reduces to showing that

$$d_q(X_n, N_n) = o(n^{q/(2-a)}), \quad n \to \infty, \tag{20}$$

for some $q \in (0,1]$.

Using recurrences (6) for X_n and (11) for N_n we obtain

$$t_n := d_q(X_n, N_n) = d_q(X'_{n-I_n}, N'_{n-(\xi \wedge n)}) \le d_q(N'_{n-I_n}, N'_{n-(\xi \wedge n)}) + d_q(X'_{n-I_n}, N'_{n-I_n})$$

$$\le d_q(N'_{n-I_n}, N'_{n-(\xi \wedge n)}) + \mathbb{E}|\widehat{X}_{n-I_n} - \widehat{N}_{n-I_n}|^q =: c_n + \sum_{k=1}^{n-1} \mathbb{P}\{I_n = n-k\}\mathbb{E}|\widehat{X}_k - \widehat{N}_k|^q,$$

for arbitrary pairs $\big((\widehat{X}_k, \widehat{N}_k)\big)_{1\le k \le n-1}$ independent of I_n such that $\widehat{X}_k \overset{d}{=} X_k$, $\widehat{N}_k \overset{d}{=} N_k$. Passing to the infimum over all such pairs leads to

$$t_n \le c_n + \sum_{k=1}^{n-1} \mathbb{P}\{I_n = n-k\}t_k. \tag{21}$$

We shall use (21) to estimate t_n.

First we find an appropriate bound for c_n. Let $(\hat{I}_n, \hat{\xi})$ be a coupling of I_n and ξ such that (recall (Inf) in Proposition 4.1) $d_q(I_n, \xi \wedge n) = \mathbb{E}|\hat{I}_n - \hat{\xi} \wedge n|^q$. Let $\big(\hat{N}_k\big)_{k\in\mathbb{N}}$ be a copy of $\big(N_k\big)_{k\in\mathbb{N}}$ independent of $(\hat{I}_n, \hat{\xi})$. Since $\big(\hat{I}_n, \hat{\xi}, (\hat{N}_k)\big)$ is a particular coupling we have

$$c_n = d_q(N'_{n-I_n}, N'_{n-(\xi \wedge n)}) \le \mathbb{E}|\hat{N}_{n-\hat{I}_n} - \hat{N}_{n-(\hat{\xi}\wedge n)}|^q.$$

Exploiting the stochastic inequality

$$N_{x+y} - N_x \overset{d}{\le} N_y, \quad x,y \in \mathbb{N}$$

yields

$$\mathbb{E}|\hat{N}_{n-\hat{I}_n} - \hat{N}_{n-\hat{\xi}\wedge n}|^q \le \mathbb{E}\hat{N}^q_{|\hat{I}_n - \hat{\xi}\wedge n|}.$$

Furthermore, we obviously have $N_n \le n$, hence

$$c_n \le \mathbb{E}|\hat{I}_n - \hat{\xi} \wedge n|^q = d_q(I_n, \xi \wedge n).$$

10

Now we invoke the Kantorovich-Rubinstein representation ((Rep) in Proposition 4.1) for d_q. Set $\mathcal{F}_{q,0} := \mathcal{F}_q \cap \{f : f(0) = 0\}$ and note that $f \in \mathcal{F}_{q,0}$ implies $|f(x)| \le |x|^q$, $x \in \mathbb{R}$. We have

$$
\begin{aligned}
c_n &\le d_q(I_n, \xi \wedge n) = \sup_{f \in \mathcal{F}_q} \left| \mathbb{E}f(I_n) - \mathbb{E}f(\xi \wedge n) \right| = \sup_{f \in \mathcal{F}_{q,0}} \left| \mathbb{E}f(I_n) - \mathbb{E}f(\xi \wedge n) \right| \\
&= \sup_{f \in \mathcal{F}_{q,0}} \left| \sum_{k=1}^{n-1} \mathbb{P}\{I_n = k\} f(k) - \sum_{k=1}^{n-1} \mathbb{P}\{\xi = k\} f(k) - f(n) \sum_{k \ge n} \mathbb{P}\{\xi = k\} \right| \\
&\le \sum_{k=1}^{n-1} \left| \mathbb{P}\{I_n = k\} - \mathbb{P}\{\xi = k\} \right| k^q + n^q \mathbb{P}\{\xi \ge n\}.
\end{aligned}
$$

For appropriate $q \in (0, 1]$ (to be specified below) such that $a + q > 1$ use Lemma 6.3 in the Appendix along with the relation $\mathbb{P}\{\xi \ge n\} = O(n^{a-2})$ to obtain the estimate $c_n = O(n^{q+a-2})$. With this bound for c_n a O-estimate for t_n follows using Lemma 6.1.

If $a \in (0, 1)$ one can take $q = 1$. Then the cited lemma applies with $\psi_n = n$ and $r_n = Mn^{a-1}$ (M large enough) and gives estimate

$$
d_q(X_n, N_n) = O(n^a),
$$

which implies (20).

For the case $a = 1$ application of the same lemma with $\psi_n = n/(\log(n+1))$ and $r_n = Mn^{q-1}$ (M large enough) leads to $t_n \le Mn^q(\log n)^{-1}$. Thus (19) holds for $q \in (0, 1/2)$. The proof is complete.

5.2 Proof of Corollaries 3.2 and 3.3.

We follow closely the proofs of Theorem 5.2 and Corollary 6.2 in [8]. In view of

$$
\frac{b \log^2 n}{n} L_n - \log n - \log \log n = \frac{\log^2 n}{n} X_n - \log n - \log \log n + \frac{\log^2 n}{n} \left(bL_n - X_n \right),
$$

it is enough to show that $\frac{\log^2 n}{n} \left(bL_n - X_n \right) \to 0$ in L_2.

Let T_j's be independent exponential variables with rates $\lambda_j, j \ge 2$. Assuming the T_j's independent of the sequence of states visited by Π_n we may identify T_j with the time Π_n spends in the state j provided this state is visited. Given the sequence of visited states is $n = i_0 > i_1 > \cdots > i_{k-1} > i_k = 1$ the total branch length L_n is distributed like $\sum_{r=0}^{k-1} i_r T_{i_r}$ for $n \in \mathbb{N} \setminus \{1\}$.

For $k \in \{1, \ldots, n\}$ and $\mathrm{i} = (i_0, \ldots, i_k)$ with $n = i_0 > i_1 > \cdots > i_{k-1} > i_k = 1$ define the events $A_{k,\mathrm{i}} := \{X_n = k, (\Pi_n(t_0), \ldots, \Pi_n(t_k)) = \mathrm{i}\}$, where $t_0 = 0$ and $t_1 < t_2 < \ldots$ are the collision epochs. We have

$$
\begin{aligned}
\mathbb{E}(bL_n - X_n)^2 &= \sum_{k,\mathrm{i}} \mathbb{P}\{A_{k,\mathrm{i}}\} \mathbb{E}\left(\sum_{r=0}^{k-1} (bi_r T_{i_r} - 1) \right)^2 \\
&= \sum_{k,\mathrm{i}} \mathbb{P}\{A_{k,\mathrm{i}}\} \left(\sum_{r=0}^{k-1} \mathbb{E}(bi_r T_{i_r} - 1)^2 + \sum_{r,s=0, r \ne s}^{k-1} \mathbb{E}(bi_r T_{i_r} - 1)(bi_s T_{i_s} - 1) \right)
\end{aligned}
$$

11

Furthermore, $\lambda_n = bn + O(\log n)$ as $n \to \infty$ for $a = 1$ and $b > 0$ (see (29)) which implies

$$|\mathbb{E}(bkT_k - 1)| = O(k^{-1}\log k) \text{ and } \mathbb{E}(bkT_k - 1)^2 = 1 + O(k^{-1}\log k).$$

Therefore,

$$
\begin{aligned}
\mathbb{E}(bL_n - X_n)^2 &\leq \sum_{k,\,i} \mathbb{P}\{A_{k,\,i}\}\Big(\sum_{r=2}^{n}\mathbb{E}(brT_r - 1)^2 + \Big(\sum_{r=2}^{n}|\mathbb{E}(brT_r - 1)|\Big)^2\Big)\\
&= \sum_{k,\,i}\mathbb{P}\{A_{k,\,i}\}\Big(n + O(\log^4 n)\Big) = n + O(\log^4 n),
\end{aligned}
$$

and the convergence in L_2 follows.

Corollary 3.3 follows from the fact that given L_n the distribution of M_n is Poisson with mean rL_n. See Corollary 6.2 in [8] for details.

5.3 Proofs of Theorem 3.4 and Corollary 3.5

Let us verify (12) by induction on $j \in \mathbb{N}$. From (6) it follows that $a_1 := \mathbb{E}X_1 = 0$ and $a_n := \mathbb{E}X_n = 1 + \sum_{m=2}^{n-1}p_{n,m}^{(1)}a_m$, $n \in \mathbb{N}\setminus\{1\}$. In the following we apply the method of sequential approximations to the sequence $(a_n)_{n\in\mathbb{N}}$. The sequence $(b_n)_{n\in\mathbb{N}}$, defined via $b_1 := 0$ and $b_n := a_n - n/\log n$ for $n \in \mathbb{N}\setminus\{1\}$, satisfies the recursion

$$b_n = a_n - \frac{n}{\log n} = 1 + \sum_{m=2}^{n-1}p_{n,m}^{(1)}\Big(\frac{m}{\log m} + b_m\Big) - \frac{n}{\log n} = q_n + \sum_{m=2}^{n-1}p_{n,m}^{(1)}b_m,$$

$n \in \mathbb{N}\setminus\{1\}$, where $q_n := 1 - n/\log n + \sum_{m=2}^{n-1}p_{n,m}^{(1)}m/\log m$, $n \in \mathbb{N}\setminus\{1\}$. By Corollary 6.7 (applied with $\alpha := 1$ and $p := 1$),

$$q_n = 1 - \frac{n}{\log n} + \Big(\frac{n}{\log n} - 1 + \frac{m_1}{\log n} + O\Big(\frac{1}{\log^2 n}\Big)\Big) = \frac{m_1}{\log n} + O\Big(\frac{1}{\log^2 n}\Big),$$

where $m_1 := c_{b,1,1} = 2 + \Psi(b)$. The sequence $(c_n)_{n\in\mathbb{N}}$, defined via $c_1 := 0$ and $c_n := b_n - m_1 n/\log^2 n$ for $n \in \mathbb{N}\setminus\{1\}$, therefore satisfies the recursion

$$c_n = b_n - m_1\frac{n}{\log^2 n} = q_n + \sum_{m=2}^{n-1}p_{n,m}^{(1)}\Big(m_1\frac{m}{\log^2 m} + c_m\Big) - m_1\frac{n}{\log^2 n} = q_n' + \sum_{m=2}^{n-1}p_{n,m}^{(1)}c_m,$$

$n \in \mathbb{N}\setminus\{1\}$, where $q_n' := q_n - m_1 n/\log^2 n + m_1\sum_{m=2}^{n-1}p_{n,m}^{(1)}m/\log^2 m$, $n \in \mathbb{N}\setminus\{1\}$. By Corollary 6.7 (applied with $\alpha := 1$ and $p := 2$),

$$q_n' = q_n - m_1\frac{n}{\log^2 n} + m_1\Big(\frac{n}{\log^2 n} - \frac{1}{\log n} + O\Big(\frac{1}{\log^2 n}\Big)\Big) = O\Big(\frac{1}{\log^2 n}\Big),$$

since $q_n = m_1/\log n + O(1/\log^2 n)$. By Lemma 6.2 (applied with $\alpha := 1$ and $p := 3$), it follows that $c_n = O(n/\log^3 n)$. Thus, (12) holds for $j = 1$. Assume now that $j \geq 2$. From $\mathbb{E}X_{I_n}^j = \mathbb{E}(X_n - 1)^j = \sum_{i=0}^{j-1}\binom{j}{i}(-1)^{j-i}\mathbb{E}X_n^i + \mathbb{E}X_n^j$ it follows that

$$a_{n,j} := \mathbb{E}X_n^j = \sum_{i=0}^{j-1}\binom{j}{i}(-1)^{j-1-i}\mathbb{E}X_n^i + \mathbb{E}X_{I_n}^j = q_{n,j} + \sum_{m=2}^{n-1}p_{n,m}^{(1)}a_{m,j},$$

12

$n \in \mathbb{N} \setminus \{1\}$, where $q_{n,j} := \sum_{i=0}^{j-1} \binom{j}{i}(-1)^{j-1-i} \mathbb{E} X_n^i$, $n \in \mathbb{N} \setminus \{1\}$. Since, by induction, for all $i < j$,

$$\mathbb{E} X_n^i = \frac{n^i}{\log^i n}\left(1 + \frac{m_i}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right),$$

it follows that (the summand for $i = j - 1$ asymptotically dominates the others)

$$q_{n,j} = \frac{jn^{j-1}}{\log^{j-1} n}\left(1 + \frac{m_{j-1}}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right).$$

Now apply the method of sequential approximations to the sequence $(a_{n,j})_{n \in \mathbb{N}}$. The sequence $(b_{n,j})_{n \in \mathbb{N}}$, defined via $b_{1,j} := 0$ and $b_{n,j} := a_{n,j} - n^j / \log^j n$ for $n \in \{2, 3, \ldots\}$, satisfies the recursion

$$b_{n,j} = q'_{n,j} + \sum_{m=2}^{n-1} p_{n,m}^{(1)} b_{m,j}, \quad n \in \{2, 3, \ldots\},$$

where $q'_{n,j} := q_{n,j} - n^j / \log^j n + \sum_{m=2}^{n-1} p_{n,m}^{(1)} m^j / \log^j m$, $n \in \{2, 3, \ldots\}$. By Corollary 6.7 (applied with $\alpha := j$ and $p := j$),

$$
\begin{aligned}
q'_{n,j} &= j\frac{n^{j-1}}{\log^{j-1} n} + jm_{j-1}\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right) - \frac{n^j}{\log^j n} \\
&\quad + \frac{n^j}{\log^j n} - j\frac{n^{j-1}}{\log^{j-1} n} + \kappa_j\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right) \\
&= jm_j\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right),
\end{aligned}
$$

where $\kappa_j := c_{b,j,j}$ and $m_j := m_{j-1} + \kappa_j / j$. The sequence $(c_{n,j})_{n \in \mathbb{N}}$, defined via $c_{1,j} := 0$ and $c_{n,j} := b_{n,j} - m_j n^j / \log^{j+1} n$ for $n \in \{2, 3, \ldots\}$, therefore satisfies the recursion

$$c_{n,j} = q''_{n,j} + \sum_{m=2}^{n-1} p_{n,m}^{(1)} c_{m,j}, \quad n \in \{2, 3, \ldots\},$$

where $q''_{n,j} := q'_{n,j} - m_j n^j / \log^{j+1} n + m_j \sum_{m=2}^{n-1} p_{n,m}^{(1)} m^j / \log^{j+1} m$, $n \in \{2, 3, \ldots\}$. By Corollary 6.7 (applied with $\alpha := j$ and $p := j + 1$),

$$
\begin{aligned}
q''_{n,j} &= jm_j\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right) - m_j\frac{n^j}{\log^{j+1} n} \\
&\quad + m_j\left(\frac{n^j}{\log^{j+1} n} - j\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right)\right) = O\left(\frac{n^{j-1}}{\log^{j+1} n}\right).
\end{aligned}
$$

By Lemma 6.2 (applied with $\alpha := j$ and $p := j + 2$), it follows that $c_{n,j} = O(n^j / \log^{j+2} n)$, which shows that (12) holds for j. The induction is complete which finishes the proof of Theorem 3.4.

We now turn to the proof of Corollary 3.5. Let us first verify that the sequence $(m_j)_{j \in \mathbb{N}_0}$, recursively defined in Theorem 3.4, satisfies the inversion formula

$$\sum_{i=0}^{j} \binom{j}{i}(-1)^{j-i} m_i = \frac{(-1)^j}{j} \mathrm{B}(b, j-1), \quad j \in \mathbb{N} \setminus \{1\}. \tag{22}$$

Using the formula $\Psi(x+1) = \Psi(x) + 1/x$, $x \in (0, \infty)$, it is readily checked that $\kappa_{j+1} - \kappa_j = 2 + \Psi(b+j)$, $j \in \mathbb{N}_0$. For all $j \in \mathbb{N}_0$ it follows that $\kappa_j = \sum_{i=0}^{j-1}(\kappa_{i+1} - \kappa_i) = \sum_{i=0}^{j-1}(2 + \Psi(b+i)) = 2j + \sum_{i=0}^{j-1} \Psi(b+i)$ and

$$m_j = \sum_{l=1}^{j}(m_l - m_{l-1}) = \sum_{l=1}^{j} \frac{\kappa_l}{l} = \sum_{l=1}^{j}\left(2 + \frac{1}{l}\sum_{i=0}^{l-1}\Psi(b+i)\right) = 2j + \sum_{i=0}^{j-1}\Psi(b+i)\sum_{l=i+1}^{j}\frac{1}{l}. \tag{23}$$

By (23), for $j \in \{2, 3, \ldots\}$,

$$\sum_{i=0}^{j} \binom{j}{i}(-1)^{j-i} m_i = \sum_{i=1}^{j} \binom{j}{i}(-1)^{j-i}\left(2i + \sum_{k=0}^{i-1}\Psi(b+k)\sum_{l=k+1}^{i}\frac{1}{l}\right)$$

$$= \sum_{i=1}^{j} \binom{j}{i}(-1)^{j-i}\sum_{k=0}^{i-1}\Psi(b+k)\sum_{l=k+1}^{i}\frac{1}{l}$$

$$= \sum_{k=0}^{j-1}\Psi(b+k)\sum_{l=k+1}^{j}\frac{1}{l}\sum_{i=l}^{j}\binom{j}{i}(-1)^{j-i} = \sum_{k=0}^{j-1}\Psi(b+k)\sum_{l=k+1}^{j}\frac{1}{l}\binom{j-1}{l-1}(-1)^{j-l}$$

$$= \frac{1}{j}\sum_{k=0}^{j-1}\Psi(b+k)\sum_{l=k+1}^{j}\binom{j}{l}(-1)^{j-l} = \frac{1}{j}\sum_{k=0}^{j-1}\Psi(b+k)\binom{j-1}{k}(-1)^{j-1-k}.$$

Plugging in $\binom{j-1}{k} = \binom{j-2}{k-1} + \binom{j-2}{k}$ and reordering with respect to $\binom{j-2}{k}$ leads to

$$\sum_{i=0}^{j} \binom{j}{i}(-1)^{j-i} m_i = \frac{1}{j}\sum_{k=0}^{j-2}(-1)^{j-2-k}\binom{j-2}{k}(\Psi(b+k+1) - \Psi(b+k))$$

$$= \frac{(-1)^j}{j}\sum_{k=0}^{j-2}(-1)^k\binom{j-2}{k}\frac{1}{b+k} = \frac{(-1)^j}{j}\mathrm{B}(b, j-1),$$

where the last equality holds, since $\sum_{k=0}^{n}(-1)^k\binom{n}{k}/(b+k) = \mathrm{B}(b, n+1)$ for all $n \in \mathbb{N}_0$, which is for example readily verified by induction on $n \in \mathbb{N}_0$. Thus, (22) is established.

Thanks to Theorem 3.4 and the inversion formula (22) the proof of Corollary 3.5 is now straightforward. Basically the same argument has for example been used by Panholzer

[27, p. 277]. Plugging in the expansion (12) for the ordinary moments shows that

$$
\begin{aligned}
\mathbb{E}(X_n - \mathbb{E}X_n)^j &= \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} \mathbb{E}X_n^i (\mathbb{E}X_n)^{j-i} \\
&= \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} \frac{n^i}{\log^i n} \left(1 + \frac{m_i}{\log n} + O\left(\frac{1}{\log^2 n} \right) \right) \left(\frac{n}{\log n} \left(1 + \frac{m_1}{\log n} + O\left(\frac{1}{\log^2 n} \right) \right) \right)^{j-i} \\
&= \frac{n^j}{\log^j n} \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} \left(1 + \frac{m_i}{\log n} + O\left(\frac{1}{\log^2 n} \right) \right) \left(1 + \frac{(j-i)m_1}{\log n} + O\left(\frac{1}{\log^2 n} \right) \right) \\
&= \frac{n^j}{\log^j n} \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} \left(1 + \frac{(j-i)m_1 + m_i}{\log n} + O\left(\frac{1}{\log^2 n} \right) \right) \\
&= \frac{n^j}{\log^j n} \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} + \frac{n^j}{\log^{j+1} n} \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} ((j-i)m_1 + m_i) + O\left(\frac{n^j}{\log^{j+2} n} \right) \\
&= \frac{n^j}{\log^{j+1} n} \frac{(-1)^j}{j} \mathrm{B}(b, j-1) + O\left(\frac{n^j}{\log^{j+2} n} \right),
\end{aligned}
$$

since, for $j \geq 2$, $\sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} = 0$, $\sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} (j-i) = 0$, and $\sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} m_i = (-1)^j / j \mathrm{B}(b, j-1)$ by (22). The proof of Corollary 3.5 is complete.

6 Appendix

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

$$
a_n = r_n + \sum_{k=0}^{n} p_{n,k} a_k, \quad n \in \mathbb{N}, \tag{24}
$$

with given $r_n \geq 0$ and given initial value $a_0 = a \geq 0$. The following result is Lemma 6.1 from [12].

Lemma 6.1. *Suppose there exists a sequence $(\psi_n)_{n \in \mathbb{N}}$ such that*

(C1) $\liminf_{n \to \infty} \psi_n \sum_{k=0}^{n} (1 - k/n) p_{n,k} > 0$,

(C2) *the sequence $(r_k \psi_k / k)_{k \in \mathbb{N}}$ is non-increasing.*

Then a_n, defined by (24), satisfies

$$
a_n = O\left(\sum_{k=1}^{n} \frac{r_k \psi_k}{k} \right), \quad n \to \infty. \tag{25}
$$

Lemma 6.2. *Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of real numbers satisfying the recursion $a_1 = 0$ and $a_n = q_n + \sum_{m=2}^{n-1} p_{n,m}^{(1)} a_m$, $n \in \mathbb{N} \setminus \{1\}$, for some given sequence $(q_n)_{n \in \mathbb{N} \setminus \{1\}}$. If $q_n = O(n^{\alpha-1} / \log^{p-1} n)$ for some given constants $\alpha \in (0, \infty)$ and $p \in [0, \infty)$, then $a_n = O(n^{\alpha} / \log^p n)$.*

15

Proof. Fix some δ such that $0 < \delta < \alpha$. Set $a'_n := |a_n|/n^\delta$ and $q'_n := |q_n|/n^\delta$. Then $q'_n \le Mn^{\alpha-1-\delta}/\log^{p-1} n =: r_n$ for some $M > 0$ and all $n \ge 2$. Further,

$$
\begin{aligned}
a'_n &\le q'_n + \sum_{m=2}^{n-1} p_{n,m}^{(1)} \frac{|a_m|}{n^\delta} \le q'_n + \sum_{m=2}^{n-1} p_{n,m}^{(1)} \frac{|a_m|}{m^\delta} \\
&= q'_n + \sum_{m=2}^{n-1} p_{n,m}^{(1)} a'_m \le r_n + \sum_{m=2}^{n-1} p_{n,m}^{(1)} a'_m.
\end{aligned}
$$

Set $\psi_n := n/\log n$, then both conditions (C1) and (C2) of Lemma 6.1 are fulfilled. Hence $a'_n = O(\sum_{k=2}^n k^{\alpha-1-\delta}/\log^p k) = O(n^{\alpha-\delta}/\log^p n)$ and $|a_n| = n^\delta a'_n = O(n^\alpha/\log^p n)$. $\qquad\square$

Lemma 6.3. *For the first decrement I_n of the Markov chain (Π_n) associated with the beta (a,b)-coalescent ($a \in (0,1]$ and $b > 0$) and a random variable ξ with distribution $(p_k^{(a)})_{k\in\mathbb{N}}$*

$$
\sum_{k=1}^{n-1} k^q |\mathbb{P}\{I_n = k\} - \mathbb{P}\{\xi = k\}| = O(n^{a+q-2}), \tag{26}
$$

whenever $0 < q \le 1$ and $q + a > 1$.

Proof. For the beta (a,b)-coalescents formula (1) reads

$$
\lambda_{n,k+1} = \int_0^1 x^{k-1}(1-x)^{n-k-1} \Lambda(\mathrm{d}x) = \frac{\mathrm{B}(a+k-1, n-k+b-1)}{\mathrm{B}(a,b)}.
$$

Using the known estimate for the gamma function (see formula (6.1.47) in [1])

$$
\left| \frac{\Gamma(n+c)}{\Gamma(n+d)} - n^{c-d} \right| \le M_{c,d} n^{c-d-1}, \quad n \ge 2, \quad c,d > -2,
$$

we obtain

$$
\begin{aligned}
\binom{n}{k+1} \lambda_{n,k+1} &= \binom{n}{k+1} \frac{\mathrm{B}(a+k-1, n-k+b-1)}{\mathrm{B}(a,b)} \tag{27} \\
&= \frac{\Gamma(n+1)\Gamma(a+k-1)\Gamma(n-k+b-1)}{\Gamma(k+2)\Gamma(n-k)\Gamma(n+a+b-2)\mathrm{B}(a,b)} \\
&= \frac{\Gamma(a+k-1)}{(k+1)!\mathrm{B}(a,b)} \left(n^{3-a-b} + O(n^{2-a-b})\right)\left((n-k)^{b-1} + O\left((n-k)^{b-2}\right)\right),
\end{aligned}
$$

uniformly for $1 \le k \le n-1$ and $n \ge 2$.

Using (5) with Λ given by (2) we infer (see also Corollary 2 in [14])

$$
\lambda_n = \frac{\Gamma(a)}{(2-a)\mathrm{B}(a,b)} n^{2-a} + O(n^{1-a}) = \frac{\Gamma(a)}{(2-a)\mathrm{B}(a,b)} n^{2-a}\left(1 + O(n^{-1})\right), \tag{28}
$$

when $a \in (0,1)$ and $b > 0$, and

$$
\lambda_n = bn + O(\log n), \tag{29}
$$

16

when $a = 1$ and $b > 0$. Hence for $0 < a < 1$, $b > 0$, $n \geq 2$ and $k = 1, \ldots, n-1$

$$
\begin{aligned}
p_{n,n-k}^{(a)} &= \frac{(2-a)\Gamma(a+k-1)}{\Gamma(a)(k+1)!} n^{1-b}\Big((n-k)^{b-1} + O\big((n-k)^{b-2}\big)\Big)\Big(1 + O(n^{-1})\Big) \\
&= p_k^{(a)}\Big((1-k/n)^{b-1} + O\big(n^{-1}(1-k/n)^{b-2}\big)\Big)\Big(1 + O(n^{-1})\Big) \\
&= p_k^{(a)}\Big((1-k/n)^{b-1} + O\big(n^{-1}(1-k/n)^{b-2}\big)\Big) \\
&= p_k^{(a)}(1-k/n)^{b-1} + O\Big(p_k^{(a)} n^{-1}(1-k/n)^{b-2}\Big).
\end{aligned}
$$

Analogously for $a = 1$

$$
\begin{aligned}
p_{n,n-k}^{(1)} &= p_k^{(1)}\Big((1-k/n)^{b-1} + O\big(n^{-1}(1-k/n)^{b-2}\big)\Big)\Big(1 + O(n^{-1}\log n)\Big) \\
&= p_k^{(1)}\Big((1-k/n)^{b-1} + O\big(n^{-1}(1-k/n)^{b-2}\big) + O\big(n^{-1}\log n(1-k/n)^{b-1}\big)\Big) \\
&= p_k^{(1)}(1-k/n)^{b-1} + O\Big(p_k^{(1)} n^{-1}(1-k/n)^{b-2}\Big) + O\Big(p_k^{(1)} n^{-1}\log n(1-k/n)^{b-1}\Big).
\end{aligned}
$$

Substituting these expansions into the left-hand side of (26) gives

$$
\begin{aligned}
\sum_{k=1}^{n-1} k^q \Big|\mathbb{P}\{I_n = k\} - \mathbb{P}\{\xi = k\}\Big| &\leq \sum_{k=1}^{n-1} p_k^{(a)} k^q \left|\Big(1 - \frac{k}{n}\Big)^{b-1} - 1\right| \\
&\quad + \frac{c_1}{n} \sum_{k=1}^{n-1} p_k^{(a)} k^q \Big(1 - \frac{k}{n}\Big)^{b-2} \\
&=: S_1(a,n) + S_2(a,n),
\end{aligned}
$$

for $0 < a < 1$, and

$$
\begin{aligned}
\sum_{k=1}^{n-1} k^q \Big|\mathbb{P}\{I_n = k\} - \mathbb{P}\{\xi = k\}\Big| &\leq S_1(1,n) + S_2(1,n) + \frac{c_2 \log n}{n} \sum_{k=1}^{n-1} p_k^{(1)} k^q \Big(1 - \frac{k}{n}\Big)^{b-1} \\
&=: S_1(1,n) + S_2(1,n) + S_3(1,n),
\end{aligned}
$$

for $a = 1$. Here and hereafter $c_1, c_2, \ldots$ denote some positive constants whose values are of no importance. Our aim is to show that $S_i(a,n) = O(n^{q+a-2})$ for $i = 1, 2$ and $S_3(1,n) = O(n^{q-1})$. By virtue of $p_k^{(a)} \leq c_3 k^{a-3}$, for all $k \in \mathbb{N}$, we infer

$$
\begin{aligned}
S_1(a,n) &\leq c_3 \sum_{k=1}^{n-1} k^{a+q-3}\left|\Big(1 - \frac{k}{n}\Big)^{b-1} - 1\right| \\
&= c_3 \sum_{k=1}^{[n/2]} k^{a+q-3}\left|\Big(1 - \frac{k}{n}\Big)^{b-1} - 1\right| + c_3 \sum_{k=[n/2]+1}^{n-1} k^{a+q-3}\left|\Big(1 - \frac{k}{n}\Big)^{b-1} - 1\right| \\
&\leq \frac{c_4}{n} \sum_{k=1}^{[n/2]} k^{a+q-2} + c_3 n^{a+q-2}\left(\frac{1}{n} \sum_{k=[n/2]+1}^{n-1} \Big(\frac{k}{n}\Big)^{a+q-3}\left|\Big(1 - \frac{k}{n}\Big)^{b-1} - 1\right|\right),
\end{aligned}
$$

17

where the inequality $|(1 - x)^q - 1| \leq c_5 x, \; x \in [0, 1/2]$ has been utilized. The expression in the parentheses converges to $\int_{1/2}^{1} x^{a+q-3}|(1-x)^{b-1} - 1|dx < \infty$. Hence $S_1(a, n) = O(n^{q+a-2})$.

Similarly

$$
\begin{aligned}
S_2(a, n) \\
&\leq \frac{c_6}{n} \sum_{k=1}^{n-1} k^{a+q-3}\left(1 - \frac{k}{n}\right)^{b-2} \\
&= \frac{c_6}{n} \sum_{k=1}^{[n/2]} k^{a+q-3}\left(1 - \frac{k}{n}\right)^{b-2} + \frac{c_6}{n} \sum_{k=[n/2]+1}^{n-1} k^{a+q-3}\left(1 - \frac{k}{n}\right)^{b-2} \\
&\leq \frac{c_6}{n} \sum_{k=1}^{[n/2]} k^{a+q-3}\left(1 - \frac{k}{n}\right)^{b-2} + c_6 \sum_{k=[n/2]+1}^{n-1} k^{a+q-3}\left(1 - \frac{k}{n}\right)^{b-1} \\
&= \frac{c_6}{n} \sum_{k=1}^{[n/2]} k^{a+q-3}\left(1 - \frac{k}{n}\right)^{b-2} + c_6 n^{a+q-2}\left(\frac{1}{n} \sum_{k=[n/2]+1}^{n-1} \left(\frac{k}{n}\right)^{a+q-3}\left(1 - \frac{k}{n}\right)^{b-1}\right) \\
&= O(n^{a+q-2})
\end{aligned}
$$

since the first term is $O(n^{-1})$ and the second term is $O(n^{a+q-2})$ by the same reasoning as for $S_1(a, n)$.

Finally

$$
\begin{aligned}
S_3(1, n) &\leq \frac{c_7 \log n}{n} \sum_{k=1}^{n-1} k^{q-2}\left(1 - \frac{k}{n}\right)^{b-1} \\
&\leq \frac{c_7 \log n}{n} \sum_{k=1}^{n-1} k^{q-2}\left|\left(1 - \frac{k}{n}\right)^{b-1} - 1\right| + \frac{c_7 \log n}{n} \sum_{k=1}^{n-1} k^{q-2} \\
&= O(n^{q-2} \log n) + O(n^{-1} \log n),
\end{aligned}
$$

in view of the estimate for $S_1(a, n)$. Therefore $S_3(1, n) = O(n^{q-1})$ and the proof is complete. □

We provide a basic lemma concerning the total rates of the beta $(1, b)$-coalescent.

Lemma 6.4. *The total rates λ_n, $n \in \mathbb{N}$, of the* beta $(1, b)$-*coalescent are explicitly given by*

$$
\lambda_n = b \sum_{k=1}^{n-1} \frac{k}{b + k - 1} = b(n - 1) - b(b - 1)(\Psi(n + b - 1) - \Psi(b)), \quad n \in \mathbb{N}. \tag{30}
$$

Moreover, the total rates have the asymptotic expansion

$$
\lambda_n = bn - b(b - 1)\log n - b + b(b - 1)\Psi(b) + O(n^{-1}), \quad n \to \infty, \tag{31}
$$

18

and the inverse of the total rate λ_n has the asymptotic expansion

$$\frac{1}{\lambda_n} = \frac{1}{bn}\left(1 + (b-1)\frac{\log n}{n} + \frac{1 - (b-1)\Psi(b)}{n} + O\left(\frac{\log^2 n}{n^2}\right)\right), \quad n \to \infty. \tag{32}$$

Proof. Eq. (30) is well known (see, for example, [17, Appendix, Eq. (19)]). The expansion (31) follows directly from (30), since $\Psi(n + b - 1) = \log n + O(n^{-1})$ as $n \to \infty$. The last assertion (32) follows from

$$\begin{aligned}
&\frac{bn}{\lambda_n} - 1 - (b-1)\frac{\log n}{n} - \frac{1 - (b-1)\Psi(b)}{n} \\
&\qquad = \frac{bn^2 - \lambda_n\big(n + (b-1)\log n + 1 - (b-1)\Psi(b)\big)}{n\lambda_n} \\
&\qquad = \frac{O(\log^2 n)}{n\lambda_n} = O\left(\frac{\log^2 n}{n^2}\right),
\end{aligned} \tag{33}$$

where the very last equality holds, since $\lambda_n \sim bn$, and the equality before follows by plugging in (31) for the term λ_n occurring in the numerator of the fraction in (33) and multiplying everything out. $\square$

The next Lemma 6.5 provides an asymptotic expansion as $n \to \infty$ for sums of the form

$$\sum_{m=2}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)\log^p m}, \quad \alpha \in \mathbb{R}, p \in [0, \infty).$$

For parameters $\alpha > 0$ we will need an even sharper version (see Lemma 6.6 below), but we start with this simpler version, which holds for arbitrary $\alpha \in \mathbb{R}$. Given the overlap with the proof of the following Lemma 6.6 and given the fact that the proof is considerably simpler than that of Lemma 6.6, the proof of Lemma 6.5 is omitted.

Lemma 6.5. *For $\alpha \in \mathbb{R}$ and $p \in [0, \infty)$, as $n \to \infty$,*

$$\sum_{m=2}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)\log^p m} = \frac{n^\alpha}{\log^p n}\left(1 - \alpha\frac{\log n}{n} + O\left(\frac{1}{n}\right)\right).$$

The following Lemma 6.6 is a sharper version of Lemma 6.5 with the cost that it holds only for $\alpha > 0$. It will turn out (see the following Corollary 6.7 and the proof of Theorem 3.4) that the expansion in Lemma 6.6 is fundamental for the analysis of the moments of the number of collisions of the beta $(1, b)$-coalescent.

Lemma 6.6. *For $\alpha \in (0, \infty)$ and $p \in [0, \infty)$,*

$$\sum_{m=2}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)\log^p m} = \frac{n^\alpha}{\log^p n}\left(1 - \alpha\frac{\log n}{n} + \frac{\alpha\Psi(\alpha) + p}{n} + O\left(\frac{1}{n\log n}\right)\right)$$

as $n \to \infty$.

Proof. Note first that

$$\sum_{m=2}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)} \left(\frac{1}{\log^p n} + p\frac{-\log(m/n)}{\log^{p+1} n} \right)$$

$$= \frac{1}{\log^p n} \sum_{m=2}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)} + \frac{p}{\log^{p+1} n} \sum_{m=2}^{n-1} \frac{m^\alpha(-\log(m/n))}{(n-m)(n-m+1)}$$

$$= \frac{1}{\log^p n} n^\alpha \left(1 - \frac{\alpha \log n}{n} + \frac{\alpha \Psi(\alpha)}{n} + O\left(\frac{1}{n \log n} \right) \right) + \frac{p}{\log^{p+1} n} \left(n^{\alpha-1} \log n + O(n^{\alpha-1}) \right)$$

$$= \frac{n^\alpha}{\log^p n} \left(1 - \frac{\alpha \log n}{n} + \frac{\alpha \Psi(\alpha)+p}{n} + O\left(\frac{1}{n \log n} \right) \right).$$

Thus, it suffices to verify that

$$\sum_{m=2}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)} \left(\frac{1}{\log^p m} - \frac{1}{\log^p n} - p\frac{-\log(m/n)}{\log^{p+1} n} \right) = O\left(\frac{n^{\alpha-1}}{\log^{p+1} n} \right). \quad (34)$$

The function $f_{np} : (1, n] \to \mathbb{R}$, defined via

$$f_{np}(x) := \frac{1}{\log^p x} - \frac{1}{\log^p n} - p\frac{-\log(x/n)}{\log^{p+1} n},$$

has derivative

$$f'_{np}(x) = \frac{p}{x} \left(\frac{1}{\log^{p+1} n} - \frac{1}{\log^{p+1} x} \right) \le 0$$

and satisfies $f_{np}(n) = 0$. Thus, $f_{np} \ge 0$. In order to verify (34) we use a decomposition method. We split up the sum on the left hand side in (34) into two parts $\sum_{m=2}^{a_n} \ldots$ and $\sum_{m=a_n+1}^{n-1} \ldots$, and handle these two parts separately. We work with the sequence $(a_n)_{n\in\mathbb{N}}$ defined via $a_1 := 1$ and $a_n := \lfloor n/\log^{p+1} n \rfloor$ for $n \ge 2$. For the first part we obtain

$$0 \le \sum_{m=2}^{a_n} \frac{m^\alpha}{(n-m)(n-m+1)} f_{np}(m) \le \sum_{m=2}^{a_n} \frac{m^\alpha}{(n-m)(n-m+1)} \frac{1}{\log^p m}$$

$$\le \frac{n^\alpha}{\log^p 2} \sum_{m=2}^{a_n} \left(\frac{1}{n-m} - \frac{1}{n-m+1} \right) = O\left(\frac{n^{\alpha-1}}{\log^{p+1} n} \right),$$

since

$$\sum_{m=2}^{a_n} \left(\frac{1}{n-m} - \frac{1}{n-m+1} \right) = \frac{1}{n-a_n} - \frac{1}{n-1} = \frac{a_n - 1}{(n-a_n)(n-1)} \sim \frac{1}{n \log^{p+1} n}.$$

Moreover, for the second part we have

$$\sum_{m=a_n+1}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)} f_{np}(m)$$

$$= \sum_{m=a_n+1}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)} \left(\frac{1}{\log^p m} - \frac{1}{\log^p n} - p\frac{-\log(m/n)}{\log^{p+1} n}\right)$$

$$= \sum_{m=a_n+1}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)} \frac{\log^p n - \log^p m - p(-\log(m/n))\log^p m/\log n}{\log^p n \log^p m}$$

$$\leq \frac{n^\alpha}{\log^p n \log^p a_n} \sum_{m=a_n+1}^{n-1} \frac{\log^p n - \log^p m + p\log(m/n)\log^p m/\log n}{(n-m)(n-m+1)}$$

$$\sim \frac{n^\alpha}{\log^{2p} n} \sum_{m=a_n+1}^{n-1} \frac{\log^p n - \log^p m + p\log(m/n)\log^p m/\log n}{(n-m)(n-m+1)},$$

since $\log a_n \sim \log n$ as $n \to \infty$. Thus, it remains to verify that

$$\sum_{m=a_n+1}^{n-1} \frac{\log^p n - \log^p m + p\log(m/n)\log^p m/\log n}{(n-m)(n-m+1)} = O\left(\frac{\log^{p-1} n}{n}\right). \qquad (35)$$

Let us distinguish the two cases $p \geq 1$ and $p < 1$. Suppose first that $p \geq 1$. Then the map $x \mapsto x^p$ is convex on $[0,\infty)$. Thus, $y^p - x^p \leq p(y-x)y^{p-1}$ for all $x, y \in [0,\infty)$ with $x \leq y$. It follows that $y^p - x^p + p(x-y)x^p/y \leq p^2(y-x)^2 y^{p-2}$ for all $x, y \in [0,\infty)$ with $x < y$. Applying this inequality with $0 \leq x := \log m < y := \log n$ yields

$$0 \leq \sum_{m=a_n+1}^{n-1} \frac{\log^p n - \log^p m + p\log(m/n)\log^p m/\log n}{(n-m)(n-m+1)}$$

$$\leq \sum_{m=1}^{n-1} \frac{\log^p n - \log^p m + p\log(m/n)\log^p m/\log n}{(n-m)(n-m+1)}$$

$$\leq \sum_{m=1}^{n-1} \frac{p^2(\log n - \log m)^2 \log^{p-2} n}{(n-m)(n-m+1)} = p^2 \log^{p-2} n \sum_{m=1}^{n-1} \frac{\log^2(m/n)}{(n-m)(n-m+1)}.$$

Note that

$$n \sum_{m=1}^{n-1} \frac{\log^2(m/n)}{(n-m)(n-m+1)} = \frac{1}{n} \sum_{m=1}^{n-1} \frac{\log^2(m/n)}{(1-m/n)(1-(m-1)/n)}$$

$$\to \int_0^1 \frac{\log^2 x}{(1-x)^2}\, dx = \Gamma(3)\zeta(2) = \frac{\pi^2}{3} \in \mathbb{R},$$

where the particular value $\Gamma(3)\zeta(2)$ of last integral is obtained by choosing $s := 2$ in the chain of equalities

$$\int_0^1 \frac{(-\log(1-u))^s}{u^2}\, du = \int_0^\infty \frac{x^s e^x}{(e^x-1)^2}\, dx = \int_0^\infty \frac{sx^{s-1}}{e^x-1}\, dx = \Gamma(s+1)\zeta(s), \quad s > 1,$$

21

which are based on the substitution $x = -\log(1-u)$, partial integration, and on formula 23.2.7 in [1]. Thus, the expression on the left hand side in (35) is even $O((\log^{p-2} n)/n)$. In particular, (35) holds. Suppose now that $p \in [0,1)$. Then the map $x \mapsto x^p$ is concave on $[0,\infty)$. Thus, $y^p - x^p \le p(y-x)x^{p-1}$ for all $x, y \in (0,\infty)$ with $x \le y$. It follows that $y^p - x^p + p(x-y)x^p/y \le p(y-x)^2 x^{p-1}/y$ for all $x, y \in (0,\infty)$ with $x \le y$. Applying this inequality with $0 < x := \log m < y := \log n$ yields

$$
\begin{aligned}
0 &\le \sum_{m=a_n+1}^{n-1} \frac{\log^p n - \log^p m + p\log(m/n)\log^p m/\log n}{(n-m)(n-m+1)} \\
&\le \sum_{m=a_n+1}^{n-1} \frac{p(\log n - \log m)^2(\log^{p-1} m)/\log n}{(n-m)(n-m+1)} \\
&\le \frac{p\log^{p-1} a_n}{\log n} \sum_{m=a_n+1}^{n-1} \frac{\log^2(m/n)}{(n-m)(n-m+1)} \\
&\le \frac{p\log^{p-1} a_n}{\log n} \sum_{m=1}^{n-1} \frac{\log^2(m/n)}{(n-m)(n-m+1)} = O\left(\frac{\log^{p-2} n}{n}\right),
\end{aligned}
$$

since $\log a_n \sim \log n$ and the last sum is $O(1/n)$ as shown above. Again, (35) holds. $\qquad\square$

The following Corollary 6.7 is essentially obtained by combining the three Lemmata 6.4, 6.5 and 6.6. It provides an asymptotic expansion for the sum $\sum_{m=2}^{n-1} p_{n,m}^{(1)} m^\alpha/\log^p m$. This expansion is a key tool for the proof of Theorem 3.4.

Corollary 6.7. *Fix $\alpha \in [1,\infty)$ and $p \in [0,\infty)$. For the* beta $(1,b)$-*coalescent with parameter $b \in (0,\infty)$,*

$$
\sum_{m=2}^{n-1} p_{n,m}^{(1)} \frac{m^\alpha}{\log^p m} = \frac{n^\alpha}{\log^p n}\left(1 - \alpha\frac{\log n}{n} + \frac{c_{b,\alpha,p}}{n} + O\left(\frac{1}{n\log n}\right)\right), \qquad n \to \infty, \qquad (36)
$$

where $c_{b,\alpha,p} := (\alpha+b-1)\Psi(\alpha+b-1)+p+1+(1-b)\Psi(b) = (\alpha+b-1)\Psi(\alpha+b)+p-(b-1)\Psi(b)$.

Remark 6.8. The following proof shows that Corollary 6.7 even holds for the slightly larger range of parameters $\alpha, b \in (0,\infty)$ satisfying $\alpha + b - 1 > 0$. However, we need Corollary 6.7 only for $\alpha \in [1,\infty)$ and $b \in (0,\infty)$, in which case $\alpha + b - 1 > 0$ automatically holds.

Proof. Let $g_{nm} := \lambda_n \mathbb{P}\{I_n = n-m\}$ denote the rate at which the block counting process moves from the state n to the state $m \in \{1,\ldots,n-1\}$. It suffices to verify that

$$
\sum_{m=2}^{n-1} g_{nm}\frac{m^\alpha}{\log^p m}
$$
$$
= b\frac{n^{\alpha+1}}{\log^p n}\left(1 - (\alpha+b-1)\frac{\log n}{n} + \frac{(\alpha+b-1)\Psi(\alpha+b-1)+p}{n} + O\left(\frac{1}{n\log n}\right)\right) \quad (37)
$$

since (36) then follows from $p_{n,m}^{(1)} = g_{nm}/\lambda_n$ by multiplying (37) with (32). Note that

$$
g_{nm} = b\frac{n!}{\Gamma(b+n-1)}\frac{1}{(n-m)(n-m+1)}\frac{\Gamma(b+m-1)}{(m-1)!}, \qquad 1 \le m < n.
$$

Since the first fraction has expansion

$$\frac{n!}{\Gamma(b+n-1)} = \frac{1}{n^{b-2}}\left(1 - \binom{b-1}{2}\frac{1}{n} + O\left(\frac{1}{n^2}\right)\right), \tag{38}$$

it hence suffices to verify that

$$\sum_{m=2}^{n-1} \frac{1}{(n-m)(n-m+1)}\frac{\Gamma(b+m-1)}{(m-1)!}\frac{m^\alpha}{\log^p m}$$
$$= \frac{n^{\alpha+b-1}}{\log^p n}\left(1 - (\alpha+b-1)\frac{\log n}{n}\right.$$
$$\left. + \frac{\binom{b-1}{2} + (\alpha+b-1)\Psi(\alpha+b-1) + p}{n} + O\left(\frac{1}{n\log n}\right)\right), \tag{39}$$

since (37) then follows by multiplying (39) with (38). Thus, it remains to verify (39). Since for all $m \in \mathbb{N}$ and all $b \in (0, \infty)$, the Pochhammer like expression $\Gamma(b+m-1)/(m-1)!$ appearing on the left hand side in (39) is bounded below and above by

$$m^{b-1} + \binom{b-1}{2}m^{b-2} \leq \frac{\Gamma(b+m-1)}{(m-1)!} \leq m^{b-1} + \binom{b-1}{2}m^{b-2} + K_b m^{b-3},$$

where $K_b := \Gamma(b) - 1 - \binom{b-1}{2}$, (39) follows by plugging in these lower and upper bounds on the left hand side in (39) and applying afterwards Lemma 6.6 with α replaced by $\alpha + b - 1 > 0$ and noting that

$$\sum_{m=2}^{n-1} \frac{m^{\alpha+b-2}}{(n-m)(n-m+1)\log^p m} = \frac{n^{\alpha+b-2}}{\log^p n}\left(1 + O\left(\frac{\log n}{n}\right)\right)$$

and that

$$\sum_{m=2}^{n-1} \frac{m^{\alpha+b-3}}{(n-m)(n-m+1)\log^p m} = O\left(\frac{n^{\alpha+b-3}}{\log^p n}\right)$$

by Lemma 6.5. The proof is complete. $\qquad\square$

References

[1] ABRAMOWITZ, M. AND STEGUN, I. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York.

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

[3] BERTOIN, J. AND PITMAN, J. (2000). Two coalescents derived from the ranges of stable subordinators. *Electron. J. Probab.* **5**, 1–17.

[4] BERESTYCKI, N. (2009). Recent progress in coalescent theory, *Ensaios Mathematicos* **16**, 1–193.

[5] BIRKNER, M., BLATH, J., CAPALDO, M., ETHERIDGE, A., MÖHLE, M., SCHWEINSBERG, J. AND WAKOLBINGER, A. (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.* **10**, 303–325.

[6] BOLTHAUSEN, E. AND SZNITMAN, A.-S. (1998). On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197**, 247–276.

[7] 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.

[8] 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.

[9] 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 Struct. Algorithms* **34**, 318–336.

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

[11] GIVENS, C. R., SHORTT, R. M. (1984). A class of Wasserstein metrics for probability distributions. *Michigan Math. J.* **31**, 231–240.

[12] GNEDIN, A., IKSANOV, A. AND MARYNYCH, A. (2011). Lambda-coalescents with dust component. *J. Appl. Probab.* **48**, 1133–1151.

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

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

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

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

[17] HUILLET, T. AND MÖHLE, M. (2011). On the extended Moran model and its relation to coalescents with multiple collisions. *Theor. Popul. Biol.* 10/2011, DOI: 10.1016/j.tpb.2011.09.004

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

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

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

[21] Johnson, O. and Samworth, R. (2005). Central limit theorem and convergence to stable laws in Mallows distance. *Bernoulli* **11**, 829–845.

[22] Kersting, G. (2012+). The asymptotic distribution of the length of Beta-coalescent trees. *Ann. Appl. Probab.*, to appear.

[23] Kingman, J. F. C. (1982). The coalescent. *Stoch. Process. Appl.* **13**, 235–248.

[24] Möhle, M. (2006). On the number of segregating sites for populations with large family sizes. *Adv. Appl. Probab.* **38**, 750–767.

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

[26] Neininger, R. and Rüschendorf, L. (2004). On the contraction method with degenerate limit equation. *Ann. Probab.* **32**, 2838–2856.

[27] Panholzer, A. (2004). Destruction of recursive trees, in: Mathematics and Computer Science, vol. III, Birkhäuser, Basel, pp. 267–280.

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

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

[30] Schweinsberg, J. (1999). A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Commun. Probab.* **5**, 1–11.

[31] Tavaré, S. (2004). Ancestral inference in population genetics. *Springer Lecture Notes in Math.*, **1837**.