Adv. Appl. Prob. **46,** 496–515 (2014)
Printed in Northern Ireland

ON ASYMPTOTICS OF THE BETA COALESCENTS

ALEXANDER GNEDIN,* *University of London*
ALEXANDER IKSANOV ** AND
ALEXANDER MARYNYCH,*** *National Taras Shevchenko University of Kyiv*
MARTIN MÖHLE,**** *University of Tübingen*

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 ∞. 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 [29] and Sagitov [30] introduced exchangeable coalescent processes with multiple collisions, also known as Λ-coalescents. A counting process associated with the Λ-coalescent is a Markov chain $\Pi_n = (\Pi_n(t))_{t\ge 0}$ with right-continuous paths, which starts with n particles, $\Pi_n(0) = n$, and terminates when a sole particle remains. The particles merge according to the rule: for each $t \ge 0$, when the number of particles is $\Pi_n(t) = m > 1$, each k tuple of them merge into one particle at probability rate

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

where Λ is a given finite measure on the unit interval. The merging of two or more particles is called a *collision*. With every collision, Π_n jumps to a smaller value. When Λ is a Dirac mass

Received 25 September 2012; revision received 3 May 2013.
* Postal address: School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, UK. Email address: a.gnedin@qmul.ac.uk
** Postal address: Faculty of Cybernetics, National Taras Shevchenko University of Kyiv, 01601 Kyiv, Ukraine. Email address: iksan@univ.kiev.ua
*** Postal address: National Taras Shevchenko University of Kyiv, 01601 Kyiv, Ukraine. Email address: marynych@unicyb.kiev.ua
**** Postal address: Mathematical Institute, University of Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany. Email address: martin.moehle@uni-tuebingen.de

at 0, the Λ-coalescent is the classical Kingman coalescent [25], in which every pair of particles merges 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

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

where $\mathrm{B}(\cdot,\cdot)$ denotes Euler's beta function. This class is amenable to analysis due to the fact that the transition rates (1.1) can be expressed in terms of $\mathrm{B}(\cdot,\cdot)$. For this reason and due to multiple connections with Lévy processes and random trees, beta coalescents have been the subject of intensive research; see [2], [3], [5], [8], [9], [17], [20], and [29]. We refer the reader 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 through 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 $r > 0$.

A principal contribution of this paper is the proof of convergence in distribution to a 1-stable law for X_n and L_n as $n \to \infty$. As in much of the previous work (see, for instance, [13] and [21]), 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.

Our second new contribution is the derivation of 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 complementary to the results on convergence in distribution. The proofs of these asymptotic expansions are based on the method of sequential approximations, similarly as in [22].

The rest of the paper is organized as follows. In Section 2 we give 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. In Section 4 we recall the definition and properties of a Wasserstein distance. In Section 5 we provide proofs of the main results. Some auxiliary lemmas are collected in Appendix A.

2. A summary of limit laws for beta coalescents

In Tables 1–3 we summarize the limit laws for X_n, L_n, and the absorption time $\tau_n := \min\{t : \Pi_n(t) = 1\}$ of the coalescent. The distributions that appear in the tables are as follows.

- $\mathcal{N}$, the standard normal distribution.
- $\mathcal{S}_\alpha$ with $1 < \alpha < 2$, the (spectrally negative) α-stable distribution with characteristic function

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

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

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

- $\mathcal{E}_\gamma(a, b)$ with $a, b, \gamma > 0$, the law of the *exponential functional* $\int_0^\infty e^{-\gamma S_{a,b}(t)}\,dt$, where $(S_{a,b}(t))_{t\geq 0}$ is a drift-free subordinator with Laplace exponent $\Phi_{a,b}(z) = \int_0^1 (1-(1-x)^z)x^{a-3}(1-x)^{b-1}\,dx,\ z \geq 0$.
- $\mathcal{G}$, the Gumbel with distribution function $x \mapsto \exp\{-e^{-x}\},\ x \in \mathbb{R}$.
- ρ, the 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
$0 < a < 1$	$b > 0$	$(1-a)n$	$(1-a)n^{1/(2-a)}$	$\mathcal{S}_{2-a}$	This paper, [7], [13], and [21] ($b = 1$)
$a = 1$	$b > 0$	$\dfrac{n\log(n\log n)}{(\log n)^2}$	$\dfrac{n}{(\log n)^2}$	$\mathcal{S}_1$	This paper and [9], [20] ($b = 1$)
$1 < a < 2$	$b > 0$	0	$\dfrac{\Gamma(a)}{2-a}n^{2-a}$	$\mathcal{E}_{2-a}(a, b)$	[14] and [18]
$a = 2$	$b > 0$	$(2r_1)^{-1}(\log n)^2$	$(3^{-1}r_1^{-3}r_2\log^3 n)^{1/2}$	$\mathcal{N}$	[14] and [22]
$a > 2$	$b > 0$	$m_1^{-1}\log n$	$(m_1^{-3}m_2\log n)^{1/2}$	$\mathcal{N}$	[14] and [15]

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	[32]
$a = 1$	$b = 1$	$\log\log n$	1	$\mathcal{G}$	[17] and [11]
$1 < a < 2$	$b > 0$	$\mathtt{m}^{-1}\log n$	$(\mathtt{m}^{-3}\mathtt{s}^2\log n)^{1/2}$	$\mathcal{N}$	[14]
$a = 2$	$b > 0$	$c_1^{-1}\log n$	$(c_1^{-3}c_2\log n)^{1/2}$	$\mathcal{N}$	[14]
$a > 2$	$b > 0$	$(\gamma m_1)^{-1}\log n$	$\gamma^{-1}(m_1^{-3}(m_2 + m_1^2)\log n)^{1/2}$	$\mathcal{N}$	[14] and [15]

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] and [32]
$0 < a < \dfrac{3-\sqrt{5}}{2}$	$b = 2 - a$	$c_1 n^a$	1	exists	[24]
$a = \dfrac{3-\sqrt{5}}{2}$	$b = 2 - a$	$c_1 n^a$	$c_2(\log n)^{1/\alpha}$	$\mathcal{S}_{2-a}$	[24]
$\dfrac{3-\sqrt{5}}{2} < a < 1$	$b = 2 - a$	$c_1 n^a$	$c_2(\beta n^{-\beta})^{1/\alpha}$	$\mathcal{S}_{2-a}$	[24]
$a = 1$	$b > 0$	$\dfrac{n\log(n\log n)}{b(\log n)^2}$	$\dfrac{n}{b(\log n)^2}$	$\mathcal{S}_1$	This paper and [8] ($b = 1$),
$a > 1$	$b > 0$	0	$\mathrm{B}(a, b)n$	$\mathcal{E}_1(a, b)$	[26], and [27]

In Table 1 $r_1 = \zeta(2, b)$ and $r_2 = 2\zeta(3, b)$, where $\zeta(\cdot, \cdot)$ is the Hurwitz zeta function; $m_1 = \Psi(a-2+b) - \Psi(b)$ and $m_2 = \Psi'(b) - \Psi'(a-2+b)$, where $\Psi(\cdot)$ is the logarithmic derivative of the gamma function.

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 [20], where the coupling of a random walk with a barrier was exploited, and the technique was further extended in [21] 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 measures Λ such that $\int_0^1 x^{-1}\Lambda(\mathrm{d}x) < \infty$ [13]–[15], and [18]. 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.

In Table 2,

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

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

For the case in which $a \in (0, 1)$ and $b > 0$, the beta (a, b)-coalescent has the property of coming down from ∞ [31], 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 of [14]. The case in which $a = 1$ and $b \neq 1$ is open; in this case the coalescent does not come down from ∞.

In Table 3, $\alpha = 2-a$, $\beta = 1+\alpha-\alpha^2$, $c_1 = \Gamma(\alpha+1)(\alpha-1)/(2-\alpha)$, and $c_2 = \Gamma(\alpha+1)(\alpha-1)^{1+\alpha^{-1}}/(\cos(\pi\alpha/2)\Gamma^{\alpha^{-1}}(2-\alpha))$.

In [26] the weak convergence of properly normalized L_n was proved for Λ-coalescents with a 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 $\lambda_{m,k}$, $2 \le k \le m$, is given by (1.1). The total transition rate from state $m \ge 2$ is

$$\lambda_m := \sum_{k=2}^{m}\binom{m}{k}\lambda_{m,k} = \int_0^1 (1-mx(1-x)^{m-1}-(1-x)^m)x^{-2}\Lambda(\mathrm{d}x). \tag{3.1}$$

The first decrement I_n of Π_n has distribution

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

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

The strong Markov property of the coalescent entails the distributional recurrences

$$X_1 = 0, \qquad X_n \stackrel{D}{=} 1 + X'_{n-I_n}, \qquad n \in \mathbb{N} \setminus \{1\}, \tag{3.2}$$

$$\tau_1 = 0, \qquad \tau_n \stackrel{D}{=} T_n + \tau'_{n-I_n}, \qquad n \in \mathbb{N} \setminus \{1\},$$

$$L_1 = 0, \qquad L_n \stackrel{D}{=} nT_n + L'_{n-I_n}, \qquad n \in \mathbb{N} \setminus \{1\}, \tag{3.3}$$

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

Letting Λ be defined as in (1.2) with $a \in (0, 1]$, define

$$p^{(a)}_{n,k} := \mathbb{P}\{I_n = n - k\}, \qquad k = 1, \dots, n-1. \tag{3.4}$$

Using the leading terms of asymptotic relations (A.3), (A.4), and (A.5) below, we infer that

$$\lim_{n\to\infty} p^{(a)}_{n,n-k} = \frac{(2-a)\Gamma(k+a-1)}{\Gamma(a)(k+1)!} =: p^{(a)}_k, \qquad k \in \mathbb{N};$$

hence (see also [7, Lemma 2.1]),

$$I_n \xrightarrow{D} \xi \quad \text{as } n \to \infty,$$

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

Consider a zero-delayed random walk $(S_n)_{n\in\mathbb{N}_0}$ defined by $S_0 := 0$ and $S_n := \xi_1 + \cdots + \xi_n$ for $n \in \mathbb{N}$, where $\xi_1, \xi_2, \dots$ are independent copies of ξ with distribution $(p^{(a)}_k)_{k\in\mathbb{N}}$, and let $(N_n)_{n\in\mathbb{N}_0}$ be the associated first passage time sequence defined by $N_n := \inf\{k \ge 0\colon S_k \ge n\}$, $n \in \mathbb{N}_0$. It is plain that

$$N_0 = 0, \qquad N_n \stackrel{D}{=} 1 + N'_{n-\xi\wedge n} = 1 + N'_{n-\xi}\, \mathbf{1}_{\{\xi<n\}}, \qquad n \in \mathbb{N}, \tag{3.5}$$

where N'_k is independent of ξ and distributed as N_k for each $k \in \mathbb{N}$. Comparing (3.2) and (3.5) we can expect that, if N_n (properly centered and normalized) converges weakly to some proper and nondegenerate 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 [13] for $a \in (0, 1)$ and $b > 0$, and in [21] for $a \in (0, 1]$ and $b = 1$ to derive the limit distribution of X_n from that of N_n. For $a \in (0, 1]$ and $b > 0$, we will use a method based on probability metrics to show the stable limits.

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* $b > 0$,

$$\frac{X_n - (1-a)n}{(1-a)n^{1/(2-a)}} \xrightarrow{D} \mathcal{S}_{2-a},$$

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

$$\frac{\log^2 n}{n} X_n - \log n - \log\log n \xrightarrow{D} \mathcal{S}_1.$$

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 [26]) of the beta $(1, b)$-coalescent.

Corollary 3.1. *For the total branch length L_n in the* beta $(1, b)$*-coalescent, we have, as $n \to \infty$,*

$$\frac{b \log^2 n}{n} L_n - \log n - \log \log n \xrightarrow{\mathrm{D}} \mathcal{S}_1.$$

Corollary 3.2. *For the number of segregating sites M_n in the* beta $(1, b)$*-coalescent, we have, as $n \to \infty$,*

$$\frac{b \log^2 n}{rn} M_n - \log n - \log \log n \xrightarrow{\mathrm{D}} \mathcal{S}_1,$$

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 3.2. *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*

$$\mathbb{E}X_n^j = \frac{n^j}{\log^j n}\left(1 + \frac{m_j}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right) \quad \text{as } n \to \infty, \tag{3.6}$$

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 further information on the coefficients m_j, $j \in \mathbb{N}$, we refer the reader to (5.5) below in the proof of the following corollary, which provides asymptotic expansions for the central moments of X_n in the beta $(1, b)$-coalescent.

Corollary 3.3. *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*

$$\mathbb{E}(X_n - \mathbb{E}X_n)^j = \frac{(-1)^j}{j} \mathrm{B}(b, j - 1) \frac{n^j}{\log^{j+1} n} + O\left(\frac{n^j}{\log^{j+2} n}\right) \quad \text{as } n \to \infty. \tag{3.7}$$

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

Remark 3.1. For $b = 1$, (3.7) reduces to the asymptotic expansion (see [28, p. 277 or Theorem 2.1 with $\alpha = 0$])

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

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

Proposition 3.1. *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*

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

where the sequence $(m_j)_{j \in \mathbb{N}_0}$ is defined as in Theorem 3.2. Moreover, for $j \in \{2, 3, \ldots\}$, the

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

jth central moment of L_n has the asymptotic expansion

$$\mathbb{E}(L_n - \mathbb{E}L_n)^j = \frac{(-1)^j}{jb^j}\mathrm{B}(b, j-1)\frac{n^j}{\log^{j+1} n} + O\bigg(\frac{n^j}{\log^{j+2} n}\bigg) \quad \textit{as } n \to \infty.$$

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

Proposition 3.1 indicates that bL_n essentially behaves as X_n, in agreement with the comparison of Theorem 3.1(ii) and Corollary 3.2. The proof of Proposition 3.1 essentially follows the same lines as the analogous proofs of Theorem 3.2 and Corollary 3.3 for X_n. Instead of the distributional recurrence (3.2) for $(X_n)_{n\in\mathbb{N}}$, we have to work with the distributional recurrence (3.3) for $(L_n)_{n\in\mathbb{N}}$. Since the expansion of $\mathbb{E}T_n = 1/\lambda_n$ is known (see Lemma A.4), the proofs concerning X_n are readily adapted for L_n. A proof of Proposition 3.1 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 $\mathbb{E}(M_n)_j = r^j\mathbb{E}L_n^j$, $j \in \mathbb{N}_0$, where $(M_n)_0 := 1$ and $(M_n)_j := M_n(M_n-1)\cdots(M_n-j+1)$ for $j \in \mathbb{N}$.

4. Probability distances χ_T and d_q

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

$$\chi_T(X, Y) = \sup_{|t|\le T} |\mathbb{E}\mathrm{e}^{\mathrm{i}tX} - \mathbb{E}\mathrm{e}^{\mathrm{i}tY}|.$$

By the continuity theorem for the characteristic functions, convergence in distribution $Z_n \xrightarrow{\mathrm{D}} Z$ 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 Wasserstein distance on $\mathcal{D}_q$ is defined by

$$d_q(X, Y) = \inf \mathbb{E}|\widehat{X} - \widehat{Y}|^q, \tag{4.1}$$

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

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

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

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

(Inf) *The infimum in (4.1) is attained for some coupling.*

(Rep) *The Kantorovich–Rubinstein representation holds, i.e.*

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

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

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

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

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

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

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

Proof. We refer the reader to [12] and [23] 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 $X + Z \stackrel{\mathrm{D}}{=} X' + Z$, $Y + Z \stackrel{\mathrm{D}}{=} Y' + Z$, and the definition of d_q entails

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

Property (Conv): the convergence of moments is easy; the rest is a consequence of Lemma 4.1 to follow.

Lemma 4.1. *For* $T > 0$ *and* $q \in (0, 1]$, *there exists a constant* $C = C_{T,q} > 0$ *such that* $\sup_{|t|\le T} |\mathbb{E}e^{itX} - \mathbb{E}e^{itY}| \le C d_q(X, Y)$.

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 $|e^{ix} - e^{iy}| = 2|\sin(x-y)/2| \le 2^{1-q} M_q |x-y|^q$ for $q \in (0, 1]$ and $x, y \in \mathbb{R}$, where $M_q := \sup_{u>0} |\sin u| u^{-q} < \infty$. Hence,

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

as required.

5. Proofs

5.1. Proof of Theorem 3.1

Suppose that $a = 1$. Set $a_n := n^{-1} \log^2 n$ and $b_n := \log n + \log\log n$ for $n \in \mathbb{N}$. It suffices to show that $\lim_{n\to\infty} \chi_T(a_n X_n - b_n, \mathcal{S}_1) = 0$ for every $T > 0$. The triangle inequality yields

$$\chi_T(a_n X_n - b_n, \mathcal{S}_1) \le \chi_T(a_n X_n - b_n, a_n N_n - b_n) + \chi_T(a_n N_n - b_n, \mathcal{S}_1).$$

The second term converges to 0 by Proposition 2 of [20] on the stable limit for the number of renewals. In view of Lemma 4.1, to prove convergence to 0 of the first term, it suffices to check that $\lim_{n\to\infty} d_q(a_n X_n - b_n, a_n N_n - b_n) = 0$ for some $q \in (0, 1]$, which in view of the properties (Hom) and (Aff) in Proposition 4.1 amounts to the estimate

$$d_q(X_n, N_n) = o(n^q \log^{-2q} n) \quad \text{as } n \to \infty. \tag{5.1}$$

Now assume that $a \in (0, 1)$. By Theorem 7 of [10] we have $\chi_T(a_n N_n - b_n, \mathcal{S}_{2-a}) \to 0$ for every $T > 0$ with $a_n := (1-a)^{-1} n^{-1/(2-a)}$ and $b_n := n^{(1-a)/(2-a)}$. By the same reasoning as above, proving Theorem 3.1 for $a \in (0, 1)$ reduces to showing that

$$d_q(X_n, N_n) = o(n^{q/(2-a)}) \quad \text{as } n \to \infty, \tag{5.2}$$

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

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

Using recurrences (3.2) for X_n and (3.5) for N_n, we obtain

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

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

$$t_n \le c_n + \sum_{k=1}^{n-1} \mathbb{P}\{I_n = n-k\} t_k. \tag{5.3}$$

We will use (5.3) 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 property (Inf) of Proposition 4.1) $d_q(I_n, \xi\wedge n) = \mathbb{E}|\hat{I}_n - \hat{\xi}\wedge n|^q$. Let $(\hat{N}_k)_{k\in\mathbb{N}}$ be a copy of $(N_k)_{k\in\mathbb{N}}$ independent of $(\hat{I}_n, \hat{\xi})$. Since $(\hat{I}_n, \hat{\xi}, (\hat{N}_k))$ is a particular coupling, we have $c_n = d_q(N'_{n-I_n}, N'_{n-(\xi\wedge n)}) \le \mathbb{E}|\hat{N}_{n-\hat{I}_n} - \hat{N}_{n-(\hat{\xi}\wedge n)}|^q$. Using the stochastic inequality $N_{x+y} - N_x \overset{\mathrm{D}}{\le} N_y$, $x, y \in \mathbb{N}$, 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, $c_n \le \mathbb{E}|\hat{I}_n - \hat{\xi}\wedge n|^q = d_q(I_n, \xi\wedge n)$. Now we invoke the Kantorovich–Rubinstein representation, property (Rep) of Proposition 4.1, for d_q. Set $\mathcal{F}_{q,0} := \mathcal{F}_q \cap \{f\colon f(0) = 0\}$, and note that $f \in \mathcal{F}_{q,0}$ implies that $|f(x)| \le |x|^q$, $x \in \mathbb{R}$. We have

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

For appropriate $q \in (0, 1]$ (to be specified below) such that $a + q > 1$, use Lemma A.3 in Appendix A 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 A.1.

If $a \in (0, 1)$, we can take $q = 1$. Then Lemma A.1 applies with $\psi_n = n$ and $r_n = Mn^{a-1}$ (large enough M) and gives the estimate $d_q(X_n, N_n) = O(n^a)$, which implies (5.2).

For the case $a = 1$, application of the same lemma with $\psi_n = n/(\log(n+1))$ and $r_n = Mn^{q-1}$ (large enough M) leads to $t_n \le Mn^q(\log n)^{-1}$. Thus, (5.1) holds for $q \in (0, \frac{1}{2})$. The proof is complete.

5.2. Proof of Corollaries 3.1 and 3.2

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

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

it is enough to show that $((\log^2 n)/n)(bL_n - X_n) \to 0$ in L_2.

Let the T_j be independent exponential variables with rates λ_j, $j \geq 2$. Assuming that the T_j are 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 that 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 as $\sum_{r=0}^{k-1} i_r T_{i_r}$ for $n \in \mathbb{N} \setminus \{1\}$.

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

$$\begin{aligned}
\mathbb{E}(bL_n - X_n)^2 &= \sum_{k,\hat{\imath}} \mathbb{P}\{A_{k,\hat{\imath}}\}\mathbb{E}\bigg(\sum_{r=0}^{k-1}(bi_r T_{i_r} - 1)\bigg)^2 \\
&= \sum_{k,\hat{\imath}} \mathbb{P}\{A_{k,\hat{\imath}}\}\bigg(\sum_{r=0}^{k-1}\mathbb{E}(bi_r T_{i_r} - 1)^2 + \sum_{r,s=0,\, r\neq s}^{k-1}\mathbb{E}(bi_r T_{i_r} - 1)(bi_s T_{i_s} - 1)\bigg).
\end{aligned}$$

Furthermore, $\lambda_n = bn + O(\log n)$ as $n \to \infty$ for $a = 1$ and $b > 0$ (see (A.5) below), which implies that $|\mathbb{E}(bkT_k - 1)| = O(k^{-1}\log k)$ and $\mathbb{E}(bkT_k - 1)^2 = 1 + O(k^{-1}\log k)$. Therefore,

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

and the convergence in L_2 follows. Corollary 3.2 follows from the fact that, given L_n, the distribution of M_n is Poisson with mean rL_n. See Corollary 6.2 of [8] for details.

5.3. Proofs of Theorem 3.2 and Corollary 3.3

Let us verify (3.6) by induction on $j \in \mathbb{N}$. From (3.2), it follows that $a_1 := \mathbb{E}X_1 = 0$ and $a_n := \mathbb{E}X_n = 1 + \sum_{m=2}^{n-1} p_{n,m}^{(1)} a_m$, $n \geq 2$. 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 $n \geq 2$, satisfies the recursion

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

for $n \geq 2$, where $q_n := 1 - n/\log n + \sum_{m=2}^{n-1} p_{n,m}^{(1)} m/\log m$, $n \geq 2$. By Corollary A.1 (applied with $\alpha = 1$ and $p = 1$),

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

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

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 $n \geq 2$, therefore satisfies the recursion

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

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

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

since $q_n = m_1/\log n + O(1/\log^2 n)$. By Lemma A.2 (applied with $\alpha = 1$ and $p = 3$), it follows that $c_n = O(n/\log^3 n)$. Thus, (3.6) 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

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

for $n \geq 2$, where $q_{n,j} := \sum_{i=0}^{j-1}\binom{j}{i}(-1)^{j-1-i}\mathbb{E}X_n^i$, $n \geq 2$. 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)

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

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 $n \geq 2$, satisfies the recursion $b_{n,j} = q_{n,j}' + \sum_{m=2}^{n-1} p_{n,m}^{(1)} b_{m,j}$, $n \geq 2$, where

$$q_{n,j}' := q_{n,j} - \frac{n^j}{\log^j n} + \sum_{m=2}^{n-1}\frac{p_{n,m}^{(1)} m^j}{\log^j m}, \qquad n \geq 2.$$

By Corollary A.1 (applied with $\alpha = j$ and $p = j$),

$$\begin{aligned} q_{n,j}' &= j\frac{n^{j-1}}{\log^{j-1} n} + j m_{j-1}\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right) - \frac{n^j}{\log^j n} \\ &\quad + \frac{n^j}{\log^j n} - j\frac{n^{j-1}}{\log^{j-1} n} + \kappa_j\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right) \\ &= j m_j\frac{n^{j-1}}{\log^j n} + O\left(\frac{n^{j-1}}{\log^{j+1} n}\right), \end{aligned}$$

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 $n \geq 2$, therefore satisfies the recursion $c_{n,j} = q_{n,j}'' + \sum_{m=2}^{n-1} p_{n,m}^{(1)} c_{m,j}$, $n \geq 2$, where $q_{n,j}'' := q_{n,j}' - m_j n^j/\log^{j+1} n + m_j\sum_{m=2}^{n-1} p_{n,m}^{(1)} m^j/\log^{j+1} m$,

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

$n \geq 2$. By Corollary A.1 (applied with $\alpha = j$ and $p = j+1$),

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

By Lemma A.2 (applied with $\alpha := j$ and with $p := j+2$), it follows that $c_{n,j} = O(n^j/\log^{j+2} n)$. Thus, (3.6) holds for j, which completes the induction and the proof of Theorem 3.2.

We now turn to the proof of Corollary 3.3. Let us first verify that the sequence $(m_j)_{j\in\mathbb{N}_0}$, recursively defined in Theorem 3.2, satisfies the inversion formula

$$\sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} m_i = \frac{(-1)^j}{j} \mathrm{B}(b, j-1), \qquad j \in \mathbb{N} \setminus \{1\}. \tag{5.4}$$

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

$$m_j = \sum_{l=1}^{j} (m_l - m_{l-1}) = \sum_{l=1}^{j} \frac{\kappa_l}{l} = \sum_{l=1}^{j} \biggl(2 + \frac{1}{l}\sum_{i=0}^{l-1} \Psi(b+i)\biggr) = 2j + \sum_{i=0}^{j-1} \Psi(b+i) \sum_{l=i+1}^{j} \frac{1}{l}. \tag{5.5}$$

By (5.5), for $j \in \{2, 3, \ldots\}$,

$$\begin{aligned} \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} m_i &= \sum_{i=1}^{j} \binom{j}{i} (-1)^{j-i} \biggl(2i + \sum_{k=0}^{i-1} \Psi(b+k) \sum_{l=k+1}^{i} \frac{1}{l}\biggr) \\ &= \sum_{i=1}^{j} \binom{j}{i} (-1)^{j-i} \sum_{k=0}^{i-1} \Psi(b+k) \sum_{l=k+1}^{i} \frac{1}{l} \\ &= \sum_{k=0}^{j-1} \Psi(b+k) \sum_{l=k+1}^{j} \frac{1}{l} \sum_{i=l}^{j} \binom{j}{i} (-1)^{j-i} \\ &= \sum_{k=0}^{j-1} \Psi(b+k) \sum_{l=k+1}^{j} \frac{1}{l} \binom{j-1}{l-1} (-1)^{j-l} \\ &= \frac{1}{j} \sum_{k=0}^{j-1} \Psi(b+k) \sum_{l=k+1}^{j} \binom{j}{l} (-1)^{j-l} \\ &= \frac{1}{j} \sum_{k=0}^{j-1} \Psi(b+k) \binom{j-1}{k} (-1)^{j-1-k}. \end{aligned}$$

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

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

$$\begin{aligned}\sum_{i=0}^{j}\binom{j}{i}(-1)^{j-i}m_i &= \frac{1}{j}\sum_{k=0}^{j-2}(-1)^{j-2-k}\binom{j-2}{k}(\Psi(b+k+1)-\Psi(b+k))\\ &= \frac{(-1)^j}{j}\sum_{k=0}^{j-2}(-1)^k\binom{j-2}{k}\frac{1}{b+k}\\ &= \frac{(-1)^j}{j}\mathrm{B}(b, j-1),\end{aligned}$$

where the last equality holds since

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

for all $n \in \mathbb{N}_0$, which is, for example, readily verified by induction on $n \in \mathbb{N}_0$. Thus, (5.4) is established.

Thanks to Theorem 3.2 and the inversion formula (5.4), the proof of Corollary 3.3 is now straightforward. Basically, the same argument has been used by, e.g. Panholzer [28, p. 277]. Substituting the expansion for the ordinary moments given in (3.6) shows that

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

since $\sum_{i=0}^{j}\binom{j}{i}(-1)^{j-i} = 0$, $\sum_{i=0}^{j}\binom{j}{i}(-1)^{j-i}(j-i) = 0$, and $\sum_{i=0}^{j}\binom{j}{i}(-1)^{j-i}m_i = (-1)^j/j\mathrm{B}(b, j-1)$ for $j \geq 2$ by (5.4). This completes the proof of Corollary 3.3.

Appendix A

For each $n \in \mathbb{N}$, let $(p_{n,k})_{0\le k\le n}$ 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

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

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

Lemma A.1. ([14, Lemma 6.1].) *Suppose that there exists a sequence $(\psi_n)_{n\in\mathbb{N}}$ such that*

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

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

Then a_n, defined by (A.1), satisfies

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

Lemma A.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 $(q_n)_{n\in\mathbb{N}\setminus\{1\}}$, where $p_{n,m}^{(1)}$ is defined via (3.4). 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 $a_n = O(n^\alpha/\log^p n)$.*

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

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

Set $\psi_n := n/\log n$. Then both conditions (C1) and (C2) hold. Hence,

$$a'_n = O\biggl(\sum_{k=2}^{n} \frac{k^{\alpha-1-\delta}}{\log^p k}\biggr) = O\biggl(\frac{n^{\alpha-\delta}}{\log^p n}\biggr) \quad \text{and} \quad |a_n| = n^\delta a'_n = O\biggl(\frac{n^\alpha}{\log^p n}\biggr).$$

Lemma A.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 $(p_k^{(a)})_{k\in\mathbb{N}}$,*

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

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

Proof. For the beta (a,b)-coalescents, (1.1) reads

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

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

Using the known estimate $|\Gamma(n+c)/\Gamma(n+d)-n^{c-d}| \le M_{c,d}n^{c-d-1}$, $n \ge 2$ and $c, d > -2$, for the gamma function (see Formula (6.1.47) of [1]), we obtain

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

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

Using (3.1) with Λ given by (1.2), we infer (see also Corollary 2 of [13]) that

$$\lambda_n = \frac{\Gamma(a)}{(2-a)\mathrm{B}(a,b)}n^{2-a} + O(n^{1-a}) = \frac{\Gamma(a)}{(2-a)\mathrm{B}(a,b)}n^{2-a}(1+O(n^{-1})) \tag{A.4}$$

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

$$\lambda_n = bn + O(\log n) \tag{A.5}$$

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

$$\begin{aligned}p^{(a)}_{n,n-k} &= \frac{(2-a)\Gamma(a+k-1)}{\Gamma(a)(k+1)!}n^{1-b}((n-k)^{b-1}+O((n-k)^{b-2}))\left(1+O\left(\frac{1}{n}\right)\right) \\ &= p^{(a)}_k\left(\left(1-\frac{k}{n}\right)^{b-1}+O\left(\frac{1}{n}\left(1-\frac{k}{n}\right)^{b-2}\right)\right)\left(1+O\left(\frac{1}{n}\right)\right) \\ &= p^{(a)}_k\left(\left(1-\frac{k}{n}\right)^{b-1}+O\left(\frac{1}{n}\left(1-\frac{k}{n}\right)^{b-2}\right)\right) \\ &= p^{(a)}_k\left(1-\frac{k}{n}\right)^{b-1}+O\left(p^{(a)}_k\frac{1}{n}\left(1-\frac{k}{n}\right)^{b-2}\right).\end{aligned}$$

Analogously, for $a = 1$,

$$\begin{aligned}p^{(1)}_{n,n-k} &= p^{(1)}_k\left(\left(1-\frac{k}{n}\right)^{b-1}+O\left(\frac{1}{n}\left(1-\frac{k}{n}\right)^{b-2}\right)\right)(1+O(n^{-1}\log n)) \\ &= p^{(1)}_k\left(\left(1-\frac{k}{n}\right)^{b-1}+O\left(\frac{1}{n}\left(1-\frac{k}{n}\right)^{b-2}\right)+O\left(\frac{1}{n}\log n\left(1-\frac{k}{n}\right)^{b-1}\right)\right) \\ &= p^{(1)}_k\left(1-\frac{k}{n}\right)^{b-1}+O\left(p^{(1)}_k\frac{1}{n}\left(1-\frac{k}{n}\right)^{b-2}\right)+O\left(p^{(1)}_k\frac{1}{n}\log n\left(1-\frac{k}{n}\right)^{b-1}\right).\end{aligned}$$

Substituting these expansions into the left-hand side of (A.2) gives

$$\begin{aligned}\sum_{k=1}^{n-1}k^q|\mathbb{P}\{I_n=k\}-\mathbb{P}\{\xi=k\}| &\le \sum_{k=1}^{n-1}p^{(a)}_k k^q\left|\left(1-\frac{k}{n}\right)^{b-1}-1\right|+\frac{c_1}{n}\sum_{k=1}^{n-1}p^{(a)}_k k^q\left(1-\frac{k}{n}\right)^{b-2} \\ &=: S_1(a,n)+S_2(a,n) \quad \text{for } 0<a<1,\end{aligned}$$

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

and

$$\sum_{k=1}^{n-1} k^q |\mathbb{P}\{I_n = k\} - \mathbb{P}\{\xi = k\}| \le S_1(1,n) + S_2(1,n) + \frac{c_2 \log n}{n} \sum_{k=1}^{n-1} p_k^{(1)} k^q \left(1 - \frac{k}{n}\right)^{b-1}$$
$$=: S_1(1,n) + S_2(1,n) + S_3(1,n) \quad \text{for } a = 1.$$

Here and hereafter $c_1, c_2, \dots$ 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 $p_k^{(a)} \le c_3 k^{a-3}$ for all $k \in \mathbb{N}$, we infer that

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

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

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

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,

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

in view of the estimate for $S_1(a,n)$. Thus, $S_3(1,n) = O(n^{q-1})$. This completes the proof.

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

Lemma A.4. *The total rates λ_n, $n \in \mathbb{N}$, of the* beta $(1, b)$*-coalescent are given by*

$$\lambda_n = b\sum_{k=1}^{n-1} \frac{k}{b+k-1} = b(n-1) - b(b-1)(\Psi(n+b-1) - \Psi(b)), \qquad n \in \mathbb{N}. \tag{A.6}$$

Moreover, the total rates have the asymptotic expansion

$$\lambda_n = bn - b(b-1)\log n - b + b(b-1)\Psi(b) + O(n^{-1}) \quad \text{as } n \to \infty, \tag{A.7}$$

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

$$\frac{1}{\lambda_n} = \frac{1}{bn}\left(1 + (b-1)\frac{\log n}{n} + \frac{1-(b-1)\Psi(b)}{n} + O\left(\frac{\log^2 n}{n^2}\right)\right) \quad \text{as } n \to \infty. \tag{A.8}$$

Proof. Equation (A.6) is known (see, for example, [19, Equation (19)]). Expansion (A.7) follows from (A.6), since $\Psi(n+b-1) = \log n + O(n^{-1})$ as $n \to \infty$. Equation (A.8) follows from

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

where the last equality holds since $\lambda_n \sim bn$, and the equality before follows by substituting the λ_n term in the numerator for the expression given in (A.7) and multiplying everything out.

The next lemma provides asymptotic expansions as $n \to \infty$ for the sum

$$s_n(p, \alpha) := \sum_{m=2}^{n-1} \frac{m^\alpha}{(n-m)(n-m+1)\log^p m}, \qquad p \in [0, \infty),\ \alpha \in \mathbb{R}. \tag{A.9}$$

Lemma A.5. *Fix $p \in [0, \infty)$. Then, as $n \to \infty$, the sum $s_n(p, \alpha)$ defined in (A.9) satisfies $s_n(p, \alpha) = O(n^\alpha/\log^p n)$ for $\alpha \in (-2, \infty)$,*

$$s_n(p, \alpha) = \frac{n^\alpha}{\log^p n}\left(1 + O\left(\frac{\log n}{n}\right)\right) \quad \text{for } \alpha \in (-1, \infty),$$

and

$$s_n(p, \alpha) = \frac{n^\alpha}{\log^p n}\left(1 - \alpha\frac{\log n}{n} + \frac{\alpha\Psi(\alpha) + p}{n} + O\left(\frac{1}{n\log n}\right)\right) \quad \text{for } \alpha \in (0, \infty).$$

For a proof of Lemma A.5, we refer the reader to [16], which is a preprint version of this article. The following corollary provides an asymptotic expansion which is a key tool in the proof of Theorem 3.2.

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

$$\sum_{m=2}^{n-1} p_{n,m}^{(1)} \frac{m^\alpha}{\log^p m} = \frac{n^\alpha}{\log^p n}\left(1 - \alpha\frac{\log n}{n} + \frac{c_{b,\alpha,p}}{n} + O\left(\frac{1}{n \log n}\right)\right) \quad \textit{as } n \to \infty, \tag{A.10}$$

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

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

$$\sum_{m=2}^{n-1} g_{nm} \frac{m^\alpha}{\log^p m} = \frac{bn^{\alpha+1}}{\log^p n}\left(1 - (\alpha + b - 1)\frac{\log n}{n} + \frac{(\alpha + b - 1)\Psi(\alpha + b - 1) + p}{n} + O\left(\frac{1}{n \log n}\right)\right), \tag{A.11}$$

since (A.10) then follows from $p_{n,m}^{(1)} = g_{nm}/\lambda_n$ by multiplying (A.11) with (A.8). 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

$$\frac{n!}{\Gamma(b + n - 1)} = \frac{1}{n^{b-2}}\left(1 - \binom{b-1}{2}\frac{1}{n} + O\left(\frac{1}{n^2}\right)\right), \tag{A.12}$$

it hence suffices to verify that

$$\sum_{m=2}^{n-1} \frac{1}{(n - m)(n - m + 1)} \frac{\Gamma(b + m - 1)}{(m - 1)!} \frac{m^\alpha}{\log^p m} = \frac{n^{\alpha+b-1}}{\log^p n}\left(1 - (\alpha + b - 1)\frac{\log n}{n} + \frac{\binom{b-1}{2} + (\alpha + b - 1)\Psi(\alpha + b - 1) + p}{n} + O\left(\frac{1}{n \log n}\right)\right), \tag{A.13}$$

since (A.11) then follows by multiplying (A.13) with (A.12). Thus, it remains to verify (A.13). 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 of (A.13) is bounded below and above by

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

where $K_b := \Gamma(b) - 1 - \binom{b-1}{2}$, (A.13) follows by substituting these lower and upper bounds into the left-hand side of (A.13), then applying the last expansion in Lemma A.5 with α replaced

https://doi.org/10.1239/aap/1401369704 Published online by Cambridge University Press

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

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

again by Lemma A.5. This completes the proof.

Acknowledgements

The authors thank an anonymous referee for helpful comments which led to a significant improvement in the presentation of the paper. A. Iksanov and A. Marynych were supported by a grant awarded by the President of Ukraine (project Φ-47/012). The research of A. Marynych was financially supported by a Free Competition Grant of the Netherlands Organization for Scientific Research (NWO).

References

[1] Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. U.S. Government Printing Office, Washington, DC.

[2] Bertoin, J. and Le Gall, J.-F. (2000). The Bolthausen–Sznitman coalescent and the genealogy of continuous-state branching processes. *Prob. Theory Relat. Fields* **117,** 249–266.

[3] Bertoin, J. and Pitman, J. (2000). Two coalescents derived from the ranges of stable subordinators. *Electron. J. Prob.* **5,** 17pp.

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

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

[6] Bolthausen, E. and Sznitman, A.-S. (1998). On Ruelle's probability cascades and an abstract cavity method. *Commun. 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. Prob.* **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 Structures Algorithms* **34,** 319–336.

[10] Feller, W. (1949). Fluctuation theory of recurrent events. *Trans. Amer. Math. Soc.* **67,** 98–119.

[11] 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.

[12] Givens, C. R. and Shortt, R. M. (1984). A class of Wasserstein metrics for probability distributions. *Michigan Math. J.* **31,** 231–240.

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

[14] Gnedin, A., Iksanov, A. and Marynych, A. (2011). On Λ-coalescents with dust component. *J. Appl. Prob.* **48,** 1133–1151.

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

[16] Gnedin, A., Iksanov, A., Marynych, A. and Moehle, M. (2012). On asymptotics of the beta-coalescents. Preprint. Available at http://uk.arxiv.org/abs/1203.3110.

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

[18] Haas, B. and Miermont, G. (2011). Self-similar scaling limits of non-increasing Markov chains. *Bernoulli* **17,** 1217–1247.

[19] Huillet, T. and Möhle, M. (2013). On the extended Moran model and its relation to coalescents with multiple collisions. *Theoret. Pop. Biol.* **87,** 5–14.

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

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

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

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

[24] Kersting, G. (2012). The asymptotic distribution of the length of beta-coalescent trees. *Ann. Appl. Prob.* **22,** 2086–2107.

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

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

[27] 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.

[28] Panholzer, A. (2004). Destruction of recursive trees. In *Mathematics and Computer Science*, Vol. III, Birkhäuser, Basel, pp. 267–280.

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

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

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

[32] Tavaré, S. (2004). Ancestral inference in population genetics. In *Lectures on Probability Theory and Statistics* (Lecture Notes Math. **1837**), Springer, Berlin, pp. 1–188.