Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 117 (2007) 1404–1421
stochastic processes and their applications
www.elsevier.com/locate/spa

Asymptotic results concerning the total branch length of the Bolthausen–Sznitman coalescent

Michael Drmota[a], Alex Iksanov[b], Martin Moehle[c,*], Uwe Roesler[d]

[a] *Institute of Discrete Mathematics and Geometry, Technical University of Vienna, Wiedner Hauptstr. 8-10, 1040 Vienna, Austria*
[b] *National Taras Shevchenko University of Kiev, 60, Volodymyrska, 01033 Kiev, Ukraine*
[c] *Mathematical Institute, University of Tuebingen, Auf der Morgenstelle 10, 72076 Tuebingen, Germany*
[d] *Mathematical Seminar, University of Kiel, Ludewig-Meyn-Str. 4, 24098 Kiel, Germany*

Received 5 January 2007; received in revised form 23 January 2007; accepted 27 January 2007
Available online 3 March 2007

Abstract

We study the total branch length L_n of the Bolthausen–Sznitman coalescent as the sample size n tends to infinity. Asymptotic expansions for the moments of L_n are presented. It is shown that $L_n/\mathrm{E}(L_n)$ converges to 1 in probability and that L_n, properly normalized, converges weakly to a stable random variable as n tends to infinity. The results are applied to derive a corresponding limiting law for the total number of mutations for the Bolthausen–Sznitman coalescent with mutation rate $r > 0$. Moreover, the results show that, for the Bolthausen–Sznitman coalescent, the total branch length L_n is closely related to X_n, the number of collision events that take place until there is just a single block. The proofs are mainly based on an analysis of random recursive equations using associated generating functions.

MSC: primary 60J25; 60F05; secondary 60C05; 05C05; 92D10; 92D15

Keywords: Asymptotic expansion; Bolthausen–Sznitman coalescent; Generating functions; Random recursive trees; Stable limit

1. Introduction and main results

Starting from the seminal work of Kingman [13,14], coalescent processes have been proven to be a powerful tool in ancestral population genetics. These processes are useful for studying the

* Corresponding author. Tel.: +49 7071 2976767; fax: +49 7071 294322.
E-mail addresses: michael.drmota@tuwien.ac.at (M. Drmota), iksan@unicyb.kiev.ua (A. Iksanov), martin.moehle@uni-tuebingen.de (M. Moehle), roesler@math.uni-kiel.de (U. Roesler).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2007.01.011

ancestral history of a sample of n particles, individuals, genes or DNA sequences chosen from a large population. In this paper we are interested in the total branch length L_n of the subclass of coalescent processes with multiple collisions, independently introduced by Pitman [21] and Sagitov [22]. These coalescent processes are also called Λ-coalescent processes, because they can be characterized via a finite measure Λ on the unit interval [0, 1]. For certain subclasses of measures Λ, the asymptotics of L_n are well known. Consider for example the Kingman coalescent, where Λ is the Dirac measure in 0. For more details about the Kingman coalescent we refer the reader to Kingman [13,14]. In this case the random variable $L_n/2 - \log n$ is asymptotically standard Gumbel distributed. An elementary proof of this result and some remarks about its history are provided in the Appendix (Lemma 7.1 and the Remark thereafter). Another class are the measures Λ satisfying $\int_{[0,1]} x^{-1}\Lambda(\mathrm{d}x) < \infty$. In this case, as n tends to infinity, L_n/n converges in distribution to a limiting variable L whose distribution coincides with that of $\int_0^\infty \mathrm{e}^{-X_t}\,\mathrm{d}t$, where $(X_t)_{t\geq 0}$ is a certain subordinator. This convergence is a slight modification of an analogous result given in [17, Proposition 5.2] for the number of mutations (segregating sites) for a Λ-coalescent with mutation.

Except for the Kingman coalescent, there is only little known about the total branch length when $\int_{[0,1]} x^{-1}\Lambda(\mathrm{d}x) = \infty$. For example, for the case when Λ is the beta $(2-\alpha, \alpha)$ distribution, it was shown in [1] that $L_n/n^{2-\alpha}$ converges in probability to a constant, whose value can be given explicitly in terms of gamma functions.

We focus in this paper on the total branch length L_n of the Bolthausen–Sznitman coalescent [5], which is the Λ-coalescent with Λ being the uniform measure on [0, 1]. The Bolthausen–Sznitman coalescent is an important process that has been studied extensively. For example, the process has connections to stable subordinators [3], the genealogy of continuous-state branching processes [2], and Derrida's generalized random energy model [6].

Section 2 briefly recalls the definition and some basic properties of the Λ-coalescent. In Section 3 we study the total branch length L_n of the Λ-coalescent. The branch length L_n satisfies a specific recursion equation (see (2)), which leads to recursions for many functionals of L_n. For example, in (8) a recursion for the jth moments $\mu_n^{(j)} := \mathrm{E}(L_n^j)$ of L_n, $j, n \in \mathbb{N}$, is provided.

From Section 4 on we focus on the Bolthausen–Sznitman coalescent. Sections 4 and 5 contain the main results of the paper. In Section 4, we modify Panholzer's approach [19], based on generating functions, to derive asymptotic expansions for the moments of L_n (see Corollary 4.3) and for the centered moments of L_n (see (28)). In particular, $\mathrm{E}(L_n) \sim n/\log n$, $\mathrm{E}(L_n^2) \sim n^2/(\log^2 n)$ and $\mathrm{Var}(L_n) \sim n^2/(2\log^3 n)$. From these results it follows immediately that $L_n/\mathrm{E}(L_n)$ converges to 1 in probability as n tends to infinity (see Corollary 4.4).

In Section 5 a weak limiting result for L_n is provided. Theorem 5.2 states that L_n, properly normalized, converges in distribution to a stable random variable with characteristic function $t \mapsto \exp(-\frac{1}{2}\pi|t| + \mathrm{i}t\log|t|)$ (see (31)). We finally apply these results in Section 6 to the Bolthausen–Sznitman coalescent with mutation rate $r > 0$ and derive corresponding convergence results for the total number S_n of mutations.

2. The Λ-coalescent process

Let $\mathcal{E}$ denote the set of all equivalence relations on $\mathbb{N} := \{1, 2, \ldots\}$. For $n \in \mathbb{N}$ let $\varrho_n : \mathcal{E} \to \mathcal{E}_n$ denote the natural restriction to the set $\mathcal{E}_n$ of all equivalence relations on $\{1, \ldots, n\}$. For a finite measure Λ on the unit interval [0, 1] let $R := (R_t)_{t\geq 0}$ be a Λ-coalescent process as introduced by Pitman [21] and Sagitov [22]. Note that R is a Markovian process with state space $\mathcal{E}$. The probabilistic structure of R depends on the measure Λ as follows. For each $n \in \mathbb{N}$ the restricted process $(\varrho_n R_t)_{t\geq 0}$ is Markovian with state space $\mathcal{E}_n$ and rates

$$q_{\xi\eta} := \begin{cases} \int_{[0,1]} (1-(1-x)^b - bx(1-x)^{b-1})x^{-2}\Lambda(\mathrm{d}x) & \text{if } \xi = \eta, \\ \int_{[0,1]} x^{b-a-1}(1-x)^{a-1}\Lambda(\mathrm{d}x) & \text{if } \xi \prec \eta, \\ 0 & \text{otherwise,} \end{cases}$$

where $a := |\eta|$ and $b := |\xi|$ are the number of classes (blocks) of $\xi \in \mathcal{E}_n$ and $\eta \in \mathcal{E}_n$ respectively, and $\xi \prec \eta$ means (by definition) that exactly $b - a + 1$ equivalence classes of ξ merge together to form one class of η, while all the other $a - 1$ classes of ξ remain unchanged. For $\Lambda = \delta_0$, the Dirac measure at 0, the process R is the Kingman coalescent [13]. For Λ being the uniform measure on [0, 1], we obtain the Bolthausen–Sznitman coalescent [5]. It is well known that the process $(|\varrho_n R_t|)_{t\geq 0}$ is a Markovian death process with rates

$$g_{ba} = \binom{b}{a-1} \int_{[0,1]} x^{b-a-1}(1-x)^{a-1}\Lambda(\mathrm{d}x), \quad 1 \leq a < b \leq n,$$

and total rates

$$g_b = \sum_{a=1}^{b-1} g_{ba} = \int_{[0,1]} \frac{1-(1-x)^b - bx(1-x)^{b-1}}{x^2}\Lambda(\mathrm{d}x), \quad 1 \leq b \leq n.$$

Let $(\mathcal{J}_r^{(n)})_{r\in\mathbb{N}_0}$ denote the jump chain of the process $(|\varrho_n R_t|)_{t\geq 0}$. Note that $\mathcal{J}_0^{(n)} \equiv n$. The first jump will be to the state k, $1 \leq k < n$, with probability

$$p_{nk} := P(I_n = k) = \frac{g_{nk}}{g_n}, \quad n, k \in \mathbb{N}, k < n, \tag{1}$$

where $I_n := \mathcal{J}_1^{(n)}$. We think of the process $(\varrho_n R_t)_{t\geq 0}$ as a random tree with n leaves having labels from 1 to n. With this interpretation, $|\varrho_n R_t|$ is the number of branches of this tree at time $t \geq 0$.

3. Total branch length

We are interested in the total branch length L_n, i.e. the sum of the length of all branches of the tree $(\varrho_n R_t)_{t\geq 0}$. It is well known [17, Eq. (10)] that L_n satisfies the recursion $L_1 = 0$ and

$$L_n = T_n + L_{I_n} = T_n + \sum_{k=1}^{n-1} 1_{\{I_n=k\}} L_k, \quad n \geq 2, \tag{2}$$

with $T_n := n\tau_n$, where τ_n is the amount of time for which the tree $(\varrho_n R_t)_{t\geq 0}$ has n branches. Note that (2) holds almost surely and not only in distribution. From the Markov property of $(\varrho_n R_t)_{t\geq 0}$ it follows that τ_n is exponentially distributed with parameter g_n. Thus, T_n is exponentially distributed with parameter $\alpha_n := g_n/n$. For $m, n \in \mathbb{N}$ with $m < n$ let $\varrho_{nm} : \mathcal{E}_n \to \mathcal{E}_m$ denote the natural restriction from $\mathcal{E}_n$ to $\mathcal{E}_m$. As $\varrho_m R_t = \varrho_{nm}\varrho_n R_t$, the tree $(\varrho_m R_t)_{t\geq 0}$ is obtained from the tree $(\varrho_n R_t)_{t\geq 0}$ by removing all branches of the tree $(\varrho_n R_t)_{t\geq 0}$ with labels $m + 1, \ldots, n$. Thus

$$L_n = L_m + R_{nm}, \quad m, n \in \mathbb{N}, m < n \tag{3}$$

almost surely, where R_{nm} denotes the sum of the lengths of all removed branches. In particular, $P(L_m \le L_n) = 1$ for $m, n \in \mathbb{N}$ with $m < n$. There is another interpretation of L_n. It is a total cost of a one-sided destruction of size n recursive trees when the toll variable T_n is exponentially distributed with parameter α_n for $n \ge 2$ and $T_1 \equiv 0$. Janson [11,12], Panholzer [19,20], and Fill, Kapur and Panholzer [8] consider similar models with non-random toll functions T_n.

For $n \in \mathbb{N}$ and $j \in \mathbb{N}_0$ let $\mu_n^{(j)} := \mathrm{E}(L_n^j)$ denote the jth moment of L_n. From (3) it follows that, for each fixed j, the sequence $(\mu_n^{(j)})_{n\in\mathbb{N}}$ is non-decreasing. Obviously, $\mu_1^{(j)} = 0$ and, by (2),

$$\mu_n^{(j)} = \sum_{i=0}^{j} \binom{j}{i} \mathrm{E}(T_n^i)\mathrm{E}(L_{I_n}^{j-i}) = \sum_{i=0}^{j} \binom{j}{i} \mathrm{E}(T_n^i) \sum_{k=1}^{n-1} p_{nk}\mu_k^{(j-i)}$$
$$= \sum_{k=1}^{n-1} p_{nk}\mu_k^{(j)} + r_n^{(j)}, \quad n \ge 2, j \in \mathbb{N}_0 \tag{4}$$

with rest term

$$r_n^{(j)} := \sum_{i=1}^{j} \binom{j}{i} \mathrm{E}(T_n^i) \sum_{k=1}^{n-1} p_{nk}\mu_k^{(j-i)}.$$

For $j \in \mathbb{N}_0$ define the generating functions

$$\mu_j(s) := \sum_{n=2}^{\infty} \mu_n^{(j)} s^n \quad \text{and} \quad r_j(s) := \sum_{n=2}^{\infty} \alpha_n r_n^{(j)} s^n, \quad 0 \le s < 1. \tag{5}$$

In the situation considered in this paper, the toll variables T_n are exponentially distributed. In this case the generating functions μ_j and r_j are related as follows.

Lemma 3.1. *Assume that $T_1 \equiv 0$ and that, for $n \ge 2$, T_n is exponentially distributed with parameter $\alpha_n > 0$. Then, for $n \ge 2$ and $j \in \mathbb{N}$,*

$$r_n^{(j)} = j\alpha_n^{-1}\mu_n^{(j-1)} \tag{6}$$

and, hence,

$$r_j(s) = j\mu_{j-1}(s), \quad j \in \mathbb{N}, 0 \le s < 1. \tag{7}$$

In particular, $r_1(s) = \mu_0(s) = \sum_{n=2}^{\infty} s^n = s^2/(1-s)$, $0 \le s < 1$.

Proof. Induction on j. For $j = 1$, Eq. (6) is obvious, as $r_n^{(1)} = \mathrm{E}(T_n) = \alpha_n^{-1}$. The step from $1, \dots, j-1$ to j works as follows. For $i \in \{2, \dots, j\}$ it follows by induction and from $\mathrm{E}(T_n^i) = i!\alpha_n^{-i}$ that

$$\binom{j}{i-1} \mathrm{E}(T_n^{i-1}) r_n^{(j-i+1)} = \binom{j}{i-1} \mathrm{E}(T_n^{i-1})(j-i+1)\alpha_n^{-1}\mu_n^{(j-i)}$$
$$= \binom{j}{i-1} \frac{j-i+1}{i} \mathrm{E}(T_n^i)\mu_n^{(j-i)}$$
$$= \binom{j}{i} \mathrm{E}(T_n^i)\mu_n^{(j-i)}.$$

Thus,

$$
\begin{aligned}
r_n^{(j)} &= \sum_{i=1}^{j-1} \binom{j}{i} \mathrm{E}(T_n^i) \sum_{k=1}^{n-1} p_{nk} \mu_k^{(j-i)} + \mathrm{E}(T_n^j) \\
&= \sum_{i=1}^{j-1} \binom{j}{i} \mathrm{E}(T_n^i)(\mu_n^{(j-i)} - r_n^{(j-i)}) + \mathrm{E}(T_n^j) \\
&= \sum_{i=1}^{j} \binom{j}{i} \mathrm{E}(T_n^i)\mu_n^{(j-i)} - \sum_{i=1}^{j-1} \binom{j}{i} \mathrm{E}(T_n^i) r_n^{(j-i)} \\
&= \sum_{i=1}^{j} \binom{j}{i} \mathrm{E}(T_n^i)\mu_n^{(j-i)} - \sum_{i=2}^{j} \binom{j}{i-1} \mathrm{E}(T_n^{i-1}) r_n^{(j-i+1)} \\
&= \binom{j}{1} \mathrm{E}(T_n)\mu_n^{(j-1)} = j\alpha_n^{-1}\mu_n^{(j-1)}.
\end{aligned}
$$

From the definition (5) of $r_j(s)$ the formula (7) follows immediately. □

Remark. The recursion (4) thus becomes $\mu_1^{(j)} = 0$, $j \in \mathbb{N}$, and

$$\mu_n^{(j)} = j\alpha_n^{-1}\mu_n^{(j-1)} + \sum_{k=2}^{n-1} p_{nk}\mu_k^{(j)}, \quad j \in \mathbb{N}, n \geq 2. \tag{8}$$

With this recursion it is possible to compute $\mu_n^{(j)}$ numerically. First, compute $\mu_1^{(1)}, \dots, \mu_n^{(1)}$ via the recursion $\mu_1^{(1)} = 0$ and

$$\mu_n^{(1)} = \alpha_n^{-1} + \sum_{k=2}^{n-1} p_{nk}\mu_k^{(1)}, \quad n \geq 2.$$

After these first moments are computed, use $\mu_1^{(2)} = 0$ and

$$\mu_n^{(2)} = 2\alpha_n^{-1}\mu_n^{(1)} + \sum_{k=2}^{n-1} p_{nk}\mu_k^{(2)}, \quad n \geq 2$$

to compute the second moments $\mu_1^{(2)}, \dots, \mu_n^{(2)}$. Repeat this procedure (using (8)) until $\mu_n^{(j)}$ is computed.

4. Total branch length of the Bolthausen–Sznitman coalescent

In the following we focus on the Bolthausen–Sznitman coalescent [5], i.e. the Λ-coalescent, where Λ is the Lebesgue measure on [0, 1]. A straightforward computation shows that $g_{nk} = n/((n-k)(n-k+1))$, $k, n \in \mathbb{N}$ with $k < n$, and that $g_n = n-1$, $n \in \mathbb{N}$. Thus, the jump chain $(\mathcal{J}_r^{(n)})_{r\in\mathbb{N}_0}$ has transition probabilities

$$p_{nk} = P(I_n = k) = \frac{g_{nk}}{g_n} = \frac{n}{(n-1)(n-k)(n-k+1)}, \quad 1 \leq k < n. \tag{9}$$

These transition probabilities coincide with those obtained by Meir and Moon [15] for the subtree size of a random recursive tree of size n, when an edge is removed at random. For $n \in \mathbb{N}$

let $h_n := \sum_{i=1}^n 1/i$ denote the nth harmonic number. Note that, for $n \geq 2$, $\mathrm{E}(n - I_n) = n(h_n - 1)/(n-1) \sim \log n$ and $\mathrm{E}((n-I_n)^2) = n(n-h_n)/(n-1) \sim n$. As n tends to infinity, the random variable $n - I_n$ converges in distribution to a limiting variable I with distribution $P(I = k) = 1/(k(k+1))$, $k \in \mathbb{N}$.

In this section we study, for arbitrary but fixed $j \in \mathbb{N}$, the asymptotics of the moments $\mu_n^{(j)} = \mathrm{E}(L_n^j)$ as n tends to infinity. Of course (see Lemmas 7.2 and 7.3 in the Appendix) Karamata's Tauberian theorem yields $\mu_n^{(1)} \sim n/\log n$ and $\mu_n^{(2)} \sim n^2/\log^2 n$, but we will not use Tauberian theorems in this section. Instead, we adapt Panholzer's [19] approach to derive (see Corollary 4.3 and the examples thereafter) asymptotic expansions for $\mu_n^{(j)}$. We start with providing a recursion for the generating functions μ_j defined in (5).

Lemma 4.1 (*Recursion for the Generating Functions* μ_j). *For* $j \in \mathbb{N}$ *and* $0 \leq s < 1$

$$\mu_j(s) = \sum_{n=2}^{\infty} \mu_n^{(j)} s^n = \frac{js}{s-1} \int_0^s \frac{\mu'_{j-1}(t)}{\log(1-t)} \mathrm{d}t. \tag{10}$$

In particular,

$$\mu_1(s) = \frac{s}{s-1} \int_0^s \frac{t(2-t)}{(1-t)^2 \log(1-t)} \mathrm{d}t, \quad 0 \leq s < 1. \tag{11}$$

Proof. Fix $j \in \mathbb{N}$. For $0 \leq s < 1$ define the auxiliary function

$$g(s) := \sum_{k=1}^{\infty} \frac{s^k}{k(k+1)} = 1 + \frac{\log(1-s)}{s} - \log(1-s).$$

It is convenient to rewrite the recursions (4) for $(\mu_n^{(j)})_{n \in \mathbb{N}}$ in the form

$$\frac{n-1}{n} \mu_n^{(j)} = \frac{n-1}{n} r_n^{(j)} + \sum_{k=1}^{n-1} \frac{\mu_{n-k}^{(j)}}{k(k+1)}, \quad n \geq 2. \tag{12}$$

Multiplication by s^n and summation over $n = 2, 3, \ldots$ leads to

$$\begin{aligned}
\mu_j(s) - \int_0^s \frac{\mu_j(t)}{t} \mathrm{d}t &= \sum_{n=2}^{\infty} \frac{n-1}{n} \mu_n^{(j)} s^n \\
&= \sum_{n=2}^{\infty} \frac{n-1}{n} r_n^{(j)} s^n + \sum_{n=2}^{\infty} s^n \sum_{k=1}^{n-1} \frac{\mu_{n-k}^{(j)}}{k(k+1)} \\
&= r_j(s) + \sum_{k=1}^{\infty} \frac{s^k}{k(k+1)} \sum_{n=k+1}^{\infty} \mu_{n-k}^{(j)} s^{n-k} \\
&= r_j(s) + g(s) \mu_j(s).
\end{aligned}$$

Taking the derivative with respect to s yields

$$\mu'_j(s) - \frac{\mu_j(s)}{s} = r'_j(s) + g'(s)\mu_j(s) + g(s)\mu'_j(s),$$

or, equivalently, $\mu_j'(s)(1-g(s)) = \mu_j(s)(g'(s)+1/s) + r_j'(s)$. Now plug in $g(s)$ and $g'(s) = -1/s - (\log(1-s))/s^2$ to conclude that

$$\mu_j'(s) = \frac{\mu_j(s)}{s(1-s)} - \frac{s r_j'(s)}{(1-s)\log(1-s)}. \tag{13}$$

Solutions of the homogeneous differential equation $f'(s) = f(s)/(s(1-s))$ are of the form $f(s) = cs/(1-s)$, $c \in \mathbb{R}$. Returning to the inhomogeneous differential equation (13) with initial value $\mu_j(0) = 0$ we see that $\mu_j(s) = c_j(s)s/(1-s)$ with

$$c_j(s) := -\int_0^s \frac{r_j'(t)}{\log(1-t)}\mathrm{d}t, \tag{14}$$

and (10) follows from (7). We have $\mu_0(s) = \sum_{n=2}^{\infty} s^n = s^2/(1-s)$, i.e. $\mu_0'(s) = s(2-s)/(1-s)^2$, and (11) follows from (10). □

For $x > 0$ let $\Psi(x) = \Gamma'(x)/\Gamma(x)$, where Γ denotes Euler's gamma function. Write $[s^n]f(s) = f_n$, if $f(s) = \sum_{n=n_0}^{\infty} s^n f_n$. In order to derive asymptotic expansions for the jth moment $\mu_n^{(j)} = \mathrm{E}(L_n^j)$, it is helpful to analyze the asymptotics of the coefficients $[s^n]c_j(s)$ of the function c_j defined in (14).

Proposition 4.2 (*Asymptotics of c_j*). *Fix $j \in \mathbb{N}$. As $n \to \infty$,*

$$[s^n]c_j(s) = j\frac{n^{j-1}}{\log^j n} + j\kappa_j \frac{n^{j-1}}{\log^{j+1} n} + O\left(\frac{n^{j-1}}{\log^{j+2} n}\right), \tag{15}$$

where the sequence $(\kappa_j)_{j\in\mathbb{N}}$ is recursively defined via $\kappa_1 := \Psi(2) = 1-\gamma$ ($\gamma \approx 0.577216$ denotes Euler's constant) and

$$\kappa_{j+1} := \kappa_j + (j+1)\Psi(j+2) - \frac{j}{j+1}(j\Psi(j+1) + \Psi(j)), \quad j \in \mathbb{N}.$$

Remark. Using the identities $\Psi(x+1) = \Psi(x) + 1/x$, $x > 0$, and $\Psi(j+1) = h_j - \gamma$, $j \in \mathbb{N}$, where h_j denotes the jth harmonic number, an induction on j yields

$$\kappa_j = (j+1)h_j - j\gamma - 1, \quad j \in \mathbb{N}. \tag{16}$$

Proof. The proof goes a similar path as the proof of Theorem 2.1 of Panholzer [19]. We will use, for $\alpha, p > 0$, the asymptotic growth of the coefficients (Panholzer [19, Eq. (19)])

$$[s^n]\frac{1}{(1-s)^\alpha(-\log(1-s))^p} = \frac{n^{\alpha-1}}{\Gamma(\alpha)\log^p n}\left(1 + \frac{p\Psi(\alpha)}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right) \tag{17}$$

and the effect on the growth of the coefficients [19, Eq. (20)] when integrating and differentiating the generating function $F(s) = \sum_{n=2}^{\infty} s^n n^\alpha/(\log^p n)$, $\alpha, p > 0$,

$$[s^n]\int_0^s F(t)\mathrm{d}t = \frac{n^{\alpha-1}}{\log^p n}\left(1 + O\left(\frac{1}{n}\right)\right), \qquad [s^n]F'(s) = \frac{n^{\alpha+1}}{\log^p n}\left(1 + O\left(\frac{1}{n}\right)\right). \tag{18}$$

We additionally use Panholzer's [19, Lemma 4.1, Eq. (16)] summation expansion: For $\alpha, \beta > -1$ and $p, q \geq 0$

$$\sum_{k=2}^{n-2} \frac{k^{\alpha}(n-k)^{\beta}}{\log^{p} k \log^{q}(n-k)} = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)} \frac{n^{\alpha+\beta+1}}{\log^{p+q} n} \times \left(1 + \frac{(p+q)\,\Psi(\alpha+\beta+2) - p\,\Psi(\alpha+1) - q\,\Psi(\beta+1)}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right). \tag{19}$$

We now verify (15) by induction on j. We have (see Proof of Lemma 4.1) $c_1(s) = \int_0^s t(2-t)/((1-t)^2(-\log(1-t)))\mathrm{d}t$. By (17),

$$\begin{aligned}
[s^n]c_1'(s) &= [s^n]\left(\frac{2s - s^2}{(1-s)^2(-\log(1-s))}\right) \\
&= 2[s^{n-1}]\frac{1}{(1-s)^2(-\log(1-s))} - [s^{n-2}]\frac{1}{(1-s)^2(-\log(1-s))} \\
&= \frac{n}{\log n} + \Psi(2)\frac{n}{\log^2 n} + O\left(\frac{n}{\log^3 n}\right)
\end{aligned}$$

and (18) yields $[s^n]c_1(s) = \frac{1}{\log n} + \frac{\Psi(2)}{\log^2 n} + O\left(\frac{1}{\log^3 n}\right)$.

Thus, (15) holds for $j = 1$. Assume now that (15) holds for some $j \in \mathbb{N}$. Then, by (18),

$$[s^n]c_j'(s) = j\frac{n^j}{\log^j n} + j\kappa_j \frac{n^j}{\log^{j+1} n} + O\left(\frac{n^j}{\log^{j+2} n}\right). \tag{20}$$

From $\mu_j(s) = c_j(s)s/(1-s)$, i.e. $\mu_j'(s) = c_j'(s)s/(1-s) + c_j(s)/(1-s)^2$ it follows that

$$-\frac{\mu_j'(s)}{\log(1-s)} = \frac{sc_j'(s)}{(1-s)(-\log(1-s))} + \frac{c_j(s)}{(1-s)^2(-\log(1-s))}.$$

We have, by (17),

$$[s^n]\frac{1}{(1-s)(-\log(1-s))} = \frac{1}{\log n} + \frac{\Psi(1)}{\log^2 n} + O\left(\frac{1}{\log^3 n}\right).$$

From (20) and (19) it follows that

$$\begin{aligned}
[s^n]\frac{sc_j'(s)}{(1-s)(-\log(1-s))} &= j\sum_{k=2}^{n-2}\frac{k^j}{\log^j k \log(n-k)} + j\sum_{k=2}^{n-2}\frac{k^j\,\Psi(1)}{\log^j k \log^2(n-k)} \\
&\quad + j\kappa_j \sum_{k=2}^{n-2}\frac{k^j}{\log^{j+1} k \log(n-k)} + O\left(\frac{n^{j+1}}{\log^{j+3} n}\right) \\
&= \frac{j}{j+1}\frac{n^{j+1}}{\log^{j+1} n}\left(1 + \frac{(j+1)\,\Psi(j+2) - j\,\Psi(j+1) - \Psi(1)}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right) \\
&\quad + \frac{j}{j+1}(\Psi(1) + \kappa_j)\frac{n^{j+1}}{\log^{j+2} n}
\end{aligned}$$

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

and (18) yields

$$[s^n]\int_0^s \frac{tc_j'(t)}{(1-t)(-\log(1-t))}\mathrm{d}t$$
$$= \frac{j}{j+1}\frac{n^{j}}{\log^{j+1} n}\left(1+\frac{(j+1)\,\Psi(j+2)-j\,\Psi(j+1)+\kappa_j}{\log n}+O\left(\frac{1}{\log^2 n}\right)\right). \tag{21}$$

By (17),

$$[s^n]\frac{1}{(1-s)^2(-\log(1-s))} = \frac{n}{\log n}+\Psi(2)\frac{n}{\log^2 n}+O\left(\frac{n}{\log^3 n}\right).$$

Hence, by (19) and (15) (for j)

$$[s^n]\frac{c_j(s)}{(1-s)^2(-\log(1-s))} = j\sum_{k=2}^{n-2}\frac{k^{j-1}(n-k)}{\log^j k\log(n-k)}$$
$$+\,j\sum_{k=2}^{n-2}\frac{k^{j-1}\,\Psi(2)(n-k)}{\log^j k\log^2(n-k)}+j\kappa_j\sum_{k=2}^{n-2}\frac{k^{j-1}(n-k)}{\log^{j+1} k\log(n-k)}+O\left(\frac{n^{j+1}}{\log^{j+3} n}\right)$$
$$= \frac{1}{j+1}\frac{n^{j+1}}{\log^{j+1} n}\left(1+\frac{(j+1)\,\Psi(j+2)-j\,\Psi(j)-\,\Psi(2)}{\log n}+O\left(\frac{1}{\log^2 n}\right)\right)$$
$$+\frac{1}{j+1}(\,\Psi(2)+\kappa_j)\frac{n^{j+1}}{\log^{j+2} n}$$
$$= \frac{1}{j+1}\frac{n^{j+1}}{\log^{j+1} n}\left(1+\frac{(j+1)\,\Psi(j+2)-j\,\Psi(j)+\kappa_j}{\log n}+O\left(\frac{1}{\log^2 n}\right)\right),$$

and (18) yields

$$[s^n]\int_0^s \frac{c(t)}{(1-t)^2(-\log(1-t))}\mathrm{d}t$$
$$= \frac{1}{j+1}\frac{n^{j}}{\log^{j+1} n}\left(1+\frac{(j+1)\,\Psi(j+2)-j\,\Psi(j)+\kappa_j}{\log n}+O\left(\frac{1}{\log^2 n}\right)\right). \tag{22}$$

Summation of (21) and (22) yields

$$[s^n]\int_0^s \frac{\mu_j'(t)}{-\log(1-t)}\mathrm{d}t$$
$$= [s^n]\int_0^s \frac{tc_j'(t)}{(1-t)(-\log(1-t))}\mathrm{d}t+[s^n]\int_0^s \frac{c_j(t)}{(1-t)^2(-\log(1-t))}\mathrm{d}t$$
$$= \frac{n^j}{\log^{j+1} n}\left(1+\frac{(j+1)\,\Psi(j+2)-\frac{j^2}{j+1}\,\Psi(j+1)-\frac{j}{j+1}\,\Psi(j)+\kappa_j}{\log n}+O\left(\frac{1}{\log^2 n}\right)\right)$$

and multiplication by $j+1$ leads to

$$[s^n]c_{j+1}(s) = [s^n]\int_0^s \frac{r'_{j+1}(t)}{-\log(1-t)}\mathrm{d}t = (j+1)[s^n]\int_0^s \frac{\mu'_j(t)}{-\log(1-t)}\mathrm{d}t$$
$$= \frac{(j+1)n^j}{\log^{j+1} n}\left(1 + \frac{\kappa_j + (j+1)\Psi(j+2) - \frac{j}{j+1}(j\Psi(j+1) + \Psi(j))}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right)$$
$$= (j+1)\frac{n^j}{\log^{j+1} n} + (j+1)\kappa_{j+1}\frac{n^j}{\log^{j+2} n} + O\left(\frac{n^j}{\log^{j+3} n}\right).$$

Thus, (15) is valid for $j+1$ and the induction is finished. □

Corollary 4.3 (*Asymptotics of the Moments of* L_n). *Fix* $j \in \mathbb{N}$. *For* $n \to \infty$, *the* j*th moment of* L_n *has the asymptotic expansion*

$$\mathrm{E}(L_n^j) = \frac{n^j}{\log^j n}\left(1 + \frac{m_j}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right), \tag{23}$$

where $m_j := \kappa_j + 1 = (j+1)h_j - j\gamma$.

Proof. We have $\mathrm{E}(L_n^j) = \mu_n^{(j)} = [s^n]\mu_j(s) = [s^n](c_j(s)s/(1-s))$. From Proposition 4.2 and (19) it follows that

$$[s^n]\left(c_j(s)\frac{s}{1-s}\right) = \sum_{k=0}^{n-1}[s^k]c_j(s)$$
$$= j\sum_{k=2}^{n-2}\frac{k^{j-1}}{\log^j k} + j\kappa_j\sum_{k=2}^{n-2}\frac{k^{j-1}}{\log^{j+1} k} + O\left(\frac{n^j}{\log^{j+2} n}\right)$$
$$= \frac{n^j}{\log^j n}\left(1 + \frac{j\Psi(j+1) - j\Psi(j)}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right) + \kappa_j\frac{n^j}{\log^{j+1} n}$$
$$= \frac{n^j}{\log^j n}\left(1 + \frac{j\Psi(j+1) - j\Psi(j) + \kappa_j}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right).$$

The corollary follows from $\Psi(j+1) - \Psi(j) = 1/j$ and from (16). □

Corollary 4.4 (*Weak Law of Large Numbers for* L_n). *As* n *tends to infinity,* $n^{-1}(\log n)L_n$ *converges in probability to* 1. *Moreover,* $L_n \to \infty$ *almost surely as* $n \to \infty$.

Proof. Fix $\varepsilon > 0$. Define $\mu_n := \mathrm{E}(L_n)$ for convenience. Tschebyscheff's inequality yields

$$P\left(\left|\frac{L_n}{\mu_n} - 1\right| \geq \varepsilon\right) = P(|L_n - \mu_n| \geq \varepsilon\mu_n) \leq \frac{\mathrm{Var}(L_n)}{\varepsilon^2\mu_n^2} = \frac{1}{\varepsilon^2}\left(\frac{\mathrm{E}(L_n^2)}{\mu_n^2} - 1\right).$$

The convergence $L_n/\mu_n \to 1$ in probability follows from $\mu_n \sim n/\log n$ and $\mathrm{E}(L_n^2) \sim n^2/\log^2 n$. There exists a subsequence $(n_k)_{k\in\mathbb{N}}$ with $L_{n_k}/\mu_{n_k} \to 1$ almost surely. In particular, $L_{n_k} \to \infty$ almost surely. Thus, $L_n \to \infty$ almost surely as $P(L_n \leq L_{n+1}) = 1$ for $n \in \mathbb{N}$. □

Table 1
First moment, second moment, and variance of L_n

n	$\mathrm{E}(L_n)$	$\mathrm{E}(L_n^2)$	$\mathrm{Var}(L_n)$
1	0	0	0
2	2	8	4
3	3	15	6
4	$\frac{34}{9} = 3.777778$	$\frac{590}{27} \approx 21.851852$	$\frac{614}{81} \approx 7.580247$
5	$\frac{40}{9} = 4.444444$	$\frac{6205}{216} \approx 28.726852$	$\frac{5815}{648} \approx 8.973765$
6	$\frac{2269}{450} = 5.042222$	$\frac{963\,571}{27\,000} \approx 35.687815$	$\frac{4\,156\,843}{405\,000} \approx 10.263810$
10	≈ 7.057879	≈ 64.777011	≈ 14.963347
100	≈ 32.441693	≈ 1183.288479	≈ 130.825020
∞	$\sim n/\log n$	$\sim n^2/\log^2 n$	$\sim n^2/(2\log^3 n)$

Remarks. It is remarkable that (23) coincides with the asymptotic expansion for the jth moment of the number X_n of collision events that take place until there is just a single block (Panholzer [19], p. 277 or Theorem 2.1. with $\alpha = 0$, Goldschmidt and Martin [9], Theorem 2.4.). Corollary 4.3 therefore indicates that, for the Bolthausen–Sznitman coalescent, the total branch length L_n is closely related to X_n. We will exploit this fact in more detail in Section 5.

Corollary 4.3 shows that $\lim_{n\to\infty} \mathrm{E}((L_n/\mathrm{E}(L_n))^j) = 1$, $j \in \mathbb{N}$. The same result holds for the sequence $(X_n)_{n\in\mathbb{N}}$ (see Panholzer [19]). The expansions for the first four moments are

$$\mathrm{E}(L_n) = \frac{n}{\log n} + (2-\gamma)\frac{n}{\log^2 n} + O\left(\frac{n}{\log^3 n}\right), \tag{24}$$

$$\mathrm{E}(L_n^2) = \frac{n^2}{\log^2 n} + \left(\frac{9}{2} - 2\gamma\right)\frac{n^2}{\log^3 n} + O\left(\frac{n^2}{\log^4 n}\right), \tag{25}$$

$$\mathrm{E}(L_n^3) = \frac{n^3}{\log^3 n} + \left(\frac{22}{3} - 3\gamma\right)\frac{n^3}{\log^4 n} + O\left(\frac{n^3}{\log^5 n}\right), \tag{26}$$

and

$$\mathrm{E}(L_n^4) = \frac{n^4}{\log^4 n} + \left(\frac{125}{12} - 4\gamma\right)\frac{n^4}{\log^5 n} + O\left(\frac{n^4}{\log^6 n}\right). \tag{27}$$

The same argument as given in [19, p. 277] yields the asymptotic expansion

$$\mathrm{E}((L_n - \mathrm{E}(L_n))^j) = \frac{(-1)^j}{j(j-1)}\frac{n^j}{\log^{j+1} n} + O\left(\frac{n^j}{\log^{j+2} n}\right), \quad j \ge 2, \tag{28}$$

for the centered moments of L_n. In particular, $\mathrm{Var}(L_n) \sim n^2/(2\log^3 n)$. The recursion presented at the end of Section 3 yields Table 1.

From (28) it follows that it is impossible to choose a sequence of positive real numbers $(b_n)_{n\in\mathbb{N}}$ such that all the moments $\mathrm{E}(((L_n - \mathrm{E}(L_n))/b_n)^j)$, $j \in \mathbb{N}$, converge as n tends to infinity. These facts indicate that the moments of L_n (and as well of X_n) do not 'encode' a possible limiting distribution in a proper way.

5. A weak convergence result for the total branch length

In the following we would like to find sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ of real numbers with $b_n > 0$ for sufficiently large n, such that $L_n^* := (L_n - a_n)/b_n$ has a non-degenerate weak limit as n tends to infinity. At a first glance it seems to be tempting to work with $a_n := \mu_n := \mathrm{E}(L_n)$ and $b_n := \sigma_n := \sqrt{\mathrm{Var}(L_n)}$. Then, by (2), $a_n = \mathrm{E}(T_n) + \mathrm{E}(a_{I_n})$ for $n \geq 2$. Thus, the sequence $(L_n^*)_{n\in\mathbb{N}}$, with the so defined a_n and b_n, would satisfy

$$L_n^* = \frac{L_{I_n} + T_n - \mu_n}{\sigma_n} = \frac{\sigma_{I_n}}{\sigma_n} L_{I_n}^* + \frac{T_n - \mathrm{E}(T_n) + \mu_{I_n} - \mathrm{E}(\mu_{I_n})}{\sigma_n}, \quad n \geq 2.$$

For $n \to \infty$, this recursion for $(L_n^*)_{n\in\mathbb{N}}$ leads to a degenerate equation which does not give any hint on the limiting behavior of the sequence $(L_n^*)_{n\in\mathbb{N}}$. Recursions with degenerate limiting equation are well known from the literature. Neininger and Rüschendorf [18] study a class of such recursions with normal limiting behavior. Theorem 2.1 in [18] is not directly applicable in our situation as the condition (10) in [18] is not satisfied. It turns out that another scaling is needed. In order to see this we have to study the random variables X_n, $n \in \mathbb{N}$, recursively defined via $X_1 := 0$ and

$$X_n := 1 + X_{I_n}, \quad n \geq 2, \tag{29}$$

where I_n is independent of $X_1, \ldots, X_{n-1}$ with distribution (9). The variable X_n can be interpreted in different ways.

(i) In the language of coalescent processes, X_n is the number of collision events that take place until there is just a single block.
(ii) In the language of random recursive trees (Panholzer [19]), X_n counts the number of removed edges (in a so-called one-sided edge-removal procedure) until the root is isolated.
(iii) In the language of Markov chains, X_n is the absorption time, i.e. the number of steps to reach the absorbing state 1, of the Markov chain $(D_r^{(n)})_{r\in\mathbb{N}_0}$, recursively defined via $D_0^{(n)} := n$ and $D_r^{(n)} := I_r(D_{r-1}^{(n)})$, $r \in \mathbb{N}$, where $I_1(k), I_2(k), \ldots$ are independent copies of I_k, $k \in \{1, \ldots, n\}$, with the convention $I_1 := 1$.

The recursion (29) is again of the form (8) in [18], but the results in [18] are not directly applicable, because I_n takes large values (close to n) with high probability. Define $a_1 := 0$, $b_1 := 1$, and, for $n \geq 2$,

$$a_n := \frac{n}{\log n} + \frac{n \log\log n}{\log^2 n}, \quad \text{and} \quad b_n := \frac{n}{\log^2 n}. \tag{30}$$

An analytic proof of the following convergence theorem is given in [7]. A probabilistic proof of the same result was found shortly later [10].

Theorem 5.1 (*Weak Convergence of Normalized X_n*). *As n tends to infinity, $(X_n - a_n)/b_n$ converges in distribution to a stable random variable X with characteristic function*

$$\mathrm{E}(\mathrm{e}^{\mathrm{i}tX}) = \exp\left(-\frac{1}{2}\pi|t| + \mathrm{i}t\log|t|\right), \quad t \in \mathbb{R}. \tag{31}$$

Remark. The distribution of $-X$ is the standard continuous Luria–Delbrueck distribution (see [16, Theorem 4.1.]).

We now present the weak convergence result for the total branch length L_n.

Theorem 5.2 (*Weak Convergence of Normalized L_n*). *As n tends to infinity, $(L_n - a_n)/b_n$ converges in distribution to a stable random variable X with characteristic function given in (31).*

Proof. Obviously, $(L_n - a_n)/b_n = (L_n - X_n)/b_n + (X_n - a_n)/b_n$. By Theorem 5.1, it suffices to verify that $(L_n - X_n)/b_n \to 0$ in probability. We even show that $(L_n - X_n)/b_n \to 0$ in L_2. For $n \geq 2$ it follows from (2) that

$$L_n = \sum_{k=2}^{n} T_k \sum_{r=0}^{\infty} 1_{\{D_r^{(n)}=k\}} = \sum_{r=0}^{\infty} T_{D_r^{(n)}} \sum_{k=2}^{n} 1_{\{D_r^{(n)}=k\}} = \sum_{r=0}^{X_n-1} T_{D_r^{(n)}},$$

as $D_r^{(n)} = 1$ for $r \geq X_n$ and $D_r^{(n)} \in \{2, \ldots, n\}$ for $0 \leq r < X_n$. For $k \in \{1, \ldots, n\}$ and $\mathbf{i} = (i_0, \ldots, i_k)$ with $n = i_0 > i_1 > \cdots > i_{k-1} > i_k = 1$ define the events $A_{k,\mathbf{i}} := \{X_n = k, (D_0^{(n)}, \ldots, D_k^{(n)}) = \mathbf{i}\}$. We have

$$\begin{aligned} \mathrm{E}((L_n - X_n)^2) &= \mathrm{E}\left(\left(\sum_{r=0}^{X_n-1}(T_{D_r^{(n)}} - 1)\right)^2\right) \\ &= \sum_{k,\mathbf{i}} P(A_{k,i})\,\mathrm{E}\left(\left(\sum_{r=0}^{k-1}(T_{i_r} - 1)\right)^2\right) \\ &= \sum_{k,\mathbf{i}} P(A_{k,i})\left(\sum_{r=0}^{k-1}\mathrm{E}((T_{i_r} - 1)^2) + \sum_{\substack{r,s=0\\ r\neq s}}^{k-1}\mathrm{E}((T_{i_r} - 1)(T_{i_s} - 1))\right). \end{aligned}$$

The random variables T_{i_r}, $r \in \{0, \ldots, k-1\}$, are independent and exponentially distributed with mean $\mathrm{E}(T_{i_r}) = i_r/(i_r - 1)$. Moreover, $i_r \geq k - r + 1$. Thus,

$$\sum_{r=0}^{k-1}\mathrm{E}(T_{i_r} - 1) = \sum_{r=0}^{k-1}\frac{1}{i_r - 1} \leq \sum_{r=0}^{k-1}\frac{1}{k-r} \leq 1 + \log k \leq 1 + \log n.$$

Furthermore, $\mathrm{E}((T_{i_r} - 1)^2) \leq \mathrm{E}((T_2 - 1)^2) = 5$. Therefore,

$$\begin{aligned} \mathrm{E}((L_n - X_n)^2) &\leq \sum_{k,\mathbf{i}} P(A_{k,\mathbf{i}})\left(\sum_{r=0}^{k-1}\mathrm{E}((T_{i_r} - 1)^2) + \left(\sum_{r=0}^{k-1}\mathrm{E}(T_{i_r} - 1)\right)^2\right) \\ &\leq \sum_{k,\mathbf{i}} P(A_{k,\mathbf{i}})(5k + (1 + \log n)^2) = 5\mathrm{E}(X_n) + (1 + \log n)^2. \end{aligned}$$

Therefore, $\mathrm{E}((L_n - X_n)^2) = O(n/\log n)$, as $\mathrm{E}(X_n) \sim n/\log n$ (see Panholzer [19], p. 277 or Theorem 2.1. with $\alpha = 0$). From the definition of b_n it finally follows that $(L_n - X_n)/b_n \to 0$ in L_2. $\square$

6. Application: Mutations

Assume that mutations occur on each branch of the coalescent tree according to a homogeneous Poisson process $(M_t)_{t\geq 0}$ with rate $r > 0$, which is independent of the coalescent

$(R_t)_{t\geq 0}$. Let S_n denote the total number of mutations on the branches of the tree $(\varrho_n R_t)_{t\geq 0}$. For $t > 0$, the variable M_t is Poisson distributed with parameter rt and has, hence, descending factorial moments $\mathrm{E}((M_t)_j) = (rt)^j$, $j \in \mathbb{N}_0$, where $(x)_0 := 1$ and $(x)_j := x(x-1)\cdots(x-j+1)$ for $j \in \mathbb{N}$ and $x \in \mathbb{R}$. From $S_n \stackrel{d}{=} M_{L_n}$ it follows that S_n has factorial moments

$$\mathrm{E}((S_n)_j) = \mathrm{E}(\mathrm{E}((M_{L_n})_j \mid L_n)) = \mathrm{E}((rL_n)^j) = r^j \mu_n^{(j)}, \quad j \in \mathbb{N}_0,$$

and, hence, moments

$$\mathrm{E}(S_n^j) = \sum_{k=0}^{j} S(j,k)\mathrm{E}((S_n)_k) = \sum_{k=0}^{j} S(j,k) r^k \mu_n^{(k)}, \quad j \in \mathbb{N}_0,$$

where the $S(j,k)$ denote the Stirling numbers of the second kind. In particular, $\mathrm{E}(S_n) = r\mathrm{E}(L_n)$ and

$$\begin{aligned}\mathrm{Var}(S_n) &= \mathrm{E}(\mathrm{Var}(M_{L_n}|L_n)) + \mathrm{Var}(\mathrm{E}(M_{L_n}|L_n))\\ &= \mathrm{E}(rL_n) + \mathrm{Var}(rL_n) = r\mathrm{E}(L_n) + r^2\mathrm{Var}(L_n).\end{aligned}$$

Corollary 6.1 (*Weak Law of Large Numbers for S_n*). *As n tends to infinity, $n^{-1}(\log n)S_n$ converges in probability to r.*

Proof. We have $L_n \to \infty$ almost surely by Corollary 4.4. Thus, $M_{L_n}/L_n \to r$ almost surely and

$$\frac{S_n}{\mathrm{E}(S_n)} \stackrel{d}{=} \frac{M_{L_n}}{rL_n}\frac{L_n}{\mathrm{E}(L_n)} \to 1$$

in probability by Corollary 4.4. The corollary follows from $\mathrm{E}(S_n) = r\,\mathrm{E}(L_n) \sim rn/\log n$. □

Corollary 6.2 (*Weak Convergence of S_n*). *Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be the sequences defined in (30). As n tends to infinity, $(S_n - ra_n)/(rb_n)$ converges in distribution to a stable random variable X with characteristic function given in (31).*

Proof. We have

$$\frac{S_n - ra_n}{rb_n} = \frac{S_n/r - L_n}{b_n} + \frac{L_n - a_n}{b_n}.$$

Thus, by Theorem 5.2, it is sufficient to verify that $Y_n := (S_n/r - L_n)/b_n$ converges to zero in probability. From $\mathrm{E}(S_n/r - L_n) = 0$ and

$$\begin{aligned}\mathrm{Var}\left(\frac{S_n}{r} - L_n\right) &= \mathrm{Var}\left(\mathrm{E}\left(\frac{M_{L_n}}{r} - L_n \,\middle|\, L_n\right)\right) + \mathrm{E}\left(\mathrm{Var}\left(\frac{M_{L_n}}{r} - L_n \,\middle|\, L_n\right)\right)\\ &= 0 + \mathrm{E}\left(\mathrm{Var}\left(\frac{M_{L_n}}{r}\,\middle|\, L_n\right)\right)\\ &= \frac{\mathrm{E}(\mathrm{Var}(M_{L_n} \mid L_n))}{r^2} = \frac{\mathrm{E}(rL_n)}{r^2} = \frac{\mathrm{E}(L_n)}{r}\end{aligned}$$

it follows that $\mathrm{E}(Y_n) = 0$ and that $\mathrm{Var}(Y_n) = \mathrm{E}(L_n)/(rb_n^2) \sim n/(rb_n^2 \log n) \to 0$ by assumption. The convergence $Y_n \to 0$ in probability follows from Tschebyscheff's inequality. □

Acknowledgements

We thank the referee for helpful comments, in particular for pointing out an error in Proposition 4.2 in a first version of the manuscript. The idea of writing the paper arose from several authors' discussions at the 'Frankfurter Stochastik Tage 2006'. The first author's research was supported by the Austrian Science Foundation FWF, grant S9604.

Appendix

In this appendix, some useful results on the total branch length L_n are collected.

Lemma 7.1. *For the Kingman coalescent ($\Lambda = \delta_0$), as n tends to infinity, $L_n/2 - \log n$ converges in distribution to a standard Gumbel distributed random variable.*

Proof. For the Kingman coalescent, $L_n = T_2 + \cdots + T_n$, where the random variables $T_2, \ldots, T_n$ are independent and T_i is exponentially distributed with parameter $\alpha_i = g_i/i = (i-1)/2$, $i \in \{2, \ldots, n\}$. Thus, L_n has distribution function

$$P(L_n \le t) = 1 - \sum_{i=2}^{n} \exp(-\alpha_i t) \prod_{\substack{j=2 \\ j \ne i}}^{n} \frac{\alpha_j}{\alpha_j - \alpha_i}, \quad t \ge 0.$$

From $\prod_{\substack{j=2 \\ j \ne i}}^{n} \frac{\alpha_j}{\alpha_j - \alpha_i} = \prod_{\substack{j=2 \\ j \ne i}}^{n} \frac{j-1}{j-i} = (-1)^i \binom{n-1}{i-1}$ and $\alpha_i = (i-1)/2$ it follows that

$$P(L_n \le t) = 1 - \sum_{i=2}^{n} (\exp(-t/2))^{i-1} (-1)^i \binom{n-1}{i-1} = (1 - \exp(-t/2))^{n-1}. \tag{32}$$

Therefore, for $x \in \mathbb{R}$ and $n \in \mathbb{N}$ such that $x + \log n \ge 0$,

$$P(L_n \le 2x + 2\log n) = (1 - \exp(-x)/n)^{n-1} \to \exp(-\exp(-x))$$

as n tends to infinity. The proof is complete, as $x \mapsto \exp(-\exp(-x))$, $x \in \mathbb{R}$, is the distribution function of the standard Gumbel distribution. □

Remark. The above proof is similar to that given in [25, Chapter 3]. A proof based on a coupling argument appeared in [24, pp. 21–23]. The Gumbel distribution arises because L_n has the same distribution as the maximum of $n-1$ independent and exponentially distributed random variables with parameter $1/2$, as can be seen from (32). This fact previously appeared in [26, pp. 255-257], and also implicitly in [23, p. 153]. The following explanation is given in [25]. Suppose we have $n-1$ exponential clocks, each going off at rate $1/2$. When there are k exponential clocks that have not yet gone off, the time one has to wait for the next one is exponential with rate $k/2$. The maximum of the $n-1$ exponential random variables is the time one has to wait for all $n-1$ clocks to go off, which is $T_2 + \cdots + T_n = L_n$.

For the Bolthausen–Sznitman coalescent, the following lemma provides an explicit formula for $\mu_n := \mathrm{E}(L_n)$ in terms of the absolute Stirling numbers of the first kind. The strict monotonicity of $(\mu_n)_{n \in \mathbb{N}}$ follows immediately. We also provide an alternative proof for the asymptotics of μ_n based on Tauberian theorems.

Lemma 7.2 (*Explicit Formula and Asymptotics of* μ_n). *For the Bolthausen–Sznitman coalescent,*

$$\mu_n = 2\sum_{i=1}^{n-1} \frac{c_i}{i!}, \quad n \in \mathbb{N}, \tag{33}$$

where

$$c_i := \sum_{j=0}^{[(i-1)/2]} \frac{s(i, 2j+1)}{2j+1} > 0, \quad i \in \mathbb{N}, \tag{34}$$

and $s(i, j)$ *denote the absolute Stirling numbers of the first kind. The sequence* $(\mu_n)_{n\in\mathbb{N}}$ *is strictly increasing with asymptotic behavior* $\mu_n \sim n/\log n$ *for* $n \to \infty$.

Proof. Substituting $t = 1 - \mathrm{e}^{-u}$ in (11) yields

$$\mu_1(s) = \frac{s}{1-s} \int_0^{-\log(1-s)} \frac{\mathrm{e}^u - \mathrm{e}^{-u}}{u} \mathrm{d}u, \quad 0 \le s < 1. \tag{35}$$

The Taylor expansion $(\mathrm{e}^u - \mathrm{e}^{-u})/u = 2\sum_{j=0}^{\infty} u^{2j}/(2j+1)!$ leads to

$$\mu_1(s) = \frac{2s}{1-s} \sum_{j=0}^{\infty} \frac{(-\log(1-s))^{2j+1}}{(2j+1)(2j+1)!}.$$

Let $s(i, j)$ denote the absolute Stirling numbers of the first kind. From

$$(-\log(1-s))^j = \left(\sum_{i=1}^{\infty} \frac{s^i}{i}\right)^j = \sum_{i=j}^{\infty} s^i \sum_{\substack{i_1,\dots,i_j=1 \\ i_1+\dots+i_j=i}}^{\infty} \frac{1}{i_1 \cdots i_j} = j! \sum_{i=j}^{\infty} \frac{s^i}{i!} s(i, j)$$

we conclude that

$$\begin{aligned} \mu_1(s) &= \frac{2s}{1-s} \sum_{j=0}^{\infty} \frac{1}{2j+1} \sum_{i=2j+1}^{\infty} \frac{s^i}{i!} s(i, 2j+1) \\ &= \frac{2s}{1-s} \sum_{i=1}^{\infty} \frac{s^i}{i!} \sum_{j=0}^{[(i-1)/2]} \frac{s(i, 2j+1)}{2j+1} \\ &= 2\left(\sum_{k=1}^{\infty} s^k\right) \sum_{i=1}^{\infty} \frac{s^i}{i!} c_i = 2\sum_{n=2}^{\infty} s^n \sum_{i=1}^{n-1} \frac{c_i}{i!}, \end{aligned}$$

with c_i defined in (34). Comparing the coefficient in front of s^n with that in $\mu_1(s) = \sum_{n=1}^{\infty} \mu_n s^n$ yields the explicit solution (33). In particular, the sequence $(\mu_n)_{n\in\mathbb{N}}$ is strictly increasing. From (35) and $\int_1^x \mathrm{e}^u/u \mathrm{d}u \sim \mathrm{e}^x/x$ for $x \to \infty$ it follows with $x = -\log(1-s)$ that

$$\mu_1(s) \sim \frac{1}{1-s} \frac{\mathrm{e}^x}{x} = -\frac{1}{(1-s)^2 \log(1-s)} = (1-s)^{-2} l(1/(1-s))$$

for $s \nearrow 1$, where $l(x) := 1/\log(x)$, $x > 0$, is slowly varying. Karamata's Tauberian theorem for power series [4, Corollary 1.7.3], applied with $\rho := 2$ and $c := 1$ in the notation of that corollary, yields $\mu_n \sim cn^{\rho-1} l(n)/\Gamma(\rho) = n/\log n$ for $n \to \infty$. □

The same method leads to the asymptotics of $\mu_n^{(2)} = \mathrm{E}(L_n^2)$.

Lemma 7.3 (*Asymptotics of* $\mu_n^{(2)}$). $\mu_n^{(2)} \sim n^2/\log^2 n$.

Proof. For $s \nearrow 1$ we have, by (13) and (7),

$$\begin{aligned}\mu_1'(s) &= \frac{\mu_1(s)}{s(1-s)} - \frac{s^2(2-s)}{(1-s)^3\log(1-s)}\\ &\sim -\frac{1}{(1-s)^3\log(1-s)} - \frac{1}{(1-s)^3\log(1-s)}\\ &= -\frac{2}{(1-s)^3\log(1-s)},\end{aligned}$$

or, equivalently, $\mu_1'(1-\mathrm{e}^{-u}) \sim 2\mathrm{e}^{3u}/u$ for $u \to \infty$. Thus,

$$\begin{aligned}\mu_2(s) &= \frac{2s}{s-1}\int_0^s \frac{\mu_1'(t)}{\log(1-t)}\mathrm{d}t\\ &= \frac{2s}{1-s}\int_0^{-\log(1-s)} \frac{\mu_1'(1-\mathrm{e}^{-u})}{u}\mathrm{e}^{-u}\mathrm{d}u\\ &\sim \frac{2}{1-s}\int_1^{-\log(1-s)} \frac{2\mathrm{e}^{2u}}{u^2}\mathrm{d}u\end{aligned}$$

for $s \nearrow 1$. From $\int_1^x \mathrm{e}^{2u}/u^2 \sim \mathrm{e}^{2x}/(2x^2)$ for $x \to \infty$ it follows with $x = -\log(1-s)$ that

$$\mu_2(s) \sim \frac{2}{1-s}\frac{\mathrm{e}^{2x}}{x^2} = \frac{2}{(1-s)^3\log^2(1-s)} = 2(1-s)^{-3}l(1/(1-s))$$

for $s \nearrow 1$, where $l(x) := 1/\log^2 x$ is slowly varying. From Section 3 we know that the sequence $(\mu_n^{(2)})_{n\in\mathbb{N}}$ is non-decreasing. Karamata's Tauberian theorem for power series [4, Corollary 1.7.3], applied with $\rho := 3$ and $c := 2$ in the notation of that corollary, yields $\mu_n^{(2)} \sim cn^{\rho-1}l(n)/\Gamma(\rho) = n^2/\log^2 n$. □

References

[1] J. Berestycki, N. Berestycki, J. Schweinsberg, 2007. Small-time behavior of beta coalescents. Ann. Inst. H. Poincare Probab. Statist. (in press).

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

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

[4] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, Cambridge University Press, 1987.

[5] E. Bolthausen, A.-S. Sznitman, On Ruelle's probability cascades and an abstract cavity method, Comm. Math. Phys. 197 (1998) 247–276.

[6] A. Bovier, I. Kurkova, Much ado about Derrida's GREM, in: E. Bolthausen, A. Bovier (Eds.), Spin Glasses, in: Lecture Notes in Mathematics, vol. 1900, Springer, Berlin, 2006.

[7] M. Drmota, A. Iksanov, M. Moehle, U. Roesler, A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree, 2006. Preprint.

[8] J.A. Fill, N. Kapur, A. Panholzer, Destruction of very simple trees, Algorithmica (2006–2007) (in press).

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

[10] A. Iksanov, M. Möhle, A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree, Electron. Comm. Probab. 12 (2007) 36–43.

[11] S. Janson, Random records and cuttings in complete binary trees, in: Mathematics and Computer Science, vol. III, Birkhäuser, Basel, 2004, pp. 241–253.

[12] S. Janson, Random cutting and records in deterministic and random trees, Random Structures Algorithms 29 (2006) 139–179.

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

[14] J.F.C. Kingman, Origins of the coalescent, Genetics 156 (2000) 1461–1463.

[15] A. Meir, J.W. Moon, Cutting down recursive trees, Math. Biosci. 21 (1974) 173–181.

[16] M. Möhle, Convergence results for compound Poisson distributions and applications to the standard Luria–Delbrück distribution, J. Appl. Probab. 42 (2005) 620–631.

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

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

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

[20] A. Panholzer, Cutting down very simple trees, Questions Math. 29 (2006) 211–227.

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

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

[23] S. Tavaré, Line of descent and genealogical processes, and their applications in population genetics models, Theor. Pop. Biol. 26 (1984) 119–164.

[24] S. Tavaré, Ancestral inference in population genetics, in: Lectures on Probability and Statistics 1–188, in: Lecture Notes in Mathematics, vol. 1837, Springer, Berlin, 2004.

[25] J. Wakeley, Coalescent Theory: An Introduction, Roberts and Company Publishers, Greenwood Village, ISBN: 0-9747077-5-9, 2007 (in press).

[26] C. Wiuf, J. Hein, Recombination as a point process along sequences, Theor. Pop. Biol. 55 (1999) 248–259.