

# On the measure division construction of $\Lambda$ -coalescents

Linglong Yuan

University Paris 13, Sorbonne Paris Cité, LAGA, CNRS, UMR 7539, F-93430, Villetaneuse, France. E-mail: yuan@math.univ-paris13.fr

Received February 5, 2013, revised September 26, 2013

**Abstract.** This paper provides a new construction of  $\Lambda$ -coalescents called “measure division construction”. This construction is pathwise and consists of dividing the characteristic measure  $\Lambda$  into several parts and adding them one by one to have a whole process. Using this construction, a “universal” normalization factor  $\mu^{(n)}$  for the randomly chosen external branch length  $T^{(n)}$  has been discovered for a class of coalescents. This class of coalescents covers processes similar to Bolthausen–Sznitman coalescent, the coalescents without proper frequencies, and also others. A conjecture is also made to extend the universality of  $\mu^{(n)}$  to more processes.

**KEYWORDS:**  $\Lambda$ -coalescent, two-type  $\Lambda$ -coalescent, branch length, measure division construction, noise measure, main measure

**AMS SUBJECT CLASSIFICATION:** Primary 60J25, 60F05, Secondary 92D15, 60K37

## 1. Introduction

### 1.1. Motivation and main results

Let  $\mathbb{N} := \{1, 2, \dots\}$ ,  $\Omega$  be a subset of  $\mathbb{N}$  and  $\pi$  a partition of  $\Omega$  such that  $|\pi| < +\infty$  ( $|\pi|$  denotes the number of blocks in  $\pi$ ). The  $\Lambda$ -coalescent process starting from  $\pi$ , introduced independently by Pitman [27] and Sagitov [28], is denoted by  $\Pi^{(\pi)} := (\Pi^{(\pi)}(t))_{t \geq 0}$ , where  $\Pi^{(\pi)}(0) = \pi$  and  $\Lambda$  is a finite measure on  $[0, 1]$ . Here we specify that a finite measure on  $[0, 1]$  can be a null measure and hence its total mass is a non-negative real value. If  $\pi = \{\{1\}, \{2\}, \dots, \{n\}\}$ , i.e., the set of first  $n$  singletons, then the process is simply denoted by  $\Pi^{(n)}$ . In this paper, we will frequently use two other notations  $\Lambda_1, \Lambda_2$  for finite measures.

We define then  $\Pi^{(1,n)}$  as the  $\Lambda_1$ -coalescent and  $\Pi^{(2,n)}$  the  $\Lambda_2$ -coalescent, both taking  $\{\{1\}, \{2\}, \dots, \{n\}\}$  as initial value.

This process  $\Pi^{(\pi)}$  is a continuous time Markov process with càdlàg trajectories taking values in the set of partitions of  $\Omega$ . More precisely: Assume that at time  $t$ ,  $\Pi^{(\pi)}(t)$  has  $b$  blocks, then after a random exponential time with parameter  $g_b$

$$g_b = \sum_{k=2}^b \binom{b}{k} \lambda_{b,k}, \quad \text{where} \quad \lambda_{b,k} = \int_0^1 x^{k-2} (1-x)^{b-k} \Lambda(dx), \quad (1.1)$$

$\Pi^{(\pi)}$  encounters a collision and the probability for a group of  $k$  ( $2 \leq k \leq b$ ) blocks to be merged into a bigger block with the other  $b - k$  blocks unchanged is

$$\frac{\lambda_{b,k}}{g_b}.$$

Then

$$p_{b,b-k+1} := \frac{\binom{b}{k} \lambda_{b,k}}{g_b} \quad (1.2)$$

is the probability to have  $b - k + 1$  blocks right after the collision. This definition gives the exchangeability of blocks. In particular, for any permutation  $\rho$  on  $\{1, 2, \dots, n\}$ ,  $\rho \circ \Pi^{(n)} \stackrel{(d)}{=} \Pi^{(n)}$ .

Remark that if  $\Lambda(\{0\}) = 0$ , we have the following formula:

$$g_b = \int_0^1 (1 - (1-x)^b - bx(1-x)^{b-1}) x^{-2} \Lambda(dx). \quad (1.3)$$

The definition shows that the law of  $\Pi^{(\pi)}$  is determined by the initial value  $\pi$  and the measure  $\Lambda$  which is hence called characteristic measure.

Notice that  $\Omega$  can be an abstract set and the coalescing mechanism works all the same. The reason why one takes  $\Omega$  as a subset of  $\mathbb{N}$  relies on its applications in the genealogies of populations. We take  $\Pi^{(n)}$  as an example where  $\Omega = \{1, 2, \dots, n\}$ . At time 0, we have  $\Pi^{(n)}(0) = \{\{1\}, \{2\}, \dots, \{n\}\}$  which is interpreted as a sample of  $n$  individuals labelled from 1 to  $n$ . If at time  $t$ ,  $\Pi^{(n)}$  has its first coalescence where  $\{1\}$  and  $\{2\}$  are merged together with the others unchanged, then  $\Pi^{(n)}(t) = \{\{1, 2\}, \{3\}, \dots, \{n\}\}$  which is interpreted as getting the MRCA (most recent common ancestor)  $\{1, 2\}$  of individuals 1 and 2 with the others unchanged at that time. Hence  $\{1, 2, \dots, n\}$  is an absorption state of  $\Pi^{(n)}$  and is the MRCA of all individuals. For more details, we refer to [23, 25] or [1, 6, 15, 20].

Let  $1 \leq m \leq n$  and  $\sigma$  the restriction from  $\{1, 2, \dots, n\}$  to  $\{1, 2, \dots, m\}$ . We have the consistency property:  $\sigma \circ \Pi^{(n)} \stackrel{(d)}{=} \Pi^{(m)}$  (see [27]). According to this

property and exchangeability of blocks, if  $\pi'$  is a subset of  $\pi$ , then the restriction of  $\Pi^{(\pi)}$  from  $\pi$  to  $\pi'$  has the same distribution as that of  $\Pi^{(\pi')}$ . We can also define  $\Pi^{(\pi)}$  when  $|\pi| = +\infty$  by using the consistency property and the definition in finite cases (see [27]).

Let  $|\Pi^{(n)}|$  be the block counting process associated to  $\Pi^{(n)}$  such that  $|\Pi^{(n)}(t)|$  is the number of blocks of  $\Pi^{(n)}(t)$  for any  $t \geq 0$ . Then it decreases from  $n$  at time 0. We denote by  $X_1^{(n)}$  the decrease of number of blocks at the first coalescence. For  $i \in \{1, \dots, n\}$ , we define

$$T_i^{(n)} := \inf \{t \geq 0 \mid \{i\} \notin \Pi_t^{(n)}\}$$

the length of the  $i$ th external branch and  $T^{(n)}$  the length of a randomly chosen external branch. By exchangeability,  $T_i^{(n)} \stackrel{(d)}{=} T^{(n)}$ . We denote by  $L_{ext}^{(n)} := \sum_{i=1}^n T_i^{(n)}$  the total external branch length of  $\Pi^{(n)}$ , and by  $L_{total}^{(n)}$  the total branch length.

The following four classes of  $\Lambda$ -coalescents have been largely studied. We give the results concerning  $T^{(n)}$ , which show a common regularity that we will discuss later.

- $\Lambda = \delta_0$ : Kingman coalescent ([23], [24]). Then  $nT^{(n)}$  is asymptotically distributed with density function  $(8/(2+x)^3)\mathbf{1}\{x \geq 0\}$  ([4], [8], [21]).
- $\Lambda = \Lambda^{leb}$ : Bolthausen–Sznitman coalescent ([5]). Here  $\Lambda^{leb}$  denotes the Lebesgue measure on  $[0, 1]$ . Then  $(\ln n)T^{(n)}$  converges in distribution to  $Exp(1)$  (we denote by  $Exp(r), r > 0$ , the exponential variable with parameter  $r$ ) [11, 17].

- $$\frac{\Lambda(dx)}{dx} = \frac{x^{a-1}(1-x)^{b-1}}{Beta(a,b)}\mathbf{1}\{0 \leq x \leq 1\}, \quad 0 < a < 1, b > 0 :$$

$Beta(a, b)$ -coalescent. Here  $Beta(\cdot, \cdot)$  denotes Euler’s beta function. Then  $n^{1-a}T^{(n)}$  converges in distribution to a random variable  $T(a, b)$  which has density function

$$\frac{\Gamma(a+b)}{(1-a)\Gamma(b)} \left(1 + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)}x\right)^{-(3-2a)/(1-a)} \mathbf{1}\{x \geq 0\}$$

(see [12]).

- $\int_0^1 x^{-1} \Lambda(dx) < +\infty$ : These processes are called coalescents without proper frequencies ([27]). One example is  $Beta(a, b)$ -coalescents with  $a > 1, b > 0$  (see [27], [29]). Then  $(\int_0^1 x^{-1} \Lambda(dx))T^{(n)}$  converges in distribution to  $Exp(1)$  ([19], [26]).

We see a common property for the last three cases concerning one external branch length which is that the normalization factor for  $T^{(n)}$  is  $\mu^{(n)} := \int_{1/n}^1 x^{-1} \times \Lambda(dx)$ . More precisely,

- Bolthausen – Sznitman coalescent: Notice that  $\mu^{(n)} = \ln n$ . Hence directly we have  $\mu^{(n)} T^{(n)} \xrightarrow{(d)} \text{Exp}(1)$ .
- *Beta*( $a, b$ )-coalescent with  $0 < a < 1, b > 0$ :

$$\begin{aligned} \mu^{(n)} &= \int_{1/n}^1 x^{-1} \Lambda(dx) = \int_{1/n}^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-2} (1-x)^{b-1} dx \\ &= \frac{\Gamma(a+b)}{(1-a)\Gamma(a)\Gamma(b)} n^{1-a} + O(1). \end{aligned}$$

Hence  $\mu^{(n)} T^{(n)}$  converges in distribution to  $T(a, b)\Gamma(a+b)/(1-a)\Gamma(a)\Gamma(b)$ .

- If  $\int_0^1 x^{-1} \Lambda(dx) < +\infty$ , then

$$\lim_{n \rightarrow +\infty} \frac{\mu^{(n)}}{\int_0^1 x^{-1} \Lambda(dx)} = 1.$$

Hence  $\mu^{(n)} T^{(n)}$  converges in distribution to  $\text{Exp}(1)$ .

Kingman coalescent can be viewed as the formal limit of *Beta*( $a, b$ )-coalescent with  $0 < a < 1, b > 0$  when  $a$  tends to 0, since the measure  $(x^{a-1} \times (1-x)^{b-1} dx / \text{Beta}(a, b)) \mathbf{1}\{0 \leq x \leq 1\}$  tends weakly to the Dirac measure on  $\{0\}$ . The normalization factor in the case of *Beta*( $a, b$ )-coalescent is  $n^{1-a}$ , and of Kingman coalescent is  $n$ . Then we see that these two factors show also some kind of continuity as  $a$  tends to 0. We can formally take  $n$  as  $\mu^{(n)}$  in the case of Kingman coalescent.

Therefore  $\mu^{(n)}$  is characteristic for the randomly chosen external branch length in those processes considered. Notice that  $\mu^{(n)}$  concerns only the measure  $\Lambda \mathbf{1}\{[1/n, 1]\}$ , so it is natural to think about the influences of measures  $\Lambda \mathbf{1}\{[1/n, 1]\}$  and  $\Lambda \mathbf{1}\{[0, 1/n]\}$  on the external branch lengths. More generally, if  $\Lambda = \Lambda_1 + \Lambda_2$ , how can we evaluate each influence on the construction of the whole  $\Lambda$ -coalescent? If  $\Lambda_1$  is “small” enough, we can imagine that  $\Pi^{(n)}$  looks like  $\Pi^{(2,n)}$  (recall that  $\Pi^{(2,n)}$  is the  $\Lambda_2$ -coalescent). In this case, we call  $\Lambda_1$  the noise measure and  $\Lambda_2$  the main measure. To separate  $\Lambda_1$  and  $\Lambda_2$ , we introduce in the next section the “measure division construction” of a  $\Lambda$ -coalescent. The idea of this construction can be at least tracked back to [2] where the authors consider also a coupling of two finite measures on  $[0, 1]$ . The difference relies on the block labelling.

The main results are as follows:

**Theorem 1.1.** *If  $\Lambda$  satisfies:*

$$\lim_{n \rightarrow +\infty} \frac{g_n}{n\mu^{(n)}} = 0, \tag{1.4}$$

then  $\mu^{(n)}T^{(n)} \xrightarrow{(d)} \text{Exp}(1)$ .

*Remark 1.1.*

- Condition (1.4) implies that  $\Lambda(\{0\}) = 0$ . Indeed, if  $\Lambda(\{0\}) > 0$ , then  $g_n \geq \binom{n}{2}\Lambda(\{0\})$  and  $\mu^{(n)} \leq n\Lambda((0, 1])$ . Then (1.4) is invalid.
- The class of coalescents satisfying condition (1.4) does not contain the  $Beta(a, b)$ -coalescents with  $0 < a < 1$  and  $b > 0$ . The following conjecture uses a description similar to condition (1.4) to include them:

**Conjecture 1.1.** *Let  $c > 0$ . If*

$$\lim_{n \rightarrow +\infty} \frac{g_n}{n\mu^{(n)}} = c,$$

then  $\mu^{(n)}T^{(n)} \xrightarrow{(d)} T_c$ , where  $T_c$  is a random variable with density  $\Gamma(2 - \alpha^*)(1 + cx)^{-\alpha^*/(\alpha^*-1)-1}\mathbf{1}\{x \geq 0\}$ . Here  $\alpha^*$  is the unique solution of the equation

$$\frac{(\alpha - 1)\Gamma(2 - \alpha)}{\alpha} = c.$$

This conjecture is true for  $Beta(a, b)$ -coalescents with  $0 < a < 1$ ,  $b > 0$ . In this case, we have  $c = (1 - a)\Gamma(a)/(2 - a)$ . The coalescents, which are more general than but similar to  $Beta(a, b)$ -coalescents with  $0 < a < 1$ ,  $b > 0$ , studied in [12] also satisfy this conjecture.

*Examples:* We give a short list of typical examples satisfying condition (1.4) which are processes without proper frequencies or similar to Bolthausen–Sztiman coalescent. Define  $\bar{\mu}^{(n)} := \int_{1/n}^1 x^{-2} \Lambda(dx)$ .

**Example 1.1.**  $\int_0^1 x^{-1} \Lambda(dx) < +\infty$ : It suffices to prove that  $\lim_{n \rightarrow +\infty} g_n/n = 0$ . Recalling the expression (1.3) of  $g_n$ , we have, for  $n \geq 2$ ,

$$\begin{aligned} \frac{g_n}{n} &= \frac{\int_0^1 (1 - (1 - x)^n - nx(1 - x)^{n-1})x^{-2} \Lambda(dx)}{n} \\ &= \frac{\int_{1/n}^1 (1 - (1 - x)^n - nx(1 - x)^{n-1})x^{-2} \Lambda(dx)}{n} \\ &\quad + \frac{\int_0^{1/n} (1 - (1 - x)^n - nx(1 - x)^{n-1})x^{-2} \Lambda(dx)}{n} \\ &\leq \frac{\bar{\mu}^{(n)}}{n} + \frac{\int_0^{1/n} n^2 \Lambda(dx)}{n}. \end{aligned} \tag{1.5}$$

The second term

$$\frac{\int_0^{1/n} n^2 \Lambda(dx)}{n} = \int_0^{1/n} n \Lambda(dx) \leq \int_0^{1/n} x^{-1} \Lambda(dx) \longrightarrow 0.$$

For the first term, let  $\varepsilon > 0$  and  $M = 1/\varepsilon$ , then

$$\begin{aligned} \frac{\bar{\mu}^{(n)}}{n} &= \frac{\int_{M/n}^1 x^{-2} \Lambda(dx)}{n} + \frac{\int_{1/n}^{M/n} x^{-2} \Lambda(dx)}{n} \\ &\leq \frac{\int_{M/n}^1 x^{-1} \Lambda(dx)}{M} + \int_{1/n}^{M/n} x^{-1} \Lambda(dx) \leq \varepsilon \int_0^1 x^{-1} \Lambda(dx) + \int_{1/n}^{M/n} x^{-1} \Lambda(dx). \end{aligned}$$

Notice that  $\varepsilon \int_0^1 x^{-1} \Lambda(dx)$  can be arbitrarily small and  $\int_{1/n}^{M/n} x^{-1} \Lambda(dx)$  tends to 0 as  $n$  tends to  $+\infty$ . Then we get that  $\bar{\mu}^{(n)}/n$  tends to 0. Hence if  $\int_0^1 x^{-1} \Lambda(dx) < +\infty$ , condition (1.4) is satisfied.

**Example 1.2.** Bolthausen–Sznitman coalescent: In this case, it is straightforward to prove that  $g_n = n - 1$  and  $\mu^{(n)} = \ln n$ , then  $\lim_{n \rightarrow +\infty} g_n/(n\mu^{(n)}) = \lim_{n \rightarrow +\infty} (n - 1)/(n \ln n) = 0$ .

**Example 1.3.**  $\Lambda$  has a density function  $f_\Lambda$  on  $[0, r)$  where  $0 < r < 1$  and there exists a positive number  $M$  such that  $f_\Lambda < M$  on  $[0, r)$ : This kind of processes can be considered as being *dominated* by the Bolthausen–Sznitman coalescent.

If  $\int_0^1 x^{-1} \Lambda(dx) < +\infty$ , we turn back to the first example. If  $\int_0^1 x^{-1} \Lambda(dx) = +\infty$ , then we have  $g_n \leq 2M(n - 1)$  for  $n$  large enough, hence

$$\limsup_{n \rightarrow +\infty} \frac{g_n}{n\mu^{(n)}} \leq \lim_{n \rightarrow +\infty} \frac{2M(n - 1)}{n\mu^{(n)}} = 0.$$

It turns out that this kind of coalescent also satisfies condition (1.4).

**Example 1.4.**  $\Lambda$  has a density function  $f_\Lambda(x) = p(\ln x^{-1})^q$  on  $[0, r)$  where  $0 < r < 1$  and  $p, q$  are positive numbers: Using (1.5), we have

$$\frac{g_n}{n\mu^{(n)}} \leq \frac{\bar{\mu}^{(n)}}{n\mu^{(n)}} + \frac{\int_0^{1/n} n^2 \Lambda(dx)}{n\mu^{(n)}}, \quad \forall n \geq 2.$$

For two real sequences  $(x_n)_{n \geq 1}, (y_n)_{n \geq 1}$ , we write  $x_n \asymp y_n$ , if there exist two positive constants  $c, C$  such that  $cy_n \leq x_n \leq Cy_n$  for  $n$  large enough. Then it is not difficult to find out that  $\mu^{(n)} \asymp (\ln n)^{q+1}$ ,  $\bar{\mu}^{(n)} \asymp n(\ln n)^q$ ,  $\int_0^{1/n} n^2 \Lambda(dx) \asymp n(\ln n)^q$ . Hence we get  $g_n/(n\mu^{(n)}) \rightarrow 0$ .

**Theorem 1.2.** *If  $\Lambda$  satisfies condition (1.4) and  $\int_0^1 x^{-1} \Lambda(dx) = +\infty$ , then we have:*

$$\mu^{(n)}(T_1^{(n)}, T_2^{(n)}, \dots, T_n^{(n)}, 0, 0, \dots) \xrightarrow{(d)} (e_1, e_2, \dots), \tag{1.6}$$

where  $(e_i)_{i \in \mathbb{N}}$  are independently distributed as  $Exp(1)$ .

*Remark 1.2.* The same result has been proved for Bolthausen–Sznitman coalescent in [11]. The authors have used a moment method. We can apply this theorem to Example 1.4 and Example 1.3 when  $\int_0^1 x^{-1} \Lambda(dx) = +\infty$ . If  $\int_0^1 x^{-1} \Lambda(dx) < +\infty$ , then (1.6) is not true and there is no more asymptotic independence (see [26]).

The following three corollaries have also been proved for Bolthausen–Sznitman coalescent (see [11, 14, 18]).

**Corollary 1.1.** *If  $\Lambda$  satisfies condition (1.4), then for any  $r \in \mathbb{R}^+$ ,*

$$\lim_{n \rightarrow +\infty} \mathbb{E}[(\mu^{(n)} T^{(n)})^r] = \mathbb{E}[e_1^r],$$

where  $e_1$  is distributed as  $Exp(1)$ . Moreover, if  $\int_0^1 x^{-1} \Lambda(dx) = +\infty$ , then for any  $k \in \mathbb{N}$  and any  $(r_1, r_2, \dots, r_k) \in \{\mathbb{R}^+\}^k$ , we have:

$$\lim_{n \rightarrow +\infty} \mathbb{E}\left[\prod_{i=1}^k (\mu^{(n)} T_i^{(n)})^{r_i}\right] = \mathbb{E}\left[\prod_{i=1}^k e_i^{r_i}\right], \tag{1.7}$$

where  $(e_i)_{1 \leq i \leq k}$  are independently distributed as  $Exp(1)$ .

**Corollary 1.2.** *If  $\Lambda$  satisfies condition (1.4) and  $\int_0^1 x^{-1} \Lambda(dx) = +\infty$ , then the total external branch length  $L_{ext}^{(n)}$  satisfies:  $\mu^{(n)} L_{ext}^{(n)} / n$  converges in  $L^2$  to 1.*

**Corollary 1.3.** *If  $\Lambda$  satisfies condition (1.4) and  $\int_0^1 x^{-1} \Lambda(dx) = +\infty$ , then the total branch length  $L_{total}^{(n)}$  satisfies:  $\mu^{(n)} L_{total}^{(n)} / n$  converges in probability to 1.*

*Remark 1.3.*

- In fact, we will prove that  $\lim_{n \rightarrow +\infty} \mathbb{E}[\mu^{(n)} L_{total}^{(n)} / n] = 1$ . Notice that Corollary 1.1 gives  $\lim_{n \rightarrow +\infty} \mathbb{E}[\mu^{(n)} L_{ext}^{(n)} / n] = 1$ . Hence we deduce this corollary using Corollary 1.2.
- If  $\int_0^1 x^{-1} \Lambda(dx) < +\infty$ , then (1.7) and Corollaries 1.2 and 1.3 are not true (see again [26]).

### 1.2. Organization

In Section 2, we introduce the main object of this paper: the measure division construction. At first, one needs to define the restriction by the smallest element which serves as a preliminary step of measure division construction. In the same section, we then introduce the two-type  $\Lambda$ -coalescent which is defined using the measure division construction. This process gives a label *primary* or *secondary* to every block and its every element of a normal  $\Lambda$ -coalescent. Using this process, we can see more clearly the coalescent times of some singletons. For a technical use, we then give a tripling to estimate the number of blocks at small times of  $\Pi^{(1,n)}$  which is related to the noise measure  $\Lambda_1$ .

In Section 3, we at first give a characterization of the condition (1.4). Then we apply the general results obtained in Section 2 to those processes satisfying (1.4). Finally, we give all the proofs for the results presented in the Section 1.

## 2. Measure division construction

### 2.1. Restriction by the smallest element

Let  $\xi_n = \{A_1, \dots, A_{|\xi_n|}\}$ ,  $\chi_n = \{B_1, \dots, B_{|\chi_n|}\}$  be two partitions of  $\{1, 2, \dots, n\}$ . We define  $s_i^A$  (resp.  $s_i^B$ ) as the smallest number in the block  $A_i$  (resp.  $B_i$ ). We define also the notation  $\xi_n \preceq \chi_n$ , if  $|\chi_n| \leq |\xi_n|$  and for any  $1 \leq i \leq |\chi_n|$ ,  $B_i = \cup_{j \in I_i} A_j$ , where  $\{I_i\}_{1 \leq i \leq |\chi_n|}$  is a partition of  $\{1, 2, \dots, |\xi_n|\}$ . Roughly speaking,  $\xi_n$  is finer than  $\chi_n$ .

If  $\xi_n \preceq \chi_n$ , we define the stochastic process  $\bar{\Pi}^{(\chi_n)}$  which is the restriction by the smallest element of  $\Pi^{(\xi_n)}$  from  $\xi_n$  to  $\chi_n$ :

- $\bar{\Pi}^{(\chi_n)}(0) = \chi_n$ ;
- For any  $t \geq 0$ , if  $\Pi^{(\xi_n)}(t) = \{D_i\}_{1 \leq i \leq |\Pi^{(\xi_n)}(t)|}$ , where  $D_i$  denotes a block, then

$$\bar{\Pi}^{(\chi_n)}(t) = \left\{ \bigcup_{s_j^B \in D_i} B_j \right\}_{1 \leq i \leq |\Pi^{(\xi_n)}(t)|},$$

where the empty sets in  $\bar{\Pi}^{(\chi_n)}(t)$  are removed.

Notice that the restriction by the smallest element is defined from path to path (see Figure 1).

**Lemma 2.1.**  $\bar{\Pi}^{(\chi_n)}$  has the same distribution as  $\Pi^{(\chi_n)}$ .

*Proof.* Every block in  $\chi_n$  is identified by its smallest element which belongs to a unique block in  $\xi_n$ . Hence for any  $B_i$  in  $\chi_n$ , there exists a unique  $A_{\tau_i}$  such that  $A_{\tau_i} \in \xi_n$ ,  $A_{\tau_i} \subset B_i$  and  $s_{\tau_i}^A = s_i^B$  with  $\tau_i \in \{1, 2, \dots, |\xi_n|\}$ . Let  $\chi'_n = \{A_{\tau_i}\}_{1 \leq i \leq |\chi_n|}$  and define a new process  $\hat{\Pi}^{(\chi'_n)}$  as follows:



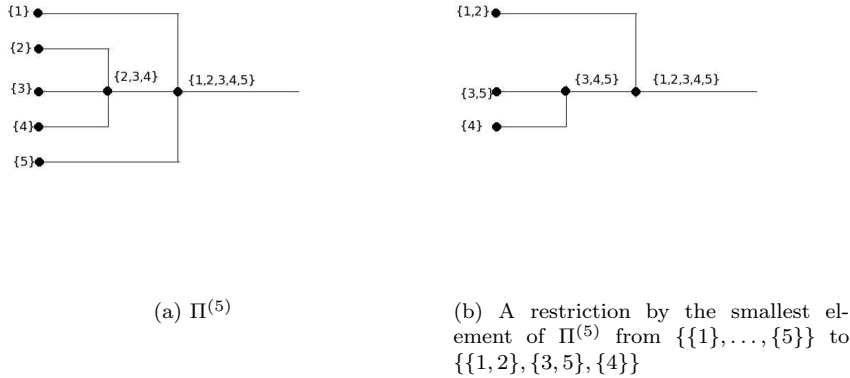


Figure 1. Restriction by the smallest element

- $\hat{\Pi}(\chi'_n)(0) = \chi'_n$ .
- For any  $t \geq 0$ , if  $\Pi^{(\xi_n)}(t) = \{D_i\}_{1 \leq i \leq |\Pi^{(\xi_n)}(t)|}$ , then

$$\hat{\Pi}(\chi'_n)(t) = \left\{ \bigcup_{s_{\tau_j}^A \in D_i} A_{\tau_j} \right\}_{1 \leq i \leq |\Pi^{(\xi_n)}(t)|},$$

where the empty sets in  $\hat{\Pi}(\chi'_n)(t)$  are removed.

It is easy to see that  $\hat{\Pi}(\chi'_n)$  is a natural restriction of  $\Pi^{(\xi_n)}$  from  $\xi_n$  to  $\chi'_n$ . By the consistency property, one gets  $\hat{\Pi}(\chi'_n) \stackrel{(d)}{=} \Pi(\chi'_n)$ . In the construction of  $\hat{\Pi}(\chi'_n)$  and  $\bar{\Pi}(\chi_n)$ , what is determinant is the smallest element in each block. Hence to obtain  $\bar{\Pi}(\chi_n)$  from  $\hat{\Pi}(\chi'_n)$ , at time 0, one needs to complete every  $A_{\tau_i}$  by some other numbers larger than  $s_{\tau_i}^A$  to get  $B_i$  and then follow the evolution of  $\hat{\Pi}(\chi'_n)$ . It turns out that  $\bar{\Pi}(\chi_n)$  is a coalescent process with initial value  $\chi_n$ . Hence we can conclude.  $\square$

### 2.2. Measure division construction

Let  $\Lambda, \Lambda_1, \Lambda_2$  be three finite measures such that  $\Lambda = \Lambda_1 + \Lambda_2$ . We denote by  $\Pi_{1,2}^{(n)} := (\Pi_{1,2}^{(n)}(t))_{t \geq 0}$  the stochastic process constructed by the measure division construction using  $\Lambda_1$  and  $\Lambda_2$ . Here the index  $(1, 2)$  is for  $\Lambda = \Lambda_1 + \Lambda_2$  with  $\Lambda_1$  called noise measure and  $\Lambda_2$  main measure. Recall that  $\Pi^{(1,n)}$  is the  $\Lambda_1$ -coalescent with  $\Pi^{(1,n)}(0) = \{\{1\}, \{2\}, \dots, \{n\}\}$ .

- Step 0: Given a realization or a path  $\Pi$  of  $\Pi^{(1,n)}$ , we set  $\Pi_{1,2}^{(n)}(t) = \Pi(t)$ , for any  $t \geq 0$ . We set also  $t_0 = 0$ .

- Step 1: Let  $t_1, t_2, \dots$  be the coalescent times after  $t_0$  of  $\Pi_{1,2}^{(n)}$  (if there is no collision after  $t_0$ , we set  $t_i = +\infty, i \geq 1$ ). Within  $[t_0, t_1)$ ,  $\Pi_{1,2}^{(n)}$  is constant. Then we run an independent  $\Lambda_2$ -coalescent with initial value  $\Pi_{1,2}^{(n)}(t_0)$  from time  $t_0$ .
  - If the  $\Lambda_2$ -coalescent has no collision on  $[t_0, t_1)$ , we pass to  $[t_1, t_2)$ . Similarly, we construct another independent  $\Lambda_2$ -coalescent with initial value  $\Pi_{1,2}^{(n)}(t_1)$  from time  $t_1$ , and so on.
  - Otherwise, we go to the next step.
- Step 2: If finally within  $[t_{i-1}, t_i)$ , the related independent  $\Lambda_2$ -coalescent has its first collision at time  $t_*$  and its value at  $t_*$  is  $\xi$ . We then modify  $(\Pi_{1,2}^{(n)}(t))_{t \geq 0}$  in the following way:
  - We change nothing for  $0 \leq t < t_*$ .
  - Let  $\Pi' = (\Pi'(t), t \geq t_*)$  be the restriction by the smallest element of  $(\Pi_{1,2}^{(n)}(t))_{t \geq t_*}$  from  $\Pi_{1,2}^{(n)}(t_*)$  to  $\xi$ . Then let  $(\Pi_{1,2}^{(n)}(t))_{t \geq t_*} = (\Pi'(t))_{t \geq t_*}$  and go to the step 1 by taking  $t_*$  as a new starting point. Notice that, due to Lemma 2.1,  $(\Pi_{1,2}^{(n)}(t))_{t \geq t_*}$  has the same distribution as a  $\Lambda_1$ -coalescent from time  $t_*$  with initial value  $\xi$ .

*Remark 2.1.*

- The measure division construction works path by path.
- If we take  $\Lambda_1 = 0$  as noise measure and  $\Lambda_2 = \Lambda$  as main measure, then  $\Pi^{(1,n)}(t) = \{\{1\}, \{2\}, \dots, \{n\}\}$  for any  $t \geq 0$  and  $\Pi_{1,2}^{(n)} \stackrel{(d)}{=} \Pi^{(n)}$ .

**Theorem 2.1.** *Let  $\Lambda, \Lambda_1$  and  $\Lambda_2$  be three finite measures and  $\Lambda = \Lambda_1 + \Lambda_2$ . Then we have  $\Pi_{1,2}^{(n)} \stackrel{(d)}{=} \Pi^{(n)}$ .*

*Proof.* Let  $t$  be a coalescent time of  $\Pi_{1,2}^{(n)}$ . We consider the time of the next coalescence and the value at that moment. In the measure division construction of  $\Pi_{1,2}^{(n)}$ , we can see appearing two independent processes with one being a  $\Lambda_1$ -coalescent with initial value  $\Pi_{1,2}^{(n)}(t)$  and the other one being a  $\Lambda_2$ -coalescent with initial value  $\Pi_{1,2}^{(n)}(t)$  from time  $t$ . The process  $\Pi_{1,2}^{(n)}$  gets the next coalescence whenever one of them first encounters a coalescence and picks up the value of the process at that moment. Then we follow the same procedure from the new coalescent time of  $\Pi_{1,2}^{(n)}$ . It is easy to see that  $\Pi_{1,2}^{(n)}$  behaves in the same way as  $\Pi^{(n)}$ . Hence we can conclude. □

*Remark 2.2.* The theorem shows that if we exchange the noise measure and the main measure, the distribution of the process is not changed and is uniquely determined by their sum.

*Remark 2.3.* The measure division construction also works for more than two measures. If there are  $k(k \geq 2)$  finite measures  $\{\Lambda_i\}_{1 \leq i \leq k}$  and  $\Lambda = \sum_{i=1}^k \Lambda_i$ , one can get a stochastic process by first giving a realization of  $\Pi^{(1,n)}$  which will be modified by  $\Lambda_2$  in the way described in the measure division construction, and then we apply  $\Lambda_3$  on the modified process, etc. The equivalence in distribution can be obtained in a recursive way.

We give a corollary to show an immediate application of the measure division construction. The following corollary is essentially the same as Lemma 3.2 in [2]. But we prove it again in our way.

**Corollary 2.1.** *Let  $\Lambda_1, \Lambda_2$  be two finite measures such that  $\Lambda_1 \leq \Lambda_2$ , then one can construct  $\Pi^{(1,n)}$  and  $\Pi^{(2,n)}$  such that  $|\Pi^{(2,n)}(t)| \leq |\Pi^{(1,n)}(t)|$  for all  $t \geq 0$ .*

*Proof.*  $\Pi^{(2,n)}$  can be regarded as the measure constructed process by imposing the measure  $\Lambda_2 - \Lambda_1$  on the paths of  $\Pi^{(1,n)}$ . Then we can deduce this corollary.  $\square$

### 2.3. Two-type $\Lambda$ -coalescents

#### 2.3.1. Definitions

Let  $\Lambda, \Lambda_1, \Lambda_2$  be three finite measures and  $\Lambda = \Lambda_1 + \Lambda_2$  and  $\Lambda_2$  satisfies  $\int_0^1 x^{-2} \Lambda_2(dx) < +\infty$ . A two-type  $\Lambda$ -coalescent, denoted by  $\tilde{\Pi}_{1,2}^{(n)}$ , is to give a label *primary* or *secondary* to every block and also to its every element at any time  $t$  of a normal  $\Lambda$ -coalescent. A block is *secondary* if and only if every element in this block is *secondary*. The construction is via the measure division construction. Let  $(\eta_i^{(2)})_{i \geq 1}$  be independent random variables following the distribution of  $(\int_0^1 x^{-2} \Lambda_2(dx))^{-1} x^{-2} \Lambda_2(dx)$ ,  $(e_i^{(2)})_{i \geq 1}$  i.i.d. copies of  $Exp(\int_0^1 x^{-2} \Lambda_2(dx))$  and  $(S_i^{(2)})_{i \geq 1} = (\sum_{j=1}^i e_j^{(2)})_{i \geq 1}$ .

Construction of a two-type  $\Lambda$ -coalescent:

- Step 0: We pick a realization or a path  $\Pi$  of  $\Pi^{(1,n)}$ . Every element and every block of  $\Pi$  at any time is labeled *primary*. We also fix independent realizations of  $(\eta_i^{(2)})_{i \geq 1}$  and  $(S_i^{(2)})_{i \geq 1}$ . Let  $\tilde{\Pi}_{1,2}^{(n)}$  be the path  $\Pi$  with labels.
- Step 1: At time  $S_1^{(2)}$ , every block of  $\tilde{\Pi}_{1,2}^{(n)}(S_1^{(2)})$  is independently marked “Head” with probability  $\eta_1^{(2)}$  and “Tail” with probability  $1 - \eta_1^{(2)}$ . Every element in a “Head” block is then labeled *secondary*. All blocks marked “Head” are merged into a bigger block, provided that there are at least two “Head”s. In this case, we use the restriction by the smallest element to modify  $\tilde{\Pi}_{1,2}^{(n)}$  at time  $S_1^{(2)}$  in the same way as in the measure division construction in Section 2.2. We still call the modified path  $\tilde{\Pi}_{1,2}^{(n)}$  and then

forward to the time  $S_2^{(2)}$  and do the same operations. This procedure can be continued until MRCA.

It is easy to verify that without labels,  $\tilde{\Pi}_{1,2}^{(n)}$  has the same distribution as  $\Pi^{(n)}$ . We call  $(S_i^{(2)})_{i \geq 1}$  the *marking times*. We define  $L_i^{(2,n)}$  as the *first marking time* of  $\{i\}$  when  $\{i\}$  is marked “Head” for the first time. Let  $L_i^{(2,n)} = +\infty$ , if  $\{i\}$  is never marked as “Head”.

*Remark 2.4.* If  $\Pi = \{\{1\}, \dots, \{n\}\}$ , then we get a coupling between  $\Lambda_2$ -coalescent and its related annihilator process (see [14]). In this case, the whole process without labels is the  $\Lambda_2$ -coalescent and the restriction to *primary* elements and blocks is the annihilator process.

### 2.3.2. Coalescent times and *first marking times*

The above construction of two-type coalescents shows that coalescences happen only at the *marking times*. This property will help us to understand the coalescent times of singletons in terms of their *first marking times*.

**Lemma 2.2.** *Let  $\Pi$  be the path of  $\Pi^{(1,n)}$  chosen at the Step 0 of the construction of two-type  $\Lambda$  coalescent. Assume that at some time  $t > 0$ ,  $\{1\} \in \Pi(t)$ ,  $|\Pi(t)| = m$  with  $2 \leq m \leq n$ . Let  $P_{1,2}^{(n,m)}(t)$  be the probability for  $\{1\}$  to be coalesced at its *first marking time* within  $[0, t)$ . Then we have*

$$P_{1,2}^{(n,m)}(t) \geq P_t^{(2,m)} := \sum_{i=1}^{+\infty} \mathbb{E}[\mathbf{1}\{S_i^{(2)} < t\} \Delta_i^{(2)} (1 - (1 - \Delta_i^{(2)})^{m-1})], \quad (2.1)$$

where  $\Delta_1^{(2)} = \eta_1^{(2)}$ ;  $\Delta_i^{(2)} = \eta_i^{(2)} \prod_{j=1}^{i-1} (1 - \eta_j^{(2)})$  for  $i > 1$ . Notice that the parameter  $n$  is hidden in  $P_t^{(2,m)}$ .

*Proof.* Let  $i_1, \dots, i_m$  be the  $m$  smallest elements respectively in each block at time  $t$  with  $1 = i_1 \leq i_2 \leq \dots \leq i_m \leq n$ .

Conditional on  $(S_i^{(2)}, \eta_i^{(2)})_{i \geq 1}$ ,  $\Delta_i^{(2)}$  is the probability for  $\{1\}$  to have its *first marking time* at  $S_i^{(2)}$  (assume that  $S_i^{(2)} \leq t$ ). To let  $\{1\}$  be coalesced at  $S_i^{(2)}$ , one needs also at least one other block marked “Head” at that time. To get a lower bound of  $P_{1,2}^{(n,m)}(t)$ , one can consider the probability to have at least one *primary* block containing one element of  $\{i_2, \dots, i_m\}$  to be marked “Head” at that time and this probability is  $1 - (1 - \Delta_i^{(2)})^{m-1}$ .  $\square$

**Lemma 2.3.** *In addition to the assumptions in the previous lemma, we assume further that for every fixed  $k$  such that  $1 \leq k \leq m$ ,  $\{i\} \in \Pi(t)$  for  $1 \leq i \leq k$ .*

Define the probability  $P_{1,2}^{(n,m,k)}(t)$  for every  $\{i\}$  to be coalesced at its first marking time within  $[0, t)$ . Then we have

$$P_{1,2}^{(n,m,k)}(t) \geq 1 - k(1 - P_t^{(2,m)}). \tag{2.2}$$

*Proof.* Let  $E = \{\forall 1 \leq i \leq k, \{i\} \in \Pi(t); |\Pi(t)| = m\}$ , which denotes the assumptions of  $\Pi(t)$  in this Lemma. Then

$$\begin{aligned} &P_{1,2}^{(n,m,k)}(t) \\ &= \mathbb{P}(\{1\}, \dots, \{k\} \text{ coalesce at their first marking times within } [0, t) \mid E) \\ &= 1 - \mathbb{P}(\text{one of } \{\{1\}, \dots, \{k\}\} \text{ does not coalesce} \\ &\quad \text{at its first marking times within } [0, t) \mid E) \\ &\geq 1 - \sum_{i=1}^k \mathbb{P}(\{i\} \text{ does not coalesce at its first marking time within } [0, t) \mid E) \\ &= 1 - k(1 - \mathbb{P}(\{1\} \text{ coalesces at its first marking time within } [0, t) \mid E)) \\ &\geq 1 - k(1 - P_t^{(2,m)}). \end{aligned}$$

The last inequality is due to the fact that

$$\mathbb{P}(\{1\} \text{ coalesces at its first marking time within } [0, t) \mid E) \geq P_t^{(2,m)},$$

which is true due to the same arguments used in the proof of the last lemma.  $\square$

If  $m, t$  are large enough such that under some assumptions, we could prove that  $P_t^{(2,m)}$  is very close to 1. Then the coalescent times are almost the *first marking times* which are easier to deal with. In Section 3.3, we will see such a situation for  $\Lambda$  satisfying condition (1.4) and  $\Lambda_1 = \Lambda \mathbf{1}\{[0, 1/n)\}$ ,  $\Lambda_2 = \Lambda \mathbf{1}\{[1/n, 1]\}$ . The following corollary studies the *first marking times* in this particular case.

**Corollary 2.2.** *Let  $t > 0$  and  $1 \leq k \leq n$ . Assume that  $\Lambda$  satisfies condition (1.4) and  $\Lambda_1 = \Lambda \mathbf{1}\{[0, 1/n)\}$ ,  $\Lambda_2 = \Lambda \mathbf{1}\{[1/n, 1]\}$ . Let  $\Pi$  be a path of  $\Pi^{(1,n)}$ . Recall that  $L_i^{(2,n)}$  is the first marking time of  $\{i\}$  for  $1 \leq i \leq n$ .*

- *If  $\{1\} \in \Pi(t/\mu^{(n)})$ , then for any  $0 \leq t_1 \leq t$ ,  $\mathbb{P}(L_1^{(2,n)} \geq (t_1/\mu^{(n)}) \mid \Pi) = \exp\{-t\}$ .*
- *Fix  $k$  such that  $1 \leq k \leq n$ . Assume  $\int_0^1 x^{-1} \Lambda(dx) = +\infty$  and  $\{i\} \in \Pi(t/\mu^{(n)})$  for any  $1 \leq i \leq k$ . Let  $0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq t$ , we then have*

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(L_i^{(2,n)} \geq \frac{t_i}{\mu^{(n)}}, \forall 1 \leq i \leq k \mid \Pi\right) = \exp\left\{-\sum_{i=1}^k t_i\right\}. \tag{2.3}$$

*Proof.* The first case is easy to see, due to the definition of  $L_1^{(2,n)}$ . For the second case, we only consider  $k = 2$ . For  $k > 2$ , the proof is similar. Assume that within  $[0, t_1/\mu^{(n)}]$ , there are  $N_1$  marking times and for  $(t_1/\mu^{(n)}, t_2/\mu^{(n)}]$ , there are  $N_2$  marking times.  $N_1$  and  $N_2$  are independently Poisson distributed with parameters respectively  $\bar{\mu}^{(n)}t_1/\mu^{(n)}$  and  $\bar{\mu}^{(n)}(t_2 - t_1)/\mu^{(n)}$  (here we have  $\bar{\mu}^{(n)} = \int_{1/n}^1 x^{-2} \Lambda(dx) = \int_{1/n}^1 x^{-2} \Lambda_2(dx)$ ). Then we get

$$\begin{aligned} & \mathbb{P}\left(L_1^{(2,n)} \geq \frac{t_1}{\mu^{(n)}}, L_2^{(2,n)} \geq \frac{t_2}{\mu^{(n)}} \mid \Pi\right) \\ &= \mathbb{E}\left[\prod_{i=1}^{N_1} (1 - \eta_i^{(2)})^2 \prod_{i=N_1+1}^{N_1+N_2} (1 - \eta_i^{(2)})\right] \\ &= \mathbb{E}\left[(1 - 2\mathbb{E}[\eta_1^{(2)}] + \mathbb{E}[(\eta_1^{(2)})^2])^{N_1}\right] \mathbb{E}\left[(1 - \mathbb{E}[\eta_1^{(2)}])^{N_2}\right] \\ &= \exp\left(\frac{\bar{\mu}^{(n)}t_1}{\mu^{(n)}}(-2\mathbb{E}[\eta_1^{(2)}] + \mathbb{E}[(\eta_1^{(2)})^2])\right) \exp\left(\frac{\bar{\mu}^{(n)}(t_2 - t_1)}{\mu^{(n)}}(-\mathbb{E}[\eta_1^{(2)}])\right), \end{aligned}$$

where the last equality is due to the probability generating function of Poisson distribution. Recall that  $\mathbb{E}[\eta_1^{(2)}] = \mu^{(n)}/\bar{\mu}^{(n)}$  and  $\mathbb{E}[(\eta_1^{(2)})^2] = (\bar{\mu}^{(n)})^{-1} \times \int_{1/n}^1 \Lambda(dx)$ . Therefore,

$$\frac{\bar{\mu}^{(n)}}{\mu^{(n)}} \mathbb{E}[\eta_1^{(2)}] = \frac{\int_{1/n}^1 x^{-1} \Lambda(dx)}{\mu^{(n)}} = 1, \quad \text{and} \quad \frac{\bar{\mu}^{(n)}}{\mu^{(n)}} \mathbb{E}[(\eta_1^{(2)})^2] = \frac{\int_{1/n}^1 \Lambda(dx)}{\mu^{(n)}} \longrightarrow 0.$$

Then we can conclude (2.3).  $\square$

#### 2.4. A tripling

We often have some results on the coalescent related to a special measure, for example, the *Beta*-coalescent. When the process is perturbed by a noise measure, we would wonder whether this damage is negligible. One example is to estimate the number of blocks of the coalescent related to the noise measure. To this aim, we use a tool of tripling.

*Tripling:* Notice that  $\Pi^{(n)}$  encounters its first collision after time  $e_1^{(n)}$ , which is a random variable. At this collision, the number of blocks is reduced to  $n - W_1^{(n)}$ , where  $W_1^{(n)}$  is a positive integer valued random variable. Then we add  $W_1^{(n)}$  new blocks (these blocks can contain any number belonging to  $\{n+1, n+2, \dots\}$ ) and consider the whole new  $n$  ones. By the consistency property, the evolution of the original  $n - W_1^{(n)}$  blocks can be embedded into that of the new  $n$  blocks, i.e. after time  $e_2^{(n)}$ , we have the collision in the new  $n$  blocks whose total number is reduced to  $n - W_2^{(n)}$  and we can calculate the distribution of the number of blocks coalesced among the original  $n - W_1^{(n)}$  blocks (we call any block containing at least one of  $\{1, 2, \dots, n\}$  as ‘‘original block’’ and

it is very possible that nothing happens for the  $n - W_1^{(n)}$  blocks). Then we add again new blocks containing different elements to have another  $n$  ones. This procedure is stopped when every element of  $\{1, 2, \dots, n\}$  is contained in one block. By the definition of  $\Lambda$ -coalescent,  $(e_i^{(n)})_{i \geq 1}$  are independent exponential random variables with parameter  $g_n$  and  $(W_i^{(n)})_{i \geq 1}$  are i.i.d. copies of  $X_1^{(n)}$ .

The above procedure gives a tripling of  $(e_i^{(n)})_{i \geq 1}$ ,  $(W_i^{(n)})_{i \geq 1}$  and  $\Pi^{(n)}$ . We define  $V_i^{(n)} := \sum_{j=1}^i e_j^{(n)}$ ,  $i \in \mathbb{N}$ . Then we have the following:

**Proposition 2.1.** *Suppose that  $(e_i^{(n)})_{i \geq 1}$ ,  $(W_i^{(n)})_{i \geq 1}$  and  $\Pi^{(n)}$  are tripled, then at any time  $t \geq 0$ , we have*

$$n - \sum_{i=0}^{N(\Lambda, n, t)} W_i^{(n)} \leq |\Pi^{(n)}(t)|, \tag{2.4}$$

where  $N(\Lambda, n, t) := \text{card}\{i \mid V_i^{(n)} \leq t\}$ , which is Poisson distributed with parameter  $g_n t$  and independent of  $(W_i^{(n)})_{i \geq 1}$ . Meanwhile,

$$\begin{aligned} \mathbb{E}[W_i^{(n)}] &= \frac{n \int_0^1 (1 - (1-x)^{n-1}) x^{-1} \Lambda(dx)}{g_n} - 1, \\ \mathbb{E}[(W_i^{(n)})^2] &= \frac{n(n-1) \int_0^1 \Lambda(dx)}{g_n} - \mathbb{E}[W_i^{(n)}]. \end{aligned} \tag{2.5}$$

*Proof.* The number of  $i$ s within  $[0, t]$  follows the Poisson distribution with parameter  $g_n t$ . Due to the tripling, at any time  $V_i^{(n)}$  with  $0 \leq V_i^{(n)} \leq t$ , the decrease of number of blocks (i.e.  $|\Pi^{(n)}(V_i^{(n)} -)| - |\Pi^{(n)}(V_i^{(n)})|$ ) of original blocks is less than or equal to  $W_i^{(n)}$ . Hence we get (2.4). Notice that  $W_i^{(n)} \stackrel{(d)}{=} X_1^{(n)}$ , then (2.5) is a consequence of two equalities in [9] with eq. (17) for the first one and p.1007 for the second one.  $\square$

### 3. Applications to coalescents satisfying condition (1.4)

#### 3.1. Characterization of condition (1.4)

Some notations for this section: Let  $\Lambda$  be a finite measure on  $[0, 1]$  and  $\Lambda_1 = \Lambda \mathbf{1}\{[0, 1/n)\}$ ,  $\Lambda_2 = \Lambda \mathbf{1}\{[1/n, 1]\}$ ;

$$\begin{aligned} \mu^{(1/y)} &= \int_y^1 x^{-1} \Lambda(dx), \\ g_{1/y} &= \int_0^1 \left(1 - (1-x)^{1/y} - \frac{1}{y} x(1-x)^{1/y-1}\right) x^{-2} \Lambda(dx) \end{aligned}$$

with  $0 < y \leq 1$ . Notice that the definitions of  $\mu^{(1/y)}$  and  $g_{1/y}$  are consistent with that of  $\mu^{(n)}$  and  $g_n$  when  $\Lambda(\{0\}) = 0$ . For any real number  $x$ , let  $\lfloor x \rfloor = \max\{y; y \in \mathbb{Z}, y \leq x\}$  and  $\lceil x \rceil = \min\{y; y \in \mathbb{Z}, y \geq x\}$ .

Here we are going to prove Theorem 1.1, Theorem 1.2, Corollary 1.1, Corollary 1.2 and Corollary 1.3. Under condition (1.4), we decompose  $\Lambda$  into  $\Lambda_2$  and  $\Lambda_1$ . The idea is to construct  $\Pi^{(n)}$  using measure division construction with noise measure  $\Lambda_1$  and main measure  $\Lambda_2$ . At first, we need to show more details implied by condition (1.4).

**Proposition 3.1.** *The following two assertions are equivalent:*

(\*)  $\Lambda$  satisfies condition (1.4);

(\*\*)  $\Lambda(\{0\}) = 0$  and there exists a càglàd (limit from right, continuous from left) function  $f : [0, 1] \rightarrow [0, 1]$ , continuous at 0 with  $f(0) = 0$  such that  $\int_0^1 \mu^{(1/x)} dx < +\infty$  and

$$\mu^{(1/y)} = \left( \int_0^1 \mu^{(1/x)} dx \right) \exp \left( \int_y^1 \frac{f(t)}{t} dt \right) (1 - f(y)), \quad 0 < y \leq 1. \quad (3.1)$$

*Proof. Part 1:* We first assume that (\*) is true. If  $\Lambda$  satisfies (1.4), then  $\Lambda(\{0\}) = 0$  due to Remark 1.1. For  $\mu^{(n)} \neq 0$ , we have

$$\frac{g_n}{n\mu^{(n)}} = \frac{\int_0^1 (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda(dx)}{n\mu^{(n)}} = I_1^{(n)} + I_2^{(n)},$$

where

$$I_1^{(n)} = \frac{\int_{1/n}^1 (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda(dx)}{n\mu^{(n)}},$$

$$I_2^{(n)} = \frac{\int_0^{1/n} (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda(dx)}{n\mu^{(n)}}.$$

Notice that for  $n$  large, using monotone property, we have

$$\frac{e-2}{2e} \frac{\int_{1/n}^1 x^{-2} \Lambda(dx)}{n\mu^{(n)}} \leq I_1^{(n)} \leq \frac{\int_{1/n}^1 x^{-2} \Lambda(dx)}{n\mu^{(n)}}$$

and

$$\frac{1}{3} \frac{n \int_0^{1/n} \Lambda(dx)}{\mu^{(n)}} \leq I_2^{(n)} \leq \frac{n \int_0^{1/n} \Lambda(dx)}{\mu^{(n)}}.$$



Hence condition (1.4) is equivalent to

$$\lim_{n \rightarrow +\infty} \frac{\int_{1/n}^1 x^{-2} \Lambda(dx)}{n\mu^{(n)}} = 0, \quad \text{and} \quad \lim_{n \rightarrow +\infty} \frac{n \int_0^{1/n} \Lambda(dx)}{\mu^{(n)}} = 0, \quad \Lambda(\{0\}) = 0. \tag{3.2}$$

Then we deduce that

$$\lim_{y \rightarrow 0+} \frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} = 0, \quad \Lambda(\{0\}) = 0. \tag{3.3}$$

Indeed, for  $1/y > 2$  and  $\mu^{(\lfloor 1/y \rfloor)} \neq 0$ , we have

$$\begin{aligned} \frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} &= \frac{\int_0^y \Lambda(dx)}{y \int_y^1 x^{-1} \Lambda(dx)} \leq \frac{\int_0^{1/\lfloor 1/y \rfloor} \Lambda(dx)}{\lfloor 1/y \rfloor^{-1} \int_{1/\lfloor 1/y \rfloor}^1 x^{-1} \Lambda(dx)} \\ &= \frac{\lfloor 1/y \rfloor}{\lfloor 1/y \rfloor} \frac{\lfloor 1/y \rfloor \int_0^{1/\lfloor 1/y \rfloor} \Lambda(dx)}{\int_{1/\lfloor 1/y \rfloor}^1 x^{-1} \Lambda(dx)} \xrightarrow{y \rightarrow 0+} 0. \end{aligned}$$

One thing to notice is that  $\lim_{y \rightarrow 0+} y\mu^{(1/y)} = 0$  is true for any finite  $\Lambda$ . In fact, for any positive number  $M$  and  $yM < 1$ , we have

$$\begin{aligned} y\mu^{(1/y)} &= y \int_y^1 x^{-1} \Lambda(dx) = y \int_{yM}^1 x^{-1} \Lambda(dx) + y \int_y^{yM} x^{-1} \Lambda(dx) \\ &\leq \frac{\int_0^1 \Lambda(dx)}{M} + \int_y^{yM} \Lambda(dx), \end{aligned}$$

where both terms can be made as small as we want by taking  $M$  large enough and  $y$  close enough to 0. Looking into details of  $(y\mu^{(1/y)})^{-1} \int_0^y \Lambda(dx)$  when  $\mu^{(1/y)} \neq 0$ , we have the following equality, using integration by parts and  $\lim_{y \rightarrow 0+} y\mu^{(1/y)} = 0$ ,

$$\frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} = \frac{\int_0^y x x^{-1} \Lambda(dx)}{y\mu^{(1/y)}} = \frac{\int_0^y \mu^{(1/x)} dx - y\mu^{(1/y+)}}{y\mu^{(1/y)}}, \tag{3.4}$$

where  $\mu^{(1/y+)} = \mu^{(1/y)} - y^{-1}\Lambda(\{y\})$ . Due to (3.3), we get that

$$1 \geq \frac{\mu^{(1/y+)}}{\mu^{(1/y)}} = 1 - \frac{\Lambda(\{y\})}{y\mu^{(1/y)}} \geq 1 - \frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} \longrightarrow 1.$$

Therefore, (3.3) and (3.4) give

$$\lim_{y \rightarrow 0+} \frac{y\mu^{(1/y)}}{\int_0^y \mu^{(1/x)} dx} = 1.$$

Notice that  $\int_0^y \mu^{(1/x)} dx \geq y\mu^{(1/y)}$  and  $\mu^{(1/y)}$  is a càglàd function. Hence there exists a càglàd function  $f : [0, 1] \rightarrow [0, 1]$ , continuous at 0 with  $f(0) = 0$  such that

$$\frac{y\mu^{(1/y)}}{\int_0^y \mu^{(1/x)} dx} = 1 - f(y). \tag{3.5}$$

Now let  $G(t) = \int_0^t \mu^{(1/x)} dx$  and any derivative is considered as left derivative. Then (3.5) becomes

$$(\ln G(t))' = \frac{G(t)'}{G(t)} = \frac{1 - f(t)}{t}.$$

Using the fundamental theorem of Newton and Leibniz which also works for càglàd functions whose primitive functions take left derivatives, one gets that for  $0 < y \leq 1$ ,

$$\ln G(1) - \ln G(y) = \int_y^1 (\ln G(t))' dt = \int_y^1 \frac{1 - f(t)}{t} dt.$$

Therefore,

$$G(y) = G(1) \exp\left(-\int_y^1 \frac{1 - f(t)}{t} dt\right).$$

By taking the left derivatives on the both sides and noticing that  $G(1) = \int_0^1 \mu^{(1/x)} dx$ , one can conclude.

*Part 2:* We now assume that (\*\*) is true. In the first part, we proved implicitly that (3.3) is equivalent to the (\*\*). Hence we will use (3.3) to prove (3.2) which is equivalent to condition (1.4) and only the first convergence in (3.2) is needed to be proved. Let  $M$  be a positive number and  $M/n \leq 1$ ,  $\mu^{(n)} \neq 0$ , then

$$\begin{aligned} \frac{\int_{1/n}^1 x^{-2} \Lambda(dx)}{n\mu^{(n)}} &= \frac{\int_{M/n}^1 x^{-2} \Lambda(dx)}{n\mu^{(n)}} + \frac{\int_{1/n}^{M/n} x^{-2} \Lambda(dx)}{n\mu^{(n)}} \\ &\leq \frac{1}{M} + 1 - \frac{\mu^{(n/M)}}{\mu^{(n)}}. \end{aligned}$$

The first term can be made as small as we want by taking  $M$  large, and the third term

$$\frac{\mu^{(n/M)}}{\mu^{(n)}} = \exp\left(-\int_{1/n}^{M/n} \frac{f(x)}{x} ds\right) \frac{1 - f(M/n)}{1 - f(1/n)}.$$

Let  $\varepsilon > 0$  and  $n$  large enough such that  $f(x) \leq \varepsilon$  on  $[0, M/n]$ . Then  $\mu^{(n/M)} \times (\mu^{(n)})^{-1} \geq \exp(-\varepsilon \ln M)(1 - \varepsilon)$ , which can be made as close as possible to 1 with  $\varepsilon$  small enough. Hence we can conclude.  $\square$

The next corollary is immediate.

**Corollary 3.1.** *If  $\Lambda$  satisfies (1.4), then*

- $\lim_{n \rightarrow +\infty} \frac{(\mu^{(n)})^k}{n} = 0, \quad \forall k > 0;$
- $\lim_{n \rightarrow +\infty} \frac{\mu^{(n)}}{\mu^{(n-M)}} = 1, \quad \forall M > 0;$
- $\lim_{n \rightarrow +\infty} \frac{\mu^{(n)}}{\mu^{(n\varepsilon)}} = 1, \quad \forall 0 < \varepsilon < 1.$

### 3.2. Properties of $\Pi^{(1,n)}$

We should next estimate the coalescent process related to the noise measure  $\Lambda_1$  which serves as a perturbation to the main measure  $\Lambda_2$ . At first, one needs a technical result.

**Lemma 3.1.** *We assume that  $\Lambda(\{0\}) = 0$ . Let  $g_n^{(1)} = \int_0^1 (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda_1(dx)$  in the spirit of (1.3). Then there exists a positive constant  $C_1$  such that for  $n$  large enough*

$$g_n^{(1)} \geq C_1 n^2 \int_0^{1/n} \Lambda_1(dx). \quad (3.6)$$

*Proof.* Let  $M > 2$ . We write

$$\begin{aligned} g_n^{(1)} &= \int_0^1 (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda_1(dx) \\ &= \int_0^{1/n} (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda_1(dx) = I_1 + I_2, \end{aligned}$$

where  $I_1 = \int_0^{1/(nM)} (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda_1(dx)$  and  $I_2 = \int_{1/(nM)}^{1/n} (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda_1(dx)$ . It is easy to see that for  $n \geq 2$ ,

$$\begin{aligned} I_1 &\geq \int_0^{1/(nM)} (n(n-1) - n(n-1)(n-2)x) \frac{1}{2} \Lambda_1(dx) \\ &\geq \int_0^{1/(nM)} \frac{n(n-1) - (n-1)(n-2)}{M} \frac{1}{2} \Lambda_1(dx) \geq \frac{1}{4} \int_0^{1/(nM)} n^2 \Lambda_1(dx). \end{aligned}$$

For the second term,

$$I_2 \geq \int_{1/(nM)}^{1/n} \left(1 - \left(1 - \frac{1}{nM}\right)^n - \frac{(1 - 1/(nM))^{n-1}}{M}\right) n^2 \Lambda_1(dx).$$

Notice that for  $n$  large, there exists a positive constant  $C(M)$  such that

$$1 - \left(1 - \frac{1}{nM}\right)^n - \frac{(1 - 1/(nM))^{n-1}}{M} \geq C(M) > 0.$$

Hence  $I_2 \geq C(M) \int_{1/(nM)}^{1/n} n^2 \Lambda_1(dx)$ . It suffices to take  $C_1 = \min\{1/4, C(M)\}$  to conclude.  $\square$

The following lemma estimates the coalescent process related to the noise measure  $\Lambda_1$ . Recall that  $\Pi^{(1,n)}$  is the  $\Lambda_1$ -coalescent process with  $\Pi^{(1,n)}(0) = \{\{1\}, \{2\}, \dots, \{n\}\}$ .

**Lemma 3.2.** *Assume that  $\Lambda$  satisfy (1.4). Then for any  $M > 0$ ,  $0 < \varepsilon \leq 1$  and  $n$  large enough, we have*

$$\mathbb{P}\left(\left|\Pi^{(1,n)}\left(\frac{M}{\mu^{(n)}}\right)\right| \leq n - n\varepsilon\right) = o(n^{-1}). \tag{3.7}$$

*Proof.* If  $\int_0^{1/n_0} \Lambda(dx) = 0$  with some  $n_0 > 1$ , then for any  $n > n_0$ ,  $\Lambda_1$  is the null measure and hence  $|\Pi^{(1,n)}(t)| = n$  for any  $t \geq 0$ , which proves this lemma. In consequence, one needs only to consider the case where  $\int_0^{1/n} \Lambda(dx) \neq 0$  for any  $n \geq 1$ . We recall  $g_n^{(1)}$  defined in Lemma 3.1. Let  $X_1^{(1,n)}$  be the decrease of the number of blocks at the first coalescence of  $\Pi^{(1,n)}$ . Thanks to Proposition 2.1 where we pick up the notations,

$$n - \sum_{i=1}^{N(\Lambda_1, n, M/\mu^{(n)})} W_i^{(n)} \leq \left|\Pi^{(1,n)}\left(\frac{M}{\mu^{(n)}}\right)\right|,$$

where  $N(\Lambda_1, n, M/\mu^{(n)})$  is Poisson distributed with parameter  $Mg_n^{(1)}/\mu^{(n)}$  independent of  $(W_i^{(n)})_{i \geq 1}$  which are i.i.d. copies of  $X_1^{(1,n)}$ . Then we have, for  $n$  large,

$$\begin{aligned} &\mathbb{P}\left(\left|\Pi^{(1,n)}\left(\frac{M}{\mu^{(n)}}\right)\right| \leq n - n\varepsilon\right) \\ &\leq \mathbb{P}\left(n - \sum_{i=1}^{N(\Lambda_1, n, M/\mu^{(n)})} W_i^{(n)} \leq n - n\varepsilon\right) \end{aligned} \tag{3.8}$$

$$\begin{aligned}
 &= \mathbb{P}\left(\sum_{i=1}^{N(\Lambda_1, n, M/\mu^{(n)})} W_i^{(n)} - \frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}] \geq n\varepsilon - \frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}]\right) \\
 &\leq \left(n\varepsilon - \frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}]\right)^{-2} \text{Var}\left(\sum_{i=1}^{N(\Lambda_1, n, M/\mu^{(n)})} W_i^{(n)}\right) \\
 &= \left(n\varepsilon - \frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}]\right)^{-2} \frac{Mg_n^{(1)}}{\mu^{(n)}}\mathbb{E}[(W_1^{(n)})^2],
 \end{aligned}$$

where the second inequality needs  $n\varepsilon - (g_n^{(1)}M/\mu^{(n)})\mathbb{E}[W_1^{(n)}] > 0$  which is justified by the following calculations: Notice that due to Proposition 2.1 and Lemma 3.1, for  $n$  large enough,

$$\begin{aligned}
 \mathbb{E}[W_1^{(n)}] + 1 &\leq \frac{n(n-1) \int_0^{1/n} \Lambda_1(dx)}{g_n^{(1)}} \leq \frac{1}{C_1}; \\
 \mathbb{E}[(W_1^{(n)})^2] &\leq \frac{n(n-1) \int_0^{1/n} \Lambda_1(dx)}{g_n^{(1)}} \leq \frac{1}{C_1},
 \end{aligned} \tag{3.9}$$

where  $C_1$  is the positive constant in Lemma 3.1.

Notice that (1.4) gives  $g_n^{(1)}/(n\mu^{(n)}) \leq g_n/(n\mu^{(n)}) \rightarrow 0$ . Then together with (3.9), we have

$$\frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}] = o(n), \quad \frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[(W_1^{(n)})^2] = o(n).$$

Hence  $n\varepsilon - (g_n^{(1)}M/\mu^{(n)})\mathbb{E}[W_1^{(n)}] \asymp n\varepsilon$ . So the inequality (3.8) is justified and one deduces that

$$\mathbb{P}\left(\left|\Pi^{(\Lambda_1, n)}\left(\frac{M}{\mu^{(n)}}\right)\right| \leq n - n\varepsilon\right) = o(n^{-1}).$$

Then we conclude (3.7). □

### 3.3. Asymptotics of $P_t^{(2,m)}, P_{1,2}^{(n,m)}(t), P_{1,2}^{(n,m,k)}(t), 2 \leq m \leq n, t \geq 0$

These terms are probabilities defined in Section 2.3.1, which measure the possibility to make one or several singletons coalesced in their *first marking times* within  $[0, t]$ . In fact, we will study  $P_{t/\mu^{(n)}}^{(2,m)}, P_{1,2}^{(n,m)}(t/\mu^{(n)}), P_{1,2}^{(n,m,k)}(t/\mu^{(n)})$ , since we want to prove that the normalization factor of the external branch length is  $\mu^{(n)}$ . We denote by “ $\ll$ ” the stochastic dominance between two real random variables. The following corollary together with the remark at the end play an important role in getting the asymptotics of the three probabilities.

**Proposition 3.2.** *Suppose that  $\Lambda$  satisfies (1.4) and define*

$$P^{(2,n)} := \lim_{t \rightarrow \infty} P_t^{(2,n)} = \sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)} (1 - (1 - \Delta_i^{(2)})^{n-1})].$$

Then

$$\lim_{n \rightarrow +\infty} P^{(2,n)} = 1. \quad (3.10)$$

*Proof.* Recall  $(\eta_i^{(2)})_{i \geq 1}$ ,  $(e_i^{(2)})_{i \geq 1}$ ,  $\{\Delta_i^{(2)}\}_{i \geq 1}$  which are associated to  $\Lambda_2$  and defined in Section 2.3. At first, we remark that  $\sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)}] = 1$ . One only needs to prove that  $\lim_{n \rightarrow +\infty} \sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n-1}] = 0$ . It is easy to see that

$$\mathbb{E}[\Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n-1}] = \mathbb{E}[\bar{\Delta}_i^{(2)} (1 - \bar{\Delta}_i^{(2)})^{n-1}],$$

where  $\bar{\Delta}_i^{(2)} = \eta_1^{(2)} \prod_{j=2}^i (1 - \eta_j^{(2)})$ . It is obvious that  $(\bar{\Delta}_i^{(2)})_{i \geq 1}$  is a Markov chain. For  $s > 0$ , we define a stopping time

$$\begin{aligned} \tau_s &= \min \{i \mid \bar{\Delta}_i^{(2)} \leq 1/s\} = \min \left\{ i \mid - \sum_{j=2}^i \ln(1 - \eta_j^{(2)}) \geq \ln s \eta_1^{(2)} \right\} \\ &= \min \left\{ i + 1 \mid - \sum_{j=1}^i \ln(1 - \eta_{j+1}^{(2)}) \geq \ln s \eta_1^{(2)} \right\}. \end{aligned}$$

Then we get

$$\begin{aligned} \sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n-1}] & \quad (3.11) \\ &= \mathbb{E} \left[ \sum_{i=1}^{+\infty} \bar{\Delta}_i^{(2)} (1 - \bar{\Delta}_i^{(2)})^{n-1} \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^{\tau_n-1} \bar{\Delta}_i^{(2)} (1 - \bar{\Delta}_i^{(2)})^{n-1} + \sum_{i=\tau_n}^{+\infty} \bar{\Delta}_i^{(2)} (1 - \bar{\Delta}_i^{(2)})^{n-1} \right]. \end{aligned}$$

Notice that  $x(1-x)^{n-1} \leq 1/n$ , if  $1/n \leq x \leq 1$  and  $x(1-x)^{n-1} \leq x$ , if  $0 \leq x \leq 1/n$ . Then (3.11) gives

$$\sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n-1}] \leq \mathbb{E} \left[ \frac{\tau_n - 1}{n} + \sum_{i=\tau_n}^{+\infty} \bar{\Delta}_i^{(2)} \right] \leq \mathbb{E} \left[ \frac{\tau_n - 1}{n} \right] + \frac{1}{\mathbb{E}[n\eta_1^{(2)}]}. \quad (3.12)$$

To calculate  $\mathbb{E}[\tau_n]$ , we use renewal theory. Let  $\mu = \mathbb{E}[-\ln(1 - \eta_1^{(2)})]$ . Depending on whether  $\mu$  is finite or not, we separate the discussion into two parts.

*Part 1:* Assume that  $\mu < +\infty$ . We denote by  $F(t)$  the distribution function and  $f(t)$  the density function of  $-\ln(1 - \eta_1^{(2)})$  and  $X$  an independent random variable with density function  $\mu^{-1}(1 - F(t))\mathbf{1}\{t \geq 0\}$ . Notice that  $X$  depends on  $n$ . Let  $\varepsilon > 0$ , then using integration by parts,

$$\mathbb{P}(0 \leq X \leq \varepsilon) = \int_0^\varepsilon \frac{1 - F(t)}{\mu} dt = \frac{\varepsilon(1 - F(\varepsilon))}{\mu} + \frac{\int_0^\varepsilon tf(t) dt}{\mu} \geq \frac{\int_0^\varepsilon tf(t) dt}{\mu}. \tag{3.13}$$

One can write  $\int_0^\varepsilon tf(t) dt$  in another way

$$\int_0^\varepsilon tf(t) dt = \frac{\int_{1/n}^{1-\exp(-\varepsilon)} -\ln(1-x)x^{-2} \Lambda(dx)}{\int_{1/n}^1 x^{-2} \Lambda(dx)}.$$

Notice that  $\mu = (\int_{1/n}^1 x^{-2} \Lambda(dx))^{-1} \int_{1/n}^1 -\ln(1-x)x^{-2} \Lambda(dx) < +\infty$ , then for  $n$  large enough, there must exist a large number  $\varepsilon_0 > 0$  such that for any  $\varepsilon \geq \varepsilon_0$ ,

$$\int_0^\varepsilon tf(t) dt \geq \frac{1}{2} \frac{\int_{1/n}^1 -\ln(1-x)x^{-2} \Lambda(dx)}{\int_{1/n}^1 x^{-2} \Lambda(dx)} = \frac{\mu}{2}.$$

Now together with (3.13), one gets that for  $n$  large enough

$$\mathbb{P}(0 \leq X \leq \varepsilon) \geq 1/2, \quad \forall \varepsilon \geq \varepsilon_0. \tag{3.14}$$

We fix  $\varepsilon \geq \varepsilon_0$  and define a new Markov chain  $(X - \sum_{j=2}^i \ln(1 - \eta_j^{(2)}))_{i \geq 1}$  and a stopping time  $\tau'_s = \min \{i \mid X - \sum_{j=1}^i \ln(1 - \eta_{j+1}^{(2)}) \geq \ln s\}$  for  $s > 0$ . It is clear from the definitions of  $\tau_s$  and  $\tau'_s$  that

$$\mathbb{E}[\tau'_{s\eta_1^{(2)}} \mid X = \varepsilon] = \mathbb{E}[\tau_{s \exp(-\varepsilon)} - 1].$$

Then

$$\begin{aligned} \mathbb{E}[\tau'_{n\eta_1^{(2)}}] &= \mathbb{E}[\tau'_{n\eta_1^{(2)}} \mathbf{1}\{0 \leq X \leq \varepsilon\}] + \mathbb{E}[\tau'_{n\eta_1^{(2)}} \mathbf{1}\{X > \varepsilon\}] \\ &\geq \mathbb{P}(0 \leq X \leq \varepsilon) \mathbb{E}[\tau_{n \exp(-\varepsilon)} - 1] + \mathbb{E}[\tau'_{n\eta_1^{(2)}} \mathbf{1}\{X > \varepsilon\}], \end{aligned}$$

which implies that

$$\mathbb{E}[\tau_{n \exp(-\varepsilon)}] \leq \frac{\mathbb{E}[\tau'_{n\eta_1^{(2)}}]}{\mathbb{P}(0 \leq X \leq \varepsilon)} + 1. \tag{3.15}$$

Due to (4.4) and (4.6) in [16, p. 369], we have

$$\mathbb{E}[\tau'_s] = \frac{\ln s}{\mu}, \quad \forall s \geq 1.$$

Notice that  $\eta_1^{(2)} \geq 1/n$ , hence  $n\eta_1^{(2)} \geq 1$ . Therefore, (3.15) gives

$$\mathbb{E}[\tau_n] \leq \frac{\mathbb{E}[\tau'_{n \exp(\varepsilon) \eta_1^{(2)}}]}{\mathbb{P}(0 \leq X \leq \varepsilon)} + 1 = \frac{\mathbb{E}[\ln(n \exp(\varepsilon) \eta^{(2)})]}{\mu \mathbb{P}(0 \leq X \leq \varepsilon)} + 1. \quad (3.16)$$

For any  $0 \leq x < 1$ , we have  $-\ln(1-x) \geq x$ , hence  $\mu \geq \mathbb{E}[\eta_1^{(2)}]$ . Then (3.16) implies

$$\frac{\mathbb{E}[\tau_n]}{n} \leq \frac{\mathbb{E}[\ln n \eta_1^{(2)}] + \varepsilon}{\mathbb{E}[n \eta_1^{(2)}] \mathbb{P}(0 \leq X \leq \varepsilon)} + \frac{1}{n}. \quad (3.17)$$

Using (3.14) and (3.12), it suffices to prove that:

$$\lim_{n \rightarrow +\infty} \mathbb{E}[n \eta_1^{(2)}] = +\infty, \quad \text{and} \quad \lim_{n \rightarrow +\infty} \frac{\mathbb{E}[\ln(n \eta_1^{(2)})]}{\mathbb{E}[n \eta_1^{(2)}]} = 0.$$

It is easy to see that, using (1.3), there exists a positive constant  $C_2$  such that

$$\mathbb{E}[n \eta_1^{(2)}] = \frac{n \int_{1/n}^1 x^{-1} \Lambda(dx)}{\bar{\mu}^{(n)}} \geq C_2 \frac{n \mu^{(n)}}{g_n},$$

for any  $n \geq 3$ . Hence  $\mathbb{E}[n \eta_1^{(2)}]$  tends to  $+\infty$  since  $\Lambda$  satisfies (1.4). For the second convergence, we fix  $M > e$ . Then,

$$\begin{aligned} \frac{\mathbb{E}[\ln(n \eta_1^{(2)})]}{\mathbb{E}[n \eta_1^{(2)}]} &= \frac{\mathbb{E}[\ln(n \eta_1^{(2)}) \mathbf{1}\{n \eta_1^{(2)} \geq M\}] + \mathbb{E}[\ln(n \eta_1^{(2)}) \mathbf{1}\{n \eta_1^{(2)} < M\}]}{\mathbb{E}[n \eta_1^{(2)}]} \\ &\leq \frac{\mathbb{E}[\ln(n \eta_1^{(2)}) \mathbf{1}\{n \eta_1^{(2)} \geq M\}]}{\mathbb{E}[n \eta_1^{(2)}]} + \frac{\ln M}{\mathbb{E}[n \eta_1^{(2)}]} \\ &\leq \frac{\mathbb{E}[\ln(n \eta_1^{(2)}) \mathbf{1}\{n \eta_1^{(2)} \geq M\}]}{\mathbb{E}[n \eta_1^{(2)} \mathbf{1}\{n \eta_1^{(2)} \geq M\}]} + \frac{\ln M}{\mathbb{E}[n \eta_1^{(2)}]} \\ &\leq \frac{\ln M}{M} + \frac{\ln M}{\mathbb{E}[n \eta_1^{(2)}]}. \end{aligned}$$

The last inequality is due to the fact that for any  $x \geq M > e$ , we have  $\ln x/x \leq \ln M/M$ . Since  $M$  can be chosen as large as we want,  $\lim_{n \rightarrow +\infty} (\mathbb{E}[n \eta_1^{(2)}])^{-1} \times \mathbb{E}[\ln(n \eta_1^{(2)})] = 0$ . Hence we can conclude.

Part 2: If  $\mu = +\infty$ . We define

$$(\bar{\eta}_i^{(2)})_{i \geq 2} := \left( \frac{1}{2} \mathbf{1}\left\{ \eta_i^{(2)} \geq \frac{1}{2} \right\} + \eta_i^{(2)} \mathbf{1}\left\{ \eta_i^{(2)} < \frac{1}{2} \right\} \right)_{i \geq 2}$$



and for  $s > 0$ ,  $\bar{\tau}_s := \min \{i + 1 \mid \sum_{j=1}^i -\ln(1 - \bar{\eta}_{j+1}^{(2)}) \geq \ln s \eta_1^{(2)}\}$ . Notice that  $\mathbb{E}[-\ln(1 - \bar{\eta}_i^{(2)})] < +\infty$ , then we return to the first case and get (3.17) by replacing  $\tau_n$  by  $\bar{\tau}_n$  and keeping the same  $\eta_1^{(2)}$  but with different  $X$  (depending on  $\bar{\eta}_i^{(2)}, i \geq 2$ ). In this setting,  $\mathbb{P}(0 \leq X \leq \ln 2) = 1$ . We see that the closer  $\bar{\eta}_i^{(2)}$  is to 1, the larger the  $-\ln(1 - \bar{\eta}_i^{(2)})$  and hence  $\tau_n \ll \bar{\tau}_n$ . Then we can conclude.  $\square$

*Remark 3.1.* For  $0 < \varepsilon < 1$ , we also have

$$\lim_{n \rightarrow +\infty} \sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n(1-\varepsilon)}] = 0. \tag{3.18}$$

The proof is all the same. The only thing different is that in place of (3.12), we have  $\sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n(1-\varepsilon)}] \leq C \mathbb{E}[(\tau_n - 1)/n + \sum_{i=\tau_n}^{+\infty} \text{bar} \Delta_i^{(2)}]$ , with  $C$  larger than 1 and depends on  $\varepsilon$ .

Now we can start to study at first  $P_{t/\mu^{(n)}}^{(2,n)}$ .

**Corollary 3.2.**

$$\lim_{t \rightarrow +\infty} \liminf_{n \rightarrow +\infty} P_{t/\mu^{(n)}}^{(2,n)} = 1. \tag{3.19}$$

*Proof.* Recall that  $\{e_i^{(2)}\}_{i \geq n}$  are i.i.d. exponential variables with parameter  $\bar{\mu}^{(n)} = \int_0^1 x^{-2} \Lambda_2(dx)$ , as defined in Section 2.3. Let  $\tau_n(t) = \max \{j : \sum_{i=1}^j e_i^{(2)} \leq t/\mu^{(n)}\}$ . Then

$$P_{t/\mu^{(n)}}^{(2,n)} = \mathbb{E} \left[ \sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} - \sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n-1} \right]. \tag{3.20}$$

Due to Proposition 3.2, we have

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[ \sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n-1} \right] \leq \lim_{n \rightarrow +\infty} \mathbb{E} \left[ \sum_{i=1}^{+\infty} \Delta_i^{(2)} (1 - \Delta_i^{(2)})^{n-1} \right] = 0.$$

Then it suffices to prove that

$$\lim_{t \rightarrow +\infty} \liminf_{n \rightarrow +\infty} \mathbb{E} \left[ \sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} \right] = 1. \tag{3.21}$$

Let  $\mathcal{E}_j = \bar{\mu}^{(n)} \sum_{i=1}^j e_i^{(2)}$ , which is the sum of  $j$  i.i.d. unit exponential variables. Let  $I_n = \bar{\mu}^{(n)}/\mu^{(n)}$ . Then

$$\tau_n(t) = \max \{j : \mathcal{E}_j \leq t I_n\}.$$

For any fixed  $0 < \beta < 1$ ,

$$\begin{aligned} & \mathbb{P}(\tau_n(t) \in [0, \beta t I_n] \cup (t I_n / \beta, +\infty)) \\ &= \mathbb{P}(\mathcal{E}_{\lceil \beta t I_n \rceil} \geq t I_n) + \mathbb{P}(\mathcal{E}_{\lfloor t I_n / \beta \rfloor} \leq t I_n) = o((t I_n)^{-1}), \end{aligned} \quad (3.22)$$

where the last equality is a large deviation result (for example, see Theorem 1.4 of [10]). Notice that  $\tau_n(t)$  is independent of  $\{\Delta_i^{(2)}\}_{i \geq 1}$ , so

$$\begin{aligned} \mathbb{E}\left[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)}\right] &= \mathbb{E}\left[1 - \left(1 - \frac{1}{I_n}\right)^{\tau_n(t)+1}\right] \\ &= \mathbb{E}\left[1 - \left(1 - \frac{1}{I_n}\right)^{\tau_n(t)+1} \mathbf{1}\left\{t I_n \beta \leq \tau_n(t) \leq t \frac{I_n}{\beta}\right\}\right] + o((t I_n)^{-1}) \\ &\geq \mathbb{E}\left[1 - \left(1 - \frac{1}{I_n}\right)^{t I_n \beta} \mathbf{1}\left\{t I_n \beta \leq \tau_n(t) \leq t \frac{I_n}{\beta}\right\}\right] + o((t I_n)^{-1}). \end{aligned}$$

Notice that  $I_n \geq 1$  and the term at the right of the above inequality satisfies

$$\lim_{t \rightarrow +\infty} \liminf_{n \rightarrow +\infty} \mathbb{E}\left[1 - \left(1 - \frac{1}{I_n}\right)^{t I_n \beta} \mathbf{1}\left\{t I_n \beta \leq \tau_n(t) \leq t \frac{I_n}{\beta}\right\}\right] + o((t I_n)^{-1}) = 1.$$

Then we can conclude (3.21). □

*Remark 3.2.* For  $0 < \varepsilon < 1$ , we also have

$$\lim_{t \rightarrow +\infty} \liminf_{n \rightarrow +\infty} P_{t/\mu^{(n)}}^{(2, \lceil n - n\varepsilon \rceil)} = 1. \quad (3.23)$$

To prove this, in the proof of this corollary, one should replace (3.20) by

$$P_{t/\mu^{(n)}}^{(2, n)} = \mathbb{E}\left[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} - \sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} (1 - \Delta_i^{(2)})^{\lceil n - n\varepsilon \rceil - 1}\right].$$

The first term satisfies (3.21). For the second term, using (3.18), we get

$$\lim_{n \rightarrow +\infty} \mathbb{E}\left[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} (1 - \Delta_i^{(2)})^{\lceil n - n\varepsilon \rceil - 1}\right] = 0.$$

Then (3.23) is proved.

The next corollary is straightforward using (2.1), (2.2) and (3.23).

**Corollary 3.3.** For any  $0 < \varepsilon < 1$ ,

$$\lim_{t \rightarrow +\infty} \liminf_{n \rightarrow +\infty} P_{1,2}^{(n, \lceil n - n\varepsilon \rceil, k)}\left(\frac{t}{\mu^{(n)}}\right) = 1, \quad \lim_{t \rightarrow +\infty} \liminf_{n \rightarrow +\infty} P_{1,2}^{(n, \lceil n - n\varepsilon \rceil)}\left(\frac{t}{\mu^{(n)}}\right) = 1.$$

**3.4. Proofs of main results**

*Proof of Theorem 1.1.* Fix  $t > 0$  and  $0 < \varepsilon < 1$ . Considering the measure division construction for two-type  $\Lambda$ -coalescents, let  $\Pi$  be the path of  $\Pi^{(1,n)}$  chosen at the step 0 and define the event

$$E' := \left\{ \left| \Pi\left(\frac{t}{\mu^{(n)}}\right) \right| \geq n - n\varepsilon \right\} \cap \left\{ \{1\} \in \Pi\left(\frac{t}{\mu^{(n)}}\right) \right\}.$$

Recall that  $\left\{ \left| \Pi^{(1,n)}(t/\mu^{(n)}) \right| \geq n - n\varepsilon \right\}$  implies that there are at least  $n - \lceil 2n\varepsilon \rceil$  singletons at time  $t/\mu^{(n)}$ . For  $n$  large enough, using the exchangeability property, we have  $\mathbb{P}(E') \geq (n - \lceil 2n\varepsilon \rceil)n^{-1}(1 - \kappa_n(t))$ , where  $\kappa_n(t) = \mathbb{P}(|\Pi^{(1,n)}(t/\mu^{(n)})| < n - n\varepsilon)$  and  $\kappa_n(t) = o(n^{-1})$  due to (3.7). For  $\varepsilon$  small enough and  $n$  large enough, we have  $\mathbb{P}(E')$  as close as we want to 1. We define another event

$$E'' := \left\{ \{1\} \text{ is coalesced at its first marking time within } [0, t) \right\}$$

Then due to (2.1) and  $P_t^{(2,m)}$  is increasing on  $m$ , we get

$$\mathbb{P}(E'' \mid E') \geq P_{t/\mu^{(n)}}^{(2, \lceil n - n\varepsilon \rceil)}. \tag{3.24}$$

Let  $0 < t_1 < t$ ,

$$\begin{aligned} \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}\right) &= \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, E' \cap E''\right) + \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, (E' \cap E'')^c\right) \\ &= \mathbb{P}\left(L_1^{(2,n)} \geq \frac{t_1}{\mu^{(n)}}, E' \cap E''\right) + \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, (E' \cap E'')^c\right). \end{aligned} \tag{3.25}$$

Corollary 2.2 tells that  $\mathbb{P}(L_1^{(2,n)} \geq t_1/\mu^{(n)} \mid E') = \exp(-t_1)$  and it has been proved that  $\mathbb{P}(E' \cap E'') = \mathbb{P}(E')\mathbb{P}(E'' \mid E')$  can be made as close as possible to 1 by taking  $\varepsilon$  small enough and  $t$  large enough and  $n$  tending to  $+\infty$ . Hence the first term of (3.25) can be made as close as we want to  $\exp(-t_1)$  and the second term is close to 0. Then we can conclude.  $\square$

*Proof of Theorem 1.2.* We prove instead for  $k \in \mathbb{N}$ :

$$\mu^{(n)}(T_1^{(n)}, T_2^{(n)}, \dots, T_k^{(n)}) \xrightarrow{(d)} (e_1, e_2, \dots, e_k), \tag{3.26}$$

which is equivalent to (1.6) (see [3, p.19]). We will give the proof for  $k = 2$  and leave the easy extension to readers. The proof is similar to that of Theorem 1.1. Let  $\Pi$  be the path of  $\Pi^{(1,n)}$  chosen at step 0. Let  $t > 0$ ,  $0 < \varepsilon < 1$  and define the event

$$F' := \left\{ \left| \Pi\left(\frac{t}{\mu^{(n)}}\right) \right| \geq n - n\varepsilon \right\} \cap \left\{ \{1\}, \{2\} \in \Pi\left(\frac{t}{\mu^{(n)}}\right) \right\}.$$

Using the same arguments, we get

$$\mathbb{P}(F') \geq \frac{\binom{n - \lceil 2n\varepsilon \rceil}{2}}{\binom{n}{2}} (1 - \kappa_n(t)).$$

We then define the event

$$F'' := \{\{1\}, \{2\} \text{ are both coalesced at their first marking times within } [0, t).\}$$

Then due to (2.2) and  $P_t^{(2,m)}$  is increasing on  $m$ , we get  $\mathbb{P}(F'' \mid F') \geq 1 - 2(1 - P_{t/\mu^{(n)}}^{(2, \lceil n - n\varepsilon \rceil)})$ , which is as close as possible to 1 for  $t$  large and  $n$  tending to  $+\infty$ .

Let  $0 \leq t_1, t_2 \leq t$ . Then

$$\begin{aligned} \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}\right) &= \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}, F' \cap F''\right) \\ &\quad + \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}, (F' \cap F'')^c\right) \\ &= \mathbb{P}\left(L_1^{(2,n)} \geq \frac{t_1}{\mu^{(n)}}, L_2^{(2,n)} \geq \frac{t_2}{\mu^{(n)}}, F' \cap F''\right) \\ &\quad + \mathbb{P}\left(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}, (F' \cap F'')^c\right). \end{aligned} \tag{3.27}$$

As shown that  $\mathbb{P}(F' \cap F'')$  can be made as close as possible to 1 by taking  $t$  large enough and  $\varepsilon$  small enough, tending  $n$  to  $+\infty$ . Then the second term in (3.27) is close to 0. Using Corollary 2.2, the first term is as close as possible to  $\exp(-t_1 - t_2)$  by tending  $n$  to  $+\infty$  with  $t$  large enough. Then we can conclude.  $\square$

*Proof of Corollary 1.1.* We prove at first the case of one external branch length. We seek to prove the uniform integrability of  $\{(\mu^{(n)} T_1^{(n)})^k, n \geq 2\}$  for any  $k \geq 0$ . One needs only to show that for any fixed  $k \in \mathbb{N}$ ,  $\sup\{\mathbb{E}[(\mu^{(n)} T_1^{(n)})^k] \mid n \geq 2\} < +\infty$  (see Lemma 4.11 of [22] and Problem 14 in Section 8.3 of [7]). Let  $M > 0$ ,  $0 < \varepsilon < 1$ ,  $\beta_n = \lfloor \Pi^{(n)}(M/\mu^{(n)}) \rfloor$  and  $n_0 = \min\{i \mid \mu^{(i)} > 0\}$ . To avoid invalid calculations, we set  $\mu^{(n)} = 1$  if  $n < n_0$ . Using the Markov property, we have

$$T_1^{(n)} \ll \frac{M}{\mu^{(n)}} + \bar{T}_1^{(\beta_n)} \mathbf{1}\left\{T_1^{(n)} \geq \frac{M}{\mu^{(n)}}\right\},$$

where  $\bar{T}_1^{(n)} \stackrel{(d)}{=} T_1^{(n)}$ ,  $n \geq 2$  and conditional on  $\beta_n$ ,  $\bar{T}_1^{(\beta_n)}$  is independent of  $\mathbf{1}\{T_1^{(n)} \geq M/\mu^{(n)}\}$ . Then for  $n\varepsilon \geq n_0$ ,

$$\begin{aligned} &\mathbb{E}[(\mu^{(n)} T_1^{(n)})^k] \\ &\leq \mathbb{E}\left[\left(M + \mu^{(n)} \bar{T}_1^{(\beta_n)} \mathbf{1}\{\mu^{(n)} T_1^{(n)} > M\}\right)^k\right] \end{aligned} \tag{3.28}$$

$$\begin{aligned}
 &\leq (2M)^k + \mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(\beta_n)}\mathbf{1}\{\mu^{(n)}T_1^{(n)} > M\})^k] \\
 &\leq (2M)^k + (\mathbb{E}[2\mu^{(n)}\bar{T}_1^{(n)}\mathbf{1}\{\beta_n = n\}])^k \\
 &\quad + \mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(\beta_n)}\mathbf{1}\{\mu^{(n)}T_1^{(n)} > M, n\varepsilon \leq \beta_n \leq n-1\})^k] \\
 &\quad + \mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(\beta_n)}\mathbf{1}\{\mu^{(n)}T_1^{(n)} > M, \beta_n < n\varepsilon\})^k] \\
 &\leq (2M)^k + \exp\left(-\frac{Mg_n}{\mu^{(n)}}\right)\mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(n)})^k] \\
 &\quad + \mathbb{P}(\mu^{(n)}T_1^{(n)} > M)\left(2\frac{\mu^{(n)}}{\mu^{(n\varepsilon)}}\right)^k \max\{\mathbb{E}[(\mu^{(j)}\bar{T}_1^{(j)})^k] \mid j \in [n\varepsilon, n-1]\} \\
 &\quad + \mathbb{P}(\beta_n < n\varepsilon)\mathbb{E}\left[\frac{\beta_n}{n}\left(2\frac{\mu^{(n)}}{\mu^{(\beta_n)}}\right)^k (\mu^{(\beta_n)}\bar{T}_1^{(\beta_n)})^k \mid \beta_n < n\varepsilon\right],
 \end{aligned}$$

where  $\exp(-Mg_n/\mu^{(n)})$  in the second term at right of the last inequality is the probability for no coalescence within  $[0, M/\mu^{(n)}]$ . The third term is due to the fact that  $\mu^{(n)}$  is an increasing function of  $n$  when  $n \geq n_0$ . The fourth term is due to exchangeability which says that the probability for  $\{1\}$  not to have coalesced at  $M/\mu^{(n)}$  when there exist only  $\beta_n$  blocks is less than  $\beta_n/n$ . One needs the following three estimates to prove the boundedness of  $(\mathbb{E}[(\mu^{(n)}T_1^{(n)})^k])_{n \geq 2}$ .

- Estimation of  $\exp(-Mg_n/\mu^{(n)})2^k$ : Notice that for  $n \geq n_0$ ,

$$\begin{aligned}
 \frac{g_n}{\mu^{(n)}} &= \frac{\int_0^1 (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda(dx)}{\int_{1/n}^1 x^{-1} \Lambda(dx)} \\
 &\geq \frac{\int_{1/n}^1 (1 - (1-x)^n - nx(1-x)^{n-1})x^{-2} \Lambda(dx)}{\int_{1/n}^1 x^{-1} \Lambda(dx)} \geq \frac{e-2}{e}.
 \end{aligned}$$

And if  $2 \leq n < n_0$ , we have  $\exp(-Mg_n/\mu^{(n)}) = \exp(-Mg_n) \xrightarrow{M \rightarrow +\infty} 0$ . Hence if  $M$  is large enough, we have, for any  $n \geq 2$ ,

$$\exp\left(-\frac{Mg_n}{\mu^{(n)}}\right)2^k \leq \frac{1}{4}. \tag{3.29}$$

- Estimation of  $\mathbb{P}(\mu^{(n)}T_1^{(n)} > M)(2\mu^{(n)}/\mu^{(n\varepsilon)})^k$ : Due to Corollary 3.1, we get  $\lim_{n \rightarrow +\infty} \mu^{(n)}/\mu^{(n\varepsilon)} = 1$ , and Theorem 1.1 gives  $\lim_{n \rightarrow +\infty} \mathbb{P}(\mu^{(n)} \times T_1^{(n)} > M) = \exp(-M)$ . Hence by taking  $M$  large enough, we have for any  $n \geq 2$ ,

$$\mathbb{P}(\mu^{(n)}T_1^{(n)} > M)\left(2\frac{\mu^{(n)}}{\mu^{(n\varepsilon)}}\right)^k \leq \frac{1}{4}. \tag{3.30}$$

- Estimation of  $\beta_n n^{-1} (2\mu^{(n)}/\mu^{(\beta_n)})^k$ ,  $\beta_n < n\varepsilon$ : Using the notations in Proposition 3.1, for  $\beta_n \geq n_0$ , we have

$$\frac{\mu^{(n)}}{\mu^{(\beta_n)}} = \exp\left(\int_{1/n}^{1/\beta_n} \frac{f(x)}{x} dx\right) \frac{1 - f(1/n)}{1 - f(1/\beta_n)}. \tag{3.31}$$

Let  $n_1 > n_0$  such that for any  $n \geq n_1$ , we have  $f(1/n) \leq 1/2k$ . Hence for any  $a, b \geq n_1$ ,  $(1 - f(a))/(1 - f(b)) \leq 2$ . This  $n_1$  can be found since  $f(1/n)$  tends to 0 as  $n$  tends to  $+\infty$ . Thus (3.31) implies, for  $\beta_n \geq n_1$ ,

$$\frac{\mu^{(n)}}{\mu^{(\beta_n)}} \leq 2\left(\frac{n}{\beta_n}\right)^{1/2k}.$$

Hence if  $n_1 \leq \beta_n < n\varepsilon$  and  $\varepsilon \leq 4^{-2k-2}$ ,

$$\frac{\beta_n}{n} \left(2 \frac{\mu^{(n)}}{\mu^{(\beta_n)}}\right)^k \leq 4^k \left(\frac{\beta_n}{n}\right)^{1/2} < 4^k (\varepsilon)^{1/2} \leq \frac{1}{4}.$$

If  $\beta_n < n_1$ , due to Corollary 3.1, one could find a large number  $n_2$  such that  $n_2 > n_1$  and for any  $n \geq n_2$

$$\frac{\beta_n}{n} \left(2 \frac{\mu^{(n)}}{\mu^{(\beta_n)}}\right)^k \leq \frac{1}{4}.$$

In total, if  $n \geq n_2$  and  $\beta_n < n\varepsilon$ , then

$$\frac{\beta_n}{n} \left(2 \frac{\mu^{(n)}}{\mu^{(\beta_n)}}\right)^k \leq \frac{1}{4}. \tag{3.32}$$

Using (3.28)–(3.30) and (3.32), we get

$$\begin{aligned} \mathbb{E}[(\mu^{(n)} T_1^{(n)})^k] &\leq \frac{4}{3} (2M)^k + \frac{1}{3} \max\{\mathbb{E}[(\mu^{(j)} \bar{T}_1^{(j)})^k] \mid j \in [n\varepsilon, n-1]\} \\ &\quad + \frac{1}{3} \mathbb{E}[(\mu^{(\beta_n)} \bar{T}_1^{(\beta_n)})^k \mid \beta_n < n\varepsilon] \\ &\leq \frac{4}{3} (2M)^k + \frac{2}{3} \max\{\mathbb{E}[(\mu^{(j)} \bar{T}_1^{(j)})^k] \mid j \leq n-1\}. \end{aligned} \tag{3.33}$$

The above inequality is valid for a large  $M$ ,  $\varepsilon = 4^{-2k-2}$  and  $n \geq n_2$ . Let  $C_3 \geq \max\{\mathbb{E}[(\mu^{(j)} T_1^{(j)})^k], 4(2M)^k \mid 2 \leq j < n_2\}$ , then for any  $n \geq 2$ ,  $C_3 \geq \mathbb{E}[(\mu^{(n)} T_1^{(n)})^k]$  using (3.33). Then we can conclude.

The case of multiple external branch lengths is merely a consequence of the case of one external branch length, the Cauchy–Schwarz inequality and also a uniform integrability argument.  $\square$

*Proof of Corollary 1.2.* Notice that  $\{T_i^{(n)}\}_{1 \leq i \leq n}$  are exchangeable. Hence Corollary 1.1 shows that

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbb{E}[\mu^{(n)} L_{ext}^{(n)} / n] &= \lim_{n \rightarrow +\infty} \mathbb{E}\left[\mu^{(n)} \frac{T_1^{(n)} + T_2^{(n)} + \dots + T_n^{(n)}}{n}\right] \\ &= \lim_{n \rightarrow +\infty} \mathbb{E}[\mu^{(n)} T_1^{(n)}] = 1, \end{aligned}$$

and

$$\begin{aligned} &\lim_{n \rightarrow +\infty} \text{Var}\left(\frac{\mu^{(n)} L_{ext}^{(n)}}{n}\right) \\ &= \lim_{n \rightarrow +\infty} \frac{\mathbb{E}[n(\mu^{(n)} T_i^{(n)})^2] + n(n-1)\mathbb{E}[(\mu^{(n)})^2 T_1^{(n)} T_2^{(n)}] - n^2(\mathbb{E}[\mu^{(n)} T_1^{(n)}])^2}{n^2} \\ &= \lim_{n \rightarrow +\infty} \frac{\text{Var}(\mu^{(n)} T_1^{(n)}) + (n-1)\text{Cov}(\mu^{(n)} T_1^{(n)}, \mu^{(n)} T_2^{(n)})}{n} = 0. \end{aligned}$$

Hence  $\mu^{(n)} L_{ext}^{(n)} / n$  converges in  $L^2$  to 1. □

Before proving Corollary 1.3, we study at first a problem of sensibility of a recurrence satisfied by  $(\mathbb{E}[T_1^{(n)}])_{n \geq 2}$ . More precisely, if  $a_n = \mathbb{E}[T_1^{(n)}]$ , then  $a_n$  satisfies a recurrence (see [11]):  $a_1 = 0$ , and for  $n \geq 2$ , we have

$$a_n = c_n + \sum_{k=1}^{n-1} p_{n,k} \frac{k-1}{n} a_k, \tag{3.34}$$

where  $(c_n)_{n \geq 2} = (1/g_n)_{n \geq 2}$  and  $p_{n,k}$  is defined in (1.2). Due to Corollary 1.1, we have  $\lim_{n \rightarrow +\infty} \mu^{(n)} a_n = 1$ . The question is as follows: what is the limit behavior of  $a_n$  if we set initially the values of  $(a_i)_{1 \leq i \leq n_0}$  with  $n_0 \geq 1$  without using (3.34) and replace  $c_n$  by  $c'_n = 1/g_n + o(1/g_n)$ ? It is answered in the next lemma.

**Lemma 3.3.** *Let  $n_0 \geq 1$  and  $(a'_i)_{1 \leq i \leq n_0}$  be  $n_0$  real numbers. For  $n > n_0$ , let*

$$a'_n = c'_n + \sum_{k=1}^{n-1} p_{n,k} \frac{k-1}{n} a'_k, \tag{3.35}$$

where  $(c'_n)_{n > n_0}$  is a sequence which satisfies  $c'_n = 1/g_n + o(1/g_n)$ . Then

$$\lim_{n \rightarrow +\infty} \mu^{(n)} a'_n = 1.$$

*Proof.* We fix  $\varepsilon > 0$  and let  $n_\varepsilon > n_0$  such that  $c'_n \leq (1 + \varepsilon)/g_n$  for  $n > n_\varepsilon$ . We set  $M = \max\{|a'_i|, a_i \mid 1 \leq i \leq n_\varepsilon\}$ .

Let us at first look at (3.34) which has the following interpretation using random walk: A walker stands initially at point  $n$ , then after time  $c_n$ , he jumps to point  $k_1$  with probability  $p_{n,k_1}$ , then after time  $[(k_1 - 1)/n]c_{k_1}$ , he jumps to  $k_2$  with probability  $p_{k_1,k_2}$ , and then after time  $[(k_1 - 1)(k_2 - 1)/nk_1]c_{k_2}$ , he jumps to the next point, etc. If he falls at point 1, then this walk is finished. It is easy to see that  $a_n$  is the expectation of the total walking time. One notices that there is a scaling effect on the walking time. More precisely, let  $l \geq 1$  and  $n = k_0 > k_1 > \dots > k_l \geq 1$  such that the walker jumps from  $k_i$  to  $k_{i+1}$  for  $0 \leq i \leq l - 1$ . Then conditional on this walking history, the remaining walking time is  $(\prod_{i=0}^{l-1} (k_{i+1} - 1)/k_i) a_{k_l}$ .

The recurrence (3.35) has the same interpretation. The difference is that one should stop the walker when he arrives at a point  $i$  within  $[1, n_0]$  and one adds a scaled value of  $a'_i$  to the walking time (notice that  $a'_i$  can be negative). To estimate  $a'_n$ , we use a Markov chain  $(W_i)_{i \geq 0}$  to couple the jumping structures of (3.34) and (3.35):  $W_0 = n$ ,

- If  $W_i = k$  with  $k \geq n_\varepsilon$ , then  $W_{i+1} = k'$  with probability  $p_{k,k'}$ , where  $1 \leq k' \leq k - 1$ ;
- If  $W_i < n_\varepsilon$ , then we set  $W_j = W_i$  for any  $j \geq i + 1$ .

Notice that the jumping dynamics of both recurrences is characterized by  $(W_i)_{i \geq 0}$  until arriving at a point within  $[1, n_\varepsilon]$ . And we also see that  $(W_i)_{i \geq 0}$  is the discrete time Markov chain related to the block counting process  $|\Pi^{(n)}|$  stopped at the first time arriving within  $[1, n_\varepsilon]$ .

Let  $\varsigma_n = \min\{i \mid W_i = W_{i+1}\}$ ,  $C_{\varsigma_n} = \prod_{i=0}^{\varsigma_n-1} (W_{i+1} - 1)/W_i$  and  $T_{\varsigma_n}$  is set to be the time to  $\varsigma_n$  of the random walk related to (3.34) and  $T'_{\varsigma_n}$  be the corresponding time related to (3.35).

By the scaling effect of  $C_{\varsigma_n}$  on the walking time, we get

$$a_n = \mathbb{E}[T_{\varsigma_n} + C_{\varsigma_n} a_{W_{\varsigma_n}}], \quad a'_n = \mathbb{E}[T'_{\varsigma_n} + C_{\varsigma_n} a'_{W_{\varsigma_n}}].$$

Due to the definitions of  $M, n_\varepsilon$ , we obtain

$$\begin{aligned} a_n - M\mathbb{E}[C_{\varsigma_n}] &\leq \mathbb{E}[T_{\varsigma_n}] \leq a_n; \\ a'_n - M\mathbb{E}[C_{\varsigma_n}] &\leq \mathbb{E}[T'_{\varsigma_n}] \leq a'_n + M\mathbb{E}[C_{\varsigma_n}]; \\ \mathbb{E}[T'_{\varsigma_n}] &\leq (1 + \varepsilon)\mathbb{E}[T_{\varsigma_n}]. \end{aligned}$$

Notice that  $\mathbb{E}[C_{\varsigma_n}] \leq n_\varepsilon/n$  and due to Corollary 3.1, we have  $\lim_{n \rightarrow +\infty} M \times \mu^{(n)}/n = 0$ . Hence  $\lim_{n \rightarrow +\infty} M\mathbb{E}[C_{\varsigma_n}]\mu^{(n)} = 0$ . Then we can conclude that for  $n$  large,  $a'_n \leq (1 + 2\varepsilon)a_n$ . In the same way, we can prove also  $a'_n \geq (1 - 2\varepsilon')a_n$  for another small positive number  $\varepsilon'$  with  $n$  large enough. Then we can deduce the lemma.  $\square$



*Proof of Corollary 1.3.* Let  $b_n = \mathbb{E}[\mu^{(n)}L_{total}^{(n)}/n]$ . Then looking at the first coalescence of the process  $\Pi^{(n)}$ , we have,

$$b_1 = 0; \quad b_n = \frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k} \frac{k\mu^{(n)}}{n\mu^{(k)}} b_k, \quad n \geq 2. \quad (3.36)$$

If for some  $k$ ,  $\mu^{(k)} = 0$ , then we set  $\mu^{(k)} = 1$  to avoid invalid calculations. To use Lemma 3.3, we write (3.36) as:

$$b_1 = 0; \quad b_n = \frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} b_k + \sum_{k=1}^{n-1} p_{n,k} \frac{(k-1)\mu^{(n)}}{n\mu^{(k)}} b_k, \quad n \geq 2. \quad (3.37)$$

We at first prove that  $\sum_{k=1}^{n-1} p_{n,k} \mu^{(n)}/(n\mu^{(k)}) = o(\mu^{(n)}/g_n)$ . Indeed, due to (2.5), let  $a = \int_0^1 (1 - (1-x)^{n-1})x^{-1} \Lambda(dx)$  and  $M > 0$ , then

$$\mathbb{P}(X_1^{(n)} \geq Ma) \leq \frac{\mathbb{E}[X_1^{(n)}]}{Ma} \leq \frac{n}{Mg_n}. \quad (3.38)$$

Using Corollary 3.1, we have

$$\limsup_{n \rightarrow +\infty} \frac{a}{n} \leq \lim_{n \rightarrow +\infty} \frac{\int_0^{1/n} (n-1) \Lambda(dx) + \mu^{(n)}}{n} = 0,$$

$\lim_{n \rightarrow +\infty} \mu^{(n)}/\mu^{(n-Ma)} = 1$ . Then for  $n$  large enough,

$$\begin{aligned} \sum_{k=1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} &= \sum_{k=1}^{\lfloor n-Ma \rfloor} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} + \sum_{k=\lfloor n-Ma \rfloor + 1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} \\ &\leq \mathbb{P}(X_1^{(n)} \geq Ma) \mathbb{E}\left[\frac{\mu^{(n)}}{n\mu^{(n-X_1^{(n)})}} \mid X_1^{(n)} \geq Ma\right] + \frac{\mu^{(n)}}{\mu^{(n-Ma)}n} \\ &\leq \frac{\mu^{(n)}}{Mg_n} \max\left\{\frac{1}{\mu^{(k)}} \mid 1 \leq k \leq n\right\} + \frac{\mu^{(n)}}{\mu^{(n-Ma)}n}, \end{aligned}$$

where the first term at the right of the the last inequality is due to (3.38) and can be made as small as we want w.r.t.  $\mu^{(n)}/g_n$  when  $M$  is large enough. Notice that  $n^{-1} = o(\mu^{(n)}/g_n)$  due to (1.4). Then the second term  $\mu^{(n)}/(\mu^{(n-Ma)}n) = o(\mu^{(n)}/g_n)$  using also  $\lim_{n \rightarrow +\infty} \mu^{(n)}/\mu^{(n-Ma)} = 1$ . Then  $\sum_{k=1}^{n-1} p_{n,k} \mu^{(n)} \times (n\mu^{(k)})^{-1} = o(\mu^{(n)}/g_n)$ .

We now only need to prove that  $(b_k)_{k \geq 2}$  are bounded, since in this case,  $\sum_{k=1}^{n-1} p_{n,k} [\mu^{(n)}/(n\mu^{(k)})] b_k = o(\mu^{(n)}/g_n)$  and we apply Lemma 3.3 to (3.37). We construct another recurrence:

$$b'_1 = 0; \quad b'_n = \frac{C\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k} \frac{(k-1)\mu^{(n)}}{n\mu^{(k)}} b'_k, \quad n \geq 2. \quad (3.39)$$

where  $C$  is a positive number. If  $C = 1$ , this is exactly a transformation of the recurrence (3.34). Let  $M'(C) = \sup\{b'_n \mid n \geq 1\}$ . Then it is easy to see that  $M'(C) = CM'(1)$ . Let  $n_0 \geq 1$ , such that for  $n \geq n_0$ , we have

$$\sum_{k=1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} M'(1) \leq \frac{1}{2} \frac{\mu^{(n)}}{g_n}.$$

Then for  $C \geq 2$ ,  $n \geq n_0$ ,

$$\frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} M'(C) \leq \frac{C\mu^{(n)}}{g_n}. \quad (3.40)$$

For  $2 \leq n < n_0$ , we set  $C$  large enough such that

$$\frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} \max\{b_i \mid 1 \leq i < n_0\} \leq \frac{C\mu^{(n)}}{g_n}. \quad (3.41)$$

Comparing the coefficients and initial values of recurrences (3.37) and (3.39) using (3.40) and (3.41), we deduce that  $b_n \leq b'_n \leq M'(C)$ . Hence we can conclude.  $\square$

### Acknowledgements

The author benefited from the support of the ‘‘Agence Nationale de la Recherche’’: ANR MANEGE (ANR-09-BLAN-0215).

The author wants to thank his supervisor, Prof. Jean-Stéphane Dhersin, for fruitful discussions and also for his careful reading of a draft version resulting in many helpful comments. The author also wants to thank Prof. Martin Möhle for a discussion on Theorem 1.1.

### References

- [1] E. ARNASON (2004) Mitochondrial cytochrome b DNA variation in the high-fecundity Atlantic cod: trans-Atlantic clines and shallow gene genealogy. *Genetics* **166** (4), 1871–1885.
- [2] J. BERESTYCKI, N. BERESTYCKI AND J. SCHWEINSBERG (2008) Small-time behavior of beta coalescents. *Ann. Inst. H. Poincaré, Probabilités et Statistiques* **44** (2), 214–238.
- [3] P. BILLINGSLEY (1995) *Probability and Measure*, 3rd ed. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York.
- [4] M.G.B. BLUM AND O. FRANÇOIS (2005) Minimal clade size and external branch length under the neutral coalescent. *Adv. Appl. Probab.* **37** (3), 647–662.

- [5] E. BOLTHAUSEN AND A. SZNITMAN (1998) On ruelle's probability cascades and an abstract cavity method. *Commun. Math. Phys.* **197** (2), 247–276.
- [6] J. BOOM, E. BOULDING AND A. BECKENBACH (1994) Mitochondrial DNA variation in introduced populations of Pacific oyster, *Crassostrea gigas*, in British Columbia. *Canadian J. Fisheries and Aquatic Sci.* **51** (7), 1608–1614.
- [7] L. BREIMAN (1992) *Probability, Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM) **7**, Pennsylvania.
- [8] A. CALIEBE, R. NEININGER, M. KRAWCZAK AND U. ROESLER (2007) On the length distribution of external branches in coalescence trees: genetic diversity within species. *Theor. Popul. Biol.* **72** (2), 245–252.
- [9] J.-F. DELMAS, J.-S. DHERSIN AND A. SIRI-JÉGOUSSE (2008) Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.* **18** (2), 997–1025.
- [10] F.D. HOLLANDER (2008) *Large Deviations*. Amer Mathematical Society **14**.
- [11] J.-S. DHERSIN AND M. MÖHLE (2013) On the external branches of coalescent processes with multiple collisions with an emphasis on the Bolthausen–Sznitman coalescent. *Electron. J. Probab.* **18** (40), 1–11.
- [12] J.-S. DHERSIN, A. SIRI-JÉGOUSSE, F. FREUND AND L. YUAN (2013) On the length of an external branch in the beta-coalescent. *Stoch. Process. Appl.* **123**, 1691–1715.
- [13] J.-S. DHERSIN AND L. YUAN (2012) Asymptotic behavior of the total length of external branches for Beta-coalescents. Preprint, [arXiv:1202.5859](https://arxiv.org/abs/1202.5859).
- [14] M. DRMOTA, A. IKSANOV, M. MÖHLE AND U. ROESLER (2007) Asymptotic results concerning the total branch length of the Bolthausen–Sznitman coalescent. *Stoch. Process. Appl.* **117** (10), 1404–1421.
- [15] B. ELDON AND J. WAKELEY (2006) Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* **172**, 2621–2633.
- [16] W. FELLER (1971) *An Introduction to Probability Theory and Its Applications*, Vol. II. 2nd ed. John Wiley & Sons Inc., New York.
- [17] F. FREUND AND M. MÖHLE (2009) On the time back to the most recent common ancestor and the external branch length of the Bolthausen–Sznitman coalescent. *Markov Processes Relat. Fields* **15** (3), 387–416.
- [18] A. GNEDIN, A. IKSANOV, A. MARYNYCH AND M. MÖHLE (2014) On asymptotics of the beta-coalescents. *Adv. Appl. Probab.* **46** (2), 496–515.
- [19] A. GNEDIN, A. IKSANOV AND M. MÖHLE (2008) On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Probab.* **45** (4), 1186–1195.
- [20] D. HEDGECOCK (1994) Does variance in reproductive success limit effective population sizes of marine organisms? In: *Genetics and Evolution of Aquatic Organisms*, A.R. Beaumont (ed.), Chapman and Hall, London, 122–134.
- [21] S. JANSON AND G. KERSTING (2011) On the total external length of the Kingman coalescent. *Electron. J. Probab.* **16**, 2203–2218.
- [22] O. KALLENBERG (2002) *Foundations of Modern Probability*. 2nd ed. Probability and its Applications. Springer-Verlag, New York.

- [23] J. KINGMAN (1982) The coalescent. *Stoch. Process. Appl.* **13** (3), 235–248.
- [24] J. KINGMAN (1982) On the genealogy of large populations. *J. Appl. Probab.* (Special Vol. **19A**), Essays in statistical science, 27–43.
- [25] J. KINGMAN (2000) Origins of the coalescent 1974–1982. *Genetics* **156** (4), 1461–1463.
- [26] M. MÖHLE (2010) Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson–Dirichlet coalescent. *Stoch. Process. Appl.* **120** (11), 2159–2173.
- [27] J. PITMAN (1999) Coalescents with multiple collisions. *Ann. Probab.* **27** (4), 1870–1902.
- [28] S. SAGITOV (1999) The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36** (4), 1116–1125.
- [29] J. SCHWEINSBERG (2000) Coalescents with simultaneous multiple collisions. *Electron. J. Probab* **5**, Paper no. 12, 50 pp.