



## On the block counting process and the fixation line of exchangeable coalescents

Florian Gaiser and Martin Möhle

Mathematisches Institut  
Eberhard Karls Universität Tübingen  
Auf der Morgenstelle 10  
72076 Tübingen, Germany  
*E-mail address:* { `florian.gaiser` , `martin.moehle` } @uni-tuebingen.de

**Abstract.** We study the block counting process and the fixation line of exchangeable coalescents. Formulas for the infinitesimal rates of both processes are provided. It is shown that the block counting process is Siegmund dual to the fixation line. For exchangeable coalescents restricted to a sample of size  $n$  and with dust we provide a convergence result for the block counting process as  $n$  tends to infinity. The associated limiting process is related to the frequencies of singletons of the coalescent. Via duality we obtain an analog convergence result for the fixation line of exchangeable coalescents with dust. The Dirichlet coalescent and the Poisson–Dirichlet coalescent are studied in detail.

### 1. Introduction

Coalescent processes have attracted the interest of many researchers over the last decades, mainly in probability theory and population genetics. Most results in coalescent theory concern coalescents with multiple collisions independently introduced by Pitman (1999) and Sagitov (1999). Less is known for the full class of exchangeable coalescents  $\Pi = (\Pi_t)_{t \geq 0}$  allowing for simultaneous multiple collisions of ancestral lineages. Note that  $\Pi$  is a Markovian process taking values in the space  $\mathcal{P}$  of partitions of  $\mathbb{N} := \{1, 2, \dots\}$ . Schweinsberg (2000a) showed that every exchangeable coalescent  $\Pi$  can be characterized by a finite measure  $\Xi$  on the infinite simplex

$$\Delta := \{x = (x_r)_{r \in \mathbb{N}} : x_1 \geq x_2 \geq \dots \geq 0, |x| := \sum_{r=1}^{\infty} x_r \leq 1\}. \quad (1.1)$$

Exchangeable coalescents are therefore called  $\Xi$ -coalescents. The aim of this article is to provide some more information on the block counting process and the fixation

---

*Received by the editors January 15, 2016; accepted August 18, 2016.*

*2010 Mathematics Subject Classification.* 60F05, 60J27, 92D15, 97K60.

*Key words and phrases.* Block counting process, coming down from infinity, Dirichlet coalescent, duality, dust, exchangeable coalescent, fixation line, Poisson–Dirichlet coalescent.

line of the  $\Xi$ -coalescent and on the relation between these two processes. We therefore briefly recall the definition of the block counting process and turn afterwards to the fixation line.

For  $t \geq 0$  let  $N_t$  denote the number of blocks of  $\Pi_t$ . It is well known that  $N := (N_t)_{t \geq 0}$  is a Markovian process with state space  $S := \mathbb{N} \cup \{\infty\}$ , called the block counting process of  $\Pi$ . We use the notation  $N_t^{(n)}$  for the number of blocks of  $\Pi_t$  restricted to a sample of size  $n \in \mathbb{N}$ . The block counting process has been studied extensively in the literature with a main focus on coalescents with multiple collisions ( $\Lambda$ -coalescents). We will revisit some of its properties throughout this article.

The definition of the fixation line is more involved. As in [Schweinsberg \(2000a\)](#) decompose  $\Xi = \Xi(\{0\})\delta_0 + \Xi_0$  with  $\delta_0$  the Dirac measure at  $0 \in \Delta$  and  $\Xi_0$  having no atom at 0. For  $x = (x_r)_{r \in \mathbb{N}} \in \Delta$  define  $(x, x) := \sum_{r=1}^{\infty} x_r^2$  and  $\nu(dx) := \Xi_0(dx)/(x, x)$ . One possible definition of the fixation line is based on the lookdown construction going back to [Donnelly and Kurtz \(1996, 1999\)](#). For some further information on the lookdown construction we refer the reader to [Birkner et al. \(2009\)](#). Imagine a population consists of countably many individuals distinguished by their levels. The level of an individual is a positive integer, and the individual at time  $t \geq 0$  at level  $i \in \mathbb{N}$  is denoted by  $(t, i)$ . In the following we use the Poisson process construction of [Schweinsberg \(2000a, Section 3\)](#). Define  $\mathbb{N}_0 := \{0, 1, 2, \dots\}$  and, for every  $x = (x_r)_{r \in \mathbb{N}} \in \Delta$ , let  $P_x$  be the law of a sequence  $\xi = (\xi_1, \xi_2, \dots)$  of independent and identically distributed  $\mathbb{N}_0$ -valued random variables with distribution  $\mathbb{P}(\xi_1 = 0) := 1 - |x|$  and  $\mathbb{P}(\xi_1 = r) = x_r$ ,  $r \in \mathbb{N}$ . Furthermore, for  $i, j \in \mathbb{N}$  with  $i < j$  let  $z_{ij}$  be the sequence  $(z_1, z_2, \dots)$  with  $z_i = z_j := 1$  and  $z_k := 0$  for  $k \notin \{i, j\}$ . Take a Poisson process on  $[0, \infty) \times \mathbb{N}_0^{\mathbb{N}}$  with intensity measure  $\lambda \otimes \mu$ , where  $\lambda$  denotes the Lebesgue measure on  $[0, \infty)$  and

$$\mu(A) := \Xi(\{0\}) \sum_{\substack{i, j \in \mathbb{N} \\ i < j}} 1_{\{z_{ij} \in A\}} + \int_{\Delta} P_x(A) \nu(dx) \quad (1.2)$$

for all measurable  $A \subseteq \mathbb{N}_0^{\mathbb{N}}$ . Each atom  $(t, x)$  corresponds to a reproduction event which is defined as follows. For  $r \in \mathbb{N}$  define  $J_r := \{j \in \mathbb{N} : \xi_j = r\}$ . For  $r \in \mathbb{N}$  and  $j \in J_r$ , the individual  $(t, j)$  is a child of the individual  $(t-, \min J_r)$ . The other lineages are shifted upwards keeping the order they had before the reproduction event. The construction of this countable infinite population model is called the lookdown construction.

We are now able to define the fixation line. Fix  $i \in \mathbb{N}$ . The levels of the offspring at time  $t \geq 0$  of the individual  $(0, i + 1)$ , that is the individual at time 0 at level  $i + 1$ , form a subset of  $\mathbb{N}$ , whose minimal element (if it exists) we denote by  $L_t^{(i)} + 1$ . If this subset is empty, we define  $L_t^{(i)} := \infty$ . For example, if at the time of a reproduction event,  $J_1 = \mathbb{N}$  and  $J_r = \emptyset$  for all  $r \in \mathbb{N} \setminus \{1\}$ , then the individual at time 0 at level 2 has no offspring at all, so in this case we have  $L_t^{(1)} = \infty$ . By construction, the process  $L^{(i)} := (L_t^{(i)})_{t \geq 0}$  has state space  $\{i, i + 1, \dots\} \cup \{\infty\}$  and non-decreasing paths. Moreover,  $L_t^{(1)} \leq L_t^{(2)} \leq \dots$ . We define  $L_t := L_t^{(1)}$  and  $L := (L_t)_{t \geq 0}$ . When  $L_t$  reaches level  $n$ , all the individuals at time  $t$  with levels  $1, \dots, n$  are offspring of the single individual  $(0, 1)$ . Therefore, the whole finite population of size  $n$  (consisting of the  $n$  individuals at time  $t$  with levels  $1, \dots, n$ )

stems from the same individual at time 0, an event called fixation in population genetics.

The process  $L$  is hence called the fixation line. This process can be traced back to Pfaffelhuber and Wakolbinger (2006) for the Kingman coalescent. For the  $\Lambda$ -coalescent the fixation line appears in Labbé (2014) and was further studied by Hénard (2013, 2015).

The block counting process and the fixation line are two rather useful processes in coalescent theory. Several functionals of the  $n$ -coalescent (the coalescent restricted to a sample of size  $n \in \mathbb{N}$ ), for example the number of jumps, the time back to the most recent common ancestor (absorption time) and the total tree length of the  $n$ -coalescent, can be expressed in terms of the block counting process  $(N_t^{(n)})_{t \geq 0}$ . We exemplarily refer the reader to Gnedin et al. (2014) for a survey on the asymptotic behavior of these functionals as  $n$  tends to infinity. The fixation line is itself an important functional of generalized Fleming–Viot processes. For the Kingman case it has for example been proven to be a powerful tool in the study of evolving coalescent trees (Pfaffelhuber and Wakolbinger, 2006; Pfaffelhuber et al., 2011). Last but not least, conditions for the  $\Xi$ -coalescent to come down from infinity are translated via Siegmund duality into conditions for explosion of the fixation line (see Remark 2.11) and may therefore help to solve the in its full generality still open problem to find a nice characterization of all measures  $\Xi$  for which the  $\Xi$ -coalescent comes down from infinity (Herriger and Möhle, 2012; Schweinsberg, 2000a).

We close the introduction by a brief summary of the organization of the article. Section 2 contains the main results. Propositions 2.1 and 2.5 provide formulas for the infinitesimal rates and the total rates of the block counting process  $N = (N_t)_{t \geq 0}$  and the fixation line  $L = (L_t)_{t \geq 0}$  for arbitrary  $\Xi$ -coalescents. Theorem 2.9 shows that the block counting process  $N$  is Siegmund dual to the fixation line  $L$ . For  $\Xi$ -coalescents with dust, Theorem 2.13 provides a convergence result for  $N$  and  $L$  when their initial state tends to infinity. The limiting processes are related to the frequencies of singletons of the coalescent and can be expressed in terms of the subordinator associated with the coalescent. Similar limiting processes appear in other frameworks of Markov chains with rare large jumps, see for example Haas and Miermont (2011) and the recent preprint of Bertoin and Kortchemski (2014). In Sections 3 and 4 we study the Dirichlet coalescent and the Poisson–Dirichlet coalescent respectively. The proofs of the results stated in Section 2 are provided in Section 5. The appendix deals with a duality relation for generalized Stirling numbers being closely related to the Siegmund duality of  $N$  and  $L$ .

## 2. Results

Throughout the article we shall use the following subsets of the infinite simplex  $\Delta$  defined in (1.1). For  $n \in \mathbb{N}$  define  $\Delta_n := \{x = (x_r)_{r \in \mathbb{N}} \in \Delta : x_1 + \dots + x_n = 1\}$ . Furthermore let  $\Delta_f := \bigcup_{n \in \mathbb{N}} \Delta_n = \{x = (x_r)_{r \in \mathbb{N}} \in \Delta : x_1 + \dots + x_n = 1 \text{ for some } n \in \mathbb{N}\}$  and  $\Delta^* := \{x \in \Delta : |x| = 1\}$ . Note that  $\Delta_1 \subset \Delta_2 \subset \dots \subset \Delta_f \subset \Delta^* \subset \Delta$  and that  $\nu(\Delta_n) \leq n\Xi(\Delta_n) < \infty$ , since  $(x, x) = \sum_{r=1}^n x_r^2 = 1/n + \sum_{r=1}^n (x_r - 1/n)^2 \geq 1/n$  for all  $n \in \mathbb{N}$  and  $x \in \Delta_n$ .

In order to state our first result it is convenient to introduce the following urn model which is essentially a version of Kingman’s paintbox construction (Kingman, 1982, Section 8) and can be also viewed as a generalization of an infinite

urn model investigated by [Hwang and Janson \(2008\)](#). Fix  $x = (x_r)_{r \in \mathbb{N}} \in \Delta$ . Recall that  $|x| := \sum_{r=1}^{\infty} x_r$  and define  $x_0 := 1 - |x|$  for convenience. Imagine a countable infinite number of boxes having labels  $r \in \mathbb{N}_0$ . Balls are allocated successively to these boxes, where it is assumed that every ball will go to box  $r \in \mathbb{N}_0$  with probability  $x_r$  independently of the other balls. If  $X_r(i, x)$  denotes the number of balls in box  $r \in \mathbb{N}_0$  after  $i \in \mathbb{N}_0$  balls have been allocated, then clearly  $(X_0(i, x), X_1(i, x), X_2(i, x), \dots)$  has an infinite multinomial distribution with parameters  $i$  and  $(x_0, x_1, x_2, \dots)$ . The random variable

$$Y(i, x) := X_0(i, x) + \sum_{r=1}^{\infty} 1_{\{X_r(i, x) \geq 1\}} \tag{2.1}$$

counts the balls contained in box 0 plus the number of other boxes which are non-empty. In the language of Kingman’s paintbox, when a ball goes to box  $r \in \mathbb{N}$ , it will be painted with color  $r$ . The box 0 plays a distinguished role. Each ball going to box 0 is painted with a new color never seen before.  $Y(i, x)$  is the number of different colors after  $i$  balls have been painted.

Propositions 2.1 and 2.5 below underline the well known fact that the process  $(Y(i, x))_{i \in \mathbb{N}_0}$  plays a fundamental role in coalescent theory. Proposition 2.1 concerns the infinitesimal rates of the block counting process  $N$ . These rates are essentially known from the literature (see, for example, [Freund and Möhle, 2009](#)). We state the result for the record and since the case  $\Xi(\Delta_f) > 0$  requires some attention.

**Proposition 2.1** (Rates of the block counting process). *Let  $\Xi$  be a finite measure on  $\Delta$  and let  $\Pi$  be a  $\Xi$ -coalescent. The block counting process  $N = (N_t)_{t \geq 0}$  of  $\Pi$  moves from state  $i \in \mathbb{N}$  to state  $j \in \mathbb{N}$  with  $j < i$  at the rate*

$$q_{ij} = \Xi(\{0\}) \binom{i}{2} \delta_{j, i-1} + \int_{\Delta} \mathbb{P}(Y(i, x) = j) \nu(dx) \tag{2.2}$$

with  $Y(i, x)$  defined via (2.1). The probability below the integral in (2.2) can be provided explicitly as  $\mathbb{P}(Y(i, x) = j) = \sum_{k=1}^j f_{ijk}(x)$ , where

$$f_{ijk}(x) := \frac{x_0^{j-k}}{(j-k)!} \sum_{\substack{i_1, \dots, i_k \in \mathbb{N} \\ i_1 + \dots + i_k = i-j+k}} \frac{i!}{i_1! \dots i_k!} \sum_{\substack{r_1, \dots, r_k \in \mathbb{N} \\ r_1 < \dots < r_k}} x_{r_1}^{i_1} \dots x_{r_k}^{i_k}. \tag{2.3}$$

The total rates are

$$\begin{aligned} q_i &:= \sum_{j=1}^{i-1} q_{ij} = \Xi(\{0\}) \binom{i}{2} + \int_{\Delta} \mathbb{P}(Y(i, x) < i) \nu(dx) \\ &= \Xi(\{0\}) \binom{i}{2} + \int_{\Delta} \left( 1 - x_0^i - \sum_{k=1}^i \binom{i}{k} x_0^{i-k} \sum_{\substack{r_1, \dots, r_k \in \mathbb{N} \\ \text{all distinct}}} x_{r_1} \dots x_{r_k} \right) \nu(dx), \end{aligned} \tag{2.4}$$

$i \in \mathbb{N}$ . Moreover,  $q_{\infty j} = \nu(\Delta_j) - \nu(\Delta_{j-1})$  for  $j \in \mathbb{N}$  ( $\Delta_0 := \emptyset$ ) and  $q_{\infty \infty} = -\nu(\Delta_f)$ .

*Remark 2.2.* From  $q_{i+1} - q_i = i \Xi(\{0\}) + \int_{\Delta} \mathbb{P}(Y(i+1, x) = i, Y(i, x) = i) \nu(dx) \geq 0$  we conclude that  $q_{i+1} \geq q_i$  with equality  $q_{i+1} = q_i$  if and only if  $\Xi(\Delta \setminus \Delta_{i-1}) = 0$ ,  $i \in \mathbb{N}$ . Thus, if  $\Xi(\Delta \setminus \Delta_f) > 0$  then the total rates  $q_i$ ,  $i \in \mathbb{N}$ , are pairwise distinct.

*Remark 2.3.* For the rates of the block counting process of the Dirichlet coalescent and the Poisson–Dirichlet coalescent we refer the reader to (3.1), (3.2) and (4.1).

*Remark 2.4.* For the  $\Lambda$ -coalescent the rate (2.2) reduces to the well known formula (see, for example, Pitman, 1999 or Möhle, 2006, Eq. (13))

$$q_{ij} = \binom{i}{j-1} \int_{[0,1]} x^{i-j-1} (1-x)^{j-1} \Lambda(dx), \quad i, j \in \mathbb{N}, j < i, \quad (2.5)$$

and the total rates are (Möhle, 2006, Eq. (14))

$$q_i = \Lambda(\{0\}) \binom{i}{2} + \int_{(0,1]} \frac{1 - (1-x)^i - ix(1-x)^{i-1}}{x^2} \Lambda(dx), \quad i \in \mathbb{N}.$$

For the  $\beta(a, b)$ -coalescent with parameters  $a, b \in (0, \infty)$  the rate (2.5) reduces to

$$q_{ij} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(i+1)}{\Gamma(i-2+a+b)} \frac{\Gamma(j-1+b)}{\Gamma(j)} \frac{\Gamma(i-j-1+a)}{\Gamma(i-j+2)}, \quad i, j \in \mathbb{N}, j < i. \quad (2.6)$$

We now state the analog result for the rates of the fixation line  $L = (L_t)_{t \geq 0}$ .

**Proposition 2.5** (Rates of the fixation line). *Let  $\Xi$  be a finite measure on  $\Delta$ . The fixation line  $L = (L_t)_{t \geq 0}$  moves from state  $i \in \mathbb{N}$  to state  $j \in \mathbb{N}$  with  $j > i$  at the rate*

$$\gamma_{ij} = \Xi(\{0\}) \binom{j}{2} \delta_{j,i+1} + \int_{\Delta} \mathbb{P}(Y(j, x) = i, Y(j+1, x) = i+1) \nu(dx) \quad (2.7)$$

with  $Y(\cdot, x)$  defined in (2.1). Moreover,  $\gamma_{i\infty} = \nu(\Delta_i)$  for all  $i \in \mathbb{N}$  and  $\gamma_{\infty\infty} = 0$ . The probability below the integral in (2.7) can be provided explicitly, namely  $\mathbb{P}(Y(j, x) = i, Y(j+1, x) = i+1) = \sum_{k=1}^i g_{ijk}(x)$ , where

$$g_{ijk}(x) := \frac{x_0^{i-k}}{(i-k)!} \sum_{\substack{i_1, \dots, i_k \in \mathbb{N} \\ i_1 + \dots + i_k = j-i+k}} \frac{j!}{i_1! \dots i_k!} \sum_{\substack{r_1, \dots, r_k \in \mathbb{N} \\ r_1 < \dots < r_k}} x_{r_1}^{i_1} \dots x_{r_k}^{i_k} \left( 1 - \sum_{l=1}^k x_{r_l} \right). \quad (2.8)$$

The total rates are  $\gamma_i := \sum_{j \in \{i+1, i+2, \dots\} \cup \{\infty\}} \gamma_{ij} = q_{i+1}$ ,  $i \in \mathbb{N}$ .

*Remark 2.6.* In general  $g_{ijk}(x)$  is not equal to  $f_{jik}(x)$  (see (2.3)) because of the additional factor  $1 - \sum_{l=1}^k x_{r_l}$  occurring on the right hand side in (2.8). This additional factor comes from the fact that in the paintbox construction, on the event that box 0 contains  $i-k$  balls and that the boxes  $r_1, \dots, r_k$  are non-empty,  $\{Y(j+1, x) = i+1\}$  corresponds to the event that ball  $j+1$  belongs to a box  $r \in \mathbb{N}_0 \setminus \{r_1, \dots, r_k\}$ , which has probability  $1 - \sum_{l=1}^k x_{r_l}$ .

*Remark 2.7.* Note that  $\gamma_{i+1} = \gamma_i$  if and only if  $\Xi(\Delta \setminus \Delta_i) = 0$ ,  $i \in \mathbb{N}$ . If  $\Xi(\Delta \setminus \Delta_f) > 0$  then the total rates  $\gamma_i$ ,  $i \in \mathbb{N}$ , are pairwise distinct.

*Remark 2.8.* For the  $\Lambda$ -coalescent the rate (2.7) reduces to

$$\gamma_{ij} = \binom{j}{i-1} \int_{[0,1]} x^{j-i-1} (1-x)^i \Lambda(dx), \quad i, j \in \mathbb{N}, i < j, \quad (2.9)$$

in agreement with Hénard (2015, Lemma 2.3). The total rates are

$$\gamma_i = q_{i+1} = \Lambda(\{0\}) \binom{i+1}{2} + \int_{(0,1]} \frac{1 - (1+ix)(1-x)^i}{x^2} \Lambda(dx), \quad i \in \mathbb{N}.$$

For the  $\Lambda$ -coalescent the equality  $\gamma_i = q_{i+1}$  was already observed by Hénard (2015).

For the Kingman coalescent ( $\Lambda = \delta_0$ ) we have  $\gamma_i = \gamma_{i,i+1} = \binom{i+1}{2}$  and  $\gamma_{ij} = 0$  for all  $j \notin \{i, i+1\}$ . Thus,  $L$  is a pure birth process with state space  $\mathbb{N} \cup \{\infty\}$  and

birth rates  $\gamma_i = \binom{i+1}{2}$ ,  $i \in \mathbb{N}$ . This process explodes, i.e.  $\mathbb{P}(L_t = \infty) > 0$  for all  $t > 0$ , since  $\sum_{i=1}^{\infty} 1/\gamma_i < \infty$ . In fact,  $L$  reaches  $\infty$  in finite time almost surely. The explosion of the fixation line for the Kingman case has for example been applied to study evolving coalescent trees (Pfaffelhuber and Wakolbinger, 2006; Pfaffelhuber et al., 2011) and related functionals (Delmas et al., 2010). For more details on explosion of  $L$  we refer the reader to Remark 2.11.

For the  $\beta(a, b)$ -coalescent with  $a, b \in (0, \infty)$  the rates (2.9) of the fixation line reduce to

$$\gamma_{ij} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(i+b)}{\Gamma(i)} \frac{\Gamma(j+1)}{\Gamma(j-1+a+b)} \frac{\Gamma(j-i-1+a)}{\Gamma(j-i+2)}, \quad i, j \in \mathbb{N}, i < j. \tag{2.10}$$

Simple formulas for the total rates  $\gamma_i$  seem to be only available for particular parameter choices of  $a$  and  $b$ . For example, for the  $\beta(2-\alpha, \alpha)$ -coalescent with parameter  $0 < \alpha < 2$ , which has attracted the interest of several researchers, the fixation line has total rates

$$\begin{aligned} \gamma_i &= \sum_{j=i+1}^{\infty} \gamma_{ij} = \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} \frac{\Gamma(i+\alpha)}{\Gamma(i)} \sum_{j=i+1}^{\infty} \frac{\Gamma(j-i+1-\alpha)}{\Gamma(j-i+2)} \\ &= \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} \frac{\Gamma(i+\alpha)}{\Gamma(i)} \frac{\Gamma(2-\alpha)}{\alpha} = \frac{\Gamma(i+\alpha)}{\Gamma(\alpha+1)\Gamma(i)} = \prod_{k=1}^{i-1} \frac{k+\alpha}{k}, \quad i \in \mathbb{N}. \end{aligned}$$

In particular,  $\gamma_i = i$  for the Bolthausen–Sznitman coalescent ( $\alpha = 1$ ). Another class of beta-coalescents for which a nice formula for the total rate  $\gamma_i$  is available is the  $\beta(3, b)$ -coalescent with parameter  $b > 0$ . In this case the fixation line has total rates

$$\begin{aligned} \gamma_i &= \sum_{j=i+1}^{\infty} \gamma_{ij} = \frac{\Gamma(3+b)}{\Gamma(3)\Gamma(b)} \frac{\Gamma(i+b)}{\Gamma(i)} \sum_{j=i+1}^{\infty} \frac{\Gamma(j+1)}{\Gamma(j+b+2)} \\ &= \frac{\Gamma(3+b)}{\Gamma(3)\Gamma(b)} \frac{\Gamma(i+b)}{\Gamma(i)} \frac{\Gamma(i+2)}{b\Gamma(i+b+2)} = \frac{(b+1)(b+2)i(i+1)}{2(i+b)(i+b+1)}, \quad i \in \mathbb{N}. \end{aligned}$$

We now turn to the duality of the block counting process  $N$  and the fixation line  $L$ . The following result (Theorem 2.9) is a reformulation and generalization of Lemma 2.1 of Hénard (2015). Note that Theorem 2.9 holds for any  $\Xi$ -coalescent. Recall the notation  $S := \mathbb{N} \cup \{\infty\}$ .

**Theorem 2.9** (Siegmond duality of  $N$  and  $L$ ). *Let  $\Pi$  be a  $\Xi$ -coalescent and let  $N = (N_t)_{t \geq 0}$  and  $L = (L_t)_{t \geq 0}$  denote the block counting process and the fixation line of  $\Pi$  respectively. Then  $N$  is dual in the sense of Liggett (2005, p. 84, Definition 3.1) to  $L$  with respect to the Siegmond duality kernel  $H : S^2 \rightarrow \{0, 1\}$  defined via  $H(i, j) := 1$  for  $i \leq j$  and  $H(i, j) := 0$  otherwise, i.e.*

$$\begin{aligned} \mathbb{P}(N_t^{(i)} \leq j) &= \mathbb{P}(N_t \leq j \mid N_0 = i) = \mathbb{E}(H(N_t, j) \mid N_0 = i) \\ &= \mathbb{E}(H(i, L_t) \mid L_0 = j) = \mathbb{P}(L_t \geq i \mid L_0 = j) = \mathbb{P}(L_t^{(j)} \geq i) \end{aligned}$$

for all  $i, j \in S$  and  $t \geq 0$ . If  $Q = (q_{ij})_{i, j \in S}$  and  $\Gamma = (\gamma_{ij})_{i, j \in S}$  denote the generator matrices of  $N$  and  $L$  respectively then  $q_{i, \leq j} := \sum_{k \in S, k \leq j} q_{ik} = \sum_{k \in S, k \geq i} \gamma_{jk} =: \gamma_{j, \geq i}$  for all  $i, j \in S$ .

*Remark 2.10.* Note that  $L$  is dual to  $N$  with respect to the transposed duality kernel  $H^\top : S^2 \rightarrow \{0, 1\}$  defined via  $H^\top(i, j) := H(j, i)$  for all  $i, j \in S$ . For  $i, j \in S$  define

$g_{ij} := \int_0^\infty \mathbb{P}(N_t^{(i)} = j) dt \in [0, \infty]$  and  $\widehat{g}_{ij} := \int_0^\infty \mathbb{P}(L_t^{(i)} = j) dt \in [0, \infty]$ . Note that  $G := (g_{ij})_{i,j \in S}$  and  $\widehat{G} := (\widehat{g}_{ij})_{i,j \in S}$  are the Green matrices of  $N$  and  $L$  respectively. The matrix  $G$  is lower left triangular whereas  $\widehat{G}$  is upper right triangular. From Theorem 2.3 it follows that

$$g_{i,>j} = \int_0^\infty \mathbb{P}(N_t^{(i)} > j) dt = \int_0^\infty \mathbb{P}(L_t^{(j)} < i) dt = \widehat{g}_{j,<i}, \quad i, j \in S,$$

where  $g_{i,>j} := \sum_{k \in S, k > j} g_{ik}$  and  $\widehat{g}_{j,<i} := \sum_{k \in S, k < i} \widehat{g}_{jk}$ ,  $i, j \in S$ . For  $i, j \in T := \mathbb{N} \setminus \{1\}$ , the set of states which are transient for  $N$  and  $L$ , it follows that

$$g_{ij} = g_{i,>j-1} - g_{i,>j} = \widehat{g}_{j-1,<i} - \widehat{g}_{j,<i} = \sum_{k \in S, k < i} (\widehat{g}_{j-1,k} - \widehat{g}_{jk})$$

and

$$\widehat{g}_{ij} = \widehat{g}_{i,<j+1} - \widehat{g}_{i,<j} = g_{j+1,>i} - g_{j,>i} = \sum_{k \in S, k > i} (g_{j+1,k} - g_{jk}).$$

*Remark 2.11.* As in Schweinsberg (2000a) we say that  $\Pi$  comes down from infinity if  $\mathbb{P}(N_t < \infty) = 1$  for all  $t > 0$  and that  $\Pi$  stays infinite if  $\mathbb{P}(N_t = \infty) = 1$  for all  $t > 0$ . Note that there exist coalescents, for example the Dirichlet coalescent studied in Section 3, that neither come down from infinity nor stay infinite. We refer the reader to Schweinsberg (2000a,b) and Herriger and Möhle (2012) for methods to determine whether a coalescent  $\Pi$  comes down from infinity or stays infinite.

We say that  $L$  does not explode if  $T_\infty := \inf\{t > 0 : L_{t-} = \infty\} = \infty$  almost surely. Note that  $L$  does not explode if and only if  $\mathbb{P}(L_t < \infty) = 1$  for all  $t \geq 0$ . By the general explosion criterion for Markov chains,  $L$  does not explode if and only if  $\sum_{n=0}^\infty 1/\gamma_{\chi_n} = \infty$  almost surely, where  $\chi = (\chi_n)_{n \in \mathbb{N}_0}$  denotes the jump chain of  $L$ . Note that  $\chi$  has transition probabilities  $p_{ij} := \gamma_{ij}/\gamma_i$ ,  $1 \leq i < j \leq \infty$ .

If  $\Pi$  comes down from infinity then  $L$  explodes. Moreover,  $\Pi$  stays infinite if and only if  $L$  does not explode.

*Proof.* If  $\Pi$  comes down from infinity, then  $\Pi$  eventually becomes absorbed almost surely, i.e.  $\lim_{t \rightarrow \infty} \mathbb{P}(N_t = 1) = 1$ . Thus, by duality,  $\mathbb{P}(L_t = \infty) = \mathbb{P}(N_t = 1) > 0$  for all sufficiently large  $t$ , i.e.  $L$  explodes.

If  $\Pi$  stays infinite, then  $\mathbb{P}(N_t = \infty) = 1$  for all  $t \geq 0$ . Thus,  $\mathbb{P}(L_t = \infty) = \mathbb{P}(N_t = 1) = 0$  for all  $t \geq 0$ , i.e.  $L$  does not explode. Conversely, suppose that  $L$  does not explode. Then, by the first statement,  $\Pi$  does not come down from infinity. Moreover, we must have  $\gamma_{i\infty} = 0$  for all  $i \in \mathbb{N}$ , because otherwise every  $L_t$  would be equal to  $\infty$  with positive probability. Since  $\gamma_{i\infty} = \nu(\Delta_i)$  it follows that  $\nu(\Delta_f) = \lim_{i \rightarrow \infty} \nu(\Delta_i) = 0$ . Thus,  $\Xi(\Delta_f) = 0$ . But under the additional assumption that  $\Xi(\Delta_f) = 0$  the coalescent does not come down from infinity if and only if the coalescent stays infinite (see Schweinsberg, 2000a).

*Remark 2.12.* Recall that  $L$  explodes if  $\Pi$  comes down from infinity. The converse holds under the additional assumption that  $\Xi(\Delta_f) = 0$ , but it does not hold in general. Examples where  $L$  explodes but  $\Pi$  does not come down from infinity are provided in Section 3 (Dirichlet coalescent).

A  $\Xi$ -coalescent  $\Pi = (\Pi_t)_{t \geq 0}$  has proper frequencies if, for all times  $t \geq 0$ , the frequency of singletons  $S_t$  of  $\Pi_t$  satisfies  $S_t = 0$  almost surely. For a precise definition of  $S_t$  we refer the reader to Möhle (2010, Section 3). Schweinsberg (2000a, Proposition 30) showed that  $\Pi$  does not have proper frequencies if and only if  $\Xi(\{0\}) = 0$

and  $\int_{\Delta} |x| \nu(dx) < \infty$ . In this case the process  $Z := (Z_t)_{t \geq 0} := (-\log S_t)_{t \geq 0}$  (with the convention  $-\log 0 := \infty$ ) is a drift-free subordinator with state space  $[0, \infty]$  and Laplace exponent

$$\Phi(\eta) = \int_{\Delta} (1 - (1 - |x|)^\eta) \nu(dx), \quad \eta \geq 0. \quad (2.11)$$

Note that  $\mathbb{E}(S_t^\eta) = \mathbb{E}(e^{-\eta Z_t}) = e^{-t\Phi(\eta)}$ ,  $\eta \geq 0$ . A coalescent without proper frequencies is also called a coalescent with dust (Gnedin et al., 2011). Theorem 2.13 below clarifies the asymptotic behavior of the block counting process  $N^{(n)}$  and the fixation line  $L^{(n)}$  as  $n \rightarrow \infty$  for  $\Xi$ -coalescents with dust. Note that we use the conventions  $e^{-\infty} := 0$ ,  $e^\infty := \infty$  and  $1/0 := \infty$ .

**Theorem 2.13.** *Let  $\Pi$  be a  $\Xi$ -coalescent with dust or, equivalently,  $\Xi(\{0\}) = 0$  and  $\int_{\Delta} |x| \nu(dx) < \infty$ . Then, the following two assertions hold.*

- a) *As  $n \rightarrow \infty$  the scaled block counting process  $(N_t^{(n)}/n)_{t \geq 0}$  converges in  $D_{[0,1]}[0, \infty)$  to the frequency of singleton process  $S = (S_t)_{t \geq 0} = (e^{-Z_t})_{t \geq 0}$ .*
- b) *As  $n \rightarrow \infty$  the scaled fixation line  $(L_t^{(n)}/n)_{t \geq 0}$  converges in  $D_{[1,\infty]}[0, \infty)$  to the reciprocal frequency of singleton process  $(1/S_t)_{t \geq 0} = (e^{Z_t})_{t \geq 0}$ .*

*Remark 2.14.* Theorem 2.13 can be stated logarithmically as follows. For  $\Xi$ -coalescents with dust, as  $n \rightarrow \infty$ , both processes  $(\log n - \log N_t^{(n)})_{t \geq 0}$  and  $(\log L_t^{(n)} - \log n)_{t \geq 0}$  converge in  $D_{[0,\infty]}[0, \infty)$  to the drift-free subordinator  $Z$  with Laplace exponent (2.11). Similar limiting processes appear in other frameworks of Markov chains with rare large jumps (see, for example, Bertoin and Kortchemski, 2014 and Haas and Miermont, 2011). Clearly, Theorem 2.13 holds for  $\Lambda$ -coalescents with dust, for example for  $\beta(a, b)$ -coalescents with  $a > 1$  and  $b > 0$ .

*Remark 2.15.* If  $\Xi$  is concentrated on  $\Delta^*$  then the coalescent has dust if and only if  $\nu$  is finite. In this case the Laplace exponent (2.11) satisfies  $\Phi(\eta) = \nu(\Delta^*)$ ,  $\eta > 0$ , so  $S_t$  has the same distribution as  $1_{\{T_f > t\}}$  for all  $t \geq 0$ , where  $T_f$  is exponentially distributed with parameter  $\nu(\Delta^*)$ . Examples are the Dirichlet coalescent studied in Section 3, the Poisson–Dirichlet coalescent (Sagitov, 2003, Section 3) and the two-parameter Poisson–Dirichlet coalescent (Möhle, 2010, Section 6). More information on the two-parameter Poisson–Dirichlet coalescent is provided in Section 4.

*Remark 2.16.* Theorem 2.13 excludes dust-free coalescents. For the Bolthausen–Sznitman coalescent we refer the reader to Kukla and Möhle (2016), Möhle (2015) and Schweinsberg (2012) for asymptotic results concerning the block counting process  $N^{(n)}$  and the fixation line  $L^{(n)}$ . For dust-free coalescents different from the Bolthausen–Sznitman coalescent we leave the asymptotic analysis of  $N^{(n)}$  and  $L^{(n)}$  for future work.

### 3. The Dirichlet coalescent

Let  $X := (X_1, \dots, X_N)$  be symmetric Dirichlet distributed with parameters  $N \in \mathbb{N}$  and  $\alpha > 0$  and let  $X_{(1)} \geq \dots \geq X_{(N)}$  denote the order statistics of  $X$ . We consider the  $\Xi$ -coalescent when the characteristic measure  $\nu$  is the distribution of  $(X_{(1)}, \dots, X_{(N)}, 0, 0, \dots)$ . We call this coalescent the Dirichlet coalescent with parameters  $N \in \mathbb{N}$  and  $\alpha > 0$ . Note that  $\nu$  is concentrated on  $\Delta_N$ . The Dirichlet coalescent neither comes down from infinity nor stays infinite. In order



to see this we may assume without loss of generality that the Dirichlet coalescent is constructed via a Poisson point process  $(e(t))_{t \geq 0}$  as described in Appendix B of Schweinsberg (2000a) with characteristic measure  $\mu$  on  $\mathbb{N}_0^{\mathbb{N}}$  defined via (1.2). Define  $A_f := \{(n_1, n_2, \dots) \in \mathbb{N}_0^{\mathbb{N}} : \{n_1, n_2, \dots\} \text{ is finite}\}$ . Note that  $P_x(A_f) = 1$  for all  $x \in \Delta_f$ . By Schweinsberg (2000a, Lemma 41),  $T_f := \inf\{t > 0 : e(t) \in A_f\}$  is exponentially distributed with parameter  $\int_{\Delta} P_x(A_f) \nu(dx) = \int_{\Delta_f} P_x(A_f) \nu(dx) = \nu(\Delta_f) = 1$ . Thus,  $\mathbb{P}(N_t = \infty) = \mathbb{P}(T_f > t) = e^{-t} \in (0, 1)$  for all  $t > 0$ , which shows that the Dirichlet coalescent neither comes down from infinity nor stays infinite. Note that the fixation line  $L = (L_t)_{t \geq 0}$  explodes, since the coalescent does not stay infinite.

In agreement with Hsu and Shiue (1998) we use the notation  $[x|y]_n := \prod_{k=0}^{n-1} (x + ky)$  and  $(x|y)_n := \prod_{k=0}^{n-1} (x - ky)$  for  $x, y \in \mathbb{R}$  and  $n \in \mathbb{N}_0$  with the convention that empty products are equal to 1. We furthermore write  $[x]_n := [x|1]_n$  and  $(x)_n := (x|1)_n$ . The proof of the following lemma is given at the end of this section.

**Lemma 3.1** (Rates of the block counting process). *The block counting process of the Dirichlet coalescent with parameters  $N \in \mathbb{N}$  and  $\alpha > 0$  has infinitesimal rates*

$$q_{ij} = \frac{(N\alpha|\alpha)_j}{[N\alpha]_i} S_{\alpha}(i, j) \quad 1 \leq j < i < \infty, \tag{3.1}$$

where  $S_{\alpha}(i, j) := S(i, j; -1, \alpha, 0)$  are the generalized Stirling numbers defined in Hsu and Shiue (1998) satisfying the recursion  $S_{\alpha}(i+1, j) = S_{\alpha}(i, j-1) + (i+\alpha j)S_{\alpha}(i, j)$ . Alternatively,

$$q_{ij} = \frac{\binom{N}{j}}{\binom{N\alpha + i - 1}{i}} \sum_{\substack{i_1, \dots, i_j \in \mathbb{N} \\ i_1 + \dots + i_j = i}} \binom{i_1 + \alpha - 1}{i_1} \dots \binom{i_j + \alpha - 1}{i_j}, \quad 1 \leq j < i < \infty. \tag{3.2}$$

Moreover,  $q_{\infty N} = 1$ ,  $q_{\infty j} = 0$  for  $j \in \mathbb{N} \setminus \{N\}$  and, hence,  $q_{\infty \infty} = -1$ .

*Remark 3.2.* The Dirichlet coalescent is closely related to the Chinese restaurant process. Imagine a restaurant with  $N \in \mathbb{N}$  tables each of infinite capacity. Customers successively enter the restaurant. When the  $(i + 1)$ th customer arrives and  $j$  tables are already occupied (by at least one person), the customer sits at an empty table with probability  $(N\alpha - j\alpha)/(N\alpha + i)$ . This corresponds to the Chinese restaurant process (see Pitman, 2006) with  $\kappa := \alpha$  and  $\theta := N\alpha$ . Let  $K_i$  denote the number of occupied tables after the  $i$ th customer has been seated. It is easily verified by induction on  $i$  that  $K_i$  has distribution  $\mathbb{P}(K_i = j) = (N\alpha|\alpha)_j S_{\alpha}(i, j)/[N\alpha]_i$ ,  $j \in \{1, \dots, i\}$ . For  $j < i$ ,  $\mathbb{P}(K_i = j)$  coincides with  $q_{ij}$  in (3.1). Note that  $K_i$  has mean

$$\mathbb{E}(K_i) = N - N \frac{[(N-1)\alpha]_i}{[N\alpha]_i}.$$

*Remark 3.3.* The Dirichlet coalescent has total rates  $q_i := \sum_{j=1}^{i-1} q_{ij} = \sum_{j=1}^{i-1} \mathbb{P}(K_i = j) = 1 - \mathbb{P}(K_i = i) = 1 - (N\alpha|\alpha)_i/[N\alpha]_i$ . Note that  $0 = q_1 < q_2 < \dots < q_N < 1 = q_{N+1} = q_{N+2} = \dots$ . Therefore the Dirichlet coalescent serves as an example that in general the total rates of a coalescent do not need to be pairwise distinct.

*Example 3.4.* For  $\alpha = 1$  the Stirling number  $S_1(i, j)$  coincides with the Lah number  $S(i, j; -1, 1, 0) = \frac{i!}{j!} \binom{i-1}{j-1}$  and we conclude that  $q_{ij} = \binom{N}{j} \binom{i-1}{j-1} / \binom{N+i-1}{i}$ . In this

case  $K_i$  has a hypergeometric distribution with parameters  $N+i-1$ ,  $N$  and  $i$ . The total rates are  $q_i = 1 - N!(N-1)!/(N-i)!/(N+i-1)!$  for  $i \leq N$  and  $q_i = 1$  for  $i > N$ .

*Example 3.5.* For  $\alpha \rightarrow \infty$  it follows that  $q_{ij} = N^{-i}(N)_j S(i, j)$ , where the  $S(i, j)$  are the usual Stirling numbers of the second kind. In this case  $K_i$  counts the number of non-empty boxes when  $i$  balls are allocated at random to  $N$  boxes. This corresponds to the Dirac  $\Xi$ -coalescent where the measure  $\nu$  assigns its total mass 1 to the single point  $x \in \Delta$  whose first  $N$  coordinates are all equal to  $1/N$ .

*Example 3.6.* For  $\alpha \rightarrow 0$  and  $N \rightarrow \infty$  such that  $N\alpha \rightarrow \theta \in (0, \infty)$  the rates  $q_{ij}$  converge to those of the Poisson–Dirichlet coalescent with parameter  $\theta$  and  $\alpha = 0$  studied in the following Section 4.

In the following we provide the asymptotics of some functionals of the Dirichlet  $n$ -coalescent when the sample size  $n$  tends to infinity. By Theorem 2.13 and Remark 2.15,  $(N_t^{(n)}/n)_{t \geq 0}$  converges in  $D_{[0,1]}[0, \infty)$  to  $(S_t)_{t \geq 0}$  as  $n \rightarrow \infty$  and  $(L_t^{(n)}/n)_{t \geq 0}$  converges in  $D_{[1,\infty]}[0, \infty)$  to  $(1/S_t)_{t \geq 0}$ , where  $S_t := 1_{\{T_f > t\}}$  and  $T_f$  is exponentially distributed with parameter 1.

Let  $C_n$  denote the number of jumps and  $\tau_n := \inf\{t > 0 : N_t = 1\}$  denote the absorption time of the Dirichlet  $n$ -coalescent. The following lemma clarifies the asymptotics of  $C_n$  and  $\tau_n$  as  $n \rightarrow \infty$ . Its proof is given at the end of this section. In the following  $R = (r_{ij})_{i,j \in S}$  denotes the transition matrix of the jump chain of the block counting process of the Dirichlet coalescent. Note that  $r_{11} = 1$ ,  $r_{ij} = q_{ij}/q_i$  for  $i \geq 2$  and  $1 \leq j < i$ , and  $r_{ij} = 0$  otherwise.

**Lemma 3.7** (Asymptotics of the number of jumps and the absorption time). *For the Dirichlet coalescent with parameter  $N \in \mathbb{N}$  and  $\alpha > 0$  the following two statements hold.*

(i) *The number of jumps  $C_n$  converges to  $C_\infty := 1 + C_N$  in distribution as  $n \rightarrow \infty$ . The limit  $C_\infty$  has distribution  $\mathbb{P}(C_\infty = k) = r_{N1}^{(k-1)}$ ,  $1 \leq k \leq N$ , where  $r_{N1}^{(k-1)}$  is the entry in row  $N$  and column 1 of the  $(k-1)$ th power  $R^{k-1}$  of the transition matrix  $R$ .*

(ii) *The absorption time  $\tau_n$  converges to  $\tau_\infty := E + \tau_N$  in distribution as  $n \rightarrow \infty$ , where  $E$  is standard exponentially distributed and independent of  $\tau_N$ . The Laplace transform  $\psi_N$  of  $\tau_N$  can be recursively computed via  $\psi_1(\lambda) = 1$  and  $\psi_n(\lambda) = q_n/(q_n + \lambda) \sum_{k=1}^{n-1} r_{nk} \psi_k(\lambda)$ ,  $n \in \{2, \dots, N\}$ ,  $\lambda \geq 0$ .*

We now turn to the fixation line  $L = (L_t)_{t \geq 0}$  of the Dirichlet coalescent.

**Lemma 3.8** (Rates of the fixation line). *The fixation line of the Dirichlet coalescent with parameters  $N \in \mathbb{N}$  and  $\alpha > 0$  has rates*

$$\gamma_{ij} = \frac{(N\alpha|\alpha)_{i+1}}{[N\alpha]_{j+1}} S_\alpha(j, i), \quad i, j \in \mathbb{N}, i < j. \tag{3.3}$$

Moreover  $\gamma_{i\infty} = 0$  for  $i \in \{1, \dots, N-1\}$  and  $\gamma_{i\infty} = 1$  for  $i \in \{N, N+1, \dots\}$ .

*Remark 3.9.* Note that  $\gamma_{ij} = \frac{N\alpha-i\alpha}{N\alpha+j} q_{ji} = \mathbb{P}(K_j = i, K_{j+1} = i+1) = \mathbb{P}(K_j \leq i, K_{j+1} > i) = \mathbb{P}(K_j \leq i) - \mathbb{P}(K_{j+1} \leq i)$ . Summation over all  $j \in \{i+1, i+2, \dots\} \cup \{\infty\}$  shows that the fixation line has total rates  $\gamma_i = q_{i+1}$ ,  $i \in \mathbb{N}$ , in agreement with Proposition 2.5.

*Example 3.10.* For  $\alpha = 1$  we obtain  $\gamma_{ij} = ((N-i)/(N+j))q_{ji} = \binom{N-1}{i} \binom{j-1}{j-i} / \binom{N+j}{N}$ .

By duality (Theorem 2.9), for all  $i, j \in \mathbb{N}$  with  $i > j$ , the two quantities

$$q_{i, \leq j} = \sum_{k=1}^j q_{ik} = \frac{1}{[N\alpha]_i} \sum_{k=1}^j (N\alpha|_{\alpha})_j S_{\alpha}(i, j) \tag{3.4}$$

and

$$\gamma_{j, \geq i} = \sum_{\substack{k \in \mathbb{N} \cup \{\infty\} \\ k \geq i}} \gamma_{jk} = \begin{cases} (N\alpha|_{\alpha})_{j+1} \sum_{k=i}^{\infty} \frac{S_{\alpha}(k, j)}{[N\alpha]_{k+1}} & \text{if } j < N \\ \gamma_{j\infty} = 1 & \text{if } j \geq N \end{cases} \tag{3.5}$$

coincide. The equality of (3.4) and (3.5) follows alternatively from Möhle (2016+, Lemma 4.1), applied to the Markov chain  $K$ , or from Lemma 6.1 in the appendix, applied with  $a := -1$ ,  $b := \alpha$ ,  $r := 0$  and  $t := N\alpha$ . Note that  $\lim_{i \rightarrow \infty} q_{i, \leq j} = \lim_{i \rightarrow \infty} \mathbb{P}(K_i \leq j) = 0$  for all  $j < N$ , since all states  $j < N$  of the Markov chain  $K$  are transient.

In the remaining part of this section we prove Lemmata 3.1, 3.7 and 3.8.

*Proof:* (of Lemma 3.1) Since the measure  $\nu$  is concentrated on  $\Delta_N$  it follows from (2.2) and (2.3) that

$$q_{ij} = \frac{i!}{j!} \sum_{\substack{i_1, \dots, i_j \in \mathbb{N} \\ i_1 + \dots + i_j = i}} \frac{\phi(i_1, \dots, i_j)}{i_1! \dots i_j!}, \tag{3.6}$$

where

$$\phi(i_1, \dots, i_j) := \int_{\Delta} \sum_{\substack{r_1, \dots, r_j \in \mathbb{N} \\ \text{all distinct}}} x_{r_1}^{i_1} \dots x_{r_j}^{i_j} \nu(dx) = \int_{\Delta} \sum_{\substack{r_1, \dots, r_j=1 \\ \text{all distinct}}}^N x_{r_1}^{i_1} \dots x_{r_j}^{i_j} \nu(dx).$$

The function below the latter integral is symmetric with respect to  $x_1, \dots, x_N$ . Thus,

$$\phi(i_1, \dots, i_j) = \sum_{\substack{r_1, \dots, r_j=1 \\ \text{all distinct}}}^N \int x_{r_1}^{i_1} \dots x_{r_j}^{i_j} D_N(\alpha)(dx_1, \dots, dx_N) = (N)_j \mathbb{E}(X_1^{i_1} \dots X_j^{i_j}),$$

where the last equality holds, since the Dirichlet distribution  $D_N(\alpha)$  is symmetric and, hence, the integrals (over each summand) are identical. The moments of the symmetric Dirichlet distribution  $D_N(\alpha)$  are well known (see, for example, Kotz et al., 2000, p. 488) to be  $\mathbb{E}(X_1^{i_1} \dots X_j^{i_j}) = [\alpha]_{i_1} \dots [\alpha]_{i_j} / [N\alpha]_i$ , where  $i := i_1 + \dots + i_j$ . Plugging all this into (3.6) leads to

$$\begin{aligned} q_{ij} &= \frac{i!}{j!} \frac{(N)_j}{[N\alpha]_i} \sum_{\substack{i_1, \dots, i_j \in \mathbb{N} \\ i_1 + \dots + i_j = i}} \frac{[\alpha]_{i_1} \dots [\alpha]_{i_j}}{i_1! \dots i_j!} \\ &= \frac{\binom{N}{j}}{\binom{N\alpha + i - 1}{i}} \sum_{\substack{i_1, \dots, i_j \in \mathbb{N} \\ i_1 + \dots + i_j = i}} \binom{i_1 + \alpha - 1}{i_1} \dots \binom{i_j + \alpha - 1}{i_j}, \end{aligned}$$

which is (3.2). Since  $(N\alpha|_j) = (N)_j\alpha^j$  and

$$S_\alpha(i, j) = \frac{i!}{j!} \sum_{\substack{i_1, \dots, i_j \in \mathbb{N} \\ i_1 + \dots + i_j = i}} \frac{1}{i_1! \cdots i_j!} \frac{\Gamma(\alpha + i_1)}{\Gamma(\alpha + 1)} \cdots \frac{\Gamma(\alpha + i_j)}{\Gamma(\alpha + 1)}$$

it follows that the right hand sides of (3.2) and (3.1) coincide. By Proposition 2.1,  $q_{\infty j} = \nu(\Delta_j) - \nu(\Delta_{j-1}) = \delta_{j,N}$  (Kronecker symbol) for all  $j \in \mathbb{N}$ , since  $\nu(\Delta_j) = 0$  for  $j < N$  and  $\nu(\Delta_j) = 1$  for  $j \geq N$ . Moreover,  $q_{\infty \infty} = -\nu(\Delta_f) = -1$ .  $\square$

*Proof:* (of Lemma 3.7) (i) Clearly,  $(C_n)_{n \in \mathbb{N}}$  satisfies the distributional recursion  $C_n \stackrel{d}{=} 1 + C_{I_n}$  with initial condition  $C_1 = 0$ , where  $I_n$  denotes the state of the jump chain of the block counting process of the Dirichlet  $n$ -coalescent with parameters  $N \in \mathbb{N}$  and  $\alpha > 0$  after its first jump. By Lemma 3.1 and the remarks thereafter,  $I_n$  has distribution  $\mathbb{P}(I_n = k) = q_{nk}/q_n = \mathbb{P}(K_n = k)/(1 - \mathbb{P}(K_n = n))$ ,  $1 \leq k < n$ . Since  $K_n \rightarrow N$  in distribution as  $n \rightarrow \infty$  we conclude that  $I_n \rightarrow N$  in distribution as  $n \rightarrow \infty$ . Thus,  $C_n \stackrel{d}{=} 1 + C_{I_n} \rightarrow 1 + C_N =: C_\infty$  in distribution as  $n \rightarrow \infty$ . Clearly,  $\mathbb{P}(C_\infty = k) = \mathbb{P}(C_N = k - 1) = r_{N1}^{(k-1)}$  for  $1 \leq k \leq N$ , where  $r_{N1}^{(k-1)}$  is the entry in row  $N$  and column 1 of the  $(k - 1)$ th power  $R^{k-1}$  of the transition matrix  $R$  of the jump chain of the block counting process.

(ii) The sequence  $(\tau_n)_{n \in \mathbb{N}}$  satisfies the distributional recursion  $\tau_n \stackrel{d}{=} E_n + \tau_{I_n}$  with initial condition  $\tau_1 = 0$ , where  $E_n$  is independent of  $I_n$  and exponentially distributed with parameter  $q_n = 1 - \mathbb{P}(K_n = n)$ . Since  $q_n \rightarrow 1$  as  $n \rightarrow \infty$  and  $I_n \rightarrow N$  in distribution as  $n \rightarrow \infty$  we conclude that  $\tau_n \stackrel{d}{=} E_n + \tau_{I_n} \rightarrow E + \tau_N$  as  $n \rightarrow \infty$ , where  $E$  is standard exponentially distributed and independent of  $\tau_N$ . From  $\tau_n \stackrel{d}{=} E_n + \tau_{I_n}$  it follows that the Laplace transform  $\psi_N$  of  $\tau_N$  can be recursively computed via  $\psi_1(\lambda) = 1$  and  $\psi_n(\lambda) = \mathbb{E}(e^{-\lambda(E_n + \tau_{I_n})}) = \mathbb{E}(e^{-\lambda E_n}) \sum_{k=1}^{n-1} r_{nk} \mathbb{E}(e^{-\lambda \tau_k}) = q_n / (q_n + \lambda \sum_{k=1}^{n-1} r_{nk} \psi_k(\lambda))$ ,  $n \in \{2, \dots, N\}$ ,  $\lambda \geq 0$ .  $\square$

*Proof:* (of Lemma 3.8) Since  $\nu$  is concentrated on  $\Delta_N$  it follows from (2.7) and (2.8) that

$$\gamma_{ij} = \frac{j!}{i!} \sum_{\substack{j_1, \dots, j_i \in \mathbb{N} \\ j_1 + \dots + j_i = j}} \frac{\psi(j_1, \dots, j_i)}{j_1! \cdots j_i!},$$

where

$$\psi(j_1, \dots, j_i) := \int_{\Delta} \sum_{\substack{r_1, \dots, r_i = 1 \\ \text{all distinct}}}^N x_{r_1}^{j_1} \cdots x_{r_i}^{j_i} \left(1 - \sum_{k=1}^i x_{r_k}\right) \nu(dx).$$

The same arguments as in the proof of Lemma 3.1 show that

$$\begin{aligned} \psi(j_1, \dots, j_i) &= (N)_i \mathbb{E} \left( X_1^{j_1} \cdots X_i^{j_i} \left(1 - \sum_{k=1}^i X_k\right) \right) \\ &= (N)_i \left( \frac{[\alpha]_{j_1} \cdots [\alpha]_{j_i}}{[N\alpha]_j} - \sum_{k=1}^i \frac{[\alpha]_{j_1} \cdots [\alpha]_{j_k+1} \cdots [\alpha]_{j_i}}{[N\alpha]_{j+1}} \right) \\ &= (N)_i \frac{[\alpha]_{j_1} \cdots [\alpha]_{j_i}}{[N\alpha]_{j+1}} \left( (N\alpha + j) - \sum_{k=1}^i (\alpha + j_k) \right) \end{aligned}$$

$$= \frac{N\alpha - i\alpha}{N\alpha + j} (N)_i \frac{[\alpha]_{j_1} \cdots [\alpha]_{j_i}}{[N\alpha]_j} = \frac{N\alpha - i\alpha}{N\alpha + j} \phi(j_1, \dots, j_i).$$

It follows that  $\gamma_{ij} = q_{ij}(N\alpha - i\alpha)/(N\alpha + j)$  and (3.3) follows from (3.1). Moreover, by Proposition 2.5,  $\gamma_{i\infty} = \nu(\Delta_i)$  for all  $i \in \mathbb{N}$ . It remains to note that  $\nu(\Delta_i) = 0$  for  $i < N$  and  $\nu(\Delta_i) = 1$  for  $i \geq N$ .  $\square$

#### 4. The Poisson–Dirichlet coalescent

If  $\Xi(\{0\}) = 0$  and if the measure  $\nu(dx) = \Xi_0(dx)/(x, x)$  is the Poisson–Dirichlet distribution with parameters  $0 \leq \alpha < 1$  and  $\theta > -\alpha$  then the  $\Xi$ -coalescent is called the two-parameter Poisson–Dirichlet coalescent (Möhle, 2010). Note that  $\nu$  is concentrated on  $\Delta^* \setminus \Delta_f$ . Since  $\int_{\Delta} |x| \nu(dx) = \nu(\Delta^*) = 1 < \infty$  it follows that this coalescent has dust and hence cannot come down from infinity. Since  $\Xi(\Delta_f) = 0$ , this coalescent stays infinite, and, hence,  $L$  does not explode. The associated block counting process has rates (see Möhle, 2010, p. 2170)

$$q_{ij} = c_{j,\alpha,\theta} \frac{\Gamma(\theta + \alpha j)}{\Gamma(\theta + i)} s_{\alpha}(i, j), \quad i > j, \tag{4.1}$$

where

$$c_{j,\alpha,\theta} := \prod_{k=1}^j \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(1 - \alpha)\Gamma(\theta + k\alpha)}$$

and

$$s_{\alpha}(i, j) := \frac{i!}{j!} \sum_{\substack{i_1, \dots, i_j \in \mathbb{N} \\ i_1 + \dots + i_j = i}} \frac{\Gamma(i_1 - \alpha) \cdots \Gamma(i_j - \alpha)}{i_1! \cdots i_j!}$$

is a kind of generalized absolute Stirling number of the first kind satisfying the recursion  $s_{\alpha}(i + 1, j) = \Gamma(1 - \alpha)s_{\alpha}(i, j - 1) + (i - \alpha j)s_{\alpha}(i, j)$ . More precisely,  $s_{\alpha}(i, j)/(\Gamma(1 - \alpha))^j$  coincides with the generalized Stirling number  $S(i, j; -1, -\alpha, 0)$  as defined in Hsu and Shiue (1998). For  $\alpha = 0$  (and hence  $\theta > 0$ ) the rate  $q_{ij}$  reduces to

$$q_{ij} = \theta^j \frac{\Gamma(\theta)}{\Gamma(\theta + i)} s(i, j), \quad i > j,$$

where  $s(i, j) := S(i, j; -1, 0, 0)$  are the (usual) absolute Stirling numbers of the first kind. For  $\theta = 0$  (and hence  $0 < \alpha < 1$ ) we obtain

$$q_{ij} = \frac{\alpha^{j-1}}{(\Gamma(1 - \alpha))^j} \frac{(j - 1)!}{(i - 1)!} s_{\alpha}(i, j) = \alpha^{j-1} \frac{(j - 1)!}{(i - 1)!} S(i, j; -1, -\alpha, 0), \quad i > j.$$

In order to compute the total rates of the block counting process of the two-parameter Poisson–Dirichlet coalescent we proceed as follows. Let  $K = (K_n)_{n \in \mathbb{N}_0}$  be a Markov chain with state space  $\mathbb{N}_0$ ,  $K_0 := 0$ ,  $K_1 := 1$  and transition probabilities  $p_k(n) := \mathbb{P}(K_{n+1} = k + 1 \mid K_n = k) := (\theta + \alpha k)/(\theta + n)$  and  $\mathbb{P}(K_{n+1} = k \mid K_n = k) = 1 - p_k(n)$  for  $n \in \mathbb{N}$  and  $k \in \{1, \dots, n\}$ . Note that  $1 \leq K_n \leq n$ ,  $n \in \mathbb{N}$ . As for the Dirichlet coalescent one may interpret  $K_n$  as the number of occupied tables in a particular Chinese restaurant process. When customer  $n + 1$  enters the restaurant and  $k$  tables are already occupied, he sits at an empty table with probability  $p_k(n)$ .

In the following it is verified by induction on  $n \in \mathbb{N}$  that  $K_n$  has distribution (see also Pitman, 2006, p. 65, Eq. (3.11))

$$\mathbb{P}(K_n = k) = c_{k,\alpha,\theta} \frac{\Gamma(\theta + \alpha k)}{\Gamma(\theta + n)} s_\alpha(n, k), \quad k \in \mathbb{N}_0.$$

For  $n = 1$  this is obvious, since  $K_1 = 1$ ,  $c_{1,\alpha,\theta} = \Gamma(\theta + 1)/\Gamma(1 - \alpha)/\Gamma(\theta + \alpha)$  and  $s_\alpha(1, 1) = \Gamma(1 - \alpha)$ . The induction step from  $n \in \mathbb{N}$  to  $n + 1$  works as follows. By the Markov property,  $\mathbb{P}(K_{n+1} = k) = p_{k-1}(n) \mathbb{P}(K_n = k - 1) + (1 - p_k(n)) \mathbb{P}(K_n = k)$ . By induction,

$$\begin{aligned} p_{k-1}(n) \mathbb{P}(K_n = k - 1) &= \frac{\theta + \alpha(k - 1)}{\theta + n} c_{k-1,\alpha,\theta} \frac{\Gamma(\theta + \alpha(k - 1))}{\Gamma(\theta + n)} s_\alpha(n, k - 1) \\ &= \frac{\Gamma(\theta + 1 + \alpha(k - 1))}{\Gamma(\theta + n + 1)} c_{k-1,\alpha,\theta} s_\alpha(n, k - 1) \\ &= \frac{\Gamma(\theta + k\alpha)\Gamma(1 - \alpha)}{\Gamma(\theta + n + 1)} c_{k,\alpha,\theta} s_\alpha(n, k - 1) \end{aligned}$$

and

$$\begin{aligned} (1 - p_k(n)) \mathbb{P}(K_n = k) &= \frac{n - \alpha k}{\theta + n} c_{k,\alpha,\theta} \frac{\Gamma(\theta + k\alpha)}{\Gamma(\theta + n)} s_\alpha(n, k) \\ &= \frac{\Gamma(\theta + k\alpha)}{\Gamma(\theta + n + 1)} c_{k,\alpha,\theta} (n - \alpha k) s_\alpha(n, k). \end{aligned}$$

Summation of these two terms yields

$$\begin{aligned} \mathbb{P}(K_{n+1} = k) &= \frac{\Gamma(\theta + \alpha k)}{\Gamma(\theta + n + 1)} c_{k,\alpha,\theta} (\Gamma(1 - \alpha) s_\alpha(n, k - 1) + (n - \alpha k) s_\alpha(n, k)) \\ &= \frac{\Gamma(\theta + \alpha k)}{\Gamma(\theta + n + 1)} c_{k,\alpha,\theta} s_\alpha(n + 1, k), \end{aligned}$$

which completes the induction. As a consequence, the block counting process of the two-parameter Poisson–Dirichlet coalescent has total rates

$$\begin{aligned} q_i &= \sum_{j=1}^{i-1} q_{ij} = \sum_{j=1}^{i-1} \mathbb{P}(K_i = j) = 1 - \mathbb{P}(K_i = i) \\ &= 1 - c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + i)} s_\alpha(i, i) = 1 - \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + i)} \prod_{k=1}^i \frac{\Gamma(\theta + 1 + (k - 1)\alpha)}{\Gamma(\theta + k\alpha)}, \end{aligned}$$

$i \in \mathbb{N}$ . For  $\alpha = 0$  the total rates reduce to  $q_i = 1 - \theta^i \Gamma(\theta)/\Gamma(\theta + i)$ ,  $i \in \mathbb{N}$ , and for  $\theta = 0$  we obtain  $q_i = 1 - \alpha^{i-1}$ ,  $i \in \mathbb{N}$ .

Let us now turn to the fixation line. The rates  $\gamma_{ij}$ ,  $i < j$ , of the fixation line are obtained in a similar vein as the rates  $q_{ij}$  as follows. Since the measure  $\nu$  is concentrated on  $\Delta^*$  it follows from (2.7) and (2.8) that

$$\gamma_{ij} = \frac{j!}{i!} \sum_{\substack{j_1, \dots, j_i \in \mathbb{N} \\ j_1 + \dots + j_i = j}} \frac{I(j_1, \dots, j_i)}{j_1! \cdots j_i!},$$

where

$$I(j_1, \dots, j_i) := \int_{\Delta} \sum_{\substack{r_1, \dots, r_i \in \mathbb{N} \\ \text{all distinct}}} x_{r_1}^{j_1} \cdots x_{r_i}^{j_i} \left( 1 - \sum_{k=1}^i x_{r_k} \right) \nu(dx).$$

By [Handa \(2009, Eq. \(2.1\)\)](#),

$$I(j_1, \dots, j_i) = \int_{\mathbb{R}^i} x_1^{j_1} \cdots x_i^{j_i} \left(1 - \sum_{k=1}^i x_k\right) \mu_i(dx_1, \dots, dx_i),$$

where  $\mu_i$  denotes the  $i$ th correlation measure associated with the Poisson–Dirichlet distribution. The density (correlation function) of  $\mu_i$  is explicitly known (see, for example, [Handa, 2009, Theorem 2.1](#)) and we obtain

$$I(j_1, \dots, j_i) = c_{i,\alpha,\theta} \int_{\Delta_i} x_1^{j_1-\alpha-1} \cdots x_i^{j_i-\alpha-1} \left(1 - \sum_{k=1}^i x_k\right)^{\theta+\alpha i} dx_1 \cdots dx_i.$$

The last integral is known (Liouville’s integration formula), and it follows that

$$I(j_1, \dots, j_i) = c_{i,\alpha,\theta} \frac{\Gamma(j_1 - \alpha) \cdots \Gamma(j_i - \alpha) \Gamma(\theta + \alpha i + 1)}{\Gamma(\theta + j + 1)},$$

$j := j_1 + \cdots + j_i > i$ . Plugging this expression into the above formula for  $\gamma_{ij}$  leads to

$$\begin{aligned} \gamma_{ij} &= c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i + 1)}{\Gamma(\theta + j + 1)} \frac{j!}{i!} \sum_{\substack{j_1, \dots, j_i \in \mathbb{N} \\ j_1 + \cdots + j_i = j}} \frac{\Gamma(j_1 - \alpha) \cdots \Gamma(j_i - \alpha)}{j_1! \cdots j_i!} \\ &= c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i + 1)}{\Gamma(\theta + j + 1)} s_\alpha(j, i), \quad i < j. \end{aligned} \tag{4.2}$$

For  $\alpha = 0$  the rate  $\gamma_{ij}$  reduces to

$$\gamma_{ij} = \theta^i \frac{\Gamma(\theta + 1)}{\Gamma(\theta + j + 1)} s(j, i), \quad i < j$$

whereas for  $\theta = 0$  we obtain

$$\gamma_{ij} = \frac{\alpha^i}{(\Gamma(1 - \alpha))^i} \frac{i!}{j!} s_\alpha(j, i) = \alpha^i \frac{i!}{j!} S(j, i; -1, -\alpha, 0), \quad i < j.$$

In particular,

$$q_{i,\leq j} = \sum_{k=1}^j q_{ik} = \frac{1}{\Gamma(\theta + i)} \sum_{k=1}^j c_{k,\alpha,\theta} \Gamma(\theta + \alpha k) s_\alpha(i, k), \quad i > j, \tag{4.3}$$

and

$$\gamma_{j,\geq i} = \sum_{k=i}^\infty \gamma_{jk} = c_{j,\alpha,\theta} \Gamma(\theta + \alpha j + 1) \sum_{k=i}^\infty \frac{s_\alpha(k, j)}{\Gamma(\theta + k + 1)}, \quad i > j. \tag{4.4}$$

The two expressions (4.3) and (4.4) are equal by duality (Theorem 2.9). The equality of (4.3) and (4.4) also follows from [Möhle \(2016+, Lemma 4.1\)](#), applied to the Markov chain  $K$ . Alternatively, one may apply Lemma 6.1 in the appendix with  $a := -1$ ,  $b := -\alpha$ ,  $r := 0$  and  $t := \theta$ . Note however that formally the case  $\theta = 0$  is not covered by Lemma 6.1. As a last option one may prove the equality of (4.3) and (4.4) directly (using the recursion for generalized Stirling numbers) as

follows. We have

$$\begin{aligned}
& q_{ki} - q_{k+1,i} \\
&= c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + k)} s_\alpha(k, i) - c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + k + 1)} s_\alpha(k + 1, i) \\
&= c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + k)} s_\alpha(k, i) \\
&\quad - c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + k + 1)} (\Gamma(1 - \alpha) s_\alpha(k, i - 1) + (k - \alpha i) s_\alpha(k, i)) \\
&= c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + k + 1)} s_\alpha(k, i) ((\theta + k) - (k - \alpha i)) \\
&\quad - c_{i,\alpha,\theta} \frac{\Gamma(\theta + \alpha i)}{\Gamma(\theta + k + 1)} \Gamma(1 - \alpha) s_\alpha(k, i - 1) \\
&= c_{i,\alpha,\theta} \frac{\Gamma(\theta + 1 + \alpha i)}{\Gamma(\theta + k + 1)} s_\alpha(k, i) - c_{i-1,\alpha,\theta} \frac{\Gamma(\theta + 1 + (i - 1)\alpha)}{\Gamma(\theta + k + 1)} s_\alpha(k, i - 1).
\end{aligned}$$

Summation over all  $i \in \{1, \dots, j\}$  yields

$$q_{k,\leq j} - q_{k+1,\leq j} = c_{j,\alpha,\theta} \frac{\Gamma(\theta + 1 + \alpha j)}{\Gamma(\theta + k + 1)} s_\alpha(k, j).$$

Another summation over all  $k \geq i$  yields

$$q_{i,\leq j} = c_{j,\alpha,\theta} \Gamma(\theta + 1 + \alpha j) \sum_{k=i}^{\infty} \frac{s_\alpha(k, j)}{\Gamma(\theta + k + 1)},$$

which shows that (4.3) and (4.4) coincide.

## 5. Proofs

*Proof:* (of Proposition 2.1) The formulas (2.2) and (2.4) for the rates  $q_{ij}$  and the total rates  $q_i$  of the block counting process are known from the literature (Freund and Möhle, 2009, Eqs. (1.2) and (1.3)). For given  $x = (x_r)_{r \in \mathbb{N}} \in \Delta$  the block counting process jumps from  $\infty$  to  $j \in \mathbb{N}$  if and only if  $x_1 + \dots + x_j = 1$  and  $x_1, \dots, x_j > 0$ , i.e. if and only if  $x \in \Delta_j \setminus \Delta_{j-1}$  with the convention  $\Delta_0 := \emptyset$ . Integration with respect to  $\nu$  yields  $q_{\infty j} = \nu(\Delta_j \setminus \Delta_{j-1}) = \nu(\Delta_j) - \nu(\Delta_{j-1})$  for all  $j \in \mathbb{N}$ . Since the generator  $Q = (q_{ij})_{i,j \in S}$  is conservative, it follows that  $q_{\infty \infty} = -\sum_{j \in \mathbb{N}} q_{\infty j} = -\sum_{j \in \mathbb{N}} (\nu(\Delta_j) - \nu(\Delta_{j-1})) = -\lim_{n \rightarrow \infty} \nu(\Delta_n) = -\nu(\Delta_f)$ .  $\square$

*Proof:* (of Proposition 2.5) We generalize the proof of Lemma 2.3 in Hénard (2015). Recall the pathwise definition of the fixation line based on the lookdown construction provided in the introduction. Assume first that  $\Xi(\{0\}) = 0$ . The fixation line jumps from  $i \in \mathbb{N}$  to  $j \in \mathbb{N}$  with  $j > i$  if and only if there exists  $k \in \{1, \dots, i\}$  and  $1 \leq r_1 < \dots < r_k$  such that

- (i) exactly  $i - k$  of the individuals  $1, \dots, j$  belong to  $J_0 := \bigcup_{r \in \mathbb{N}} J_r$ ,
- (ii) for every  $l \in \{1, \dots, k\}$  at least one of the individuals  $1, \dots, j$  belongs to  $J_{r_l}$  and
- (iii) the individual  $j + 1$  does not belong to  $J_{r_1} \cup \dots \cup J_{r_k}$ .



For fixed  $x = (x_r)_{r \in \mathbb{N}} \in \Delta$  this event has probability

$$\binom{j}{i-k} (1 - |x|)^{i-k} \sum_{\substack{i_1, \dots, i_k \in \mathbb{N} \\ i_1 + \dots + i_k = j - (i-k)}} \frac{(j - (i-k))!}{i_1! \dots i_k!} x_{r_1}^{i_1} \dots x_{r_k}^{i_k} \left(1 - \sum_{l=1}^k x_{r_l}\right).$$

Summing this probability over all  $k \in \{1, \dots, i\}$  and  $1 \leq r_1 < \dots < r_k$  and integrating with respect to the law  $\nu$  yields  $\gamma_{ij} = \int_{\Delta} \sum_{k=1}^i g_{ijk}(x) \nu(dx)$  with  $g_{ijk}(x)$  defined in (2.8). From the definition of  $Y(\cdot, x)$  in (2.1) it follows that  $\sum_{k=1}^i g_{ijk}(x) = \mathbb{P}(Y(j, x) = i, Y(j+1, x) = i+1)$ .

If  $\Xi(\{0\}) > 0$  then the rate  $\gamma_{ij}$  increases by  $\Xi(\{0\}) \binom{j}{2} \delta_{j,i+1}$ , since  $\binom{j}{2} \delta_{j,i+1}$  is the rate at which the fixation line of the Kingman coalescent jumps from  $i$  to  $j$ . Thus (2.7) is established. Similarly, given  $x = (x_r)_{r \in \mathbb{N}} \in \Delta$ , the fixation line jumps from  $i \in \mathbb{N}$  to  $\infty$  if and only if  $x_1 + \dots + x_i = 1$ , i.e. if and only if  $x \in \Delta_i$ . Integration with respect to  $\nu$  yields  $\gamma_{i\infty} = \nu(\Delta_i)$ ,  $i \in \mathbb{N}$ . Clearly,  $\gamma_{\infty\infty} = 0$ , since the state  $\infty$  is absorbing.

It remains to determine the total rates  $\gamma_i$ ,  $i \in \mathbb{N}$ . From the definition of  $Y(\cdot, x)$  in (2.1) via the paintbox construction it follows that  $Y(j+1, x) - Y(j, x) \in \{0, 1\}$  for all  $j \in \mathbb{N}$  and  $x \in \Delta$ . Thus, for all  $i, j \in \mathbb{N}$ ,  $\mathbb{P}(Y(j, x) = i, Y(j+1, x) = i+1) = \mathbb{P}(Y(j, x) \leq i, Y(j+1, x) > i) = \mathbb{P}(Y(j, x) \leq i) - \mathbb{P}(Y(j+1, x) \leq i)$ . Summation over all  $j \in \mathbb{N}$  with  $j > i$  yields

$$\begin{aligned} \sum_{\substack{j \in \mathbb{N} \\ j > i}} \mathbb{P}(Y(j, x) = i, Y(j+1, x) = i+1) &= \sum_{\substack{j \in \mathbb{N} \\ j > i}} (\mathbb{P}(Y(j, x) \leq i) - \mathbb{P}(Y(j+1, x) \leq i)) \\ &= \mathbb{P}(Y(i+1, x) \leq i) - \lim_{k \rightarrow \infty} \mathbb{P}(Y(k, x) \leq i) \\ &= \begin{cases} \mathbb{P}(Y(i+1, x) \leq i) & \text{if } x \in \Delta \setminus \Delta_i, \\ 0 & \text{if } x \in \Delta_i, \end{cases} \end{aligned}$$

since  $\mathbb{P}(Y(k, x) \leq i) \rightarrow 0$  as  $k \rightarrow \infty$  if  $x \in \Delta \setminus \Delta_i$  and  $\mathbb{P}(Y(k, x) \leq i) = 1$  for all  $k \in \mathbb{N}$  if  $x \in \Delta_i$ . Thus, the total rates of the fixation line are

$$\begin{aligned} \gamma_i &= \gamma_{i\infty} + \sum_{\substack{j \in \mathbb{N} \\ j > i}} \gamma_{ij} \\ &= \gamma_{i\infty} + \Xi(\{0\}) \binom{i+1}{2} + \int_{\Delta \setminus \Delta_i} \mathbb{P}(Y(i+1, x) \leq i) \nu(dx) \\ &= \Xi(\{0\}) \binom{i+1}{2} + \int_{\Delta} \mathbb{P}(Y(i+1, x) \leq i) \nu(dx), \quad i \in \mathbb{N}. \end{aligned}$$

A comparison with the total rate  $q_i$  of the block counting process shows that  $\gamma_i = q_{i+1}$ ,  $i \in \mathbb{N}$ . □

*Proof:* (of Theorem 2.9) From  $Y(k+1, x) - Y(k, x) \in \{0, 1\}$  for all  $k \in \mathbb{N}$  and  $x \in \Delta$  we conclude that

$$\begin{aligned} \mathbb{P}(Y(k, x) = j, Y(k+1, x) = j+1) &= \mathbb{P}(Y(k, x) \leq j, Y(k+1, x) > j) \\ &= \mathbb{P}(Y(k, x) \leq j) - \mathbb{P}(Y(k+1, x) \leq j) \end{aligned}$$

for all  $j, k \in \mathbb{N}$  and  $x \in \Delta$ . Integration with respect to  $\nu$  and taking the formula (2.7) for the rates of the fixation line into account, it follows for all  $j, k \in \mathbb{N}$  with  $j < k$  that

$$\begin{aligned} \gamma_{jk} &= \Xi(\{0\}) \binom{k}{2} \delta_{k,j+1} + \int_{\Delta} \mathbb{P}(Y(k, x) = j, Y(k + 1, x) = j + 1) \nu(dx) \\ &= \Xi(\{0\}) \binom{k}{2} \delta_{j,k-1} + \int_{\Delta} (\mathbb{P}(Y(k, x) \leq j) - \mathbb{P}(Y(k + 1, x) \leq j)) \nu(dx) \\ &= \sum_{l=1}^j \left( \Xi(\{0\}) \binom{k}{2} \delta_{l,k-1} + \int_{\Delta} \mathbb{P}(Y(k, x) = l) \nu(dx) \right. \\ &\quad \left. - \Xi(\{0\}) \binom{k+1}{2} \delta_{l,k} - \int_{\Delta} \mathbb{P}(Y(k + 1, x) = l) \nu(dx) \right) \\ &= \sum_{l=1}^j (q_{kl} - q_{k+1,l}) = q_{k,\leq j} - q_{k+1,\leq j}, \quad j, k \in \mathbb{N}, j < k, \end{aligned}$$

where the second last equality holds by (2.2). Let  $i, j \in \mathbb{N}$  with  $i > j$ . Summing over all  $k \in \mathbb{N}$  with  $k \geq i$  yields

$$\sum_{\substack{k \in \mathbb{N} \\ k \geq i}} \gamma_{jk} = \sum_{\substack{k \in \mathbb{N} \\ k \geq i}} (q_{k,\leq j} - q_{k+1,\leq j}) = q_{i,\leq j} - \lim_{k \rightarrow \infty} q_{k,\leq j} = q_{i,\leq j} - \nu(\Delta_j).$$

The last equality holds since, by (2.2),

$$\begin{aligned} q_{k,\leq j} &= \sum_{l=1}^j q_{kl} = \Xi(\{0\}) \binom{k}{2} \delta_{j,k-1} + \int_{\Delta} \mathbb{P}(Y(k, x) \leq j) \nu(dx) \\ &\rightarrow \int_{\Delta} 1_{\Delta_j}(x) \nu(dx) = \nu(\Delta_j) \end{aligned}$$

as  $k \rightarrow \infty$  by dominated convergence. Note that  $\mathbb{P}(Y(k, x) \leq j) \leq \mathbb{P}(Y(j + 1, x) \leq j)$  for all  $k > j$  and that the dominating map  $x \mapsto \mathbb{P}(Y(j + 1, x) \leq j)$  is  $\nu$ -integrable.

Since  $\gamma_{j\infty} = \nu(\Delta_j)$  it follows that

$$q_{i,\leq j} = \sum_{\substack{k \in S \\ k \geq i}} \gamma_{jk} = \gamma_{j,\geq i} \tag{5.1}$$

for all  $i, j \in \mathbb{N}$  with  $i > j$ . Eq. (5.1) holds as well for  $i, j \in \mathbb{N}$  with  $i \leq j$  since in this case both sides in (5.1) are equal to 0. Moreover,  $q_{i,\leq \infty} = 0 = \gamma_{\infty,\geq i}$  for all  $i \in S$  and  $q_{\infty,\leq j} = \nu(\Delta_j) = \gamma_{j\infty} = \gamma_{j,\geq \infty}$  for all  $j \in \mathbb{N}$ . Thus, (5.1) holds for all  $i, j \in S$ .

Let  $Q = (q_{ij})_{i,j \in S}$  and  $\Gamma = (\gamma_{ij})_{i,j \in S}$  denote the generator matrices of  $N$  and  $L$  respectively and let  $H = (h_{ij})_{i,j \in S}$  denote the matrix with entries  $h_{ij} := 1$  for  $i \leq j$  and  $h_{ij} := 0$  for  $i > j$ . Since  $(QH)_{ij} = \sum_{k \in S, k \leq j} q_{ik} h_{kj} = \sum_{k \in S, k \leq j} q_{ik} = q_{i,\leq j}$  and  $(H\Gamma^T)_{ij} = \sum_{k \in S, k \geq i} h_{ik} \gamma_{jk} = \sum_{k \in S, k \geq i} \gamma_{jk} = \gamma_{j,\geq i}$  it follows that  $QH = H\Gamma^T$ . It follows that  $Q^k H = H(\Gamma^T)^k$  for all  $k \in \mathbb{N}$  and, hence,  $e^{tQ} H = H(e^{t\Gamma})^T$  for all  $t \geq 0$ . Since  $e^{tQ}$  and  $e^{t\Gamma}$  are the transition matrices of the block counting process  $N = (N_t)_{t \geq 0}$  and the fixation line  $L = (L_t)_{t \geq 0}$  respectively, this shows that  $N$  is Siegmund dual to  $L$  with respect to the kernel  $H$ .  $\square$

*Remark 5.1.* For  $\Lambda$ -coalescents Lemma 2.1 of [Hénard \(2015\)](#) essentially states the Siegmund duality of  $N$  and  $L$  and Lemma 2.4 of [Hénard \(2015\)](#) is a reformulation of this duality in terms of the generators of  $N$  and  $L$ .

*Proof:* (of Theorem 2.13) For  $n \in \mathbb{N}$  and  $i \in \{1, \dots, n\}$  let  $\tau_{n,i} := \inf\{t > 0 : i \text{ is not a singleton of } \Pi_t^{(n)}\}$  denote the length of the  $i$ th external branch of  $\Pi^{(n)}$ . For every  $i \in \mathbb{N}$  the sequence  $(\tau_{n,i})_{n \geq i}$  is non-increasing in  $n$  with  $\tau_{n,i} \searrow \tau_i$  almost surely as  $n \rightarrow \infty$ , where  $\tau_i := \inf\{t > 0 : \{i\} \text{ is not a singleton of } \Pi_t\}$  denotes the length of the  $i$ th external branch of  $\Pi$ . Let  $t_1, \dots, t_k \geq 0$ . Given  $S_{t_1}, \dots, S_{t_k}$  the conditional probability that  $i$  is still a singleton at time  $t_i$  for all  $i \in \{1, \dots, k\}$  is  $S_{t_1} \cdots S_{t_k}$ . Thus,  $\mathbb{P}(\tau_1 > t_1, \dots, \tau_k > t_k) = \mathbb{E}(S_{t_1} \cdots S_{t_k})$ . In particular,  $\mathbb{P}(\tau_1 > t, \dots, \tau_k > t) = \mathbb{E}(S_t^k)$ .

*Proof of part a).* For  $n \in \mathbb{N}$  and  $t \geq 0$  decompose  $N_t^{(n)} = E_t^{(n)} + I_t^{(n)}$ , where  $E_t^{(n)} := \sum_{i=1}^n 1_{\{\tau_{n,i} > t\}}$  and  $I_t^{(n)} := N_t^{(n)} - E_t^{(n)}$  denotes the number of singleton and non-singleton blocks of  $\Pi_t^{(n)}$  respectively. We think of  $E_t^{(n)}$  and  $I_t^{(n)}$  as the number of ‘external’ and ‘internal’ blocks of  $\Pi_t^{(n)}$  and proceed similarly as in the proof of Theorem 3 of [Möhle \(2010\)](#). For  $t \geq 0$  and  $n, k \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{E}((E_t^{(n)})^k) &= \mathbb{E}((1_{\{\tau_{n,1} > t\}} + \cdots + 1_{\{\tau_{n,n} > t\}})^k) \\ &= \sum_{\substack{k_1, \dots, k_n \in \mathbb{N}_0 \\ k_1 + \dots + k_n = k}} \frac{k!}{k_1! \cdots k_n!} \mathbb{E}(1_{\{\tau_{n,1} > t\}}^{k_1} \cdots 1_{\{\tau_{n,n} > t\}}^{k_n}) \\ &= \sum_{j=1}^k \binom{n}{j} \sum_{\substack{k_1, \dots, k_j \in \mathbb{N} \\ k_1 + \dots + k_j = k}} \mathbb{E}(1_{\{\tau_{n,1} > t\}}^{k_1} \cdots 1_{\{\tau_{n,j} > t\}}^{k_j}), \end{aligned}$$

where the last equality holds since the random variables  $\tau_{n,i}$ ,  $i \in \{1, \dots, n\}$ , are exchangeable. Thus,

$$\mathbb{E}((E_t^{(n)})^k) = \sum_{j=1}^k (n)_j S(k, j) \mathbb{P}(\tau_{n,1} > t, \dots, \tau_{n,j} > t), \quad t \geq 0, n, k \in \mathbb{N}, \quad (5.2)$$

where  $(n)_j := n(n-1) \cdots (n-j+1)$  and  $S(\cdot, \cdot)$  denote the Stirling numbers of the second kind. Dividing by  $n^k$ , letting  $n \rightarrow \infty$  and noting that  $(n)_j/n^k \rightarrow \delta_{j,k}$  it follows that

$$\lim_{n \rightarrow \infty} \mathbb{E}((E_t^{(n)}/n)^k) = \mathbb{P}(\tau_1 > t, \dots, \tau_k > t) = \mathbb{E}(S_t^k), \quad k \in \mathbb{N}. \quad (5.3)$$

Since  $0 \leq E_t^{(n)}/n \leq 1$  and  $0 \leq S_t \leq 1$  the convergence (5.3) of moments implies the convergence  $E_t^{(n)}/n \rightarrow S_t$  in distribution as  $n \rightarrow \infty$ . In order to show that  $N_t^{(n)}/n \rightarrow S_t$  in distribution as  $n \rightarrow \infty$  it remains to verify that  $I_t^{(n)}/n \rightarrow 0$  in distribution as  $n \rightarrow \infty$ . In the following it is verified that the latter convergence even holds in  $L^1$ . Each internal branch is generated by a collision. Thus, if  $C_n$  denotes the total number of collisions, the inequality  $\mathbb{E}(I_t^{(n)}) \leq \mathbb{E}(C_n)$  holds. The assumption that the coalescent has dust ensures that  $C_n/n \rightarrow 0$  in  $L^1$  by Lemma 4.1 of [Freund and Möhle \(2009\)](#). Thus,  $I_t^{(n)}/n \rightarrow 0$  in  $L^1$  as  $n \rightarrow \infty$ , which yields the desired convergence  $N_t^{(n)}/n \rightarrow S_t$  in distribution as  $n \rightarrow \infty$ . Thus, the convergence of the one-dimensional distributions is established.

Let us now turn to the proof of the convergence in  $D_{[0,1]}[0, \infty)$ . Let  $(T_t^{(n)})_{t \geq 0}$  and  $(T_t)_{t \geq 0}$  denote the semigroups of  $(N_t^{(n)}/n)_{t \geq 0}$  and  $(S_t)_{t \geq 0}$  respectively. By [Ethier and Kurtz \(1986, p. 172, Theorem 2.11\)](#), applied with state spaces  $E := [0, 1]$  and  $E_n := \{j/n : j \in \{1, \dots, n\}\}$ ,  $n \in \mathbb{N}$ , and with the maps  $\eta_n : E_n \rightarrow E$  and  $\pi_n : B(E) \rightarrow B(E_n)$  defined via  $\eta_n(x) := x$  for all  $x \in E_n$  and  $\pi_n f(x) := f(x)$  for all  $f \in B(E)$  and all  $x \in E_n$ , it suffices to verify that for every  $t \geq 0$  and  $f \in C(E)$ ,

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} |T_t^{(n)} \pi_n f(x) - \pi_n T_t f(x)| = 0.$$

Obviously,  $T_t^{(n)} \pi_n f(x) = \mathbb{E}(\pi_n f(N_{s+t}^{(n)}/n) | N_s^{(n)}/n = x) = \mathbb{E}(f(N_t^{(nx)}/n))$  and  $\pi_n T_t f(x) = T_t f(x) = \mathbb{E}(f(xS_t))$ . Thus, we have to verify that

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} |\mathbb{E}(f(N_t^{(nx)}/n) - \mathbb{E}(f(xS_t)))| = 0.$$

Since the polynomials are dense in  $C(E)$  it suffices to verify the latter equation for monomials  $f(x) = x^k$ , so we have to prove that

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} |\mathbb{E}((N_t^{(nx)})^k)/n^k - x^k \mathbb{E}(S_t^k)| = 0, \quad k \in \mathbb{N}, t \geq 0.$$

Using the decomposition  $N_t^{(nx)} = E_t^{(nx)} + I_t^{(nx)}$  and the facts that  $I_t^{(nx)} \leq C_{nx} \leq C_n$  and that  $C_n/n \rightarrow 0$  in  $L^1$  (see [Lemma 4.1 of Freund and Möhle, 2009](#)), it suffices to show that

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} |\mathbb{E}((E_t^{(nx)})^k)/n^k - x^k \mathbb{E}(S_t^k)| = 0.$$

By [\(5.2\)](#),  $\mathbb{E}((E_t^{(nx)})^k) = \sum_{j=1}^k (nx)_j S(k, j) \mathbb{P}(\tau_{nx,1} > t, \dots, \tau_{nx,j} > t)$ . Thus, it suffices to verify that

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} \left| \frac{(nx)_k}{n^k} \mathbb{P}(\tau_{nx,1} > t, \dots, \tau_{nx,k} > t) - x^k \mathbb{E}(S_t^k) \right| = 0.$$

Since  $(nx)_k/n^k \rightarrow x^k$  as  $n \rightarrow \infty$  uniformly on  $[0, 1]$  it remains to prove that

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} x^k |\mathbb{P}(\tau_{nx,1} > t, \dots, \tau_{nx,k} > t) - \mathbb{E}(S_t^k)| = 0.$$

This is seen as follows. Choose a sequence  $(\varepsilon_n)_{n \in \mathbb{N}}$  satisfying  $\varepsilon_n \rightarrow 0$  and  $n\varepsilon_n \rightarrow \infty$  (for example  $\varepsilon_n := n^{-1/2}$ ) and distinguish the two cases  $x \in E_n \cap [0, \varepsilon_n]$  and  $x \in E_n \cap (\varepsilon_n, 1]$ . Clearly,

$$\sup_{x \in E_n \cap [0, \varepsilon_n]} x^k |\mathbb{P}(\tau_{nx,1} > t, \dots, \tau_{nx,k} > t) - \mathbb{E}(S_t^k)| \leq 2\varepsilon_n^k \rightarrow 0, \quad n \rightarrow \infty.$$

Moreover, since  $p_{t,k}(m) := \mathbb{P}(\tau_{m,1} > t, \dots, \tau_{m,k} > t)$  is non-increasing in  $m$  ( $\geq k$ ) and, hence,  $p_{t,k}(m) \geq \mathbb{P}(\tau_1 > t, \dots, \tau_k > t) = \mathbb{E}(S_t^k)$  for all  $m \geq k$ , it follows for all  $n \in \mathbb{N}$  with  $n\varepsilon_n \geq k$  that

$$\begin{aligned} & \sup_{x \in E_n \cap (\varepsilon_n, 1]} x^k |\mathbb{P}(\tau_{nx,1} > t, \dots, \tau_{nx,k} > t) - \mathbb{E}(S_t^k)| \\ & \leq \sup_{x \in E_n \cap (\varepsilon_n, 1]} |p_{t,k}(nx) - \mathbb{E}(S_t^k)| \leq p_{t,k}(\lfloor n\varepsilon_n \rfloor) - \mathbb{E}(S_t^k) \rightarrow 0 \end{aligned}$$

as  $n \rightarrow \infty$ , where  $\lfloor n\varepsilon_n \rfloor := \max\{z \in \mathbb{Z} : z \leq n\varepsilon_n\}$ . The proof of part a) is complete.

*Proof of part b).* We have to verify that  $(L_t^{(n)}/n)_{t \geq 0}$  converges in  $D_{[1,\infty]}[0, \infty)$  to  $(1/S_t)_{t \geq 0}$  as  $n \rightarrow \infty$ . Define  $\varphi : D_{[0,1]}[0, \infty) \rightarrow D_{[1,\infty]}[0, \infty)$  via  $\varphi(x) :=$

$(1/x_t)_{t \geq 0}$  for all  $x = (x_t)_{t \geq 0} \in D_{[0,1]}[0, \infty)$  with the convention  $1/0 := \infty$ . Since the transformation  $\varphi$  is continuous we will (equivalently) verify that  $(n/L_t^{(n)})_{t \geq 0}$  converges in  $D_{[0,1]}[0, \infty)$  to  $(S_t)_{t \geq 0}$  as  $n \rightarrow \infty$ .

Let  $F_n := \{n/j : j \in \{n, n+1, \dots\} \cup \{0\}\}$  and  $F := [0, 1]$  denote the state spaces and  $(U_t^{(n)})_{t \geq 0}$  and  $(U_t)_{t \geq 0}$  the semigroups of  $(n/L_t^{(n)})_{t \geq 0}$  and  $(S_t)_{t \geq 0}$  respectively. Define  $\pi_n : B(F) \rightarrow B(F_n)$  via  $\pi_n f(x) := f(x)$ ,  $f \in B(F)$ ,  $x \in F_n$ . By [Ethier and Kurtz \(1986, p. 172, Theorem 2.11\)](#) it suffices to verify that for all  $t \geq 0$  and all  $f \in C(F)$ ,

$$\lim_{n \rightarrow \infty} \sup_{x \in F_n} |U_t^{(n)} \pi_n f(x) - \pi_n U_t f(x)| = 0.$$

Obviously,  $U_t^{(n)} \pi_n f(x) = \mathbb{E}(\pi_n f(n/L_{s+t}^{(n)}) | n/L_s^{(n)} = x) = \mathbb{E}(f(n/L_t^{(n/x)}))$  and  $\pi_n U_t f(x) = U_t f(x) = \mathbb{E}(f(xS_t))$ . Thus, we have to verify that

$$\lim_{n \rightarrow \infty} \sup_{x \in F_n} |\mathbb{E}(f(n/L_t^{(n/x)})) - \mathbb{E}(f(xS_t))| = 0.$$

Since the polynomials are dense in  $C(F)$  it suffices to verify the latter equation for monomials  $f(x) = x^k$ , so we have to prove that  $\lim_{n \rightarrow \infty} \sup_{x \in F_n} |\mathbb{E}((n/L_t^{(n/x)})^k) - x^k \mathbb{E}(S_t^k)| = 0$  for all  $t \geq 0$  and  $k \in \mathbb{N}$ . In the following it is even shown that

$$\lim_{n \rightarrow \infty} \sup_{x \in [0,1]} |\mathbb{E}((n/L_t^{(\lfloor n/x \rfloor)})^k) - x^k \mathbb{E}(S_t^k)| = 0, \quad t \geq 0, k \in \mathbb{N}. \tag{5.4}$$

For  $m \in \mathbb{N}$ ,  $t \geq 0$  and  $y \in (0, 1]$ , it follows by duality ([Theorem 2.9](#) applied with  $i := \lceil m/y \rceil := \min\{z \in \mathbb{Z} : z \geq m/y\}$  and  $j := m$ ) that

$$\begin{aligned} \mathbb{P}(m/L_t^{(m)} \leq y) &= \mathbb{P}(L_t^{(m)} \geq m/y) = \mathbb{P}(L_t^{(m)} \geq \lceil m/y \rceil) \\ &= \mathbb{P}(N_t^{(\lceil m/y \rceil)} \leq m) = \mathbb{P}\left(\frac{N_t^{(\lceil m/y \rceil)}}{\lceil m/y \rceil} \leq \frac{m}{\lceil m/y \rceil}\right). \end{aligned}$$

Since  $N_t^{(m)}/m \rightarrow S_t$  in distribution as  $m \rightarrow \infty$  by part a) of [Theorem 2.13](#), which is already proven, and since  $m/\lceil m/y \rceil \rightarrow y$  as  $m \rightarrow \infty$ , we conclude that  $\lim_{m \rightarrow \infty} \mathbb{P}(m/L_t^{(m)} \leq y) = \mathbb{P}(S_t \leq y)$ , if  $y \in (0, 1]$  is a continuity point of the distribution function of  $S_t$ . The point  $y = 0$  has to be treated separately. For all  $m \in \mathbb{N}$  we have  $\mathbb{P}(m/L_t^{(m)} \leq 0) = \mathbb{P}(L_t^{(m)} = \infty) = \lim_{n \rightarrow \infty} \mathbb{P}(L_t^{(m)} \geq n) = \lim_{n \rightarrow \infty} \mathbb{P}(N_t^{(n)} \leq m) = \lim_{n \rightarrow \infty} \mathbb{P}(N_t^{(n)}/n \leq m/n) = \mathbb{P}(S_t \leq 0)$ , if  $y = 0$  is a continuity point of the distribution function of  $S_t$ . The pointwise convergence of the distribution functions implies the convergence  $m/L_t^{(m)} \rightarrow S_t$  in distribution as  $m \rightarrow \infty$ .

Fix  $x \in (0, 1]$ . Replacing  $m$  by  $\lfloor n/x \rfloor$  it follows by an application of Slutsky's theorem that  $n/L_t^{(\lfloor n/x \rfloor)} \rightarrow xS_t$  in distribution as  $n \rightarrow \infty$ . This convergence obviously holds as well for  $x = 0$  with the convention  $n/0 := \infty$ . Noting that the map  $y \mapsto y^k$  is bounded and continuous on  $[0, 1]$  we conclude that

$$\lim_{n \rightarrow \infty} \mathbb{E}((n/L_t^{(\lfloor n/x \rfloor)})^k) = x^k \mathbb{E}(S_t^k), \quad t \geq 0, k \in \mathbb{N}, x \in [0, 1].$$

In order to see that this pointwise convergence holds even uniformly for all  $x \in [0, 1]$  we proceed as follows. Fix  $t \geq 0$ ,  $k \in \mathbb{N}$  and  $n \in \mathbb{N}$ . By the pathwise construction of the fixation line, we have  $L_t^{(1)} \leq L_t^{(2)} \leq \dots$ . It follows that the map  $x \mapsto \mathbb{E}((n/L_t^{(\lfloor n/x \rfloor)})^k)$  is non-decreasing on  $[0, 1]$ . Clearly, the limiting map

$x \mapsto x^k \mathbb{E}(S_t^k)$  is non-decreasing, bounded and continuous on  $[0, 1]$ . Thus, the pointwise convergence holds even uniformly for all  $x \in [0, 1]$ . Note that the proof that this pointwise convergence holds even uniformly works the same as the proof that pointwise convergence of distribution functions is uniform if the limiting distribution function is continuous (Pólya, 1920, Satz I). Therefore, (5.4) is established. The proof is complete.  $\square$

### 6. Appendix

We establish a sort of duality relation for generalized Stirling numbers. Let  $a, b, r \in \mathbb{R}$  and suppose that  $t \notin \{0, a, 2a, 3a, \dots\}$  such that we can define

$$q_{ij} := \frac{(t-r|b)_j}{(t|a)_i} S(i, j), \quad i, j \in \mathbb{N}_0,$$

where  $(t|a)_i := \prod_{k=0}^{i-1} (t-ak)$  and the coefficients  $S(i, j) := S(i, j; a, b, r)$  are the generalized Stirling numbers as defined in Hsu and Shiue (1998). The recursion  $S(i+1, j) = S(i, j-1) + (jb-ia+r)S(i, j)$  for the generalized Stirling numbers (see Hsu and Shiue, 1998, Theorem 1) obviously transforms into the recursion

$$q_{i+1, j} = \frac{t-r-(j-1)b}{t-ia} q_{i, j-1} + \frac{jb-ia+r}{t-ia} q_{ij} \tag{6.1}$$

for the quantities  $q_{ij}$  ( $q_{i, -1} := 0$ ). Note that  $\sum_{j=0}^\infty q_{ij} = 1$  for all  $i \in \mathbb{N}_0$ . For  $i, j \in \mathbb{N}_0$  we define  $q_{i, \leq j} := \sum_{k=0}^j q_{ik}$ .

**Lemma 6.1.** *Fix  $j \in \mathbb{N}_0$  and suppose that the limit  $\lim_{k \rightarrow \infty} q_{k, \leq j}$  exists. Then, for all  $i \in \mathbb{N}_0$ ,*

$$q_{i, \leq j} - \lim_{k \rightarrow \infty} q_{k, \leq j} = \sum_{k=i}^\infty \frac{t-r-jb}{t-ka} q_{kj}. \tag{6.2}$$

*Remark 6.2.* At a first glance Lemma 6.1 looks somewhat technical and does not seem to have many applications. Indeed, in general it seems to be not straightforward to verify the existence of the limit  $\lim_{k \rightarrow \infty} q_{k, \leq j}$  and to determine this limit (if it exists). However, for particular parameter choices (for instance for  $a \leq 0, b > 0, r = 0$  and  $t > 0$  an integer multiple of  $b$ ), the  $q_{ij}$  turn out to be non-negative. In this case there exists a random variable  $K_i$  with distribution  $\mathbb{P}(K_i = j) = q_{ij}, j \in \mathbb{N}_0$ . Based on the recursion (6.1) the sequence  $K := (K_i)_{i \in \mathbb{N}_0}$  can be even constructed such that  $K$  is a Markov chain with initial state  $K_0 = 0$  satisfying  $K_{i+1} - K_i \in \{0, 1\}$  for all  $i \in \mathbb{N}_0$ . Hence,  $q_{i, \leq j} = \mathbb{P}(K_i \leq j)$  is non-increasing in  $i$ , which ensures the existence of the limit  $\lim_{i \rightarrow \infty} q_{i, \leq j}$ . If all states  $0, 1, \dots, j$  of the chain  $K$  are transient, then  $\lim_{i \rightarrow \infty} q_{i, \leq j} = 0$  and (6.2) reduces to

$$\sum_{k=0}^j q_{ik} = \sum_{k=i}^\infty \frac{t-r-jb}{t-ka} q_{kj}. \tag{6.3}$$

Roughly speaking, (6.3) is a sort of analytic reformulation of a particular Siegmund duality. For typical examples we refer the reader to the equality of (3.4) and (3.5) for the Dirichlet coalescent and to the equality of (4.3) and (4.4) for the Poisson-Dirichlet coalescent.

*Proof:* (of Lemma 6.1) The proof is purely analytic and straightforward. For all  $k, i \in \mathbb{N}_0$  we have

$$\begin{aligned} q_{ki} - q_{k+1,i} &= q_{ki} - \frac{t-r-(i-1)b}{t-ka} q_{k,i-1} - \frac{ib-ka+r}{t-ka} q_{ki} \\ &= \frac{t-r-ib}{t-ka} q_{ki} - \frac{t-r-(i-1)b}{t-ka} q_{k,i-1}. \end{aligned}$$

Summation over all  $i \in \{0, \dots, j\}$  yields  $q_{k,\leq j} - q_{k+1,\leq j} = ((t-r-jb)/(t-ka))q_{kj}$ . Another summation over all  $k \geq i$  yields the result.  $\square$

### Acknowledgements

The authors thank Jonas Kukla for helpful discussions and comments and an anonymous referee for a careful reading of the manuscript leading to an improvement of the presentation.

### References

- J. Bertoin and I. Kortchemski. Self-similar scaling limits of Markov chains on the positive integers. *ArXiv Mathematics e-prints* (2014). [arXiv: 1412.1068](#).
- M. Birkner, J. Blath, M. Möhle, M. Steinrücken and J. Tams. A modified lookdown construction for the Xi-Fleming-Viot process with mutation and populations with recurrent bottlenecks. *ALEA Lat. Am. J. Probab. Math. Stat.* **6**, 25–61 (2009). [MR2485878](#).
- J.-F. Delmas, J.-S. Dhersin and A. Siri-Jegousse. On the two oldest families for the Wright-Fisher process. *Electron. J. Probab.* **15**, no. 26, 776–800 (2010). [MR2653183](#).
- P. Donnelly and T. G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.* **24** (2), 698–742 (1996). [MR1404525](#).
- P. Donnelly and T. G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.* **27** (1), 166–205 (1999). [MR1681126](#).
- S. N. Ethier and T. G. Kurtz. *Markov Processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York (1986). ISBN 0-471-08186-8. [MR838085](#).
- F. Freund and M. Möhle. On the number of allelic types for samples taken from exchangeable coalescents with mutation. *Adv. in Appl. Probab.* **41** (4), 1082–1101 (2009). [MR2663237](#).
- A. Gnedin, A. Iksanov, A. Marynych and M. Möhle. On asymptotics of the beta coalescents. *Adv. in Appl. Probab.* **46** (2), 496–515 (2014). [MR3215543](#).
- A. Gnedin, Alexander Iksanov and Alexander Marynych. On  $\Lambda$ -coalescents with dust component. *J. Appl. Probab.* **48** (4), 1133–1151 (2011). [MR2896672](#).
- B. Haas and G. Miermont. Self-similar scaling limits of non-increasing Markov chains. *Bernoulli* **17** (4), 1217–1247 (2011). [MR2854770](#).
- K. Handa. The two-parameter Poisson-Dirichlet point process. *Bernoulli* **15** (4), 1082–1116 (2009). [MR2597584](#).
- O. Hénard. Change of measure in the lookdown particle system. *Stochastic Process. Appl.* **123** (6), 2054–2083 (2013). [MR3038498](#).
- O. Hénard. The fixation line in the  $\Lambda$ -coalescent. *Ann. Appl. Probab.* **25** (5), 3007–3032 (2015). [MR3375893](#).

- P. Herriger and M. Möhle. Conditions for exchangeable coalescents to come down from infinity. *ALEA Lat. Am. J. Probab. Math. Stat.* **9** (2), 637–665 (2012). [MR3069379](#).
- L. C. Hsu and P. J.-S. Shiue. A unified approach to generalized Stirling numbers. *Adv. in Appl. Math.* **20** (3), 366–384 (1998). [MR1618435](#).
- H.-K. Hwang and S. Janson. Local limit theorems for finite and infinite urn models. *Ann. Probab.* **36** (3), 992–1022 (2008). [MR2408581](#).
- J. F. C. Kingman. On the genealogy of large populations. *J. Appl. Probab.* (Special Vol. 19A), 27–43 (1982). [MR633178](#).
- S. Kotz, N. Balakrishnan and N. L. Johnson. *Continuous Multivariate Distributions. Vol. 1*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. Wiley-Interscience, New York, second edition (2000). ISBN 0-471-18387-3. [MR1788152](#).
- J. Kukla and M. Möhle. On the block counting process and the fixation line of the Bolthausen–Sznitman coalescent. *ArXiv Mathematics e-prints* (2016). [arXiv:1604.04514](#).
- C. Labbé. From flows of  $\Lambda$ -Fleming–Viot processes to lookdown processes via flows of partitions. *Electron. J. Probab.* **19**, no. 55, 49 (2014). [MR3227064](#).
- T. M. Liggett. *Interacting Particle Systems*. Classics in Mathematics. Springer-Verlag, Berlin (2005). ISBN 3-540-22617-6. Reprint of the 1985 original. [MR2108619](#).
- M. Möhle. On sampling distributions for coalescent processes with simultaneous multiple collisions. *Bernoulli* **12** (1), 35–53 (2006). [MR2202319](#).
- M. Möhle. Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson–Dirichlet coalescent. *Stochastic Process. Appl.* **120** (11), 2159–2173 (2010). [MR2684740](#).
- M. Möhle. The Mittag-Leffler process and a scaling limit for the block counting process of the Bolthausen–Sznitman coalescent. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** (1), 35–53 (2015). [MR3333734](#).
- M. Möhle. Hitting probabilities for the Greenwood model and relations to near constancy oscillation (2016+). To appear in *Bernoulli*.
- P. Pfaffelhuber and A. Wakolbinger. The process of most recent common ancestors in an evolving coalescent. *Stochastic Process. Appl.* **116** (12), 1836–1859 (2006). [MR2307061](#).
- P. Pfaffelhuber, A. Wakolbinger and H. Weisshaupt. The tree length of an evolving coalescent. *Probab. Theory Related Fields* **151** (3-4), 529–557 (2011). [MR2851692](#).
- J. Pitman. Coalescents with multiple collisions. *Ann. Probab.* **27** (4), 1870–1902 (1999). [MR1742892](#).
- J. Pitman. *Combinatorial Stochastic Processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin (2006). ISBN 978-3-540-30990-1. [MR2245368](#).
- G. Pólya. Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem. *Math. Z.* **8** (3-4), 171–181 (1920). [MR1544437](#).
- S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36** (4), 1116–1125 (1999). [MR1742154](#).
- S. Sagitov. Convergence to the coalescent with simultaneous multiple mergers. *J. Appl. Probab.* **40** (4), 839–854 (2003). [MR2012671](#).



- 
- J. Schweinsberg. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5**, no. 12, 50 (2000a). [MR1781024](#).
- J. Schweinsberg. A necessary and sufficient condition for the  $\Lambda$ -coalescent to come down from infinity. *Electron. Comm. Probab.* **5**, 1–11 (electronic) (2000b). [MR1736720](#).
- J. Schweinsberg. Dynamics of the evolving Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **17**, no. 91, 50 (2012). [MR2988406](#).