ALEA, Lat. Am. J. Probab. Math. Stat. **18**, 1195–1220 (2021)
DOI: 10.30757/ALEA.v18-44

The rate of convergence of the block counting process of exchangeable coalescents with dust

Martin Möhle

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

Abstract. Exchangeable coalescents with dust are studied. The rate of convergence as the sample size tends to infinity of the scaled block counting process to the frequency of singleton process is determined. This rate is expressed in terms of a certain Bernstein function. The proofs are based on Taylor expansions of the infinitesimal generators and semigroups and involve a particular concentration inequality arising in the context of Karlin's infinite urn model. The rate of convergence is calculated for several examples of coalescents.

1. Introduction

Exchangeable coalescents (Ξ-coalescents) have been the subject of extensive studies over the last decades. These are continuous-time Markovian processes $\Pi = (\Pi_t)_{t\geq 0}$ with values in the space $\mathcal{P}$ of partitions of $\mathbb{N} := \{1, 2, \ldots\}$. As time proceeds, blocks merge together to form larger blocks. For fundamental information on these processes we refer the reader to Möhle and Sagitov (2001) and Schweinsberg (2000a). Exchangeable coalescents are characterized by a finite measure Ξ on the infinite simplex $\Delta := \{u = (u_i)_{i\in\mathbb{N}} : u_1 \geq u_2 \geq \cdots \geq 0, |u| \leq 1\}$, where $|u| := \sum_{i\in\mathbb{N}} u_i$ for $u = (u_i)_{i\in\mathbb{N}} \in \Delta$. In general these processes allow for simultaneous multiple collisions of blocks. The subclass of coalescents with multiple collisions (Λ-coalescents) is obtained if Ξ is concentrated on $\{u \in \Delta : u_2 = 0\}$. In this case the coalescent is characterized by the finite measure Λ on $[0, 1]$ defined via $\Lambda(B) := \Xi(B \times \{0\} \times \{0\} \times \cdots)$ for all Borel sets $B \subseteq [0, 1]$. Coalescents with multiple collisions have been independently introduced by Pitman (1999) and Sagitov (1999). It is convenient to decompose $\Xi = \Xi(\{0\})\delta_0 + \Xi_0$, where Ξ_0 has no mass at $0 := (0, 0, \ldots) \in \Delta$. We focus on the subclass of Ξ-coalescents with dust. By definition, a Ξ-coalescent $\Pi = (\Pi_t)_{t\geq 0}$ has no dust if, for all times $t \geq 0$, the frequency S_t of singletons of Π_t satisfies $\mathbb{P}(S_t = 0) = 1$. For a precise definition of S_t we refer the reader to Section 3 of Möhle (2010). By the criterion of Schweinsberg (2000a),

Received by the editors March 13th, 2020; accepted April 1st, 2021.
2010 *Mathematics Subject Classification.* 60F05, 60J27, 60J90, 60G09, 92D15, 47D07.
Key words and phrases. Bernstein function, block counting process, concentration inequality, dust, exchangeable coalescent, rate of convergence, subordinator.

Π has dust if and only if

$$\Xi(\{0\}) \;=\; 0 \quad \text{and} \quad \alpha \;:=\; \int_\Delta |u|\,\nu(\mathrm{d}u) \;<\; \infty, \tag{1.1}$$

where $\nu(\mathrm{d}u) := \Xi_0(\mathrm{d}u)/(u,u)$ with $(u,u) := \sum_{i\in\mathbb{N}} u_i^2$ for $u = (u_i)_{i\in\mathbb{N}} \in \Delta$. It is well-known (see, for example, Eq. (8) of Möhle, 2010) that $Z = (Z_t)_{t\geq 0}$, defined via $Z_t := -\log S_t$ for all $t \geq 0$ (with the convention $-\log 0 := \infty$), is a drift-free subordinator with state space $[0,\infty]$, initial value $Z_0 = 0$ and Laplace exponent

$$\Phi(q) \;:=\; \int_\Delta \big(1-(1-|u|)^q\big)\,\nu(\mathrm{d}u), \qquad q \geq 0. \tag{1.2}$$

The subordinator Z has Lévy measure $\varrho := \nu_T$, where ν_T denotes the image measure of ν under the transformation $T : \Delta \to [0,\infty]$ defined via $T(u) := -\log(1-|u|)$ for all $u \in \Delta$. Note that S_t takes values in $E := [0,1]$ and has moments $\mathbb{E}(S_t^q) = e^{-t\Phi(q)}$, $t, q \geq 0$. The semigroup $(T_t)_{t\geq 0}$ of $S = (S_t)_{t\geq 0}$ is given by

$$T_t f(x) \;:=\; \mathbb{E}(f(S_{s+t})\,|\,S_s = x) \;=\; \mathbb{E}(f(xS_t)), \qquad t \geq 0, f \in B(E), x \in E, \tag{1.3}$$

where $B(E)$ denotes the space of bounded measurable real valued functions on E. The infinitesimal generator A of S satisfies

$$Af(x) \;=\; \int_\Delta \big(f(x(1-|u|)) - f(x)\big)\,\nu(\mathrm{d}u), \qquad f \in C^2(E), x \in E, \tag{1.4}$$

where $C^2(E)$ denotes the space of twice continuously differentiable real valued functions on E. Note that $C^2(E)$ is a core for A (see the remark after the proof of Corollary 2.3).

Coalescents with dust have been studied in a number of papers (see, for example, Gaiser and Möhle, 2016, Gnedin et al., 2011, Haas and Miermont, 2011 and Möhle, 2010). The latter article mainly focusses on the asymptotics as $n \to \infty$ of the total branch length of Ξ-coalescents with dust restricted to a sample of size n. In the more recent work of Gaiser and Möhle (2016) the asymptotic behavior of the block counting process of a coalescent with dust restricted to a sample of size n is analyzed via the method of moments. Motivated from approximation theory it is natural to ask for the speed of this convergence. This article provides detailed information on the speed of convergence of the corresponding infinitesimal generators and semigroups of the block counting process, which to the best of the author's knowledge has not been addressed in the literature so far.

The article is organized as follows. Section 2 contains the main results concerning the block counting process. Theorem 2.1 provides the speed of convergence for the generator of the scaled block counting process as the sample size n tends to infinity. The rate $r(n)$ of convergence (see Eq. (2.4)) turns out to be related to a particular Bernstein function $\widetilde{\Phi}$ defined in (2.3). A Lévy–Khintchine representation for $\widetilde{\Phi}$ is provided in Lemma 4.6. For a couple of important examples of coalescents the Bernstein function $\widetilde{\Phi}$ is computed in Section 3 showing that in these examples the rate $r(n)$ of convergence is typically of order $n^{-\beta}$ for some $\beta \in (0,1]$ or of order $(\log n)/n$. Proofs are deferred to Section 4. The proofs are essentially based on Taylor expansions of the corresponding infinitesimal generators and semigroups, but they involve in addition certain concentration inequalities occurring in the context of Karlin's infinite urn model (Karlin, 1967), where balls are allocated independently to an infinite number of boxes. These concentration inequalities are provided and verified in Sections 5 and 6 for Λ-coalescents and Ξ-coalescents respectively.

2. Results

For a Ξ-coalescent $\Pi = (\Pi_t)_{t\geq 0}$ and $n \in \mathbb{N}$ we denote by $\Pi_t^{(n)}$ the restriction of Π_t to $\mathcal{P}_n$, the set of partitions of $\{1,\ldots,n\}$. For $t \geq 0$ and $n \in \mathbb{N}$ let $N_t^{(n)}$ be the number of blocks of $\Pi_t^{(n)}$. The

process $(N_t^{(n)})_{t\geq 0}$ is called the block counting process of the restricted coalescent $(\Pi_t^{(n)})_{t\geq 0}$. For Ξ-coalescents with dust it is known (see Theorem 2.13 a) of Gaiser and Möhle (2016) or Corollary 2.4) that 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 $(S_t)_{t\geq 0}$ as n tends to infinity. We are interested in the speed of this convergence. For this purpose let us briefly consider the semigroups and generators of the involved processes. The scaled block counting process $(N_t^{(n)}/n)_{t\geq 0}$ has semigroup

$$\begin{aligned} T_t^{(n)} f(x) &= \mathbb{E}\left(f\left(\frac{N_{s+t}^{(n)}}{n}\right) \,\middle|\, \frac{N_s^{(n)}}{n} = x \right) \\ &= \mathbb{E}\left(f\left(\frac{N_t^{(nx)}}{n}\right)\right) = \sum_{j=1}^{nx} f\left(\frac{j}{n}\right) \mathbb{P}(N_t^{(nx)} = j) \end{aligned}$$

for $t \geq 0$, $n \in \mathbb{N}$, $x \in E_n := \{k/n : k \in \{1,\ldots,n\}\}$ and $f : E_n \to \mathbb{R}$. Thus, the corresponding infinitesimal generator A_n satisfies

$$\begin{aligned} A_n f(x) &:= \lim_{t\to 0} \frac{T_t^{(n)} f(x) - f(x)}{t} = \lim_{t\to 0} \sum_{j=1}^{nx} \left(f\left(\frac{j}{n}\right) - f(x)\right) \frac{\mathbb{P}(N_t^{(nx)} = j)}{t} \\ &= \sum_{j=1}^{nx-1} \left(f\left(\frac{j}{n}\right) - f(x)\right) q_{nx,j}, \qquad n \in \mathbb{N}, f : E_n \to \mathbb{R}, x \in E_n, \end{aligned}$$

where $q_{i,j} := \lim_{t\to 0} t^{-1}\mathbb{P}(N_t^{(i)} = j)$, $i,j \in \{1,\ldots,n\}$ with $i > j$, are the infinitesimal rates of the block counting process $(N_t^{(n)})_{t\geq 0}$. By Möhle (2010, p. 2162), the infinitesimal rates $q_{i,j}$, $i > j$, have the form $q_{i,j} = \int_\Delta \mathbb{P}(Y(i,u) = j)\, \nu(\mathrm{d}u)$, where (see Möhle, 2010, Eq. (4))

$$Y(i,u) := X_0(i,u) + \sum_{r\in\mathbb{N}} 1_{\{X_r(i,u)>0\}}, \tag{2.1}$$

and $(X_0(i,u), X_1(i,u), X_2(i,u), \ldots)$ has an infinite multinomial distribution with parameters i and $(1-|u|, u_1, u_2, \ldots)$. The occurrence of this infinite multinomial distribution comes from a putting balls into boxes experiment, originally introduced by Karlin (1967) and further investigated by several authors, see for example Dutko (1989), Gnedin et al. (2007) and Ben-Hamou et al. (2017). In this experiment each ball is allocated independently of all other balls into box r with probability u_r, $r \in \mathbb{N}_0 := \{0,1,\ldots\}$, where $u_0 := 1-|u|$. This experiment can as well be viewed as an urn version of Kingman's paintbox construction (Kingman, 1978a,b) of exchangeable random partitions. Most articles concentrate on the number $K_i := \sum_{r\in\mathbb{N}} 1_{\{X_r(i,u)>0\}}$ of occupied boxes after i balls have been allocated. In the context of coalescent theory however, $Y(i,u)$, defined in (2.1), is the important random variable, which counts all balls allocated in box 0 (the singletons) plus the number of all other occupied boxes (the number of multiple collisions). The above formula for $A_n f(x)$ can hence be written as

$$\begin{aligned} A_n f(x) &= \sum_{j=1}^{nx-1} \left(f\left(\frac{j}{n}\right) - f(x)\right) \int_\Delta \mathbb{P}(Y(nx,u) = j)\, \nu(\mathrm{d}u) \\ &= \int_\Delta \left(\mathbb{E}\left(f\left(\frac{Y(nx,u)}{n}\right)\right) - f(x)\right) \nu(\mathrm{d}u) \end{aligned} \tag{2.2}$$

for $n \in \mathbb{N}$, $f : E_n \to \mathbb{R}$ and $x \in E_n$. The proofs later on will draw heavily from this representation of the generator A_n.

Particular integrals over the characterizing measure ν play an important role in the theory of Ξ-coalescents. For example, the finiteness of the integral $\int_\Delta |u|\, \nu(\mathrm{d}u)$ in (1.1) characterizes the

dust property and the integral in (1.2) characterizes the distribution of the associated frequency of singleton process $S = (S_t)_{t\geq 0}$. We shall see shortly that another family of integrals, defined via

$$\widetilde{\Phi}(q) \;:=\; \int_\Delta \sum_{i\in\mathbb{N}} \left(1-(1-u_i)^q\right) \nu(\mathrm{d}u), \qquad q \geq 0, \tag{2.3}$$

plays an important role when studying the speed of convergence as $n \to \infty$ of the scaled block counting process of the Ξ-coalescent. Note that $\widetilde{\Phi}(0) = 0$, $\widetilde{\Phi}(1) = \int_\Delta |u|\, \nu(\mathrm{d}u) = \Phi(1)$ and that $\widetilde{\Phi}(\infty) := \lim_{q\to\infty} \widetilde{\Phi}(q)$ might be finite or infinite. By Lemma 7.1 in the appendix, $\Phi(n) \leq \widetilde{\Phi}(n)$ for all $n \in \mathbb{N}$. Lemma 4.2 provided in Section 4 shows that $\widetilde{\Phi}$ is infinitely often differentiable on $(0,\infty)$ with $(-1)^{k-1}\widetilde{\Phi}^{(k)}(q) \geq 0$ for all $k \in \mathbb{N}$ and $q > 0$. Therefore, $\widetilde{\Phi}$, restricted to $(0,\infty)$, is a Bernstein function in the sense of Schilling et al. (2012, Definition 3.1). A Lévy–Khintchine representation for $\widetilde{\Phi}$ is provided in Lemma 4.6. For the subclass of Λ-coalescents with dust, $\widetilde{\Phi} = \Phi$ coincides with the Laplace exponent (1.2) of the subordinator Z associated with the coalescent. This equality of $\widetilde{\Phi}$ and Φ however does not hold for arbitrary Ξ-coalescents with dust. In Section 3 the functions Φ and $\widetilde{\Phi}$ are calculated for several examples of coalescents. Let us introduce the function $r : (0,\infty) \to (0,\infty)$ via

$$r(q) \;:=\; \frac{\widetilde{\Phi}(q)}{q}, \qquad q > 0. \tag{2.4}$$

For natural reasons becoming clear immediately we call r the rate function. Lemma 4.5 shows that r is non-increasing on $(0,\infty)$ with $r(q) \to 0$ as $q \to \infty$. Our main result, Theorem 2.1 below, shows that, for the Ξ-coalescent with dust, $r(n)$ is the rate of convergence of the generator of the scaled block counting process to that of the frequency of singleton process. The proof of Theorem 2.1, provided in Section 4, is based on Taylor expansions of the generators and on crucial concentration inequalities for $Y(n,u)$ defined via (2.1). These concentration inequalities are provided in Sections 5 and 6 for Λ-coalescents and Ξ-coalescents respectively. Recall that $E := [0,1]$ and $E_n := \{k/n : k \in \{1,\ldots,n\}\}$ $(\subseteq E)$ for $n \in \mathbb{N}$. In the following $\pi_n : B(E) \to B(E_n)$ is defined via $\pi_n f(x) := f(x)$ for all $f \in B(E)$ and $x \in E_n$.

Theorem 2.1. *(Rate of convergence of the generator of the scaled block counting process) Let $\Pi = (\Pi_t)_{t\geq 0}$ be a Ξ-coalescent with dust and let A_n and A be the generators of the scaled block counting process $(N_t^{(n)}/n)_{t\geq 0}$, $n \in \mathbb{N}$, and the frequency of singleton process $(S_t)_{t\geq 0}$ respectively. Then, for all $n \in \mathbb{N}$ and $f \in C^2(E)$, the space of twice continuously differentiable real valued functions on E,*

$$\|A_n\pi_n f - \pi_n A f\| \;:=\; \sup_{x\in E_n} |A_n\pi_n f(x) - \pi_n A f(x)| \;\leq\; C_f r(n), \tag{2.5}$$

where $C_f := \|f'\| + 2\|f''\|$ and the rate $r(n)$ is defined via (2.4).

Remark 2.2. (Optimality) The rate $r(n)$ in (2.5) is optimal, which is seen as follows. Consider the identity polynomial $\mathrm{id}(x) := x$, $x \in E$. For $n \in \mathbb{N}$, $x \in E_n$ and $u \in \Delta$,

$$\begin{aligned}\mathbb{E}\left(\mathrm{id}\left(\frac{Y(nx,u)}{n}\right)\right) - \mathrm{id}(x(1-|u|)) \;&=\; \frac{\mathbb{E}(Y(nx,u))}{n} - x(1-|u|)\\ &=\; \sum_{i\in\mathbb{N}} \frac{1-(1-u_i)^{nx}}{n}.\end{aligned}$$

Integration with respect to the measure ν and taking (1.4) and (2.2) into account yields

$$A_n\pi_n\mathrm{id}(x) - \pi_n A\,\mathrm{id}(x) \;=\; \frac{1}{n}\int_\Delta \sum_{i\in\mathbb{N}} \left(1-(1-u_i)^{nx}\right) \nu(\mathrm{d}u) \tag{2.6}$$

for $n \in \mathbb{N}$ and $x \in E_n$. This expression is non-decreasing in $x \in E_n$ and hence takes its maximum at $x = 1 \in E_n$. Thus, the supremum over all $x \in E_n$ is

$$\|A_n \pi_n \mathrm{id} - \pi_n A\,\mathrm{id}\| \;=\; \frac{1}{n}\int_\Delta \sum_{i\in\mathbb{N}} \big(1-(1-u_i)^n\big)\,\nu(\mathrm{d}u) \;=\; r(n).$$

In this sense the rate $r(n)$ in Theorem 2.1 is optimal. As a side effect the calculations show that $C_{\mathrm{id}} = 1$ is the best possible (smallest) constant in Theorem 2.1. However, we do not claim that the constant C_f in Theorem 2.1 is optimal in general. For example (see the remark after the proof of Theorem 2.1), for the Λ-coalescent, Eq. (2.5) even holds with the optimized constant $C_f := \|f'\| + \|f''\|$. Finding the optimal constant C_f is a subtile problem related to the optimization of certain concentration inequalities, which does not seem to be straightforward to solve.

The following corollary provides an analog result on the level of semigroups.

Corollary 2.3. *In the situation of Theorem 2.1 let $(T_t^{(n)})_{t\geq 0}$ and $(T_t)_{t\geq 0}$ denote the semigroups of the scaled block counting process $(N_t^{(n)}/n)_{t\geq 0}$ and the frequency of singleton process $(S_t)_{t\geq 0}$ respectively. Then, for all $t \geq 0$, $n \in \mathbb{N}$ and $f \in C^2(E)$,*

$$\|T_t^{(n)}\pi_n f - \pi_n T_t f\| \;\leq\; tC_f r(n), \tag{2.7}$$

where C_f is the constant from Theorem 2.1 and the rate $r(n)$ is defined via (2.4).

The following weak convergence result is known from the literature, see Gaiser and Möhle (2016, Theorem 2.13 a)). The proof in Gaiser and Möhle (2016) is based on the method of moments. We provide alternative short proofs based on the bounds provided in Theorem 2.1 and Corollary 2.3.

Corollary 2.4. *(Convergence of the scaled block counting process) In the situation of Theorem 2.1, as $n \to \infty$, 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 $S = (S_t)_{t\geq 0} = (e^{-Z_t})_{t\geq 0}$.*

3. Examples

In this section the functions Φ and $\widetilde{\Phi}$ and their asymptotics are computed for several examples. The first two examples (beta coalescent and NLG-coalescent) are Λ-coalescents. In this case $\widetilde{\Phi}$ coincides with Φ. The other five examples are true Ξ-coalescents (with simultaneous multiple collisions) where $\widetilde{\Phi}$ in general differs from Φ. In the following $\Psi := (\log\Gamma)' = \Gamma'/\Gamma$ denotes the digamma function (logarithmic derivative of the gamma function) and $\gamma := -\Gamma'(1) \approx 0.577216$ the Euler–Mascheroni constant.

Example 3.1. (beta coalescent) Let $\Lambda = \beta(a,b)$ be the beta distribution with parameters $a, b \in (0,\infty)$. From (1.1) it follows that the beta coalescent has dust if and only if $a > 1$. Clearly, $\widetilde{\Phi} = \Phi$, so it suffices to compute Φ. For $a \in (1,\infty)\setminus\{2\}$ the Laplace exponent Φ is given by

$$\begin{aligned}\Phi(q) &= \frac{1}{\mathrm{B}(a,b)}\int_0^1 \big(1-(1-u)^q\big)u^{a-3}(1-u)^{b-1}\,\mathrm{d}u \\ &= \frac{\Gamma(a+b)}{(a-1)(a-2)}\left(\frac{a+b-2}{\Gamma(a+b-1)} - \frac{a+b+q-2}{\Gamma(a+b+q-1)}\frac{\Gamma(b+q)}{\Gamma(b)}\right), \qquad q \geq 0,\end{aligned}$$

where $\mathrm{B}(.,.)$ denotes the beta function. The case $a = 2$ has to be treated separately. In this case $\Phi(q) = b(b+1)(\Psi(b+q)-\Psi(b))$, $q \geq 0$. Clearly, in all cases $r(q) = \Phi(q)/q$, $q > 0$. For the particular case $a + b = 2$, a class of coalescents studied extensively in the literature, the Laplace exponent considerably simplifies to

$$\Phi(q) \;=\; \frac{1}{b(1-b)}\frac{q\Gamma(b+q)}{\Gamma(q+1)\Gamma(b)}, \qquad q \geq 0, b \in (0,1).$$

As $q \to \infty$,

$$\Phi(q) \sim \begin{cases} \dfrac{\Gamma(a+b)}{(a-1)(2-a)\Gamma(b)} q^{2-a} & \text{if } 1 < a < 2, \\ b(b+1)\log q & \text{if } a = 2, \\ \dfrac{(a+b-1)(a+b-2)}{(a-1)(a-2)} & \text{if } a > 2. \end{cases} \tag{3.1}$$

In particular $r(q) = \Phi(q)/q \to 0$ as $q \to \infty$, in agreement with Lemma 4.5.

Example 3.2. (NLG-coalescent) Let Λ be the negative logarithmic gamma distribution with parameters $\alpha, \varrho > 0$ having density $u \mapsto \alpha^\varrho u^{\alpha-1}(-\log u)^{\varrho-1}/\Gamma(\varrho)$, $u \in (0,1)$. We use the notation $\Lambda = \mathrm{NLG}(\alpha, \varrho)$ for this distribution. As for the beta coalescents this family of Λ-coalescents contains the Bolthausen–Sznitman coalescent ($\alpha = \varrho = 1$) and interpolates between the Kingman coalescent ($\alpha \to 0$ or $\varrho \to \infty$) and the star-shaped coalescent ($\alpha \to \infty$ or $\varrho \to 0$). Schweinsberg (2000b, Example 14) briefly considers the case $\alpha = 1$ and $\varrho = 2$, where Λ has density $u \mapsto -\log u$, $u \in (0,1)$. It is readily checked that $\int u^{-1}\Lambda(\mathrm{d}u) = (\alpha/(\alpha-1))^\varrho < \infty$ for $\alpha > 1$ and $\int u^{-1}\Lambda(\mathrm{d}u) = \infty$ for $\alpha \le 1$. By (1.1), the $\mathrm{NLG}(\alpha, \varrho)$-coalescent has dust if and only if $\alpha > 1$, which is assumed in the following. As for all Λ-coalescents, $\widetilde{\Phi} = \Phi$, so it suffices to consider Φ. By (1.2),

$$\Phi(q) = \frac{\alpha^\varrho}{\Gamma(\varrho)} \int_0^1 \big(1 - (1-u)^q\big) u^{\alpha-3}(-\log u)^{\varrho-1}\,\mathrm{d}u, \qquad q \ge 0. \tag{3.2}$$

Note that Φ is the Laplace exponent of a drift-free subordinator Z with state space $[0,\infty)$ and Lévy measure $\alpha^\varrho(\Gamma(\varrho))^{-1}(1-e^{-x})^{\alpha-3}(-\log(1-e^{-x}))^{\varrho-1}e^{-x}\,\mathrm{d}x$, $x \in (0,\infty)$. Binomial series expansion $1-(1-u)^q = \sum_{j=1}^\infty \binom{q}{j}(-1)^{j-1}u^j$, $0 < u < 1$, $q \ge 0$, leads to the alternating summation formula

$$\begin{aligned} \Phi(q) &= \frac{\alpha^\varrho}{\Gamma(\varrho)} \sum_{j=1}^\infty \binom{q}{j}(-1)^{j-1} \int_0^1 u^{j+\alpha-3}(-\log u)^{\varrho-1}\,\mathrm{d}u \\ &= \alpha^\varrho \sum_{j=1}^\infty \binom{q}{j} \frac{(-1)^{j-1}}{(j+\alpha-2)^\varrho}. \end{aligned}$$

For $\varrho = 1$, $\mathrm{NLG}(\alpha, 1)$ coincides with the beta distribution $\beta(\alpha, 1)$, in which case $\Phi(q)$ can be expressed in terms of gamma functions for $\alpha \in (1,\infty)\setminus\{2\}$ and in terms of the digamma function for $\alpha = 2$ (see Example 3.1). For integer $\varrho \in \mathbb{N}$ similar formulas for $\Phi(q)$ can be derived involving higher derivatives of the gamma function, but these formulas become more and more involved for large integer ϱ. For example, for $\varrho = 2$ and $\alpha > 2$ the well-known formula $\int_0^1 u^{a-1}(1-u)^{b-1}(-\log u)\,\mathrm{d}u = \mathrm{B}(a,b)(\Psi(a+b) - \Psi(a))$, $a, b > 0$, applied with $a := \alpha - 2$ and $b \in \{1, q+1\}$, leads to $\Phi(q) = \alpha^2/(\alpha-2)^2 - \alpha^2\mathrm{B}(\alpha-2, q+1)(\Psi(q+\alpha-1) - \Psi(\alpha-2))$. It does not seem to be possible to express the integral in (3.2) more explicitly for general α, ϱ and q. We thus focus on the asymptotics of $\Phi(q)$ as $q \to \infty$. Lemma 7.3 provided in the appendix, applied with $a := \alpha - 1 > 0$, $b := q$ and $c := \varrho$, shows that, as $q \to \infty$,

$$\Phi(q) \sim \begin{cases} \dfrac{\alpha^\varrho}{\Gamma(\varrho)} \dfrac{\Gamma(\alpha-1)}{2-\alpha} q^{2-\alpha}(\log q)^{\varrho-1} & \text{if } 1 < \alpha < 2, \\ \dfrac{2^\varrho}{\Gamma(\varrho+1)}(\log q)^\varrho & \text{if } \alpha = 2, \\ \left(\dfrac{\alpha}{\alpha-2}\right)^\varrho & \text{if } \alpha > 2. \end{cases} \tag{3.3}$$

For $\varrho = 1$ the asymptotics (3.3) coincides with the asymptotics (3.1) of $\Phi(q)$ as $q \to \infty$ for the $\beta(\alpha, 1)$-coalescent studied in the previous Example 3.1, which must be so, since $\mathrm{NLG}(\alpha, 1) = \beta(\alpha, 1)$. As a side effect, Lemma 7.3 is also useful to clarify the coming down from infinity problem for the

NLG(α,ϱ)-coalescent for arbitrary parameters $\alpha,\varrho>0$. For information on the coming down from infinity property we refer the reader to Schweinsberg (2000b).

Proposition 3.3. *Let $\alpha,\varrho>0$. The* NLG(α,ϱ)*-coalescent comes down from infinity if and only if $\alpha<1$ or $\alpha=1$ and $\varrho>1$.*

Remark 3.4. For $\alpha=1$ and $\varrho=2$ this result was obtained by Schweinsberg (2000b, Example 14).

Proof of Proposition 3.3: We have

$$\begin{aligned}\eta_n &:= n\int_{[0,1]}\big(1-(1-u)^{n-1}\big)u^{-1}\,\Lambda(\mathrm{d}u)\\ &= n\frac{\alpha^\varrho}{\Gamma(\varrho)}\int_0^1\big(1-(1-u)^{n-1}\big)u^{\alpha-2}(-\log u)^{\varrho-1}\,\mathrm{d}u.\end{aligned}$$

From Lemma 7.3, applied with $a:=\alpha$, $b:=n-1$ and $c:=\varrho$, it follows that, as $n\to\infty$,

$$\eta_n \sim \begin{cases}\dfrac{\alpha^\varrho}{\Gamma(\varrho)}\dfrac{\Gamma(\alpha)}{1-\alpha}n^{2-\alpha}(\log n)^{\varrho-1} & \text{if } 0<\alpha<1,\\ \dfrac{1}{\Gamma(\varrho+1)}n(\log n)^\varrho & \text{if } \alpha=1,\\ \left(\dfrac{\alpha}{\alpha-1}\right)^\varrho n & \text{if } \alpha>1.\end{cases}$$

Thus, $\sum_{n=2}^\infty \eta_n^{-1}<\infty$ if and only if $\alpha<1$ or $\alpha=1$ and $\varrho>1$. The statement of Proposition 3.3 therefore follows from the criterion of Schweinsberg (2000b, Corollary 2). □

Example 3.5. (Dirac coalescent) Assume that the characterizing measure $\nu=\delta_a$ is the Dirac measure at some given point $a=(a_i)_{i\in\mathbb{N}}\in\Delta\backslash\{0\}$. The corresponding coalescent has dust, since $\int|u|\,\nu(\mathrm{d}u)=|a|\le 1<\infty$. By (1.2) and (2.3),

$$\Phi(q) \;=\; 1-(1-|a|)^q \quad\text{and}\quad \widetilde{\Phi}(q) \;=\; \sum_{i\in\mathbb{N}}(1-(1-a_i)^q), \qquad q\ge 0.$$

Note that Φ is the Laplace exponent of a drift-free subordinator Z with Lévy measure $\varrho:=\delta_{-\log(1-|a|)}$ whereas $\widetilde{\Phi}$ is the Laplace exponent of a drift-free subordinator $\widetilde{Z}$ with Lévy measure $\widetilde{\varrho}:=\sum_{i\in\mathbb{N}}\delta_{-\log(1-a_i)}$. Clearly $\Phi(q)\to 1$ as $q\to\infty$, since $|a|>0$. The asymptotic behavior of $\widetilde{\Phi}(q)$ as $q\to\infty$ is difficult to describe in general and heavily depends on the point $a\in\Delta\setminus\{0\}$. Define $a_0:=1-|a|$ and consider Karlin's infinite urn model (Karlin, 1967), where balls are independently allocated to an infinite number of boxes with probability a_i for each ball to be allocated in box $i\in\mathbb{N}_0$. Then, $\widetilde{\Phi}(n)=\sum_{i\in\mathbb{N}}(1-(1-a_i)^n)$ is the mean number of non-empty boxes (disregarding box 0) after $n\in\mathbb{N}$ balls have been allocated. It is well-known (see, for example, Gnedin et al., 2007) that, for particular choices of the point $a\in\Delta\setminus\{0\}$, this mean number can oscillate with n. Under additional regularity assumptions the asymptotic behavior of $\widetilde{\Phi}(q)$ as $q\to\infty$ can be determined as follows. Let $\delta_a:=\sum_{i\in\mathbb{N}}\delta_{a_i}$ denote the counting measure of a. A straightforward calculation shows that $\widetilde{\Phi}(q)=q\int_0^\infty e^{-qx}u(x)\,\mathrm{d}x$ for $q>0$, where $u(x):=\delta_a([1-e^{-x},1])$ for all $x\ge 0$, so $\widetilde{\Phi}(q)/q$ coincides with the Laplace transform of u. Define $U(x):=\int_0^x u(t)\,\mathrm{d}t$ for all $x\ge 0$. If

$$U(x) \;\sim\; x^{1-\alpha}\ell(1/x), \qquad x\to 0, \tag{3.4}$$

for some $\alpha\in[0,1]$ and some function ℓ slowly varying at infinity, then a Tauberian argument (apply for example Theorem 1.7.1' of Bingham et al. (1987) with $\varrho:=1-\alpha$) yields

$$\widetilde{\Phi}(q) \;\sim\; \Gamma(2-\alpha)q^\alpha\ell(q), \qquad q\to\infty. \tag{3.5}$$

In the following three examples are provided corresponding to the cases $\alpha=0$ (slow variation), $0<\alpha<1$ and $\alpha=1$ (rapid variation).

(i) Assume that $a_i = cp^i$, $i \in \mathbb{N}$, for some $p \in (0,1)$ and some $c \in (0,(1-p)/p]$. Then $u(x) = \lfloor(\log(1-e^{-x}) - \log c)/\log p\rfloor \sim \mu\log(1/x)$ as $x \to 0$ and $U(x) \sim \mu x\log(1/x)$ as $x \to 0$ with constant $\mu := -1/\log p \in (0,\infty)$. Thus, (3.4) holds with $\alpha := 0$ and $\ell(y) := \mu\log y$ and, by (3.5), $\widetilde{\Phi}(q) \sim \ell(q) = \mu\log q$ as $q \to \infty$. The asymptotic behavior of $\widetilde{\Phi}(q)$ can be alternatively derived without exploiting Tauberian theorems via

$$\begin{aligned}\widetilde{\Phi}(q) &= \sum_{i\in\mathbb{N}}(1-(1-cp^i)^q) \sim \int_0^\infty (1-(1-cp^t)^q)\,\mathrm{d}t = \mu\int_0^c \frac{1-(1-x)^q}{x}\,\mathrm{d}x \\ &\sim \mu\int_0^1 \frac{1-(1-x)^q}{x}\,\mathrm{d}x = \mu(\Psi(q+1)+\gamma) \sim \mu\log q, \qquad q\to\infty.\end{aligned}$$

(ii) Assume that $a_i = ci^{-\beta}$, $i \in \mathbb{N}$, for some $\beta \in (1,\infty)$ and some $c \in (0, 1/\zeta(\beta)]$, where $\zeta(\beta) := \sum_{i\in\mathbb{N}} i^{-\beta}$. Define $\alpha := 1/\beta \in (0,1)$. Then $u(x) = \lfloor c^\alpha/(1-e^{-x})^\alpha\rfloor \sim c^\alpha x^{-\alpha}$ as $x \to 0$ and $U(x) \sim c^\alpha x^{1-\alpha}/(1-\alpha)$ as $x \to 0$. Thus, (3.4) holds with $\ell(y) := c^\alpha/(1-\alpha)$ and, by (3.5), $\widetilde{\Phi}(q) \sim \Gamma(1-\alpha)c^\alpha q^\alpha$ as $q \to \infty$. Alternatively,

$$\widetilde{\Phi}(q) = \sum_{i\in\mathbb{N}}(1-(1-ci^{-1/\alpha})^q) \sim \int_{c^\alpha}^\infty (1-(1-ct^{-1/\alpha})^q)\,\mathrm{d}t$$

and the substitution $x = ct^{-1/\alpha}$ yields

$$\widetilde{\Phi}(q) \sim c^\alpha\int_0^1 \frac{\alpha(1-(1-x)^q)}{x^{\alpha+1}}\,\mathrm{d}x = c^\alpha\left(\frac{\Gamma(q+1)\Gamma(1-\alpha)}{\Gamma(q+1-\alpha)} - 1\right) \sim \Gamma(1-\alpha)c^\alpha q^\alpha$$

as $q \to \infty$.

(iii) Assume that $a_1 = 0$ and $a_i = c/(i(\log i)^\beta)$, $i \in \mathbb{N}\backslash\{1\}$, for some $\beta \in (1,\infty)$ and some sufficiently small constant $c > 0$ such that $|a| \le 1$. A technical but straightforward computation shows that $u(x) = \lfloor y/(\beta W(y^{1/\beta}/\beta))^\beta\rfloor$, where $y := y(x) := c/(1-e^{-x})$ and W denotes the Lambert W function satisfying $W(z)e^{W(z)} = z$. From $W(t) \sim \log t$ as $t \to \infty$ we conclude that $u(x) \sim c/(x(-\log x)^\beta)$ as $x \to 0$. Therefore, $U(x) \sim c/((\beta-1)(-\log x)^{\beta-1})$ as $x \to 0$. Thus, (3.4) holds with $\alpha := 1$ and $\ell(y) := c/((\beta-1)(\log y)^{\beta-1})$ and, by (3.5),

$$\widetilde{\Phi}(q) \sim q\ell(q) = \frac{c}{\beta-1}\frac{q}{(\log q)^{\beta-1}}, \qquad q\to\infty.$$

Example 3.6. (Dirichlet coalescent) Let $X = (X_1,\ldots,X_N)$ have a symmetric Dirichlet distribution $D_N(\alpha)$ with parameters $N \in \mathbb{N}$ and $\alpha > 0$ and let $X_{(1)} \ge \cdots \ge X_{(N)}$ denote the order statistics of X. The Dirichlet coalescent with parameters $N \in \mathbb{N}$ and $\alpha > 0$ has been studied in Section 3 of Gaiser and Möhle (2016). In this case the measure $\nu(\mathrm{d}u) = \Xi_0(\mathrm{d}u)/(u,u)$ is the distribution of $(X_{(1)},\ldots,X_{(N)},0,0,\ldots)$. Since ν is concentrated on $\Delta_N := \{u = (u_i)_{i\in\mathbb{N}} \in \Delta : u_1+\cdots+u_N = 1\}$, it follows that the Laplace exponent (1.2) of the associated subordinator satisfies $\Phi(0) = 0$ and $\Phi(q) = \nu(\Delta) = 1$ for $q > 0$. The formula (2.3) for $\widetilde{\Phi}(q)$ turns into

$$\widetilde{\Phi}(q) = \int_{\Delta_N}\sum_{i=1}^N \big(1-(1-u_i)^q\big)\,\nu(\mathrm{d}u), \qquad q\ge 0.$$

The function below the integral is symmetric with respect to $u_1,\ldots,u_N$. Thus,

$$\widetilde{\Phi}(q) = \int\sum_{i=1}^N \big(1-(1-u_i)^q\big)D_N(\alpha)(\mathrm{d}u_1,\ldots,\mathrm{d}u_N) = N\mathbb{E}(1-(1-X_1)^q), \qquad q\ge 0.$$

If $N = 1$ then $X_1 \equiv 1$ and hence $\widetilde{\Phi} = \Phi$. Assume now that $N \in \mathbb{N}\setminus\{1\}$. Since X_1 is beta distributed with parameters α and $N\alpha - \alpha$ and hence satisfies

$$\begin{aligned}\mathbb{E}((1-X_1)^q) &= \frac{1}{\mathrm{B}(\alpha, N\alpha-\alpha)}\int_0^1 (1-x)^q x^{\alpha-1}(1-x)^{N\alpha-\alpha-1}\,\mathrm{d}x \\ &= \frac{\mathrm{B}(\alpha, N\alpha-\alpha+q)}{\mathrm{B}(\alpha, N\alpha-\alpha)} = \frac{\Gamma(N\alpha)\Gamma(N\alpha-\alpha+q)}{\Gamma(N\alpha-\alpha)\Gamma(N\alpha+q)},\end{aligned}$$

it follows that

$$\widetilde{\Phi}(q) = N\left(1 - \frac{\Gamma(N\alpha)\Gamma(N\alpha-\alpha+q)}{\Gamma(N\alpha-\alpha)\Gamma(N\alpha+q)}\right) \to N, \qquad q \to \infty.$$

Note that $\widetilde{\Phi}$ is the Laplace exponent of a drift-free subordinator $\widetilde{Z}$ with state space $[0,\infty)$ and Lévy measure $\widetilde{\varrho}(\mathrm{d}x) = N(\mathrm{B}(\alpha, N\alpha-\alpha))^{-1}(1-e^{-x})^{\alpha-1}e^{-(N\alpha-\alpha)x}\mathrm{d}x$, $x \in (0,\infty)$. The Laplace exponent $\widetilde{\Phi}$ does not coincide with Φ.

Example 3.7. (Poisson–Dirichlet coalescent) The two-parameter Poisson–Dirichlet coalescent is the Ξ-coalescent where the measure $\nu(\mathrm{d}u) = \Xi_0(\mathrm{d}u)/(u,u)$ is the two-parameter Poisson–Dirichlet distribution with parameters $0 \le \alpha < 1$ and $\theta > -\alpha$. This coalescent has been introduced in Möhle (2010) and for the particular case $\alpha = 0$ in Section 3 of Sagitov (2003). The associated block counting process and its Siegmund dual fixation line have been investigated in Section 4 of Gaiser and Möhle (2016). Since ν is concentrated on $\Delta^* := \{u \in \Delta : |u| = 1\}$ it follows that $\int_\Delta |u|\,\nu(\mathrm{d}u) = \nu(\Delta^*) = 1 < \infty$, which implies that this coalescent has dust. Moreover, as in Example 3.6, $\Phi(0) = 0$ and $\Phi(q) = \nu(\Delta) = 1$ for $q > 0$. By Handa (2009, Eq. (2.1)), applied with $n := 1$ and $f(x) := 1-(1-x)^q$,

$$\widetilde{\Phi}(q) = \int_\Delta \sum_{i\in\mathbb{N}} f(u_i)\,\nu(\mathrm{d}u) = \int_{\mathbb{R}} \big(1-(1-x)^q\big)\,\mu_1(\mathrm{d}x), \qquad q \ge 0,$$

where μ_1 denotes the first correlation measure associated with the Poisson–Dirichlet distribution. The density (correlation function) of μ_1 is explicitly known (see, for example, Handa, 2009, Theorem 2.1) and it follows that

$$\widetilde{\Phi}(q) = c_{1,\alpha,\theta}\int_0^1 \big(1-(1-x)^q\big)x^{-\alpha-1}(1-x)^{\theta+\alpha-1}\,\mathrm{d}x, \qquad q \ge 0,$$

with normalizing constant $c_{1,\alpha,\theta} := \mathrm{B}(1-\alpha, \theta+\alpha) = \Gamma(\theta+1)/(\Gamma(1-\alpha)\Gamma(\theta+\alpha))$. Note that $\widetilde{\Phi}$ is the Laplace exponent of a drift-free subordinator $\widetilde{Z}$ with Lévy measure $\widetilde{\varrho}(\mathrm{d}x) = c_{1,\alpha,\theta}(1-e^{-x})^{-\alpha-1}e^{-(\theta+\alpha)x}\mathrm{d}x$, $x \in (0,\infty)$. In particular, $\widetilde{\Phi}$ does not coincide with Φ. In order to compute $\widetilde{\Phi}$ explicitly assume first that $\alpha > 0$. Then the integral above can be expressed in terms of gamma functions which leads after some straightforward manipulation to

$$\widetilde{\Phi}(q) = \frac{\theta+q}{\alpha}\frac{\Gamma(\theta+\alpha+q)\Gamma(\theta+1)}{\Gamma(\theta+1+q)\Gamma(\theta+\alpha)} - \frac{\theta}{\alpha}, \qquad q \ge 0.$$

In particular $\widetilde{\Phi}(q) \sim \Gamma(\theta+1)/\Gamma(\theta+\alpha)q^\alpha/\alpha$ as $q \to \infty$. Assume now that $\alpha = 0$. Then

$$\widetilde{\Phi}(q) = \theta\int_0^1 \big(1-(1-x)^q\big)x^{-1}(1-x)^{\theta-1}\,\mathrm{d}x = \theta\big(\Psi(\theta+q) - \Psi(\theta)\big), \qquad q \ge 0.$$

Because of the occurrence of the Ψ-function in its Laplace exponent we call $\widetilde{Z}$ the Ψ-subordinator with parameter θ. Note that $\widetilde{\Phi}(q) \sim \theta\log q$ as $q \to \infty$.

Example 3.8. (Dirichlet–Kingman coalescent) Let $a_1, a_2, \ldots$ be positive real numbers satisfying $a := \sum_{n\in\mathbb{N}} a_n < \infty$. For any finite non-empty subset J of $\mathbb{N}$ define $a_J := \sum_{j\in J} a_j$ and let P_J denote the probability measure on $\Delta_J := \{x_J := (x_j)_{j\in J} \in [0,1]^{|J|} : |x_J| := \sum_{j\in J} x_j \le 1\}$ with density

$$f_J(x_J) \;:=\; \Gamma(a)\frac{(1-|x_J|)^{a-a_J-1}}{\Gamma(a-a_J)}\prod_{j\in J}\frac{x_j^{a_j-1}}{\Gamma(a_j)}, \qquad x_J = (x_j)_{j\in J} \in \Delta_J,$$

with respect to Lebesgue measure on Δ_J. For any $J \subset \mathbb{N}$ with $1 < |J| < \infty$ and any $j \in J$ a straightforward calculation shows that $\int_0^{1-|x_H|} f_J(x_J)\,\mathrm{d}x_j = f_H(x_H)$ for all $x_H \in \Delta_H$, where $H := J \setminus \{j\}$. Thus, the family of probability measures $(P_J)_J$ is consistent. The projective limit P of $(P_J)_J$ is called the Kingman–Dirichlet distribution with parameter $(a_n)_{n\in\mathbb{N}}$. For more information on this distribution we refer the reader to Huillet and Martinez (2008, Section 4). Let $X_1, X_2, \ldots$ be random variables with joint distribution P and let $X_{(1)} \ge X_{(2)} \cdots$ denote these random variables in decreasing order. The Dirichlet–Kingman coalescent, introduced in Example 6.1 on p. 549 of Huillet and Möhle (2011), is the Ξ-coalescent where the characterizing measure $\nu(\mathrm{d}u) = \Xi_0(\mathrm{d}u)/(u,u)$ is (by definition) the distribution of $(X_{(1)}, X_{(2)}, \ldots)$. For this coalescent a $(i_1, \ldots, i_j)$-merger, $j, i_1, \ldots, i_j \in \mathbb{N}$, occurs at the rate

$$\begin{aligned}\varphi(i_1,\ldots,i_j) \;&:=\; \int_\Delta \sum_{\substack{r_1,\ldots,r_j\in\mathbb{N}\\ \text{all distinct}}} u_{r_1}^{i_1}\cdots u_{r_j}^{i_j}\,\nu(\mathrm{d}u) \;=\; \sum_{\substack{r_1,\ldots,r_j\in\mathbb{N}\\ \text{all distinct}}} \mathbb{E}(X_{(r_1)}^{i_1}\cdots X_{(r_j)}^{i_j})\\ &=\; \sum_{\substack{r_1,\ldots,r_j\in\mathbb{N}\\ \text{all distinct}}} \mathbb{E}(X_{r_1}^{i_1}\cdots X_{r_j}^{i_j}) \;=\; \sum_{\substack{r_1,\ldots,r_j\in\mathbb{N}\\ \text{all distinct}}} \frac{[a_{r_1}]_{i_1}\cdots[a_{r_j}]_{i_j}}{[a]_{i_1+\cdots+i_j}},\end{aligned}$$

where $[x]_0 := 1$ and $[x]_i := x(x+1)\cdots(x+i-1)$ for $x \in \mathbb{R}$ and $i \in \mathbb{N}$.
If $a_1 = \cdots = a_N = \alpha$ for some $\alpha > 0$ and some $N \in \mathbb{N}$, then in the limit $a_n \to 0$ for all $n > N$ we are back to the Dirichlet model discussed in Example 3.6.
As in the previous example, since ν is concentrated on Δ^*, this coalescent has dust and the Laplace exponent Φ of the associated subordinator satisfies $\Phi(q) = 1$ for $q > 0$. The formula (2.3) for $\widetilde{\Phi}(q)$ turns into $\widetilde{\Phi}(q) = \sum_{i\in\mathbb{N}}(1 - \mathbb{E}((1-X_{(i)})^q)) = \sum_{i\in\mathbb{N}}(1 - \mathbb{E}((1-X_i)^q))$. Since $1 - X_i$ is beta distributed with parameters $a - a_i$ and a_i, it follows that

$$\widetilde{\Phi}(q) \;=\; \sum_{i\in\mathbb{N}}\left(1 - \frac{\Gamma(a)\Gamma(a-a_i+q)}{\Gamma(a-a_i)\Gamma(a+q)}\right), \qquad q \ge 0. \tag{3.6}$$

Alternatively, by Lemma 4.6 and (4.8), $\widetilde{\Phi}(q) = \int_{[0,1]}(1-(1-x)^q)\,\mu(\mathrm{d}x)$, $q \ge 0$, where $\mu(B) := \int_\Delta \sum_{i\in\mathbb{N}} 1_B(u_i)\,\nu(\mathrm{d}u) = \sum_{i\in\mathbb{N}} \mathbb{E}(1_B(X_i)) = \sum_{i\in\mathbb{N}} \mathbb{P}(X_i \in B)$ for all Borel sets $B \subseteq [0,1]$. Note that μ is the intensity measure of the point process $\sum_{i\in\mathbb{N}} \delta_{X_i}$.
In order to see that $\widetilde{\Phi}(q) \sim a\log q$ as $q \to \infty$ let us rewrite (3.6) for $q > 1$ in the form

$$\frac{\widetilde{\Phi}(q)}{\log q} \;=\; \int_{\mathbb{N}} f_q(i)\,\delta_{\mathbb{N}}(\mathrm{d}i), \qquad q > 1, \tag{3.7}$$

where $\delta_{\mathbb{N}}$ denotes the counting measure on $\mathbb{N}$ and

$$f_q(i) \;:=\; \frac{1}{\log q}\left(1 - \frac{\Gamma(a)\Gamma(1-a_i+q)}{\Gamma(a-a_i)\Gamma(a+q)}\right), \qquad q > 1, i \in \mathbb{N}.$$

From

$$1 - \frac{\Gamma(a)\Gamma(a-a_i+q)}{\Gamma(a-a_i)\Gamma(a+q)} \;\sim\; a_i(\Psi(a+q) - \Psi(a)) \;\sim\; a_i\log q, \qquad q \to \infty,$$

it follows for all $i \in \mathbb{N}$ that $f_q(i) \to a_i =: f(i)$ as $q \to \infty$.
Fix $q_0 > 1$. In the following we find a dominating map for the functions f_q, $q \ge q_0$. For $q \ge 1$ the

Pochhammer map $x \mapsto [x]_q = \Gamma(x+q)/\Gamma(x)$ is convex on $(0,\infty)$. (For $q \geq 1$ even the map $x^{-q}[x]_q$ is completely monotone and hence convex on $(0,\infty)$ by Theorem 2.4 of Ismail et al. (1986), applied with $a := q$ and $b := 0$.) Replacing x by $a-x$ and multiplying by the factor $\Gamma(a)/\Gamma(a+q)$ it follows that

$$h(x) \;:=\; \frac{\Gamma(a)}{\Gamma(a+q)}[a-x]_q \;=\; \frac{\Gamma(a)\Gamma(a-x+q)}{\Gamma(a-x)\Gamma(a+q)}$$

is convex on $[0,a)$, which yields the inequality $h(x) - h(0) \geq h'(0)x$, $0 \leq x < a$. Noting that $h(0) = 1$ and $h'(0) = \Psi(a) - \Psi(a+q)$ we obtain

$$1 - \frac{\Gamma(a)\Gamma(a+q-x)}{\Gamma(a-x)\Gamma(a+q)} \;\leq\; (\Psi(a+q) - \Psi(a))x, \quad a > 0, q \geq 1, 0 \leq x < a.$$

Applying this inequality with $x := a_i$ $(< a)$ yields for all $i \in \mathbb{N}$ and all $q \geq q_0$ that

$$f_q(i) \;\leq\; \frac{\Psi(a+q) - \Psi(a)}{\log q} a_i \;\leq\; K a_i,$$

where $K := K_a := \sup_{q \geq q_0}(\Psi(a+q) - \Psi(a))/\log q < \infty$, since $q_0 > 1$ and $\Psi(a+q) \sim \log q$ as $q \to \infty$. The dominating map $i \mapsto K a_i$ is integrable with respect to the counting measure on $\mathbb{N}$. Thus, by dominated convergence, the integral in (3.7) converges to $\int_{\mathbb{N}} f(i)\,\delta_{\mathbb{N}}(\mathrm{d}i) = \sum_{i \in \mathbb{N}} a_i = a$ as $q \to \infty$. Thus, $\widetilde{\Phi}(q) \sim a \log q$ as $q \to \infty$.

Example 3.9. (Symmetric coalescent) Let $(m_k)_{k \in \mathbb{N}}$ be a sequence of non-negative real numbers satisfying $\sum_{k \in \mathbb{N}} m_k/k < \infty$. Suppose that ν assigns for each $k \in \mathbb{N}$ mass m_k to the point $u^{(k)} := (1/k, \ldots, 1/k, 0, 0, \ldots) \in \Delta^\star$. The assumption $\sum_{k \in \mathbb{N}} m_k/k < \infty$ ensures that the measure Ξ is finite as required. Except for the fact that we exclude a Kingman part for simplicity, i.e. $\Xi(\{0\}) = 0$, this class of coalescents has been recently studied in González Casanova et al. (2021+). For this coalescent a $(i_1, \ldots, i_j)$-merger, $j, i_1, \ldots, i_j \in \mathbb{N}$, occurs at the rate

$$\varphi(i_1, \ldots, i_j) \;:=\; \int_\Delta \sum_{\substack{r_1, \ldots, r_j = 1 \\ \text{all distinct}}}^{k} u_{r_1}^{i_1} \cdots u_{r_j}^{i_j}\, \nu(\mathrm{d}u) \;=\; \sum_{k \in \mathbb{N}} (k)_j \left(\frac{1}{k}\right)^i m_k \;=\; \sum_{k \geq j} \frac{(k)_j}{k^i} m_k,$$

where $i := i_1 + \cdots + i_j$, $(x)_0 := 0$ and $(x)_j := x(x-1)\cdots(x-j+1)$ for $x \in \mathbb{R}$ and $j \in \mathbb{N}$. In this symmetric situation, $\mathbb{P}(Y(i, u^{(k)}) = j)$ is the probability to obtain j non-empty boxes when $i \in \mathbb{N}$ balls are allocated independently and uniformly to $k \in \mathbb{N}$ boxes. This probability is given by (a formula well-known from several textbooks, see for example Durrett (2019, p. 172), and already known by Kolchin et al., 1978, Eqs. (1) and (2))

$$\mathbb{P}(Y(i, u^{(k)}) = j) \;=\; S(i,j) \frac{(k)_j}{k^i}, \tag{3.8}$$

where $S(i,j) := (1/j!) \sum_{r=0}^{j} (-1)^{j-r} \binom{j}{r} r^i$, $i, j \in \mathbb{N}_0$, denote the Stirling numbers of the second kind. Thus, the block counting process moves from state $i \in \mathbb{N}$ to state $j \in \mathbb{N}$ with $j < i$ at the rate

$$q_{i,j} \;=\; \int_\Delta \mathbb{P}(Y(i,u) = j)\, \nu(\mathrm{d}u) \;=\; \sum_{k \in \mathbb{N}} \mathbb{P}(Y(i, u^{(k)}) = j) m_k \;=\; S(i,j) \sum_{k \in \mathbb{N}} \frac{(k)_j}{k^i} m_k.$$

The total rates are

$$\begin{aligned} q_i \;&:=\; \sum_{j=1}^{i-1} q_{i,j} \;=\; \sum_{k \in \mathbb{N}} \frac{m_k}{k^i} \sum_{j=1}^{i-1} S(i,j)(k)_j \\ &=\; \sum_{k \in \mathbb{N}} \frac{m_k}{k^i}(k^i - (k)_i) \;=\; \sum_{k \in \mathbb{N}} m_k \left(1 - \frac{(k)_i}{k^i}\right), \quad i \in \mathbb{N}. \end{aligned}$$

For all $k \in \mathbb{N}$ and $i, j \in \mathbb{N}$ with $i < j$,

$$\mathbb{P}(Y(j, u^{(k)}) = i, Y(j+1, u^{(k)}) = i+1) = \frac{(k)_i}{k^j} S(j,i) \frac{k-i}{k} = \frac{(k)_{i+1}}{k^{j+1}} S(j,i).$$

The Siegmund dual fixation line therefore moves from state $i \in \mathbb{N}$ to state $j \in \mathbb{N}$ with $j > i$ at the rate (see Gaiser and Möhle, 2016, Proposition 2.5)

$$\gamma_{i,j} \;=\; \sum_{k\in\mathbb{N}} \mathbb{P}(Y(j, u^{(k)}) = i, Y(j+1, u^{(k)}) = i+1) m_k \;=\; S(j,i) \sum_{k\in\mathbb{N}} \frac{(k)_{i+1}}{k^{j+1}} m_k.$$

and the total rates are $\gamma_i := \sum_{j>i} \gamma_{i,j} = q_{i+1}$, $i \in \mathbb{N}$. The well-known (see Gaiser and Möhle, 2016, Theorem 2.9) Siegmund duality relations $\sum_{l=1}^{j} q_{i,l} =: q_{i,\le j} = \gamma_{j,\ge i} := \sum_{l=i}^{\infty} \gamma_{j,l}$, $i, j \in \mathbb{N}$, are hence equivalent to the combinatorial identities (see also Möhle, 2018, Eq. (19))

$$\frac{1}{k^i} \sum_{l=1}^{j} S(i,l)(k)_l \;=\; (k)_{j+1} \sum_{l=i}^{\infty} S(l,j) \frac{1}{k^{l+1}}, \qquad i, j, k \in \mathbb{N},$$

for the Stirling numbers of the second kind. Note that the left hand side is a finite sum whereas the right hand side is an infinite series. An analog combinatorial identity for the Stirling numbers of the first kind is provided in Möhle (2018, Eq. (22)). Since the time of the first coalescence of the associated Ξ-coalescent is exponentially distributed with parameter $\sum_{k\in\mathbb{N}} \nu(\{u^{(k)}\}) = \sum_{k\in\mathbb{N}} m_k = \nu(\Delta)$ it follows that this coalescent comes down from infinity if and only if $\nu(\Delta) = \infty$. Since ν is concentrated on Δ^* this coalescent has dust if and only if $\int_\Delta |u| \nu(\mathrm{d}u) = \nu(\Delta) = \sum_{k\in\mathbb{N}} m_k < \infty$. Having dust and not coming down from infinity are therefore the same properties. In this case (2.3) turns into

$$\begin{aligned}
\widetilde{\Phi}(q) \;&=\; \sum_{k\in\mathbb{N}} m_k \sum_{i=1}^{k} \left(1 - \left(1 - \frac{1}{k}\right)^q\right) \;=\; \sum_{k\in\mathbb{N}} k m_k \left(1 - \left(1 - \frac{1}{k}\right)^q\right) \\
&=\; \int_{(0,1]} \left(1 - (1-x)^q\right) \widetilde{\nu}(\mathrm{d}x), \qquad q \ge 0,
\end{aligned}$$

where the measure $\widetilde{\nu}$ (see Lemma 4.6) is discrete and assigns for each $k \in \mathbb{N}$ mass $k m_k$ to the point $1/k$. These formulas for $\widetilde{\Phi}(q)$ do not seem to simplify much further except for particular choices of the sequence $(m_k)_{k\in\mathbb{N}}$. If $\sum_{k\in\mathbb{N}} k m_k < \infty$ then $\widetilde{\Phi}(\infty) := \lim_{q\to\infty} \widetilde{\Phi}(q) = \sum_{k\in\mathbb{N}} k m_k < \infty$. If $\sum_{k\in\mathbb{N}} k m_k = \infty$ then $\widetilde{\Phi}(\infty) = \infty$, but the asymptotics of $\widetilde{\Phi}(q)$ as $q \to \infty$ is not easy to determine in general. For example, let $\alpha > 0$ and assume that $m_k := k^{-\alpha}$ for $k \in \mathbb{N}$. Then the associated coalescent has dust if and only if $\alpha > 1$. For $\alpha > 2$, $\widetilde{\Phi}(\infty) = \sum_{k\in\mathbb{N}} k^{1-\alpha} = \zeta(\alpha - 1) < \infty$, where ζ denotes the zeta function. For $\alpha \in (1, 2]$, as $q \to \infty$,

$$\begin{aligned}
\widetilde{\Phi}(q) \;&\sim\; \int_1^\infty t^{1-\alpha} \left(1 - \left(1 - \frac{1}{t}\right)^q\right) \mathrm{d}t \;=\; \int_0^1 \frac{1 - (1-x)^q}{x^{3-\alpha}} \,\mathrm{d}x \\
&=\; \begin{cases} \Psi(q+1) + \gamma \;\sim\; \log q & \text{if } \alpha = 2, \\ \dfrac{1}{2-\alpha} \left(\dfrac{\Gamma(q+1)\Gamma(\alpha-1)}{\Gamma(q+\alpha-1)} - 1 \right) \;\sim\; \dfrac{\Gamma(\alpha-1)}{2-\alpha} q^{2-\alpha} & \text{if } \alpha \in (1,2). \end{cases}
\end{aligned}$$

For example, for $\alpha = 3/2$, $\widetilde{\Phi}(q) \sim 2\sqrt{\pi q}$ as $q \to \infty$.

4. Proofs concerning the block counting process

Before we prove Theorem 2.1 we derive fundamental properties of the functions $\widetilde{\Phi}$ and r. We start with an analysis of the function $g : [0,\infty) \times \Delta \to [0,\infty)$, defined via

$$g(q,u) \;:=\; \sum_{i\in\mathbb{N}} \big(1-(1-u_i)^q\big), \qquad q \geq 0, u \in \Delta. \tag{4.1}$$

Lemma 4.1. *For every $u \in \Delta$ the function $g(.,u)$ is infinitely often differentiable on $(0,\infty)$ with derivatives*

$$g^{(k)}(q,u) \;:=\; \frac{\partial^k}{\partial^k q} g(q,u) \;=\; -\sum_{i\in\mathbb{N}} (1-u_i)^q (\log(1-u_i))^k, \qquad k \in \mathbb{N}, q > 0. \tag{4.2}$$

Proof: (of Lemma 4.1) We proceed similar as in the proof of Lemma 3.1 of Herriger and Möhle (2012). For $u = (1,0,0,\ldots) \in \Delta$, $g(q,u) = 1 - 0^q = 1 - \delta_{0,q}$ has a discontinuity at $q = 0$ but is infinitely often differentiable on $(0,\infty)$ and all derivatives vanish there. Thus, (4.2) obviously holds for $u = (1,0,0,\ldots)$. In the following it is therefore assumed that $(1,0,0,\ldots) \neq u \in \Delta$. Then $u_i < 1$ for all $i \in \mathbb{N}$, so we can take the logarithm of $1-u_i$. Note that $\sum_{i\in\mathbb{N}}(-\log(1-u_i)) < \infty$, since $-\log(1-t) = \sum_{n=1}^{\infty} t^n/n \leq t\sum_{n=0}^{\infty} t^n = t/(1-t) \leq 2t$ for all $t \in [0,\frac{1}{2}]$. In the following (4.2) is verified by induction on $k \in \mathbb{N}$.

For $m \in \mathbb{N}$ define $g_m : [0,\infty) \times \Delta \to \mathbb{R}$ via

$$g_m(q,u) \;:=\; \sum_{i=1}^{m} \big(1-(1-u_i)^q\big), \qquad q \geq 0, u \in \Delta.$$

Applying the inequality $t \leq -\log(1-t)$, $t \in [0,1)$, with $t := 1-(1-u_i)^q \in [0,1)$ it follows that $1-(1-u_i)^q \leq -q\log(1-u_i)$, $i \in \mathbb{N}$, $q \geq 0$. Therefore, for all $q \geq 0$,

$$|g(q,u) - g_m(q,u)| \;=\; \sum_{i>m} \big(1-(1-u_i)^q\big) \;\leq\; q\sum_{i>m} \big(-\log(1-u_i)\big) \;\to\; 0$$

as $m \to \infty$, so $g_m(.,u) \to g(.,u)$ uniformly on any compact set $K \subseteq [0,\infty)$. Furthermore, each g_m is differentiable with respect to q with continuous derivative $g_m'(q,u) = \sum_{i=1}^{m}(1-u_i)^q(-\log(1-u_i))$. From

$$\begin{aligned} |g_m'(q,u) - \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i))| \;&\leq\; \sum_{i>m}(1-u_i)^q(-\log(1-u_i)) \\ &\leq\; \sum_{i>m}(-\log(1-u_i)) \;\to\; 0 \end{aligned}$$

as $m \to \infty$ it follows that $g_m'(q,u) \to \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i))$ uniformly for all $q \in [0,\infty)$. Therefore, $g(.,u)$ is differentiable with derivative $g'(q,u) = \lim_{m\to\infty} g_m'(q,u) = \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i))$, $q \in [0,\infty)$. Thus, (4.2) holds for $k = 1$.

In order to prepare the induction step define $g_{m,k} : [0,\infty) \times \Delta \to \mathbb{R}$, $m,k \in \mathbb{N}$, via

$$g_{m,k}(q,u) \;:=\; -\sum_{i=1}^{m}(1-u_i)^q(\log(1-u_i))^k, \qquad q \geq 0, u \in \Delta.$$

The induction step from $k \in \mathbb{N}$ to $k+1$ works similarly as follows. Fix $k \in \mathbb{N}$. By the induction hypothesis,

$$\begin{aligned} &|g_{m,k}(q,u) - g^{(k)}(q,u)| \\ &= \left| -\sum_{i=1}^{m}(1-u_i)^q(\log(1-u_i))^k + \sum_{i\in\mathbb{N}}(1-u_i)^q(\log(1-u_i))^k \right| \\ &\le \sum_{i>m}(1-u_i)^q(-\log(1-u_i))^k \ \le \ \sum_{i>m}(-\log(1-u_i))^k \ \to \ 0 \end{aligned}$$

as $m \to \infty$, which shows that $g_{m,k}(.,u) \to g^{(k)}(.,u)$ uniformly on $[0,\infty)$. Furthermore each $g_{m,k}$ is differentiable with respect to q with continuous derivative $g'_{m,k}(q,u) = -\sum_{i=1}^{m}(1-u_i)^q(1-u_i)^{k+1} = g_{m,k+1}(q,u)$. From

$$\begin{aligned} |g'_{m,k}(q,u) + \sum_{i\in\mathbb{N}}(1-u_i)^q(\log(1-u_i))^{k+1}| \ &\le \ \sum_{i>m}(1-u_i)^q(-\log(1-u_i))^{k+1} \\ &\le \ \sum_{i>m}(-\log(1-u_i))^{k+1} \ \to \ 0 \end{aligned}$$

as $m \to \infty$ it follows that $g'_{m,k}(q,u) \to -\sum_{i\in\mathbb{N}}(1-u_i)^q(\log(1-u_i))^{k+1}$ uniformly for all $q \in [0,\infty)$. Therefore, $g^{(k)}(.,u)$ is differentiable with derivative

$$\begin{aligned} g^{(k+1)}(q,u) \ &= \ (g^{(k)})'(q,u) \ = \ \lim_{m\to\infty} g'_{m,k}(q,u) \\ &= \ \lim_{m\to\infty} g_{m,k+1}(q,u) \ = \ -\sum_{i\in\mathbb{N}}(1-u_i)^q(\log(1-u_i))^{k+1}. \end{aligned}$$

The induction is complete. □

Lemma 4.2. *The function $\widetilde{\Phi}$, defined via (2.3), is infinitely often differentiable on $(0,\infty)$ with derivatives*

$$\widetilde{\Phi}^{(k)}(q) \ = \ -\int_\Delta \sum_{i\in\mathbb{N}}(1-u_i)^q(\log(1-u_i))^k \, \nu(\mathrm{d}u), \qquad k \in \mathbb{N}, q > 0. \tag{4.3}$$

Remark 4.3. In general $\widetilde{\Phi}$ is not differentiable at 0, which is seen by choosing the measure ν such that $\int_\Delta |u| \, \nu(\mathrm{d}u) < \infty$ but $\int_\Delta \sum_{i\in\mathbb{N}}(-\log(1-u_i)) \, \nu(\mathrm{d}u) = \infty$. Such an example (even a Λ-coalescent where Λ has no mass at 1) is provided at the end of Remark 4.3 in Herriger and Möhle (2012). In this example the measure Λ assigns for each $m \in \mathbb{N}$ mass m^{-2} to $u_m := 1 - e^{-m}$. Therefore,

$$\widetilde{\Phi}(q) \ = \ \Phi(q) \ = \ \sum_{m=1}^{\infty} \frac{1-(1-u_m)^q}{m^2 u_m^2} \ = \ \sum_{m=1}^{\infty} \frac{1-e^{-mq}}{m^2(1-e^{-m})^2}, \qquad q \ge 0.$$

Thus, $\widetilde{\Phi}$ is a real-valued continuous function on $[0,\infty)$ and differentiable on $(0,\infty)$ with derivative $\widetilde{\Phi}'(q) = \sum_{m=1}^{\infty} e^{-mq}/(m(1-e^{-m})^2)$, $q > 0$, but $\widetilde{\Phi}$ is not differentiable at $q = 0$.

Remark 4.4. If ν has positive mass, say $k > 0$, at $(1,0,0,\ldots) \in \Delta$, then $\widetilde{\Phi}$ has even a discontinuity (a jump of size k) at $q = 0$.

Proof: (of Lemma 4.2) By definition,

$$\widetilde{\Phi}(q) \ = \ \int_\Delta g(q,u) \, \nu(\mathrm{d}u), \qquad q \ge 0, \tag{4.4}$$

where g is defined via (4.1). Having Lemma 4.1 in mind it suffices to verify that

$$\widetilde{\Phi}^{(k)}(q) \ = \ \int_\Delta g^{(k)}(q,u) \, \nu(\mathrm{d}u), \qquad k \in \mathbb{N}, q > 0. \tag{4.5}$$

We prove (4.5) by induction on $k \in \mathbb{N}$.

It is readily checked that for every $q > 0$ there exists a constant $c_q \in (0,\infty)$ such that $(1-t)^q(-\log(1-t)) \le c_q t$ for all $t \in [0,1)$. (Note that this statement does not hold for $q = 0$.) Applying this inequality with $t := u_i < 1$ it follows from Lemma 4.1 that $|g'(q,u)| = \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i)) \le \sum_{i\in\mathbb{N}} c_q u_i = c_q|u|$, which implies that for every $q > 0$ the map $g'(q,.)$ is ν-integrable. Moreover, choosing some $q_0 \in (0,\infty)$ it follows that $|g'(q,u)| = \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i)) \le \sum_{i\in\mathbb{N}}(1-u_i)^{q_0}(-\log(1-u_i)) = |g'(q_0,u)| \le c_{q_0}|u| =: h(u)$ for all $q \ge q_0$, where the dominating map h is ν-integrable. By the differentiation lemma, (4.4) is differentiable with respect to $q \in (0,\infty)$ and it is allowed to take the derivative below the integral, which shows that (4.5) holds for $k = 1$.

The induction step from $k \in \mathbb{N}$ to $k+1$ works essentially in the same way as follows. Let $k \in \mathbb{N}$. Again it is not hard to check that for every $q > 0$ there exists a constant $c_{k,q} \in (0,\infty)$ such that $(1-t)^q(-\log(1-t))^k \le c_{k,q}t$ for all $t \in [0,1)$. Applying this inequality with $t := u_i < 1$ it follows from Lemma 4.1 that $|g^{(k)}(q,u)| = \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i))^k \le c_{k,q}|u|$, which implies that for every $q > 0$ the map $g^{(k)}(q,.)$ is ν-integrable. Moreover, choosing some $q_0 \in (0,\infty)$ it follows that $|g^{(k)}(q,u)| = \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i))^k \le \sum_{i\in\mathbb{N}}(1-u_i)^{q_0}(-\log(1-u_i))^k = |g^{(k)}(q_0,u)| \le c_{k,q_0}|u| =: h_k(u)$ for all $q \ge q_0$, where the dominating map h_k is ν-integrable. By the differentiation lemma, (4.5) is differentiable with respect to $q \in (0,\infty)$ and it is allowed to take the derivative below the integral, which shows that (4.5) holds with k replaced by $k+1$. The induction is complete. □

Let us now turn to properties of the rate function r.

Lemma 4.5. *Let Ξ be a measure on Δ satisfying (1.1). Then the rate function r, defined via (2.4), is infinitely often differentiable on $(0,\infty)$ with first derivative*

$$\begin{aligned} r'(q) &= \frac{q\widetilde{\Phi}'(q) - \widetilde{\Phi}(q)}{q^2} \\ &= \int_\Delta \sum_{i\in\mathbb{N}} \frac{(1-u_i)^q - 1 - (1-u_i)^q \log((1-u_i)^q)}{q^2}\,\nu(\mathrm{d}u), \qquad q > 0. \end{aligned}$$

In particular, r is non-increasing on $(0,\infty)$. Moreover, $r(q) \to 0$ as $q \to \infty$.

Proof: (of Lemma 4.5) With $\widetilde{\Phi}$ also r is infinitely often differentiable on $(0,\infty)$ with first derivative

$$\begin{aligned} r'(q) &= \frac{\mathrm{d}}{\mathrm{d}q}\frac{\widetilde{\Phi}(q)}{q} = \frac{q\widetilde{\Phi}'(q) - \widetilde{\Phi}(q)}{q^2} \\ &= \frac{q\int_\Delta \sum_{i\in\mathbb{N}}(1-u_i)^q(-\log(1-u_i))\,\nu(\mathrm{d}u) - \int_\Delta \sum_{i\in\mathbb{N}}(1-(1-u_i)^q)\,\nu(\mathrm{d}u)}{q^2} \\ &= \int_\Delta \sum_{i\in\mathbb{N}} \frac{(1-u_i)^q - 1 - (1-u_i)^q \log((1-u_i)^q)}{q^2}\,\nu(\mathrm{d}u) \le 0, \qquad q > 0, \end{aligned}$$

since $t - 1 - t\log t \le 0$ for all $t \in [0,1]$. In particular, r is non-increasing on $(0,\infty)$.

In order to see that $r(q) \to 0$ as $q \to \infty$ fix $u = (u_i)_{i\in\mathbb{N}} \in \Delta$ and define $s_q(i) := (1-(1-u_i)^q)/q$, $q > 0$, $i \in \mathbb{N}$. Note that $s_q(i) \le 1/q \to 0$ as $q \to \infty$ and $s_q(i) \le u_i$ for all $i \in \mathbb{N}$ and $q \ge 1$, where the map $i \mapsto u_i$ is integrable with respect to the counting measure $\delta_\mathbb{N}$ on $\mathbb{N}$. Thus, by dominated convergence, for every $u \in \Delta$,

$$\sum_{i\in\mathbb{N}} \frac{1-(1-u_i)^q}{q} = \int_\mathbb{N} s_q(i)\,\delta_\mathbb{N}(\mathrm{d}i) \to 0$$

as $q \to \infty$. Moreover, from $s_q(i) \le u_i$ for all $q \ge 1$ and $i \in \mathbb{N}$ we conclude that

$$\sum_{i\in\mathbb{N}} \frac{1-(1-u_i)^q}{q} = \sum_{i\in\mathbb{N}} s_q(i) \le \sum_{i\in\mathbb{N}} u_i = |u|$$

for all $q \geq 1$ and $u \in \Delta$. Since the dominating map $u \mapsto |u|$ is ν-integrable, it follows that $r(q) \to 0$ as $q \to \infty$ again by dominated convergence. $\square$

The following lemma provides a Lévy–Khintchine representation for $\widetilde{\Phi}$ and shows that the integral over the infinite simplex Δ in (2.3) can be replaced by a simpler integral over the half-open unit interval $(0,1]$.

Lemma 4.6. *Let Ξ be a measure on Δ satisfying (1.1). Then there exists a unique measure $\widetilde{\nu}$ on $(0,1]$ having total mass $\widetilde{\nu}((0,1]) = \widetilde{\Phi}(\infty) \in [0,\infty]$ and satisfying $\int_{(0,1]} x\,\widetilde{\nu}(\mathrm{d}x) < \infty$ such that*

$$\widetilde{\Phi}(q) \;=\; \int_{(0,1]} \big(1-(1-x)^q\big)\,\widetilde{\nu}(\mathrm{d}x), \qquad q \geq 0. \tag{4.6}$$

Remark 4.7. The following proofs show that $\widetilde{\nu}(\{1\}) = \Xi(\{(1,0,0,\ldots)\})$. In particular, if the Ξ-coalescent has no star-shaped part, i.e. $\Xi(\{(1,0,0,\ldots)\}) = 0$, then $\widetilde{\Phi}$ is the Laplace exponent of a drift-free subordinator $\widetilde{Z}$ with state space $[0,\infty)$ and Lévy measure $\widetilde{\varrho} := \widetilde{\nu}_T$, where $\widetilde{\nu}_T$ denotes the image of the measure $\widetilde{\nu}$ under the transformation $T : (0,1) \to (0,\infty)$ defined via $T(x) := -\log(1-x)$ for all $x \in (0,1)$.

Proof: (of Lemma 4.6) Two proofs of (4.6) are provided. The first proof is measure theoretic and provides an explicit formula (see (4.8)) for the measure $\widetilde{\nu}$ in terms of the measure ν. The second proof is based on a Bernstein function argument. Both proofs are relatively short.

Proof 1. Consider the measure space $(\Delta, \mathcal{B}(\Delta), \nu)$ and the measurable space $(S, \mathcal{S}) := ([0,1], \mathcal{B}([0,1]))$. The point process $\xi : \Delta \times \mathcal{S} \to \mathbb{N}_0 \cup \{\infty\}$, defined via $\xi(u,B) := \sum_{i\in\mathbb{N}} 1_B(u_i)$ for all $u \in \Delta$ and $B \in \mathcal{S}$, has intensity measure

$$\mu(B) \;:=\; \mathbb{E}_\nu(\xi(.,B)) \;:=\; \int_\Delta \xi(u,B)\,\nu(\mathrm{d}u) \;=\; \int_\Delta \sum_{i\in\mathbb{N}} 1_B(u_i)\,\nu(\mathrm{d}u), \qquad B \in \mathcal{S},$$

and, hence, Campbell's formula

$$\begin{aligned} \mathbb{E}_\nu\bigg(\int_{[0,1]} f(x)\,\xi(.,\mathrm{d}x)\bigg) \;&=\; \int_\Delta \int_{[0,1]} f(x)\,\xi(u,\mathrm{d}x)\,\nu(\mathrm{d}u) \\ &=\; \int_\Delta \sum_{i\in\mathbb{N}} f(u_i)\,\nu(\mathrm{d}u) \;=\; \int_{[0,1]} f(x)\,\mu(\mathrm{d}x) \end{aligned} \tag{4.7}$$

holds for $f = 1_B$ and hence for all nonnegative measurable functions $f : [0,1] \to [0,\infty]$. Fix $q \geq 0$. Choosing $f(x) := f_q(x) := 1-(1-x)^q$ shows that

$$\widetilde{\Phi}(q) \;=\; \int_\Delta \sum_{i\in\mathbb{N}} f_q(u_i)\,\nu(\mathrm{d}u) \;=\; \int_{[0,1]} f_q(x)\,\mu(\mathrm{d}x) \;=\; \int_{(0,1]} (1-(1-x)^q)\,\widetilde{\nu}(\mathrm{d}x),$$

where $\widetilde{\nu}$ denotes the restriction of μ to $(0,1]$, i.e.

$$\widetilde{\nu}(B) \;:=\; \mu(B) \;=\; \int_\Delta \sum_{i\in\mathbb{N}} 1_B(u_i)\,\nu(\mathrm{d}u) \tag{4.8}$$

for all Borel sets $B \subseteq (0,1]$. Thus (4.6) holds. $\square$

Proof 2. By Lemma 4.2, $\widetilde{\Phi}$ is infinitely often differentiable on $(0,\infty)$ with derivatives (4.3). In particular, $(-1)^{k-1}\widetilde{\Phi}^{(k)}(q) = \int_\Delta \sum_{i\in\mathbb{N}} (1-u_i)^q (-\log(1-u_i))^k\,\nu(\mathrm{d}u) \geq 0$ for all $k \in \mathbb{N}$ and $q > 0$. Thus, $\widetilde{\Phi}$, restricted to $(0,\infty)$, is a Bernstein function in the sense of Definition 3.1 of Schilling et al. (2012). By Schilling et al. (2012, Theorem 3.2), $\widetilde{\Phi}$ has a Lévy–Khintchine representation, i.e. there

exist constants $a, b \geq 0$ and a unique measure $\widetilde{\varrho}$ on $(0,\infty)$ satisfying $\int_{(0,\infty)}(1\wedge y)\widetilde{\varrho}(\mathrm{d}y) < \infty$ such that

$$\widetilde{\Phi}(q) \;=\; a + bq + \int_{(0,\infty)}(1-e^{-qy})\widetilde{\varrho}(\mathrm{d}y) \;=\; a + bq + \int_{(0,1)}\left(1-(1-x)^q\right)\widetilde{\nu}(\mathrm{d}x), \qquad q>0,$$

where $\widetilde{\nu} := \widetilde{\varrho}_\varphi$ is the image of $\widetilde{\varrho}$ under the transformation $\varphi : (0,\infty) \to (0,1)$ defined via $\varphi(y) := 1-e^{-y}$ for all $y \in (0,\infty)$.

Assume now first that ν has no mass at $(1,0,0,\ldots) \in \Delta$. Then $\widetilde{\Phi}$ is continuous at $q=0$ and from $\widetilde{\Phi}(0)=0$ it follows that $a=0$. Moreover, $b = \lim_{q\to\infty}\widetilde{\Phi}(q)/q = \lim_{q\to\infty} r(q) = 0$ by Lemma 4.5. Thus,

$$\widetilde{\Phi}(q) \;=\; \int_{(0,1)}\left(1-(1-x)^q\right)\widetilde{\nu}(\mathrm{d}x), \qquad q\geq 0. \tag{4.9}$$

Note that this formula obviously holds for $q=0$, since in this case both sides in (4.9) are equal to 0.

If ν has mass $k := \nu(\{(1,0,0,\ldots)\}) = \Xi(\{(1,0,0,\ldots)\}) \in [0,\infty)$ at $(1,0,0,\ldots) \in \Delta$, then extend the measure $\widetilde{\nu}$ on $(0,1)$ to a measure on $(0,1]$, again denoted by $\widetilde{\nu}$, via $\widetilde{\nu}(\{1\}) := k$, and note that one has to add to the above expression (4.9) for $\widetilde{\Phi}(q)$ the part

$$\int_{\{(1,0,0,\ldots)\}}\sum_{i\in\mathbb{N}}(1-(1-u_i)^q)\,\nu(\mathrm{d}u) \;=\; (1-0^q)k \;=\; \int_{\{1\}}(1-(1-x)^q)\,\widetilde{\nu}(\mathrm{d}x),$$

which shows that (4.6) holds.

The total mass $\widetilde{\nu}((0,1]) = \widetilde{\Phi}(\infty)$ of the measure $\widetilde{\nu}$ might be infinite, but its first moment $\int_{(0,1]} x\,\widetilde{\nu}(\mathrm{d}x) = \widetilde{\Phi}(1)$ is finite. Clearly, the measure $\widetilde{\nu}$ is uniquely determined by the values (4.6). □

Remark 4.8. The measure $\widetilde{\Lambda}(\mathrm{d}x) := x^2\widetilde{\nu}(\mathrm{d}x)$ satisfies

$$\int_{(0,1]} x^{-1}\widetilde{\Lambda}(\mathrm{d}x) \;=\; \int_{(0,1]} x\,\widetilde{\nu}(\mathrm{d}x) \;<\; \infty,$$

which implies that the corresponding $\widetilde{\Lambda}$-coalescent has dust. In this way we can define for each Ξ-coalescent with dust a $\widetilde{\Lambda}$-coalescent (allowing only for multiple collisions) with dust such that both, the Ξ-coalescent and the $\widetilde{\Lambda}$-coalescent, have the same function $\widetilde{\Phi}$. If the $\widetilde{\Lambda}$-coalescent is in a partition with $b \in \mathbb{N}\setminus\{1\}$ blocks then each possible merger of $k \in \{2,\ldots,b\}$ blocks into a single block is occurring at the rate

$$\begin{aligned}\widetilde{\lambda}_{b,k} \;&=\; \int_{(0,1]} x^k(1-x)^{b-k}\widetilde{\nu}(\mathrm{d}x) \;=\; \int_{[0,1]} x^k(1-x)^{b-k}\mu(\mathrm{d}x)\\ &=\; \int_\Delta \sum_{i\in\mathbb{N}} u_i^k(1-u_i)^{b-k}\,\nu(\mathrm{d}u), \qquad 2\leq k\leq b,\end{aligned} \tag{4.10}$$

where the last equality holds by Campbell's formula (4.7). The rates $\widetilde{\lambda}_{b,k}$ can hence be expressed in terms of the characterizing measure $\nu(\mathrm{d}u) = \Xi(\mathrm{d}u)/(u,u)$ of the exchangeable coalescent via (4.10).

We now prove Theorem 2.1.

Proof of Theorem 2.1: Fix $f \in C^2(E)$, $n \in \mathbb{N}$, $x \in E_n$ and $u \in \Delta$. Define $u_0 := 1-|u|$ and $\widetilde{Y} := \widetilde{Y}(nx,u) := Y(nx,u)/(nx) - u_0$ for convenience. Note that $-1 \leq \widetilde{Y} \leq 1$, so $\widetilde{Y}$ may take negative values with positive probability, but the mean $\mathbb{E}(\widetilde{Y}) = \sum_{i\in\mathbb{N}}(1-(1-u_i)^{nx})/(nx) \in [0,1]$

is non-negative. By Taylor expansion,

$$\begin{aligned} & f\left(\frac{Y(nx,u)}{n}\right) - f(xu_0) \\ &= f'(xu_0)\left(\frac{Y(nx,u)}{n} - xu_0\right) + f''(\xi)\left(\frac{Y(nx,u)}{n} - xu_0\right)^2 \\ &= f'(xu_0)x\widetilde{Y} + f''(\xi)x^2\widetilde{Y}^2 \end{aligned}$$

for some $\xi = \xi(n,x,u)$ taking values between $Y(nx,u)/n$ and xu_0, Taking expectation yields

$$\mathbb{E}\left(f\left(\frac{Y(nx,u)}{n}\right)\right) - f(xu_0) \;=\; f'(xu_0)x\mathbb{E}(\widetilde{Y}) + x^2\mathbb{E}(f''(\xi)\widetilde{Y}^2)$$

and taking the absolute value leads to

$$\left|\mathbb{E}\left(f\left(\frac{Y(nx,u)}{n}\right)\right) - f(xu_0)\right| \;\le\; C_1x\mathbb{E}(\widetilde{Y}) + C_2x^2\mathbb{E}(\widetilde{Y}^2) \;\le\; C_1x\mathbb{E}(\widetilde{Y}) + C_2x\mathbb{E}(\widetilde{Y}^2),$$

where $C_1 := \|f'\| < \infty$ and $C_2 := \|f''\| < \infty$. By Lemma 6.1 provided in Section 6, the concentration inequality $\mathbb{E}(\widetilde{Y}^2) \le 2\mathbb{E}(\widetilde{Y})$ holds. Thus, if we define $C := C_f := C_1 + 2C_2$, then

$$\begin{aligned} \left|\mathbb{E}\left(f\left(\frac{Y(nx,u)}{n}\right)\right) - f(xu_0)\right| &\le Cx\mathbb{E}(\widetilde{Y}) = C\sum_{i\in\mathbb{N}} \frac{1-(1-u_i)^{nx}}{n} \\ &\le C\sum_{i\in\mathbb{N}} \frac{1-(1-u_i)^{n}}{n}. \end{aligned}$$

From the integral representations (2.2) and (1.4) for the generators A_n and A we conclude that

$$\begin{aligned} |A_n\pi_n f(x) - \pi_n Af(x)| &= \left|\int_\Delta \left(\mathbb{E}\left(f\left(\frac{Y(nx,u)}{n}\right)\right) - f(xu_0)\right)\nu(\mathrm{d}u)\right| \\ &\le \int_\Delta \left|\mathbb{E}\left(f\left(\frac{Y(nx,u)}{n}\right)\right) - f(xu_0)\right|\nu(\mathrm{d}u) \\ &\le C\int_\Delta \sum_{i\in\mathbb{N}} \frac{1-(1-u_i)^n}{n}\,\nu(\mathrm{d}u) = Cr(n), \end{aligned}$$

which is the desired result. □

Remark 4.9. For Λ-coalescents the sharper inequality $\mathbb{E}(\widetilde{Y}^2) \le \mathbb{E}(\widetilde{Y})$ holds (see Proposition 5.1 in Section 5) and the previous proof can hence be performed with the optimized constant $C := C_1 + C_2$.

Proof: (of Corollary 2.3) Fix $t \ge 0$ and $f \in C^2(E)$. By (1.3), $T_tf(x) = \mathbb{E}(f(xS_t))$, $x \in E$. An application of the differentiation lemma shows that T_tf is twice differentiable with derivatives $(T_tf)'(x) = \mathbb{E}(S_tf'(xS_t))$ and $(T_tf)''(x) = \mathbb{E}(S_t^2f''(xS_t))$, $x \in E$. Moreover, exploiting the fact that $\|f''\| < \infty$ it follows by dominated convergence that $(T_tf)''$ is continuous. Thus, $T_tf \in C^2(E)$ showing that $T_t(C^2(E)) \subseteq C^2(E)$.

From estimates for semigroups (see, for example, Ethier and Kurtz, 1986, p. 29, Eq. (6.2)) we conclude that, for all $n \in \mathbb{N}$, $\|T_t^{(n)}\pi_nf - \pi_nT_tf\| \le \int_0^t \|A_n\pi_nT_sf - \pi_nAT_sf\|\,\mathrm{d}s \le r(n)\int_0^t C_{T_sf}\,\mathrm{d}s$, where the last inequality holds by Theorem 2.1, applied to the function $T_sf \in C^2(E)$ instead of f. The result follows, since $C_{T_sf} = \|(T_sf)'\| + 3\|(T_sf)''\| \le \|f'\|\,\mathbb{E}(S_s) + 3\|f''\|\,\mathbb{E}(S_s^2) \le \|f'\| + 2\|f''\| = C_f$. □

Remark 4.10. The previous proof shows that $T_t(C^2(E)) \subseteq C^2(E)$ for all $t \ge 0$. Since $C^2(E)$ is dense in $B(E)$, the standard core theorem (see, for example, Ethier and Kurtz, 1986, p. 17, Proposition 3, applied with $L := B(E)$ and $D_0 := D := C^2(E)$) ensures that $C^2(E)$ is a core for the generator A of the semigroup $(T_t)_{t\ge0}$.

Proof of Corollary 2.4: Two proofs of Corollary 2.4 are provided. The first proof is based on the generator bound in Theorem 2.1. The second proof is rather short and based on the corresponding semigroup bound in Corollary 2.3.

Proof 1. (via generators) If g_k denotes the monomial defined via $g_k(x) := x^k$ for all $x \in E$, then, by (1.4), $Ag_k(x) = \int_\Delta ((x(1-|u|))^k - x^k)\,\nu(\mathrm{d}u) = x^k \int_\Delta ((1-|u|)^k - 1)\,\nu(\mathrm{d}u) = -\Phi(k)x^k$, where Φ denotes the Laplace exponent (1.2) of the associated subordinator. In particular, $AD_k \subseteq D_k$ for all $k \in \mathbb{N}$, where D_k denotes the finite-dimensional space of all polynomials $f : E \to \mathbb{R}$ of degree less than or equal to k. The space $D := \bigcup_{k\in\mathbb{N}} D_k$ of all polynomials is dense in $L := C(E)$. Exploiting the fact that $AD_k \subseteq D_k$ for all $k \in \mathbb{N}$ it follows as in the proof of Theorem 3.5 on p. 19 of Ethier and Kurtz (1986) that $(\lambda - A)(D) = D$ for some $\lambda \in (0, \infty)$. In particular, the range $(\lambda - A)(D) = D$ is dense in L. Thus, by Ethier and Kurtz (1986, p. 17, Proposition 3.1), D is a core for A. From Theorem 2.1 and Lemma 4.5 it follows that $\lim_{n\to\infty} \|A_n\pi_n f - \pi_n Af\| = 0$ for all $f \in D$. From Ethier and Kurtz (1986, p. 28, Theorem 6.1) we conclude that $\lim_{n\to\infty} \|T_t^{(n)}\pi_n f - \pi_n T_t f\| = 0$ for all $f \in L$ and $t \geq 0$. The statement therefore follows from Ethier and Kurtz (1986, p. 232, Corollary 8.7), applied with $G_n := E_n$. □

Proof 2. (via semigroups) Corollary 2.3 yields $\lim_{n\to\infty} \|T_t^{(n)}\pi_n f - \pi_n T_t f\| = 0$ for all $t \geq 0$ and $f \in C^2(E)$, and hence also for all $f \in L := C(E)$, since $C^2(E)$ is dense in L. The statement therefore follows again from Ethier and Kurtz (1986, p. 232, Corollary 8.7), applied with $G_n := E_n$. □

5. A phenomenon of concentration

In this section we restrict our attention to the subclass of coalescents with multiple collisions (Λ-coalescents). In fact, what follows depends only on the random variables $Y(n, u)$ defined via (2.1), so we are dealing with a problem which arises in the context of Kingman's paintbox construction (Kingman, 1978a,b) or, equivalently, in Karlin's putting balls into boxes experiment (Karlin, 1967) further studied by Dutko (1989) and Gnedin et al. (2007). We are interested in the scaled random variables

$$\widetilde{Y}(n, u) \;:=\; \frac{Y(n, u)}{n} - (1 - u) \qquad n \in \mathbb{N}, u \in [0, 1], \tag{5.1}$$

where $Y(n, u) := X_0(n, u) + 1_{\{X_1(n,u)>0\}}$ (see Eq. (2.1)). We will prove in this section (see Proposition 5.1 (ii)) that these random variables satisfy for all $n \in \mathbb{N}$ and $u \in [0, 1]$ the concentration inequality

$$\mathbb{E}((\widetilde{Y}(n, u))^2) \;\leq\; \mathbb{E}(\widetilde{Y}(n, u)). \tag{5.2}$$

Similar results that the second moment is bounded by the expectation are well-known for the number of occupied boxes in Karlin's model (see, for example, p. 153 of Gnedin et al., 2007). The inequality (5.2) looks rather simple. However, we have not been able to find a short proof. In particular, we have not been able to find a probabilistic proof of this inequality. The following analytic proof is based on explicit calculations and carefully designed bounds. The proof shows that the inequality (5.2) is more involved as it seems to be at a first glance. The conjecture that the same inequality holds not only for Λ-coalescents but as well for the full class of all Ξ-coalescents will be discussed in the following Section 6.

Proposition 5.1. *For all $n \in \mathbb{N}$ and $u \in [0, 1]$,*

$$\mathbb{E}(Y(n, u)) \;=\; n(1 - u) + 1 - (1 - u)^n \tag{5.3}$$

and

$$\mathrm{Var}(Y(n, u)) \;=\; nu(1 - u) + (1 - u)^n\big(1 - (1 - u)^n\big) - 2nu(1 - u)^n. \tag{5.4}$$

Moreover, the scaled random variable $\widetilde{Y}(n, u) := Y(n, u)/n - (1-u)$ has the following two properties.

(i) *For every $u \in [0, 1]$ the mean $\mathbb{E}(\widetilde{Y}(n, u)) = (1 - (1 - u)^n)/n$ is non-increasing in $n \in \mathbb{N}$.*

(ii) *For all* $n \in \mathbb{N}$ *and* $u \in [0,1]$ *the concentration inequality* $\mathbb{E}((\widetilde{Y}(n,u))^2) \leq \mathbb{E}(\widetilde{Y}(n,u))$ *holds.*

Proof: We suppress the parameters n and u and write Y for $Y(n,u)$ and X_i for $X_i(n,u)$ for convenience. We have $Y = X_0 + 1_{\{X_1>0\}}$, where (X_0, X_1) has a multinomial distribution with parameters $n \in \mathbb{N}$ and $(1-u, u)$. Taking expectation and noting that $\mathbb{E}(X_0) = n(1-u)$ and $\mathbb{E}(1_{\{X_1>0\}}) = 1 - \mathbb{P}(X_1 = 0) = 1 - (1-u)^n$ yields (5.3). Eq. (5.4) follows from $\mathrm{Var}(Y) = \mathrm{Var}(X_0) + \mathrm{Var}(1_{\{X_1>0\}}) + 2\mathrm{Cov}(X_0, 1_{\{X_1>0\}})$ and the explicit formulas $\mathrm{Var}(X_0) = nu(1-u)$, $\mathrm{Var}(1_{\{X_1>0\}}) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_1 > 0) = (1-u)^n(1-(1-u)^n)$ and

$$\begin{aligned}
\mathrm{Cov}(X_0, 1_{\{X_1>0\}}) &= \mathrm{Cov}(X_0, 1 - 1_{\{X_1=0\}}) = -\mathrm{Cov}(X_0, 1_{\{X_1=0\}}) \\
&= \mathbb{E}(X_0)\mathbb{E}(1_{\{X_1=0\}}) - \mathbb{E}(X_0 1_{\{X_1=0\}}) \\
&= n(1-u)(1-u)^n - n\mathbb{P}(X_1 = 0) \\
&= n(1-u)^{n+1} - n(1-u)^n = -nu(1-u)^n.
\end{aligned}$$

From (5.3) it follows that $\mathbb{E}(\widetilde{Y}) = \mathbb{E}(Y)/n - (1-u) = (1-(1-u)^n)/n$. In order to see that this expression is non-increasing in $n \in \mathbb{N}$ we verify that the map $m(x) := (1-(1-u)^x)/x$, $x \in (0,\infty)$, is non-increasing on $(0,\infty)$. We have $m'(x) = ((1-u)^x - 1 - (1-u)^x \log((1-u)^x))/x^2 \leq 0$, since $t - 1 - t\log t \leq 0$ for all $t \in [0,1]$. Thus, the map m is non-increasing on $(0,\infty)$.

It remains to verify (ii). From (5.3) and (5.4) we conclude that

$$\begin{aligned}
\mathbb{E}(\widetilde{Y}^2) &= \mathrm{Var}(\widetilde{Y}) + (\mathbb{E}(\widetilde{Y}))^2 = \frac{\mathrm{Var}(Y)}{n^2} + \left(\frac{1-(1-u)^n}{n}\right)^2 \\
&= \frac{u(1-u)}{n} + \frac{(1-u)^n(1-(1-u)^n)}{n^2} - \frac{2u(1-u)^n}{n} + \frac{(1-(1-u)^n)^2}{n^2} \\
&= \frac{u(1-u)}{n} - \frac{2u(1-u)^n}{n} + \frac{1-(1-u)^n}{n^2}.
\end{aligned}$$

Therefore, the function $p_n : [0,1] \to \mathbb{R}$, defined via $p_n(u) := \mathbb{E}(\widetilde{Y}) - \mathbb{E}(\widetilde{Y}^2)$, is a polynomial of degree $n+1$ of the form

$$p_n(u) = \frac{1-(1-u)^n}{n} - \frac{u(1-u)}{n} + \frac{2u(1-u)^n}{n} - \frac{1-(1-u)^n}{n^2}. \tag{5.5}$$

In particular $p_1(u) = u(1-u) \geq 0$. In order to check that $p_n(u)$ is nonnegative we can therefore assume that $n \geq 2$. We have $p_n(0) = 0$ and $p_n(1) = 1/n - 1/n^2 \geq 0$. A straightforward calculation shows that

$$p_n'(u) = (2u-1)\left(\frac{1-(1-u)^{n-1}}{n} - (1-u)^{n-1}\right).$$

It follows that p_n has a local minimum at $u = 1/2$ (independent of n). (The polynomial p_n has as well a local maximum at $u = 1 - (\frac{1}{n+1})^{\frac{1}{n-1}}$, which is however not important in our context.) Since the value $p_n(1/2) = (3n - 4 + (1/2)^{n-2})/(4n^2)$ of p_n at this minimum is still positive, it follows that $p_n(u) \geq 0$. Thus, the concentration inequality $\mathbb{E}(\widetilde{Y}^2) \leq \mathbb{E}(\widetilde{Y})$ holds. □

Remark 5.2. The previous proof shows that the polynomial p_n in (5.5) is nonnegative on $[0,1]$. By (a linear transformed variant of) Lasserre (2010, p. 24, Theorem 2.8) there exists $d \in \mathbb{N}_0$ and coefficients $c_{ij} = c_{ij}(n) \geq 0$, $i,j \in \mathbb{N}_0$ with $i + j \leq d$, such that $p_n(u) = \sum_{i+j\leq d} c_{ij} u^i (1-u)^j$. It does not seem to be straightforward to determine the coefficients c_{ij} explicitly.

Remark 5.3. From part (ii) of Proposition 5.1 it follows that the quantity $D(\widetilde{Y}) := \mathrm{Var}(\widetilde{Y})/\mathbb{E}(\widetilde{Y})$, sometimes called the index of dispersion in the literature (see, for example, Cox and Lewis, 1966, p. 72, Eq. (3)), never exceeds 1.

6. The phenomenon of concentration for Ξ-coalescents

It is natural to conjecture that the concentration inequality (5.2) holds not only for the class of Λ-coalescents but as well for the full class of all Ξ-coalescents. Unfortunately, we have not been able to prove this conjecture. In this last section the less strict inequality $\mathbb{E}((\widetilde{Y}(n,u))^2) \le 2\mathbb{E}(\widetilde{Y}(n,u))$ is verified and some information is provided, which could be helpful in proving the conjecture that $\mathbb{E}((\widetilde{Y}(n,u))^2) \le \mathbb{E}(\widetilde{Y}(n,u))$. The main results on the random variable $Y(n,u)$, defined via (2.1), are collected in the following lemma.

Lemma 6.1. *For all $n \in \mathbb{N}$ and $u \in \Delta$,*

$$\mathbb{E}(Y(n,u)) \;=\; n(1-|u|) + \sum_{i\in\mathbb{N}} \big(1-(1-u_i)^n\big) \tag{6.1}$$

and

$$\begin{aligned}\mathrm{Var}(Y(n,u)) \;&=\; n|u|(1-|u|) + \sum_{i\in\mathbb{N}} (1-u_i)^n\big(1-(1-u_i)^n\big)\\ &\quad -2n(1-|u|)\sum_{i\in\mathbb{N}} u_i(1-u_i)^{n-1} - 2\sum_{\substack{i,j\in\mathbb{N}\\ i<j}} \big((1-u_i)^n(1-u_j)^n - (1-u_i-u_j)^n\big).\end{aligned} \tag{6.2}$$

Moreover, $\widetilde{Y}(n,u) := Y(n,u)/n - (1-|u|)$ satisfies the concentration inequality

$$\mathbb{E}((\widetilde{Y}(n,u))^2) \;\le\; 2\,\mathbb{E}(\widetilde{Y}(n,u)) \;=\; \frac{2}{n}\sum_{i\in\mathbb{N}} \big(1-(1-u_i)^n\big), \qquad n\in\mathbb{N}, u\in\Delta. \tag{6.3}$$

Remark 6.2. Write $\widetilde{Y} := \widetilde{Y}(n,u)$ for convenience. Eq. (6.3) in particular implies that $\mathrm{Var}(\widetilde{Y}) \le 2\mathbb{E}(\widetilde{Y})$. For the number $K_n := \sum_{i\in\mathbb{N}} 1_{\{X_i(n,u)>0\}}$ of occupied boxes (disregarding box 0) similar statements that the variance $\mathrm{Var}(K_n)$ is bounded by its expectation $\mathbb{E}(K_n)$ are stated on p. 153 of Gnedin et al. (2007).

Proof: Recall the notation $|u| = \sum_{i\in\mathbb{N}} u_i$ and $u_0 := 1-|u|$ for $u = (u_1, u_2, \ldots) \in \Delta$. We suppress the parameters n and u and write Y for $Y(n,u)$ and X_i for $X_i(n,u)$, $i \in \mathbb{N}$, for convenience. We have $Y = X_0 + \sum_{i\in\mathbb{N}} 1_{\{X_i>0\}}$, where $(X_0, X_1, \ldots)$ has an infinite multinomial distribution with parameters $n \in \mathbb{N}$ and $(u_0, u_1, \ldots)$. Taking expectation and noting that $\mathbb{E}(X_0) = nu_0$ and $\mathbb{E}(1_{\{X_i>0\}}) = 1 - \mathbb{P}(X_i = 0) = 1-(1-u_i)^n$ yields (6.1). Eq. (6.2) follows from

$$\begin{aligned}\mathrm{Var}(Y) \;&=\; \mathrm{Var}(X_0) + \sum_{i\in\mathbb{N}} \mathrm{Var}(1_{\{X_i>0\}})\\ &\quad +2\sum_{i\in\mathbb{N}} \mathrm{Cov}(X_0, 1_{\{X_i>0\}}) + 2\sum_{i<j} \mathrm{Cov}(1_{\{X_i>0\}}, 1_{\{X_j>0\}})\end{aligned}$$

and the explicit formulas $\mathrm{Var}(X_0) = nu_0(1-u_0)$, $\mathrm{Var}(1_{\{X_i>0\}}) = \mathbb{P}(X_i = 0)\mathbb{P}(X_i > 0) = (1-u_i)^n(1-(1-u_i)^n)$,

$$\begin{aligned}\mathrm{Cov}(X_0, 1_{\{X_i>0\}}) \;&=\; \mathrm{Cov}(X_0, 1-1_{\{X_i=0\}}) \;=\; -\mathrm{Cov}(X_0, 1_{\{X_i=0\}})\\ &=\; \mathbb{E}(X_0)\mathbb{E}(1_{\{X_i=0\}}) - \mathbb{E}(X_0 1_{\{X_i=0\}})\\ &=\; \mathbb{E}(X_0)\mathbb{P}(X_i=0) - \sum_{j=1}^{n} j\binom{n}{j} u_0^j (1-u_0-u_i)^{n-j}\\ &=\; nu_0(1-u_i)^n - nu_0(1-u_i)^{n-1} \;=\; -nu_0u_i(1-u_i)^{n-1}\end{aligned}$$

and

$$\begin{aligned}\mathrm{Cov}(1_{\{X_i>0\}}, 1_{\{X_j>0\}}) &= \mathrm{Cov}(1-1_{\{X_i=0\}}, 1-1_{\{X_j=0\}})\\ &= \mathrm{Cov}(1_{\{X_i=0\}}, 1_{\{X_j=0\}})\\ &= \mathbb{P}(X_i=X_j=0)-\mathbb{P}(X_i=0)\mathbb{P}(X_j=0)\\ &= (1-u_i-u_j)^n-(1-u_i)^n(1-u_j)^n\end{aligned}$$

for all $i,j\in\mathbb{N}$ with $i\neq j$. In order to prove (6.3) we write $\widetilde{Y}=X_0/n-u_0+K_n/n$, where $K_n:=\sum_{i\in\mathbb{N}}1_{\{X_i>0\}}$ denotes the number of occupied boxes (disregarding box 0). We have

$$\mathbb{E}(\widetilde{Y}^2) \;=\; \mathbb{E}\bigg(\Big(\frac{X_0}{n}-u_0\Big)^2\bigg)+2\,\mathbb{E}\bigg(\Big(\frac{X_0}{n}-u_0\Big)\frac{K_n}{n}\bigg)+\mathbb{E}\bigg(\Big(\frac{K_n}{n}\Big)^2\bigg).$$

The first expectation on the right hand side is equal to the variance of X_0/n. The second expectation on the right hand side is non-positive since the covariance $\mathrm{Cov}(X_0,K_n)=\sum_{i\in\mathbb{N}}\mathrm{Cov}(X_0,1_{\{X_i>0\}})=-nu_0\sum_{i\in\mathbb{N}}u_i(1-u_i)^{n-1}$ of X_0 and K_n is non-positive. We therefore obtain the upper bound

$$\mathbb{E}(\widetilde{Y}^2) \;\leq\; \mathrm{Var}\Big(\frac{X_0}{n}\Big)+\mathbb{E}\bigg(\Big(\frac{K_n}{n}\Big)^2\bigg).$$

Using that $u_i=1-(1-u_i)\leq 1-(1-u_i)^n$, $n\in\mathbb{N}$, the first quantity on the right hand side is bounded by

$$\mathrm{Var}\Big(\frac{X_0}{n}\Big)=\frac{|u|(1-|u|)}{n}\leq\frac{1}{n}\sum_{i\in\mathbb{N}}u_i\leq\frac{1}{n}\sum_{i\in\mathbb{N}}\big(1-(1-u_i)^n\big)=\mathbb{E}\Big(\frac{K_n}{n}\Big)=\mathbb{E}(\widetilde{Y}).$$

From $0\leq K_n/n\leq 1$ we conclude as well that

$$\mathbb{E}\bigg(\Big(\frac{K_n}{n}\Big)^2\bigg) \;\leq\; \mathbb{E}\Big(\frac{K_n}{n}\Big) \;=\; \mathbb{E}(\widetilde{Y}),$$

which yields the desired inequality $\mathbb{E}(\widetilde{Y}^2)\leq 2\mathbb{E}(\widetilde{Y})$. □

Problem 6.3. For $u_2=u_3=\cdots=0$ (which corresponds to the Λ-coalescent case), the sharper inequality $\mathbb{E}(\widetilde{Y}^2)\leq\mathbb{E}(\widetilde{Y})$ holds (see Proposition 5.1 in Section 5). It is hence natural to conjecture that this sharper inequality $\mathbb{E}(\widetilde{Y}^2)\leq\mathbb{E}(\widetilde{Y})$ holds as well for arbitrary $u=(u_r)_{r\in\mathbb{N}}\in\Delta$ (which corresponds to the Ξ-coalescent case). We have neither been able to verify this conjecture nor to find a counterexample. In order to verify this conjecture one would need to show the nonnegativity of the function $p_n:\Delta\to\mathbb{R}$, defined via

$$p_n(u) \;:=\; \mathbb{E}(\widetilde{Y})-\mathbb{E}(\widetilde{Y}^2) \;=\; \mathbb{E}(\widetilde{Y})-(\mathbb{E}(\widetilde{Y}))^2-\mathrm{Var}(\widetilde{Y}) \;=\; \mathbb{E}(\widetilde{Y})-(\mathbb{E}(\widetilde{Y}))^2-\frac{\mathrm{Var}(Y)}{n^2}.$$

Noting that $\mathbb{E}(\widetilde{Y}) = \frac{1}{n}\sum_{i\in\mathbb{N}}(1-(1-u_i)^n)$ and, hence, $(\mathbb{E}(\widetilde{Y}))^2 = \frac{1}{n^2}\sum_{i\in\mathbb{N}}(1-(1-u_i)^n)(1-(1-u_i)^n) + \frac{1}{n^2}\sum_{i\neq j}(1-(1-u_i)^n)(1-(1-u_j)^n)$ we obtain via (6.2) for $p_n(u)$ the explicit formula

$$\begin{aligned}
p_n(u) &= \frac{1}{n}\sum_{i\in\mathbb{N}}(1-(1-u_i)^n) - \frac{1}{n^2}\sum_{i\in\mathbb{N}}(1-(1-u_i)^n)(1-(1-u_i)^n) \\
&\quad -\frac{1}{n^2}\sum_{\substack{i,j\in\mathbb{N}\\ i\neq j}}(1-(1-u_i)^n)(1-(1-u_j)^n) \\
&\quad -\frac{|u|(1-|u|)}{n} - \frac{1}{n^2}\sum_{i\in\mathbb{N}}(1-u_i)^n(1-(1-u_i)^n) \\
&\quad +\frac{2(1-|u|)}{n}\sum_{i\in\mathbb{N}}u_i(1-u_i)^{n-1} \\
&\quad +\frac{1}{n^2}\sum_{\substack{i,j\in\mathbb{N}\\ i\neq j}}((1-u_i)^n(1-u_j)^n - (1-u_i-u_j)^n) \\
&= \frac{1}{n}\sum_{i\in\mathbb{N}}(1-(1-u_i)^n) - \frac{|u|(1-|u|)}{n} + \frac{2(1-|u|)}{n}\sum_{i\in\mathbb{N}}u_i(1-u_i)^{n-1} \\
&\quad -\frac{1}{n^2}\sum_{i\in\mathbb{N}}(1-(1-u_i)^n) \\
&\quad +\frac{1}{n^2}\sum_{\substack{i,j\in\mathbb{N}\\ i\neq j}}((1-u_i)^n + (1-u_j)^n - 1 - (1-u_i-u_j)^n).
\end{aligned}$$

In particular $p_1(u) = |u|(1-|u|) \geq 0$. Note that p_n is a symmetric function with respect to the coordinates u_i, $i \in \mathbb{N}$. We have not been able to prove for general $n \in \mathbb{N}$ that $p_n(u) \geq 0$ and leave this conjecture as an open problem for future research.

7. Appendix

Lemma 7.1. *For all $n \in \mathbb{N}$ and $u = (u_i)_{i\in\mathbb{N}} \in \Delta$ the inequality*

$$1-(1-|u|)^n \;\leq\; \sum_{i\in\mathbb{N}}(1-(1-u_i)^n) \tag{7.1}$$

holds. In particular, $\Phi(n) \leq \widetilde{\Phi}(n)$ for all $n \in \mathbb{N}$.

Proof: Induction on n. Clearly (7.1) holds for $n = 1$. The induction step from $n \in \mathbb{N}$ to $n+1$ works as follows. By the induction hypothesis,

$$\begin{aligned}
1-(1-|u|)^{n+1} &= |u| + (1-|u|)(1-(1-|u|)^n) \\
&\leq |u| + (1-|u|)\sum_{i\in\mathbb{N}}(1-(1-u_i)^n) \\
&= |u| + \sum_{i\in\mathbb{N}}(1-|u|)(1-(1-u_i)^n) \;\leq\; |u| + \sum_{i\in\mathbb{N}}(1-u_i)(1-(1-u_i)^n) \\
&= |u| + \sum_{i\in\mathbb{N}}(1-u_i-(1-u_i)^{n+1}) \;=\; \sum_{i\in\mathbb{N}}(1-(1-u_i)^{n+1}),
\end{aligned}$$

where the last equality holds, since $|u| = \sum_{i\in\mathbb{N}} u_i$ by definition. Integration of both sides of (7.1) with respect to the measure ν yields $\Phi(n) \leq \widetilde{\Phi}(n)$ for all $n \in \mathbb{N}$. □

The following two asymptotic results for beta like integrals with an additional logarithmic factor are of general interest. Lemma 7.3 is used in Example 3.2 (NLG-coalescent).

Lemma 7.2. *Let $a, b, c > 0$. Then, as $b \to \infty$,*

$$\int_0^1 u^{a-1}(1-u)^b(-\log u)^{c-1}\,\mathrm{d}u \;\sim\; \mathrm{B}(a,b)(\log b)^{c-1} \;\sim\; \Gamma(a)b^{-a}(\log b)^{c-1}, \tag{7.2}$$

where B *and* Γ *denote the beta function and the gamma function respectively.*

Proof: The integral is monotone (decreasing) in $b \geq 0$. It hence suffices to consider integer $b \in \mathbb{N}$. Binomial expansion $(1-u)^b = \sum_{j=0}^b \binom{b}{j}(-u)^j$ turns the integral into

$$\sum_{j=0}^b \binom{b}{j}(-1)^j \int_0^1 u^{j+a-1}(-\log u)^{c-1}\,\mathrm{d}u \;=\; \Gamma(c)\sum_{j=0}^b \binom{b}{j}\frac{(-1)^j}{(j+a)^c},$$

which is asymptotically equal to $\Gamma(a)b^{-a}(\log b)^{c-1}$ as $b \to \infty$ (see, for example, Rubinstein, 2013, Theorem 1.7). □

Lemma 7.3. *Let $a, b, c > 0$. Then, as $b \to \infty$,*

$$\int_0^1 u^{a-2}\big(1-(1-u)^b\big)(-\log u)^{c-1}\,\mathrm{d}u \;\sim\; \begin{cases} \dfrac{\Gamma(a)}{1-a}b^{1-a}(\log b)^{c-1} & \text{if } 0<a<1, \\ \dfrac{(\log b)^c}{c} & \text{if } a=1, \\ \dfrac{\Gamma(c)}{(a-1)^c} & \text{if } a>1. \end{cases} \tag{7.3}$$

Proof: Let $I(b)$ denote the integral on the left hand side in (7.3). Assume first that $0 < a \leq 1$. Since $I(b)$ is monotone (increasing) in $b \geq 0$ it suffices to consider integer $b \in \mathbb{N}$. Using $1-(1-u)^b = u\sum_{k=0}^{b-1}(1-u)^k$ it follows that

$$I(b) \;=\; \sum_{k=0}^{b-1}\int_0^1 u^{a-1}(1-u)^k(-\log u)^{c-1}\,\mathrm{d}u.$$

The integral below the sum is asymptotically equal to $\Gamma(a)(\log k)^{c-1}/k^a$ as $k \to \infty$ by Lemma 7.2. Taking $a \leq 1$ and $c > 0$ into account we conclude that, as $b \to \infty$,

$$I(b) \;\sim\; \Gamma(a)\sum_{k=1}^{b-1}\frac{(\log k)^{c-1}}{k^a} \;\sim\; \begin{cases} \Gamma(a)(\log b)^{c-1}\frac{b^{1-a}}{1-a} & \text{if } 0<a<1, \\ \frac{(\log b)^c}{c} & \text{if } a=1, \end{cases}$$

since $\int_2^b(\log x)^{c-1}/x^a\,\mathrm{d}x \sim (\log b)^{c-1}b^{1-a}/(1-a)$ for $0<a<1$ and, for $a=1$, $\int_2^b(\log x)^{c-1}/x\,\mathrm{d}x \sim (\log b)^c/c$ as $b \to \infty$. Alternatively, the case $a=1$ may be handled as follows. By partial integration, $I(b) = (b/c)\int_0^1(1-u)^{b-1}(-\log u)^c\,\mathrm{d}u \sim (\log b)^c/c$ as $b \to \infty$ by Lemma 7.2, applied with $a := 1$ and c replaced by $c+1$.

If $a > 1$ then dominated convergence yields $I(b) \to \int_0^1 u^{a-2}(-\log u)^{c-1}\,\mathrm{d}u = \Gamma(c)/(a-1)^c$ as $q \to \infty$. □

References

Ben-Hamou, A., Boucheron, S., and Ohannessian, M. I. Concentration inequalities in the infinite urn scheme for occupancy counts and the missing mass, with applications. *Bernoulli*, **23** (1), 249–287 (2017). MR3556773.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (1987). ISBN 0-521-30787-2. MR898871.

Cox, D. R. and Lewis, P. A. W. *The statistical analysis of series of events*. Methuen & Co., Ltd., London; John Wiley & Sons, Inc., New York (1966). MR0199942.

Durrett, R. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge (2019). ISBN 978-1-108-47368-2. MR3930614.

Dutko, M. Central limit theorems for infinite urn models. *Ann. Probab.*, **17** (3), 1255–1263 (1989). MR1009456.

Ethier, S. N. and Kurtz, T. G. *Markov processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York (1986). ISBN 0-471-08186-8. MR838085.

Gaiser, F. and Möhle, M. On the block counting process and the fixation line of exchangeable coalescents. *ALEA Lat. Am. J. Probab. Math. Stat.*, **13** (2), 809–833 (2016). MR3546382.

Gnedin, A., Hansen, B., and Pitman, J. Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. *Probab. Surv.*, **4**, 146–171 (2007). MR2318403.

Gnedin, A., Iksanov, A., and Marynych, A. On Λ-coalescents with dust component. *J. Appl. Probab.*, **48** (4), 1133–1151 (2011). MR2896672.

González Casanova, A., Miró Pina, V., and Siri-Jégousse, A. The symmetric coalescent and Wright–Fisher models with bottlenecks (2021+). To appear in *Ann. Appl. Probab.*

Haas, B. and Miermont, G. Self-similar scaling limits of non-increasing Markov chains. *Bernoulli*, **17** (4), 1217–1247 (2011). MR2854770.

Handa, K. The two-parameter Poisson-Dirichlet point process. *Bernoulli*, **15** (4), 1082–1116 (2009). MR2597584.

Herriger, P. and Möhle, M. Conditions for exchangeable coalescents to come down from infinity. *ALEA Lat. Am. J. Probab. Math. Stat.*, **9** (2), 637–665 (2012). MR3069379.

Huillet, T. and Martinez, S. Dirichlet-Kingman partition revisited. *Far East J. Theor. Stat.*, **24** (1), 1–33 (2008). MR2426289.

Huillet, T. and Möhle, M. Population genetics models with skewed fertilities: a forward and backward analysis. *Stoch. Models*, **27** (3), 521–554 (2011). MR2827443.

Ismail, M. E. H., Lorch, L., and Muldoon, M. E. Completely monotonic functions associated with the gamma function and its q-analogues. *J. Math. Anal. Appl.*, **116** (1), 1–9 (1986). MR837337.

Karlin, S. Central limit theorems for certain infinite urn schemes. *J. Math. Mech.*, **17**, 373–401 (1967). MR0216548.

Kingman, J. F. C. Random partitions in population genetics. *Proc. Roy. Soc. London Ser. A*, **361** (1704), 1–20 (1978a). MR526801.

Kingman, J. F. C. The representation of partition structures. *J. London Math. Soc. (2)*, **18** (2), 374–380 (1978b). MR509954.

Kolchin, V. F., Sevast'yanov, B. A., and Chistyakov, V. P. *Random allocations*. V. H. Winston & Sons, Washington, D.C.; distributed by Halsted Press [John Wiley & Sons], New York-Toronto, Ont.-London (1978). ISBN 0-470-99394-4. MR0471016.

Lasserre, J. B. *Moments, positive polynomials and their applications*, volume 1 of *Imperial College Press Optimization Series*. Imperial College Press, London (2010). ISBN 978-1-84816-445-1; 1-84816-445-9. MR2589247.

Möhle, M. 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.

Möhle, M. Hitting probabilities for the Greenwood model and relations to near constancy oscillation. *Bernoulli*, **24** (1), 316–332 (2018). MR3706759.

Möhle, M. and Sagitov, S. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, **29** (4), 1547–1562 (2001). MR1880231.

Pitman, J. Coalescents with multiple collisions. *Ann. Probab.*, **27** (4), 1870–1902 (1999). MR1742892.

Rubinstein, M. O. Identities for the Hurwitz zeta function, Gamma function, and L-functions. *Ramanujan J.*, **32** (3), 421–464 (2013). MR3130657.

Sagitov, S. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, **36** (4), 1116–1125 (1999). MR1742154.

Sagitov, S. Convergence to the coalescent with simultaneous multiple mergers. *J. Appl. Probab.*, **40** (4), 839–854 (2003). MR2012671.

Schilling, R. L., Song, R., and Vondraček, Z. *Bernstein functions. Theory and applications*, volume 37 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition (2012). ISBN 978-3-11-025229-3; 978-3-11-026933-8. MR2978140.

Schweinsberg, J. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, **5**, Paper no. 12, 50 (2000a). MR1781024.

Schweinsberg, J. A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.*, **5**, 1–11 (2000b). MR1736720.