ALEA, Lat. Am. J. Probab. Math. Stat. 19, 641–664 (2022)
DOI: 10.30757/ALEA.v19-25

Scaling limits for the block counting process and the fixation line for a class of Λ-coalescents

Martin Möhle and Benedict Vetter

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

Abstract. We provide scaling limits for the block counting process and the fixation line of Λ-coalescents as the initial state n tends to infinity under the assumption that the measure Λ on $[0,1]$ satisfies $\int_{[0,1]} u^{-1}|\Lambda - b\lambda|(\mathrm{d}u) < \infty$ for some $b \geq 0$. Here λ denotes the Lebesgue measure on $[0,1]$. The main result states that the block counting process, properly transformed, converges in the Skorohod space to a generalized Ornstein–Uhlenbeck process as n tends to infinity. The result is applied to beta coalescents with parameters 1 and $b > 0$. We split the generators into two parts by additively decomposing Λ into a 'Bolthausen–Sznitman part' $b\lambda$ and a 'dust part' $\Lambda - b\lambda$ and then prove the uniform convergence of both parts separately.

1. Introduction

The Λ-coalescent, independently introduced by Pitman (1999) and Sagitov (1999), is a Markov process $\Pi = (\Pi_t)_{t\geq 0}$ with càdlàg paths, values in the space of partitions of $\mathbb{N} := \{1, 2, \ldots\}$, starting at time $t = 0$ from the partition $\{\{1\}, \{2\}, \ldots\}$ of $\mathbb{N}$ into singletons, whose behavior is fully determined by a finite measure Λ on the Borel subsets of $[0, 1]$. If the process is in a state with $k \geq 2$ blocks, any particular $j \in \{2, \ldots, k\}$ blocks merge at the rate

$$\lambda_{k,j} = \int_{[0,1]} u^{j-2}(1-u)^{k-j}\Lambda(\mathrm{d}u).$$

The reader is referred to Berestycki (2009) for a survey of Λ-coalescents. Unless $\Lambda(\{1\}) > 0$, Π_t has either infinitely many blocks for all $t > 0$ almost surely or finitely many blocks for all $t > 0$ almost surely. The Λ-coalescent is said to stay infinite in the first case and to come down from infinity in the second. An atom of Λ at 1 corresponds to the rate of jumping to the trivial and absorbing partition consisting only of the block $\mathbb{N}$. For $t \geq 0$ let $N_t^{(n)}$ denote the number of blocks of the restriction $\Pi_t^{(n)} := \{B \cap [n] | B \in \Pi_t, B \cap [n] \neq \emptyset\}$ of Π_t to $[n] := \{1, \ldots, n\}$. The *block counting*

Received by the editors August 26th, 2021; accepted January 14th, 2022.
2020 *Mathematics Subject Classification.* 60J90, 60J27.
Key words and phrases. Block counting process; Bolthausen–Sznitman coalescent; coalescent; dust; fixation line; generalized Ornstein–Uhlenbeck process; time-inhomogeneous process; weak convergence.

process $N^{(n)} := (N_t^{(n)})_{t\geq 0}$ is a $[n]$-valued Markov process that jumps from state $k \geq 2$ to state $j \in \{1,\ldots,k-1\}$ at the rate

$$q_{k,j} \;=\; \binom{k}{j-1}\int_{[0,1]} u^{k-j-1}(1-u)^{j-1}\Lambda(\mathrm{d}u).$$

Clearly, $N^{(n)}$ starts in n at time $t=0$, has decreasing paths and eventually reaches the absorbing state 1. This work's main objective is to analyze the limiting behavior of the block counting process of Λ-coalescents that stay infinite as the initial state n tends to infinity by determining suitable scaling constants. The question of the existence of scaling constants for which non-trivial limits can be obtained is answered in the literature for coalescents with dust, i.e., (see Pitman, 1999; Schweinsberg, 2000a) for measures Λ that satisfy

$$\int_{[0,1]} u^{-1}\Lambda(\mathrm{d}u) \;<\; \infty, \qquad \Lambda(\{0\}) \;=\; \Lambda(\{1\}) \;=\; 0, \tag{1.1}$$

and for the Bolthausen–Sznitman coalescent (Bolthausen and Sznitman, 1998), where $\Lambda=\lambda$ is the uniform distribution on $[0,1]$, an example of a dust-free coalescent that stays infinite. The respective convergence results are recalled in Section 2, where they are stated as Propositions 2.1 and 2.2. This work provides unified proofs of Propositions 2.1 and 2.2 and extends the convergence results by combining both proofs. The main result (Theorem 2.3) covers Λ-coalescents for which there exists some $b\geq 0$ such that

$$\int_{[0,1]} u^{-1}|\Lambda-b\lambda|(\mathrm{d}u) \;<\; \infty,$$

which can be understood that Λ is the sum of a 'Bolthausen–Sznitman part' $b\lambda$ and a 'dust part' $\Lambda-b\lambda$. Here $|\Lambda-b\lambda|$ denotes the total variation of the signed measure $\Lambda-b\lambda$. The assumption includes Λ-coalescents where $\Lambda=\beta(1,b)$ is the beta distribution with parameters 1 and $b>0$. The main result states that

$$(\log N_t^{(n)} - e^{-bt}\log n)_{t\geq 0}$$

converges in the Skorohod space $D_{\mathbb{R}}[0,\infty)$ as n tends to infinity. The limiting process is influenced by both the 'Bolthausen–Sznitman part' and the 'dust part'. The logarithmic version of the convergence result has the advantage of putting the limiting process in Theorem 2.3 to the class of generalized Ornstein–Uhlenbeck processes, which have been extensively studied in the literature. In Limic and Talarczyk (2015), a work concerning the small-time behavior of the block counting process for a broad class of Λ-coalescents that come down from infinity, a generalized Ornstein–Uhlenbeck process also appears in a limit theorem for the block counting process. Regarding generalized Ornstein–Uhlenbeck processes, the interested reader is referred to Sato and Yamazato (1984).

The fixation line $L=(L_t)_{t\geq 0}$ is a $\mathbb{N}$-valued Markov process that jumps from state $k\in\mathbb{N}$ to state $j\in\{k+1,k+2,\ldots\}$ at the rate

$$\gamma_{k,j} \;=\; \binom{j}{j-k+1}\int_{[0,1]} u^{j-k-1}(1-u)^{k}\Lambda(\mathrm{d}u).$$

The fixation line is the 'time-reversal' of the block counting process, in the sense that the hitting times $\inf\{t\geq 0 | N_t^{(n)}\leq m\}$ and $\inf\{t\geq 0|L_t^{(m)}\geq n\}$ share the same distribution, see Hénard (2015, Lemma 2.1). Here the upper index '(m)' denotes the initial state $L_0^{(m)}=m$ at time $t=0$. Equivalently, the process L is Siegmund-dual (Siegmund, 1976) to the block counting process, i.e., (see Kukla and Möhle, 2018)

$$\mathbb{P}(L_t^{(m)}\geq n) \;=\; \mathbb{P}(N_t^{(n)}\leq m), \qquad m,n\in\mathbb{N}, t\geq 0. \tag{1.2}$$

For a thorough definition of the fixation line see Hénard (2015) and the references therein. Theorem 2.4 states that $(\log L_t^{(n)} - e^{bt}\log n)_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as the initial value n tends to infinity.

The article is organized as follows. In Section 2 the two known convergence results for the block counting process of coalescents with dust (Proposition 2.1) and the Bolthausen–Sznitman coalescent (Proposition 2.2) are recalled and the main result (Theorem 2.3) is stated. In Section 3 well-known results concerning generalized Ornstein–Uhlenbeck processes are applied to our setting. In particular, the generator of the limiting process is determined. In Section 4 the main result is applied to beta coalescents with parameter 1 and $b > 0$. The line of proof is as follows. First, we prove Propositions 2.1 and 2.2 in Sections 5 and 6 by showing the uniform convergence of the generators of the logarithm of the scaled block counting processes. The decomposition of Λ into the uniform distribution multiplied by a constant and a measure that corresponds to a coalescent with dust is transferred to the generators. This enables us to use relations obtained in Sections 5 and 6 to prove Theorem 2.3 in Section 7. Two proofs of Theorem 2.4 are given in Section 8.

Notation. Let E be a complete separable metric space. The Banach space $B(E)$ of bounded measurable functions $f: E\to\mathbb{R}$ is equipped with the usual supremum norm $\|f\| := \sup_{x\in E}|f(x)|$ and the Banach subspace $\widehat{C}(E)\subset B(E)$ consists of all continuous functions vanishing at infinity. If $E\subseteq\mathbb{R}^d$ for some $d\in\mathbb{N}$, then $C_k(E)$ denotes the space of k-times continuously differentiable functions. A Feller semigroup $(T_t)_{t\geq 0}$ is strongly continuous on $\widehat{C}(E)$, i.e., $\lim_{t\to 0}||T_t f - f|| = 0$ for each $f\in\widehat{C}(E)$, and satisfies $T_t(\widehat{C}(E))\subseteq\widehat{C}(E)$ for each $t\geq 0$. The generators corresponding to Feller semigroups, usually denoted by A, are understood to be defined on a dense subspace of $\widehat{C}(E)$. The Borel-σ-field on $\mathbb{R}$ is denoted by $\mathcal{B}$ and λ denotes Lebesgue measure on $([0,1],\mathcal{B}\cap[0,1])$. For a measure space $(\Omega,\mathcal{F},\mu)$ and $p>0$ the space of measurable functions $f:\Omega\to\mathbb{R}$ with $\int_\Omega |f|^p \mathrm{d}\mu < \infty$ is denoted by $L^p(\mu)$ or, in short, L^p.

2. Results

Throughout the article Λ is a finite non-zero measure on $([0,1],\mathcal{B}\cap[0,1])$. Additionally, it is assumed that $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$, because coalescents in this article shall stay infinite and an atom at 0 would imply that the coalescent comes down from infinity and an atom at 1 would imply that the block counting process $N^{(n)}$ is almost surely in state 1 for all $n\in\mathbb{N}$ after a random finite time not depending on n.

First, the two known results mentioned in the introduction are presented. A block $B\in\Pi_t$ of size $|B| = 1$ is called a singleton. The number of singletons in $[n]$ divided by n converges to the frequency of singletons as n tends to infinity, and if the frequency of singletons is strictly positive, the coalescent is said to have dust. A necessary and sufficient conditon for coalescents to have dust is given by Eq. (1.1). For further results on Λ-coalescents with dust see Gnedin et al. (2011) and Gaiser and Möhle (2016). Proposition 2.1 below has been established in Gaiser and Möhle (2016) and Möhle (2021). In both articles the processes have non-logarithmic form and the blocks of the coalescent are allowed to even merge simultaneously. The limiting process is the logarithm of the frequency of singletons process as described in Pitman (1999, Proposition 26). In Möhle (2021) the uniform convergence of the generators has been proven and a rate of convergence has been determined. In this article the uniform convergence of the generators is going to be proven as well, but with different techniques. In Gaiser and Möhle (2016) the convergence of the corresponding semigroups has been shown, which is equivalent to the convergence of the generators on a core. The proof is carried out, since parts are needed in order to verify Theorem 2.3.

Proposition 2.1 (dust case)**.** *Suppose that* $\int_{[0,1]} u^{-1}\Lambda(\mathrm{d}u) < \infty$. *Then the time-homogeneous Markov process* $X^{(n)} := (X_t^{(n)})_{t\geq 0} := (\log N_t^{(n)} - \log n)_{t\geq 0}$ *converges in* $D_{\mathbb{R}}[0,\infty)$ *as* $n\to\infty$ *to a* *limiting process $X = (X_t)_{t\geq 0}$ with initial value $X_0 = 0$ and semigroup $(T_t)_{t\geq 0}$ given by*

$$T_t f(x) \;:=\; \mathbb{E}(f(X_{s+t})|X_s = x) \;=\; \mathbb{E}(f(x + X_t)), \qquad x \in \mathbb{R}, f \in B(\mathbb{R}), s,t \geq 0, \tag{2.1}$$

where X_t has characteristic function $\mathbb{E}(\exp(ivX_t)) = \exp(t\psi(v))$, $v \in \mathbb{R}, t \geq 0$, with

$$\psi(v) \;=\; \int_{[0,1]} ((1-u)^{iv} - 1)u^{-2}\Lambda(\mathrm{d}u), \qquad v \in \mathbb{R}. \tag{2.2}$$

Observe that $-X$ is a pure-jump subordinator with characteristic exponent $v \mapsto \psi(-v)$, $v \in \mathbb{R}$.

The block counting process of the Bolthausen–Sznitman coalescent, has been treated in Möhle (2015) and Kukla and Möhle (2018). Both works show that the semigroup of $(N_t^{(n)}/n^{e^{-t}})_{t\geq 0}$ converges on a dense subset of $B([0,\infty))$ to the semigroup of the Mittag–Leffler process as n tends to infinity, hence the processes converge in $D_{[0,\infty)}[0,\infty)$. Taking logarithms does not spoil the convergence. If $f \in \widehat{C}(\mathbb{R})$, then $f \circ \log \in \widehat{C}([0,\infty))$, and the semigroup and hence the generator $A^{(n)}$ of the logarithm of the scaled block counting process $X^{(n)} = (\log N_t^{(n)} - e^{-t}\log n)_{t\geq 0}$ converge as well. We prove the convergence of $A^{(n)}$ in Section 6 directly. Since the scaling depends on t, the process $X^{(n)}$ is time-inhomogeneous, and Kukla and Möhle (2018) have used the time-space process in order to transfer the question of convergence to time-homogeneous Markov processes. The time-space process is revisited in Section 6. By constructing the Bolthausen–Sznitman coalescent from a random recursive tree, it is shown in Goldschmidt and Martin (2005) and Baur and Bertoin (2015) that $N_t^{(n)}/n^{e^{-t}}$ converges almost surely as n tends to infinity for each $t \geq 0$. Since λ is the particular beta distribution with both parameters equal to 1, the following result is the case $b = 1$ of Example 4.2 provided in Section 4.

Proposition 2.2 (Bolthausen–Sznitman case)**.** *Suppose that $\Lambda = \lambda$. Then the time-inhomgeneous Markov process $X^{(n)} := (X_t^{(n)})_{t\geq 0} := (\log N_t^{(n)} - e^{-t}\log n)_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to the time-homogeneous Markov process $X = (X_t)_{t\geq 0}$ with initial value $X_0 = 0$ and semigroup $(T_t)_{t\geq 0}$ given by*

$$T_t f(x) \;:=\; \mathbb{E}(f(X_{s+t})|X_s = x) \;=\; \mathbb{E}(f(e^{-t}x + X_t)), \qquad x \in \mathbb{R}, f \in B(\mathbb{R}), s,t \geq 0,$$

where X_t has characteristic function $\phi_t(v) := \mathbb{E}(\exp(ivX_t)) = \Gamma(1 + iv)/\Gamma(1 + ie^{-t}v)$, $v \in \mathbb{R}, t \geq 0$.

For $b \geq 0$ define the possibly signed measure Λ_D on $\mathcal{B} \cap [0,1]$ via $\Lambda_D(B) := \Lambda(B) - b\lambda(B)$, $B \in \mathcal{B} \cap [0,1]$. Hahn's decomposition theorem states the existence of some set $A \in \mathcal{B} \cap [0,1]$ such that $\Lambda_D^+(B) := \Lambda_D(B \cap A)$, $B \in \mathcal{B} \cap [0,1]$, and $\Lambda_D^-(B) := -\Lambda_D(B \cap A^c)$, $B \in \mathcal{B} \cap [0,1]$, define nonnegative measures. The two nonnegative measures Λ_D^+ and Λ_D^- constitute the Jordan decomposition of Λ_D. By using this decomposition, one can integrate with respect to the signed measure Λ_D by defining $\int f\mathrm{d}\Lambda_D := \int f\mathrm{d}\Lambda_D^+ - \int f\mathrm{d}\Lambda_D^-$ for $f \in L^1(\Lambda_D^+) \cap L^1(\Lambda_D^-)$. The total variation $|\Lambda_D|$ of Λ_D is given by $|\Lambda_D| := \Lambda_D^+ + \Lambda_D^-$. The assumption of Theorem 2.3 below is the following.

Assumption A. There exists $b \geq 0$ such that $\int_{[0,1]} u^{-1}|\Lambda_D|(\mathrm{d}u) < \infty$, i.e., $\int_{[0,1]} u^{-1}\Lambda_D^+(\mathrm{d}u) < \infty$ and $\int_{[0,1]} u^{-1}\Lambda_D^-(\mathrm{d}u) < \infty$.

Assumption A implies that $b = \lim_{\varepsilon\to 0} \varepsilon^{-1}\Lambda((0,\varepsilon))$, see Lemma 9.1 a) in the appendix. In particular, if Assumption A holds, then the constant b is uniquely determined by the measure Λ. Schweinsberg's criterion (Schweinsberg, 2000b) shows that the Λ-coalescent does not come down from infinity under Assumption A, see Lemma 9.1 b). Moreover, the Λ-coalescent is dust-free if and only if $b > 0$. Assumption A is for example satisfied, if Λ has density $f \in C_1([0,1])$ with respect to λ for which $\lim_{u\searrow 0} f'(u)$ exists and is finite. In this case, $b = f(0)$.

Suppose that Λ satisfies Assumption A. Let $\Gamma(z) := \int_0^\infty u^{z-1}e^{-u}\mathrm{d}u$, $\mathrm{Re}(z) > 0$, denote the gamma function and $\Psi(z) := (\log\Gamma)'(z) = \Gamma'(z)/\Gamma(z)$, $\mathrm{Re}(z) > 0$, the digamma function. Define

$$a \;:=\; b(1+\Psi(1)) - \int_{[0,1]} u^{-1}\Lambda_D(\mathrm{d}u) \tag{2.3}$$

and the infinitely divisible characteristic exponent $\psi : \mathbb{R} \to \mathbb{C}$ via

$$\psi(v) \;:=\; iav + \int_{[0,1]} ((1-u)^{iv} - 1 + ivu)u^{-2}\Lambda(\mathrm{d}u), \qquad v \in \mathbb{R}. \tag{2.4}$$

Formally, the constant b of Assumption A only appears in the drift part of ψ. Note however that b is uniquely determined by Λ. In this sense b depends on Λ and therefore also influences (via Λ) the jump part of ψ. Substituting $g : (0,1) \to \mathbb{R}$, $g(u) := \log(1-u)$, $u \in (0,1)$, shows that

$$\psi(v) \;=\; iav + \int_{(-\infty,0)} (e^{ivu} - 1 + iv(1-e^u))\varrho(\mathrm{d}u), \qquad v \in \mathbb{R},$$

where the measure ϱ, defined via

$$\varrho(A) \;:=\; \int_{g^{-1}(A)} u^{-2}\Lambda(\mathrm{d}u), \qquad A \in \mathcal{B}, \tag{2.5}$$

satisfies $\int_{\mathbb{R}}(u^2 \wedge 1)\varrho(\mathrm{d}u) < \infty$ and $\varrho(\{0\}) = 0$. Hence, ϱ is a Lévy measure, $e^{\psi(v)}$, $v \in \mathbb{R}$, is the characteristic function of an infinitely divisible distribution and the process described by Eqs. (2.6) and (2.7) below belongs to the class of generalized Ornstein–Uhlenbeck processes. Due to $\varrho((0,\infty)) = 0$, the limiting process in Theorem 2.3 has only negative jumps. Compensation of small jumps occurs if and only if $b \neq 0$. Further properties of the limiting process are presented in Section 3.

Theorem 2.3. *Suppose that Λ satisfies Assumption A. Then the possibly time-inhomogeneous Markov process $X^{(n)} := (X_t^{(n)})_{t\geq 0} := (\log N_t^{(n)} - e^{-bt}\log n)_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to the time-homogeneous Markov process $X = (X_t)_{t\geq 0}$ with initial value $X_0 = 0$ and semigroup $(T_t)_{t\geq 0}$ given by*

$$T_tf(x) \;:=\; \mathbb{E}(f(X_{s+t})|X_s = x) \;=\; \mathbb{E}(f(e^{-bt}x + X_t)), \qquad x \in \mathbb{R}, f \in B(\mathbb{R}), s,t \geq 0, \tag{2.6}$$

where X_t has characteristic function ϕ_t given by

$$\phi_t(v) \;=\; \exp\biggl(\int_0^t \psi(e^{-bs}v)\mathrm{d}s\biggr), \qquad v \in \mathbb{R}, t \geq 0, \tag{2.7}$$

and ψ is given by (2.4).

The dust case and the Bolthausen–Sznitman case arise from Assumption A as follows. If $\int_{[0,1]} u^{-1}\Lambda(\mathrm{d}u) < \infty$, then Assumption A holds with $b = 0$. Thus, $a = -\int_{[0,1]} u^{-1}\Lambda(\mathrm{d}u)$, the definitions (2.2) and (2.4) for ψ coincide and Proposition 2.1 and Theorem 2.3 describe the same limiting result. For $\Lambda = \lambda$, Assumption A holds with $b = 1$ and without a dust part. In this case, $a = 1 + \Psi(1)$ and the underlying Lévy measure ϱ has density f with respect to Lebesgue measure on $\mathbb{R}\setminus\{0\}$ given by $f(u) := e^u(1-e^u)^{-2}$ for $u < 0$ and $f(u) := 0$ for $u > 0$. The connection between Proposition 2.2 and Theorem 2.3 in the Bolthausen–Sznitman case is clarified in Section 4.

A convergence result for the fixation line can be stated analogously to Theorem 2.3; see also Gaiser and Möhle (2016, Theorem 2.13 b)) for the case $b = 0$.

Theorem 2.4. *Suppose that Λ satisfies Assumption A. Then the possibly time-inhomogeneous Markov process $Y^{(n)} := (Y_t^{(n)})_{t\geq 0} := (\log L_t^{(n)} - e^{bt}\log n)_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$*

to the time-homogeneous Markov process $Y = (Y_t)_{t\geq 0}$ with initial value $Y_0 = 0$ and semigroup $(T_t)_{t\geq 0}$ given by

$$T_t f(y) \ := \ \mathbb{E}(f(Y_{s+t})|Y_s = y) \ = \ \mathbb{E}(f(e^{bt}y + Y_t)), \qquad y \in \mathbb{R}, f \in B(\mathbb{R}), s, t \geq 0, \tag{2.8}$$

where Y_t has characteristic function χ_t given by

$$\chi_t(w) \ = \ \exp\bigg(\int_0^t \psi(-e^{bs}w)\mathrm{d}s\bigg), \qquad w \in \mathbb{R}, t \geq 0, \tag{2.9}$$

and ψ is given by (2.4).

Remark 2.5. The process defined by (2.8) and (2.9) is a generalized Ornstein–Uhlenbeck process with underlying characteristic exponent $v \mapsto \psi(-v)$, $v \in \mathbb{R}$, but with nonnegative drift.

Remark 2.6. Let the random variable S_t have characteristic function ϕ_t, given by (2.7), for $t \geq 0$, and let $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ denote the processes defined in Theorems 2.3 and 2.4, respectively. Conditional on $X_s = x$, X_{t+s} is distributed as $e^{-bt}x + S_t$ for all $x \in \mathbb{R}$. Note that $Y_t \stackrel{\mathrm{d}}{=} -e^{bt}X_t \stackrel{\mathrm{d}}{=} -e^{bt}S_t$ and that conditional on $Y_s = y$, Y_{t+s} is distributed as $e^{bt}y - e^{bt}S_t$. Hence,

$$\mathbb{P}(e^{Y_{t+s}} \geq x | e^{Y_s} = y) = \mathbb{P}(y^{e^{bt}} e^{-e^{bt}S_t} \geq x) = \mathbb{P}(x^{e^{-bt}} e^{S_t} \leq y) = \mathbb{P}(e^{X_{t+s}} \leq y | e^{X_s} = x)$$

for all $x, y, s, t \geq 0$, i.e., e^Y is Siegmund-dual to e^X (see Siegmund, 1976) parallel to the Siegmund-duality of the block counting process and the fixation line.

Remark 2.7. For the Bolthausen–Sznitman case, the convergence result corresponding to Theorem 2.4 is stated in Kukla and Möhle (2018, Theorem 3.1 b)) in non-logarithmic form. The fixation line of the Bolthausen–Sznitman coalescent is a continuous-time discrete state space branching process in which the offspring distribution has probability generating function $f(s) = s + (1 - s)\log(1 - s)$, $s \in [0, 1]$. The limiting process described in Theorem 2.4 is the logarithm of Neveu's continuous-state branching process. By Proposition 2.2, the characteristic functions χ_t of the marginal distributions are given by (see Kukla and Möhle, 2018, Eq. (19))

$$\chi_t(w) \ = \ \phi_t(-e^t w) \ = \ \frac{\Gamma(1 - ie^{bt}w)}{\Gamma(1 - iw)}, \qquad w \in \mathbb{R}, t \geq 0.$$

3. The limiting process

Standard computations (see Sato, 1999, Lemma 17.1) show that ϕ_t, given by (2.7), is the characteristic function of an infinitely divisible distribution for each $t \geq 0$ without Gaussian component and Lévy measure ϱ_t given by

$$\varrho_t(A) = \int_{(-\infty,0)} \int_0^t 1_A(e^{-bs}u)\mathrm{d}s\varrho(\mathrm{d}u), \qquad A \in \mathcal{B}, t \geq 0.$$

Sato and Yamazato (1984, Theorem 3.1) provide a formula for the generator corresponding to the semigroup $(T_t)_{t\geq 0}$ given by (2.6).

Lemma 3.1. *Suppose that Λ satisfies Assumption A. Let ψ be given by (2.4), ϕ_t be defined by (2.7) and let the random variable X_t have characteristic function ϕ_t for each $t \geq 0$. The family of operators $(T_t)_{t\geq 0}$ defined by (2.6) is a Feller semigroup. Let D denote the space of twice differentiable functions $f : \mathbb{R} \to \mathbb{R}$ such that $f, f', f'' \in \widehat{C}(\mathbb{R})$ and such that the map $x \mapsto xf'(x)$, $x \in \mathbb{R}$, belongs to $\widehat{C}(\mathbb{R})$. Then D is a core for the generator A corresponding to $(T_t)_{t\geq 0}$ and*

$$Af(x) \ = \ f'(x)(a - bx) + \int_{[0,1]} (f(x + \log(1 - u)) - f(x) + uf'(x))u^{-2}\Lambda(\mathrm{d}u) \tag{3.1}$$

for $x \in \mathbb{R}$ and $f \in D$, where a is given by (2.3).

Proof: Substituting $g:(0,1)\to\mathbb{R}$, $g(u):=\log(1-u)$, $u\in(0,1)$, shows that (3.1) is an integro-differential operator of the form (1.1) of Sato and Yamazato (1984) with dimension $d=1$. In Sato and Yamazato (1984), operators of this form are initially considered as acting on the space C_c^2 of twice differentiable functions with compact support (see the explanations after Eq. (1.2) in Sato and Yamazato (1984)), but Step 3 of the proof of Sato and Yamazato (1984, Theorem 3.1) shows that (3.1) even holds for functions $f\in D$ ($\supset C_c^2$). Note that the space D is denoted by F_1 in Sato and Yamazato (1984). The fact that D is a core for A is only a different phrasing of the claim in Step 5 of the proof of Sato and Yamazato (1984, Theorem 3.1). □

The limiting process's generator in the dust case ($b=0$) is given by

$$Af(x)\;=\;\int_{[0,1]}(f(x+\log(1-u))-f(x))u^{-2}\Lambda(\mathrm{d}u),\qquad x\in\mathbb{R},$$

in agreement with Eq. (3.1).

The limiting process in Theorem 2.3 arises as the solution to a certain stochastic differential equation. For the remainder of this section, $b>0$ is fixed and ψ is allowed to be the characteristic exponent of an arbitrary infinitely divisible distribution on $\mathbb{R}$, except for Lemmata 3.2 and 3.3, which are applications of results known from the literature to the coalescent setting. Let the Lévy process $L=(L_t)_{t\geq 0}$ with characteristic functions $\mathbb{E}(e^{ivL_t})=e^{t\psi(v)}$, $v\in\mathbb{R}, t\geq 0$, be adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ which satisfies the usual hypotheses. In particular, $L_{t+s}-L_s$ is independent of $\mathcal{F}_s$ for all $s,t\geq 0$. The Langevin equation with Lévy noise instead of a Brownian motion

$$\mathrm{d}X_t\;=\;-bX_t\mathrm{d}t+\mathrm{d}L_t,\qquad t\geq 0,\tag{3.2}$$

with initial value $X_0=x$ has an unique $(\mathcal{F}_t)_{t\geq 0}$-adapted solution $X=(X_t)_{t\geq 0}$ with càdlàg paths. The solution to (3.2) or the corresponding semigroup are hence called generalized Ornstein–Uhlenbeck or Ornstein–Uhlenbeck type process or semigroup. It holds that

$$X_t\;=\;e^{-bt}x+\int_0^t e^{-b(t-s)}\mathrm{d}L_s,\qquad t\geq 0.\tag{3.3}$$

Various constructions for the stochastic integral in (3.3) are possible, e.g., in Applebaum (2009, Sections 6.3 and 6.2) the stochastic integral is the Itô-integral with respect to semimartingales. The process X is a stochastically continuous Markov process and the corresponding semigroup is given by (2.6), where the characteristic functions ϕ_t of X_t are given by (2.7) with underlying infinitely divisible characteristic exponent ψ for $t\geq 0$.

Generalized Ornstein–Uhlenbeck processes bear a close connection to self-decomposable distributions. A real-valued random variable S is called self-decomposable if for every $\alpha\in[0,1]$ there exists a random variable S_α independent of S such that S has the same distribution as $\alpha S+S_\alpha$. If ϕ is the characteristic function of S, then S is self-decomposable if and only if $v\mapsto\phi(v)/\phi(\alpha v)$, $v\in\mathbb{R}$, is the characteristic function of a real-valued random variable for every $\alpha\in[0,1]$. A distribution μ on $\mathbb{R}$ or its characteristic function ϕ is said to be self-decomposable if there exists a self-decomposable random variable with distribution μ. Suppose that the Lévy measure ϱ of the characteristic exponent ψ satisfies

$$\int_{\{|u|>1\}}\log(1+|u|)\varrho(\mathrm{d}u)<\infty.\tag{3.4}$$

According to Sato and Yamazato (1984, Theorems 4.1 and 4.2), X_t converges in distribution as $t\to\infty$ to the unique stationary distribution μ of X. The distribution μ is self-decomposable. Conversely, every self-decomposable distribution can be obtained as the stationary distribution of a generalized Ornstein–Uhlenbeck process. If (3.4) does not hold, then there exists no stationary distribution. The following lemma is an application of Sato and Yamazato (1984, Theorems 4.1 and 4.2) to this article's coalescent setting.

Lemma 3.2. *Suppose that Λ satisfies Assumption A with $b > 0$ and let $X = (X_t)_{t\geq 0}$ be as in Theorem 2.3. If $\int_{(\varepsilon,1)} \log\log(1-u)^{-1}\Lambda(\mathrm{d}u) < \infty$ for some $1 - e^{-1} < \varepsilon < 1$, then X_t converges in distribution as $t \to \infty$ to the unique stationary distribution μ of X. The distribution μ is self-decomposable with characteristic function ϕ given by*

$$\phi(v) \;=\; \exp\biggl(\int_0^\infty \psi(e^{-bs}v)\mathrm{d}s\biggr), \qquad v \in \mathbb{R}.$$

The characteristic function ϕ_t of X_t satisfies $\phi_t(v) = \phi(v)/\phi(e^{-bt}v)$, $v \in \mathbb{R}$.

If $\int_{(\varepsilon,1)} \log\log(1-u)^{-1}\Lambda(\mathrm{d}u) = \infty$ for $0 < \varepsilon < 1$, then, for every l,

$$\lim_{t\to\infty} \sup_{x\in\mathbb{R}} \sup_{y\in\mathbb{R}} \mathbb{P}(|e^{-bt}x + X_t - y| \leq l) \;=\; 0.$$

The process has no stationary distribution.

The criterion of Shiga (1990, Theorem 1.1) for transience and recurrence complements Lemma 3.2.

Lemma 3.3. *Suppose that Λ satisfies Assumption A with $b > 0$ and let $X = (X_t)_{t\geq 0}$ be as in Theorem 2.3. Then X is irreducible in $\mathbb{R}$. Let $\varepsilon \in [1 - e^{-1}, 1)$ and define $g_\Lambda(y) := \int_{(\varepsilon,1)}(1 - e^{y\log(1-u)})u^{-2}\Lambda(\mathrm{d}u)$, $y \in [0, 1]$. If the integral*

$$\int_0^1 z^{-1} \exp\biggl(-\int_z^1 \frac{g_\Lambda(y)}{by}\mathrm{d}y\biggr)\mathrm{d}z \tag{3.5}$$

is finite, then X is transient, i.e., it holds that $\mathbb{P}(\lim_{t\to\infty} |X_t| = \infty | X_0 = x) = 1$ for every $x \in \mathbb{R}$. If the integral (3.5) is infinite, then X is recurrent, i.e., there exists $a \in \mathbb{R}$ such that $\mathbb{P}(\liminf_{t\to\infty} |X_t - a| = 0 | X_0 = a) = 1$.

Note that the limiting process X or, more precisely, its semigroup $(T_t)_{t\geq 0}$ belongs to the class of Mehler semigroups (Bogachev et al., 1996), as is true for all generalized Ornstein–Uhlenbeck processes, since $\phi_{t+s}(v) = \phi_t(e^{-bs}v)\phi_s(v)$, $v \in \mathbb{R}$, for $s, t \geq 0$.

4. Beta coalescents

The beta distribution $\beta(a, b)$ with parameters $a, b > 0$ has density $u \mapsto \Gamma(a + b)/(\Gamma(a)\Gamma(b))\, u^{a-1}(1 - u)^{b-1}$, $u \in (0, 1)$, with respect to Lebesgue measure on $(0, 1)$. Beta coalescents, for which $\Lambda = \beta(a, b)$ for some $a, b > 0$, have been extensively studied in the literature due to the easy computability of the jump rates

$$q_{k,j} \;=\; \frac{\Gamma(a + b)\Gamma(k + 1)\Gamma(j - 1 + b)\Gamma(k - j - 1 + a)}{\Gamma(a)\Gamma(b)\Gamma(k - 2 + a + b)\Gamma(j)\Gamma(k - j + 2)}, \qquad j \in \{1, \ldots, k - 1\}, k \geq 2. \tag{4.1}$$

The $\beta(a, b)$-coalescent comes down from infinity if and only if $0 < a < 1$ (Schweinsberg, 2000b, Example 15), and has dust if and only if $a > 1$.

For $a = 1$ the beta coalescent is dust-free and does not come down from infinity. From the observation stated below Assumption A we conclude that Assumption A is satisfied with the same constant b. The 'dust part' $\Lambda - b\lambda$ has possibly negative density $u \mapsto b((1 - u)^{b-1} - 1)$, $u \in (0, 1)$, with respect to Lebesgue measure on $(0, 1)$. The computations of a and ψ in the proof of the following proposition are based on Gauß' representation (see Whittaker and Watson, 1996, p. 247)

$$\Psi(z) \;=\; \int_0^\infty \biggl(\frac{e^{-u}}{u} - \frac{e^{-zu}}{1 - e^{-u}}\biggr)\mathrm{d}u, \qquad \mathrm{Re}(z) > 0,$$

of the digamma function.

Proposition 4.1. *Suppose that $\Lambda = \beta(1,b)$ with $b>0$. Let a, ψ and ϱ be given by (2.3), (2.4) and (2.5), respectively. Then ϱ has density f with respect to Lebesgue measure on $(-\infty,0)$ given by $f(u) := be^{bu}(1-e^u)^{-2}$, $u<0$,*

$$a \;=\; b(1+\Psi(b)) \tag{4.2}$$

and

$$\psi(v) \;=\; b\big((1-b)\Psi(b)-(1-b-iv)\Psi(b+iv)\big), \qquad v\in\mathbb{R}. \tag{4.3}$$

Proof: It can be easily verified that ϱ has density as stated in the proposition. Eq. (4.2) follows from

$$\begin{aligned}\int_{[0,1]} u^{-1}(\Lambda - b\lambda)\mathrm{d}u \;&=\; b\int_0^1 u^{-1}((1-u)^{b-1}-1)\mathrm{d}u\\ &=\; b\int_0^\infty\Big(\frac{e^{-bu}}{1-e^{-u}}-\frac{e^{-u}}{1-e^{-u}}\Big)\mathrm{d}u \;=\; b(\Psi(1)-\Psi(b)).\end{aligned}$$

Next, note that

$$\Psi(b)-\Psi(b+iv) \;=\; \int_0^\infty (e^{-ivu}-1)\frac{e^{-bu}}{1-e^{-u}}\mathrm{d}u, \qquad v\in\mathbb{R}. \tag{4.4}$$

Integration by parts yields

$$\begin{aligned}iv(\Psi(b+iv)-\Psi(b)) \;&=\; \int_0^\infty (iv-ive^{-ivu})\frac{e^{-bu}}{1-e^{-u}}\mathrm{d}u\\ =\; (ivu+e^{-ivu}-1)\frac{e^{-bu}}{1-e^{-u}}\Big|_{u=0}^{u=\infty} &- \int_0^\infty (ivu+e^{-ivu}-1)\Big(\frac{-be^{-bu}}{1-e^{-u}}-\frac{e^{-bu}}{(1-e^{-u})^2}e^{-u}\Big)\mathrm{d}u\\ =\; \int_0^\infty (e^{-ivu}-1+ivu)&\frac{e^{-bu}}{(1-e^{-u})^2}\big(1-(1-b)(1-e^{-u})\big)\mathrm{d}u, \qquad v\in\mathbb{R}.\end{aligned}$$

Hence,

$$\begin{aligned}&(1-b)\Psi(b)-(1-b-iv)\Psi(b+iv)\\ &=\; iv\Psi(b)+(1-b)(\Psi(b)-\Psi(b+iv))+iv(\Psi(b+iv)-\Psi(b))\\ &=\; iv\Psi(b)+(1-b)\int_0^\infty (e^{-ivu}-1)\frac{e^{-bu}}{1-e^{-u}}\mathrm{d}u\\ &\;+\;\int_0^\infty (e^{-ivu}-1+ivu)\frac{e^{-bu}}{(1-e^{-u})^2}\big(1-(1-b)(1-e^{-u})\big)\mathrm{d}u\\ &=\; iv\Psi(b)+\int_0^\infty (e^{-ivu}-1+ivu)\frac{e^{-bu}}{(1-e^{-u})^2}\mathrm{d}u - iv(1-b)\int_0^\infty u\frac{e^{-bu}}{1-e^{-u}}\mathrm{d}u\\ &=\; iv\big(\Psi(b)-(1-b)\Psi'(b)\big)+b^{-1}\int_{\mathbb{R}\setminus\{0\}}(e^{ivu}-1-ivu)\varrho(\mathrm{d}u)\\ &=\; iv\Big(\Psi(b)-(1-b)\Psi'(b)+b^{-1}\int_{\mathbb{R}\setminus\{0\}}(e^u-1-u)\varrho(\mathrm{d}u)\Big)\\ &\;+\; b^{-1}\int_{\mathbb{R}\setminus\{0\}}(e^{ivu}-1+iv(1-e^u))\varrho(\mathrm{d}u).\end{aligned}$$

The calculation

$$\begin{aligned}
&-(1-b)\Psi'(b)+b^{-1}\int_{\mathbb{R}\setminus\{0\}}(e^u-1-u)\varrho(\mathrm{d}u)\\
&=\int_0^\infty\big(-(1-b)u(1-e^{-u})+e^{-u}-1+u\big)\frac{e^{-bu}}{(1-e^{-u})^2}\mathrm{d}u \;=\; -\frac{e^{-bu}}{1-e^{-u}}u\Big|_{u=0}^{u=\infty}=1
\end{aligned}$$

and multiplication with b complete the proof of (4.3). $\square$

Example 4.2. Suppose that $\Lambda=\beta(1,b)$ with $b>0$. Then Assumption A is satisfied with the same constant b. According to Theorem 2.3 the process $(\log N_t^{(n)}-e^{-bt}\log n)_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n\to\infty$ to a Markov process $X=(X_t)_{t\geq 0}$ with initial value $X_0=0$ and semigroup $(T_t)_{t\geq 0}$ given by

$$T_tf(x):=\mathbb{E}(f(X_{s+t})|X_s=x)=\mathbb{E}(f(e^{-bt}x+X_t)),\qquad x\in\mathbb{R},f\in B(\mathbb{R}),s,t\geq 0,$$

where X_t has characteristic function ϕ_t given by (2.7).

Since $\int_{(1-e^{-1},1)}\log\log(1-u)^{-1}\Lambda(\mathrm{d}u)=\int_{1-e^{-1}}^1\log\log(1-u)^{-1}b(1-u)^{b-1}\mathrm{d}u<\infty$, the logarithmic moment condition of Lemma 3.2 is satisfied and X_t converges in distribution as $t\to\infty$ to the unique stationary distribution μ of X. The distribution μ is self-decomposable with characteristic function ϕ given by

$$\phi(v)=\exp\left(\int_0^\infty\psi(e^{-bs}v)\mathrm{d}s\right)=\exp\left((1-b)\int_0^v\frac{\Psi(b)-\Psi(b+iu)}{u}\mathrm{d}u\right)\frac{\Gamma(b+iv)}{\Gamma(b)},\quad v\in\mathbb{R}.\quad(4.5)$$

In the last step equation (4.3) and the fact that $\Psi(z)=(\log\Gamma(z))'$, $\mathrm{Re}(z)>0$, have been used. The characteristic function ϕ_t of X_t is hence given by

$$\phi_t(v)=\frac{\phi(v)}{\phi(e^{-bt}v)}=\exp\left((1-b)\int_{e^{-bt}v}^v\frac{\Psi(b)-\Psi(b+iu)}{u}\mathrm{d}u\right)\frac{\Gamma(b+iv)}{\Gamma(b+ie^{-bt}v)},\quad v\in\mathbb{R},t\geq 0.$$

Similarly to the convergence above, $(N_t^{(n)}/n^{e^{-bt}})_{t\geq 0}$ converges in $D_{[0,\infty)}[0,\infty)$ to $(\exp(X_t))_{t\geq 0}$ as $n\to\infty$.

The following is an attempt to describe μ and the distribution of X_t. If Z has a gamma distribution with parameters b and 1, i.e., Z has density $u\mapsto u^{b-1}e^{-u}(\Gamma(b))^{-1}$, $u>0$, with respect to Lebesgue measure on $(0,\infty)$, then $\log Z$ has the self-decomposable characteristic function $v\mapsto\Gamma(b+iv)/\Gamma(b)$, $v\in\mathbb{R}$, see Steutel and van Harn (2004, V, Example 9.18), which implies that the map $v\mapsto\Gamma(b+iv)/\Gamma(b+ie^{-bt}v)$, $v\in\mathbb{R}$, is the characteristic function of a real-valued random variable for every $t\geq 0$. As long as $b<1$, the function $u\mapsto(1-b)(\Psi(b)-\Psi(b+iu))$, $u\in\mathbb{R}$, which appears in the first factor on the right-hand side of (4.5), is the characteristic exponent of the negative of a drift-free subordinator, whose Lévy measure has density $u\mapsto(1-b)e^{-bu}(1-e^{-u})^{-1}$, $u>0$, with respect to Lebesgue measure on $(0,\infty)$, cf. (4.4). In particular, it is the characteristic exponent of an infinitely divisible distribution, and if Z has characteristic function $v\mapsto\exp((1-b)(\Psi(b)-\Psi(b+iv))$, $v\in\mathbb{R}$, then $\mathbb{E}(\log(1+|Z|))<\infty$. By Steutel and van Harn (2004, V, Theorem 6.7), the first factor on the right-hand side of (4.5) is a self-decomposable characteristic function as well, and

$$v\mapsto\exp\left((1-b)\int_{e^{-bt}v}^v\frac{\Psi(b)-\Psi(b+iu)}{u}\mathrm{d}u\right),\qquad v\in\mathbb{R},$$

is the characteristic function of a real-valued random variable for each $t\geq 0$. The arguments that allow the decomposition of ϕ_t into the product of two characteristic functions fail for $b>1$.

We shortly return to the Bolthausen–Sznitman coalescent. Recall that the Bolthausen–Sznitman coalescent is the particular beta coalescent with driving measure $\Lambda=\beta(1,1)$. Proposition 4.1 with $b=1$ states that $\psi(v)=iv\Psi(1+iv)$, $v\in\mathbb{R}$. Example 4.2 with $b=1$ entails the convergence of the limiting process's marginal distributions as $t \to \infty$ to a self-decomposable distribution with characteristic function $\phi(v) = \Gamma(1+iv)$, $v \in \mathbb{R}$. Let Z have an exponential distribution with parameter 1. Then $\log Z$ is the negative of a Gumbel distributed random variable and has characteristic function ϕ, see e.g. Steutel and van Harn (2004, V, Example 9.15). Hence, $-X_t$ converges in distribution as $t \to \infty$ to the Gumbel distribution. Moreover,

$$\phi_t(v) = \exp\left(\int_0^\infty \psi(e^{-s}v)\mathrm{d}s\right) = \frac{\Gamma(1+iv)}{\Gamma(1+ie^{-t}v)}, \qquad v \in \mathbb{R}, t \geq 0,$$

which connects Proposition 2.2 and Theorem 2.3.

5. Proof of Proposition 2.1

In this section Λ satisfies the dust condition $\int_{[0,1]} u^{-1}\Lambda(\mathrm{d}u) < \infty$ in addition to the general assumption $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$. Let $E_n := \{x \in \mathbb{R} | e^x n \in [n]\}$ denote the state space of $X^{(n)} = (X_t^{(n)})_{t\geq 0} = (\log N_t^{(n)} - \log n)_{t \geq 0}$ for each $n \in \mathbb{N}$. By defining $k := k(x,n) := e^x n \in [n]$ for $x \in E_n$ and $n \in \mathbb{N}$, we can represent the generator $A^{(n)}$ of $X^{(n)}$ as

$$A^{(n)}f(x) \;=\; \sum_{j=1}^{k-1}(f(x+\log \tfrac{j}{k}) - f(x))q_{k,j}, \qquad x \in E_n, f \in \widehat{C}(\mathbb{R}), n \in \mathbb{N}.$$

The process $X = (X_t)_{t\geq 0}$ defined by (2.1) and (2.2) is a Feller process in $\widehat{C}(\mathbb{R})$. Let A denote the generator. From Sato (1999, Theorem 31.5) it follows that the space $\widehat{C}_2(\mathbb{R})$ of twice differentiable functions $f \in C_2(\mathbb{R})$ with $f, f', f'' \in \widehat{C}(\mathbb{R})$ is a core for A and

$$Af(x) \;=\; \int_{[0,1]} (f(x+\log(1-u)) - f(x))u^{-2}\Lambda(\mathrm{d}u), \qquad x \in \mathbb{R}, f \in \widehat{C}_2(\mathbb{R}).$$

The idea to prove the uniform convergence of the generators is the following: write the jump rates as values of a distribution depending on k (with some minor adjustments) whose limiting behavior as $k \to \infty$ can be determined. The generators $A^{(n)}$ and A can then be written as the mean of random variables and classical weak convergence results can be applied.

Proof: (of Proposition 2.1) Let $f \in \widehat{C}_2(\mathbb{R})$. Define $h : [0,1]\times\mathbb{R} \to \mathbb{R}$ via $h(u,x) := u^{-1}(f(x+\log(1-u)) - f(x))$, $u \in (0,1)$, $h(0,x) := \lim_{u\searrow 0} h(u,x) = -f'(x)$ and $h(1,x) := \lim_{u\nearrow 1} h(u,x) = -f(x)$ for $x \in \mathbb{R}$. Differentiating $s \mapsto f(x+\log(1-us))$, $s \in (0,1)$, leads to

$$f(x+\log(1-u)) - f(x) \;=\; -u\int_0^1 \frac{f'(x+\log(1-us))}{1-us}\mathrm{d}s, \qquad u \in [0,1), x \in \mathbb{R}.$$

Thus,

$$h(u,x) \;=\; -\int_0^1 \frac{f'(x+\log(1-us))}{1-us}\mathrm{d}s, \qquad u \in [0,1), x \in \mathbb{R},$$

and h stays bounded even as u tends to 0. Define

$$S(k,x) \;:=\; \sum_{j=1}^{k-1}(f(x+\log\tfrac{j}{k}) - f(x))q_{k,j}, \quad I(x) \;:=\; \int_{[0,1]} h(u,x)u^{-1}\Lambda(\mathrm{d}u), \qquad k \in \mathbb{N}, x \in \mathbb{R}. \tag{5.1}$$

Obviously, $A^{(n)}f(x) = S(k,x)$ for $x \in E_n$ and $n \in \mathbb{N}$ and $I(x) = Af(x)$ for $x \in \mathbb{R}$. Substituting $k-j$ for j and the definition of h yield

$$\begin{aligned} S(k,x) &= \sum_{j=1}^{k-1}(f(x+\log(1-\tfrac{j}{k}))-f(x))q_{k,k-j} \\ &= \sum_{j=1}^{k-1} h(\tfrac{j}{k},x)\frac{j}{k}\binom{k}{j+1}\int_{[0,1]} u^{j-1}(1-u)^{k-j-1}\Lambda(\mathrm{d}u) \\ &= \sum_{j=0}^{k-1} h(\tfrac{j}{k},x)\frac{j}{j+1}\binom{k-1}{j}\int_{[0,1]} u^{j-1}(1-u)^{k-j-1}\Lambda(\mathrm{d}u), \qquad k\in\mathbb{N}, x\in\mathbb{R}. \end{aligned}$$

Set $c := \int_{[0,1]} u^{-1}\Lambda(\mathrm{d}u) \in (0,\infty)$ and define the probability measure Q on $([0,1],\mathcal{B}\cap[0,1])$ via $\mathrm{Q}(A) := c^{-1}\int_A u^{-1}\Lambda(\mathrm{d}u)$, $A \in \mathcal{B}\cap[0,1]$. Let the random variables Z_k, $k \in \mathbb{N}$, have distribution given by

$$\mathbb{P}(Z_k = j) = \binom{k-1}{j}\int_{[0,1]} u^j(1-u)^{k-1-j}\mathrm{Q}(\mathrm{d}u), \qquad j \in \{0,\ldots,k-1\},$$

i.e., Z_k has a mixed binomial distribution with sample size $k-1$ and random success probability Q. Let the random variable Z have distribution Q. Then

$$S(k,x) = c\mathbb{E}\big((1-(Z_k+1)^{-1})h(Z_k/k,x)\big), \qquad k\in\mathbb{N}, x\in\mathbb{R},$$

and $I(x) = c\mathbb{E}(h(Z,x))$, $x \in \mathbb{R}$. It is straightforward to check that $Z_k/k \to Z$ in distribution as $k \to \infty$, e.g., by verifying the convergence of the cumulative distribution functions (cdf) on the set of continuity points of the cdf of Z. In particular, $\lim_{k\to\infty}\mathbb{P}(Z_k \le C) = Q(0) = 0$ for every $C > 0$ and, hence, $\lim_{k\to\infty}\mathbb{E}((Z_k+1)^{-1}) = 0$. Since h is bounded and $f, f' \in \widehat{C}(\mathbb{R})$ are uniformly continuous, the family of functions $\{h(\cdot,x)|x \in \mathbb{R}\}$ is equicontinuous on $[\delta, 1-\delta]$ for every $0 < \delta < 1/2$ and uniformly bounded on $[0,1]$. From Lemma 9.4 it follows that $\mathbb{E}(h(Z_k/k,x)) \to \mathbb{E}(h(Z,x))$ uniformly in $x \in \mathbb{R}$ as $k \to \infty$, thus

$$\lim_{k\to\infty}\sup_{x\in\mathbb{R}}|S(k,x)-I(x)| = 0. \tag{5.2}$$

From $\lim_{x\to-\infty} h(Z,x) = 0$ a.s., the fact that h is bounded and the dominated convergence theorem it follows that

$$\lim_{x\to-\infty}|I(x)| = c\lim_{x\to-\infty}|\mathbb{E}(h(Z,x))| = 0. \tag{5.3}$$

Since $f \in \widehat{C}(\mathbb{R})$, $\lim_{x\to-\infty} S(k,x) = 0$ for any $k \in \mathbb{N}$. Due to (5.2) and (5.3),

$$\lim_{x\to-\infty}\sup_{k\in\mathbb{N}}|S(k,x)| = 0. \tag{5.4}$$

As $n \to \infty$, $k = k(x,n) = e^x n \to \infty$ or $x \to -\infty$. For example, for $n \in \mathbb{N}$ and $x \in E_n$, either $k \ge n^{1/2}$ or $x < -\frac{1}{2}\log n$. Distinguishing the two cases leads to

$$\begin{aligned} &\lim_{n\to\infty}\sup_{x\in E_n}|A^{(n)}f(x)-Af(x)| \\ &\le \lim_{k\to\infty}\sup_{x\in\mathbb{R}}|S(k,x)-I(x)| + \lim_{x\to-\infty}\sup_{k\in\mathbb{N}}|S(k,x)| + \lim_{x\to-\infty}|I(x)| = 0. \end{aligned} \tag{5.5}$$

By Ethier and Kurtz (1986, I, Theorem 6.1 and IV, Theorem 2.5), $X^{(n)} \to X$ in $D_\mathbb{R}[0,\infty)$ as $n \to \infty$. □

Remark 5.1. The generator $A^{(n)}$ converges even if $\Lambda(\{1\})>0$. In this case the atom at 1 can be split off from Λ such that $q_{k,j}=\binom{k}{j-1}\int_{[0,1)}u^{k-j-1}(1-u)^{j-1}\Lambda|_{[0,1)}(\mathrm{d}u)+\Lambda(\{1\})1_{\{1\}}(j)$, $j\in\{1,\ldots,k-1\}, k\geq 2$, where the first summand are the jump rates of the block counting process corresponding to the restriction $\Lambda|_{[0,1)}$ of Λ to $[0,1)$, i.e., a measure with no atom at 1. Thus,

$$A^{(n)}f(x)\;=\;S(k,x)+(f(\log n^{-1})-f(x))\Lambda(\{1\}),\qquad x\in E_n, f\in\widehat{C}(\mathbb{R}), n\in\mathbb{N},$$

where the jump rates in $S(k,x)$ correspond to $\Lambda|_{[0,1)}$, and

$$Af(x)\;=\;I(x)+h(1,x)\Lambda(\{1\})\;=\;I(x)-f(x)\Lambda(\{1\}),\qquad x\in(-\infty,0],$$

where $I(x)=\int h(u,x)\Lambda|_{[0,1)}(\mathrm{d}u)$, $x\in\mathbb{R}$. The additional term corresponds to the killing of the subordinator $-X$ at the rate $\Lambda(\{1\})$. Since $f\in\widehat{C}(\mathbb{R})$, $\lim_{n\to\infty}\sup_{x\in E_n}|(f(\log n^{-1})-f(x))\Lambda(\{1\})+f(x)\Lambda(\{1\})|=\Lambda(\{1\})\lim_{n\to\infty}|f(\log n^{-1})|=0$, i.e., the additional term converges, and again (5.5) holds true.

Remark 5.2. The approach to the convergence of the generators is related to Bernstein polynomials. The $(k-1)$-th Bernstein polynomial

$$\sum_{j=0}^{k-1}h(\tfrac{j}{k-1},x)\binom{k-1}{j}u^j(1-u)^{k-1-j}$$

of $h(\cdot,x)$ converges uniformly in $u\in[0,1]$ to $h(u,x)$ as $k\to\infty$, if $x\in\mathbb{R}$ is fixed.

6. Proofs concerning the Bolthausen–Sznitman coalescent

In this section $\Lambda=\lambda$ is the Lebesgue measure on $[0,1]$. Define $\alpha:=\alpha(t):=e^{-t}$, $t\geq 0$. The process $X^{(n)}=(X_t^{(n)})_{t\geq 0}=(\log N_t^{(n)}-\alpha\log n)_{t\geq 0}$ is a time-inhomogeneous Markov process. In order to prove the convergence in $D_{\mathbb{R}}[0,\infty)$ to X we want to show the uniform convergence of the generators. Typical convergence results are stated for time-homogeneous Markov processes and in order to use these we are going to introduce the time-space process.

6.1. *Time-space process: semigroup and generator.* Define the time-space processes $\widetilde{X}:=(t,X_t)_{t\geq 0}$ and $\widetilde{X}^{(n)}:=(t,X_t^{(n)})_{t\geq 0}$ for $n\in\mathbb{N}$. It is known (see, for example Revuz and Yor, 1999, p. 85, Exercise (1.10) or Böttcher, 2014) that $\widetilde{X}^{(n)}$ and $\widetilde{X}$ are time-homogeneous Markov processes (and exist on a new probability space). In the following the tilde symbol indicates the time-space setting. Let $\widetilde{E}_n:=\{(s,x)\in[0,\infty)\times\mathbb{R}|e^xn^{\alpha(s)}\in[n]\}$ denote the state space of $\widetilde{X}^{(n)}$, $\widetilde{E}:=[0,\infty)\times\mathbb{R}$ denote the state space of $\widetilde{X}$ and define $k:=k(s,x,n):=e^xn^{\alpha(s)}\in\mathbb{N}$ for $(s,x)\in\widetilde{E}_n$ and $n\in\mathbb{N}$. Given $f\in B(\widetilde{E})$ and $s\geq 0$, denote the function $x\mapsto f(s,x)$, $x\in\mathbb{R}$, by $\pi f(s,x)$. The limiting process X already is time-homogeneous. Recall that D, the space of twice differentiable functions $f:\mathbb{R}\to\mathbb{R}$ such that f,f',f'' and the map $x\mapsto xf'(x)$, $x\in\mathbb{R}$, belong to $\widehat{C}(\mathbb{R})$, is a core for the generator A of the semigroup $(T_t)_{t\geq 0}$ corresponding to X. The semigroup $(\widetilde{T}_t)_{t\geq 0}$ of $\widetilde{X}$, given by

$$\widetilde{T}_tf(s,x)\;=\;\mathbb{E}(f(s+t,X_{s+t})|X_s=x)\;=\;\mathbb{E}(f(s+t,\alpha(t)x+X_t)),\quad (s,x)\in\widetilde{E}, f\in B(\widetilde{E}), t\geq 0,$$

is a Feller semigroup. Let $\widetilde{D}$ denote the space of functions $f\in\widehat{C}(\widetilde{E})$ of the form $f(s,x)=\sum_{i=1}^l g_i(s)h_i(x)$ with $l\in\mathbb{N}, h_i\in D$ and $g_i\in C_1([0,\infty))$ such that $g_i,g_i'\in\widehat{C}([0,\infty))$ for $i=1,\ldots,l$. Proposition 9.6 states that $\widetilde{D}$ is a core for the generator $\widetilde{A}$ of $(\widetilde{T}_t)_{t\geq 0}$ and

$$\widetilde{A}f(s,x)\;=\;\frac{\partial}{\partial s}f(s,x)+A\pi f(s,x),\qquad (s,x)\in\widetilde{E}, f\in\widetilde{D}.\tag{6.1}$$

The 'semigroup' $(T_{s,t}^{(n)})_{s,t\geq 0}$ of $X^{(n)}$ is given by

$$\begin{aligned} T_{s,t}^{(n)}f(x) &:= \mathbb{E}(f(X_{s+t}^{(n)})|X_s^{(n)}=x) \;=\; \mathbb{E}(f(\log N_{s+t}^{(n)}-\alpha(s+t)\log n)|N_s^{(n)}=k)\\ &= \mathbb{E}(f(\log N_t^{(k)}-\alpha(s+t)\log n)), \qquad (s,x)\in\widetilde{E}_n, f\in B(\mathbb{R}), t\geq 0. \end{aligned}$$

The 'generator' $(A_s^{(n)})_{s\geq 0}$ of $(T_{s,t}^{(n)})_{s,t\geq 0}$ is given by

$$\begin{aligned} A_s^{(n)}f(x) &:= \lim_{t\to 0} t^{-1}(T_{s,t}^{(n)}f(x)-f(x))\\ &= \lim_{t\to 0} t^{-1}\Big(\mathbb{E}\Big(f(\log N_t^{(k)}-\alpha(s+t)\log n)\Big)-f(x)\Big)\\ &= -f'(x)\alpha'(s)\log n+\sum_{j=1}^{k-1}(f(x+\log\tfrac{j}{k})-f(x))q_{k,j}, \quad (s,x)\in\widetilde{E}_n. \end{aligned} \tag{6.2}$$

Here $f\in C_1(\mathbb{R})$ such that $f,f'\in\widehat{C}(\mathbb{R})$. The semigroup $(\widetilde{T}_t^{(n)})_{t\geq 0}$ of $\widetilde{X}^{(n)}$, given by

$$\begin{aligned} \widetilde{T}_t^{(n)}(s,x) &:= \mathbb{E}(f(s+t,X_{s+t})|X_s^{(n)}=x)\\ &= \mathbb{E}(f(s+t,\log N_t^{(k)}-\alpha(s+t)\log n)), \qquad (s,x)\in\widetilde{E}_n, f\in B(\widetilde{E}_n), t\geq 0, n\in\mathbb{N}, \end{aligned}$$

is a Feller semigroup on $\widehat{C}(\widetilde{E}_n)$ for every $n\in\mathbb{N}$. On $\widetilde{D}$, or more precisely, for the restriction of $f\in\widetilde{D}$ to $\widetilde{E}_n$, the generator $\widetilde{A}^{(n)}$ of $\widetilde{T}^{(n)}$ is given by

$$\widetilde{A}^{(n)}f(s,x) \;=\; \frac{\partial}{\partial s}f(s,x)+A_s^{(n)}\pi f(s,x), \qquad (s,x)\in\widetilde{E}_n, n\in\mathbb{N}. \tag{6.3}$$

6.2. *Proof of Proposition 2.2.*

Proof: (of Proposition 2.2) Recall that $\Lambda=\lambda$. Let $f\in D$. The approach to the proof is the same as in Section 5, but the function $u\mapsto f(x+\log(1-u))$, $u\in[0,1]$, demands second order approximation like in the integral part of the limiting generator (3.1). Define $h:[0,1]\times\mathbb{R}\to\mathbb{R}$ via $h(u,x):=u^{-2}(f(x+\log(1-u))-f(x)+uf'(x))$, $u\in(0,1)$, $h(0,x):=\lim_{u\searrow 0}h(u,x)=2^{-1}(f''(x)-f'(x))$ and, since $f\in\widehat{C}(\mathbb{R})$, $h(1,x):=\lim_{u\nearrow 1}h(u,x)=f'(x)-f(x)$ for $x\in\mathbb{R}$. Taylor's theorem applied to $u\mapsto f(x+\log(1-u))$, $u<1$, with evaluation point $u=0$ and exact integral remainder yields

$$\begin{aligned} h(u,x) &= u^{-2}\int_0^u\frac{u-s}{(1-s)^2}(f''(x+\log(1-s))-f'(x+\log(1-s)))\mathrm{d}s\\ &= \int_0^1\frac{1-s}{(1-us)^2}(f''(x+\log(1-us))-f'(x+\log(1-us)))\mathrm{d}s, \quad u\in[0,1), x\in\mathbb{R}. \end{aligned}$$

The latter formula of $h(u,x)$ shows that h is bounded even as u tends to 0. Putting $k=k(s,x,n)=e^x n^{\alpha(s)}$ in (6.2) yields

$$A_s^{(n)}f(x) \;=\; f'(x)R(k,x)+S(k,x), \qquad (s,x)\in\widetilde{E}_n, n\in\mathbb{N},$$

where

$$R(k,x) \;:=\; \log k-\sum_{j=1}^{k-1}\tfrac{k-j}{k}q_{k,j}-x, \qquad k\in\mathbb{N}, x\in\mathbb{R}, \tag{6.4}$$

and

$$S(k,x) \;:=\; \sum_{j=1}^{k-1}(f(x+\log\tfrac{j}{k})-f(x)+\tfrac{k-j}{k}f'(x))q_{k,j}, \qquad k\in\mathbb{N}, x\in\mathbb{R}. \tag{6.5}$$

Further define $I(x) := \int_{[0,1]} h(u,x)\Lambda(\mathrm{d}u)$, $x\in\mathbb{R}$, and observe that $Af(x)=f'(x)(1+\Psi(1)-x)+I(x)$ for $x\in\mathbb{R}$.

By Eq. (4.1) with $a=b=1$, $\frac{k-j}{k}q_{k,j}=(k-j+1)^{-1}$, $j\in\{1,\ldots,k-1\}, k\geq 2$. Hence, $\sum_{j=1}^{k-1}\frac{k-j}{k}q_{k,j}=\sum_{j=2}^{k}j^{-1}$ for $k\geq 2$. Recall that $\alpha(s)=e^{-s}$ for $s\geq 0$ and $k=k(s,x,n)=e^x n^{\alpha(s)}$ for $(s,x)\in\widetilde{E}_n$ and $n\in\mathbb{N}$. As $n\to\infty$, $k\to\infty$ or $x\to-\infty$. Fix $T>0$. E.g., if $s\in[0,T]$, then either $k\geq n^{\alpha(T+\delta)}$ or $x<-\alpha(T)(1-\alpha(\delta))\log n$, where $\delta>0$ is a constant. The well-known asymptotics of the harmonic numbers states that $\sup_{x\in\mathbb{R}}|R(k,x)-(1+\Psi(1)-x)|=|\log k-\sum_{j=1}^k j^{-1}-\Psi(1)|\to 0$ as $k\to\infty$. Clearly, $\lim_{x\to-\infty}|f'(x)|=0$. Dividing the state space as above therefore implies

$$\lim_{n\to\infty}\sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]}|f'(x)||R(k,x)-(1+\Psi(1)-x)| \;=\; 0. \tag{6.6}$$

In the next step the uniform convergence of $S(k,x)$ to $I(x)$ is shown. Substituting $k-j-1$ for j in (6.5) yields

$$\begin{aligned} S(k,x) &= \sum_{j=0}^{k-2}\big(f(x+\log(1-\tfrac{j+1}{k}))-f(x)+\tfrac{j+1}{k}f'(x)\big)q_{k,k-j-1}\\ &= \sum_{j=0}^{k-2}h(\tfrac{j+1}{k},x)\frac{(j+1)^2}{k^2}\binom{k}{j+2}\int_{[0,1]}u^j(1-u)^{k-2-j}\Lambda(\mathrm{d}u)\\ &= \frac{k-1}{k}\sum_{j=0}^{k-2}h(\tfrac{j+1}{k},x)\frac{j+1}{j+2}\binom{k-2}{j}\int_{[0,1]}u^j(1-u)^{k-2-j}\Lambda(\mathrm{d}u), \qquad k\in\mathbb{N}, x\in\mathbb{R}. \end{aligned}$$

Set $c:=\Lambda([0,1])\in(0,\infty)$ and define the probability measure Q on $([0,1],\mathcal{B}\cap[0,1])$ as $Q:=c^{-1}\Lambda$. Let the random variables Z_k, $k\in\mathbb{N}$, have distribution given by

$$\mathbb{P}(Z_k=j) \;=\; \binom{k-2}{j}\int_{[0,1]}u^j(1-u)^{k-2-j}\mathrm{Q}(\mathrm{d}u), \qquad j\in\{0,\ldots,k-2\},$$

i.e., Z_k has a mixed binomial distribution with sample size $k-2$ and random success probability Q. Let Z have distribution Q. Then

$$S(k,x) \;=\; c(1-k^{-1})\mathbb{E}\big((1-(Z_k+2)^{-1})h((Z_k+1)/k,x)\big), \qquad k\in\mathbb{N}, x\in\mathbb{R},$$

and $I(x)=c\mathbb{E}(h(Z,x))$, $x\in\mathbb{R}$. It is easy to check that $(Z_k+1)/k\to Z$ in distribution as $k\to\infty$. The family of functions $\{h(\cdot,x)|x\in\mathbb{R}\}$ is equicontinuous on $[\delta,1-\delta]$ for every $0<\delta<1/2$ and uniformly bounded on $[0,1]$. Due to $Q(\{0\})=c^{-1}\Lambda(\{0\})=0$, $Z_k\to\infty$ a.s. as $k\to\infty$, thus $\lim_{k\to\infty}\mathbb{E}(1/(Z_k+2))=0$ and the additional factor $1-(Z_k+2)^{-1}$ in the mean above can be omitted when considering the limit of $S(k,x)$ as $k\to\infty$. From Lemma 9.4 it follows that

$$\lim_{k\to\infty}\sup_{x\in\mathbb{R}}|S(k,x)-I(x)| \;=\; 0. \tag{6.7}$$

From $\lim_{x\to-\infty}h(Z,x)=0$ a.s., the fact that the functions $h(\cdot,x)$, $x\in\mathbb{R}$, are uniformly bounded and the dominated convergence theorem it follows that

$$\lim_{x\to-\infty}|I(x)| \;=\; c\lim_{x\to-\infty}|\mathbb{E}(h(Z,x))| \;=\; 0. \tag{6.8}$$

Since $f,f'\in\widehat{C}(\mathbb{R})$, $\lim_{x\to-\infty}S(k,x)=0$ for any $k\in\mathbb{N}$ and, in view of (6.7) and (6.8),

$$\lim_{x\to-\infty}\sup_{k\in\mathbb{N}}|S(k,x)| \;=\; 0. \tag{6.9}$$

As seen in the proof of Proposition 2.1, Eqs. (6.7)-(6.9) imply

$$\lim_{n\to\infty}\sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]}|S(k,x)-I(x)| \;=\; 0. \tag{6.10}$$

By (6.6), $\lim_{n\to\infty}\sup_{(s,x)\in\widetilde{E}_n,s\in[0,T]}|A_s^{(n)}f(x)-Af(x)|=0$. Due to (6.1) and (6.3),

$$\lim_{n\to\infty}\sup_{(s,x)\in\widetilde{E}_n,s\in[0,T]}|\widetilde{A}^{(n)}f(s,x)-\widetilde{A}f(s,x)| \;=\; 0$$

for every function f belonging to the core $\widetilde{D}$ and each $T>0$. From Ethier and Kurtz (1986, IV, Corollary 8.7) it follows that $\widetilde{X}^{(n)}\to\widetilde{X}$ in $D_{\widetilde{E}}[0,\infty)$, hence $X^{(n)}\to X$ in $D_{\mathbb{R}}[0,\infty)$ as $n\to\infty$. □

Remark 6.1. Note that Z_k has a discrete uniform distribution on $\{0,\ldots,k-2\}$ and Z has a continuous uniform distribution on $(0,1)$, since $\Lambda=\lambda$.

Remark 6.2. Put $\gamma(k):=\sum_{j=1}^{k-1}(k-j)q_{k,j}=\sum_{j=2}^{k}(j-1)\binom{k}{j}\lambda_{k,j}$ for $k\geq 2$. Among dust-free Λ-coalescents that do not come down from infinity the proof works for the Bolthausen–Sznitman coalescent due to the asymptotics $\gamma(k)/k=\log k-\Psi(1)-1+O(k^{-1})$ as $k\to\infty$. For other measures Λ the asymptotics of $\gamma(k)/k$ might be difficult to determine. In the proof of Proposition 2.2 the fact that $\Lambda=\lambda$ is only used to verify (6.6). Eq. (6.10) holds true more generally for finite measures Λ on $[0,1]$ with $\Lambda(\{0\})=\Lambda(\{1\})=0$ and therefore we wrote Λ and Q instead of the Bolthausen–Sznitman coalescent's driving measure λ.

7. Proof of Theorem 2.3

In this section Λ satisfies Assumption A. We continue to use the time-space setting and the notation of Subsection 6.1 with α replaced by $\alpha:=\alpha(t):=e^{-bt}$, $t\geq 0$. Define $\Lambda_D:=\Lambda-b\lambda$ and let Λ_D^+,Λ_D^- denote the nonnegative measures constituting the Jordan decomposition $\Lambda_D=\Lambda_D^+-\Lambda_D^-$ of Λ_D. The decomposition of Λ into a 'Bolthausen–Sznitman part' $b\lambda$ and a 'dust part' Λ_D is transferred to the jump rates and the generator. Proving Theorem 2.3 now only requires to suitable arrange equations already obtained in Sections 5 and 6. To be precise, the results of Section 5 are applied to the summands $\Lambda_D^{\pm}$ of Λ_D, but we omit this detail in the following.

Proof: (of Theorem 2.3) Let $q_{k,j}^{\lambda},q_{k,j}^{D,+}$ and $q_{k,j}^{D,-}$ denote the rates of the block counting process corresponding to λ,Λ_D^+ and Λ_D^-, respectively, and define $q_{k,j}^{D}:=q_{k,j}^{D,+}-q_{k,j}^{D,-}$ for $j\in\{1,\ldots,k\}$ and $k\in\mathbb{N}$. Obviously, $q_{k,j}=bq_{k,j}^{\lambda}+q_{k,j}^{D}$. Recall that $k=k(s,x,n)=e^x n^{\alpha(s)}\in\mathbb{N}$ for $(s,x)\in\widetilde{E}_n$ and $n\in\mathbb{N}$. From (6.2) it follows that the 'generator' $A_s^{(n)}$ of $X^{(n)}=(X_t^{(n)})_{t\geq 0}=(\log N_t^{(n)}-\alpha(t)\log n)_{t\geq 0}$ is given by

$$A_s^{(n)}f(x) \;=\; bR(k,x)f'(x)+bS_{BS}(k,x)+S_D(k,x), \qquad (s,x)\in\widetilde{E}_n, n\in\mathbb{N},$$

where

$$\begin{aligned}
R(k,x) \;&:=\; \log k-\sum_{j=1}^{k-1}\tfrac{k-j}{k}q_{k,j}^{\lambda}-x,\\
S_{BS}(k,x) \;&:=\; \sum_{j=1}^{k-1}(f(x+\log\tfrac{j}{k})-f(x)+\tfrac{k-j}{k}f'(x))q_{k,j}^{\lambda},\\
S_D(k,x) \;&:=\; \sum_{j=1}^{k-1}(f(x+\log\tfrac{j}{k})-f(x))q_{k,j}^{D},
\end{aligned}$$

are defined as in (6.4), (6.5) and (5.1) for $k \in \mathbb{N}$ and $x \in \mathbb{R}$, and $f \in C_1(\mathbb{R})$ such that $f, f' \in \widehat{C}(\mathbb{R})$. By Lemma 3.1 and Eq. (2.3), the generator A of $X = (X_t)_{t\geq 0}$ can be written as

$$\begin{aligned} Af(x) &= b(1+\Psi(1)-x)f'(x) + b\int_{[0,1]} \frac{f(x+\log(1-u)) - f(x) + uf'(x)}{u^2}\lambda(\mathrm{d}u) \\ &\quad + \int_{[0,1]} \frac{f(x+\log(1-u)) - f(x)}{u^2}\Lambda_D(\mathrm{d}u), \qquad x \in \mathbb{R}, f \in D. \end{aligned}$$

From Eqs. (6.6), (6.10) and (5.2)-(5.4) it follows that $\lim_{n\to\infty} \sup_{(s,x)\in \widetilde{E}_n, s\in[0,T]} |A_s^{(n)} f(x) - Af(x)| = 0$ for $f \in D$. Due to (6.1) and (6.3),

$$\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} |\widetilde{A}^{(n)} f(s,x) - \widetilde{A}f(s,x)| \;=\; 0$$

for every $f \in \widetilde{D}$ and $T > 0$. By Proposition 9.6, the space $\widetilde{D}$ is a core for $\widetilde{A}$. Thus, it follows from Ethier and Kurtz (1986, IV, Corollary 8.7) that $\widetilde{X}^{(n)} \to \widetilde{X}$ in $D_{\widetilde{E}}[0,\infty)$, hence $X^{(n)} \to X$ in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$. □

8. Proof of Theorem 2.4

In this section Λ satisfies Assumption A. The process $Y^{(n)} = (Y_t^{(n)})_{t\geq 0} = (\log L_t^{(n)} - e^{bt}\log n)_{t\geq 0}$ is a possibly time-inhomogeneous Markov process, depending on whether $b > 0$ or not, hence we set up the time-space framework. We provide two proofs. Using Theorem 2.3 and Siegmund-duality, in the first proof the convergence of the one-dimensional distributions and subsequently the uniform convergence of the semigroups is shown. The second proof, in which the uniform convergence of generators is shown, resembles previous ones.

Proof: (First proof of Theorem 2.4) For $x \in \mathbb{R}$ and $t \geq 0$ define $m := \lceil e^y n^{e^{bt}} \rceil \in \mathbb{N}$. If $\varrho_t((-\infty,0)) = \int_{[0,1]} u^{-2}\Lambda(\mathrm{d}u) = \infty$, then X_t has a continuous distribution for every $t > 0$. Eq. (1.2) and Theorem 2.3 imply that

$$\begin{aligned} \mathbb{P}(Y_t^{(n)} \geq y) &= \mathbb{P}(L_t^{(n)} \geq m) \;=\; \mathbb{P}(N_t^{(m)} \leq n) \;=\; \mathbb{P}(X_t^{(m)} \leq \log n - e^{-bt}\log m) \\ &\longrightarrow \mathbb{P}(X_t \leq -e^{-bt}y) \;=\; \mathbb{P}(-e^{bt}X_t \geq y), \qquad y \in \mathbb{R}, t \geq 0, \end{aligned} \tag{8.1}$$

as $n \to \infty$. If $\int_{[0,1]} u^{-2}\Lambda(\mathrm{d}u) < \infty$, then the dust condition is satisfied. Hence, $b = 0$ and (8.1) holds true for $-y$ in the set C_{X_t} of continuity points of X_t. Since $Y_t \stackrel{\mathrm{d}}{=} -e^{bt}X_t$ with $b = 0$, $\lim_{n\to\infty} \mathbb{P}(-Y_t^{(n)} \leq -y) = \mathbb{P}(-Y_t \leq -y)$ for every $-y \in C_{X_t} = C_{-Y_t}$. Thus, $Y_t^{(n)}$ converges in distribution to Y_t as $n \to \infty$ for every $t \geq 0$.

Define the time-space processes $\widetilde{Y}^{(n)} := (t, Y_t^{(n)})_{t\geq 0}$, $n \in \mathbb{N}$, and $\widetilde{Y} := (t, Y_t)_{t\geq 0}$. The processes $\widetilde{Y}^{(n)}$ and $\widetilde{Y}$ are time-homogeneous Markov processes with state spaces $\widetilde{E}_n = \{(s,y) | s \geq 0, e^y n^{e^{bs}} \in \{n, n+1, \ldots\}\}$ and $\widetilde{E} = [0,\infty) \times \mathbb{R}$ and semigroups $(\widetilde{T}_t^{(n)})_{t\geq 0}$ and $(\widetilde{T}_t)_{t\geq 0}$. Define $k := k(s,y,n) := e^y n^{e^{bs}} \in \{n, n+1, \ldots\}$ for $(s,y) \in \widetilde{E}_n$ and $n \in \mathbb{N}$. Then

$$\begin{aligned} \widetilde{T}_t^{(n)} f(s,y) &= \mathbb{E}(f(s+t, Y_{s+t}^{(n)}) | Y_s^{(n)} = y) \;=\; \mathbb{E}(f(s+t, \log L_t^{(k)} - e^{b(t+s)}\log n)) \\ &= \mathbb{E}(f(s+t, e^{bt}y + Y_t^{(k)})), \qquad (s,y) \in \widetilde{E}_n, f \in B(\widetilde{E}), t \geq 0, n \in \mathbb{N}. \end{aligned}$$

Fix $t > 0$ and first let $f \in B(\widetilde{E})$ be of the form $f(s,y) = g(s)h(y)$, $(s,y) \in \widetilde{E}$, where $g \in B([0,\infty))$ and $h \in \widehat{C}(\mathbb{R})$. Clearly, $\widetilde{T}_t^{(n)} f(s,y) = g(s+t)\mathbb{E}(h(e^{bt}y + Y_t^{(k)}))$, $(s,y) \in \widetilde{E}_n, n \in \mathbb{N}$, and $\widetilde{T}_t f(s,y) = \mathbb{E}(f(s+t, Y_{s+t}) | Y_s = y) = g(s+t)T_t h(y) = g(s+t)\mathbb{E}(h(e^{bt}y + Y_t))$, $(s,y) \in \widetilde{E}$, where the distribution of Y_t is defined by its characteristic function χ_t, given by (2.9). Note that h is uniformly continuous and bounded. For $y \in \mathbb{R}$ define the function $h_y : \mathbb{R} \to \mathbb{R}$ via $h_y(x) := h(e^{bt}y + x)$,

$x \in \mathbb{R}$. The family of functions $\{h_y | y \in \mathbb{R}\}$ is equicontinuous and uniformly bounded. From the weak convergence of $Y_t^{(k)}$ to Y_t as $k \to \infty$ and Ranga Rao (1962, Theorem 3.1) it follows that $\lim_{k\to\infty} \sup_{y\in\mathbb{R}} |\mathbb{E}(h(e^{bt}y + Y_t^{(k)})) - \mathbb{E}(h(e^{bt}y + Y_t))| = 0$. Since $k = e^y n^{e^{bs}} \geq n$ for $(s,y) \in \widetilde{E}_n$ and $n \in \mathbb{N}$, $\lim_{n\to\infty} \sup_{(s,y)\in\widetilde{E}_n} |\mathbb{E}(h(e^{bt}y + Y_t^{(k)})) - \mathbb{E}(h(e^{bt}y + Y_t))| = 0$. Thus,

$$\lim_{n\to\infty} \sup_{(s,y)\in\widetilde{E}_n} |\widetilde{T}_t^{(n)} f(s,y) - \widetilde{T}_t f(s,y)| \;=\; 0. \tag{8.2}$$

By linearity, (8.2) holds for the algebra of functions $f \in B(\widetilde{E})$ of the form $f(s,y) = \sum_{i=1}^l g_i(s)h_i(y)$, $(s,y) \in \widetilde{E}$, where $l \in \mathbb{N}, g_i \in B([0,\infty))$ and $h_i \in \widehat{C}(\mathbb{R})$ for $i = 1,\ldots,l$. This algebra of functions separates points and vanishes nowhere. According to the Stone–Weierstrass theorem for locally compact spaces (see e.g. de Branges, 1959) it is a dense subset of $B(\widetilde{E})$. Hence, (8.2) holds true for $f \in B(\widetilde{E})$. Ethier and Kurtz (1986, IV, Theorem 2.11) states that $\widetilde{Y}^{(n)} \to \widetilde{Y}$ in $D_{\widetilde{E}}[0,\infty)$, hence $Y^{(n)} \to Y$ in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$. □

The process Y defined by (2.8) and (2.9) is a generalized Ornstein–Uhlenbeck process (with nonnegative linear drift) as in Sato and Yamazato (1984). The underlying infinitely divisible distribution has characteristic exponent $v \mapsto \psi(-v)$, $v \in \mathbb{R}$. According to Sato and Yamazato (1984, Theorem 3.1), D is a core for the corresponding generator A and

$$Af(y) \;=\; f'(y)(-a+by) + \int_{[0,1]} (f(y - \log(1-u)) - f(y) - uf'(y))u^{-2}\Lambda(\mathrm{d}u) \tag{8.3}$$

for $y \in \mathbb{R}$ and $f \in D$; comparatively see Lemma 3.1 and its proof.

Proof: (Second proof of Theorem 2.4) The ‘generator’ $(A_s^{(n)})_{s\geq 0}$ of $Y^{(n)}$ is given by

$$A_s^{(n)} f(y) \;=\; -f'(y)be^{bs}\log n + \sum_{j > e^y n^{e^{bs}}} (f(\log j - e^{bs}\log n) - f(y))\gamma_{e^y n^{e^{bs}},j}, \quad (s,y) \in \widetilde{E}_n, n \in \mathbb{N}.$$

Here $f \in C_1(\mathbb{R})$ such that $f, f' \in \widehat{C}(\mathbb{R})$. Putting $k := k(s,y,n) := e^y n^{e^{bs}}$ for $(s,y) \in \widetilde{E}_n$ and $n \in \mathbb{N}$ yields

$$A_s^{(n)} f(y) \;=\; bf'(y)(-\log k + y) + \sum_{j=1}^\infty (f(y + \log(1+\tfrac{j}{k})) - f(y))\gamma_{k,k+j}, \quad (s,y) \in \widetilde{E}_n, n \in \mathbb{N}.$$

Define $\Lambda_D := \Lambda - b\lambda$ and let Λ_D^+, Λ_D^- denote the nonnegative measures constituting the Jordan decomposition $\Lambda_D = \Lambda_D^+ - \Lambda_D^-$ of Λ_D. Let $\gamma_{k,j}^\lambda, \gamma_{k,j}^{D,+}$ and $\gamma_{k,j}^{D,-}$ denote the jump rates of the fixation line corresponding to λ, Λ_D^+ and Λ_D^-, respectively, and define $\gamma_{k,j}^D := \gamma_{k,j}^{D,+} - \gamma_{k,j}^{D,-}$ for $j \in \{k, k+1, \ldots\}$ and $k \in \mathbb{N}$. Then $\gamma_{k,k+j} = b\gamma_{k,k+j}^\lambda + \gamma_{k,k+j}^D$, $k \in \mathbb{N}, j \in \mathbb{N}_0$, and

$$A_s^{(n)} f(y) \;=\; bf'(y)R(k,y) + bS_{BS}(k,y) + S_D(k,y), \qquad (s,y) \in \widetilde{E}_n, n \in \mathbb{N}, \tag{8.4}$$

where

$$R(k,y) \;:=\; -\log k + y + \sum_{j=1}^k \tfrac{j}{k}\gamma_{k,k+j}^\lambda, \qquad k \in \mathbb{N}, y \in \mathbb{R},$$

$$S_{BS}(k,y) \;:=\; \sum_{j=1}^\infty (f(y + \log(1+\tfrac{j}{k})) - f(y) - \tfrac{j}{k}1_{[0,1]}(\tfrac{j}{k})f'(y))\gamma_{k,k+j}^\lambda, \qquad k \in \mathbb{N}, y \in \mathbb{R},$$

$$S_D(k,y) \;:=\; \sum_{j=1}^\infty (f(y + \log(1+\tfrac{j}{k})) - f(y))\gamma_{k,k+j}^D, \qquad k \in \mathbb{N}, y \in \mathbb{R},$$

and $f \in C_1(\mathbb{R})$ such that $f, f' \in \widehat{C}(\mathbb{R})$. Using the decomposition of Λ on Eq. (8.3) yields

$$Af(y) \;=\; bf'(y)(-1-\Psi(1)+y)+bI_{BS}(y)+I_D(y), \qquad y\in\mathbb{R}, f\in D,$$

where

$$I_{BS}(y) \;:=\; \int_{[0,1]}(f(y-\log(1-u))-f(y)-uf'(y))u^{-2}\lambda(\mathrm{d}u), \qquad y\in\mathbb{R},$$
$$I_D(y) \;:=\; \int_{[0,1]}(f(y-\log(1-u))-f(y))u^{-2}\Lambda_D(\mathrm{d}u), \qquad y\in\mathbb{R}.$$

Let $f\in D$. In the Bolthausen–Sznitman coalescent, $\gamma^\lambda_{k,k+j}=k/(j(j+1))$ for $k,j\in\mathbb{N}$ and hence $\sum_{j=1}^k\frac{j}{k}\gamma^\lambda_{k,k+j}=\sum_{j=1}^k(j+1)^{-1}=H_{k+1}-1=\log k-1-\Psi(1)+o(1)$ as $k\to\infty$. Here H_k denotes the k-th harmonic number for $k\in\mathbb{N}$. Thus,

$$\lim_{k\to\infty}\sup_{y\in\mathbb{R}}|R(k,y)-(-1-\Psi(1)+y)| \;=\; 0. \tag{8.5}$$

The function $h_{BS}:[0,1]\times\mathbb{R}\to\mathbb{R}$, defined via $h_{BS}(u,y):=u^{-2}(f(y-\log(1-u))-f(y)-\frac{u}{1-u}1_{[0,1/2]}(u)f'(y))$, $u\in[0,1], y\in\mathbb{R}$, is bounded. Let the random variables Z_k, $k\in\mathbb{N}$, have distribution given by

$$\mathbb{P}(Z_k=j) \;=\; \binom{k+j-2}{j-1}\int_{[0,1]}u^{j-1}(1-u)^k\lambda(\mathrm{d}u), \qquad j,k\in\mathbb{N},$$

i.e., Z_k-1 has a mixed negative binomial distribution. Observe that $h_{BS}(1-(1+\frac{j}{k})^{-1},y)=(\frac{j}{k+j})^{-2}(f(y+\log(1+\frac{j}{k}))-f(y)-\frac{j}{k}1_{[0,1]}(\frac{j}{k})f'(y))$, $y\in\mathbb{R}$, and $\gamma^\lambda_{k,k+j}=(\frac{j}{k+j})^{-2}(1-(k+j)^{-1})(1-(j+1)^{-1})\mathbb{P}(Z_k=j)$ for $j,k\in\mathbb{N}$. Hence,

$$S_{BS}(k,y) \;=\; \mathbb{E}\Big(h_{BS}(1-(1+Z_k/k)^{-1},y)\Big(1-\frac{1}{k+Z_k}\Big)\Big(1-\frac{1}{Z_k+1}\Big)\Big).$$

Let Z have uniform distribution on $(0,1)$. Then $I_{BS}(y)=\mathbb{E}(h_{BS}(Z,y))$ for $y\in\mathbb{R}$ due to $\int_{[0,1]}u^{-2}(u-\frac{u}{1-u}1_{[0,1/2]}(u))\lambda(\mathrm{d}u)=\int_0^{1/2}-(1-u)^{-1}\mathrm{d}u+\int_{1/2}^1u^{-1}\mathrm{d}u=0$. The function $g:(0,\infty)\to(0,1)$, defined via $g(u):=1-(1+u)^{-1}$, $u\in(0,\infty)$, is bounded and continuous. Since $Z_k/k\to Z/(1-Z)$ in distribution as $k\to\infty$, $1-(1+Z_k/k)^{-1}=g(Z_k/k)\to g(Z/(1-Z))=Z$ in distribution as $k\to\infty$. In particular, the random variables have values in $[0,1]$. When considering the limit $k\to\infty$, the factor $(1-(k+Z_k)^{-1})(1-(Z_k+1)^{-1})$ has no influence on $S_{BS}(k,y)$. From Lemma 9.4 it follows that

$$\lim_{k\to\infty}\sup_{y\in\mathbb{R}}|S_{BS}(k,y)-I_{BS}(y)| \;=\; 0. \tag{8.6}$$

The measure Λ_D is real-valued. Eq. (8.7) below can be proven when Λ_D is replaced by Λ_D^+ and Λ_D^- in this paragraph, and then holds for Λ_D by linearity. The function $h_D:[0,1]\times\mathbb{R}\to\mathbb{R}$, defined via $h_D(u,y):=u^{-1}(f(y-\log(1-u))-f(y))$, $u\in[0,1], y\in\mathbb{R}$, is bounded. By assumption, $c:=\int_{[0,1]}u^{-1}\Lambda_D(\mathrm{d}u)<\infty$. As long as $c>0$, define the probability measure Q on $([0,1],\mathcal{B}\cap[0,1])$ via $Q(A):=c^{-1}\int_A u^{-1}\Lambda_D(\mathrm{d}u)$, $A\in\mathcal{B}\cap[0,1]$, and let the random variables Z_k, $k\in\mathbb{N}$, have distribution given by

$$\mathbb{P}(Z_k=j) \;=\; \binom{k+j-1}{j}\int_{[0,1]}u^j(1-u)^kQ(\mathrm{d}u), \qquad j\in\mathbb{N}_0, k\in\mathbb{N},$$

i.e., Z_k has a mixed negative binomial distribution. Observe that $h_D(1-(1+\frac{j}{k})^{-1},y) = (f(y+\log(1+\frac{j}{k}))-f(y))\frac{k+j}{j}$, $y\in\mathbb{R}$, and $\gamma^D_{k,k+j} = c\frac{k+j}{j}(1-(1+j)^{-1})\mathbb{P}(Z_k=j)$ for $j,k\in\mathbb{N}$. Hence,

$$\begin{aligned} S_D(k,y) &= \sum_{j=0}^{\infty}(f(y+\log(1+\tfrac{j}{k}))-f(y))\gamma^D_{k,k+j} \\ &= c\mathbb{E}(h_D(1-(1+Z_k/k)^{-1},y)(1-(1+Z_k)^{-1})), \qquad k\in\mathbb{N}, y\in\mathbb{R}. \end{aligned}$$

Let the random variable Z have distribution Q. In particular, $I_D(y) = c\mathbb{E}(h_D(Z,y))$, $y\in\mathbb{R}$. According to Lemma 9.4 and since $1-(1+Z_k/k)^{-1}$ converges in distribution to Z as $k\to\infty$,

$$\lim_{k\to\infty}\sup_{y\in\mathbb{R}}|\mathbb{E}(h_D(1-(1+Z_k/k)^{-1},y))-\mathbb{E}(h_D(Z,y))| = 0.$$

Thus,

$$\lim_{k\to\infty}\sup_{y\in\mathbb{R}}|S_D(k,y)-I_D(y)| = 0. \tag{8.7}$$

Note that Eq. (8.7) holds true for $c=0$ as well.

Taking into account that $k=e^y n^{e^{bs}}\ge n$ for $(s,y)\in\widetilde{E}_n$ and $n\in\mathbb{N}$, Eqs. (8.4)-(8.7) imply

$$\lim_{n\to\infty}\sup_{(s,y)\in\widetilde{E}_n}|A_s^{(n)}f(y)-Af(y)| = 0.$$

The time-space variant of Ethier and Kurtz (1986, IV, Corollary 8.7) as implemented in the proof of Theorem 2.3 yields the desired convergence of $Y^{(n)}$ to Y in $D_{\mathbb{R}}[0,\infty)$ as $n\to\infty$. □

9. Appendix

Lemma 9.1. *Suppose that Λ satisfies Assumption A. Then the following statements hold.*

a) $b=\lim_{\varepsilon\to 0}\varepsilon^{-1}\Lambda((0,\varepsilon))$.

b) *The Λ-coalescent does not come down from infinity.*

Proof: a) If the condition $\int_{[0,1]}u^{-1}\Lambda(\mathrm{d}u)<\infty$ for dust is given, then Assumption A is satisfied with $b=0$ and, by dominated convergence,

$$\frac{\Lambda((0,\varepsilon))}{\varepsilon} \le \int_{(0,\varepsilon)}u^{-1}\Lambda(\mathrm{d}u) \longrightarrow 0, \qquad \varepsilon\to 0.$$

Hence, a) holds for coalescents with dust. Now suppose that Λ satisfies Assumption A. Define $\Lambda_D := \Lambda - b\lambda$ and let Λ_D^+ and Λ_D^- denote the nonnegative measures constituting the Jordan decomposition $\Lambda_D=\Lambda_D^+-\Lambda_D^-$ of Λ_D. By assumption and the first part of the proof, $\lim_{\varepsilon\to 0}\varepsilon^{-1}\Lambda_D^{\pm}((0,\varepsilon))=0$. From the decomposition $\Lambda=b\lambda+\Lambda_D^+-\Lambda_D^-$ it follows that

$$\frac{\Lambda((0,\varepsilon))}{\varepsilon} = b+\frac{\Lambda_D^+((0,\varepsilon))}{\varepsilon}-\frac{\Lambda_D^-((0,\varepsilon))}{\varepsilon} \longrightarrow b, \qquad \varepsilon\to 0.$$

b) Let $|\Lambda_D|=\Lambda_D^++\Lambda_D^-$ denote the total variation of Λ_D. Define $\eta_k^\Lambda := k\sum_{j=0}^{k-2}\int_{[0,1]}(1-u)^j\Lambda(\mathrm{d}u)$ and $\eta_k^{b\lambda}$ and $\eta_k^{|\Lambda_D|}$ similarly with $b\lambda$ and $|\Lambda_D|$ in place of Λ for $k\ge 2$. By assumption,

$$\lim_{k\to\infty}k^{-1}\eta_k^{|\Lambda_D|} = \int_{[0,1]}u^{-1}|\Lambda_D|(\mathrm{d}u)<\infty.$$

From

$$(k\log k)^{-1}\eta_k^{b\lambda} = b(\log k)^{-1}\sum_{j=0}^{k-2}\int_0^1(1-u)^j\mathrm{d}u = b(\log k)^{-1}\sum_{j=0}^{k-2}(j+1)^{-1}\longrightarrow b, \qquad k\to\infty,$$

it follows that $\eta_k^{b\lambda}+\eta_k^{|\Lambda_D|}\sim bk\log k$ as $k\to\infty$. Due to $\Lambda\le b\lambda+|\Lambda_D|$, it holds that $\eta_k^{\Lambda}\le\eta_k^{b\lambda}+\eta_k^{|\Lambda_D|}$ for $k\ge 2$. Hence,

$$\sum_{k=2}^{\infty}\left(\eta_k^{\Lambda}\right)^{-1}\ge\sum_{k=2}^{\infty}\left(\eta_k^{b\lambda}+\eta_k^{|\Lambda_D|}\right)^{-1} \;=\; \infty.$$

The claim b) then follows from Schweinsberg's criterion (Schweinsberg, 2000b, Corollary 2). □

Remark 9.2. Any converse statements of Lemma 9.1 do not hold: neither a) nor b) nor a) and b) together imply that Assumption A holds, which can be seen by looking at the measure Λ having density f with respect to Lebesgue measure given by $f(u):=(-\log u)^{-1}$ for $0<u<1/2$ and $f(u):=0$ otherwise.

The following lemma is a generalization of the integral criterion of convergence in distribution and is applied in Sections 5-8 to prove the uniform convergence of generators. In the statement the notion of equicontinuity is used, whose definition is first recalled.

Definition 9.3. A family F of functions $f:E\to\mathbb{R}$ on a metric space E with metric d is called equicontinuous if for every $\varepsilon>0$ there exists $\delta>0$ such that $|f(x)-f(y)|<\varepsilon$ for all $f\in F$ and $x,y\in E$ with $d(x,y)<\delta$. The family F is called equicontinuous on a subset $V\subseteq E$ if the family $\{f|_V|f\in F\}$ is equicontinuous. Here $f|_V$ denotes the restriction of f to V.

Lemma 9.4. *Let $X,X_1,X_2,\ldots$ be random variables on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ with values in $[0,1]$ such that $\mathbb{P}(X=0)=\mathbb{P}(X=1)=0$ and $X_n\to X$ in distribution as $n\to\infty$. Suppose that the family F of functions $f:[0,1]\to\mathbb{R}$ is uniformly bounded on $[0,1]$, i.e., $M:=\sup_{f\in F}\sup_{x\in[0,1]}|f(x)|<\infty$, and equicontinuous on $[\delta,1-\delta]$ for every $0<\delta<1/2$. In particular, $f\in F$ is bounded and continuous on $(0,1)$. Then*

$$\lim_{n\to\infty}\sup_{f\in F}|\mathbb{E}(f(X_n))-\mathbb{E}(f(X))| \;=\; 0.$$

Proof: Let $\varepsilon>0$ be arbitrary. The assumption $\mathbb{P}(X=0)=\mathbb{P}(X=1)=0$ and the convergence of X_n to X in distribution as $n\to\infty$ provide the existence of $0<\delta<1/2$ and $n_0\in\mathbb{N}$ such that $\mathbb{P}(X_n\notin[\delta,1-\delta])<\varepsilon/(4M)$ for $n\ge n_0$ and $\mathbb{P}(X\notin[\delta,1-\delta])<\varepsilon/(4M)$. For $f\in F$ define $\tilde f:[0,1]\to\mathbb{R}$ via $\tilde f(u):=f(\delta)$, $0\le u\le\delta$, $\tilde f(u):=f(u)$, $\delta\le u\le 1-\delta$, and $\tilde f(u):=f(1-\delta)$, $1-\delta\le u\le 1$. Then $\{\tilde f|f\in F\}$ is bounded (by M) and equicontinuous on $[0,1]$. Ranga Rao (1962, Theorem 3.1) yields

$$\lim_{n\to\infty}\sup_{f\in F}|\mathbb{E}(\tilde f(X_n))-\mathbb{E}(\tilde f(X))| \;=\; 0.$$

From

$$\begin{aligned}|\mathbb{E}(f(X_n))-\mathbb{E}(f(X))| \;&\le\; \mathbb{E}(|f(X_n)-\tilde f(X_n)|)\\ &\quad+|\mathbb{E}(\tilde f(X_n))-\mathbb{E}(\tilde f(X))|+\mathbb{E}(|\tilde f(X)-f(X)|)\\ &\le\; 2M\mathbb{P}(X_n\notin[\delta,1-\delta])\\ &\quad+2M\mathbb{P}(X\notin[\delta,1-\delta])+|\mathbb{E}(\tilde f(X_n))-\mathbb{E}(\tilde f(X))|, \qquad n\in\mathbb{N}, f\in F,\end{aligned}$$

it follows that $\lim_{n\to\infty}\sup_{f\in F}|\mathbb{E}(f(X_n))-\mathbb{E}(f(X))|\le\varepsilon$. Since $\varepsilon>0$ is arbitrary, the proof is complete. □

Remark 9.5. In Ranga Rao (1962, Theorem 3.1) the state space is more generally a separable metric space, but equicontinuity of F is required to hold on the whole state space.

Let E be a complete separable metric space and equip $\widetilde E:=[0,\infty)\times E$ with the product metric. The following proposition treats the generator of time-space processes of time-homogeneous Feller processes.

Proposition 9.6. *Suppose that $(T_t)_{t\ge 0}$ is a Feller semigroup on $\widehat{C}(E)$ with generator A and that D is a core for A. For $f\in\widehat{C}(\widetilde{E})$ and $s\in[0,\infty)$ let $\pi f(s,x)$ denote the function $x\mapsto f(s,x)$, $x\in E$. The semigroup $(\widetilde{T}_t)_{t\ge 0}$, defined via*

$$\widetilde{T}_t f(s,x) \ := \ T_t\pi f(s+t,x), \qquad (s,x)\in\widetilde{E}, f\in B(\widetilde{E}), t\ge 0,$$

is a Feller semigroup on $\widehat{C}(\widetilde{E})$. Let $\widetilde{D}$ denote the space of functions $f\in\widehat{C}(\widetilde{E})$ of the form $f(s,x)=\sum_{i=1}^l g_i(s)h_i(x)$, $(s,x)\in\widetilde{E}$, where $l\in\mathbb{N}, h_i\in D$ and $g_i\in C_1([0,\infty))$ such that $g_i,g_i'\in\widehat{C}([0,\infty))$ for $i=1,\ldots,l$. Then $\widetilde{D}$ is a core for the generator $\widetilde{A}$ of $(\widetilde{T}_t)_{t\ge 0}$ and

$$\widetilde{A}f(s,x) \ = \ \frac{\partial}{\partial s}f(s,x)+A\pi f(s,x), \qquad (s,x)\in\widetilde{E}, f\in\widetilde{D}. \tag{9.1}$$

Proof: Observe that all functions involved in the proof are bounded and uniformly continuous. Clearly, the right-hand side of (9.1) lies in $\widehat{C}(\widetilde{E})$. The core D is a dense subset of $\widehat{C}(E)$. Hence $\widetilde{D}$ is a dense subset of the space D_0 of functions $f\in\widehat{C}(\widetilde{E})$ of the form $f(s,x)=\sum_{i=1}^l g_i(s)h_i(x)$, $(s,x)\in\widetilde{E}$, where $l\in\mathbb{N}, h_i\in\widehat{C}(E)$ and $g_i\in\widehat{C}([0,\infty))$ for $i=1,\ldots,l$. The algebra D_0 separates points and vanishes nowhere. The Stone–Weierstrass theorem for locally compact spaces (e.g. de Branges, 1959) ensures that D_0 is a dense subset of $\widehat{C}(\widetilde{E})$. In de Branges (1959) the theorem is stated for complex-valued functions, but it remains true for real-valued functions. To see this, let $f\in\widehat{C}(E)\subseteq\widehat{C}(E,\mathbb{C})$ be arbitrary. By the theorem, there exist a sequence $(k_n)_{n\in\mathbb{N}}\subseteq\widehat{C}(E,\mathbb{C})$ such that $\lim_{n\to\infty}||k_n-f||=0$. Then $f_n:=\mathrm{Re}(k_n)\in\widehat{C}(E)$, $n\in\mathbb{N}$, and $\|f_n-f\|\le\|k_n-f\|\to 0$ as $n\to\infty$. Thus, $\widetilde{D}$ is a dense subset of $\widehat{C}(\widetilde{E})$ as well. If $h\in D$ and $g\in C_1([0,\infty))$ such that $g,g'\in\widehat{C}([0,\infty))$, then

$$t^{-1}(\widetilde{T}_t g(s)h(x)-g(s)h(x)) \ = \ t^{-1}(g(s+t)-g(s))h(x)+g(s+t)t^{-1}(T_th(x)-h(x))$$

converges uniformly in $(s,x)\in\widetilde{E}$ to $g'(s)h(x)+g(s)Ah(x)$ as $t\searrow 0$, thus $\widetilde{D}$ lies in the domain of $\widetilde{A}$ and (9.1) holds true. By the same argument as above, the space D_1 of functions $f\in\widehat{C}(\widetilde{E})$ of the form $f(s,x)=\sum_{i=1}^l g_i(s)h_i(x)$, $(s,x)\in\widetilde{E}$, where $g_i(s)=c_i\exp(-a_is)$, $s\in[0,\infty)$ with $c_i\in\mathbb{R}$ and $a_i>0$ and $h_i\in D$ for $i=1,\ldots,l$, is a dense subset of $\widehat{C}(\widetilde{E})$. By Hille–Yosida theory (see e.g. Ethier and Kurtz, 1986, I, Proposition 3.1) it now suffices to show that the image of $\lambda I-\widetilde{A}|_{\widetilde{D}}$ is a dense subspace of $\widehat{C}(\widetilde{E})$ for some $\lambda>0$ in order to prove that $\widetilde{D}$ is a core for $\widetilde{A}$. Here I denotes the identity map on $\widehat{C}(E)$ or $\widehat{C}(\widetilde{E})$. Let $\varepsilon>0$ and $f\in\widehat{C}(\widetilde{E})$ be arbitrary. By density of D_1 in $\widehat{C}(\widetilde{E})$, there exists $f_1\in D_1$ of the form $f_1(s,x)=\sum_{i=1}^l g_i(s)h_i(x)$, $(s,x)\in\widetilde{E}$, such that $\|f_1-f\|<\varepsilon/2$. Since D is a core for A, the image of $\lambda I-A|_D$ is a dense subset of $\widehat{C}(E)$ for every $\lambda>0$, in particular for $\lambda+a_i$ in place of λ. Hence there exists $r_i\in D$ such that $\|(\lambda+a_i)r_i-Ar_i-h_i\|<\varepsilon/(2l\|g_i\|)$ for $i=1,\ldots,l$. Clearly, the function $(s,x)\mapsto\sum_{i=1}^l g_i(s)r_i(x)$, $(s,x)\in\widetilde{E}$, belongs to $\widetilde{D}$ and, by (9.1),

$$\begin{aligned}\|(\lambda I-\widetilde{A})\sum_{i=1}^l g_i(s)r_i(x)-f(s,x)\| \ &\le \ \|(\lambda I-\widetilde{A})\sum_{i=1}^l g_i(s)r_i(x)-\sum_{i=1}^l g_i(s)h_i(x)\|+\|f_1-f\|\\ &\le \ \sum_{i=1}^l\|g_i((\lambda+a_i)r_i-Ar_i-h_i)\|+\varepsilon/2 \ \le \ \varepsilon.\end{aligned}$$

In the second last step it is used that $g_i'(s)=-a_ig_i(s)$, $s\in[0,\infty)$ for $i=1,\ldots,l$. Since $\varepsilon>0$ is arbitrary, the proof is complete. □

Remark 9.7. The last part of the proof of Proposition 9.6 can be simplified under the additional assumption that $T_tD\subseteq D$ for every $t>0$. Then $\widetilde{T}_t\widetilde{D}\subseteq\widetilde{D}$ for every $t\ge 0$ and the claim follows by applying the core theorem (Ethier and Kurtz, 1986, I, Proposition 3.3).

References

Applebaum, D. *Lévy processes and stochastic calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition (2009). ISBN 978-0-521-73865-1. MR2512800.

Baur, E. and Bertoin, J. The fragmentation process of an infinite recursive tree and Ornstein-Uhlenbeck type processes. *Electron. J. Probab.*, **20**, no. 98, 20 (2015). MR3399834.

Berestycki, N. *Recent progress in coalescent theory*, volume 16 of *Ensaios Matemáticos*. Sociedade Brasileira de Matemática, Rio de Janeiro (2009). ISBN 978-85-85818-40-1. MR2574323.

Bogachev, V. I., Röckner, M., and Schmuland, B. Generalized Mehler semigroups and applications. *Probab. Theory Related Fields*, **105** (2), 193–225 (1996). MR1392452.

Bolthausen, E. and Sznitman, A.-S. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, **197** (2), 247–276 (1998). MR1652734.

Böttcher, B. Feller evolution systems: generators and approximation. *Stoch. Dyn.*, **14** (3), 1350025, 15 (2014). MR3213184.

de Branges, L. The Stone-Weierstrass theorem. *Proc. Amer. Math. Soc.*, **10**, 822–824 (1959). MR113131.

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

Goldschmidt, C. and Martin, J. B. Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.*, **10**, no. 21, 718–745 (2005). MR2164028.

Hénard, O. The fixation line in the Λ-coalescent. *Ann. Appl. Probab.*, **25** (5), 3007–3032 (2015). MR3375893.

Kukla, J. and Möhle, M. On the block counting process and the fixation line of the Bolthausen-Sznitman coalescent. *Stochastic Process. Appl.*, **128** (3), 939–962 (2018). MR3758343.

Limic, V. and Talarczyk, A. Second-order asymptotics for the block counting process in a class of regularly varying Λ-coalescents. *Ann. Probab.*, **43** (3), 1419–1455 (2015). MR3342667.

Möhle, M. 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.

Möhle, M. The rate of convergence of the block counting process of exchangeable coalescents with dust. *ALEA Lat. Am. J. Probab. Math. Stat.*, **18** (2), 1195–1220 (2021). MR4282186.

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

Ranga Rao, R. Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.*, **33**, 659–680 (1962). MR137809.

Revuz, D. and Yor, M. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition (1999). ISBN 3-540-64325-7. MR1725357.

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

Sato, K.-i. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (1999). ISBN 0-521-55302-4. MR1739520.

Sato, K.-i. and Yamazato, M. Operator-self-decomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Process. Appl.*, **17** (1), 73–100 (1984). MR738769.

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.

Shiga, T. A recurrence criterion for Markov processes of Ornstein-Uhlenbeck type. *Probab. Theory Related Fields*, **85** (4), 425–447 (1990). MR1061937.

Siegmund, D. The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probability*, **4** (6), 914–924 (1976). MR431386.

Steutel, F. W. and van Harn, K. *Infinite divisibility of probability distributions on the real line*, volume 259 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York (2004). ISBN 0-8247-0724-9. MR2011862.

Whittaker, E. T. and Watson, G. N. *A course of modern analysis. An introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions*. Cambridge Mathematical Library. Cambridge University Press, Cambridge (1996). ISBN 0-521-58807-3. MR1424469.