On Some Scaling Limits for Branching and Coalescent Processes

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Benedict Vetter
aus Reutlingen

Tübingen
2022

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	19.04.2023
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Martin Möhle
2. Berichterstatter:	Prof. Dr. Martin Zerner

Contents

Notation and symbols

Sets and numbers

$\mathbb{N}$	Set of all positive integers
$\mathbb{N}_0$	Set of all non-negative integers
$\mathbb{N}^*$	$\mathbb{N} \cup \{\infty\}$
$\mathbb{Z}$	Set of all integers
$\mathbb{R}$	Set of all real numbers
$\|A\|$	Cardinality of the set A
$[n]$	The set $\{1, \ldots, n\}$ for $n \in \mathbb{N}$
1_A	Indicator function of the set A
$f(x) = o(g(x))$	Two functions f and g satisfy $\lim_{x\to\infty} f(x)/g(x) = 0$
$f(x) \sim g(x)$ as $x \to \infty$	Two functions f and g satisfy $\lim_{x\to\infty} f(x)/g(x) = 1$
δ_{ij}	Kronecker symbol
$a \wedge b$	The minimum of $a, b \in \mathbb{R}$

Partitions

$\mathcal{P}_A$	Set of all partitions of the set A
$\mathcal{P}_\infty$	Set of all partitions of $\mathbb{N}$
$\mathcal{P}_n$	Set of all partitions of $[n]$ for $n \in \mathbb{N}$
$\|\pi\|$	Number of blocks of the partition $\pi \in \mathcal{P}_A$

Probability Theory

$\mathbb{P}(A)$	Probability of the event A
$\mathbb{P}(A \mid B)$	Conditional probability of A given B
$\mathbb{E}(X)$	Mean of the random variable X
$\mathrm{Var}(X)$	Variance of the random variable X
$\mathrm{Cov}(X, Y)$	Covariance of the random variables X and Y
ε_a	Dirac measure at a
$N(\mu, \sigma^2)$	Normal distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \geq 0$ ($N(\mu, 0) = \varepsilon_\mu$)
λ	Lebesgue measure

Zusammenfassung

In dieser Dissertation werden zwei bekannte Prozesse aus dem Gebiet der mathematischen Populationsgenetik untersucht. Dabei handelt es sich zum einen um eine Klasse von Verzweigungsprozessen und zum anderen um partitionswertige, austauschbare Coalescents.

Die Dissertation basiert auf drei wissenschaftlichen Arbeiten, die hier in zeitlicher Reihenfolge Article I, II und III genannt werden. Article I behandelt die Verzweigungsprozesse und Article II und III behandeln die Coalescents.

Coalescents beschreiben die Genealogie von Populationen und lassen sich leicht am Beispiel des Kingman-Coalescent erklären, der als erstes untersucht wurde. Man nehme eine Population bestehend aus den Individuen $1, \dots, n$ und fasse jedes Individuum i als den einelementigen Block $\{i\}$ auf. Je zwei Blöcke $\{i\}, \{j\}$ verschmelzen nun nach einer exponentialverteilten Zeit zum Block $\{i, j\}$, unabhängig von anderen Paaren von Blöcken. Nach dem ersten Verschmelzen befindet sich der Coalescent in einem Zustand bestehend aus $n - 1$ Blöcken, nämlich einem zweielementigen Block und $n - 2$ einelementigen Blöcken. Erneut verschmelzen je zwei Blöcke nach einer exponentialverteilten Zeit. Mit fortlaufender Zeit verschmelzen jeweils zwei Blöcke bis der Prozess in der Partition endet, die nur aus dem Block $\{1, \dots, n\}$ besteht. Dieser Prozess lässt sich verallgemeinern, indem erlaubt wird, dass mehr als zwei Blöcke zu einem verschmelzen und dass mehrere Verschmelzungen zeitgleich stattfinden dürfen. Als Voraussetzung fordern wir nur ein Neutralität in der Form der Annahme, dass der Prozess austauschbar ist.

Wir beweisen einen Grenzwertsatz für die Anzahl der Blöcke im Coalescent, der aus der Literatur bekannte Resultate erweitert. Sei $N_t^{(n)}$ die Anzahl der Blöcke zur Zeit $t \geq 0$ in einem Coalescent, der mit n Individuen startet. Wir nennen $(N_t^{(n)})_{t\geq 0}$ den *block counting process*. In Article II und III suchen wir nach Bedingungen, welche die Existenz einer Folge $(v(n,t))_{n\in\mathbb{N}}$ garantieren für die $N_t^{(n)}/v(n,t)$ für $n \to \infty$ konvergiert. Die zentralen Resul-

tate von Article II sind in Article III enthalten. Beide Arbeiten zeigen die Konvergenz des skalierten block counting process im Skorohod Raum der càdlàg Pfade. Article II liefert Resultate für eine Klasse von Λ-Coalescents, deren wichtigster Vertreter als zugrunde liegendes Maß die Beta-Verteilung $\beta(1, b)$ mit Parametern 1 und $b > 0$ besitzt. Article III befasst sich mit der größeren Klasse von Ξ-Coalescents. Neben dieser Verallgemeinerung sind auch die Voraussetzungen des zentralen Konvergenzsatzes schwächer. Die Beweise bestehen aus dem Nachweis der uniformen Konvergenz der zugehörigen Generatoren. Der Grenzprozess ist ein verallgemeinerter Ornstein-Uhlenbeck-Prozess. Für die sogenannte *fixation line* werden analoge Ergebnisse gezeigt.

Der in Article I betrachtete Verzweigungsprozess ist ein Populationsmodell in dem jedes Individuum eine exponentialverteilte Zeit lebt, unabhängig von anderen Individuen, und im Moment seines Todes eine zufällige Zahl an Nachkommen bekommt, die wiederum unabhängig von den Lebensdauern und den Nachkommenszahlen anderer Individuen ist. Die Verteilung der Nachkommenszahl ist für alle Individuen identisch. Sei $Z_t^{(n)}$ die Größe der Population zur Zeit $t \geq 0$, wenn die Population zu Beginn die Größe n hat. Wir untersuchen die Frage wann es Folgen $(a(n,t))_{n\in\mathbb{N}}$ und $(b(n,t))_{n\in\mathbb{N}}$ derart gibt, dass

$$\frac{Z_t^{(n)} - b(n,t)}{a(n,t)}$$

für $n \to \infty$ konvergiert. Aufgrund der Verzweigungseigenschaft lässt sich diese Frage unter Verwendung des zentralen Grenzwertsatzes oder der Theorie über stabile Verteilungen beantworten. Darüber hinaus beweisen wir die Konvergenz im Skorohod Raum und stellen Zusammenhänge zwischen den notwendigen und hinreichenden Bedingungen für die Existenz der normalisierenden Folgen $(a(n,t))_{n\in\mathbb{N}}$ und $(b(n,t))_{n\in\mathbb{N}}$ für alle $t \geq 0$ und der Nachkommensverteilung des Verzweigungsprozesses dar.

Die Ergebnisse für Coalescents und Verzweigungsprozesse haben einen Schnittpunkt. Der Bolthausen-Sznitman-Coalescent wird von den Resultaten aus Article II und III abgedeckt. Gleichzeitig ist die fixation line im Bolthausen-Sznitman-Coalescent ein Verzweigungsprozess, der den Voraussetzungen aus Article I genügt. Theorem II.2.4 in Article II besagt in diesem Fall dasselbe wie Theorem I.2.8 in Article I in logarithmischer Form. Diese Verbindung ist bekannt, ebenso wie der zugehörige Konvergenzsatz [19]. Der Versuch diese Resultate zu verallgemeinern, einerseits auf der Seite der Coalescents, andererseits auf der Seite der Verzweigungsprozesse, war der Ursprung der in dieser Dissertation behandelten Fragestellungen.

Summary

In this dissertation two known processes from the field of mathematical population genetics are treated. The two processes are a class of branching processes and partition-valued, exchangeable coalescents.

The dissertation is based on three scientific articles, which are called Article I, II and III here. Article I treats the branching processes and Article II and III treat the coalescents.

Coalescents describe the genealogy of populations. The Kingman coalescent is the coalescent studied first and a fitting example that can be used in order to explain more general coalescent processes. Take a population that consists of the individuals $1, \dots, n$ and consider each individual i as the singleton block $\{i\}$. Every pair of blocks $\{i\}, \{j\}$ merges to be the block $\{i, j\}$ after an exponentially distributed time, independently of other pairs of blocks. After the first merger, the coalescent is in a state that consists of one block of size 2 and $n-2$ singleton blocks. Again each pair of blocks merges into one after an exponentially distributed time. This procedure repeats until the coalescent reaches the absorbing state that consists only of the block $\{1, \dots, n\}$. The process can be generalized by allowing multiple blocks to merge into one and more than one merger to occur at the same time. The only requirement is a neutrality assumption in the form that the process needs to be exchangeable.

We provide scaling limits for the number of blocks in a coalescent, which extends known results from the literature. Let $N_t^{(n)}$ be the number of blocks after time $t \geq 0$ in a coalescent that started with n individuals. We call $(N_t^{(n)})_{t\geq 0}$ the *block counting process*. In Article II and III, we search for conditions that guarantee the existence of a sequence $(v(n,t))_{n\in\mathbb{N}}$ for which $N_t^{(n)}/v(n,t)$ converges as $n \to \infty$ for all $t \geq 0$. The main results of Article II are contained in Article III. Both works show the convergence of the scaled block counting process in the space of càdlàg paths endowed with the Skorohod topology. Article II covers a class of Λ-coalescents, whose most important member has the underlying measure $\Lambda = \beta(1, b)$, the beta

distribution with parameters 1 and $b > 0$. Article III treats the larger class of Ξ-coalescents. Moreover, the assumptions of the main result in Article III are less strict. The main results are proved by showing the uniform convergence of the corresponding generators. The limiting process is an Ornstein–Uhlenbeck type process. Analogous results are stated for the so-called *fixation line*.

The branching process treated in Article I is a population model in which each individual lives for an exponentially distributed time, independently of other individuals, and at the moment of its death the individual produces a random number of offspring, which is independent of the lifetimes and the offspring numbers of other individuals. Let $Z_t^{(n)}$ denote the size of the population after time $t \geq 0$ if the population started from n individuals at time $t = 0$. We investigate the question of the existence of sequences $(a(n,t))_{n\in\mathbb{N}}$ and $(b(n,t))_{n\in\mathbb{N}}$ for which

$$\frac{Z_t^{(n)} - b(n,t)}{a(n,t)}$$

converges as $n \to \infty$. Due to the branching property, the question can be answered by utilizing the central limit theorem or the theory about stable distributions. We additionally prove the convergence in the Skorohod space und show relations between the necessary and sufficient conditions for the existence of normalizing sequences $(a(n,t))_{n\in\mathbb{N}}$ and $(b(n,t))_{n\in\mathbb{N}}$ for all $t \geq 0$ and the offspring distribution of the branching process.

The results for branching and coalescent processes overlap in one point. The Bolthausen–Sznitman coalescent is covered by the results of Article II and III. The fixation line in the Bolthausen–Sznitman coalescent is also a branching processes that satisfies the assumptions of Article I. In this case Theorem II.2.4 in Article II states the same as Theorem I.2.8 in Article I, although in logarithmic form. This relation has been known, just as the respective convergence results [19]. The attempt to generalize this convergence result, on the side of branching processes on one hand and on the side of coalescents on the other has been the origin of the questions pursued in this thesis.

List of Manuscripts

This thesis is based on the following three articles.

Article I Möhle, M. and Vetter, B. (2021) Asymptotics of continuous-time discrete state space branching processes for large initial state. *Markov Process. Related Fields* **27**, no. 1, 1–24.

Article II Möhle, M. and Vetter, B. (2022) Scaling limits for the block counting process and the fixation line for a class of Λ-coalescents. *ALEA Lat. Am. J. Probab. Math. Stat.* **19**, no. 1, 641–664.

Article III Möhle, M. and Vetter, B. (2023) Scaling limits for a class of regular Ξ-coalescents. *Stochastic Process. Appl.* **162**, 387–422.

Chapter 1

Introduction

In this chapter, the two stochastic models are introduced in whose area of research the three articles on which this thesis is based lie. First some standard vocabulary and notation on random partitions is fixed. Then the definitions of exchangeable coalescents, the block counting process and the properties "coming down from infinity" and "dust" are given and fundamental results are listed. The known, explicit construction of exchangeable coalescents by use of a Poisson point process is included. The function γ that, roughly speaking, describes the expected size of a jump of the block counting process is assigned to its own section and is repeatedly remarked on throughout this chapter due to its importance in Article III. The fixation line is introduced as the Siegmund-dual of the block counting process. At the end of this chapter, Markov branching processes are covered. We try to be concise and focus on the definitions and notions that are needed in order to understand the content of Chapter 2.

1.1 Exchangeable random partitions

Let A be a set. A *partition* π of A is an equivalence relation on A. The set A is partitioned into non-empty disjoint subsets, namely the equivalence classes of π, whose union equals A. The equivalence classes of this relation, called *blocks*, are used to determine a partition in the following. Let $\mathcal{P}_A$ denote the set of partitions of A. We write $|\pi|$ for the number of blocks of $\pi \in \mathcal{P}_A$ and $|B|$ for the size of some block B of π. A block B of size $|B| = 1$ is called *singleton*. We can generate new partitions by exchanging elements of the blocks or by restricting $\pi \in \mathcal{P}_A$ to a subset $A' \subseteq A$. For a permutation σ of A, the blocks of $\sigma\pi \in \mathcal{P}_A$ are defined to be all the sets $\sigma B := \{\sigma i : i \in B\}$, where B is a block of π. The restriction $\pi|_{A'} \in \mathcal{P}_{A'}$ of π to A' comprises all non-empty intersections $B \cap A'$ with blocks B of π.

For the sets $\mathbb{N} := \{1, 2, \ldots\}$ and $[n] := \{1, \ldots, n\}$ with $n \in \mathbb{N}$ we use the notation $\mathcal{P}_\infty := \mathcal{P}_\mathbb{N}$, $\mathcal{P}_n := \mathcal{P}_{[n]}$ and, in addition, $[\infty] := \mathbb{N}$. For $\pi \in \mathcal{P}_m$ with $m \in \mathbb{N}^* := \mathbb{N} \cup \{\infty\}$ and $n \in \mathbb{N}^*$ with $n \leq m$ we also define $\pi^{(n)} := \pi|_{[n]}$. Listing the blocks in the increasing order of their smallest element provides an unique representation of the partition $\pi \in \mathcal{P}_n$. We write $\pi = \{B_1, B_2, \ldots\}$ when π consists of the blocks $B_1, B_2, \ldots$ and $\min B_i \leq \min B_j$ for all $i \leq j$, where the set stands for $\{B_1, \ldots, B_{|\pi|}\}$ if $|\pi| < \infty$ and for the infinitely large set if $|\pi| = \infty$.

The following definition is going to enable to compactly describe the transition between states in coalescent processes. Let $n \in \mathbb{N}^*$, $\pi = \{B_1, B_2, \ldots\} \in \mathcal{P}_n$ and $\pi' = \{B'_1, B'_2, \ldots\} \in \mathcal{P}_{|\pi|}$. Then the *coagulation* of π by π' is defined as

$$\operatorname{coag}(\pi, \pi') := \left\{ \bigcup_{j \in B'_1} B_j, \bigcup_{j \in B'_2} B_j, \ldots \right\} \in \mathcal{P}_n.$$

The definition can be extended by admitting $\pi' \in \mathcal{P}_m$ with $m > |\pi|$ and setting $\operatorname{coag}(\pi, \pi') := \operatorname{coag}(\pi, (\pi')^{(|\pi|)})$. During the coagulation the blocks of π merge. Thus, π is a refinement of $\operatorname{coag}(\pi, \pi')$, meaning that each block of π is a subset of some block of $\operatorname{coag}(\pi, \pi')$. On the other hand, if π is a refinement of some partition π'', then there exists π' such that $\pi'' = \operatorname{coag}(\pi, \pi')$, and π' becomes unique when the partition is restricted to $[|\pi|]$. For example, if $n = 10$, $\pi = \{\{1, 3, 5\}, \{2, 10\}, \{4\}, \{6\}, \{7\}, \{8\}, \{9\}\}$ and $\pi' = \{\{1, 2\}, \{5, 6, 7\}, \{3\}, \{4\}\}$, then $\operatorname{coag}(\pi, \pi') = \{\{1, 2, 3, 5, 10\}, \{4\}, \{6\}, \{7, 8, 9\}\}$.

The space $\mathcal{P}_\infty$, endowed with the metric

$$d(\pi, \pi') := \left(\sup \left\{ k \in \mathbb{N} : \pi^{(k)} = (\pi')^{(k)} \right\} \right)^{-1}, \qquad \pi, \pi' \in \mathcal{P}_\infty,$$

(and the convention $\infty^{-1} = 0$) is Polish. Due to the metric, we are able to equip $\mathcal{P}_\infty$ with the Borel-σ-algebra and can define a *random partition* of $\mathbb{N}$ as a $\mathcal{P}_\infty$-valued random variable. The σ-algebra is generated by the countably many open balls $\{\pi \in \mathcal{P}_\infty : \pi^{(n)} = \pi'\}$ with $n \in \mathbb{N}$ and $\pi' \in \mathcal{P}_n$. For $n \in \mathbb{N}$ a random partition of $[n]$ is a $\mathcal{P}_n$-valued random variable, where the σ-algebra on the finite set $\mathcal{P}_n$ is the power set of $\mathcal{P}_n$. It is straightforward to verify that the σ-algebra on $\mathcal{P}_\infty$ is generated by the restriction maps $\mathcal{P}_\infty \to \mathcal{P}_n, \pi \mapsto \pi^{(n)}$, with $n \in \mathbb{N}$.

The class of random partitions introduced next are the ones whose distributions do not change when elements are swapped between blocks.

Definition 1.1.1. Let $n \in \mathbb{N}$. A random partition Π of $[n]$ is said to be *exchangeable* if Π and $\sigma\Pi$ have the same distribution for all permutations σ of $[n]$. A random partition Π of $\mathbb{N}$ is said to be *exchangeable* if the restriction $\Pi^{(n)}$ is an exchangeable random partition of $[n]$ for all $n \in \mathbb{N}$.

1.2 Exchangeable coalescents

Coalescents are partition-valued processes where blocks merge over time. The below given, precise definition of the object with which we work in Articles II and III goes back to Bertoin and Le Gall [5].

Definition 1.2.1. An *exchangeable coalescent* is a process $\Pi = (\Pi_t)_{t\geq 0}$ with state space $\mathcal{P}_\infty$ and càdlàg paths that satisfies the following properties:

(i) $\Pi_0 = \{\{1\}, \{2\}, \ldots\}$.

(ii) Π is a time-homogeneous Markov process and, for all $t \geq 0$, there exists an exchangeable random partition π_t of $\mathbb{N}$ such that, for all $s \geq 0$, the law of Π_{s+t} conditional on Π_s is the law of the coagulation of Π_s by π_t.

Remark. Coagulating the partition $\{\{1\}, \{2\}, \ldots\} \in \mathcal{P}_\infty$ by π_t results in π_t itself. Choosing $s = 0$ in Definition 1.2.1 (ii) hence shows that π_t and Π_t have the same distribution for all $t \geq 0$. In particular, Π_t is exchangeable for all $t \geq 0$.

Remark. Coalescents satisfying Definition 1.2.1 (i) are called *standard* in the literature. We restrict our attention to standard exchangeable coalescents and omit the additional term, since any non-standard exchangeable coalescent can be easily deduced from a standard one. For example, if $\Pi = (\Pi_t)_{t\geq 0}$ is a standard exchangeable coalescent and $\pi \in \mathcal{P}_\infty$, then $(\text{coag}(\pi, \Pi_t))_{t\geq 0}$ is a $\mathcal{P}_\infty$-valued Markov process with càdlàg paths, intial state π and the same transition probabilities as Π.

Let $\Pi = (\Pi_t)_{t\geq 0}$ be an exchangeable coalescent. For $n \in \mathbb{N}$ the restriction $\Pi^{(n)} := (\Pi_t^{(n)})_{t\geq 0}$ to $[n]$ has finite state space $\mathcal{P}_n$. We call $\Pi^{(n)}$ exchangeable coalescent too or restricted coalescent or n-coalescent if we want to put emphasis on the "size". Applying the fact that the σ-algebra on $\mathcal{P}_\infty$ is generated by the family of open balls $\{\pi \in \mathcal{P}_\infty : \pi^{(n)} = \pi_n\}$ with $\pi_n \in \mathcal{P}_n$ and $n \in \mathbb{N}$, which is also closed upon intersecting, it can be shown that instead of Definition 1.2.1 (ii) it can equivalently be required that, for all $n \in \mathbb{N}$, $\Pi^{(n)}$ is a Markov chain with transition probabilities

$$\mathbb{P}\big(\Pi^{(n)}_{s+t} = \pi' \,\big|\, \Pi^{(n)}_s = \pi\big) \;=\; \mathbb{P}\big(\mathrm{coag}(\pi, \pi_t) = \pi'\big), \qquad (1.1)$$

$s, t \geq 0, \pi, \pi' \in \mathcal{P}_n, n \in \mathbb{N}$, where π_t is the same random variable as in Definition 1.2.1. The path property from Definition 1.2.1 transfers too. As an immediate consequence of the definition of the metric d on $\mathcal{P}_\infty$, Π has càdlàg paths if and only if each restriction $\Pi^{(n)}$ has càdlàg paths.

The class of exchangeable coalescents was independently introduced by Schweinsberg [30] and by Möhle and Sagitov [23]. Möhle and Sagitov obtained exchangeable coalescents as limits of ancestral processes in a population model in discrete time with a constant, finite size and an exchangeable reproduction law (Cannings model [8, 9]). Schweinsberg labeled these processes "coalescents with simultaneous multiple collisions", refering to the fact that multiple blocks merge into a single one at certain times and that such mergers are allowed to occur at the same time. For $j, k_1, \ldots, k_j \in \mathbb{N}$ with $k_1 \geq k_2 \geq \ldots \geq k_j$ and $k_1 \geq 2$ we call a merging event during which the n-coalescent $\Pi^{(n)}$ moves from state $\pi \in \mathcal{P}_n$ to the state $\mathrm{coag}(\pi, \pi_t)$ a $(k_1, \ldots, k_j)$-collision if $|\pi| = k := k_1 + \ldots + k_j$ and $\pi_t^{(k)}$ consists of j blocks of sizes $k_1, \ldots, k_j$. Clearly, the assumption that π_t be exchangeable in (1.1) implies that the transition probabilities from π to the outcome of a $(k_1, \ldots, k_j)$-collision only depend on $k_1, \ldots, k_j$ and the time that has passed, and not on the sizes of the blocks of π or on the integers that are contained in the blocks. Thus, the rate at which a $(k_1, \ldots, k_j)$-collision occurs is equal to some fixed value, denoted by $\phi_j(k_1, \ldots, k_j)$, which also does not depend on n. In [30], coalescents with simultaneous multiple collisions are defined in the same way as exchangeable coalescents but Definition 1.2.1 (ii) is replaced by

(ii)' For all $n \in \mathbb{N}$, $\Pi^{(n)}$ is a Markov chain for which any possible $(k_1, \ldots, k_j)$-collision occurs at a rate equal to some fixed value $\phi_j(k_1, \ldots, k_j)$.

In difference to [30] we do not treat singleton blocks of π_t separately and hence use a differing, adjusted notation. It is known that

Schweinsberg's definition of coalescents with simultaneous multiple collisions coincides with Definition 1.2.1. In fact, the argument just sketched shows that Definition 1.2.1 implies (ii)'. It is straightforward to prove that (ii)' in turn implies (1.1) and, consequently, Definition 1.2.1 (ii), where the random variable π_t has the same distribution as Π_t.

An advantage of Schweinsberg's definition is the fact that explicit formulas for the rates $\phi_j(k_1, \ldots, k_j)$ are known. Schweinsberg proved an one-to-one correspondence between exchangeable coalescents and finite measures Ξ on the infinite simplex $\Delta := \{(u_1, u_2, \ldots) : u_1 \geq u_2 \geq \ldots, \sum_{i\geq 1} u_i \leq 1\}$, see [30, Theorem 2]. For this reason exchangeable coalescents are also referred to as Ξ-coalescents.

Let Ξ be a finite measure on Δ and decompose Ξ in the form $\Xi = a\varepsilon_0 + \Xi_0$, where ε_0 denotes the Dirac measure at $0 := (0, 0, \ldots) \in \Delta$ and $a := \Xi(\{0\})$. Moreover, we use the notation $|u| := \sum_{i\geq 1} u_i$ and $(u, u) := \sum_{i\geq 1} u_i^2$ for $u \in \Delta$ and define the measure ν via $\nu(\mathrm{d}u) := \Xi_0(\mathrm{d}u)/(u, u)$. Then there exists an exchangeable coalescent such that, for all $j, k_1, \ldots, k_j \in \mathbb{N}$ with $k_1 \geq 2$, $k_1 \geq k_2 \geq \ldots \geq k_j$, the rate $\phi_j(k_1, \ldots, k_j)$ equals

$$a1_{\{r=1,k_1=2\}} + \int_\Delta \sum_{l=0}^{s} \binom{s}{l} (1-|u|)^{s-l} \sum_{i_1\neq \ldots \neq i_{r+l}} u_{i_1}^{k_1} \cdots u_{i_{r+l}}^{k_{r+l}} \, \nu(\mathrm{d}u), \quad (1.2)$$

where $s := |\{i \in [j] : k_i = 1\}|$ and $r := j - s$. In Section 1.3, a possible construction of an exchangeable coalescent whose rates are given by (1.2) is presented for every finite measure Ξ on Δ. Conversely, for every exchangeable coalescent there exists a finite measure Ξ on Δ such that the rates are given by (1.2).

Shortly before the class of Ξ-coalescents was introduced, Pitman [26] and Sagitov [27] independently defined a class of coalescents whose merging events only allow for multiple blocks to merge into a single one but not more than one merger to occur at the same time. These processes, called "coalescent with multiple collisions", entail an one-to-one correspondence with finite measures Λ on $[0, 1]$ and are hence also called Λ-coalescents. We identify a Λ-coalescent with the exchangeable coalescent whose underlying measure Ξ is defined via $\Xi(B \times \{0\} \times \{0\} \times \ldots) := \Lambda(B)$ for Borel-measurable $B \subseteq [0, 1]$ and $\Xi(\{u \in \Delta : u_2 > 0\}) := 0$. In particular, Λ-coalescents constitute a subclass of exchangeable coalescents. For Λ-coalescents the rates (1.2) simplify notably. If the Λ-coalescent is in a state with k blocks, any j blocks merge at the rate

$$\lambda_{k,j} := \phi_{k-j+1}(j,1,\ldots,1) = a1_{\{j=2\}} + \int_{[0,1]} u^{j-2}(1-u)^{k-j}\,\Lambda(\mathrm{d}u),$$

$k \in \mathbb{N}, 1 \leq j < k$.

In order to give a complete account of the development of exchangeable coalescents we want to remark that two famous Λ-coalescents have been studied before the appearance of Pitman's and Sagitov's works: the Kingman coalescent [18] whose underlying measure $\Lambda = \varepsilon_0$ is the Dirac measure at $0 \in [0,1]$ with rates $\lambda_{k,j} = 1_{\{j=k-1\}}$ and the Bolthausen–Sznitman coalescent [7] whose underlying measure $\Lambda = \lambda$ is Lebesgue measure on $[0,1]$ with rates $\lambda_{k,j} = (j-2)!(k-j)!/(k-1)!$ for $1 \leq j < k$.

To conclude this section, we want to clarify the terminology we use. In the following we speak of exchangeable coalescents as coalescents. If the underlying measure is important, we add it as prefix. We write "Λ-coalescent" when statements concern only the class of coalescents with multiple collisions, and we write "Ξ-coalescent" if we want to stress that we are in the more general setting.

1.3 Poisson point process construction

In this section, we give an outline of the Poisson point process construction of coalescents going back to Schweinsberg; for details we refer the reader to [30]. The construction provides an explanatory, probabilistic view on the rates $\phi_j(k_1,\ldots,k_j)$, defined via (1.2), in terms of an urn model.

Let Ξ be a finite, non-zero measure on Δ. For $u = (u_1, u_2, \ldots) \in \Delta$ define $u_0 := 1 - |u|$ and "urns" as the tilings $J_0 := [0, u_0), J_1 := [u_0, u_0 + u_1), J_2 := [u_0 + u_1, u_1 + u_2 + u_3)$ of lengths $u_0, u_1, u_2, \ldots$ of the interval $[0,1)$. Let the "balls" be independent and identically distributed (i.i.d.) random variables $U_1, U_2, \ldots$ that have an uniform distribution on $[0,1)$. For $i \in \mathbb{N}$ and $u \in \Delta$ define the random variable $Z_i(u)$ by setting $Z_i(u) := j$ for $j \geq 1$ if the i-th ball lands in the urn J_j, i.e., $U_i \in J_j$ and $Z_i(u) := -i$ if $U_i \in J_0$. For every $u \in \Delta$, $Z_1(u), Z_2(u), \ldots$ are independent random variables with $\mathbb{P}(Z_i(u) = j) = u_j$ and $\mathbb{P}(Z_i(u) = -i) = u_0$ for $j \geq 1$ and $i \in \mathbb{N}$. We equip $\mathbb{Z}^\infty$ with the product topology and denote the distribution of the $\mathbb{Z}^\infty$-valued random variable $(Z_1(u), Z_2(u), \ldots)$ by $\mathbb{P}_u$. Let z_{ij} be the sequence $(z_1, z_2, \ldots) \in \mathbb{Z}^\infty$ such that $z_i = z_j = 1$ and $z_k = -k$ for

$k \notin \{i, j\}$. Define the measure L on $\mathbb{Z}^\infty$ via

$$L(A) := a \sum_{1 \le i < j} 1_{\{z_{ij} \in A\}} + \int_\Delta \mathbb{P}_u(A)\, \nu(\mathrm{d}u) \qquad (1.3)$$

for all measurable $A \subseteq \mathbb{Z}^\infty$. Define $A_n := \{(z_1, z_2, \ldots) \in \mathbb{Z}^\infty : z_1, \ldots, z_n \text{ not all distinct}\}$ for $n \ge 2$ and $A_\infty := \cup_{n \ge 2} A_n = \{(z_1, z_2, \ldots) \in \mathbb{Z}^\infty : z_i = z_j \text{ for some } i \ne j\}$. Then $L(A_\infty^c) = \int_\Delta \mathbb{P}(Z_i(u) \ne Z_j(u)\ \forall\, i \ne j)\, \nu(\mathrm{d}u) = 0$ and it can be shown that $L(A_n) < \infty$ for $n \ge 2$. As the union of A_n for all $n \ge 2$ and A_∞^c equals $\mathbb{Z}^\infty$, L is σ-finite and, hence, a proper Poisson point process ξ on $[0, \infty) \times \mathbb{Z}^\infty$ with intensity measure $\lambda \otimes L$ exists, i.e., there exist $([0, \infty) \times \mathbb{Z}^\infty)$-valued random variables $(T_1, \xi_1), (T_2, \xi_2), \ldots$ and a $\mathbb{N}^*$-valued random variable K such that $\xi = \sum_{i=1}^K \varepsilon_{(T_i, \xi_i)}$. In the following we refer to any pair (T_i, ξ_i) as a point of the Poisson point process ξ and denote a generic point by (t, z).

Now we define a family of processes $\Pi_n := (\Pi_{n,t})_{t \ge 0}$ with state spaces $\mathcal{P}_n$ and rates $\phi_j(k_1, \ldots, k_j)$, given by (1.2). Fix $n \in \mathbb{N}$. Let $\Pi_{n,0}$ be the partition of $[n]$ into singletons. For any $T > 0$ the Poisson point process ξ has only finitely many points (t, z) with $z \in A_n$ and $0 \le t \le T$ and none of them share the same first entry almost surely. In particular, we can order all points (t, z) with $z \in A_n$ by their first entry. Let Π_n have càdlàg paths in the finite state space $\mathcal{P}_n$ with jumps only at times $t \in [0, \infty)$, where t is the first entry in a point (t, z) of ξ with $z \in A_n$. The second entry z determines the outcome of the jump corresponding to (t, z). Regard $z \in \mathbb{Z}^\infty$ as the function $i \mapsto z_i$, $i \in \mathbb{N}$. Ignoring the empty sets, the inverse images $z^{-1}(j) := \{i \in \mathbb{N} : z_i = j\}$, $j \in \mathbb{Z}$, define a partition of $\mathbb{N}$, say $\pi(z)$. During the jump corresponding to (t, z), let Π_n move from its current state $\Pi_{n,t-}$ to $\Pi_{n,t} := \mathrm{coag}(\Pi_{n,t-}, \pi(z))$. Given $\Pi_{n,t-} = \{B_1, \ldots, B_k\}$, each block of the resulting partition $\Pi_{n,t}$ is an union of blocks of $\Pi_{n,t-}$ such that B_i and B_j are contained in the same block of $\Pi_{n,t}$ if and only if $z_i = z_j$. Note that if $z_i < 0$, then B_i does not merge with any other block.

From the definition of Poisson point processes it follows that Π_n is a time-homogeneous Markov chain that jumps from $\pi \in \mathcal{P}_n$ to $\pi' \in \mathcal{P}_n$ at the rate $L(A_{\pi,\pi'})$, where $A_{\pi,\pi'} := \{z \in \mathbb{Z}^\infty : \mathrm{coag}(\pi, \pi(z)) = \pi'\}$. Suppose that $\pi = \{B_1, \ldots, B_k\}$. As seen above, $A_{\pi,\pi'}$ is the set of $z \in \mathbb{Z}^\infty$ for which $z_i = z_j$ if and only if B_i and B_j are contained in the same block of π' for all $i, j \in [k]$. We can describe the probability $\mathbb{P}_u(A_{\pi,\pi'})$ appearing in (1.3) in terms of the urn model: $\mathbb{P}_u(A_{\pi,\pi'})$ is the probability of the event that if $1 \leq i \neq j \leq k$, then the balls U_i and U_j are in the same urn but not in urn J_0, i.e., $U_i, U_j \in J_m$ for some $m \geq 1$ if and only if B_i and B_j are contained in the same block of π'. Bearing that in mind, it is straightforward to see that $\mathbb{P}_u(A_{\pi,\pi'})$ is equal to the integrand in (1.2), provided that π' is the result of a coagulation of π by a partition of $[k]$, whose blocks have sizes $k_1, \ldots, k_j$. If only two blocks of π merge into one, i.e., $k_1 = 2$ and $j = k - 1$, then the value a is added to the integral in (1.3). Consequently, the rates of Π_n are given by (1.2).

The family of processes Π_n, $n \in \mathbb{N}$, is consistent in the sense that if $n < m$, then Π_n and the restriction $\Pi_m^{(n)} := (\Pi_{m,t}^{(n)})_{t \geq 0}$ of Π_m to $[n]$ are the same process. In order to see this note that every jump of Π_n, which corresponds to some point (t, z) with $z \in A_n$, results in a jump of Π_m, since $A_n \subseteq A_m$, and the coagulating mechanism is the same when only the integers from $[n]$ are taken into account. Also, a jump of Π_m corresponding to a point (t, z) with $z \in A_m \setminus A_n$, i.e., a point which is not involved in the construction of Π_n does not inflict the integers from $[n]$, which implies that $\Pi_m^{(n)}$ does not change during this jump.

Thus, the $\mathcal{P}_\infty$-valued process $\Pi = (\Pi_t)_{t \geq 0}$ is well-defined by saying that, for all $t \geq 0$, Π_t is the partition $\pi \in \mathcal{P}_\infty$ such that $\pi^{(n)} = \Pi_{n,t}$ for all $n \in \mathbb{N}$. Clearly, $\Pi_t^{(n)} = \Pi_{n,t}$ for all $n \in \mathbb{N}$ and $t \geq 0$. The rates of $\Pi^{(n)} := (\Pi_t^{(n)})_{t \geq 0}$ are also given by (1.2). In particular, $\Pi^{(n)}$ satisfies Condition (ii)' from Section 1.2. By construction, Π_0 only consists of singletons. Moreover, Π has càdlàg paths, since it is equivalent to demand that $\Pi^{(n)}$ has càdlàg paths for all $n \in \mathbb{N}$. Thus, as shown in the previous section, Π is an exchangeable coalescent according to Definition 1.2.1.

If Λ is a finite measure on $[0, 1]$, which we identify with a certain measure Ξ on Δ, we only have the two urns J_0 and J_1. For a point (t, z) of ξ it holds that $z_i = -i$ or $z_i = 1$ for all $i \in \mathbb{N}$. During a jump of the restricted Λ-coalescent $\Pi^{(n)}$ corresponding to the point (t, z), the blocks B_i of $\Pi_{t-}^{(n)}$ for which $z_i = 1$ merge into a single block and all other blocks remain unchanged. Hence, collisions of multiple blocks are allowed but cannot occur simultaneously. Given that $\Pi^{(n)}$ contains the blocks $B_1, \ldots, B_k$ at time $t \geq 0$, any fixed choice $B_{i_1}, \ldots, B_{i_j}$ of blocks merges into a single one at the rate $\phi_{k-j+1}(j, 1, \ldots, 1)$. The integrand in (1.2) is the probability of the event that, for all $1 \leq i \leq k$, the i-th ball lands in urn J_1 if $i \in$

$\{i_1, \ldots, i_j\}$ and in the other urn if $i \notin \{i_1, \ldots, i_j\}$. This event can also be described as follows: consider a coin that shows heads with probabiltiy $u \in (0, 1]$, flip the coin for each block and merge all blocks for which the coin toss shows heads. Then the integrand in (1.2) is the probability of the event that exactly the coins flipped for the blocks $B_{i_1}, \ldots, B_{i_j}$ show heads.

1.4 The block counting process

Let $N_t := |\Pi_t|$ and $N_t^{(n)} := |\Pi_t^{(n)}|$ denote the number of blocks in a coalescent $(\Pi_t)_{t\geq 0}$ and in its restriction to $[n]$ for $n \in \mathbb{N}$ after time $t \geq 0$, respectively.

Definition 1.4.1. In an abuse of notation, we call the processes $N := (N_t)_{t\geq 0}$ and $N^{(n)} := (N_t^{(n)})_{t\geq 0}$ *block counting process.*

The block counting process has non-increasing càdlàg paths and state space $\mathbb{N}^*$ or $[n]$. Of course, one might ask if N_t is even finite. First we set this question aside and answer it in Section 1.6 as good as it is known in the literature. In this section, we only consider the process $N^{(n)}$ for $n \in \mathbb{N}$. We will see that $N^{(n)}$ is a Markov process. The distribution is uniquely determined by the rates. We aim to represent the rates in a way that makes them accessible to computation. We are especially interested in the limiting behavior of the rates as the values of the states tend to infinity. To this end, we return to the Poisson point process construction.

Fix $n \in \mathbb{N}$. The coalescent $\Pi^{(n)}$ jumps at times $t \geq 0$, where t is the first entry in a point (t, z) of the Poisson point process with intensity measure $\lambda \otimes L$ and $z = (z_1, z_2, \ldots) \in A_n$. The measure L on $\mathbb{Z}^\infty$ is defined via (1.3). During the jump at time t, blocks are coagulated by the partition $\pi(z)$, which is induced by the sets $z^{-1}(j) = \{i \in \mathbb{N} : z_i = j\}$, $j \in \mathbb{Z}$. More precisely, $\Pi^{(n)}$ moves from state $\Pi_{t-}^{(n)}$ to the state $\Pi_t^{(n)} = \text{coag}(\Pi_{t-}^{(n)}, \pi(z))$. Moreover, if $N_{t-}^{(n)} = |\Pi_{t-}^{(n)}| = k$, then

$$N_t^{(n)} = \left|\Pi_t^{(n)}\right| = \left|\pi(z)^{(k)}\right| = |\{z_1, \ldots, z_k\}|,$$

i.e., given that the block counting process $N^{(n)}$ is in the state k before the jump corresponding to (t, z), the state of $N^{(n)}$ after the jump is the number of distinct values assumed by the numbers $z_1, \ldots, z_k$.

Suppose that $\Pi^{(n)}$ is in state $\pi \in \mathcal{P}_n$ and $k := |\pi|$. The set of states that consist of exactly $j < k$ blocks to which $\Pi^{(n)}$ can jump to is given by $\{\text{coag}(\pi, \pi') : \pi' \in \mathcal{P}_k, |\pi'| = j\}$. Recall that $A_{\pi,\text{coag}(\pi,\pi')} =$

$\{z \in \mathbb{Z}^\infty : \pi(z)^{(k)} = \pi'\}$ and define

$$A_{k,j} \;:=\; \bigcup_{\pi' \in \mathcal{P}_k, |\pi'|=j} A_{\pi,\mathrm{coag}(\pi,\pi')} \;=\; \big\{z \in \mathbb{Z}^\infty : |\{z_1,\ldots,z_k\}| = j\big\}$$

as the set of $z \in \mathbb{Z}^\infty$ leading to a transition from π to a partition that consists of j blocks for $1 \le j \le k$. The notation "$A_{k,j}$" is justified, since the set depends on π only via $k = |\pi|$. The rate at which $\Pi^{(n)}$ moves from state π to a partition that consists of j blocks is given by $L(A_{k,j})$. By using the fact that this term only depends on k, it can be shown that $N^{(n)}$ is a Markov chain which, for $1 \le j < k$, jumps from state k to state j at the rate $q_{k,j} := L(A_{k,j})$.

Note that $\mathbb{P}_u(A_{k,j})$ is the probability of the event that the entries of $(Z_1(u),\ldots,Z_k(u))$ assume j distinct values. Any ball in urn J_0 leads to a negative, new value. The number of distinct, positive values is the number of urns other than J_0 to which at least one of the first k balls has been allocated. Thus, it makes sense to define $X_i(k,u) := \sum_{j=1}^k 1_{J_i}(U_j)$ as the number of balls in urn J_i after $k \in \mathbb{N}_0 := \{0,1,\ldots\}$ throws and

$$Y(k,u) \;:=\; X_0(k,u) \;+\; \sum_{i\ge1} 1_{\{X_i(k,u)\ge1\}}, \qquad k \in \mathbb{N}_0, u \in \Delta, \quad (1.4)$$

as the sum of the number of balls in urn J_0 and the number of all other occupied urns. Finally, we obtain $\mathbb{P}_u(A_{k,j}) = \mathbb{P}(Y(k,u) = j)$ for $u \in \Delta$ and, by (1.3),

$$q_{k,j} \;=\; a\binom{k}{2}1_{\{j=k-1\}} \;+\; \int_\Delta \mathbb{P}(Y(k,u)=j)\,\nu(\mathrm{d}u), \qquad (1.5)$$

$1 \le j < k, k \ge 2$. The representation (1.5) for the rates of the block counting process plays a major role in the proofs of the main convergence results in Article III. A comparable representation for the rates for Λ-coalescents is defined in Article II.

The asymptotics of $Y(k,u)$ as $k \to \infty$ is of special interest. First observe that, for $i \in \mathbb{N}_0$, the random variable $X_i(k,u)$ has a binomial distribution with parameters k and u_i. In particular, $X_0(k,u)/k \to 1 - |u|$ almost surely as $k \to \infty$. For $u \in \Delta$ let $K(k,u) := \sum_{i\ge1} 1_{\{X_i(k,u)\ge1\}}$ denote the number of occupied urns other than J_0 after $k \in \mathbb{N}$ throws with mean $\phi(k,u) := \mathbb{E}(K(k,u))$. Then $Y(k,u) = X_0(k,u) + K(k,u)$ and $\phi(k,u) = \sum_{i\ge1}(1-(1-u_i)^k)$. Lemma III.5.1 shows that $K(k,u)/\phi(k,u) \to 1$ almost surely as

$k \to \infty$. From $\phi(k,u)/k \to 0$ it follows that (see Lemma III.5.2)

$$\frac{Y(k,u)}{k} \;\to\; 1-|u|, \qquad k \to \infty,$$

almost surely for each $u \in \Delta$.

Explicit formulas for the rates of the block counting process are known, but notationally complicated. It holds that (see [13, Eq. (1.3)] or [14, Proposition 2.1])

$$q_{k,j} \;=\; a\binom{k}{2}1_{\{j=k-1\}} \;+\; \int_\Delta \sum_{1\le i\le j} f_{kji}(u)\,\nu(\mathrm{d}u),$$

$1 \le j < k, k \ge 2$, where

$$f_{kji}(u) \;:=\; \sum_{\substack{k_1,\ldots,k_i\in\mathbb{N}\\ k_1+\cdots+k_i=k-j+i}} \frac{k!}{(j-i)!k_1!\cdots k_i!}(1-|u|)^{j-i} \sum_{\substack{l_1,\ldots,l_i\in\mathbb{N}\\ l_1<\cdots<l_i}} u_{l_1}^{k_1}\cdots u_{l_i}^{k_i}$$

for $i \in \{1,\ldots,j\}$ and $u \in \Delta$.

1.5 The function γ

Now we introduce a function known from the literature. See, for example, [17, 20] for Ξ-coalescents and [4, 10, 11, 21] for Λ-coalescents. The relevance first became apparent in Schweinsberg's work [29], where the function is used in the formulation of a necessary and sufficient condition for the Λ-coalescent to come down from infinity.

Recall that $a = \Xi(\{0\})$ and $\nu(\mathrm{d}u) = (u,u)^{-1}\Xi_0(\mathrm{d}u)$, where Ξ_0 is defined via $\Xi = a\varepsilon_0 + \Xi_0$. Define $\gamma : [0,\infty) \to \mathbb{R}$ via

$$\gamma(x) \;:=\; a\binom{x}{2} \;+\; \int_\Delta \sum_{i\ge1}\big((1-u_i)^x - 1 + xu_i\big)\nu(\mathrm{d}u), \tag{1.6}$$

$x \in [0,\infty)$. Some properties of γ are collected in Lemma III.2.1. In particular, the function γ is continuous and asymptotically grows at least of order x but not faster than of order x^2 as $x \to \infty$. On $[1,\infty)$, the functions γ and $x \mapsto \gamma(x)/x$ are strictly increasing. Moreover, by (1.5),

$$\begin{aligned}
\sum_{j=1}^{k-1}(k-j)q_{k,j} \;&=\; a\binom{k}{2} \;+\; \int_\Delta \mathbb{E}(k-Y(k,u))\,\nu(\mathrm{d}u)\\
&=\; a\binom{k}{2} \;+\; \int_\Delta \Big(k - k(1-|u|) - \sum_{i\ge1}(1-(1-u_i)^k)\Big)\nu(\mathrm{d}u)\\
&=\; \gamma(k), \qquad k \in \mathbb{N},
\end{aligned}$$

yielding the interpretation of the quantity γ as the expected change of the block counting process: if the block counting process is in state $k \in [n]$ at time t, then after dt units of time

$$\mathbb{E}\big(N_{t+\mathrm{d}t}^{(n)} \big| N_t^{(n)} = k\big) \;=\; k - \gamma(k)\mathrm{d}t.$$

Besides Schweinsberg's condition for the coalescent to come down from infinity the function γ can be used to formulate a necessary and sufficient condition for the coalescent to have dust, see (1.13) below.

1.6 Coming down from infinity

In this section we give the definition of "coming down from infinity" and one important condition for this characteristic. We also summarize results of Berestycki, Berestycki and Limic' work [4, 20, 21] concerning the small-time behavior of the block counting process of coalescents that come down from infinity, because there are parallels to the results of Article III.

The first coalescent studied was the Kingman coalescent, the Λ-coalescent with underlying measure $\Lambda = \varepsilon_0$. One remarkable result Kingman proved states that the number of blocks is finite for all times $t > 0$ almost surely. For the Bolthausen–Sznitman coalescent, the Λ-coalescent whose underlying measure is the uniform distribution on $[0, 1]$, the number of blocks is infinite for all times $t > 0$ almost surely. Later Pitman showed that all Λ-coalescents either satisfy

$$\mathbb{P}(N_t < \infty \;\forall\; t > 0) \;=\; 1 \tag{1.7}$$

or

$$\mathbb{P}(N_t = \infty \;\forall\; t \geq 0) \;=\; 1, \tag{1.8}$$

provided that $\Lambda(\{1\}) = 0$ [26, Proposition 23]. The same dichotomy is true for the larger class of Ξ-coalescents. Define $\Delta_f := \{u \in \Delta : u_1 + \ldots + u_n = 1 \text{ for some } n \in \mathbb{N}\}$ and suppose that $\Xi(\Delta_f) = 0$. Again, either (1.7) or (1.8) holds true [30, Lemma 31]. In the first case the coalescent is said to *come down from infinity* or, abbreviating, to be *cdi* and in the second to *stay infinite.*

Naturally, the question about conditions to decide between the two properties arises. Schweinsberg proved that the Λ-coalescent comes down from infinity if and only if $\Lambda(\{1\}) = 0$ and

$$\int_c^\infty \frac{\mathrm{d}u}{\gamma(u)} \tag{1.9}$$

is finite for some (and hence all) $c > 1$ [29, Theorem 1]. Clearly, if $\Lambda(\{1\}) = 0$ and the integral (1.9) is infinite for some $c > 1$, then the Λ-coalescent stays infinite. One of the two implications holds true for Ξ-coalescents but additional assumptions are needed for the inversion. If $\Xi(\Delta_f) = 0$ and the integral (1.9) is finite for some $c > 1$, then the Ξ-coalescent comes down from infinity [30, Proposition 32]. Now define $\Delta_\varepsilon := \{u \in \Delta : |u| \leq 1 - \varepsilon\}$ for $\varepsilon > 0$. If $\Xi(\Delta_f) = 0$, $\nu(\Delta \setminus \Delta_\varepsilon) < \infty$ for each $\varepsilon > 0$ and the integral (1.9) is infinite for some $c > 1$, then the Ξ-coalescent stays infinite [30, Proposition 33]. Schweinsberg's criteria in fact use $\sum_{k=2}^{\infty}(\gamma(k))^{-1}$ instead of (1.9), but in the original version the sum can be replaced by (1.9), since, as a consequence of γ being positive and monotonous, one is finite if and only if the other is finite. Further work on the topic "coming down from infinity" has been done in [17].

The case $\Xi(\Delta_f) > 0$ is described with the use of the Poisson point process construction, e.g., in [30, Section 5.5]. Let Ξ be a finite measure on Δ and decompose Ξ into $\Xi = \Xi_1 + \Xi_2$, where Ξ_1 is the restriction of Ξ to Δ_f and $\Xi_2 := \Xi - \Xi_1$. Replacing Ξ by Ξ_1 and Ξ_2 in (1.3) yields two measures L_{Ξ_1} and L_{Ξ_2} on $\mathbb{Z}^\infty$. Suppose that ξ_1 and ξ_2 are two independent proper Poisson point processes with intensity measures $\lambda \otimes L_{\Xi_1}$ and $\lambda \otimes L_{\Xi_2}$. As in Section 1.3, we can construct a Ξ_i-coalescent by merging blocks according to the points of ξ_i for $i \in \{1, 2\}$. Of course, $\xi := \xi_1 + \xi_2$ is a Poisson point process with intensity measure $\lambda \otimes L$. We say that a point (t, z) of ξ comes from Δ_f if (t, z) is a point of ξ_1. Note that the set $\{z_1, z_2, \ldots\}$ is finite for every point (t, z) of ξ that comes from Δ_f, since the urn J_0 is not present and there are only finitely many other urns for $u \in \Delta_f$. Thus, only finitely many blocks remain after a coagulation of the Ξ-coalescent corresponding to a point (t, z) of ξ that comes from Δ_f. In particular, the Ξ-coalescent does not stay infinite when $\Xi_1(\Delta) = \Xi(\Delta_f) > 0$. It can be shown that the Ξ-coalescent comes down from infinity if $\nu(\Delta_f) = \infty$, so $L_{\Xi_1}(\mathbb{Z}^\infty) = \infty$, or if the Ξ_2-coalescent comes down from infinity. If the Ξ_2-coalescent stays infinite and $\nu(\Delta_f) < \infty$, then the Ξ-coalescent neither comes down from infinity nor stays infinite. For Λ-coalescents the atom at 1 is the rate at which the Λ-coalescent jumps to the absorbing partition consisting only of the block $\mathbb{N}$.

The coalescents in this work start with an infinite number of blocks. If the coalescent comes down from infinity, then the number of blocks N_t after time $t > 0$ is finite but tends to infinity as $t \to 0+$. In

[4], Berestycki et al. determine a rate of divergence for Λ-coalescents, which they call "speed" of coming down from infinity. Almost at the same time, Limic showed similar results for Ξ-coalescents [20]. Suppose that the integral (1.9) is finite for $c > 1$. Then the function $v : (0, \infty) \to (0, \infty)$, determined as the solution to

$$t \;=\; \int_{v(t)}^{\infty} \frac{\mathrm{d}u}{\gamma(u)}, \qquad t > 0, \tag{1.10}$$

is well-defined. Further assume that $\Xi(\Delta_f) = 0$ and that the regularity condition

$$\int_{\Delta} |u|^2 \, \nu(\mathrm{d}u) \;<\; \infty \tag{1.11}$$

is given. Then

$$\frac{N_t}{v(t)} \to 1, \qquad t \to 0+, \tag{1.12}$$

almost surely [4, 20]. For Λ-coalescents the convergence (1.12) also holds true in L_p spaces for $p \geq 1$ [4, Theorem 2]. Observe that, given the regularity condition (1.11), the assumption of the finiteness of the integral (1.9) and $\Xi(\Delta_f) = 0$ pose no actual restriction, since both conditions are necessary for the coalescent to come down from infinity. As every Λ-coalescent satisfies the regularity condition (1.11), the prerequisites of [4, Theorems 1 and 2] reduce to the simple fact that the Λ-coalescent comes down from infinity.

Note that a slightly different variant of γ is used in [4, 20]. But as shown in [21, Lemma 2.2], the speed $v(t)$ is asymptotically equivalent if γ is replaced by the different variant in (1.10). Thus, (1.12) still holds true.

1.7 Coalescents with dust

Let $\Pi = (\Pi_t)_{t \geq 0}$ be a coalescent and, for $t \geq 0$, let $B_1(t), B_2(t), \ldots$ denote the blocks of Π_t. A variant of de Finetti's Theorem (see, e.g., [18, Theorem 2] or [30, Lemma 40]) implies that the asymptotic frequencies

$$f_j(t) \;:=\; \lim_{n \to \infty} \frac{|B_j(t) \cap [n]|}{n}$$

exist for each $j \geq 1$ and $t \geq 0$ almost surely. As a consequence of Kingman's correspondence between exchangeable random partitions of $\mathbb{N}$ and finite measures on Δ (see [18] or [30, Appendix A]), the blocks of Π_t are either singletons or infinitely large. In fact, the infinitely large blocks have positive asymptotic frequency. Moreover, the number of singletons is either infinite or there are none. Define $S_t := 1 - \sum_{j\geq 1} f_j(t)$, $t \geq 0$, to be the frequency of singletons. In the literature, a coalescent is said to have *proper frequencies* if $S_t = 0$ for all $t \geq 0$ almost surely. Coalescents are said to have *dust*, if they do not have proper frequencies, i.e., if there exist singletons at some time $t > 0$. Given $\Xi(\Delta_f) = 0$, the presence of singletons in a coalescent does not depend on the time, meaning that the number of singletons is either infinite for all times $t > 0$ almost surely or Π_t has no singletons for all $t > 0$ almost surely. If $\Xi(\Delta_f) > 0$, then the coalescent jumps to a state with only finitely many, infinitely large blocks after a finite random time and no singletons remain.

It was first shown for Λ-coalescents by Pitman [26, Proposition 26] and later for the larger class of Ξ-coalescents by Schweinsberg [30, Proposition 30] that the coalescent has dust if and only if

$$a \;=\; 0 \qquad \text{and} \qquad \mu \;:=\; \int_\Delta |u|\, \nu(\mathrm{d}u) \;<\; \infty.$$

Note that the function $x \mapsto \gamma(x)/x$, $x \geq 1$, is non-decreasing, converges to μ, and $\gamma(x)/x \geq a(x-1)/2$ for $x \geq 1$. Hence, the coalescent has dust if and only if

$$\lim_{x\to\infty} \gamma(x)/x \;<\; \infty. \tag{1.13}$$

Of course, a coalescent $\Pi = (\Pi_t)_{t\geq 0}$ with dust does not come down from infinity, and if $\Xi(\Delta_f) = 0$, then Π even stays infinite. An illustrative picture is provided in Article III.

1.8 The fixation line

The fixation line was introduced first by Pfaffelhuber und Wakolbinger [25] in order to study the genealogy back to the most recent common ancestor of the population in a continuous-time Moran model, whose backwards genealogy is the Kingman coalescent. A construction of the fixation line based on the Lookdown-model is feasible for the full class of Ξ-coalescents, see Gaiser and Möhle [14]. Earlier Hénard [16] properly defined and studied the fixation line of Λ-coalescents. Here it suffices to define the fixation line as the Siegmund-dual [31] of the block counting process: the fixation line $L = (L_t)_{t\geq 0}$ is a Markov process that satisfies (see [14])

$$\mathbb{P}\big(L_t^{(m)} \geq n\big) \;=\; \mathbb{P}\big(N_t^{(n)} \leq m\big), \qquad m, n \in \mathbb{N}, t \geq 0, \tag{1.14}$$

where the upper index "(m)" denotes the initial state $L_0^{(m)} = m$ at time $t = 0$. For a thorough definition of the fixation line see [14] or [16] and the references therein.

From (1.14) we can conclude that the fixation line has non-decreasing paths. The fixation line explodes if and only if the block counting process becomes finite, so if and only if the coalescent comes down from infinity. We also obtain the following formula for the rates $\gamma_{k,j} := \lim_{t\to 0+} t^{-1}(\mathbb{P}(L_t^{(k)} = j) - \delta_{kj})$. By (1.5),

$$\begin{aligned}\gamma_{k,j} &= \sum_{1\le i\le k} (q_{j,i} - q_{j+1,i}) \\ &= a\binom{j}{2}1_{\{j=k+1\}} + \int_\Delta \mathbb{P}\big(Y(j,u) \le k, Y(j+1,u) > k\big)\nu(\mathrm{d}u)\end{aligned}$$

for $j, k \in \mathbb{N}$ with $j > k$, and note that only upward jumps are possible. For Λ-coalescents we have

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

for $j, k \in \mathbb{N}$ with $j > k$.

1.9 Markov branching processes

The Bienaymé–Galton–Watson process is a process $Z = (Z_n)_{n\in\mathbb{N}_0}$ that satisfies the recursion

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{i,n}, \qquad n \in \mathbb{N}_0, \tag{1.15}$$

where $\xi_{i,n}$, $i \in \mathbb{N}, n \in \mathbb{N}_0$, are i.i.d. $\mathbb{N}_0$-valued random variables. The process is usually interpreted as the size of an evolving population. Start with a number $Z_0 \in \mathbb{N}$ of individuals. Each individual lives for one unit of time and at the moment of its death the individual produces a random number of children, independently of the others. The children also live for one unit of time and at the moment of their death each of them produces new children, which again live for one unit of time before they reproduce, and so on. The number of children of an individual is independent of the offspring sizes of all previously living individuals and the distribution of the number of children is the same for all individuals. Label the individuals that are alive at the same time by $1, 2, \ldots$ and suppose that the offspring size of individual i that lives after n units of time is given by $\xi_{i,n}$. In this case Z_n, defined via (1.15), is equal to the size of the population after n units of time.

The classical model has been adapted in various ways. We will focus on the modification in which the lifetime of each individual has an exponential distribution with parameter $a \in (0, \infty)$, independently of the lifetimes of other individuals and the offspring sizes, and the number of children which are born at the end of a lifetime are still i.i.d. $\mathbb{N}_0$-valued random variables. Let $Z_t^{(n)}$ denote the number of individuals that are alive at time $t \geq 0$ in a population that started from $n \in \mathbb{N}$ individuals at time $t = 0$. The assumptions about the lifetime guarantee that $Z^{(n)} := (Z_t^{(n)})_{t \geq 0}$ is a time-homogeneous Markov chain. The class of processes $Z^{(n)}$ we have just constructed are the well known continuous-time Markov branching processes with discrete state space $\mathbb{N}_0 \cup \{\infty\}$. In the remainder of this section and in Article I, $Z^{(n)}$ denotes a continuous-time Markov branching process and is shortly referred to as branching process. We exemplary refer the reader to the book by Athreya and Ney [2, Ch. 3] for results about such branching processes.

From the construction it follows that the *branching property* holds, stating that $Z^{(n)}$ has the same distribution as the sum of n independent copies of $Z := (Z_t)_{t \geq 0} := Z^{(1)}$. In order to understand the behavior of the branching process it therefore suffices to analyze Z, like it is done in most of the literature.

Let ξ be the size of the offspring of an arbitrary individual from the population with mean $m := \mathbb{E}(\xi)$ and probability generating function (pgf) f, given by $f(s) := \mathbb{E}(s^{\xi})$, $s \in [0, 1]$. Also, let $u(s) := a(f(s) - s)$, $s \in [0, 1]$, and $\lambda := u'(1) = a(m - 1)$.

From now on we exclude the uninteresting case $\mathbb{P}(\xi = 1) = 1$. Then the process almost surely either ends up in the absorbing state 0 or $Z_t \to \infty$ as $t \to \infty$. The event that the process eventually reaches the state 0 is called *extinction*. The probability of extinction, denoted by q, is the smallest solution to the equation $f(s) = s$ in $[0, 1]$. The branching process is said to be *critical* (*subcritical*, *supercritical*) if $m = 1$ $(m < 1, m > 1)$. In the subcritical and critical case it holds that $q = 1$ and in the supercritical case $q < 1$. The branching process does not explode, meaning that $\mathbb{P}(Z_t < \infty) = 1$ for all $t \geq 0$ if and only if

$$\int_{1-\varepsilon}^{1} \frac{\mathrm{d}s}{s - f(s)} = \infty$$

for some $1 - q > \varepsilon > 0$. Note that the process does not explode if $m < \infty$.

Of course, the distribution of Z is fully determined by the expected lifetime a^{-1} and the offspring pgf f. The one-dimensional distributions are accessible through the pgf $F(., t)$ of Z_t, defined via $F(s,t) := \mathbb{E}(s^{Z_t})$, $s \in [0,1], t \geq 0$, and uniquely determined by the boundary condition $F(s,0) = s$ and the forward and backward equation

$$\frac{\partial}{\partial t} F(s,t) \;=\; u(s) \frac{\partial}{\partial s} F(s,t) \quad \text{and} \quad \frac{\partial}{\partial t} F(s,t) \;=\; u(F(s,t)), \quad (1.16)$$

$t \geq 0, s \in [0,1]$. By the definition of u, the backward equation is equivalent to

$$at \;=\; \int_s^{F(s,t)} \frac{\mathrm{d}u}{f(u) - u} \qquad (1.17)$$

for $t \geq 0$ as long as $s \neq q$. Moreover, let $m(t) := \mathbb{E}(Z_t)$ denote the expected size of the population after time $t \geq 0$. Then $m(t) = e^{\lambda t}$ for $t \geq 0$ and, as a consequence of the branching property, we obtain that the process $(Z_t/m(t))_{t \geq 0}$ is a martingale, which is used in some of the proofs of Article I.

Chapter 2

Discussion of the results

Articles II and III provide scaling limits for the block counting process of coalescents that stay infinite when the initial sample size n tends to infinity. Analogous results can be stated for the fixation line. The main results are presented in Section 2.1. Article I provides scaling limits for branching processes, which are presented in Section 2.2. The connection between the two problems is explained at the end of Subsection 2.2.3.

2.1 Scaling limits for the block counting process

The key assumption of the results for the block counting process and the fixation line is formulated in terms of the function γ introduced in Section 1.5. A subsequent discussion provides a list of conditions that are equivalent to the key assumption. In the second subsection, the results are stated.

2.1.1 The key assumption

Define

$$\kappa \; := \; \lim_{x\to\infty} x\gamma''(x) \; \in \; [0,\infty] \tag{2.1}$$

whenever this limit exists. Given that the limit in (2.1) exists, we call κ *asymptotic curvature* or *curvature parameter*. Note that the limit does not exist for all coalescents. The key assumption of Article III is the following.

$$\text{The limit } \kappa \text{ in (2.1) exists and is finite.} \tag{2.2}$$

The following comments should help to put the key assumption into perspective. For coalescents with dust the limit exists and $\kappa = 0$, though there exist dust-free coalescents for which $\kappa = 0$, see Example III.3.3. Given that the regularity condition (1.11) is satisfied, the key assumption implies that the coalescent does not come down from infinity. If additionally $\nu(\Delta_f) = 0$, then the coalescent stays infinite.

The existence of the limit κ can be linked to the behavior of Ξ near zero with the use of a Tauberian theorem. Define the function $G : [0,1] \to [0,\infty)$ via

$$G(t) \;:=\; \int_\Delta \sum_{i\geq 1} 1_{[0,t]}(u_i)u_i^2\,\nu(\mathrm{d}u), \qquad t \in [0,1].$$

Proposition III.2.3 states that (2.2) is equivalent to

$$\kappa \;=\; \lim_{t\to 0+} t^{-1}G(t). \tag{2.3}$$

For Λ-coalescents (2.3) simplifies to

$$\kappa \;=\; \lim_{t\to 0+} t^{-1}\Lambda([0,t]). \tag{2.4}$$

Eq. (2.4) essentially means that Λ behaves like Lebesgue measure near 0. From (2.4) it can easily be seen that the Bolthausen–Sznitman coalescent satisfies the key assumption with $\kappa = 1$. As a generalization, the $\beta(1,b)$-coalescent, the particular Λ-coalescent whose underlying measure $\Lambda = \beta(1,b)$ has density function $u \mapsto b(1-u)^{b-1}$, $u \in [0,1]$, satisfies the key assumption with $\kappa = b$ for all $b > 0$. The $\beta(1,b)$-coalescent is treated in detail in Article II (Example II.4.2).

The assumptions of the main results of Article II are more restrictive. First of all we only consider Λ-coalescents. The main result (Theorem II.2.3) states the convergence of the scaled block counting process under the assumption that (see Assumption A in Article II)

$$\int_{[0,1]} u^{-1}\,(\Lambda - \kappa\lambda)(\mathrm{d}u) \;<\; \infty \tag{2.5}$$

for some $\kappa \geq 0$, where λ denotes Lebesgue measure on $[0,1]$. Additionally it is assumed that $\Lambda(\{1\}) = 0$. In Article II, the constant κ is denoted by b but we use κ here, since the two constants in (2.2) and (2.5) coincide. As shown in Lemma II.9.1, (2.5) in fact implies (2.4) and, hence, (2.2). The converse implication is not true. Article III provides an example of a Λ-coalescent that satisfies the key assumption although not being covered in Article II (Example III.3.3).

The main results in Articles II and III are proved by verifying that the generator of the transformed block counting process converges uniformly. At the time of writing Article II, it wasn't fully understood that (2.2) is the suitable assumption under which the uniform convergence can be shown. By Proposition III.2.2, the key assumption (2.2) holds true if and only if

$$\frac{\gamma(x)}{x} \;=\; \kappa \log x \;+\; \log \ell(x), \qquad x \geq 0, \tag{2.6}$$

for some slowly varying function $\ell : [0, \infty) \to (0, \infty)$. By the definition of slowly varying functions, it is equivalent to require that the limit

$$\lim_{x \to \infty} \left(\frac{\gamma(xy)}{xy} - \frac{\gamma(x)}{x} \right)$$

exists for each $y > 0$. This condition is used in the proof of the main result (cf. (2.9)).

2.1.2 Results

Recall that $N^{(n)} = (N_t^{(n)})_{t \geq 0}$ denotes the block counting process with initial state n. In Articles II and III, we investigate the existence of scaling constants $v(n,t)$ for which the process $X^{(n)} := (X_t^{(n)})_{t \geq 0}$, defined via

$$X_t^{(n)} \;:=\; \log N_t^{(n)} \;-\; \log v(n,t), \qquad t \geq 0, n \in \mathbb{N}, \tag{2.7}$$

converges to a non-degenerate limiting process as $n \to \infty$ in the space $D_{\mathbb{R}}[0, \infty)$ of càdlàg paths endowed with the Skorohod topology. We consider the logarithmic form for technical reasons only. The convergence results are presented in non-logarithmic form in Section III.2.5.

We are primarily interested in coalescents that do not come down from infinity. In the literature, the problem is solved for coalescents with dust and for the Bolthausen–Sznitman coalescent. In [14], it is shown that for coalescents with dust the scaling $v(n,t) := n$, $t \geq 0, n \in \mathbb{N}$, can be chosen; also see [22]. In [19], it is shown that for the Bolthausen–Sznitman coalescent the scaling $v(n,t) := n^{e^{-t}}$, $t \geq 0, n \in \mathbb{N}$, can be chosen. Earlier Goldschmidt and Martin [15] and Baur and Bertoin [3] even showed the almost sure convergence of $N_t^{(n)}/n^{e^{-t}}$ as $n \to \infty$ for all $t \geq 0$. Both convergence results, the dust case and the Bolthausen–Sznitman case, are reshown in Articles II and III, although our methods differ and the convergence result in [14] for coalescents with dust is slightly more general.

The scaling function for the block counting process can be defined without any assumptions on γ (or Ξ). One might compare the definition to (1.10). Define $v : [1, \infty) \times [0, \infty) \to [1, \infty)$ via

$$v(1,t) \;:=\; 1 \quad \text{and} \quad \int_{v(x,t)}^{x} \frac{\mathrm{d}u}{\gamma(u)} \;=\; t, \qquad x > 1, t \geq 0. \tag{2.8}$$

Proposition III.2.4 shows that the scaling function is well-defined. The proposition also states that $v(x, .)$ is the solution to the initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t} v(x,t) \;=\; -\gamma(v(x,t)), \quad t \geq 0, \qquad v(x,0) \;=\; x,$$

for all $x \geq 1$. The definition of the scaling function thus makes sense, since $\gamma(k)$ is the expected rate of decrease of the block counting process if the coalescent currently is in a state with k blocks.

Now assume that the key assumption (2.2) is satisfied with $\kappa \geq 0$. Then there exist slowly varying functions $\ell_t : [1, \infty) \to (0, \infty)$ such that (see Proposition III.2.5)

$$v(x,t) \;=\; x^{e^{-\kappa t}} \ell_t(x), \qquad x \geq 1, t \geq 0.$$

Put $\Delta^* := \{u \in \Delta : |u| = 1\}$ and further assume that $\nu(\Delta^*) = 0$ and that the regularity condition (1.11) holds true. The main result (Theorem III.2.7) states that the process $X^{(n)}$, defined via (2.7), converges in $D_{\mathbb{R}}[0, \infty)$ as $n \to \infty$ for the particular scaling function v, defined via (2.8). The limiting process, denoted by X, is an Ornstein–Uhlenbeck type process that can be characterized as follows. Define $\psi : \mathbb{R} \to \mathbb{C}$ via

$$\psi(x) \;:=\; \int_{\Delta} \big((1 - |u|)^{ix} - 1 + ix|u|\big) \nu(\mathrm{d}u), \qquad x \in \mathbb{R}.$$

The assumptions $\nu(\Delta^*) = 0$ and (1.11) imply that ψ is the characteristic exponent of an infinitely divisible distribution. The limiting process $X = (X_t)_{t \geq 0}$ is a real-valued Markov process with initial value $X_0 = 0$ and semigroup

$$\mathbb{E}(f(X_{s+t}) \,|\, X_s = x) \;=\; \mathbb{E}(f(e^{-\kappa t} x + X_t)),$$

$x \in \mathbb{R}, f \in B(\mathbb{R}), s, t \geq 0$, where the marginal distributions are determined by the characteristic functions

$$\mathbb{E}(e^{ixX_t}) \;=\; \exp\bigg(\int_0^t \psi(e^{-\kappa s} x) \,\mathrm{d}s \bigg), \qquad x \in \mathbb{R}, t \geq 0.$$

The Markov process X is alternatively determined by the corresponding generator A, given by

$$\begin{aligned} Af(x) &= -\kappa x f'(x) \\ &\quad + \int_\Delta \big(f(x+\log(1-|u|)) - f(x) + |u| f'(x)\big)\nu(\mathrm{d}u), \end{aligned}$$

$x \in \mathbb{R}$, for suitable functions f (see Lemma III.4.1 or [28, Theorem 3.1]). We are able to write the generator corresponding to $X^{(n)}$ in a comparable way. The generators $(A_s^{(n)})_{s\geq 0}$ of the (time-inhomogeneous) Markov process $X^{(n)}$ are defined via $A_s^{(n)} f(x) := \lim_{t\to 0+}(\mathbb{E}(f(X_{s+t}^{(n)}) \,|\, X_s^{(n)} = x) - f(x))$ for $s \geq 0$. Write $k := k(s,x,n) := e^x v(n,s)$ for $(s,x) \in \widetilde{E}_n$ and $n \in \mathbb{N}$. Then, by (1.5),

$$\begin{aligned} A_s^{(n)} f(x) &= f'(x)\Big(\tfrac{\gamma(e^{-x}k)}{e^{-x}k} - \tfrac{\gamma(k)}{k}\Big) \qquad (2.9) \\ &\quad + \int_\Delta \mathbb{E}\Big(f\big(x+\log\tfrac{Y(k,u)}{k}\big) - f(x) + \big(1-\tfrac{Y(k,u)}{k}\big)f'(x)\Big)\nu(\mathrm{d}u), \end{aligned}$$

where the random variables $Y(k,u)$ are defined in Section 1.4. By using the fact that $Y(k,u)/k \to 1-|u|$ almost surely as $k \to \infty$, it can be shown that $A_s^{(n)}$ uniformly converges to A, implying the desired convergence of $X^{(n)}$ to X as $n \to \infty$.

An analogous convergence result can be stated for the fixation line $(L_t^{(n)})_{t\geq 0}$ with initial state $L_0^{(n)} = n$; see [14] for coalescents with dust and [19] for the Bolthausen–Sznitman coalescent. The scaling $w(x,t)$ for the fixation line is defined as the inverse of v, in the sense that $w(.,t)$ is the inverse of $v(.,t)$ for all $t \geq 0$. In order for $w(x,t)$ to be defined for all $x \geq 1$ the assumption $\int_c^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$ for some $c > 1$ is necessary. Given the key assumption (2.2), the map $w(.,t)$ is regularly varying with index $e^{\kappa t}$ for each $t \geq 0$, since the property of regular variation transfers to the inverse function. There exist slowly varying functions $\ell_t^{\#} : [1,\infty) \to (0,\infty)$ such that

$$w(x,t) = x^{e^{\kappa t}} \ell_t^{\#}(x), \qquad x \geq 1, t \geq 0.$$

Assume that $\nu(\Delta^*) = 0$ and that the regularity condition (1.11) is given. Under these additional assumptions, the requirement $\int_c^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$ simply means that the fixation line be non-exploding. Theorem III.2.10 states that the process $Y^{(n)} := (Y_t^{(n)})_{t\geq 0}$, defined via

$$Y_t^{(n)} := \log L_t^{(n)} - \log w(n,t)$$

converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$.

2.2 Scaling limits for Markov branching processes

Recall that $Z^{(n)} = (Z_t^{(n)})_{t\geq 0}$ denotes a branching process with initial state $Z_0^{(n)} = n$, offspring distribution ξ and expected lifetime a^{-1}. Moreover, define $Z := (Z_t)_{t\geq 0} := (Z_t^{(1)})_{t\geq 0}$, f as the offspring pgf, given by $f(s) := \mathbb{E}(s^\xi)$, $s \in [0,1]$, and $F(s,t)$ as the pgf of Z_t, given by $F(s,t) = \mathbb{E}(s^{Z_t})$, $s \in [0,1], t \geq 0$. We also use the notation $m := \mathbb{E}(\xi)$ and $m(t) := \mathbb{E}(Z_t)$ for the mean of ξ and Z_t for $t \geq 0$.

In Article I, we pursue the question of the existence of constants $a(n,t)$ and $b(n,t)$ for $n \in \mathbb{N}$ and $t \geq 0$ for which the process $X^{(n)} := (X_t^{(n)})_{t\geq 0}$, defined via

$$X_t^{(n)} := \frac{Z_t^{(n)} - b(n,t)}{a(n,t)}, \qquad t \geq 0, n \in \mathbb{N}, \tag{2.10}$$

converges in $D_\mathbb{R}[0,\infty)$ to a non-degenerate limiting process as $n \to \infty$. The branching property states that $Z^{(n)}$ has the same distribution as n independent copies of Z. The problem of the convergence of one-dimensional distributions of $X^{(n)}$ can hence be solved by utilizing the theory of stable distributions and their domains of attraction. The necessary and sufficient conditions for convergence are well known and are here expressed in terms of the pgf $F(.,t)$.

We distinguish three cases, whose prerequisites differ in the finiteness of the offspring distribution's first and second moment. Article I contains three convergence results for the processes $X^{(n)}$, one in each case, and results that relate particular asymptotics of $F(.,t)$ to the offspring pgf f. Examples for all three regimes are given in Article I. A short discussion of scaling limits for explosive Markov branching processes is included too.

2.2.1 The finite variance case

Assume that $\mathbb{E}(\xi^2) < \infty$. Then $\mathbb{E}(Z_t^2) < \infty$ for all $t \geq 0$. According to the central limit theorem we can choose $a(n,t) := \sqrt{n}$ and $b(n,t) := nm(t)$ in (2.10) in order for the one-dimensional distributions of $X^{(n)}$ to converge. In fact, we apply a multivariate version of the central limit theorem in order to obtain the convergence of the finite-dimensional distributions. Afterwards the convergence in $D_\mathbb{R}[0,\infty)$ is established by using a criterion by Aldous for martingales [1]. Theorem I.2.1 states that the process $X^{(n)} := (X_t^{(n)})_{t\geq 0}$, defined via

$$X_t^{(n)} := \frac{Z_t^{(n)} - nm(t)}{\sqrt{n}}, \qquad n \in \mathbb{N}, t \geq 0,$$

converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to a continuous Gaussian Markov process.

2.2.2 The finite mean infinite variance case

Assume that $m := \mathbb{E}(\xi) < \infty$ and $\mathbb{E}(\xi^2) = \infty$. In order for the one-dimensional distributions of the process $X^{(n)}$, defined via (2.10), to converge we need Z_t to belong to the domain of attraction of an α-stable law with $\alpha \in [1,2]$ for all $t \geq 0$. Moreover, the parameter α cannot depend on t. Indeed, if Z_t belongs to the domain of attraction of an α-stable law, then $\mathbb{E}(Z_t^\beta) < \infty$ for all $\beta < \alpha$ and if in addition $\alpha < 2$, then $\mathbb{E}(Z_t^\beta) = \infty$ for all $\beta > \alpha$. But the finiteness of moments of Z_t does not depend on the time, meaning that, for each $r \geq 1$, $\mathbb{E}(Z_t^r) < \infty$ for all $t \geq 0$ if and only if $\mathbb{E}(\xi^r) < \infty$. Therefore, α does not depend on t.

Article I does not cover the case $\alpha = 1$. Combining results by Feller [12, Ch. XVII.5] and by Bingham and Doney [6, Theorem A] shows that a necessary and sufficient condition for Z_t to belong to the domain of attraction of an α-stable law with $\alpha \in (1,2]$ is given by

$$1 - F(s,t) = m(t)(1-s) - (1-s)^\alpha \ell_t((1-s)^{-1}), \tag{2.11}$$

$s \in [0,1), t \geq 0$, where the function $\ell_t : [1,\infty) \to \mathbb{R}$ varies slowly for each $t \geq 0$.

In the finite mean infinite variance case we assume that there exist $\alpha \in (1,2]$ and a slowly varying function $\ell : [1,\infty) \to (0,\infty)$ such that

$$1 - f(s) \ = \ m(1-s) \ - \ (1-s)^\alpha \ell((1-s)^{-1}), \qquad s \in [0,1).$$

Then we can derive from the forward and backward equation, see (1.16) and (1.17), that (2.11) holds true with the slowly varying function satisfying $\ell_t(x) = c(t)\ell(x)(1 + o(1))$ as $x \to \infty$, where

$$c(t) := \begin{cases} at & \text{if } m = 1 \text{ (critical case)}, \\ \dfrac{m(\alpha t) - m(t)}{(\alpha - 1)(m-1)} & \text{if } m \neq 1 \text{ (non-critical case)}. \end{cases}$$

Theorem I.2.3 states that the process $X^{(n)} := (X_t^{(n)})_{t \geq 0}$, defined via

$$X_t^{(n)} := \frac{Z_t^{(n)} - nm(t)}{a_n}, \qquad n \in \mathbb{N}, t \geq 0,$$

converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to a time-inhomogeneous Ornstein–Uhlenbeck type process for a suitably chosen scaling sequence $(a_n)_{n\in\mathbb{N}}$. As we have seen, the convergence of the one-dimensional distribution follows from the theory about stable distributions and their domains of attraction. The convergence in $D_{\mathbb{R}}[0,\infty)$ is obtained by showing that the semigroups converge uniformly on a sufficiently large set.

2.2.3 The infinite mean case with non-explosion

Assume that $m = \infty$ and that the branching process still does not explode. The distribution of Z_t belongs to the domain of attraction of an $\alpha(t)$-stable law if and only if

$$\mathbb{P}(Z_t > x) \ \sim \ x^{-\alpha(t)}\Gamma(1-\alpha(t))^{-1}\ell_t(x), \qquad x \to \infty, \tag{2.12}$$

for some slowly varying functions $\ell_t : [0,\infty) \to (0,\infty)$. Here the gamma function only has a corrective meaning. The assumption $m = \infty$ implies that necessarily $\alpha(t) \in (0,1]$. In difference to the previous section, $\alpha(t)$ depends on $t \geq 0$. Again, the case $\alpha(t) = 1$ is excluded. For $\alpha(t) < 1$ it follows from Bingham and Doney [6, Theorem A] that the tail behavior (2.12) is equivalent to

$$1 - F(s,t) \ = \ (1-s)^{\alpha(t)}\ell_t((1-s)^{-1}), \qquad s \in [0,1). \tag{2.13}$$

Define the function $\ell : [1,\infty) \to (0,\infty)$ via

$$1 - f(s) \ = \ (1-s)\ell((1-s)^{-1}), \qquad s \in [0,1).$$

Lemma I.2.6 shows that (2.13) is satisfied for every $t > 0$ with $\alpha(t) \in (0,1)$ and slowly varying functions $\ell_t : [1,\infty) \to (0,\infty)$ if and only if there exists $A \in (0,\infty)$ such that

$$\ell(x) \ \sim \ A\log x, \qquad x \to \infty. \tag{2.14}$$

In this case $\alpha(t) = e^{-aAt}$ for all $t \geq 0$. The proof utilizes the backward equation. If in addition the limit

$$B \ := \ \lim_{x\to\infty}(\ell(x) - A\log x) \ \in \ \mathbb{R} \tag{2.15}$$

exists, then

$$\lim_{x\to\infty}\ell_t(x) = \exp\left(\frac{B-1}{A}(1-\alpha(t))\right), \qquad t \geq 0.$$

Under the assumption that (2.14) and (2.15) hold true, Theorem I.2.8 states that the process $X^{(n)} := (X_t^{(n)})_{t\geq 0}$, defined via

$$X_t^{(n)} \ := \ n^{-1/\alpha(t)}Z_t^{(n)}, \qquad t \geq 0, n \in \mathbb{N},$$

converges in $D_{[0,\infty)}[0,\infty)$ as $n \to \infty$ to a continuous-state branching process.

We are now able to describe the connection between Article I on one hand and Articles II and III on the other. The three articles state the same result for one particular example. The Bolthausen–Sznitman coalescent satisfies the assumptions of Articles II and III. The fixation line of the Bolthausen–Sznitman coalescent is also a branching process with expected lifetime $a^{-1} = 1$ and offspring pgf $f(s) = s+(1-s)\log(1-s)$, $s \in [0,1]$. Hence, Eq. (2.14) and Eq. (2.15) hold true with $\ell(x) = \log x + 1$, $x \geq 1$, $A = 1$ and $B = 1$, and the results of Article I in the infinite mean case are applicable. In this case Theorem II.2.4 and Theorem I.2.8 both state that

$$Z_t^{(n)}/n^{e^{-t}}$$

converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$. The limiting process is Neveu's continuous-state branching process [24]. Also see Example I.2.9 or Theorem III.2.11 (although with a scaling that differs by a constant not depending on n) and [19, Theorem 3.1 b)] for a more detailed analysis of the fixation line of the Bolthausen–Sznitman coalescent.

Bibliography

[1] Aldous, D. (1989) Stopping times and tightness II. *Ann. Probab.* **17**, no. 2, 586–595.

[2] Athreya, K. B. and Ney, P. E. (1972) *Branching Processes.* Die Grundlehren der Mathematischen Wissenschaften **196**, Springer.

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

[4] Berestycki, J., Berestycki, N. and Limic, V. (2010) The Λ-coalescent speed of coming down from infinity. *Ann. Probab.* **38**, no. 1, 207–233.

[5] Beroin, J. and Le Gall, J.-F. (2003) Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields* **128**, no. 2, 261–288.

[6] Bingham, N. H. and Doney, R. A. (1974) Asymptotic properties of supercritical branching processes I: The Galton-Watson process. *Adv. Appl. Probab.* **6**, 711–731.

[7] Bolthausen, E. and Sznitman, A.-S. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197**, no. 2, 247–276.

[8] Cannings, C. (1974) The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Adv. Appl. Probab.* **6**, no. 2, 260–290.

[9] Cannings, C. (1975) The latent roots of certain Markov chains arising in genetics: a new approach, II. Further haploid models. *Adv. Appl. Probab.* **7**, no. 2, 264–282.

[10] Diehl, C. and Kersting, G. (2019) Tree lengths for general Λ-coalescents and the asymptotic site frequency spectrum around the Bolthausen–Sznitman coalescent. *Ann. Appl. Probab.* **29**, no. 5, 2700–2743.

[11] Diehl, C. and Kersting, G. (2019) External branch lengths of Λ-coalescents without a dust component. *Electron. J. Probab.* **24**, no. 134, 36pp.

[12] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications.* Volume 2, Second Edition, Wiley, New York.

[13] Freund, F. and M. Möhle, M. (2009) On the number of allelic types for samples taken from exchangeable coalescents with mutation. *Adv. in Appl. Probab.* **41**, no. 4, 1082–1101.

[14] Gaiser, F. and Möhle, M. (2016) On the block counting process and the fixation line of exchangeable coalescents. *ALEA Lat. Am. J. Probab. Math. Stat.* **13**, no. 2, 809–833.

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

[16] Hénard, O. (2015) The fixation line in the Λ-coalescent *Ann. Appl. Prob.* **25**, no. 5, 3007–3032.

[17] Herriger, P. and Möhle, M. (2012) Conditions for exchangeable coalescents to come down from infinity. *ALEA Lat. Am. J. Probab. Math. Stat.* **9**, no. 2, 637–665.

[18] Kingman, J. F. C. The coalescent. *Stochastic Process. Appl.* **13**, no. 3, 235–248.

[19] Kukla, J. and Möhle, M. (2018) On the block counting process and the fixation line of the Bolthausen–Sznitman coalescent. *Stochastic Process. Appl.* **128**, no. 3, 939–962.

[20] Limic, V. (2010) On the speed of coming down from infinity for Ξ-coalescent processes. *Electron. J. Probab.* **15**, no. 8, 217–240.

[21] Limic, V. and Talarczyk, A. (2015) Second order asymptotics for the block counting process in a class of regularly varying Λ-coalescents. *Ann. Probab.* **43**, no. 3, 1419–1455.

[22] Möhle, M. (2021) The rate of convergence of the block counting process of exchangeable coalescents with dust. *ALEA Lat. Am. J. Probab. Math. Stat.* **18**, no. 2, 1195–1220.

[23] Möhle, M. and Sagitov, S. (2001) A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, no. 4, 1547–1562.

[24] Neveu, J. (1992) A continuous-state branching process in relation with the GREM model of spin glass theory. Rapport interne no 267, École Polytechnique.

[25] Pfaffelhuber, P. and Wakolbinger, A. (2006) The process of most recent common ancestors in an evolving coalescent. *Stochastic Process. Appl.* **116**, no. 12, 1836–1859.

[26] Pitman, J. (1999) Coalescents with multiple collisions. *Ann. Probab.* **27**, no. 4, 1870–1902.

[27] Sagitov, S. (1999) The general coalescent with asynchronous mergers of ancestral lines *J. Appl. Prob.* **36**, no. 4, 1116–1125.

[28] Sato, K. and Yamazato, M. (1984) Operator-selfdecomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type. *Stochastic Process. Appl.* **17**, no. 1, 73–100.

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

[30] Schweinsberg, J. (2000) Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5**, no. 12, 50pp.

[31] Siegmund, D. (1976) The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probability* **4**, no. 6, 914–924.

Article I

Asymptotics of continuous-time discrete state space branching processes for large initial state

Möhle, M. and Vetter, B. (2021) Markov Process. Related Fields 27, no. 1, 1–42.

Abstract.

Scaling limits for continuous-time branching processes with discrete state space are provided as the initial state tends to infinity. Depending on the finiteness or non-finiteness of the mean and/or the variance of the offspring distribution, the limits are in general time-inhomogeneous Gaussian processes, time-inhomogeneous generalized Ornstein–Uhlenbeck type processes or continuous-state branching processes. We also provide transfer results showing how specific asymptotic relations for the probability generating function of the offspring distribution carry over to those of the one-dimensional distributions of the branching process.

Keywords: Branching process; generalized Mehler semigroup; Neveu's continuous-state branching process; Ornstein–Uhlenbeck type process; self-decomposability; stable law; time-inhomogeneous process; weak convergence

2020 Mathematics Subject Classification: Primary 60J80; 60F05 Secondary 60F17; 60G50; 60J27

I.1 Introduction

Suppose that the lifetime of each individual in some population is exponentially distributed with a given parameter $a \in (0, \infty)$ and that at the end of its life each individual gives birth to $k \in \mathbb{N}_0 := \{0, 1, \ldots\}$ individuals with probability p_k, independently of the rest of the population. Assuming that the population consists of $n \in \mathbb{N} := \{1, 2, \ldots\}$ individuals at time $t = 0$ we denote with $Z_t^{(n)}$ the random number of individuals alive at time $t \geq 0$. The process $Z^{(n)} := (Z_t^{(n)})_{t \geq 0}$ is a classical continuous-time branching process with discrete state space $\mathbb{N}_0 \cup \{\infty\}$ and initial state $Z_0^{(n)} = n$. These processes have been studied extensively in the literature. For fundamental properties of these processes we refer the reader to the classical books of Harris [18, Chapter V], Athreya and Ney [3, Chapter III] and Sewastjanow [39]. Define $Z_t := Z_t^{(1)}$ and $Z := Z^{(1)}$ for convenience. By the branching property, $Z^{(n)}$ is distributed as the sum of n independent copies of Z. The literature thus mainly focuses on the situation $n = 1$ and most results focus on the asymptotic behavior of these processes as the time t tends to infinity.

In contrast we are interested in the asymptotic behavior of $Z^{(n)}$ as the initial state n tends to infinity. To the best of the authors knowledge this question has not been discussed rigorously in the literature for continuous-time discrete state space branching processes. Related questions for discrete-time Bienaymé–Galton–Watson processes have been studied extensively in the literature (see, for example, Lamperti [25, 26] or Green [17]), however in this situation time is usually scaled as well, so these approaches differ from the continuous-time case. The article of Sagitov [38] contains related results, however the critical case is considered and again an additional time scaling is used.

The asymptotics as the initial state n tends to infinity may in some sense be viewed as a non-natural question in branching process theory, however this question has fundamental applications, for example in coalescent theory. It is well known that the block counting process of any exchangeable coalescent, restricted to a sample of size n, has a Siegmund dual process, called the fixation line. For the Bolthausen–Sznitman coalescent the fixation line is (see, for example, [23]) a continuous-time discrete state space branching process $Z^{(n)}$ with offspring distribution $p_k = 1/(k(k-1))$, $k \in \{2, 3 \ldots\}$. In this context, the parameter n is the sample size and hence the question about its asymptotic behavior when the sample size n becomes large is natural and important. In fact, this example was the starting point to become interested in the asymptotical behavior of branching processes for large initial value.

The convergence results are provided in Section I.2. We provide a convergence result for the finite variance case (Theorem I.2.1), another result for the situation when the process has still finite mean but infinite variance (Theorem I.2.3) and for the situation when even the mean is infinite but the process still does not explode in finite time (Theorem I.2.8). The limiting processes arising in Theorem I.2.1 are (time-inhomogeneous) Gaussian processes whereas those in Theorem I.2.3 are (time-inhomogeneous) Ornstein–Uhlenbeck type processes. In Theorem I.2.8 continuous-state branching processes arise in the limit as $n \to \infty$. For all three regimes typical examples are provided. The basic idea to obtain convergence results of this form is relatively obvious. For fixed time t, since $Z_t^{(n)}$ is a sum of n independent copies of Z_t, we can apply central limit theorems, leading to the convergence of the one-dimensional distributions. We refer the reader exemplary to the books of Petrov [31, 32] and Ibragimov and Linnik [20] and the article of Geluk and De Haan [15] for classical limiting results on sums of independent and identically distributed random variables. However, we prove not only convergence of the marginals or the finite-dimensional distributions. We provide functional limiting results for the sequence of processes $(Z^{(n)})_{n\in\mathbb{N}}$. Their proofs require some additional efforts. We think that the arising limiting processes are quite interesting. For example, since the centering or scaling of the space in Theorem I.2.1 and Theorem I.2.3 in general explicitly depends on the time t, the limiting processes are in general time-inhomogeneous.

The convergence results are as well based on crucial transfer results showing how particular asymptotic relations for the probability generating function (pgf) of the offspring distribution carry over to the pgf of Z_t. Results of this form are for example provided in Lemma I.2.2, Lemma I.2.6 and Lemma I.2.7 and are of its own interest. Despite the fact that there is a vast literature on continuous-time branching processes, we have not been able to trace these results.

Throughout the article ξ denotes a random variable taking values in $\mathbb{N}_0$ with probability $p_k := \mathbb{P}(\xi = k)$, $k \in \mathbb{N}_0$. For a space E equipped with a σ-algebra we denote with $B(E)$ the space of all bounded measurable functions $g : E \to \mathbb{R}$. For a topological space X and $K \in \{\mathbb{R}, \mathbb{C}\}$ we denote by $\widehat{C}(X, K)$ the space of continuous functions $g : X \to K$ vanishing at infinity and also write $\widehat{C}(X)$ for $\widehat{C}(X, \mathbb{R})$.

I.2 Results

Let f denote the pgf of ξ, i.e., $f(s) := \mathbb{E}(s^\xi) = \sum_{k\geq 0} p_k s^k$ and define $u(s) := a(f(s) - s)$ for $s \in [0, 1]$. Let $r \geq 1$. It is well known (see, for example, Athreya and Ney [3, p. 111, Corollary 1]) that $m_r(t) := \mathbb{E}(Z_t^r) < \infty$ for all $t > 0$ if and only if $\mathbb{E}(\xi^r) = \sum_{k\geq 0} k^r p_k < \infty$. Moreover $m(t) := m_1(t) = e^{\lambda t}$ with $\lambda := u'(1-) = a(\mathbb{E}(\xi) - 1)$ and

$$m_2(t) = \begin{cases} \tau^2 \lambda^{-1} e^{\lambda t}(e^{\lambda t} - 1) + e^{\lambda t} & \text{if } \lambda \neq 0, \\ \tau^2 t + 1 & \text{if } \lambda = 0, \end{cases} \qquad \text{(I.1)}$$

with $\tau^2 := u''(1-) = af''(1-) = a\mathbb{E}(\xi(\xi-1))$. Note that (I.1) slightly corrects Eq. (5) on p. 109 in [3], which accidently provides the formula for the second descending factorial moment $\mathbb{E}(Z_t(Z_t-1))$ instead of the second moment $\mathbb{E}(Z_t^2)$. In particular, if $m_2(t) < \infty$, then

$$\sigma^2(t) \;:=\; \mathrm{Var}(Z_t) \;=\; \begin{cases} (\tau^2-\lambda)e^{\lambda t}(e^{\lambda t}-1)/\lambda & \text{if } \lambda \neq 0, \\ \tau^2 t & \text{if } \lambda = 0. \end{cases}$$

I.2.1 The finite variance case

Assume that the second moment $\mathbb{E}(\xi^2) = \sum_{k\geq 0} k^2 p_k$ of the offspring distribution is finite or, equivalently, that $\mathrm{Var}(Z_t) < \infty$ for all $t \geq 0$. In the following $a \wedge b := \min\{a,b\}$ denotes the minimum of $a, b \in \mathbb{R}$. We furthermore use for $\mu \in \mathbb{R}$ and $\sigma^2 \geq 0$ the notation $N(\mu,\sigma^2)$ for the normal distribution with mean μ and variance σ^2 with the convention that $N(\mu,0)$ is the Dirac measure at μ. Our first fluctuation result (Theorem I.2.1) clarifies the asymptotic behavior of $Z_t^{(n)}$ as the initial state n tends to infinity. The proof of Theorem I.2.1 is provided in Section I.3.

Theorem I.2.1. *If $\mathbb{E}(\xi^2) < \infty$ or, equivalently, if $\sigma^2(t) := \mathrm{Var}(Z_t) < \infty$ for all $t \geq 0$, then, as $n \to \infty$, the process $X^{(n)} := (X_t^{(n)})_{t\geq 0}$, defined via*

$$X_t^{(n)} \;:=\; \frac{Z_t^{(n)} - nm(t)}{\sqrt{n}} \;=\; \frac{Z_t^{(n)} - ne^{\lambda t}}{\sqrt{n}}, \qquad n \in \mathbb{N}, t \geq 0, \qquad \text{(I.2)}$$

converges in $D_{\mathbb{R}}[0,\infty)$ to a continuous Gaussian Markov process $X = (X_t)_{t\geq 0}$ with $X_0 = 0$ and covariance function $(s,t) \mapsto \mathrm{Cov}(X_s, X_t) = \mathbb{E}(X_s X_t) = m(|s-t|)\sigma^2(s \wedge t)$, $s, t \geq 0$.

Remarks.

1. (Continuity of X) Let $s, t \geq 0$ and $x \in \mathbb{R}$. Conditional on $X_s = x$ the random variable $X_{s+t} - X_s$ has a normal distribution with mean $\mu := xm(t) - x = x(m(t)-1)$ and variance $v^2 := m(s)\sigma^2(t)$. Thus, $\mathbb{E}((X_{s+t}-X_s)^4 \,|\, X_s = x) = 3v^4 + 6\mu^2 v^2 + \mu^4 = 3m^2(s)\sigma^4(t) + 6x^2(m(t)-1)^2 m(s)\sigma^2(t) + x^4(m(t)-1)^4$ or, equivalently,

$$\mathbb{E}((X_{s+t}-X_s)^4 \,|\, X_s) \;=\; 3m^2(s)\sigma^4(t) + 6X_s^2(m(t)-1)^2 m(s)\sigma^2(t) + X_s^4(m(t)-1)^4.$$

Taking expectation yields

$$\begin{aligned}\mathbb{E}((X_{s+t}-X_s)^4) &= 3m^2(s)\sigma^4(t)+6\mathbb{E}(X_s^2)(m(t)-1)^2m(s)\sigma^2(t)\\ &\quad+\ \mathbb{E}(X_s^4)(m(t)-1)^4\\ &= 3m^2(s)\sigma^4(t)+6\sigma^2(s)(m(t)-1)^2m(s)\sigma^2(t)\\ &\quad+\ 3\sigma^4(s)(m(t)-1)^4.\end{aligned}$$

From this formula it follows that for every $T>0$ there exists a constant $K=K(T)\in(0,\infty)$ such that $\mathbb{E}((X_s-X_t)^4)\le K(s-t)^2$ for all $s,t\in[0,T]$. By Kolmogorov's continuity theorem (see, for example, Kallenberg [22, p. 57, Theorem 3.23]) we can therefore assume that X has continuous paths.

2. (Generator) For $\lambda\neq 0$ the Gaussian process X is time-inhomogeneous. Note that $T_{s,t}g(x):=\mathbb{E}(g(X_{s+t})\,|\,X_s=x)=\mathbb{E}(g(xm(t)+\sqrt{m(s)}X_t))$, $s,t\ge 0$, $g\in B(\mathbb{R})$, $x\in\mathbb{R}$. Let $C^2(\mathbb{R})$ denote the space of real-valued twice continuously differentiable functions on $\mathbb{R}$. For $s\ge 0$, $g\in C^2(\mathbb{R})$ and $x\in\mathbb{R}$ it follows that

$$A_sg(x)\ :=\ \lim_{t\to 0}\frac{T_{s,t}g(x)-g(x)}{t}\ =\ \lambda xg'(x)\ +\ \frac{\sigma^2}{2}m(s)g''(x),$$

where $\sigma^2:=\lim_{t\to 0}\sigma^2(t)/t=\tau^2-\lambda=a\mathbb{E}((\xi-1)^2)$. For $\lambda=0$ (critical case) the process X is a time-homogeneous Brownian motion with generator $Ag(x)=(\tau^2/2)g''(x)$, $g\in C^2(\mathbb{R})$, $x\in\mathbb{R}$, where $\tau^2=a\mathrm{Var}(\xi)$.

3. (Doob–Meyer decomposition) Define the process $C:=(C_t)_{t\ge 0}$ via $C_t:=\lambda\int_0^t X_s\,\mathrm{d}s$, $t\ge 0$. Let $\mathcal{F}_t:=\sigma(X_s,s\le t)$, $t\ge 0$. For all $0\le s\le t$,

$$\begin{aligned}\mathbb{E}(C_t-C_s\,|\,\mathcal{F}_s) &= \lambda\mathbb{E}\bigg(\int_s^t X_u\,\mathrm{d}u\,\bigg|\,\mathcal{F}_s\bigg)\ =\ \lambda\int_s^t\mathbb{E}(X_u\,|\,\mathcal{F}_s)\,\mathrm{d}u\\ &= \lambda\int_s^t m(u-s)X_s\,\mathrm{d}u\ =\ X_s\int_s^t\lambda e^{\lambda(u-s)}\,\mathrm{d}u\\ &= X_s(e^{\lambda(t-s)}-1)\ =\ X_sm(t-s)-X_s\\ &= \mathbb{E}(X_t\,|\,\mathcal{F}_s)-X_s\ =\ \mathbb{E}(X_t-X_s\,|\,\mathcal{F}_s).\end{aligned}$$

Thus, the compensated process $M:=(M_t)_{t\ge 0}:=(X_t-C_t)_{t\ge 0}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\ge 0}$. For $\lambda=0$ the process X itself is hence a martingale. Clearly, $X=M+C$ is the Doob–Meyer decomposition of X. The process C is not monotone, but decomposes into $C=C^+-C^-$, where $C^+:=$

$(C_t^+)_{t\geq 0}$ and $C^- := (C_t^-)_{t\geq 0}$, defined via $C_t^+ := \lambda \int_0^t X_s^+ \, \mathrm{d}s$ and $C_t^- := \lambda \int_0^t X_s^- \, \mathrm{d}s$ for all $t \geq 0$, both have non-decreasing paths.

4. (Positive semi-definiteness) The limiting process X in Theorem I.2.1 is Gaussian. For any finite number k of time points $0 \leq t_1 < \cdots < t_k < \infty$ it follows that $(X_{t_1}, \ldots, X_{t_k})$ has a multivariate normal distribution with positive semi-definite covariance matrix $\Sigma := (\sigma_{i,j})_{i,j\in\{1,\ldots,k\}}$ having entries $\sigma_{i,j} = \mathrm{Cov}(X_{t_i}, X_{t_j}) = m(|t_i - t_j|)\sigma^2(t_i \wedge t_j)$, $i, j \in \{1, \ldots, k\}$. For $\lambda = 0$ (critical case) it follows that the matrix $(t_i \wedge t_j)_{i,j\in\{1,\ldots,k\}}$ is positive semi-definite. For further properties of such min and max matrices and related meet and join matrices we refer the reader exemplary to Bhatia [4, 5] and Mattila and Haukkanen [27, 28]. For $\lambda \neq 0$ (non-critical case) it follows that the matrix $(e^{\lambda|t_i - t_j|} e^{\lambda(t_i\wedge t_j)}(e^{\lambda(t_i\wedge t_j)} - 1)/\lambda)_{i,j\in\{1,\ldots,k\}}$ is positive semi-definite.

Examples. (i) Let ξ be geometrically distributed with parameter $p \in (0, 1)$. Define $q := 1 - p$. Then all descending factorial moments $\mathbb{E}((\xi)_j) = j!(q/p)^j$, $j \in \mathbb{N}_0$, are finite. Theorem I.2.1 is hence applicable with $\lambda = a(\mathbb{E}(\xi) - 1) = a(q/p - 1)$ and $\tau^2 = a\mathbb{E}((\xi)_2) = 2a(q/p)^2$. For $p = 1/2$ (critical case) the process X is a Brownian motion with generator $Af(x) = af''(x)$, $f \in C^2(\mathbb{R})$, $x \in \mathbb{R}$.

(ii) If ξ is Poisson distributed with parameter $\mu \in (0, \infty)$, then again all descending factorial moments $\mathbb{E}((\xi)_j) = \mu^j$, $j \in \mathbb{N}_0$, are finite. Theorem I.2.1 is applicable with $\lambda = a(\mathbb{E}(\xi) - 1) = a(\mu - 1)$ and $\tau^2 = a\mathbb{E}((\xi)_2) = a\mu^2$. For $\mu = 1$ (critical case) the process X is a Brownian motion with generator $Af(x) = (a/2)f''(x)$, $f \in C^2(\mathbb{R})$, $x \in \mathbb{R}$.

(iii) Let $a_1, a_2 \geq 0$ with $a_1 + a_2 > 0$. Theorem I.2.1 is applicable for birth and death processes with rates na_1 and na_2 for birth and death respectively if the process is in state n. In this case we have $a = a_1 + a_2$, $f(s) = (a_2 + a_1 s^2)/a$, $u(s) = a_2 + a_1 s^2 - as$, $\lambda = a_1 - a_2$ and $\tau^2 = 2a_1$. For $a_1 = a_2$ (critical case) the process X is a Brownian motion with generator $Af(x) = a_1 f''(x)$, $f \in C^2(\mathbb{R})$, $x \in \mathbb{R}$.

I.2.2 The finite mean infinite variance case

In this subsection it is assumed that $m := \mathbb{E}(\xi) < \infty$. Since f is convex on $[0, 1]$, the inequality $1 - f(s) \leq m(1 - s)$ holds for all $s \in [0, 1]$. In order to state appropriate limiting results it is usual to control the difference between $m(1 - s)$ and $1 - f(s)$. A typical assumption of this form is the following.

Assumption A. There exists a constant $\alpha \in (1,2]$ and a function $L : [1,\infty) \to (0,\infty)$ slowly varying (at infinity) such that

$$1 - f(s) \;=\; m(1-s) \;-\; (1-s)^{\alpha} L((1-s)^{-1}), \qquad s \in [0,1). \quad \text{(I.3)}$$

Since f is differentiable, Assumption A in particular implies that L is differentiable. Define $F(s,t) := \mathbb{E}(s^{Z_t})$ for $s \in [0,1]$ and $t \geq 0$. The following lemma clarifies the structure of $F(s,t)$ under Assumption A. Recall that $m(t) := \mathbb{E}(Z_t) = e^{\lambda t} < \infty$.

Lemma I.2.2. *If the offspring pgf f satisfies Assumption A, then, for every $t \geq 0$,*

$$1 - F(s,t) \;=\; m(t)(1-s) \;-\; c(t)(1-s)^{\alpha} L((1-s)^{-1})(1+o(1)), \quad \text{(I.4)}$$

$s \to 1-$*, where*

$$c(t) \;:=\; \begin{cases} at & \text{if } \lambda = 0, \\ \dfrac{m(\alpha t) - m(t)}{(\alpha-1)(m-1)} \;=\; ae^{\lambda t}\dfrac{e^{\lambda(\alpha-1)t} - 1}{(\alpha-1)\lambda} & \text{if } \lambda \neq 0. \end{cases} \quad \text{(I.5)}$$

Remark. Although we are in this subsection mainly interested in the infinite variance case, Lemma I.2.2 holds in particular for the finite variance case. In this case expansion of f for $s \to 1-$ shows that (I.3) holds with $\alpha = 2$ and $L((1-s)^{-1}) \sim f''(1-)/2 = \mathbb{E}(\xi(\xi-1))/2$ as $s \to 1-$. Moreover, $c(t)f''(1-) = \mathbb{E}(Z_t(Z_t-1)) = F''(1-,t)$, where $F''(s,t)$ denotes the second derivative of $F(s,t)$ with respect to s.

In the following we are however interested in the infinite variance situation, so we assume that $\mathbb{E}(\xi^2) = \infty$. We are now able to state our second main convergence result.

Theorem I.2.3. *Assume that $m := \mathbb{E}(\xi) < \infty$ and $\mathbb{E}(\xi^2) = \infty$. Suppose that Assumption A holds. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers satisfying $a_n \geq 1$ for all sufficiently large n and $L(a_n) \sim a_n^{\alpha}/(\alpha n)$ as $n \to \infty$. Then the process $X^{(n)} := (X_t^{(n)})_{t\geq 0}$, defined via*

$$X_t^{(n)} \;:=\; \frac{Z_t^{(n)} - nm(t)}{a_n}, \qquad n \in \mathbb{N}, t \geq 0,$$

converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to a limiting process $X = (X_t)_{t\geq 0}$ with state space $\mathbb{R}$ and initial state $X_0 = 0$, whose distribution is characterized as follows. Conditional on $X_s = x$ the random variable X_{s+t} is distributed as $xm(t) + (m(s))^{1/\alpha}X_t$, where X_t is α-stable with *characteristic function* $u \mapsto \mathbb{E}(e^{iuX_t}) = \exp(c(t)(-iu)^\alpha/\alpha)$, $s,t \geq 0$, $u \in \mathbb{R}$, *and Laplace transform* $\eta \mapsto \mathbb{E}(e^{-\eta X_t}) = \exp(c(t)\eta^\alpha/\alpha)$, $\eta, t \geq 0$. *Note that* $\mathbb{E}(X_t) = 0$, $t \geq 0$. *The variance of* X_t *is equal to* $c(t)$ *for* $\alpha = 2$ *whereas* $\mathrm{Var}(X_t) = \infty$ *for* $t > 0$ *and* $\alpha \in (1,2)$.

Remark. As in Theorem I.2.1 the limiting process X in Theorem I.2.3 is time-homogeneous if and only if $\lambda = 0$. We have $T_{s,t}g(x) := \mathbb{E}(g(X_{s+t}) \,|\, X_s = x) = \mathbb{E}(g(xm(t) + (m(s))^{1/\alpha}X_t))$ for $s,t \geq 0$, $g \in B(\mathbb{R})$ and $x \in \mathbb{R}$. Note that $T_{s,t}g(x)$ is well-defined even for some functions g which are not bounded. For example, for Laplace test functions of the form $g = g_\eta$, defined via $g_\eta(x) := e^{-\eta x}$ for all $x \in \mathbb{R}$ and $\eta \geq 0$, we obtain the explicit formula

$$\begin{aligned}
A_s g_\eta(x) &:= \lim_{t \to 0} \frac{T_{s,t}g_\eta(x) - g_\eta(x)}{t} = \lim_{t \to 0} \frac{e^{-m(t)\eta x + c(t)m(s)\eta^\alpha/\alpha} - e^{-\eta x}}{t} \\
&= \lim_{t \to 0} \Big(- m'(t)\eta x + c'(t)m(s)\frac{\eta^\alpha}{\alpha}\Big) e^{-m(t)\eta x + c(t)m(s)\eta^\alpha/\alpha} \\
&= \Big(- m'(0+)\eta x + c'(0+)m(s)\frac{\eta^\alpha}{\alpha}\Big) e^{-\eta x} \\
&= \Big(- \lambda \eta x + am(s)\frac{\eta^\alpha}{\alpha}\Big) e^{-\eta x}, \qquad s, \eta \geq 0, x \in \mathbb{R}. \qquad \text{(I.6)}
\end{aligned}$$

For $\alpha = 2$ and $g \in C^2(\mathbb{R})$ it follows from (I.6) that

$$A_s g(x) := \lim_{t \to 0} \frac{T_{s,t}g(x) - g(x)}{t} = \lambda x g'(x) + \frac{a}{2} m(s) g''(x),$$

$s \geq 0, x \in \mathbb{R}$ showing that for $\alpha = 2$ the process X has the same structure as in Theorem I.2.1 with σ^2 replaced by the constant a.

Now assume that $\alpha \in (1,2)$. Then, from (I.6), a straightforward calculation based on the formula

$$\int_0^\infty \frac{e^{-\eta h} - 1 + \eta h}{h^{\alpha+1}} \,\mathrm{d}h = \frac{\Gamma(2-\alpha)}{\alpha(\alpha-1)} \eta^\alpha = \Gamma(-\alpha)\eta^\alpha,$$

$\eta \geq 0, \alpha \in (1,2)$, yields

$$A_s g(x) = \lambda x g'(x) + am(s)\frac{\alpha - 1}{\Gamma(2-\alpha)} \int_0^\infty \frac{g(x+h) - g(x) - hg'(x)}{h^{\alpha+1}} \,\mathrm{d}h,$$

$s \geq 0, x \in \mathbb{R}$, first for $g = g_\eta$ and, hence, for other classes of functions g, for example for $g \in C_c^2(\mathbb{R})$. These formulas for the semigroup and the generator show that X is a time-inhomogeneous Ornstein–Uhlenbeck type process [40]. For fundamental results on such processes and related generalized Mehler semigroups we refer the reader to [8].

For $\alpha = 2$ we have $a_n^\alpha/n \sim \alpha L(a_n) \to \infty$ as $n \to \infty$, in contrast to the situation in Theorem I.2.1, where $a_n = \sqrt{n}$ and, hence, $a_n^2/n = 1$. For $\alpha = 2$ the limiting random variable X_t has a normal distribution with mean 0 and variance $c(t)$ given via (I.5) with $\alpha = 2$.

Two examples are provided, one with $\alpha = 2$ and the other with $\alpha \in (1, 2)$. In the first example the underlying branching process is supercritical whereas in the second example it is critical. In the first example $F(s, t)$ can be expressed in terms of the Lambert W function. In the second example $F(s, t)$ is known explicitly.

Example I.2.4. Suppose that $p_k = 4/((k-1)k(k+1))$ for $k \in \{2, 3, \ldots\}$, i.e., $f(s) = \sum_{k=2}^\infty p_k s^k = 2s^{-1}(1-s)^2(-\log(1-s)) - 2 + 3s$, $s \in (0, 1)$. Note that (I.3) holds with $\alpha = 2$, $m := \mathbb{E}(\xi) = 3$, $L(1) := 2$ and $L(x) := 2(\log x)/(1 - 1/x)$ for $x > 1$. Clearly, $L(x) \sim 2 \log x$ as $x \to \infty$. Moreover, $\lambda = 2a$, $m(t) := \mathbb{E}(Z_t) = e^{2at}$ and $\mathrm{Var}(Z_t) = \infty$ for $t > 0$. The sequence $(a_n)_{n\in\mathbb{N}}$, defined via $a_1 := 1$ and $a_n := \sqrt{2n \log n}$ for $n \in \mathbb{N} \setminus \{1\}$, satisfies $L(a_n) \sim 2 \log a_n \sim \log n = a_n^2/(2n)$ as $n \to \infty$. By Theorem I.2.3, the process $((Z_t^{(n)} - ne^{2at})/\sqrt{2n \log n})_{t\ge0}$ converges in $D_\mathbb{R}[0, \infty)$ as $n \to \infty$ to a time-inhomogeneous process $X = (X_t)_{t\ge0}$ with distribution as described in Theorem I.2.3. In particular, for every $t > 0$, the random variable X_t has a normal distribution with mean 0 and variance $c(t) = \frac{1}{2}e^{2at}(e^{2at} - 1)$. The pgf $F(., t)$ of Z_t is computed as follows. From the backward equation

$$
\begin{aligned}
t &= \int_s^{F(s,t)} \frac{1}{u(x)}\,\mathrm{d}x \\
&= \frac{1}{a}\int_s^{F(s,t)} \frac{x}{2(1-x)((x-1)\log(1-x) - x)}\,\mathrm{d}x \;=\; \frac{1}{2a}[v(x)]_s^{F(s,t)}
\end{aligned}
$$

with $v(x) := \log(1 - x) - \log(x + (1 - x) \log(1 - x))$, $x \in (0, 1)$, we conclude that

$$F(s, t) \;=\; v^{-1}(2at + v(s)), \tag{I.7}$$

where $v^{-1} : \mathbb{R} \to (0, 1)$ denotes the inverse of v, which turns out to be of the form $v^{-1}(y) = 1 + 1/W(h)$, where $h := -\exp(-1 - e^{-y}) \in (-1/e, 0)$ and $W = W_{-1}$ denotes the lower branch of the Lambert W function satisfying $W(h)e^{W(h)} = h$ and being real-valued on $[-1/e, 0)$. Expansion of (I.7) shows that

$$F(s, t) = 1-e^{2at}(1-s)+e^{2at}(e^{2at}-1)(1-s)^2 \log((1-s)^{-1})+O((1-s)^2),$$

$s \to 1-$, in agreement with (I.4), since $c(t) = \frac{1}{2}e^{2at}(e^{2at} - 1)$ and $L(x) \sim 2 \log x$ as $x \to \infty$.

Example I.2.5. Let $\alpha \in (1,2)$ and $b \in (0,1/\alpha]$. Assume that $f(s) = s + b(1-s)^\alpha$, $s \in [0,1]$. Note that $p_0 = b$, $p_1 = 1 - b\alpha$ and $p_k = b(-1)^k\binom{\alpha}{k}$ for $k \in \{2,3,\ldots\}$. In particular, $p_k \sim b/(\Gamma(-\alpha)k^{\alpha+1})$ as $k \to \infty$. Moreover, $f'(s) = 1 - b\alpha(1-s)^{\alpha-1}$ and, therefore, $m := \mathbb{E}(\xi) = f'(1-) = 1$. Thus, the underlying branching process is critical, the extinction probability is $q = 1$ and (I.3) holds with $L \equiv b$. Note that $u(s) = ab(1-s)^\alpha$. Theorem I.2.3 is applicable with $a_n := n^{1/\alpha}$. It follows that $(n^{-1/\alpha}(Z_t^{(n)} - n))_{t\ge 0}$ converges in $D_\mathbb{R}[0,\infty)$ as $n \to \infty$ to a process X with distribution as described in Theorem I.2.3. In particular, for every $t \ge 0$, the random variable X_t has characteristic function $u \mapsto \exp(-abt(-iu)^\alpha)$, $u \in \mathbb{R}$. From

$$\begin{aligned} at &= \int_s^{F(s,t)} \frac{1}{f(x)-x}\,\mathrm{d}x \;=\; \int_s^{F(s,t)} \frac{1}{b(1-x)^\alpha}\,\mathrm{d}x \\ &= \frac{(1-F(s,t))^{1-\alpha} - (1-s)^{1-\alpha}}{b(\alpha-1)} \end{aligned}$$

it follows that $F(s,t) = 1 - ((\alpha-1)abt + (1-s)^{1-\alpha})^{1/(1-\alpha)}$. Generating functions of this form can be traced back at least to Zolotarev [47, Section 5]. Note that

$$1-F(s,t) \;=\; (1-s) - abt(1-s)^\alpha + \frac{\alpha}{2}(abt)^2(1-s)^{2\alpha-1} + O((1-s)^{3\alpha-2}),$$

$s \to 1-$, in agreement with (I.4), since $c(t) = at$ and $L \equiv b$.

I.2.3 The infinite mean case with non-explosion

In this subsection it is assumed that $m := \mathbb{E}(\xi) = \infty$ or, equivalently, that $m(t) := \mathbb{E}(Z_t) = \infty$ for all $t > 0$. In order to state the result it is convenient to define the function $L : [1,\infty) \to (0,\infty)$ via

$$L(x) \;:=\; x(1 - f(1-x^{-1})), \qquad x \ge 1. \tag{I.8}$$

The substitution $s = 1 - x^{-1}$ shows that this definition is equivalent to

$$1 - f(s) \;=\; (1-s)L((1-s)^{-1}), \qquad s \in [0,1). \tag{I.9}$$

Non-explosion is assumed throughout this subsection, which is equivalent to (see, for example, Harris [18, Chapter V, Section 9, p. 106, Theorem 9.1])

$$\int_\varepsilon^1 \frac{1}{s-f(s)}\,\mathrm{d}s \;=\; \int_{(1-\varepsilon)^{-1}}^\infty \frac{1}{x(L(x)-1)}\,\mathrm{d}x \;=\; \infty$$

for all $\varepsilon \in (q,1)$, where q denotes the extinction probability. For the theory of stable distributions and their domains of attraction we refer the reader to Geluk and de Haan [15]. For the moment let $t > 0$ be fixed. Then $Z_t^{(n)}$, suitably normalized, converges in distribution as $n \to \infty$ to a non-degenerate limit, i.e., Z_t is in the domain of attraction of a stable law if and only if the following condition is satisfied. There exists $\alpha(t) \in (0,1]$ and a slowly varying function $L_t : [1,\infty) \to (0,\infty)$ such that

$$\mathbb{P}(Z_t > x) \;\sim\; x^{-\alpha(t)} L_t(x), \qquad x \to \infty. \tag{I.10}$$

And, if $\alpha(t) = 1$, then $L_t(x) \to \infty$ as $x \to \infty$. In this subsection only the case $\alpha(t) < 1$ is investigated. Recall that $F(s,t) = \mathbb{E}(s^{Z_t})$ for $s \in [0,1]$ and $t \geq 0$. It follows from Bingham and Doney [6] that (I.10) is then equivalent to

$$1 - F(s,t) \;=\; (1-s)^{\alpha(t)} L_t((1-s)^{-1}), \qquad s \in [0,1), \tag{I.11}$$

where, to be precise, the function L_t of (I.11) replaces $\Gamma(1-\alpha(t))L_t$. Then,

$$\alpha(t) \;=\; \frac{\log \frac{1-F(s,t)}{L_t((1-s)^{-1})}}{\log(1-s)}, \qquad t \geq 0, s \in [0,1). \tag{I.12}$$

Since L_t is slowly varying and hence satisfies $\log L_t(x)/\log x \to 0$ as $x \to \infty$, it follows from (I.12) that

$$\alpha(t) \;=\; \lim_{s \to 1-} \frac{\log(1 - F(s,t))}{\log(1-s)}, \qquad t \geq 0. \tag{I.13}$$

In particular, $\alpha(t)$ is uniquely determined by the pgf $F(.,t)$. Note that (I.11) always holds for $t = 0$ with $\alpha(0) = 1$ and $c(0) = 1$ because of the boundary condition $F(s,0) = s$.

Suppose (I.11) holds for all $t \geq 0$. From the iteration formula $F(s,t+u) = F(F(s,t),u)$ it follows that

$$\begin{aligned}
&(1-s)^{\alpha(t+u)} L_{t+u}((1-s)^{-1}) \\
&\quad = \; 1 - F(s,t+u) \; = \; 1 - F(F(s,t),u) \\
&\quad = \; (1 - F(s,t))^{\alpha(u)} L_u((1-F(s,t))^{-1}) \\
&\quad = \; (1-s)^{\alpha(t)\alpha(u)} L_t^{\alpha(u)}((1-s)^{-1}) L_u((1-s)^{-\alpha(t)} L_t^{-1}((1-s)^{-1})),
\end{aligned}$$

$s \in [0,1)$. Since all terms depending on $L_.$ are slowly varying, $\alpha(.)$ has to be multiplicative, i.e., $\alpha(t+u) = \alpha(t)\alpha(u)$ for all $t,u \geq 0$. The map $k : [0,\infty) \to [0,\infty)$, defined via $k(t) := -\log \alpha(t)$ for all $t \geq 0$, is hence additive, so it satisfies the Cauchy functional equation. By Aczel [1, p. 34, Theorem 1], $k(t) = Ct$ and, hence, $\alpha(t) = e^{-Ct}$ for all $t \geq 0$, where $C := k(1) = -\log \alpha(1) \in [0,\infty)$. Clearly, either

$\alpha(t) = 1$ for all $t \geq 0$ or $\alpha(t) < 1$ for all $t > 0$, depending on whether $C = 0$ or $C > 0$. Also, the map $t \mapsto L_t(x)$, $t \geq 0$, is continuously differentiable and satisfies

$$L_{t+u}((1-s)^{-1}) \;=\; L_t^{\alpha(u)}((1-s)^{-1})L_u((1-s)^{-\alpha(t)}L_t^{-1}((1-s)^{-1})),$$

$t, u \geq 0, s \in [0,1)$, or $L_{t+u}(x) = L_t^{\alpha(u)}(x)L_u(x^{\alpha(t)}L_t^{-1}(x))$ for all $t, u \geq 0$ and all $x \geq 1$. The following result (Lemma I.2.6) relates (I.11) to the offspring pgf f. The map $s \mapsto L((1-s)^{-1}) = \frac{1-f(s)}{1-s}$, $s \in [0,1]$, has derivative $s \mapsto \frac{1}{1-s}(\frac{1-f(s)}{1-s} - f'(s))$, which is strictly positive on $[0,1)$, since f is strictly convex. Thus, L is strictly increasing on $[1,\infty)$. We also have $L(x) \to \infty$ as $x \to \infty$, since $m = \infty$. The proof of Lemma I.2.6 is provided in Section I.5.

Lemma I.2.6. *If $m := f'(1-) = \infty$, then the following conditions are equivalent.*

(i) For every $t > 0$ there exists $\alpha(t) \in (0,1)$ and a slowly varying function $L_t : [1,\infty) \to (0,\infty)$ such that (I.11) holds.

(ii) For every $t > 0$ the limit

$$\alpha(t) \;:=\; \lim_{s\to 1-} \alpha(s,t) \;\in\; (0,1)$$

exists, where $\alpha(s,t) := (1-s)(\frac{\partial}{\partial s}F(s,t))/(1-F(s,t))$ for all $s \in [0,1)$.

(iii) The limit

$$A \;:=\; \lim_{x\to\infty} \frac{L(x)}{\log x} \;=\; \lim_{s\to 1-} \frac{1-f(s)}{(1-s)\log((1-s)^{-1})} \qquad \text{(I.14)}$$

exists in $(0,\infty)$.

In this case $\alpha(t) = e^{-aAt}$ for all $t \geq 0$.

Remark. Note that

$$\begin{aligned} aA \;&=\; a\lim_{s\to 1+} \frac{f(s)-1}{(1-s)\log(1-s)} \;=\; \lim_{s\to 1-} \frac{u(s)-a(1-s)}{(1-s)\log(1-s)} \\ &=\; \lim_{s\to 1-} \frac{u(s)}{(1-s)\log(1-s)}. \end{aligned}$$

Thus, $\alpha(t) = e^{-aAt}$ can be alternatively computed from the function u.

Suppose that $m=\infty$ and that the limit $A:=\lim_{x\to\infty} L(x)/\log x \in (0,\infty)$ in Lemma I.2.6 exists. Recall that, by Lemma I.2.6, the existence of the limit A is equivalent to the existence of constants $\alpha(t)\in(0,1)$ and of slowly varying functions L_t such that (I.11) holds, i.e., $1-F(s,t)=(1-s)^{\alpha(t)}L_t((1-s)^{-1})$. In the following we focus on the particular situation that the limit

$$\beta(t) := \lim_{x\to\infty} L_t(x) = \lim_{s\to 1-} L_t((1-s)^{-1}) = \lim_{s\to 1-} \frac{1-F(s,t)}{(1-s)^{\alpha(t)}} \quad \text{(I.15)}$$

exists in $(0,\infty)$ for each $t\geq 0$. We already know that $\alpha(t)=e^{-aAt}$. If (I.15) holds, then we must have $A>0$, since otherwise $\alpha(t)=1$ and hence $\beta(t)=m(t)=\infty$, in contradiction to (I.15). The following result relates (I.15) to the offspring pgf f and provides an explicit formula for $\beta(t)$. The proof of Lemma I.2.7 is provided in Section I.5.

Lemma I.2.7. *Suppose that $m=\infty$ and that (I.14) holds. If the limit $B:=\lim_{x\to\infty}(L(x)-A\log x)\in\mathbb{R}$ exists, then (I.15) holds for all $t\geq 0$. In this case*

$$\beta(t) \;=\; \exp\left(\frac{B-1}{A}(1-\alpha(t))\right), \qquad t\geq 0. \quad \text{(I.16)}$$

We are now able to provide the third main convergence result. In the following the notation $E:=[0,\infty)$ is used.

Theorem I.2.8. *Suppose that $m=\infty$ and let L be defined via (I.8) such that (see (I.9)) the relation $1-f(s)=(1-s)L((1-s)^{-1})$ holds for all $s\in[0,1)$. Assume that both limits*

$$A \;:=\; \lim_{x\to\infty}\frac{L(x)}{\log x} \;\in\; (0,\infty) \quad \textit{and} \quad B \;:=\; \lim_{x\to\infty}(L(x)-A\log x) \;\in\; \mathbb{R}$$

exist. For $t\geq 0$ define

$$\alpha(t) \;:=\; e^{-aAt} \quad \textit{and} \quad \beta(t) \;:=\; \exp\left(\frac{B-1}{A}(1-\alpha(t))\right). \quad \text{(I.17)}$$

Then, as $n\to\infty$, the scaled process $X^{(n)}:=(X_t^{(n)})_{t\geq 0}$, defined via

$$X_t^{(n)} \;:=\; n^{-1/\alpha(t)}Z_t^{(n)}, \qquad t\geq 0,$$

converges in $D_E[0,\infty)$ to a limiting continuous-state branching process $X=(X_t)_{t\geq 0}$, whose distribution is characterized as follows.

(i) For every $t\geq 0$ the marginal random variable X_t is $\alpha(t)$-stable with Laplace transform $\lambda\mapsto\exp(-\beta(t)\lambda^{\alpha(t)})$, $\lambda\geq 0$.

(ii) The semigroup $(T_t)_{t\geq 0}$ of X satisfies $T_tg(x) = \mathbb{E}(g(x^{1/\alpha(t)}X_t))$, $x, t \geq 0$, $g \in B(E)$, i.e., conditional on $X_s = x$ the random variable X_{s+t} has the same distribution as $x^{1/\alpha(t)}X_t$.

The proof of Theorem I.2.8 is provided in Section I.5. We now provide three examples. In the first two examples the distribution of Z_t is known explicitly.

Example I.2.9. Assume that ξ has distribution $p_k := \mathbb{P}(\xi = k) := 1/(k(k-1))$, $k \in \{2, 3, \ldots\}$. Note that $\xi = \lfloor X \rfloor$, where X has density $x \mapsto 1/(x-1)^2$, $x \geq 2$, so X has a shifted Pareto distribution with parameter 1. Then, $f(s) = s+(1-s)\log(1-s) = 1-(1-s)L((1-s)^{-1})$ with $L(x) := 1 + \log x$ and $u(s) := a(f(s) - s) = a(1-s)\log(1-s)$. Note that $A := \lim_{x\to\infty} L(x)/\log x = 1$ and $B := \lim_{x\to\infty}(L(x) - \log x) = 1$. From the backward equation $(\partial/\partial t)F(s,t) = u(F(s,t))$ it follows that

$$\begin{aligned} t &= \int_s^{F(s,t)} \frac{1}{u(x)}\,\mathrm{d}x = \frac{1}{a}\left[-\log(-\log(1-x))\right]_s^{F(s,t)} \\ &= \frac{1}{a}\log\Big(\frac{\log(1-s)}{\log(1-F(s,t))}\Big). \end{aligned}$$

Thus, $F(s,t) = 1-(1-s)^{e^{-at}}$ showing that Z_t is Sibuya distributed (see, for example, Christoph and Schreiber [10, Eq. (2)]) with parameter e^{-at}. The Sibuya distribution and similar distributions occur for example in Gnedin [16, p. 84, Eq. (9)], Huillet and Möhle [19, p. 9], Iksanov and Möhle [21, p. 225] and Pitman [33, p. 84, Eq. (18)], [34, p. 70, Eq. (3.38)]. We conclude that (I.15) holds with $\alpha(t) := e^{-at}$ and $\beta(t) := 1$. By Theorem I.2.8, as $n \to \infty$, the scaled process $X^{(n)} := (Z_t^{(n)}/n^{e^{at}})_{t\geq 0}$ converges in $D_E[0,\infty)$ to a limiting process $X = (X_t)_{t\geq 0}$ such that X_t has Laplace transform $\lambda \mapsto \exp(-\lambda^{e^{-at}})$, $\lambda \geq 0$, and the semigroup $(T_t)_{t\geq 0}$ of X satisfies $T_tg(x) = \mathbb{E}(g(x^{e^{at}}X_t))$, $x, t \geq 0$, $g \in B(E)$. We identify $(X_{t/a})_{t\geq 0}$ as Neveu's continuous-state branching process [29]. For $a = 1$ this example coincides with [23, Theorem 2.1 b)] stating that the fixation line of the Bolthausen–Sznitman n-coalescent, properly scaled, converges as $n \to \infty$ to Neveu's continuous-state branching process.

Example I.2.10. Example I.2.9 is easily generalized as follows. Fix two constants $b > 0$ and $c \geq 0$ with $b + c \leq 1$ and assume that $p_0 := c$, $p_1 := 1 - b - c$ and $p_k := b/(k(k-1))$ for $k \geq 2$. Then, $f(s) = s + (1-s)(c + b\log(1-s)) = 1-(1-s)(1-c-b\log(1-s))$, $u(s) = a(f(s) - s) = a(1-s)(c + b\log(1-s))$ and $L(x) = 1 - c +$

$b\log x$. For $b = 1$ and $c = 0$ we are back in Example I.2.9. Note that $A := \lim_{x\to\infty} L(x)/\log x = b > 0$ and $B := \lim_{x\to\infty}(L(x) - b\log x) = 1-c \in (0,1]$. The same argument as in Example I.2.9 leads to $F(s,t) = 1-(1-s)^{e^{-abt}}\exp(cb^{-1}(e^{-abt}-1))$. Thus, Theorem I.2.8 is applicable with $\alpha(t) := e^{-abt}$ and $\beta(t) := \exp(cb^{-1}(e^{-abt}-1))$, $t \geq 0$. Clearly, these formulas for $\alpha(t)$ and $\beta(t)$ are in agreement with those from Lemma I.2.6 and Lemma I.2.7, namely $\alpha(t) = e^{-aAt} = e^{-abt}$ and $\beta(t) = \exp((B-1)A^{-1}(1-\alpha(t))) = \exp(cb^{-1}(e^{-abt}-1))$, $t \geq 0$.

Example I.2.11. (Discrete Luria–Delbrück distribution) Assume that ξ has a discrete Luria–Delbrück distribution with parameter $b \in (0,\infty)$, i.e., $f(s) = (1-s)^{b(1-s)/s}$, $s \in (0,1)$. Note that $f(0) = e^{-b}$ and $f(s) = 1-(1-s)L((1-s)^{-1})$ for $s \in [0,1)$, where $L(1) := 1-e^{-b}$ and $L(x) := x(1-x^{b/(1-x)})$ for $x \in (1,\infty)$. Note that $A := \lim_{x\to\infty} L(x)/\log x = b$ and $B := \lim_{x\to\infty}(L(x) - b\log x) = 0$. Let $q = q(b)$ denote the extinction probability, i.e., the smallest fixed point of f in the interval $[0,1]$. For all $\varepsilon \in (q,1)$,

$$\int_\varepsilon^1 \frac{1}{s-f(s)}\,\mathrm{d}s \;=\; \int_{(1-\varepsilon)^{-1}}^\infty \frac{1}{x(L(x)-1)}\,\mathrm{d}x \;=\; \infty,$$

since $L(x) \sim b\log x$ as $x \to \infty$. By the explosion criterion, the associated branching process $Z = (Z_t)_{t\geq 0}$ does not explode. The functions $\alpha(.)$ and $\beta(.)$ are obtained as follows. By Lemma I.2.6, $\alpha(t) = e^{-aAt} = e^{-abt}$, $t \geq 0$. Furthermore,

$$\beta(t) \;=\; \exp\left(\frac{B-1}{A}(1-\alpha(t))\right) \;=\; \exp\left(\frac{e^{-abt}-1}{b}\right), \qquad t \geq 0.$$

By Theorem I.2.8, as $n \to \infty$, the scaled process $X^{(n)} := (Z_t^{(n)}/n^{e^{abt}})_{t\geq 0}$ converges in $D_E[0,\infty)$ to a limiting process $X = (X_t)_{t\geq 0}$ such that X_t has Laplace transform $\lambda \mapsto \exp(-\beta(t)\lambda^{e^{-abt}})$, $\lambda \geq 0$, and the semigroup $(T_t)_{t\geq 0}$ of X satisfies $T_t g(x) = \mathbb{E}(g(x^{e^{abt}}X_t))$, $x,t \geq 0$, $g \in B(E)$.

The previous three examples are summarized in the following table.

Example	Example I.2.9	Example I.2.10	Example I.2.11
Parameters	—	$b>0$, $c\geq 0$, $b+c\leq 1$	$0<b<\infty$
pgf $f(s)$	$s+(1-s)\log(1-s)$	$s+(1-s)(c+b\log(1-s))$	$(1-s)^{b(1-s)/s}$
$L(x)$	$1+\log x$	$1-c+b\log x$	$x(1-x)^{b/(1-x)}$
$\alpha(t)$	e^{-at}	e^{-abt}	e^{-abt}
$\beta(t)$	1	$\exp(cb^{-1}(e^{-abt}-1))$	$\exp((e^{-abt}-1)/b)$

Remark. Theorem I.2.8 does not cover the situation when the limit $A := \lim_{x\to\infty} L(x)/\log x$ is either 0 or ∞. We leave the analysis of the two boundary cases $A = 0$ and $A = \infty$ for future work, but provide two concrete examples.

Example I.2.12. An example satisfying $A = 0$ (and $\mathbb{E}(\xi) = \infty$) is obtained as follows. Define $L(1) := 1$, $L(x) := 1 + \log\log x - \log(1 - 1/x)$ for $x > 1$ and $f(s) := 1-(1-s)L((1-s)^{-1})$ for $s \in [0, 1)$. Clearly, $L(x) \sim \log\log x$ as $x \to \infty$. Hence, $A := \lim_{x\to\infty} L(x)/\log x = 0$. In the following it is clarified that f is a pgf. It is not hard to check that the function $g : [0, 1) \to \mathbb{R}$, defined via $g(0) := 0$ and $g(s) := \log(-\log(1-s))-\log s$ for $s \in (0, 1)$, has the Taylor expansion $g(s) = \sum_{n\ge1} g_n s^n$ with coefficients $g_n := (n!n)^{-1} \int_0^1 [x]_n \,\mathrm{d}x$, $n \in \mathbb{N}$, where $[x]_n := x(x+1)\cdots(x+n-1)$, i.e., $g_1 = 1/2$, $g_2 = 5/24$, $g_3 = 1/8$, and so on. Thus, $f(s) = 1-(1-s)(1+g(s)) = s-(1-s)g(s)$ has the Taylor expansion $f(s) = \sum_{n\ge1} p_n s^n$ with coefficients $p_1 = 1 - g_1 = 1/2$ and

$$\begin{aligned} p_n \;&=\; g_{n-1} - g_n \;=\; \frac{1}{(n-1)(n-1)!}\int_0^1 [x]_{n-1}\,\mathrm{d}x \;-\; \frac{1}{nn!}\int_0^1 [x]_n\,\mathrm{d}x \\ &=\; \frac{1}{n!}\int_0^1 [x]_{n-1}\Big(\frac{1-x}{n} + \frac{1}{n-1}\Big)\,\mathrm{d}x, \qquad n \in \mathbb{N}\setminus\{1\}. \end{aligned}$$

In particular, $p_n > 0$ for all $n \in \mathbb{N}$. Thus, f is the pgf of some (offspring) random variable ξ taking values in $\mathbb{N}$. Note that $\mathbb{E}(\xi) = \infty$, since $\lim_{x\to\infty} L(x) = \infty$. From $L(x) \sim \log\log x$ as $x \to \infty$ it follows that the associated continuous-time branching process $Z = (Z_t)_{t\ge0}$ does not explode.

The asymptotics of p_n as $n \to \infty$ is obtained as follows. It is easily seen that $f''(s) \sim (1 - s)^{-1}\ell((1 - s)^{-1})$ as $s \to 1-$, where $\ell(u) := 1/\log u$. Moreover, the sequence

$$\begin{aligned} a_n \;&:=\; [s^n]f''(s) \;=\; (n+1)(n+2)p_{n+2} \\ &=\; \frac{1}{n!}\int_0^1 [x]_{n+1}\Big(\frac{1-x}{n+2} + \frac{1}{n+1}\Big)\,\mathrm{d}x, \quad n \in \mathbb{N}_0, \end{aligned}$$

is strictly decreasing, since, by straightforward calculations,

$$a_{n-1} - a_n \;=\; \frac{1}{n!(n+2)}\int_0^1 [x]_n(1-x)(2-x)\,\mathrm{d}x \;>\; 0, \qquad n \in \mathbb{N}.$$

From Karamata's Tauberian theorem for power series (apply, for example, Bingham, Goldie and Teugels [7, p. 40, Corollary 1.7.3] with $A := f''$, $c := 1$ and $\rho := 1$) it follows that $a_n = [s^n]f''(s) \sim \ell(n) = 1/\log n$ as $n \to \infty$. Thus, $p_n \sim 1/(n^2 \log n)$ as $n \to \infty$.

Example I.2.13. A fruitful example satisfying $A = \infty$ is the following. Define $L(x) := 1 + (1 + \log x)\log(1 + \log x)$ for $x \ge 1$ and $f(s) :=$

$1-(1-s)L((1-s)^{-1})$, i.e.,

$$f(s) \;=\; s \;-\; (1-s)(1-\log(1-s))\log(1-\log(1-s)), \qquad s \in [0,1).$$

Clearly, $f(1-) = 1 =: f(1)$ and $f'(s) = (-\log(1-s))\log(1-\log(1-s))$, $s \in [0,1)$. By Lemma I.6.5 provided in the appendix, f' is absolutely monotone and $f(0) = f'(0) = f''(0) = 0$, which implies that f is the pgf of some (offspring) random variable ξ taking values in $\{3,4,\ldots\}$. Note that $A = \infty$ implies $\lim_{x\to\infty} L(x) = \infty$, which is equivalent to $\mathbb{E}(\xi) = \infty$. Nevertheless, the associated continuous-time branching process $Z = (Z_t)_{t\ge 0}$ does not explode. The pgf $F(.,t)$ of Z_t is even explicitly known. More precisely, solving the backward equation

$$\begin{aligned} at \;&=\; \int_{F(s,t)}^{s} \frac{1}{u-f(u)}\,\mathrm{d}u \\ &=\; \int_{F(s,t)}^{s} \frac{1}{(1-u)(1-\log(1-u))\log(1-\log(1-u))}\,\mathrm{d}u \\ &=\; [\log(\log(1-\log(1-u)))]_{F(s,t)}^{s} \;=\; \log\frac{\log(1-\log(1-s))}{\log(1-\log(1-F(s,t)))} \end{aligned}$$

yields the solution $F(s,t) = 1-\exp(1-(1-\log(1-s))^{e^{-at}})$, $s \in [0,1)$, $t \ge 0$. In particular, for each $\alpha \in (0,1)$, the map $s \mapsto 1-\exp(1-(1-\log(1-s))^{\alpha})$, $s \in [0,1)$, is a pgf, which does not seem to be straightforward to verify directly.

I.2.4 The explosive case

We briefly comment on the situation when the branching process may explode in finite time. Note that explosion implies that $A := \lim_{x\to\infty} L(x)/\log x = \infty$. Thus, Theorem I.2.8 is not applicable. We have $F(1,t) < 1$ for all $t > 0$. For $t \ge 0$ let $G(.,t)$ denote the pgf of Z_t conditioned on $Z_t < \infty$, i.e.,

$$G(s,t) \;:=\; \frac{F(s,t)}{F(1,t)}, \qquad s \in [0,1], t \ge 0,$$

In this situation a convergence result in the spirit of the previous theorems, but with F replaced by G, is obtained as follows. For $t > 0$ we have $\mathbb{E}(Z_t \,|\, Z_t < \infty) = G'(1-,t) = F'(1-,t)/F(1,t) = \infty$. Thus, it is natural to assume that $1-G(s,t) = (1-s)^{\alpha(t)}L_t((1-s)^{-1})$ for some $\alpha(t) \in (0,1]$ and some slowly varying function L_t. Now assume furthermore that the limits

$$\beta(t) \;:=\; \lim_{x\to\infty} L_t(x) \;\in\; (0,\infty), \qquad t \ge 0,$$

exist. Then $\alpha(t) < 1$ for all $t > 0$. Now, for $t \geq 0$ and $n \in \mathbb{N}$, choose $a_n(t)$ such that $L_t(a_n(t)) \sim (a_n(t))^{\alpha(t)}/(n\alpha(t))$ as $n \to \infty$. Then $Z_t^{(n)}/a_n(t)$, conditioned on $Z_t < \infty$, converges to X_t in distribution as $n \to \infty$, where X_t has Laplace transform $\lambda \mapsto \exp(-\beta(t)\lambda^{\alpha(t)})$, $\lambda \geq 0$. Example I.2.14 below, going back at least to Sewastjanow [39, Chapter 1, Section 8, Example 6], turns out to be in that regime.

Example I.2.14. Suppose that ξ is Sibuya distributed with parameter $\alpha \in (0,1)$, i.e., $f(s) = 1-(1-s)^\alpha$, $s \in [0,1]$. Note that f has the Taylor expansion $f(s) = \sum_{n\geq 1} p_n s^n$ with coefficients

$$p_n \;:=\; \binom{\alpha}{n}(-1)^{n-1} \;=\; \frac{\alpha}{\Gamma(1-\alpha)}\frac{\Gamma(n-\alpha)}{\Gamma(1-\alpha)}, \qquad n \in \mathbb{N}.$$

In particular, $p_n \sim (\alpha/\Gamma(1-\alpha))n^{-\alpha-1}$ as $n \to \infty$. Moreover, $f(s) = 1-(1-s)R((1-s)^{-1})$, where $R(x) := x^{1-\alpha}$ is regularly varying of index $1-\alpha$. The backward equation

$$\begin{aligned} at \;&=\; \int_s^{F(s,t)} \frac{1}{f(x)-x}\,\mathrm{d}x \;=\; \int_s^{F(s,t)} \frac{1}{1-x-(1-x)^\alpha}\,\mathrm{d}x \\ &=\; \left[\frac{-\log(1-(1-x)^{1-\alpha})}{1-\alpha}\right]_s^{F(s,t)} \\ &=\; \frac{1}{1-\alpha}\log\frac{1-(1-s)^{1-\alpha}}{1-(1-F(s,t))^{1-\alpha}}, \qquad t \geq 0, \end{aligned}$$

yields the explicit solution (see [39, p. 26, Eq. (19)])

$$F(s,t) \;=\; 1 \;-\; \Big(1-e^{-(1-\alpha)at}(1-(1-s)^{1-\alpha})\Big)^{\frac{1}{1-\alpha}}, \qquad \text{(I.18)}$$

$s \in [0,1], t \geq 0$. We have $\mathbb{P}(Z_t = \infty) = 1-F(1,t) = (1-e^{-(1-\alpha)at})^{\frac{1}{1-\alpha}}$ for $t \geq 0$, so $0 < \mathbb{P}(Z_t = \infty) < 1$ for all $t > 0$. The time $T := \inf\{t > 0 : Z_t = \infty\}$ of explosion satisfies $\mathbb{P}(T < \infty) = \lim_{t\to\infty}\mathbb{P}(Z_t = \infty) = 1$, so Z explodes in finite time almost surely. Note that T has mean

$$\begin{aligned} \mathbb{E}(T) \;&=\; \int_0^\infty \mathbb{P}(T > t)\,\mathrm{d}t \;=\; \int_0^\infty \mathbb{P}(Z_t < \infty)\,\mathrm{d}t \\ &=\; \int_0^\infty (1-(1-e^{-(1-\alpha)at})^{\frac{1}{1-\alpha}})\,\mathrm{d}t. \end{aligned}$$

The substitution $x = 1-e^{-(1-\alpha)at}$ yields

$$\mathbb{E}(T) \;=\; \frac{1}{a(1-\alpha)}\int_0^1 \frac{1-x^{\frac{1}{1-\alpha}}}{1-x}\,\mathrm{d}x \;=\; \frac{1}{a(1-\alpha)}\bigg(\Psi\Big(\frac{2-\alpha}{1-\alpha}\Big)+\gamma\bigg),$$

where $\Psi = \Gamma'/\Gamma$ denotes the logarithmic derivative of the gamma function and γ is the Euler–Mascheroni constant.

Let $t > 0$ in the following. Expansion of (I.18) yields

$$F(s,t) = F(1,t) - \frac{1}{1-\alpha}(1-e^{-(1-\alpha)at})^{\frac{\alpha}{1-\alpha}} e^{-(1-\alpha)at}(1-s)^{1-\alpha} + O((1-s)^{2(1-\alpha)}), \quad s \to 1-. \tag{I.19}$$

Rewriting (I.19) in the form

$$\begin{aligned} 1 - G(s,t) &= 1 - \frac{F(s,t)}{F(1,t)} \\ &= \frac{(1-e^{-(1-\alpha)at})^{\frac{\alpha}{1-\alpha}} e^{-(1-\alpha)at}}{(1-\alpha)(1-(1-e^{-(1-\alpha)at})^{\frac{1}{1-\alpha}})}(1-s)^{1-\alpha} \\ &\quad + O((1-s)^{2(1-\alpha)}), \qquad s \to 1-, \end{aligned}$$

yields $\alpha(t) = 1 - \alpha$ for all $t > 0$ and

$$\beta(t) := \lim_{x\to\infty} L_t(x) = \frac{(1-e^{-(1-\alpha)at})^{\frac{\alpha}{1-\alpha}} e^{-(1-\alpha)at}}{(1-\alpha)(1-(1-e^{-(1-\alpha)at})^{\frac{1}{1-\alpha}})}, \qquad t > 0.$$

Thus, the sequence $a_n(t) := (n\alpha(t)\beta(t))^{1/\alpha(t)}$ satisfies $L_t(a_n(t)) \sim (a_n(t))^{\alpha(t)}/(n\alpha(t))$ as $n \to \infty$ and it follows that $X_t^{(n)} := Z_t^{(n)}/a_n(t)$, conditioned on $Z_t < \infty$, converges to X_t in distribution as $n \to \infty$, where X_t has Laplace transform $\lambda \mapsto \exp(-\beta(t)\lambda^{\alpha(t)})$, $\lambda \geq 0$.

We leave the study of further examples of branching processes with explosion similar to those of Example I.2.14 to the interested reader. One may for instance study the pgf $f(s) := \frac{2}{\pi}\arcsin s$, $s \in [0,1]$, occurring in Pakes [30, p. 276, Example 4.5]. A further example is the offspring distribution $p_k = \frac{\sqrt{\pi}}{4}\Gamma(k)/\Gamma(k+3/2)$, $k \in \mathbb{N}$, in which case the offspring pgf has the form $f(s) = 1 - \sqrt{(1-s)/s}\arcsin\sqrt{s}$.

Let us finally discuss the situation when

$$1 - G(s,t) = (1-s)L_t((1-s)^{-1}), \qquad t \geq 0, \tag{I.20}$$

for some slowly varying function L_t. Note that (see, for example, Bingham and Doney [6, Theorem A]) (I.20) is equivalent to $\sum_{k=0}^n \mathbb{P}(Z_t > k \,|\, Z_t < \infty) \sim L_t(n)$ as $n \to \infty$, which is Condition (ii) in Rogozin's relative stability theorem (see, for example, Bingham, Goldie and Teugels [7, Theorem 8.8.1]). Let $(a_n(t))_{n\in\mathbb{N}}$ be a sequence such that $L_t(a_n(t)) \sim a_n(t)/n$ as $n \to \infty$. Then, by Theorem 8.8.1 of [7], $Z_t^{(n)}/a_n(t)|_{Z_t<\infty} \to 1$ in probability as $n \to \infty$.

Thus, in this situation we cannot have a non-degenerate limit. The following example fits into this regime. In this example the limits

$$\gamma(t) \;:=\; \lim_{x\to\infty} \frac{L_t(x)}{\log x} \;\in\; (0,\infty), \qquad t\ge 0,$$

exist.

Example I.2.15. Define $f(0) := 0$, $f(1) := 1$ and

$$f(s) \;:=\; 1 \;+\; \frac{s}{\log(1-s)}, \qquad s\in(0,1).$$

It is easily seen that f has the Taylor expansion $f(s) = \sum_{n\ge 1} p_n s^n$ with positive coefficients

$$p_n \;:=\; (-1)^{n-1}\int_0^1 \binom{x}{n}\,\mathrm{d}x \;=\; \frac{1}{n!}\int_0^1 x\frac{\Gamma(n-x)}{\Gamma(1-x)}\,\mathrm{d}x \;>\; 0, \qquad n\in\mathbb{N}.$$

Thus, f is the pgf of some random variable ξ taking values in $\mathbb{N}$. Note that $p_n = (-1)^{n-1}b_n/n!$ for all $n\in\mathbb{N}$, where $b_n := \int_0^1 (x)_n\,\mathrm{d}x$ denotes the n-th Bernoulli number of the second kind (see, e.g., Roman [37, p. 114]). Here $(x)_n := x(x-1)\cdots(x-n+1)$, $n\in\mathbb{N}$, denotes the n-th descending factorial of $x\in\mathbb{R}$. From $p_0 = 0$ it follows that the associated continuous-time branching process $Z = (Z_t)_{t\ge 0}$ has extinction probability $q = 0$. Note that $f(s) = 1-(1-s)R((1-s)^{-1})$, where $R(x) := (x-1)/\log x$, $x>1$, is regularly varying of index 1. For all $\varepsilon\in(q,1) = (0,1)$,

$$\begin{aligned}\int_\varepsilon^1 \frac{1}{s-f(s)}\,\mathrm{d}s \;&=\; \int_\varepsilon^1 \frac{1}{s-1-\frac{s}{\log(1-s)}}\,\mathrm{d}s\\ &=\; [\log(s+(1-s)\log(1-s))]_\varepsilon^1\\ &=\; -\log(\varepsilon+(1-\varepsilon)\log(1-\varepsilon)) \;<\; \infty,\end{aligned}$$

which shows that Z explodes. It is also known (see, for example, Flajolet and Sedgewick [14, p. 387]) that $p_n \sim 1/(n\log^2 n)$ as $n\to\infty$. Thus, p_n tends slower to 0 than in Example I.2.14. In this sense Z is strongly explosive. The backward equation is

$$\begin{aligned}at \;&=\; \int_s^{F(s,t)} \frac{1}{f(u)-u}\,\mathrm{d}u \;=\; \int_s^{F(s,t)} \frac{1}{1-u+\frac{u}{\log(1-u)}}\,\mathrm{d}u\\ &=\; [-\log(u+(1-u)\log(1-u))]_s^{F(s,t)}\\ &=\; \log\frac{s+(1-s)\log(1-s)}{F(s,t)+(1-F(s,t))\log(1-F(s,t))}\end{aligned}$$

or, equivalently,

$$\begin{aligned} F(s,t) \;+\; (1-F(s,t))\log(1-F(s,t)) \\ =\; e^{-at}(s+(1-s)\log(1-s)) \;=:\; h(s,t). \end{aligned}$$

It is straightforward to check that this equation has the solution

$$F(s,t) \;=\; 1 \;-\; \exp\left(1+W\left(\frac{h(s,t)-1}{e}\right)\right), \qquad s\in[0,1), t\geq 0,$$

where $W=W_{-1}$ denotes the lower branch of the Lambert W function satisfying $W(h)e^{W(h)}=h$ and being real-valued on $[-1/e,0)$. Note that $\mathbb{P}(Z_t=\infty)=1-F(1,t)=\exp(1+W((e^{-at}-1)/e))$ for $t\geq 0$, so $0<\mathbb{P}(Z_t=\infty)<1$ for $t>0$. The time $T:=\inf\{t>0:Z_t=\infty\}$ of explosion satisfies $\mathbb{P}(T<\infty)=\lim_{t\to\infty}\mathbb{P}(Z_t=\infty)=\exp(1+W(-1/e))=\exp(0)=1$, so Z explodes in finite time almost surely. Note that T has mean

$$\mathbb{E}(T)=\int_0^\infty \mathbb{P}(Z_t<\infty)\,\mathrm{d}t=\int_0^\infty\left(1-\exp\left(1+W\left(\frac{e^{-at}-1}{e}\right)\right)\right)\mathrm{d}t.$$

The substitution $x=1-e^{-at}$ ($\Rightarrow t=-\frac{1}{a}\log(1-x)$ and $\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{a(1-x)}$) leads to

$$\mathbb{E}(T) \;=\; \frac{1}{a}\int_0^1\frac{1-\exp(1+W(-x/e))}{1-x}\,\mathrm{d}x.$$

The function below the integral has a singularity at $x=1$. From $1+W(-x/e)\sim\sqrt{2(1-x)}$ as $x\to 1$ it follows that the function below the integral behaves asymptotically as $\sqrt{2/(1-x)}$ as $x\to 1-$, which yields $\mathbb{E}(T)<\infty$. Numerical computations show that $\mathbb{E}(T)\approx 2.45/a$.

Let $G(s,t):=F(s,t)/F(1,t)$ denote the pgf of Z_t conditioned on $Z_t<\infty$. A somewhat tedious but straightforward calculation shows that $1-G(s,t)=(1-s)L_t((1-s)^{-1})$, where L_t is slowly varying with

$$\gamma(t) \;:=\; \lim_{x\to\infty}\frac{L_t(x)}{\log x} \;=\; \frac{w}{(w+1)(1-(w+1)e^{at})}$$

with $w:=W(\frac{e^{-at}-1}{e})$. For $t\geq 0$ let $(a_n(t))_{n\in\mathbb{N}}$ be a sequence such that $L_t(a_n(t))\sim a_n(t)/n$ as $n\to\infty$. Then, as explained before, for every $t\geq 0$, conditional on $Z_t<\infty$, $Z_t^{(n)}/a_n(t)\to 1$ in probability as $n\to\infty$. A concrete sequence $(a_n(t))_{n\in\mathbb{N}}$ is $a_n(t):=\gamma(t)n\log n$, since, in this case, $L_t(a_n(t)) = L_t(\gamma(t)n\log n) \sim L_t(n\log n) \sim \gamma(t)\log(n\log n)\sim\gamma(t)\log n = a_n(t)/n$ as $n\to\infty$.

I.3 Proof of Theorem I.2.1

The proof of Theorem I.2.1 is quite natural and can be summarized as follows. An application of the multivariate central limit theorem yields the convergence of the finite-dimensional distributions. The convergence in $D_{\mathbb{R}}[0,\infty)$ is then established using a criterion of Aldous [2]. The following proof is relatively short and elegant.

Proof. (of Theorem I.2.1) Let us compute for $s,t\ge 0$ the covariance of Z_s and Z_{s+t}. For $k\in\mathbb{N}_0$,

$$\begin{aligned}
&\mathbb{E}((Z_s-m(s))(Z_{s+t}-m(s+t))\,|\,Z_s=k)\\
&\quad= (k-m(s))\mathbb{E}(Z_t^{(k)}-m(s+t))\\
&\quad= (k-m(s))(km(t)-m(s)m(t)) \;=\; m(t)(k-m(s))^2.
\end{aligned}$$

Thus, $\mathbb{E}((Z_s-m(s))(Z_{s+t}-m(s+t))\,|\,Z_s)=m(t)(Z_s-m(s))^2$ almost surely. Taking expectation yields $\mathrm{Cov}(Z_s,Z_{s+t})=m(t)\mathrm{Var}(Z_s)=m(t)\sigma^2(s)$.

In order to verify the convergence $X^{(n)}\stackrel{\mathrm{fd}}{\to}X$ of the finite-dimensional distributions fix $k\in\mathbb{N}$ and $0\le t_1<\cdots<t_k<\infty$, define the $\mathbb{R}^k$-valued random variable $Y:=(Z_{t_1}-m(t_1),\ldots,Z_{t_k}-m(t_k))$ and let $Y_1,Y_2,\ldots$ be independent copies of Y. By the branching property, $(X_{t_1}^{(n)},\ldots,X_{t_k}^{(n)})=((Z_{t_1}^{(n)}-nm(t_1))/\sqrt{n},\ldots,(Z_{t_k}^{(n)}-nm(t_k))/\sqrt{n})$ has the same distribution as $(Y_1+\cdots+Y_n)/\sqrt{n}$, which by the multivariate central limit theorem (see, for example, [44, p. 16, Example 2.18]) converges in distribution as $n\to\infty$ to a centered normal distribution $N(0,\Sigma)$ with covariance matrix $\Sigma=(\sigma_{i,j})_{1\le i,j\le k}$ having entries $\sigma_{i,j}:=\mathbb{E}((Z_{t_i}-m(t_i))(Z_{t_j}-m(t_j)))=\mathrm{Cov}(Z_{t_i},Z_{t_j})=m(|t_i-t_j|)\sigma^2(t_i\wedge t_j)$. Thus, the convergence $X^{(n)}\stackrel{\mathrm{fd}}{\to}X$ holds.

The convergence $X^{(n)}\to X$ in $D_{\mathbb{R}}[0,\infty)$ is achieved as follows. Define the processes $M^{(n)}:=(M_t^{(n)})_{t\ge 0}$, $n\in\mathbb{N}$, and $M:=(M_t)_{t\ge 0}$ via

$$M_t^{(n)} \;:=\; \frac{X_t^{(n)}}{m(t)} \;=\; \sqrt{n}\biggl(\frac{Z_t^{(n)}}{nm(t)}-1\biggr) \quad\text{and}\quad M_t \;:=\; \frac{X_t}{m(t)},$$

$n\in\mathbb{N},t\ge 0$ Then, $M,M^{(1)},M^{(2)},\ldots$ are martingales and M is continuous, since the Gaussian process X is continuous and $m(.)$ is continuous. Since $\mathbb{E}((M_t^{(n)})^2)=\mathrm{Var}(M_t^{(n)})=\mathrm{Var}(Z_t^{(n)})/(n(m(t))^2)=\sigma^2(t)/(m(t))^2<\infty$ does not depend on $n\in\mathbb{N}$, we conclude that, for each $t\ge 0$, the family $\{M_t^{(n)}:n\in\mathbb{N}\}$ is uniformly integrable. The convergence $M^{(n)} \to M$ in $D_{\mathbb{R}}[0,\infty)$ therefore follows from Aldous' criterion [2, Proposition 1.2]. Since the map $t \mapsto m(t)$ is continuous and deterministic it follows by multiplication with $m(t)$ that $X^{(n)} \to X$ in $D_{\mathbb{R}}[0,\infty)$. $\square$

I.4 Proofs concerning Theorem I.2.3

This section contains the proofs of Lemma I.2.2 and Theorem I.2.3.

Proof. (of Lemma I.2.2) The proof distinguishes the critical and non-critical case. Both cases are handled with different techniques. The representation in the critical case (for age-dependent branching processes) follows via an equivalence for the extinction probability from a combination of the results of Slack [41, Theorem 1] and Vatutin [45, Theorem 1]. The following more elementary proof (see Case 1) is based on the backward equation and does not use extinction probabilities.

Case 1. ($\lambda = 0$) Let $t \geq 0$. In the critical case the backward equation is

$$at \;=\; \int_s^{F(s,t)} \frac{1}{f(x)-x}\,\mathrm{d}x \;=\; \int_s^{F(s,t)} \frac{1}{(1-x)^\alpha L((1-x)^{-1})}\,\mathrm{d}x,$$

$s \in [0,1]$. Since the map $x \mapsto f(x) - x$ is non-negative and non-increasing on $[0,1]$ it follows that

$$\frac{F(s,t)-s}{(1-s)^\alpha L((1-s)^{-1})} \;\leq\; at \;\leq\; \frac{F(s,t)-s}{(1-F(s,t))^\alpha L((1-F(s,t))^{-1})}$$

and, hence,

$$\begin{aligned}\limsup_{s\to 1-} \frac{F(s,t)-s}{(1-s)^\alpha L((1-s)^{-1})} \;&\leq\; at\\ &\leq\; \liminf_{s\to 1-} \frac{F(s,t)-s}{(1-F(s,t))^\alpha L((1-F(s,t))^{-1})}\\ &=\; \liminf_{s\to 1-} \frac{F(s,t)-s}{(1-s)^\alpha L((1-s)^{-1})},\end{aligned}$$

where the last equality holds, since $1 - F(s,t) \sim 1 - s$ as $s \to 1-$. Thus, $\lim_{s\to 1-}(F(s,t)-s)/((1-s)^\alpha L((1-s)^{-1})) = at$.

Case 2. ($\lambda \neq 0$) Fix $t \geq 0$. Set $h_1(s) := (1-s)m(t) - (1-F(s,t))$ and $h_2(s) := (1-s)^\alpha L((1-s)^{-1})$ for $s \in [0,1)$. We have to verify that $\lim_{s\to 1-} h_1(s)/h_2(s) = c(t)$, where $c(t)$ is defined in (I.5). By the forward and backward equation, $h_1'(s) = -m(t) + \frac{\partial}{\partial s}F(s,t) =$

$-m(t) + (f(F(s,t)) - F(s,t))/(f(s) - s)$. Moreover, $h_2'(s) = (1 - s)^{\alpha-1}L((1-s)^{-1})(L'((1-s)^{-1})(1-s)^{-1}/L((1-s)^{-1}) - \alpha)$. From Assumption (I.3), the asymptotic relation $1 - F(s,t) \sim m(t)(1-s)$ as $s \to 1-$ and $(m(t))^\alpha = m(\alpha t)$ it follows that

$$\begin{aligned}
m(\alpha t) - m(t) &= \lim_{s\to 1-}\Bigg(\frac{(1-F(s,t))m - (1-f(F(s,t)))}{(1-s)^\alpha L((1-s)^{-1})} \\
&\quad - m(t)\frac{(1-s)m - (1-f(s))}{(1-s)^\alpha L((1-s)^{-1})}\Bigg) \\
&= \lim_{s\to 1-}\Bigg((1-m)\frac{(1-s)m(t) - (1-F(s,t))}{(1-s)^\alpha L((1-s)^{-1})} \\
&\quad + \frac{m(t)(1-f(s)-(1-s)) - (1-f(F(s,t))) + (1-F(s,t))}{(1-s)^\alpha L((1-s)^{-1})}\Bigg) \\
&= \lim_{s\to 1-}\Bigg((1-m)\frac{(1-s)m(t) - (1-F(s,t))}{(1-s)^\alpha L((1-s)^{-1})} \\
&\quad + \frac{-m(t)(f(s)-s) + f(F(s,t)) - F(s,t)}{(1-s)^\alpha L((1-s)^{-1})}\Bigg) \\
&= (m-1)\lim_{s\to 1-}\Bigg(\alpha\frac{-m(t)(f(s)-s) + (f(F(s,t)) - F(s,t))}{\alpha(m-1)(1-s)^\alpha L((1-s)^{-1})} \\
&\quad - \frac{(1-s)m(t) - (1-F(s,t))}{(1-s)^\alpha L((1-s)^{-1})}\Bigg) \\
&=: (m-1)\lim_{s\to 1-}\left(\alpha\frac{h_1'(s)}{h_2'(s) + R(s)} - \frac{h_1(s)}{h_2(s)}\right). \qquad \text{(I.21)}
\end{aligned}$$

Using

$$\frac{(1-m)(1-s)}{f(s)-s} = \frac{1-m}{1-m+(1-s)^{\alpha-1}L((1-s)^{-1})},$$

we see that $R(s)$ is given by

$$\begin{aligned}
R(s) &= -\alpha(1-s)^{\alpha-1}L((1-s)^{-1})\frac{1-m}{1-m+(1-s)^{\alpha-1}L((1-s)^{-1})} \\
&\quad - (1-s)^{\alpha-1}L((1-s)^{-1})\left(\frac{L'((1-s)^{-1})(1-s)^{-1}}{L((1-s)^{-1})} - \alpha\right) \\
&= \alpha(1-s)^{\alpha-1}L((1-s)^{-1})\Bigg(1 - \frac{1-m}{1-m+(1-s)^{\alpha-1}L((1-s)^{-1})} \\
&\quad - \frac{L'((1-s)^{-1})(1-s)^{-1}}{\alpha L((1-s)^{-1})}\Bigg).
\end{aligned}$$

In order to see that $\lim_{x\to\infty} xL'(x)/L(x) = 0$ we proceed as follows. Define $U : [1,\infty) \to (0,\infty)$ via $U(x) := m - x(1 - f(1-1/x)) =$

$x^{1-\alpha}L(x)$ for $x \geq 1$, where the last equality holds by (I.3). Note that $U(x) = \int_x^\infty u(y)\,\mathrm{d}y$, where $u : [1,\infty) \to (0,\infty)$ is defined via $u(x) := -U'(x) = 1 - f(1-1/x) - f'(1-1/x)/x$. The function u is non-increasing, since $u'(x) = -f''(1-1/x)/x^3 \leq 0$ by the convexity of f. From a variant of the monotone density theorem (see, for example, Bingham, Goldie and Teugels [7, p. 39, Theorem 1.7.2] and the comments thereafter) for integrals of the form $U(x) = \int_x^\infty u(y)\,\mathrm{d}y$ it follows that $u(x) \sim (\alpha-1)x^{-\alpha}L(x)$ as $x \to \infty$. Thus, $\lim_{x\to\infty} xU'(x)/U(x) = 1-\alpha$. Noting that $xU'(x)/U(x) = 1-\alpha+xL'(x)/L(x)$ we conclude that $\lim_{x\to\infty} xL'(x)/L(x) = 0$. Applying this relation with $x := (1-s)^{-1}$ yields

$$\begin{aligned}\lim_{s\to1-}\frac{R(s)}{h_2'(s)} &= \lim_{s\to1-}\frac{\alpha\left(1-\frac{1-m}{1-m+(1-s)^{\alpha-1}L((1-s)^{-1})}-\frac{L'((1-s)^{-1})(1-s)^{-1}}{\alpha L((1-s)^{-1})}\right)}{\frac{L'((1-s)^{-1})(1-s)^{-1}}{L((1-s)^{-1})}-\alpha}\\ &= 0. \qquad\qquad (I.22)\end{aligned}$$

The three quantities $h_1(s)$, $h_2(s)$ and $(m(\alpha t)-m(t))/(m-1)$ are non-negative, so, by (I.21), necessarily $\liminf_{s\to1-} h_1'(s)/(h_2'(s)+R(s)) \geq 0$, leading to the boundary $h_1'(s)/(h_2'(s)+R(s)) \geq (1-\delta)h_1'(s)/h_2'(s)$ for any $0<\delta<(\alpha-1)/\alpha$ and sufficiently large s. Then

$$\frac{m(\alpha t)-m(t)}{m-1} \geq \limsup_{s\to1-}\left(\alpha(1-\delta)\frac{h_1'(s)}{h_2'(s)}-\frac{h_1(s)}{h_2(s)}\right),$$

and the second part of Lemma I.6.2 provides

$$\limsup_{s\to1-}\frac{h_1(s)}{h_2(s)} \leq \frac{m(\alpha t)-m(t)}{m-1}. \qquad (I.23)$$

Now (I.21), (I.22) and (I.23) yield

$$\begin{aligned}\frac{m(\alpha t)-m(t)}{m-1} &= \lim_{s\to1-}\Bigg(\left(\alpha\frac{h_1'(s)}{h_2'(s)}-\frac{h_1(s)}{h_2(s)}\right)\frac{h_2'(s)}{h_2'(s)+R(s)}\\ &\qquad -\frac{h_1(s)}{h_2(s)}\frac{R(s)}{h_2'(s)+R(s)}\Bigg)\\ &= \lim_{s\to1-}\left(\alpha\frac{h_1'(s)}{h_2'(s)}-\frac{h_1(s)}{h_2(s)}\right).\end{aligned}$$

The claim follows again from Lemma I.6.2 in the appendix. Note that Lemma I.6.2 is applicable in both cases due to Lemma I.6.1. □

Proof. (of Theorem I.2.3) The proof is divided into four parts. The first part establishes the convergence of the one-dimensional distributions. The second and third part give two auxiliary results, one is about the normalizing sequence $(a_n)_{n\in\mathbb{N}}$ and the other is a kind of upper bound for the process, used in the final part to conclude the convergence in $D_{\mathbb{R}}[0,\infty)$.

Part 1. (Convergence of the one-dimensional distributions) Fix $t \in [0,\infty)$, define $Y := Z_t$ for convenience and let $Y_1, Y_2, \ldots$ be independent copies of Y. By Lemma I.2.2, Eq. (I.4) holds.

First assume that $\alpha \in (1,2)$. Then, by Bingham and Doney [6, Theorem A], applied with $n = 1$, Eq. (I.4) is equivalent to $\mathbb{P}(Y > x) \sim c(t)(-\Gamma(1-\alpha))^{-1}L(x)x^{-\alpha}$, $x \to \infty$. In particular, the map $x \mapsto \mathbb{P}(Y > x)$ is regularly varying (at infinity) of index $-\alpha$. By Theorem 1 (ii) $\Rightarrow$ (i) of Geluk and de Haan [15] (note that $p = 1$ since $Y \geq 0$), it follows that the cumulative distribution function of Y is in the domain of attraction of an α-stable distribution. The results on p. 174 in [15] on the choice of the normalizing sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ furthermore show that, if we choose $(a_n)_{n\in\mathbb{N}}$ such that $L(a_n) \sim a_n^\alpha/(\alpha n)$ as $n \to \infty$ and $b_n := n\mathbb{E}(Y)/a_n = nm(t)/a_n$, then $(Z_t^{(n)} - nm(t))/a_n \stackrel{d}{=} (Y_1 + \cdots + Y_n)/a_n - b_n \to X_t$ in distribution as $n \to \infty$, where X_t is α-stable with characteristic function $u \mapsto \exp(c(t)(-iu)^\alpha/\alpha)$, $u \in \mathbb{R}$. Thus, the convergence of the one-dimensional distributions holds.

The case $\alpha = 2$ is handled similarly by noting that (I.4) is then equivalent (see [6]) to $\mathbb{E}(1_{\{Y\leq x\}}Y^2) \sim 2c(t)L(x)$ as $x \to \infty$ such that we can apply Theorem 2 of [15].

Part 2. (Asymptotic relation for $(a_n)_{n\in\mathbb{N}}$) Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be an arbitrary sequence of positive real numbers converging to zero as $n \to \infty$. For $n \in \mathbb{N}$ and $T > 0$ define $S_{n,T} := [-\varepsilon_n n/a_n, \varepsilon_n n/a_n] \times [0,T]$, where $(a_n)_{n\in\mathbb{N}}$ is the normalizing sequence satisfying $a_n/(L(a_n))^{1/\alpha} \sim (\alpha n)^{1/\alpha}$ as $n \to \infty$. Bojanić and Seneta [9, p. 308] provide the existence of another slowly varying function L^* such that $a_n \sim (\alpha n)^{1/\alpha}L^*(n^{1/\alpha})$ as $n \to \infty$. Set $h(n) := (\alpha n)^{1/\alpha}L^*(n^{1/\alpha})/a_n$ for $n \in \mathbb{N}$ and $h(r) := h(\lfloor r \rfloor)$ for $r \in \mathbb{R}$, $r \geq 1$. Then the asymptotic relation simply means $\lim_{r\to\infty} h(r) = 1$. From

$$\lim_{n\to\infty} \inf_{(x,s)\in S_{n,T}} (nm(s) + xa_n) \;=\; \infty \tag{I.24}$$

it follows that $\sup_{(x,s)\in S_{n,T}} |h(nm(s) + xa_n) - 1| \to 0$ as $n \to \infty$. Furthermore, $\lim_{n\to\infty} \sup_{(x,s)\in S_{n,T}} |xa_n/n| \leq \lim_{n\to\infty} \varepsilon_n = 0$ implies $\lim_{n\to\infty} \sup_{(x,s)\in S_{n,T}} |(m(s) + xa_n/n)^{1/\alpha} - (m(s))^{1/\alpha}| = 0$ as well as, using the uniform convergence theorem for slowly varying functions (see, for example, Bingham, Goldie and Teugels [7, Theorem 1.2.1] or Bojanić and Seneta [9])

$$\lim_{n\to\infty} \sup_{(x,s)\in S_{n,T}} \left| \frac{L^*(n^{1/\alpha}(m(s)+xa_n/n)^{1/\alpha})}{L^*(n^{1/\alpha})} - 1 \right| = 0.$$

Having bounded limits, the listed uniformly convergent sequences are uniformly bounded and thus their product converges uniformly again, yielding

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{(x,s)\in S_{n,T}} \left| \frac{a_{nm(s)+xa_n}}{a_n} - (m(s))^{1/\alpha} \right| \\
&= \lim_{n\to\infty} \sup_{(x,s)\in S_{n,T}} \left| \frac{h(n)}{h(nm(s)+xa_n)} \frac{L^*((nm(s)+xa_n)^{1/\alpha})}{L^*(n^{1/\alpha})} \right. \\
&\qquad \left. \cdot \left(m(s) + \frac{xa_n}{n} \right)^{1/\alpha} - (m(s))^{1/\alpha} \right| \\
&= 0. \qquad\qquad (I.25)
\end{aligned}$$

Part 3. (Kind of upper bound for $X_t^{(n)}$) In this part it is shown that for each $T > 0$ there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of positive real numbers with $\lim_{n\to\infty} \varepsilon_n = 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{t\in[0,T]} |X_t^{(n)}| \geq \frac{\varepsilon_n n}{a_n} \right) = 0. \qquad (I.26)$$

Let $\delta := 0$ if $m < 1$ and $\delta := T$ if $m \geq 1$. Then, for any sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of positive real numbers,

$$\mathbb{P}\left(\sup_{t\in[0,T]} |X_t^{(n)}| \geq \frac{\varepsilon_n n}{a_n} \right) \leq \mathbb{P}\left(\sup_{t\in[0,T]} \left| \frac{a_n X_t^{(n)}}{m(t)} \right| \geq \frac{\varepsilon_n n}{m(\delta)} \right).$$

Applying Doob's submartingale inequality to the martingale $(a_n X_t^{(n)}/m(t))_{t\geq 0} = (Z_t^{(n)}/m(t) - n)_{t\geq 0}$ yields

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[0,T]} \left| \frac{a_n X_t^{(n)}}{m(t)} \right| \geq \frac{\varepsilon_n n}{m(\delta)} \right) &\leq \frac{m(\delta)}{\varepsilon_n n} \mathbb{E}\left(\left| \frac{Z_T^{(n)}}{m(T)} - n \right| \right) \\
&= \frac{m(\delta)}{m(T)} \frac{1}{\varepsilon_n} \mathbb{E}\left(\left| \frac{Z_T^{(n)}}{n} - m(T) \right| \right).
\end{aligned}$$

By the law of large numbers, the latter expectation converges to 0 as $n \to \infty$. Thus, the sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ can be chosen such that $\lim_{n\to\infty} \varepsilon_n = 0$ and such that the right-hand side still converges to 0, which implies that (I.26) holds for the particular sequence $(\varepsilon_n)_{n\in\mathbb{N}}$.

Part 4. (Convergence in $D_{\mathbb{R}}[0,\infty)$) In general, the processes $X^{(n)}$ and X are time-inhomogeneous. Let $Y^{(n)} := (X_t^{(n)}, t)_{t\ge 0}$ and $Y := (X_t, t)_{t\ge 0}$ denote the space-time processes of $X^{(n)}$ and X, respectively. According to Revuz and Yor [36, p. 85, Exercise (1.10)], the processes $Y^{(n)}$ and Y are time-homogeneous Markov processes with state space $S := \mathbb{R}\times[0,\infty)$. Recall that $S_{n,T} = [-\varepsilon_n n/a_n, \varepsilon_n n/a_n]\times[0,T]$, where $(\varepsilon_n)_{n\in\mathbb{N}}$ is the sequence used in Part 3. In terms of $Y^{(n)}$, (I.26) is simply

$$\lim_{n\to\infty} \mathbb{P}\big(Y_t^{(n)} \in S_{n,T}, 0\le t\le T\big) \;=\; 1. \tag{I.27}$$

Corollary 8.7 on p. 232 of Ethier and Kurtz [12] states that (I.27) jointly with the uniform convergence of the semigroups on the restricted area $S_{n,T}$ implies the convergence of $Y^{(n)}$ to Y in $D_S[0,\infty)$, hence the desired convergence of $X^{(n)}$ to X in $D_{\mathbb{R}}[0,\infty)$. Thus, it remains to show that for each $f\in\widehat{C}(S)$, the space of real-valued continuous functions on S vanishing at infinity, and $t\in[0,T]$

$$\lim_{n\to\infty}\sup_{(x,s)\in S_{n,T}} |\widetilde{T}_t^{(n)} f(x,s) - \widetilde{T}_t f(x,s)| \;=\; 0, \tag{I.28}$$

where $(\widetilde{T}_t^{(n)})_{t\ge 0}$ and $(\widetilde{T}_t)_{t\ge 0}$ denote the semigroups of $Y^{(n)}$ and Y, respectively, i.e., $\widetilde{T}_t^{(n)} f(x,s) = \mathbb{E}(f(X_{s+t}^{(n)}, s+t)\,|\,X_s^{(n)} = x)$ and $\widetilde{T}_t f(x,s) = \mathbb{E}(f(X_{s+t}, s+t)\,|\,X_s = x)$ for all $f\in\widehat{C}(S)$ and $(x,s)\in S$. By Lemma I.6.4, the space of all maps of the form $(x,s)\mapsto\sum_{i=1}^l g_i(x)h_i(s)$ with $l\in\mathbb{N}$, $g_i\in\widehat{C}(\mathbb{R})$ and $h_i\in\widehat{C}([0,\infty))$ is dense in $\widehat{C}(S)$. Hence it suffices to show (I.28) for $f = gh$ with $g\in\widehat{C}(\mathbb{R})$ and $h\in\widehat{C}([0,\infty))$, in which case

$$\begin{aligned}\widetilde{T}_t^{(n)} f(x,s) &= h(s+t)\mathbb{E}(g(X_{s+t}^{(n)})\,|\,X_s^{(n)} = x)\\ &= h(s+t)\mathbb{E}\bigg(g\bigg(\frac{a_k}{a_n}X_t^{(k)} + xm(t)\bigg)\bigg)\end{aligned}$$

for $(x,s)\in S$, where $k := k(n,s,x) := nm(s) + xa_n$, and

$$\begin{aligned}\widetilde{T}_t f(x,s) &= h(s+t)\mathbb{E}(g(X_{s+t}),\,|\,X_s = x)\\ &= h(s+t)\mathbb{E}(g(m(s)^{1/\alpha}X_t + xm(t))),\quad (x,s)\in S.\end{aligned}$$

Let $\varepsilon > 0$. Choose $C > 0$ such that $\sup_{n\in\mathbb{N}}\mathbb{P}(|X_t^{(n)}| > C) < \varepsilon$.

Splitting the mean along the event $A_k := \{|X_t^{(k)}| \le C\}$ yields

$$
\begin{aligned}
\sup_{(x,s)\in S_{n,T}} & |\widetilde{T}_t^{(n)} f(x,s) - \widetilde{T}_t f(x,s)| \\
&= \sup_{(x,s)\in S_{n,T}} h(s+t) \bigg| \mathbb{E}\bigg(g\bigg(\frac{a_k}{a_n} X_t^{(k)} + xm(t) \bigg)\bigg) \\
&\qquad - \mathbb{E}\big(g\big((m(s))^{1/\alpha} X_t + xm(t)\big)\big) \bigg| \\
&\le \|h\| \bigg(\sup_{(x,s)\in S_{n,T}} \big|\mathbb{E}\big(g((m(s))^{1/\alpha} X_t^{(k)} + xm(t))\big) \\
&\qquad - \mathbb{E}\big(g\big((m(s))^{1/\alpha} X_t + xm(t)\big)\big)\big| + 2\|g\|\varepsilon \\
&\qquad + \sup_{(x,s)\in S_{n,T}} \mathbb{E}\bigg(1_{A_k} \bigg| g\bigg(\frac{a_k}{a_n} X_t^{(k)} + xm(t) \bigg) \\
&\qquad - g\big((m(s))^{1/\alpha} X_t^{(k)} + xm(t)\big) \bigg| \bigg)\bigg).
\end{aligned}
$$

The second last supremum converges to 0 as $n \to \infty$ due to Lemma I.6.3 and since $k \to \infty$ as $n \to \infty$ by (I.24). The last supremum also converges to 0 due to (I.25) together with the uniform continuity of g. Since $\varepsilon > 0$ can be chosen arbitrarily, (I.28) holds, which completes the proof. $\square$

I.5 Proofs concerning Theorem I.2.8

This section contains the proofs of Lemma I.2.6, Lemma I.2.7 and Theorem I.2.8.

Proof. (of Lemma I.2.6) Fix $t \ge 0$. By Theorem 2 or Corollary 2.2 of Lamperti [24], applied with $x := 1 - s$ to the function $x \mapsto 1 - F(1-x,t)$, (I.11) holds if and only if

$$\lim_{s \to 1-} \alpha(s,t) = \alpha(t), \tag{I.29}$$

where

$$
\begin{aligned}
\alpha(s,t) := \frac{(1-s)\frac{\partial}{\partial s} F(s,t)}{1 - F(s,t)} &= \frac{f(F(s,t)) - F(s,t)}{1 - F(s,t)} \frac{1-s}{f(s)-s} \\
&= \frac{L((1-F(s,t))^{-1}) - 1}{L((1-s)^{-1}) - 1}
\end{aligned}
$$

for all $s \in (0,1)$. Thus (i) and (ii) are equivalent. By the backward equation,

$$at \;=\; \int_s^{F(s,t)} \frac{1}{f(u)-u}\,\mathrm{d}u \;=\; \int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}} \frac{1}{x(L(x)-1)}\,\mathrm{d}x. \qquad \text{(I.30)}$$

Also, note that

$$\begin{aligned}\log\frac{1}{\alpha(s,t)} \;&=\; \log(L((1-s)^{-1})-1) - \log(L((1-F(s,t))^{-1})-1)\\ &=\; \int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}} \frac{L'(x)}{L(x)-1}\,\mathrm{d}x.\end{aligned}$$

(iii) $\Rightarrow$ (ii): Applying integration by parts to (I.30) yields

$$\begin{aligned}at \;&=\; \left[\frac{\log x}{L(x)-1}\right]_{x=(1-F(s,t))^{-1}}^{x=(1-s)^{-1}}\\ &\quad+\int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}} \frac{\log x}{L(x)-1}\frac{L'(x)}{L(x)-1}\,\mathrm{d}x. \qquad \text{(I.31)}\end{aligned}$$

Let $\varepsilon > 0$ be arbitrary. Since $L(x)/\log x \to A > 0$ as $x \to \infty$, there exists $K > 0$ such that $1-\varepsilon \le A\log x/(L(x)-1) \le 1+\varepsilon$ for all $x \ge K$. But, if s is sufficiently close to 1, both inequalities hold on the interval where it is integrated above in (I.31), implying that $Aat = \lim_{s\to 1-}\log(\alpha(s,t))^{-1}$, which is exactly (I.29).

(i) $\Rightarrow$ (iii): Assume that (I.11) holds for all $t \ge 0$. By (I.13),

$$\alpha(t) \;=\; \lim_{s\to 1-}\frac{\log(1-F(s,t))}{\log(1-s)}.$$

As already seen before Lemma I.2.6, there exists $C \ge 0$ such that $\alpha(t) = e^{-Ct}$. Thus,

$$\begin{aligned}Ct \;&=\; -\lim_{s\to 1-}\log\frac{\log(1-F(s,t))}{\log(1-s)}\\ &=\; \lim_{s\to 1-}\int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}} \frac{1}{x\log x}\,\mathrm{d}x. \qquad \text{(I.32)}\end{aligned}$$

Division of (I.32) by (I.30) leads to

$$A \;:=\; \frac{C}{a} \;=\; \lim_{s\to 1-}\frac{\int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}}\frac{1}{x\log x}\,\mathrm{d}x}{\int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}}\frac{1}{x(L(x)-1)}\,\mathrm{d}x}.$$

Now exploit the monotonicity of $\log x$ and $L(x)$ to conclude that

$$
\begin{aligned}
A &\le \liminf_{s\to 1-} \frac{\frac{1}{\log((1-F(s,t))^{-1})}\int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}} \frac{1}{x}\,\mathrm{d}x}{\frac{1}{L((1-s)^{-1})-1}\int_{(1-F(s,t))^{-1}}^{(1-s)^{-1}} \frac{1}{x}\,\mathrm{d}x}\\
&= \liminf_{s\to 1-} \frac{L((1-s)^{-1})}{\log((1-F(s,t))^{-1})} \;=\; \frac{1}{\alpha(t)}\liminf_{s\to 1-}\frac{L((1-s)^{-1})}{\log((1-s)^{-1})}.
\end{aligned}
$$

Similarly, $A \ge \alpha(t)\limsup_{s\to 1-} L((1-s)^{-1})/\log((1-s)^{-1})$. Letting $t\to 0+$ yields $A = \lim_{s\to 1-} L((1-s)^{-1})/\log((1-s)^{-1})$, which is (iii) and completes the proof. □

Proof. (of Lemma I.2.7) By assumption, the function $H(x) := L(x) - 1 - A\log x$, $x \ge 1$, satisfies $\lim_{x\to\infty} H(x) = B-1$. Moreover, $\beta(t)$, defined via (I.16), satisfies

$$
\log\beta(t) \;=\; a\int_0^t (B-1-A\log\beta(s))\,\mathrm{d}s, \qquad t\ge 0. \tag{I.33}
$$

Computing the derivative of $L_t(x)$ with respect to t provides a representation for $L_t(x)$ similar to (I.33), namely

$$
\begin{aligned}
\frac{\partial}{\partial t}L_t(x) &= \frac{\partial}{\partial t}\big(x^{\alpha(t)}(1-F(1-x^{-1},t))\big)\\
&= x^{\alpha(t)}\alpha'(t)(\log x)(1-F(1-x^{-1},t))\\
&\quad - x^{\alpha(t)}a(f(F(1-x^{-1},t))-F(1-x^{-1},t))\\
&= ax^{\alpha(t)}(1-F(1-x^{-1},t))\bigg(\frac{1-f(F(1-x^{-1},t))}{1-F(1-x^{-1},t)}\\
&\quad - \frac{1-F(1-x^{-1},t)}{1-F(1-x^{-1},t)} - A\alpha(t)\log x\bigg)\\
&= aL_t(x)\Big(L((1-F(1-x^{-1},t))^{-1}) - 1 - A\log x^{\alpha(t)}\Big)\\
&= aL_t(x)\Big(L(x^{\alpha(t)}L_t^{-1}(x)) - 1 - A\log x^{\alpha(t)}\Big)\\
&= aL_t(x)\Big(H(x^{\alpha(t)}L_t^{-1}(x)) - A\log L_t(x)\Big), \qquad t\ge 0, x\ge 1.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\log L_t(x) &= \int_0^t \frac{\frac{\partial}{\partial s}L_s(x)}{L_s(x)}\,\mathrm{d}s\\
&= a\int_0^t \big(H(x^{\alpha(s)}L_s^{-1}(x)) - A\log L_s(x)\big)\mathrm{d}s, \qquad \text{(I.34)}
\end{aligned}
$$

$t\ge 0$. Let $t>0$ be fixed and $\varepsilon>0$ be arbitrary. If $1-x^{-1}>q$, where q denotes the extinction probability, then the map $s\to x^{\alpha(s)}L_s^{-1}(x) =$

$(1-F(1-x^{-1},s))^{-1}$ is non-increasing. Hence, $|H(x^{\alpha(s)}L_s^{-1}(x))-(B-1)|<\varepsilon$ for all $s\in[0,t]$ and sufficiently large x. By (I.33) and (I.34),

$$|\log L_t(x)\ -\ \log\beta(t)|\ \le\ a\varepsilon t\ +\ aA\int_0^t|\log L_s(x)-\log\beta(s)|\,\mathrm{d}s.$$

By Gronwall's inequality,

$$\begin{aligned}|\log L_t(x)-\log\beta(t)|\ &\le\ a\varepsilon t\ +\ aA\int_0^t a\varepsilon s\exp\Big(\int_s^t aA\,\mathrm{d}\sigma\Big)\mathrm{d}s\\ &\le\ a\varepsilon t\Big(1+\int_0^t aA\exp(aA(t-s))\,\mathrm{d}s\Big)\\ &=\ a\varepsilon t\exp(aAt).\end{aligned}$$

Since $\varepsilon>0$ can be chosen arbitrarily small, the claim $\lim_{x\to\infty}L_t(x)=\beta(t)$ follows. □

Proof. (of Theorem I.2.8) The proof is divided into two steps. First the assumption (I.15) is used to establish the convergence of the one-dimensional distributions. Afterwards it is shown with some general weak convergence machinery for Markov processes that the convergence of the one-dimensional distributions is already sufficient for convergence in $D_E[0,\infty)$, where $E:=[0,\infty)$.

Step 1. (Convergence of the one-dimensional distributions) Fix $\lambda,t\ge 0$. Define $s_n:=\exp(-\lambda n^{-1/\alpha(t)})$, $n\in\mathbb{N}$. Note that $s_n\to 1$ as $n\to\infty$. We have $\mathbb{E}(\exp(-\lambda X_t^{(n)}))=\mathbb{E}(\exp(-\lambda n^{-1/\alpha(t)}Z_t^{(n)}))=(\mathbb{E}(\exp(-\lambda n^{-1/\alpha(t)}Z_t)))^n=(F(s_n,t))^n$. Taking the logarithm yields

$$\begin{aligned}\log\mathbb{E}(\exp(-\lambda X_t^{(n)}))\ &=\ n\log(1-(1-F(s_n,t)))\\ &\sim\ -n(1-F(s_n,t))\ \sim\ -n\beta(t)(1-s_n)^{\alpha(t)}\end{aligned}$$

as $n\to\infty$ by (I.15). Since $1-s_n=1-\exp(-\lambda n^{-1/\alpha(t)})\sim\lambda n^{-1/\alpha(t)}$ as $n\to\infty$ it follows that the latter expression is asymptotically equal to $-n\beta(t)(\lambda n^{-1/\alpha(t)})^{\alpha(t)}=-\beta(t)\lambda^{\alpha(t)}$. Therefore, $\lim_{n\to\infty}\mathbb{E}(\exp(-\lambda X_t^{(n)}))=\exp(-\beta(t)\lambda^{\alpha(t)})=\mathbb{E}(\exp(-\lambda X_t))$. This pointwise convergence of the Laplace transforms implies the convergence $X_t^{(n)}\to X_t$ in distribution as $n\to\infty$.

Step 2. (Convergence in $D_E[0,\infty)$) We proceed as in the proof of [23, Theorem 2.1]. For $n\in\mathbb{N}$ and $t\ge 0$ define $E_{n,t}:=\{j/n^{1/\alpha(t)}:j\in\mathbb{N}_0\}$. In general the process $X^{(n)}$ is time-inhomogeneous. Let $Y^{(n)}:=(X_t^{(n)},t)_{t\ge0}$ and $Y:=(X_t,t)_{t\ge0}$ denote the space-time processes of $X^{(n)}$ and X, respectively. Note that $Y^{(n)}$ has state space

$S_n := \{(j/n^{1/\alpha(t)}, t) : j \in \mathbb{N}_0, t \geq 0\} = \bigcup_{t\geq 0}(E_{n,t} \times \{t\})$ and Y has state space $S := [0,\infty)^2$. According to Revuz and Yor [36, p. 85, Exercise (1.10)] the process $Y^{(n)}$ is time-homogeneous. Define $\pi_n : B(S) \to B(S_n)$ via $\pi_n g(x,s) := g(x,s)$ for $g \in B(S)$ and $(x,s) \in S_n$. In the following it is shown that $Y^{(n)}$ converges in $D_S[0,\infty)$ to Y as $n \to \infty$. Note that this convergence implies the desired convergence of $X^{(n)}$ in $D_E[0,\infty)$ to X as $n \to \infty$. For $\lambda, \mu > 0$ define the test function $g_{\lambda,\mu}$ via $g_{\lambda,\mu}(x,s) := e^{-\lambda x - \mu s}$, $(x,s) \in S$. By [23, Proposition 5.4], it suffices to verify that for every $t \geq 0$ and $\lambda, \mu > 0$,

$$\lim_{n\to\infty} \sup_{s\geq 0} \sup_{x\in E_{n,s}} |U_t^{(n)}\pi_n g_{\lambda,\mu}(x,s) - \pi_n U_t g_{\lambda,\mu}(x,s)| \;=\; 0, \qquad \text{(I.35)}$$

where $U_t^{(n)} : B(S_n) \to B(S_n)$ is defined via $U_t^{(n)} g(x,s) := \mathbb{E}(g(X_{s+t}^{(n)}, s+t)\,|\,X_s^{(n)} = x)$, $g \in B(S_n)$, $s \geq 0$, $x \in E_{n,s}$. Note that $(U_t^{(n)})_{t\geq 0}$ is the semigroup of $Y^{(n)}$.

Fix $t \geq 0$ and $\lambda, \mu > 0$. For all $n \in \mathbb{N}$, $s \geq 0$ and $x \in E_{n,s}$,

$$\begin{aligned}
U_t^{(n)}\pi_n g_{\lambda,\mu}(x,s) &= \mathbb{E}(\pi_n g_{\lambda,\mu}(X_{s+t}^{(n)}, s+t)\,|\,X_s^{(n)} = x)\\
&= \mathbb{E}(\exp(-\lambda X_{s+t}^{(n)} - \mu(s+t))\,|\,X_s^{(n)} = x)\\
&= e^{-\mu(s+t)}\mathbb{E}(\exp(-\lambda n^{-1/\alpha(s+t)} Z_{s+t)}^{(n)})\,|\,Z_s^{(n)} = xn^{1/\alpha(s)})\\
&= e^{-\mu(s+t)}\mathbb{E}(\exp(-\lambda n^{-1/\alpha(s+t)} Z_t^{(xn^{1/\alpha(s)})}))
\end{aligned}$$

and

$$\begin{aligned}
\pi_n U_t g_{\lambda,\mu}(x,s) &= U_t g_{\lambda,\mu}(x,s)\\
&= \mathbb{E}(\exp(-\lambda X_{s+t} - \mu(s+t))\,|\,X_s = x)\\
&= e^{-\mu(s+t)}\mathbb{E}(\exp(-\lambda X_{s+t})\,|\,X_s = x)\\
&= e^{-\mu(s+t)}\mathbb{E}(\exp(-\lambda x^{1/\alpha(t)} X_t)).
\end{aligned}$$

Thus, one has to verify that

$$\lim_{n\to\infty} \sup_{s\geq 0} \sup_{x\in E_{n,s}} e^{-\mu(s+t)}\Big|\mathbb{E}(\exp(-\lambda n^{-1/\alpha(s+t)} Z_t^{(xn^{1/\alpha(s)})})) - \mathbb{E}(\exp(-\lambda x^{1/\alpha(t)} X_t))\Big| \;=\; 0.$$

We will even verify that

$$\lim_{n\to\infty} \sup_{s\geq 0} \sup_{x>0} \Big|\mathbb{E}(\exp(-\lambda n^{-1/\alpha(s+t)} Z_t^{(\lfloor xn^{1/\alpha(s)}\rfloor)})) - \mathbb{E}(\exp(-\lambda x^{1/\alpha(t)} X_t))\Big| \;=\; 0.$$

Since $\alpha(s+t) = \alpha(s)\alpha(t)$, the quantity inside the absolute values depends on n and s only via $n^{1/\alpha(s)}$. Since $n^{1/\alpha(s)}$ is non-decreasing in s it follows that the convergence for fixed $s \geq 0$ is slower as s is smaller. So the slowest convergence holds for $s = 0$ ($\Rightarrow \alpha(s) = 1$). Thus it suffices to verify that for every $t \geq 0$ and $\lambda > 0$

$$\lim_{n\to\infty} \sup_{x>0} |\mathbb{E}(\exp(-\lambda n^{-1/\alpha(t)} Z_t^{(\lfloor xn \rfloor)})) - \mathbb{E}(\exp(-\lambda x^{1/\alpha(t)} X_t))| \;=\; 0.$$

The map $x \mapsto \mathbb{E}(\exp(-\lambda x^{1/\alpha(t)} X_t))$ is bounded, continuous and non-increasing. Since $Z_t^{(1)} \leq Z_t^{(2)} \leq \cdots$ almost surely it follows by Pólya's theorem [35, Satz I] that it suffices to verify the above convergence pointwise for every $x > 0$. Defining $k := \lfloor xn \rfloor$ it is readily seen that this is equivalent to the convergence of the one-dimensional distributions $X_t^{(k)} = k^{-1/\alpha(t)} Z_t^{(k)} \to X_t$ in distribution as $k \to \infty$, $t \geq 0$. But the convergence of the one-dimensional distributions holds by Step 1. The proof is complete. □

I.6 Appendix

In this appendix five auxiliary results are provided. Lemma I.6.1 and Lemma I.6.2 below are used in the proof of Lemma I.2.2. Lemma I.6.1 provides an asymptotic statement for Laplace transforms and generating functions, respectively. Lemma I.6.2 is a version of L'Hospital's rule, which is stated for completeness.

Lemma I.6.1. *Let ξ be a non-negative real-valued random variable with $m := \mathbb{E}(\xi) < \infty$. Suppose that the cumulative distribution function F of ξ satisfies $1 - F(x) \leq Cx^{-\alpha}$ for all $x \geq 0$ for some $C < \infty$ and $\alpha > 1$. Then, for every $\varepsilon \in [0, \min(\alpha - 1, 1))$,*

$$\lim_{\lambda\to 0+} \frac{1 - \varphi(\lambda) + \lambda m}{\lambda^{1+\varepsilon}} \;=\; 0, \tag{I.36}$$

where φ denotes the Laplace transform of ξ. If ξ takes only values in $\mathbb{N}_0$, then, for the same range of values of ε as above,

$$\lim_{s\to 1-} \frac{(1-s)m - (1 - f(s))}{(1-s)^{1+\varepsilon}} \;=\; 0, \tag{I.37}$$

where f denotes the pgf of ξ.

Remark. The tail condition is satisfied if $\mathbb{E}(\xi^\alpha) < \infty$, since, by Markov's inequality, $1 - F(x) = \mathbb{P}(\xi^\alpha > x^\alpha) \leq x^{-\alpha}\mathbb{E}(\xi^\alpha)$.

Proof. (of Lemma I.6.1) Applying the well known formula $\mathbb{E}(g(\xi)) = g(0) + \int_0^\infty g'(x)(1-F(x))\,\mathrm{d}x$, $g \in C^1([0,\infty))$, to the function $g(x) := e^{-\lambda x} - 1 + \lambda x$ yields

$$\begin{aligned}\frac{\varphi(\lambda) - 1 + \lambda m}{\lambda^{1+\varepsilon}} &= \frac{1}{\lambda^\varepsilon}\int_0^\infty (1-F(x))(1-e^{-\lambda x})\,\mathrm{d}x \\ &\le \int_0^1 \frac{1-e^{-\lambda x}}{\lambda^\varepsilon}\,\mathrm{d}x \;+\; C\int_1^\infty \frac{1-e^{-\lambda x}}{(\lambda x)^\varepsilon x^{\alpha-\varepsilon}}\,\mathrm{d}x.\end{aligned}$$

Since $\varepsilon < 1$, $\lim_{\lambda\to 0}(1-e^{-\lambda x})/\lambda^\varepsilon = 0$ and the first integral converges to 0 by the dominated convergence theorem. Since $(1-e^{-\lambda x})/(\lambda x)^\varepsilon$ is bounded uniformly in λ and x, and $\alpha - \varepsilon > 1$, the dominated convergence theorem is again applicable and the second integral converges to 0. If ξ takes only values in $\mathbb{N}_0$, then (I.37) follows from (I.36) via the substitution $\lambda := -\log s$, $s \in (0,1)$, and the fact that $-\log s = (1-s) + O((1-s)^2)$ as $s \to 1$. □

The situation in the following lemma is the one of L'Hospital's rule.

Lemma I.6.2. *Let $c, x_0 \in [-\infty, \infty]$. Let $f, g : I \to \mathbb{R}$ be continuously differentiable on an open interval I containing x_0 or having x_0 as a limit point if the limit is one-sided. Assume further that $g'(x) \neq 0$ for all $x \in I \setminus \{x_0\}$. Let $\alpha \in \mathbb{R} \setminus \{1\}$. If either*

$$\lim_{x\to x_0} g^{1-1/\alpha}(x) \;=\; \lim_{x\to x_0} f(x)/g^{1/\alpha}(x) \;=\; 0$$

or

$$\lim_{x\to x_0} g^{1-1/\alpha}(x) \;=\; \lim_{x\to x_0} f(x)/g^{1/\alpha}(x) \;=\; \infty,$$

and

$$\lim_{x\to x_0}(\alpha f'(x)/g'(x) - f(x)/g(x)) \;=\; c, \tag{I.38}$$

then $\lim_{x\to x_0} f(x)/g(x) = c(\alpha-1)^{-1}$. If the limit (I.38) does not exist, it still holds that

$$\begin{aligned}\liminf_{x\to x_0}\left(\alpha\frac{f'(x)}{g'(x)} - \frac{f(x)}{g(x)}\right) &\le \liminf_{x\to x_0}(\alpha-1)\frac{f(x)}{g(x)} \\ &\le \limsup_{x\to x_0}(\alpha-1)\frac{f(x)}{g(x)} \\ &\le \limsup_{x\to x_0}\left(\alpha\frac{f'(x)}{g'(x)} - \frac{f(x)}{g(x)}\right).\end{aligned}$$

Proof. A straightforward computation shows that

$$\frac{1}{\alpha-1}\left(\alpha\frac{f'(x)}{g'(x)}-\frac{f(x)}{g(x)}\right)=\frac{f'(x)g^{-1/\alpha}(x)-(1/\alpha)g^{-1-1/\alpha}(x)g'(x)f(x)}{(1-1/\alpha)g^{-1/\alpha}g'(x)},$$

where the numerator and the denominator are the derivatives of $f(x)g^{-1/\alpha}(x)$ and $g^{1-1/\alpha}(x)$, respectively. Thus the convergence of the left-hand side to $c(\alpha-1)^{-1}\in[-\infty,\infty]$ implies $\lim_{x\to x_0}f(x)/g(x)=\lim_{x\to x_0}(f(x)/g^{1/\alpha}(x))/g^{1-1/\alpha}(x)=c(\alpha-1)^{-1}$. □

The following two results are needed in the proof of Theorem I.2.3. Lemma I.6.3 contains a statement on uniform weak convergence. The last result (Lemma I.6.4) provides a certain dense subset of $\widehat{C}(\mathbb{R}\times[0,\infty))$.

Lemma I.6.3. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued random variables converging weakly to a real-valued random variable X. Then, for every bounded and continuous function $f:\mathbb{R}\to\mathbb{R}$ and $A,B>0$,*

$$\lim_{n\to\infty}\sup_{|a|\le A,|b|\le B}|\mathbb{E}(f(aX_n+b))-\mathbb{E}(f(aX+b))| \;=\; 0. \qquad \text{(I.39)}$$

If $f\in\widehat{C}(\mathbb{R})$, then (I.39) even holds if the supremum is taken over $[-A,A]\times\mathbb{R}$ instead of $[-A,A]\times[-B,B]$.

Proof. For $n\in\mathbb{N}$ define $g_n:\mathbb{R}^2\to\mathbb{R}$ via $g_n(a,b):=\mathbb{E}(f(aX_n+b))$, $a,b\in\mathbb{R}$, and g similarly with X_n replaced by X. Fix $A,B>0$. Obtaining pointwise convergence of g_n to g from weak convergence, (I.39) follows, in view of the Arzelà–Ascoli theorem, from the uniform equicontinuity of $\{g_n:n\in\mathbb{N}\}$ on $K:=[-A,A]\times[-B,B]$, that is, for every $\varepsilon>0$ there exists $\delta>0$ such that $\max\{|a-a'|,|b-b'|\}<\delta$ implies $|g_n(a,b)-g_n(a',b')|<\varepsilon$ for all $n\in\mathbb{N}$ and all $(a,b),(a',b')\in K$.

Let $\varepsilon>0$. By Prohorov's theorem, the family of distributions of the weakly convergent sequence $(X_n)_{n\in\mathbb{N}}$ is tight. Thus, there exists $C\in(0,\infty)$ such that $\sup_{n\in\mathbb{N}}\mathbb{P}(|X_n|>C)<\varepsilon$ and $\mathbb{P}(|X|>C)<\varepsilon$. Using the uniform continuity of f on K, choose $\delta>0$ such that $|x-y|<\delta(C+1)$ implies $|f(x)-f(y)|<\varepsilon$. Consequently,

$$\begin{aligned}|g_n(a,b)-g_n(a',b')| \;&=\; |\mathbb{E}(f(aX_n+b))-\mathbb{E}(f(a'X_n+b'))|\\ &\le\; 2\varepsilon\|f\| \;+\; \mathbb{E}(1_{\{|X_n|\le C\}}|f(aX_n+b)-f(a'X_n+b')|)\\ &\le\; 2\varepsilon\|f\| \;+\; \varepsilon\end{aligned}$$

for $(a,b),(a',b')\in K$ with $\max\{|a-a'|,|b-b'|\}<\delta$, proving the first statement.

If $f \in \widehat{C}(\mathbb{R})$, then there exists $L > 0$ such that $|f(x)| < \varepsilon$ for all $|x| > L$. In particular, (I.39) holds for $B := AC + L$. On the remaining area $[-A, A] \times (\mathbb{R} \setminus [-B, B])$ all the functions g_n and g are sufficiently small. More precisely, if $|a| \leq A$ and $|b| > B$, then $|aX_n + b| > L$ on the event $\{|X_n| \leq C\}$, hence

$$\begin{aligned} |g_n(a,b)| &= |\mathbb{E}(f(aX_n + b))| \\ &\leq \varepsilon\|f\| + \mathbb{E}(1_{\{|X_n| \leq C\}}|f(aX_n + b)|) \leq \varepsilon\|f\| + \varepsilon \end{aligned}$$

for all $n \in \mathbb{N}$, and, similarly, $|g(a,b)| \leq \varepsilon\|f\| + \varepsilon$, which proves the additional statement. □

Lemma I.6.4. *Let $S := \mathbb{R} \times [0, \infty)$. The space of functions $f : S \to \mathbb{R}$ of the form $f(x,y) = \sum_{i=1}^{l} g_i(x)h_i(y)$ with $l \in \mathbb{N}$, $g_1, \ldots, g_l \in \widehat{C}(\mathbb{R})$ and $h_1, \ldots, h_l \in \widehat{C}([0,\infty))$ is dense in $\widehat{C}(S)$.*

Proof. Two proofs are provided. The first proof is elementary and constructive. The second proof exploits the Stone–Weierstrass theorem for locally compact spaces.

Proof 1. (elementary) Each $f \in \widehat{C}(S)$ can be transformed (with the additional definition $f(\pm\infty, y) := 0$ for all $y \in [0,\infty)$ and $f(x,\infty) := 0$ for all $x \in \mathbb{R}$) into a map $\widetilde{f} \in C([0,1]^2)$ satisfying $\widetilde{f}(0,y) = \widetilde{f}(1,y) = \widetilde{f}(x,1) = 0$ for all $x, y \in [0,1]$ via

$$\widetilde{f}(x,y) := f\left(\frac{1}{1-x} - \frac{1}{x}, \frac{y}{1-y}\right), \qquad x, y \in [0,1]^2.$$

Thus, it suffices to verify that the space D of functions $f : [0,1]^2 \to \mathbb{R}$ of the form $f(x,y) = \sum_{i=1}^{l} g_i(x)h_i(y)$ with $l \in \mathbb{N}$, $g_1, \ldots, g_l \in D_1 := \{g \in C([0,1]) : g(0) = g(1) = 0\}$ and $h_1, \ldots, h_l \in D_2 := \{h \in C([0,1]) : h(1) = 0\}$ is dense in $\{f \in C([0,1]^2) : f(0,y) = f(1,y) = f(x,1) = 0 \text{ for all } x, y \in [0,1]\}$. This is seen as follows. Let $m \in \mathbb{N}$. For $i \in \{0, \ldots, m\}$ define $x_i := i/m$ and $g_i : [0,1] \to [0,1]$ via

$$g_i(x) := (1 - m|x - x_i|)\, 1_{\{|x - x_i| \leq 1/m\}}, \qquad x \in [0,1].$$

Note that $g_0, \ldots, g_m$ form a partition of unity, i.e., $\sum_{i=0}^{m} g_i(x) = 1$ for all $x \in [0,1]$. Moreover, $g_1, \ldots, g_{m-1} \in D_1$. In the same manner define $y_j := j/m$ and $h_j : [0,1] \to [0,1]$ via $h_j(y) := (1 - m|y - y_j|)1_{\{|y - y_j| \leq 1/m\}}$ for all $j \in \{0, \ldots, m\}$. Again, $h_0, \ldots, h_m$ form a partition of unity, i.e., $\sum_{j=0}^{m} h_j(y) = 1$ for all $y \in [0,1]$. Moreover, $h_0, \ldots, h_{m-1} \in D_2$. Now define $f_m : [0,1]^2 \to \mathbb{R}$ via

$$f_m(x,y) := \sum_{i,j=0}^{m} f(x_i, y_j) g_i(x) h_j(y) = \sum_{i=1}^{m-1} \sum_{j=0}^{m-1} f(x_i, y_j) g_i(x) h_j(y),$$

$x, y \in [0,1]$, where the last equality holds, since $f(0,y) = f(1,y) = f(x,1) = 0$ for all $x, y \in [0,1]$. From $g_1, \ldots, g_{m-1} \in D_1$ and $h_0, \ldots, h_{m-1} \in D_2$ it follows that $f_m \in D$. It remains to verify that $\lim_{m\to\infty} \|f_m - f\| = 0$. Let $\varepsilon > 0$. Since f is uniformly continuous on $[0,1]^2$, there exists $\delta = \delta(\varepsilon) > 0$ such that $|f(x', y') - f(x,y)| < \varepsilon$ for all $x, y, x', y' \in [0,1]$ with $|x - x'| < \delta$ and $|y - y'| < \delta$. For all $x, y \in [0,1]$ it follows from $\sum_{i,j=0}^m g_i(x)h_j(y) = 1$ that

$$\begin{aligned} |f_m(x,y) - f(x,y)| &= \left|\sum_{i,j=0}^m \big(f(x_i, y_j) - f(x,y)\big) g_i(x) h_j(y)\right| \\ &\leq \sum_{i,j=0}^m |f(x_i, y_j) - f(x,y)| g_i(x) h_j(y). \end{aligned}$$

Now for each $(x,y) \in [0,1]^2$ there exist $i_0, j_0 \in \{0, \ldots, m-1\}$ (depending on x and y) such that $x_{i_0} \leq x \leq x_{i_0+1}$ and $y_{j_0} \leq y \leq y_{j_0+1}$. Since $g_i(x) = 0$ for all $i \in \{0, \ldots, m\} \setminus \{i_0, i_0+1\}$ and $h_j(y) = 0$ for all $j \in \{0, \ldots, m\} \setminus \{j_0, j_0+1\}$, we conclude that

$$\begin{aligned} &|f_m(x,y) - f(x,y)| \\ &\quad \leq |f(x_{i_0}, y_{i_0}) - f(x,y)| + |f(x_{i_0}, y_{j_0+1}) - f(x,y)| \\ &\qquad + |f(x_{i_0+1}, y_{j_0}) - f(x,y)| + |f(x_{i_0+1}, y_{j_0+1}) - f(x,y)| \leq 4\varepsilon \end{aligned}$$

for all $m \in \mathbb{N}$ with $m > 1/\delta$. Thus, $\lim_{m\to\infty} \|f_m - f\| = 0$. □

Proof 2. (using the Stone–Weierstrass theorem) The space of functions $f : S \to \mathbb{R}$ of the form $f(x,y) = \sum_{i=1}^l g_i(x)h_i(y)$ with $l \in \mathbb{N}$, $g_1, \ldots, g_l \in \widehat{C}(\mathbb{R})$ and $h_1, \ldots, h_l \in \widehat{C}([0,\infty))$ is a subalgebra of $\widehat{C}(S)$ which separates points and vanishes nowhere, whence is dense in $\widehat{C}(S)$ by the Stone–Weierstrass theorem (see, for example, [11]). In [11] the theorem is stated for complex-valued functions, but it remains true for real-valued functions. To see this, let $f \in \widehat{C}(S) \subseteq \widehat{C}(S, \mathbb{C})$ be arbitrary. By the theorem, there exist $g_1, g_2, \ldots \in \widehat{C}(S, \mathbb{C})$ such that $\lim_{n\to\infty} \|g_n - f\| = 0$. Then $f_n := \mathrm{Re}(g_n) \in \widehat{C}(S)$, $n \in \mathbb{N}$, and $\|f_n - f\| \leq \|g_n - f\| \to 0$ as $n \to \infty$. □

The last result (Lemma I.6.5) states that a particular function is absolutely monotone. This result is needed in Example I.2.13, but is as well of its own interest. For general information on absolutely and completely monotonic functions we refer the reader to (Appendix A, §4 of) Steutel and van Harn [42] and Chapter IV of Widder [46].

Lemma I.6.5. *The function* $g(z) := (-\log(1-z)) \log(1 - \log(1-z))$ *is absolutely monotone on* $[0,1)$.

Remark. The following proof of Lemma I.6.5 is based on a particular integral representation (see (I.41)) for the coefficient $g_n := g^{(n)}(0)/n!$ in front of z^n in the Taylor expansion of $g(z)$ around 0. Concrete calculations of the coefficients g_n via the summation formula (I.40) or the integral representation (I.41) yield

$$g(z) \;=\; z^2+\frac{1}{2}z^3+\frac{1}{2}z^4+\frac{3}{8}z^5+\frac{247}{720}z^6+\frac{7}{24}z^7+\frac{535}{2016}z^8+\frac{2051}{8640}z^9+O(z^{10}).$$

Proof. (of Lemma I.6.5) We have to verify that $g_n \geq 0$ for all $n \in \mathbb{N}_0$. Clearly, $g_0 = g(0) = 0$ and $g = u \circ v$, where $u(x) := x\log(x+1)$ for $x \geq 0$ and $v(z) := -\log(1-z)$ for $0 \leq z < 1$. The functions u and v have derivatives $u'(x) = \log(x+1) + x/(x+1)$, $x \geq 0$,

$$u^{(k)}(x) \;=\; (-1)^k\bigg(\frac{(k-2)!}{(x+1)^{k-1}} + \frac{(k-1)!}{(x+1)^k}\bigg), \qquad k \in \mathbb{N}\setminus\{1\}, x \geq 0,$$

and $v^{(m)}(z) = (m-1)!/(1-z)^m$, $m \in \mathbb{N}$, $0 \leq z < 1$. Note that v is absolutely monotone, but u is not absolutely monotone, which explains why the statement of Lemma I.6.5 is less simple as it seems at a first glance. By Faà di Bruno's formula,

$$\begin{aligned}
g^{(n)}(z) \;&=\; \sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}_0\\ 1k_1+2k_2+\cdots+nk_n=n}} \frac{n!}{k_1!\cdots k_n!}u^{(k_1+\cdots+k_n)}(v(z))\prod_{m=1}^{n}\bigg(\frac{v^{(m)}(z)}{m!}\bigg)^{k_m}\\
&=\; \frac{1}{(1-z)^n}\sum_{k=1}^{n}u^{(k)}(v(z))\sum_{\substack{1k_1+2k_2+\cdots+nk_n=n\\ k_1+\cdots+k_n=k}}\frac{n!}{k_1!\cdots k_n!1^{k_1}\cdots n^{k_n}}\\
&=\; \frac{1}{(1-z)^n}\sum_{k=1}^{n}u^{(k)}(v(z))|s(n,k)|, \qquad n \in \mathbb{N}, 0 \leq z < 1,
\end{aligned}$$

where $s(.,.)$ denote the Stirling numbers of the first kind. Thus,

$$\begin{aligned}
g_n \;:=\; \frac{g^{(n)}(0)}{n!} \;&=\; \frac{1}{n!}\sum_{k=1}^{n}u^{(k)}(0)|s(n,k)|\\
&=\; \frac{1}{n!}\sum_{k=2}^{n}(-1)^k\frac{k!}{k-1}|s(n,k)|, \qquad n \in \mathbb{N}. \qquad \text{(I.40)}
\end{aligned}$$

Plugging in $k!/(k-1) = (k-2)! + (k-1)! = \int_0^\infty t^k(1/t+1/t^2)e^{-t}\,\mathrm{d}t$ and interchanging the sum with the integral yields

$$g_n \;=\; \frac{1}{n!}\int_0^\infty \sum_{k=2}^{n}(-t)^k|s(n,k)|\bigg(\frac{1}{t}+\frac{1}{t^2}\bigg)e^{-t}\,\mathrm{d}t, \qquad n \in \mathbb{N}.$$

Applying the relation $\sum_{k=0}^n x^k|s(n,k)| = [x]_n := x(x+1)\cdots(x+n-1)$ to the point $x := -t$ and noting that $s(n,0)=0$ and $s(n,1)=(n-1)!$ for $n\in\mathbb{N}$ yields

$$g_n \;=\; \frac{1}{n!}\int_0^\infty \big([-t]_n + t(n-1)!\big)\left(\frac{1}{t}+\frac{1}{t^2}\right)e^{-t}\,\mathrm{d}t, \qquad n\in\mathbb{N}.$$

Noting that $[-t]_n/n! = (-1)^n\binom{t}{n}$ shows that g_n has the integral representation

$$g_n \;=\; \int_0^\infty \left((-1)^n\binom{t}{n}+\frac{t}{n}\right)\left(\frac{1}{t}+\frac{1}{t^2}\right)e^{-t}\,\mathrm{d}t, \quad n\in\mathbb{N}. \tag{I.41}$$

In particular, $g_1=0$. For $n\in\{2,4,\ldots\}$ the map $t\mapsto(-1)^n\binom{t}{n}+\frac{t}{n} = \binom{t}{n}+\frac{t}{n}$ is non-decreasing and hence non-negative on $[0,\infty)$ implying that $g_n\ge 0$ for even n. Assume now that $n\in\{3,5,\ldots\}$. Then the map $t\mapsto(-1)^n\binom{t}{n}+\frac{t}{n}=\frac{t}{n}-\binom{t}{n}$, $t>0$, has a single root at $t=n$ and is positive for $t\in(0,n)$ and negative for $t\in(n,\infty)$. Decompose $g_n = I_1 - I_2$, where

$$I_1 \;:=\; \int_0^n \left(\frac{t}{n}-\binom{t}{n}\right)\left(\frac{1}{t}+\frac{1}{t^2}\right)e^{-t}\,\mathrm{d}t$$

and

$$I_2 \;:=\; \int_n^\infty \left(\binom{t}{n}-\frac{t}{n}\right)\left(\frac{1}{t}+\frac{1}{t^2}\right)e^{-t}\,\mathrm{d}t.$$

The map $t\mapsto(\frac{t}{n}-\binom{t}{n})(\frac{1}{t}+\frac{1}{t^2})$ takes on the interval $(0,n-1]$ its minimum value $1/(n-1)$ at the right most point $n-1$. For $n\in\{3,5,\ldots\}$ we hence obtain for I_1 the lower bound

$$\begin{aligned} I_1 \;&\ge\; \int_0^{n-1}\left(\frac{t}{n}-\binom{t}{n}\right)\left(\frac{1}{t}+\frac{1}{t^2}\right)e^{-t}\,\mathrm{d}t\\ &\ge\; \frac{1}{n-1}\int_0^{n-1}e^{-t}\,\mathrm{d}t \;=\; \frac{1-e^{-(n-1)}}{n-1} \;\ge\; \frac{1}{n}. \end{aligned}$$

For I_2 we obtain the upper bound

$$\begin{aligned} I_2 \;&\le\; \int_n^\infty \binom{t}{n}\left(\frac{1}{t}+\frac{1}{t^2}\right)e^{-t}\,\mathrm{d}t\\ &\le\; \int_n^\infty \frac{t^n}{n!}\frac{2}{t}e^{-t}\,\mathrm{d}t \;=\; \frac{2}{n}\int_n^\infty \mathbb{P}(N_t=n-1)\,\mathrm{d}t, \end{aligned}$$

where N_t is Poisson distributed with parameter t. Applying the formula $\int_n^\infty\mathbb{P}(N_t=k)\,\mathrm{d}t = \mathbb{P}(N_n\le k)$ with $k=n-1$ leads to $I_2\le(2/n)\mathbb{P}(N_n\le n-1)\le 1/n$, since $\mathbb{P}(N_n\le n-1)$ is increasing in n with $\lim_{n\to\infty}\mathbb{P}(N_n\le n-1)=1/2$; see, for example, Teicher

[43]. From $I_1 \geq 1/n$ and $I_2 \leq 1/n$ it follows that $g_n = I_1 - I_2 \geq 0$ for $n \in \{3, 5, \ldots\}$. In summary, $g_n \geq 0$ for all $n \in \mathbb{N}_0$. □

Remark. (Asymptotics of g_n) It is easily seen that $g'(z) \sim (1-z)^{-1}L((1-z)^{-1})$ as $z \to 1-$ in $\Delta \setminus \{1\}$, where $L(u) := \log\log u$ and Δ is defined as in Flajolet and Odlyzko [13, Eq. (2.5)]. By Theorem 5 of [13], applied with $\alpha := -1$ and f replaced by g', $[z^n]g'(z) \sim L(n) = \log\log n$ as $n \to \infty$. From $[z^n]g'(z) = (n+1)g_{n+1}$ it follows that $g_n \sim n^{-1}\log\log n$ as $n \to \infty$.

Bibliography

[1] Aczel, J. (1966) *Lectures on Functional Equations and Their Applications.* Academic Press, New York.

[2] Aldous, D. (1989) Stopping times and tightness II. *Ann. Probab.* **17**, 586–595.

[3] Athreya, K. B. and Ney, P. E. (1972) *Branching Processes.* Die Grundlehren der Mathematischen Wissenschaften **196**, Springer, New York.

[4] Bhatia, R. (2006) Infinitely divisible matrices. *Amer. Math. Monthly* **113**, 221–235.

[5] Bhatia, R. (2011) Min matrices and mean matrices. *Math. Intelligencer* **33**, 22–28.

[6] Bingham, N. H. and Doney, R. A. (1974) Asymptotic properties of supercritical branching processes I: The Galton-Watson process. *Adv. Appl. Probab.* **6**, 711–731.

[7] Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[8] Bogachev, V. I., Röckner, M. and Schmuland, B. (1996) Generalized Mehler semigroups and applications. *Probab. Theory Relat. Fields* **105**, 193–225.

[9] Bojanić, R. and Seneta, E. (1971) Slowly varying functions and asymptotic relations. *J. Math. Anal. Appl.* **34**, 302–315.

[10] Christoph, G. and Schreiber, K. (1998) Discrete stable random variables. *Statist. Probab. Lett.* **37**, 243–247.

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

[12] Ethier, S. N. and Kurtz, T. G. (1986) *Markov Processes, Characterization and Convergence.* Wiley, New York.

[13] Flajolet, P. and Odlyzko, A. (1990) Singularity analysis of generating functions. *SIAM J. Disc. Math.* **3**, 216–240.

[14] Flajolet, P. and Sedgewick, R. (2009) *Analytic Combinatorics.* Cambridge University Press, Cambridge.

[15] Geluk, J. L. and de Haan, L. (2000) Stable probability distributions and their domains of attraction: a direct approach. *Probab. Math. Statist.* **20**, 169–188.

[16] Gnedin, A. (2010) A species sampling model with finitely many types. *Electron. Commun. Probab.* **15**, 79–88.

[17] Green, P. J. (1977) Conditional limit theorems for general branching processes. *J. Appl. Probab.* **14**, 451–463.

[18] Harris, T. E. (1963) *The Theory of Branching Processes.* Die Grundlehren der Mathematischen Wissenschaften **119**, Springer, Berlin.

[19] Huillet, T. and Möhle, M. (2013) On the extended Moran model and its relation to coalescents with multiple collisions. *Theor. Popul. Biol.* **87**, 5–14.

[20] Ibragimov, I. A. and Linnik, Y. V. (1971) *Independent and Stationary Sequences of Random Variables.* Wolters–Noordhoff, Groningen.

[21] Iksanov, A. and Möhle, M. (2008) On the number of jumps of random walks with a barrier. *Adv. Appl. Probab.* **40**, 206–228.

[22] Kallenberg, O. (2002) *Foundations of Modern Probability.* 2nd Edition, Probability and Its Applications, Springer, New York.

[23] Kukla, J. and Möhle, M. (2018) On the block counting process and the fixation line of the Bolthausen–Sznitman coalescent. *Stoch. Process. Appl.* **128**, 939–962.

[24] Lamperti, J. (1958) An occupation time theorem for a class of stochastic processes. *Trans. Amer. Math. Soc.* **88**, 380–387.

[25] Lamperti, J. (1967a) Continuous state branching processes. *Bull. Amer. Math. Soc.* **73**, 382–386.

[26] Lamperti, J. (1967b) The limit of a sequence of branching processes. *Z. W. verw. Geb.* **7**, 271–288.

[27] Mattila, M. and Haukkanen, P. (2014) On the positive definiteness and eigenvalues of meet and join matrices. *Discrete Math.* **326**, 9–19.

[28] Mattila, M. and Haukkanen, P. (2016) Studying the various properties of MIN and MAX matrices - elementary vs. more advanced methods. *Spec. Matrices* **4**, 101–109.

[29] Neveu, J. (1992) A continuous-state branching process in relation with the GREM model of spin glass theory. Rapport interne no 267, École Polytechnique.

[30] Pakes, A. G. (1997) On the recognition and structure of probability generating functions. *Classical and Modern Branching Processes*, *IMA Vol. Math. Appl.* **84**, Springer, New York, 263–284.

[31] Petrov, V. V. (1975) *Sums of Independent Random Variables.* Springer, New York.

[32] Petrov, V. V. (1995) Limit Theorems of Probability Theory. Oxford University Press, Oxford.

[33] Pitman, J. (1997) Partition structures derived from Brownian motion and stable subordinators. *Bernoulli* **3**, 79–96.

[34] Pitman, J. (2006) *Combinatorial Stochastic Processes.* Lecture Notes in Mathematics **1875**, Springer, Berlin.

[35] Pólya, G. (1920) Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem. *Math. Z.* **8**, 171–181.

[36] Revuz, D. and Yor, M. (2005) *Continuous Martingales and Brownian Motion.* Grundlehren der mathematischen Wissenschaften **293**, Corrected Third Printing of the Third Edition. Springer, Berlin.

[37] Roman, S. (1984) *The Umbral Calculus.* Academic Press, Orlando.

[38] Sagitov, S. (1995) A key limit theorem for critical branching processes. *Stoch. Process. Appl.* **56**, 87–100.

[39] Sewastjanow, B. A. (1975) *Verzweigungsprozesse.* Mathematische Monographien **34**. Akademie-Verlag, Berlin.

[40] Shun-Xiang, O. and Röckner, M. (2016) Time inhomogeneous generalized Mehler semigroups and skew convolution equations. *Forum Math.* **28**, 339–376.

[41] Slack, R. S. (1972) Further notes on branching processes with mean 1. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **25**, 31–38.

[42] Steutel F. W. and Van Harn, K. (2004) *Infinite Divisibility of Probability Distributions on the Real Line*. Marcel Dekker, Inc., New York.

[43] Teicher, H. (1955) An inequality on Poisson probabilities. *Ann. Math. Statist.* **26**, 147–149.

[44] van der Vaart, A. W. (1998) *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics 3, Cambridge University Press, Cambridge.

[45] Vatutin, V. A. (1979) A new limit theorem for a critical Bellman-Harris branching process. *Mat. Sb.* **109**, 440–452.

[46] Widder, D. V. (1941) *The Laplace Transform.* Princeton University Press, Princeton.

[47] Zolotarev, V. M. (1957) More exact statements of several theorems in the theory of branching processes. *Theory Probab. Appl.* **2**, 245–253.

Article II

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

Möhle, M. and Vetter, B. (2021) ALEA Lat. Am. J. Probab. Math. Stat. 19, no. 1, 641–664.

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.

Keywords: Block counting process; Bolthausen–Sznitman coalescent; coalescent; dust; fixation line; generalized Ornstein–Uhlenbeck process; time-inhomogeneous process; weak convergence

2020 Mathematics Subject Classification: Primary 60J90 Secondary 60J27

II.1 Introduction

The Λ-coalescent, independently introduced by Pitman [17] and Sagitov [20], 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 [3] 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 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 [17, 24]) for measures Λ that satisfy

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

and for the Bolthausen–Sznitman coalescent [5], 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 II.2, where they are stated as Propositions II.2.1 and II.2.2. This work provides unified proofs of Propositions II.2.1 and II.2.2 and extends the convergence results by combining both proofs. The main result (Theorem II.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 II.2.3 to the class of generalized Ornstein–Uhlenbeck processes, which have been studied extensively in the literature. In [14], 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 [22].

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 [12, 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 [26] to the block counting process, i.e., (see [13])

$$\mathbb{P}(L_t^{(m)} \geq n) \;=\; \mathbb{P}(N_t^{(n)} \leq m), \qquad m, n \in \mathbb{N}, t \geq 0. \tag{II.2}$$

For a thorough definition of the fixation line see [12] and the references therein. Theorem II.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 II.2 the two known convergence results for the block counting process of coalescents with dust (Proposition II.2.1) and the Bolthausen–Sznitman coalescent (Proposition II.2.2) are recalled and the main result (Theorem II.2.3) is stated. In Section II.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 II.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 II.2.1 and II.2.2 in Sections II.5 and II.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 II.5 and II.6 to prove Theorem II.2.3 in Section II.7. Two proofs of Theorem II.2.4 are given in Section II.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.

II.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. (II.1). For further results on Λ-coalescents with dust see [10] and [9]. Proposition II.2.1 below has been established in [9] and [16]. 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

[17, Proposition 26]. In [16] 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 [9] 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 II.2.3.

Proposition II.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*

$$T_t f(x) \;:=\; \mathbb{E}(f(X_{s+t})|X_s = x) \;=\; \mathbb{E}(f(x + X_t)), \tag{II.3}$$

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

$$\psi(v) \;=\; \int_{[0,1]} ((1-u)^{iv} - 1)u^{-2}\,\Lambda(\mathrm{d}u), \qquad v \in \mathbb{R}. \tag{II.4}$$

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 [15] and [13]. 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 II.6 directly. Since the scaling depends on t, the process $X^{(n)}$ is time-inhomogeneous, and in [13] the time-space process is used in order to transfer the question of convergence to time-homogeneous Markov processes. The time-space process is revisited in Section II.6. By constructing the Bolthausen–Sznitman coalescent from a random recursive tree, it is shown in [11] and [2] 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 II.4.2 provided in Section II.4.

Proposition II.2.2 (Bolthausen–Sznitman case). *Suppose that $\Lambda = \lambda$. Then the time-inhomogeneous 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*

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

$x\in\mathbb{R}, f\in B(\mathbb{R}), s,t\geq 0$, 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 non-negative measures. The two non-negative 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 II.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 II.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 [23] shows that the Λ-coalescent does not come down from infinity under Assumption A, see Lemma II.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

$$a \;:=\; b(1+\Psi(1)) \;-\; \int_{[0,1]} u^{-1}\,\Lambda_D(\mathrm{d}u) \tag{II.5}$$

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

$$\psi(v) \;:=\; iav \;+\; \int_{[0,1]} ((1-u)^{iv} - 1 + ivu)u^{-2}\,\Lambda(\mathrm{d}u), \quad v\in\mathbb{R}. \tag{II.6}$$

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

$$\varrho(A) \;:=\; \int_{g^{-1}(A)} u^{-2}\,\Lambda(\mathrm{d}u), \qquad A\in\mathcal{B}, \tag{II.7}$$

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. (II.8) and (II.9) below belongs to the class of generalized Ornstein–Uhlenbeck processes. Due to $\varrho((0,\infty))=0$, the limiting process in Theorem II.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 II.3.

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

$$T_tf(x) \;:=\; \mathbb{E}(f(X_{s+t})|X_s=x) \;=\; \mathbb{E}(f(e^{-bt}x+X_t)), \tag{II.8}$$

$x\in\mathbb{R}, f\in B(\mathbb{R}), s,t\ge 0$, where X_t has characteristic function ϕ_t given by

$$\phi_t(v) \;=\; \exp\biggl(\int_0^t \psi(e^{-bs}v)\,\mathrm{d}s\biggr), \qquad v\in\mathbb{R}, t\ge 0, \tag{II.9}$$

and ψ is given by (II.6).

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 (II.4) and (II.6) for ψ coincide and Proposition II.2.1 and Theorem II.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 II.2.2 and Theorem II.2.3 in the Bolthausen–Sznitman case is clarified in Section II.4.

A convergence result for the fixation line can be stated analogously to Theorem II.2.3; see also [9, Theorem 2.13 b)] for the case $b=0$.

Theorem II.2.4. *Suppose that Λ satisfies Assumption A. Then the possibly time-inhomogeneous Markov process $Y^{(n)} := (Y_t^{(n)})_{t\ge 0} := (\log L_t^{(n)} - e^{bt}\log n)_{t\ge 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n\to\infty$ to the time-homogeneous Markov process $Y=(Y_t)_{t\ge 0}$ with initial value $Y_0=0$ and semigroup $(T_t)_{t\ge 0}$ given by*

$$T_tf(y) \;:=\; \mathbb{E}(f(Y_{s+t})|Y_s=y) \;=\; \mathbb{E}(f(e^{bt}y+Y_t)), \qquad \text{(II.10)}$$

$y\in\mathbb{R}, f\in B(\mathbb{R}), s,t\ge 0$, where Y_t has characteristic function χ_t given by

$$\chi_t(w) \;=\; \exp\left(\int_0^t \psi(-e^{bs}w)\,\mathrm{d}s\right), \qquad w\in\mathbb{R}, t\ge 0, \qquad \text{(II.11)}$$

and ψ is given by (II.6).

Remark II.2.5. The process defined by (II.10) and (II.11) is a generalized Ornstein–Uhlenbeck process with underlying characteristic exponent $v\mapsto\psi(-v)$, $v\in\mathbb{R}$, but with non-negative drift.

Remark II.2.6. Let the random variable S_t have characteristic function ϕ_t, given by (II.9), for $t\ge 0$, and let $X=(X_t)_{t\ge 0}$ and $Y=(Y_t)_{t\ge 0}$ denote the processes defined in Theorems II.2.3 and II.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,

$$\begin{aligned}\mathbb{P}(e^{Y_{t+s}}\ge x\,|\,e^{Y_s}=y) \;&=\; \mathbb{P}(y^{e^{bt}}e^{-e^{bt}S_t}\ge x)\\ &=\; \mathbb{P}(x^{e^{-bt}}e^{S_t}\le y) \;=\; \mathbb{P}(e^{X_{t+s}}\le y\,|\,e^{X_s}=x)\end{aligned}$$

for all $x,y,s,t\ge 0$, i.e., e^Y is Siegmund-dual to e^X (see [26]) parallel to the Siegmund-duality of the block counting process and the fixation line.

Remark II.2.7. For the Bolthausen–Sznitman case, the convergence result corresponding to Theorem II.2.4 is stated in [13, 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 II.2.4 is the logarithm of Neveu's continuous-state branching process. By Proposition II.2.2, the characteristic functions χ_t of the marginal distributions are given by (see [13, Eq. (19)])

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

II.3 The limiting process

Standard computations (see [21, Lemma 17.1]) show that ϕ_t, given by (II.9), 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 [22, Theorem 3.1] provide a formula for the generator corresponding to the semigroup $(T_t)_{t\geq 0}$, given by (II.8).

Lemma II.3.1. *Suppose that Λ satisfies Assumption A. Let ψ be given by* (II.6)*, ϕ_t be defined by* (II.9) *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* (II.8) *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*

$$\begin{aligned} Af(x) \;=\; & f'(x)(a-bx) \\ & + \int_{[0,1]} (f(x+\log(1-u)) - f(x) + uf'(x))u^{-2}\,\Lambda(\mathrm{d}u) \end{aligned} \tag{II.12}$$

for $x \in \mathbb{R}$ and $f \in D$, where a is given by (II.5).

Proof. Substituting $g : (0,1) \to \mathbb{R}$, $g(u) := \log(1-u)$, $u \in (0,1)$, shows that (II.12) is an integro-differential operator of the form (1.1) of [22] with dimension $d = 1$. In [22], 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 [22]), but Step 3 of the proof of [22, Theorem 3.1] shows that (II.12) even holds for functions $f \in D$ ($\supset C_c^2$). Note that the space D is denoted by F_1 in [22]. The fact that D is a core for A is only a different phrasing of the claim in Step 5 of the proof of [22, 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. (II.12).

The limiting process in Theorem II.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 II.3.2 and II.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

$$\mathrm{d}X_t \;=\; -bX_t\mathrm{d}t \;+\; \mathrm{d}L_t, \qquad t \geq 0, \tag{II.13}$$

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 (II.13) or the corresponding semigroup are hence called generalized Ornstein–Uhlenbeck or Ornstein–Uhlenbeck type process or semigroup. It holds that

$$X_t \;=\; e^{-bt}x \;+\; \int_0^t e^{-b(t-s)}\,\mathrm{d}L_s, \qquad t \geq 0. \tag{II.14}$$

Various constructions for the stochastic integral in (II.14) are possible, e.g., in [1, 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 (II.8), where the characteristic functions ϕ_t of X_t are given by (II.9) 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

$$\int_{\{|u|>1\}} \log(1+|u|)\,\varrho(\mathrm{d}u) \;<\; \infty. \tag{II.15}$$

According to [22, 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 (II.15) does not hold, then there exists no stationary distribution. The following lemma is an application of [22, Theorems 4.1 and 4.2] to this article's coalescent setting.

Lemma II.3.2. *Suppose that Λ satisfies Assumption A with $b > 0$ and let $X = (X_t)_{t\geq 0}$ be as in Theorem II.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*

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

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.

Shiga's criterion [25, Theorem 1.1] for transience and recurrence complements Lemma II.3.2.

Lemma II.3.3. *Suppose that* Λ *satisfies Assumption A with* $b > 0$ *and let* $X = (X_t)_{t\geq 0}$ *be as in Theorem II.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*

$$\int_0^1 z^{-1}\exp\Big(-\int_z^1 \frac{g_\Lambda(y)}{by}\,\mathrm{d}y\Big)\mathrm{d}z \tag{II.16}$$

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* (II.16) *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 [4], 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 $s, t \geq 0$.

II.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

$$q_{k,j} = \frac{\Gamma(a+b)\Gamma(k+1)\Gamma(j-1+b)\Gamma(k-j-1+a)}{\Gamma(a)\Gamma(b)\Gamma(k-2+a+b)\Gamma(j)\Gamma(k-j+2)}, \tag{II.17}$$

$j \in \{1,\ldots,k-1\}, k \geq 2$. The $\beta(a,b)$-coalescent comes down from infinity if and only if $0 < a < 1$ [23, 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 e.g. [28, p. 247])

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

of the digamma function.

Proposition II.4.1. *Suppose that* $\Lambda = \beta(1,b)$ *with* $b > 0$. *Let* a, ψ *and* ϱ *be given by* (II.5), (II.6) *and* (II.7), *respectively. Then* ϱ *has density* f *with respect to Lebesgue measure on* $(-\infty, 0)$ *given by* $f(u) := be^{bu}(1-e^{u})^{-2}$, $u < 0$,

$$a \;=\; b(1+\Psi(b)) \tag{II.18}$$

and

$$\psi(v) \;=\; b\big((1-b)\Psi(b) - (1-b-iv)\Psi(b+iv)\big), \qquad v \in \mathbb{R}. \tag{II.19}$$

Proof. It can be easily verified that ϱ has density as stated in the proposition. Eq. (II.18) follows from

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

Next, note that

$$\Psi(b) - \Psi(b+iv) \;=\; \int_0^\infty (e^{-ivu}-1)\frac{e^{-bu}}{1-e^{-u}}\,\mathrm{d}u, \qquad v \in \mathbb{R}. \tag{II.20}$$

Integration by parts yields

$$\begin{aligned}
iv(\Psi(b+iv)-\Psi(b)) \;&=\; \int_0^\infty (iv - ive^{-ivu})\frac{e^{-bu}}{1-e^{-u}}\,\mathrm{d}u \\
&=\; (ivu + e^{-ivu} - 1)\frac{e^{-bu}}{1-e^{-u}}\bigg|_{u=0}^{u=\infty} \\
&\quad - \int_0^\infty (ivu + e^{-ivu}-1)\left(\frac{-be^{-bu}}{1-e^{-u}} - \frac{e^{-bu}}{(1-e^{-u})^2}e^{-u}\right)\mathrm{d}u \\
&=\; \int_0^\infty (e^{-ivu} - 1 + ivu)\frac{e^{-bu}}{(1-e^{-u})^2}\big(1-(1-b)(1-e^{-u})\big)\mathrm{d}u,
\end{aligned}$$

for $v \in \mathbb{R}$. Hence,

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

The calculation

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

and multiplication with b complete the proof of (II.19). □

Example II.4.2. Suppose that $\Lambda = \beta(1,b)$ with $b > 0$. Then Assumption A is satisfied with the same constant b. According to Theorem II.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

$$T_t f(x) \;:=\; \mathbb{E}(f(X_{s+t})|X_s = x) \;=\; \mathbb{E}(f(e^{-bt}x + X_t)),$$

$x \in \mathbb{R}, f \in B(\mathbb{R}), s,t \geq 0$, where X_t has characteristic function ϕ_t given by (II.9).

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 II.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

$$\begin{aligned}\phi(v) &= \exp\biggl(\int_0^\infty \psi(e^{-bs}v)\,\mathrm{d}s\biggr) \qquad\qquad \text{(II.21)}\\ &= \exp\biggl((1-b)\int_0^v \frac{\Psi(b)-\Psi(b+iu)}{u}\,\mathrm{d}u\biggr)\frac{\Gamma(b+iv)}{\Gamma(b)}, \qquad v\in\mathbb{R}.\end{aligned}$$

In the last step Eq. (II.19) 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

$$\begin{aligned}\phi_t(v) &= \frac{\phi(v)}{\phi(e^{-bt}v)}\\ &= \exp\biggl((1-b)\int_{e^{-bt}v}^v \frac{\Psi(b)-\Psi(b+iu)}{u}\,\mathrm{d}u\biggr)\frac{\Gamma(b+iv)}{\Gamma(b+ie^{-bt}v)},\end{aligned}$$

$v \in \mathbb{R}, t \geq 0$. 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 [27, 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 (II.21), 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. (II.20). 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 [27, V, Theorem 6.7], the first factor on the right-hand side of (II.21) is a self-decomposable characteristic function as well, and

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

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 II.4.1 with $b = 1$ states that $\psi(v) = iv\Psi(1 + iv)$, $v \in \mathbb{R}$. Example II.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. [27, 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 II.2.2 and Theorem II.2.3.

II.5 Proof of Proposition II.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

$$A^{(n)}f(x) \;=\; \sum_{j=1}^{k-1}(f(x + \log \tfrac{j}{k}) - f(x))q_{k,j},$$

$x \in E_n, f \in \widehat{C}(\mathbb{R}), n \in \mathbb{N}$. The process $X = (X_t)_{t\geq 0}$ defined by (II.3) and (II.4) is a Feller process in $\widehat{C}(\mathbb{R})$. Let A denote the generator. From [21, 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

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

$x \in \mathbb{R}, f \in \widehat{C}_2(\mathbb{R})$. 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 II.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

$$f(x+\log(1-u)) - f(x) \;=\; -u\int_0^1 \frac{f'(x+\log(1-us))}{1-us}\,\mathrm{d}s,$$

$u \in [0,1), x \in \mathbb{R}$. 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

$$S(k,x) \;:=\; \sum_{j=1}^{k-1}(f(x+\log\tfrac{j}{k}) - f(x))q_{k,j} \tag{II.22}$$

and

$$I(x) \;:=\; \int_{[0,1]} h(u,x)u^{-1}\,\Lambda(\mathrm{d}u) \tag{II.23}$$

for $k \in \mathbb{N}, x \in \mathbb{R}$. 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

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

$k \in \mathbb{N}, x \in \mathbb{R}$. 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

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

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 II.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

$$\lim_{k\to\infty}\sup_{x\in\mathbb{R}}|S(k,x)-I(x)| \;=\; 0. \tag{II.24}$$

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

$$\lim_{x\to-\infty}|I(x)| \;=\; c\lim_{x\to-\infty}|\mathbb{E}(h(Z,x))| \;=\; 0. \tag{II.25}$$

Since $f\in\widehat{C}(\mathbb{R})$, $\lim_{x\to-\infty}S(k,x) = 0$ for any $k\in\mathbb{N}$. Due to (II.24) and (II.25),

$$\lim_{x\to-\infty}\sup_{k\in\mathbb{N}}|S(k,x)| \;=\; 0. \tag{II.26}$$

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

$$\begin{aligned}\lim_{n\to\infty}\sup_{x\in E_n}|A^{(n)}f(x)-Af(x)| \;&\le\; \lim_{k\to\infty}\sup_{x\in\mathbb{R}}|S(k,x)-I(x)|\\ &\quad+\;\lim_{x\to-\infty}\sup_{k\in\mathbb{N}}|S(k,x)| \;+\; \lim_{x\to-\infty}|I(x)| \;=\; 0.\end{aligned} \tag{II.27}$$

By [8, I, Theorem 6.1 and IV, Theorem 2.5], $X^{(n)}\to X$ in $D_{\mathbb{R}}[0,\infty)$ as $n\to\infty$. □

Remark II.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 $q_{k,j} = \binom{k}{j-1}\int_{[0,1)}u^{k-j-1}(1-u)^{j-1}\,\Lambda|_{[0,1)}(\mathrm{d}u)+\Lambda(\{1\})1_{\{1\}}(j)$, $j\in\{1,\ldots,k-1\}$, $k\ge 2$, 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,

$$A^{(n)}f(x) \;=\; S(k,x)+(f(\log n^{-1})-f(x))\Lambda(\{1\})$$

$x \in E_n, f \in \widehat{C}(\mathbb{R}), n \in \mathbb{N}$, where the jump rates in $S(k,x)$ correspond to $\Lambda|_{[0,1)}$, and

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

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 (II.27) holds true.

Remark II.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.

II.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.

II.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, e.g., [19, p. 85, Exercise (1.10)] or [6]) 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 $\widetilde{E}_n := \{(s,x) \in [0,\infty) \times \mathbb{R} : e^x n^{\alpha(s)} \in [n]\}$ denote the state space of $\widetilde{X}^{(n)}$, $\widetilde{E} := [0,\infty) \times \mathbb{R}$ denote the state space of $\widetilde{X}$ and define $k := k(s,x,n) := e^x n^{\alpha(s)} \in \mathbb{N}$ 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

$$\widetilde{T}_t f(s,x) \;=\; \mathbb{E}(f(s+t, X_{s+t})|X_s = x) \;=\; \mathbb{E}(f(s+t, \alpha(t)x + X_t))$$

for $(s,x) \in \widetilde{E}$, $f \in B(\widetilde{E})$ and $t \geq 0$, 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 II.9.6 states that $\widetilde{D}$ is a core for the generator $\widetilde{A}$ of $(\widetilde{T}_t)_{t\geq 0}$ and

$$\widetilde{A}f(s,x) \;=\; \frac{\partial}{\partial s}f(s,x) \;+\; A\pi f(s,x), \qquad (s,x) \in \widetilde{E}, f \in \widetilde{D}. \quad \text{(II.28)}$$

The "semigroup" $(T_{s,t}^{(n)})_{s,t\geq 0}$ of $X^{(n)}$ is given by

$$\begin{aligned} T_{s,t}^{(n)} f(x) \;&:=\; \mathbb{E}(f(X_{s+t}^{(n)})|X_s^{(n)} = x) \\ &=\; \mathbb{E}(f(\log N_{s+t}^{(n)} - \alpha(s+t)\log n)|N_s^{(n)} = k) \\ &=\; \mathbb{E}(f(\log N_t^{(k)} - \alpha(s+t)\log n)), \end{aligned}$$

$(s,x) \in \widetilde{E}_n, f \in B(\mathbb{R}), t \geq$. The "generator" $(A_s^{(n)})_{s\geq 0}$ of $(T_{s,t}^{(n)})_{s,t\geq 0}$ is given by

$$\begin{aligned} A_s^{(n)} f(x) \;&:=\; \lim_{t\to 0} t^{-1}(T_{s,t}^{(n)} f(x) - f(x)) \\ &=\; \lim_{t\to 0} t^{-1}\Big(\mathbb{E}\Big(f(\log N_t^{(k)} - \alpha(s+t)\log n)\Big) - f(x)\Big) \\ &=\; -f'(x)\alpha'(s)\log n + \sum_{j=1}^{k-1}(f(x+\log \tfrac{j}{k}) - f(x))q_{k,j}, \end{aligned} \quad \text{(II.29)}$$

$(s,x) \in \widetilde{E}_n$ 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

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

$(s,x) \in \widetilde{E}_n, f \in B(\widetilde{E}_n), t \geq 0, n \in \mathbb{N}$, 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

$$\widetilde{A}^{(n)} f(s,x) \;=\; \frac{\partial}{\partial s}f(s,x) \;+\; A_s^{(n)}\pi f(s,x), \quad \text{(II.30)}$$

$(s,x) \in \widetilde{E}_n, n \in \mathbb{N}$.

II.6.2 Proof of Proposition II.2.2

Proof. (of Proposition II.2.2) Recall that $\Lambda = \lambda$. Let $f \in D$. The approach to the proof is the same as in Section II.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 (II.12). 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

$$h(u,x) = u^{-2} \int_0^u \frac{u-s}{(1-s)^2}(f''(x+\log(1-s)) - f'(x+\log(1-s)))\mathrm{d}s$$
$$= \int_0^1 \frac{1-s}{(1-us)^2}(f''(x+\log(1-us)) - f'(x+\log(1-us)))\mathrm{d}s,$$

$u \in [0,1), x \in \mathbb{R}$. 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 (II.29) 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

$$R(k,x) \;:=\; \log k \;-\; \sum_{j=1}^{k-1} \tfrac{k-j}{k} q_{k,j} \;-\; x, \qquad k \in \mathbb{N}, x \in \mathbb{R}, \qquad \text{(II.31)}$$

and

$$S(k,x) \;:=\; \sum_{j=1}^{k-1} (f(x+\log \tfrac{j}{k}) - f(x) + \tfrac{k-j}{k} f'(x)) q_{k,j}, \qquad \text{(II.32)}$$

$k \in \mathbb{N}, x \in \mathbb{R}$. 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. (II.17) with $a = b = 1$, $\frac{k-j}{k} q_{k,j} = (k-j+1)^{-1}$, $j \in \{1,\dots,k-1\}, k \geq 2$. 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

$$\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} |f'(x)||R(k,x) - (1+\Psi(1)-x)| \;=\; 0. \qquad \text{(II.33)}$$

In the next step the uniform convergence of $S(k,x)$ to $I(x)$ is shown. Substituting $k-j-1$ for j in (II.32) yields

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

$k \in \mathbb{N}, x \in \mathbb{R}$. 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

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

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

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

$k \in \mathbb{N}, x \in \mathbb{R}$, 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 II.9.4 it follows that

$$\lim_{k\to\infty} \sup_{x\in\mathbb{R}} |S(k,x) - I(x)| \;=\; 0. \qquad \text{(II.34)}$$

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

$$\lim_{x\to-\infty} |I(x)| \;=\; c \lim_{x\to-\infty} |\mathbb{E}(h(Z,x))| \;=\; 0. \qquad \text{(II.35)}$$

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 (II.34) and (II.35),

$$\lim_{x\to-\infty} \sup_{k\in\mathbb{N}} |S(k,x)| \;=\; 0. \tag{II.36}$$

As seen in the proof of Proposition II.2.1, Eqs. (II.34)-(II.36) imply

$$\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} |S(k,x) - I(x)| \;=\; 0. \tag{II.37}$$

By (II.33), $\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} |A_s^{(n)} f(x) - Af(x)| = 0$. Due to (II.28) and (II.30),

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

for every function f belonging to the core $\widetilde{D}$ and each $T > 0$. From [8, 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 II.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 II.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 II.2.2 the fact that $\Lambda = \lambda$ is only used to verify (II.33). Eq. (II.37) 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.

II.7 Proof of Theorem II.2.3

In this section Λ satisfies Assumption A. We continue to use the time-space setting and the notation of Subsection II.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 II.2.3 now only requires to suitable arrange equations already obtained in Sections II.5 and II.6. To be precise, the results of Section II.5 are applied to the summands $\Lambda_D^\pm$ of Λ_D, but we omit this detail in the following.

Proof. (of Theorem II.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 (II.29) 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

$$A_s^{(n)} f(x) \;=\; bR(k,x)f'(x) \;+\; bS_{BS}(k,x) \;+\; S_D(k,x),$$

$(s,x) \in \widetilde{E}_n, n \in \mathbb{N}$, where

$$R(k,x) \;:=\; \log k \;-\; \sum_{j=1}^{k-1} \tfrac{k-j}{k} q_{k,j}^\lambda \;-\; x,$$

$$S_{BS}(k,x) \;:=\; \sum_{j=1}^{k-1} (f(x+\log\tfrac{j}{k}) - f(x) + \tfrac{k-j}{k} f'(x)) q_{k,j}^\lambda,$$

$$S_D(k,x) \;:=\; \sum_{j=1}^{k-1} (f(x+\log\tfrac{j}{k}) - f(x)) q_{k,j}^D,$$

are defined as in (II.31), (II.32), (II.22) and (II.23) 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 II.3.1 and Eq. (II.5), the generator A of $X = (X_t)_{t\geq 0}$ can be written as

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

From (II.33), (II.37) and (II.24)-(II.26) 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 (II.28) and (II.30),

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

for every $f \in \widetilde{D}$ and $T > 0$. By Proposition II.9.6, the space $\widetilde{D}$ is a core for $\widetilde{A}$. Thus, it follows from [8, 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 $n \to \infty$. □

II.8 Proof of Theorem II.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 II.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 II.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. (II.2) and Theorem II.2.3 imply that

$$\begin{aligned}\mathbb{P}(Y_t^{(n)} \geq y) &= \mathbb{P}(L_t^{(n)} \geq m)\\ &= \mathbb{P}(N_t^{(m)} \leq n) \;=\; \mathbb{P}(X_t^{(m)} \leq \log n - e^{-bt}\log m)\\ &\longrightarrow \mathbb{P}(X_t \leq -e^{-bt}y) \;=\; \mathbb{P}(-e^{bt}X_t \geq y),\end{aligned} \qquad \text{(II.38)}$$

$y \in \mathbb{R}, t \geq 0$, 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 (II.38) 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

$$\begin{aligned}\widetilde{T}_t^{(n)} f(s,y) &= \mathbb{E}(f(s+t, Y_{s+t}^{(n)}) \,|\, Y_s^{(n)} = y)\\ &= \mathbb{E}(f(s+t, \log L_t^{(k)} - e^{b(t+s)}\log n))\\ &= \mathbb{E}(f(s+t, e^{bt}y + Y_t^{(k)})),\end{aligned}$$

$(s,y) \in \widetilde{E}_n, f \in B(\widetilde{E}), t \geq 0, n \in \mathbb{N}$. 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 (II.11). 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 [18, 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,

$$\lim_{n\to\infty} \sup_{(s,y)\in\widetilde{E}_n} |\widetilde{T}_t^{(n)} f(s,y) - \widetilde{T}_t f(s,y)| \ = \ 0. \qquad \text{(II.39)}$$

By linearity, Eq. (II.39) 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., [7]) it is a dense subset of $B(\widetilde{E})$. Hence, (II.39) holds true for $f \in B(\widetilde{E})$. [8, 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 (II.10) and (II.11) is a generalized Ornstein–Uhlenbeck process (with non-negative linear drift) as in [22]. The underlying infinitely divisible distribution has characteristic exponent $v \mapsto \psi(-v)$, $v \in \mathbb{R}$. According to [22, Theorem 3.1], D is a core for the corresponding generator A and

$$\begin{aligned} Af(y) \ &= \ f'(y)(-a+by) \qquad \text{(II.40)} \\ &\quad + \int_{[0,1]} (f(y - \log(1-u)) - f(y) - uf'(y))u^{-2}\,\Lambda(\mathrm{d}u) \end{aligned}$$

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

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

$$\begin{aligned} A_s^{(n)} f(y) \ &= \ -f'(y)be^{bs}\log n \\ &\quad + \sum_{j > e^y n^{e^{bs}}} (f(\log j - e^{bs}\log n) - f(y))\gamma_{e^y n^{e^{bs}},j} \end{aligned}$$

for $(s,y)\in \widetilde{E}_n$ and $n\in\mathbb{N}$. 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

$$\begin{aligned} A_s^{(n)}f(y) &= bf'(y)(-\log k+y)\\ &\quad + \sum_{j=1}^{\infty}(f(y+\log(1+\tfrac{j}{k}))-f(y))\gamma_{k,k+j} \end{aligned}$$

for $(s,y)\in\widetilde{E}_n$ and $n\in\mathbb{N}$. Define $\Lambda_D:=\Lambda-b\lambda$ and let Λ_D^+,Λ_D^- denote the non-negative 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

$$A_s^{(n)}f(y) \;=\; bf'(y)R(k,y) \;+\; bS_{BS}(k,y) \;+\; S_D(k,y), \qquad \text{(II.41)}$$

$(s,y)\in\widetilde{E}_n, n\in\mathbb{N}$, where

$$\begin{aligned} R(k,y) &:= -\log k \;+\; y \;+\; \sum_{j=1}^{k}\tfrac{j}{k}\gamma_{k,k+j}^{\lambda},\\ S_{BS}(k,y) &:= \sum_{j=1}^{\infty}(f(y+\log(1+\tfrac{j}{k}))-f(y)-\tfrac{j}{k}1_{[0,1]}(\tfrac{j}{k})f'(y))\gamma_{k,k+j}^{\lambda},\\ S_D(k,y) &:= \sum_{j=1}^{\infty}(f(y+\log(1+\tfrac{j}{k}))-f(y))\gamma_{k,k+j}^{D}, \end{aligned}$$

for $k\in\mathbb{N},y\in\mathbb{R}$ and $f\in C_1(\mathbb{R})$ such that $f,f'\in\widehat{C}(\mathbb{R})$. Using the decomposition of Λ on Eq. (II.40) yields

$$Af(y) \;=\; bf'(y)(-1-\Psi(1)+y) \;+\; bI_{BS}(y) \;+\; I_D(y),$$

$y\in\mathbb{R}, f\in D$, where

$$\begin{aligned} I_{BS}(y) &:= \int_{[0,1]}(f(y-\log(1-u))-f(y)-uf'(y))u^{-2}\,\lambda(\mathrm{d}u),\\ I_D(y) &:= \int_{[0,1]}(f(y-\log(1-u))-f(y))u^{-2}\,\Lambda_D(\mathrm{d}u) \end{aligned}$$

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

$$\lim_{k\to\infty}\sup_{y\in\mathbb{R}}|R(k,y)-(-1-\Psi(1)+y)| \;=\; 0. \qquad \text{(II.42)}$$

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,

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

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 $\int_{[0,1]} u^{-2}(u - \frac{u}{1-u}1_{[0,1/2]}(u))\,\lambda(\mathrm{d}u) = \int_0^{1/2} -(1-u)^{-1}\,\mathrm{d}u + \int_{1/2}^1 u^{-1}\,\mathrm{d}u = 0$. 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 II.9.4 it follows that

$$\lim_{k\to\infty} \sup_{y\in\mathbb{R}} |S_{BS}(k,y) - I_{BS}(y)| \;=\; 0. \tag{II.43}$$

The measure Λ_D is real-valued. Eq. (II.44) 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

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

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,

$$S_D(k,y) \;=\; \sum_{j=0}^{\infty}(f(y+\log(1+\tfrac{j}{k})) - f(y))\gamma^D_{k,k+j}$$
$$= \; c\mathbb{E}\Big(h_D(1-(1+Z_k/k)^{-1}, y)\big(1-(1+Z_k)^{-1}\big)\Big),$$

$k \in \mathbb{N}, y \in \mathbb{R}$. 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 II.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,

$$\lim_{k\to\infty}\sup_{y\in\mathbb{R}} |S_D(k,y) - I_D(y)| \;=\; 0. \qquad \text{(II.44)}$$

Note that Eq. (II.44) holds true for $c = 0$ as well.

Taking into account that $k = e^y n^{e^{bs}} \geq n$ for $(s,y) \in \widetilde{E}_n$ and $n \in \mathbb{N}$, Eqs. (II.41)-(II.44) imply

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

The time-space variant of [8, IV, Corollary 8.7] as implemented in the proof of Theorem II.2.3 yields the desired convergence of $Y^{(n)}$ to Y in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$. □

II.9 Appendix

Lemma II.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,

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

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

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

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 [23, Corollary 2]. □

Remark II.9.2. Any converse statements of Lemma II.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 II.5-II.8 to prove the uniform convergence of generators. In the statement the notion of equicontinuity is used, whose definition is first recalled.

Definition II.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 II.9.4. *Let $X, X_1, X_2, \dots$ 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 $n \geq n_0$ and $\mathbb{P}(X \notin [\delta, 1-\delta]) < \varepsilon/(4M)$. For $f \in F$ define $\tilde{f} : [0,1] \to \mathbb{R}$ via $\tilde{f}(u) := f(\delta)$, $0 \leq u \leq \delta$, $\tilde{f}(u) := f(u)$, $\delta \leq u \leq 1-\delta$, and $\tilde{f}(u) := f(1-\delta)$, $1-\delta \leq u \leq 1$. Then $\{\tilde{f} : f \in F\}$ is bounded (by M) and equicontinuous on $[0,1]$. [18, 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

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

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

Remark II.9.5. In [18, 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 II.9.6. *Suppose that $(T_t)_{t\geq 0}$ 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 $(\widetilde{T}_t)_{t\geq 0}$*, defined via*

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

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* $i=1,\dots,l$*. Then* $\widetilde{D}$ *is a core for the generator* $\widetilde{A}$ *of* $(\widetilde{T}_t)_{t\geq 0}$ *and*

$$\widetilde{A}f(s,x) \;=\; \frac{\partial}{\partial s}f(s,x) \;+\; A\pi f(s,x), \qquad (s,x)\in\widetilde{E}, f\in\widetilde{D}. \quad \text{(II.45)}$$

Proof. Observe that all functions involved in the proof are bounded and uniformly continuous. Clearly, the right-hand side of (II.45) 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 $i=1,\dots,l$. The algebra D_0 separates points and vanishes nowhere. The Stone–Weierstrass theorem for locally compact spaces (e.g. [7]) ensures that D_0 is a dense subset of $\widehat{C}(\widetilde{E})$. In [7] 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 $\|f_n-f\|\leq\|k_n-f\|\to 0$ 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

$$\begin{aligned} t^{-1}(\widetilde{T}_t g(s)h(x)-g(s)h(x)) \;&=\; t^{-1}(g(s+t)-g(s))h(x) \\ &\quad+\; g(s+t)t^{-1}(T_t h(x)-h(x)) \end{aligned}$$

converges uniformly in $(s,x)\in\widetilde{E}$ to $g'(s)h(x)+g(s)Ah(x)$ as $t\to 0+$, thus $\widetilde{D}$ lies in the domain of $\widetilde{A}$ and (II.45) 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 $g_i(s)=c_i\exp(-a_i s)$, $s\in[0,\infty)$ with $c_i\in\mathbb{R}$ and $a_i>0$ and $h_i\in D$ for $i=1,\dots,l$, is a dense subset of $\widehat{C}(\widetilde{E})$. By Hille–Yosida theory (see, e.g., [8, 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 (II.45),

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

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

Remark II.9.7. The last part of the proof of Proposition II.9.6 can be simplified under the additional assumption that $T_t D \subseteq D$ for every $t > 0$. Then $\widetilde{T}_t \widetilde{D} \subseteq \widetilde{D}$ for every $t \geq 0$ and the claim follows by applying the core theorem [8, I, Proposition 3.3].

Bibliography

[1] Applebaum, D. (2009) *Lévy Processes and Stochastic Calculus.* Cambridge University Press, Cambridge.

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

[3] Berestycki, N. (2009) *Recent progress in coalescent theory.* Sociedade Brasileira de Matemática, Rio de Janeiro.

[4] Bogachev, V. I., Röckner, M. and Schmuland, B. (1996) Generalized Mehler semigroups and applications. *Probab. Theory Related Fields* **105**, no. 2, 193–225.

[5] Bolthausen, E. and Sznitman, A.-S. (1998) On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197**, no. 2, 247–276.

[6] Böttcher, B. (2014) Feller evolution systems: generators and approximation. *Stoch. Dyn.* **14**, no. 3, 15 pp.

[7] de Branges, L. (1959) The Stone–Weierstrass theorem. *Proc. Amer. Math. Soc.* **10**, no. 5, 822–824.

[8] Ethier, S. N. and Kurtz, T. G. (1986) *Markov Processes, Characterization and Convergence.* Wiley.

[9] Gaiser, F. and Möhle, M. (2016) On the block counting process and the fixation line of exchangeable coalescents. *ALEA Lat. Am. J. Probab. Math. Stat.* **13**, no. 2, 809–833.

[10] Gnedin, A., Iksanov, A. and Marynych, A. (2011) On Λ-coalescents with dust component. *J. Appl. Probab.* **48**, no. 4, 1133–1151.

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

[12] Hénard, O. (2015) The fixation line in the Λ-coalescent. *Ann. Appl. Probab.* **25**, no. 5, 3007–3032.

[13] Kukla, J. and Möhle, M. (2018) On the block counting process and the fixation line of the Bolthausen–Sznitman coalescent. *Stochastic Process. Appl.* **128**, no. 3, 939–962.

[14] Limic, V. and Talarczyk, A. (2015) Second order asymptotics for the block counting process in a class of regularly varying Λ-coalescents. *Ann. Probab.* **43**, no. 3, 1419–1455.

[15] Möhle, M. (2015) 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**, no. 1, 35–53.

[16] Möhle, M. (2021) The rate of convergence of the block counting process of exchangeable coalescents with dust. *ALEA Lat. Am. J. Probab. Math. Stat.* **18**, no. 2, 1195–1220.

[17] Pitman, J. (1999) Coalescents with multiple collisions. *Ann. Probab.* **27**, no. 4, 1870–1902.

[18] Ranga Rao, R. (1962) Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33**, no. 2, 659–680.

[19] Revuz, D. and Yor, M. (1999) *Continuous Martingales and Brownian Motion.* Springer-Verlag, Berlin.

[20] Sagitov, S. (1999) The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36**, no. 4, 1116–1125.

[21] Sato, K.-I. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

[22] Sato, K.-I. and Yamazato, M. (1984) Operator-selfdecomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type. *Stochastic Process. Appl.* **17**, no. 1, 73–100.

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

[24] Schweinsberg, J. (2000) Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5**, no. 12, 1–50.

[25] Shiga, T. (1990) A recurrence criterion for Markov processes of Ornstein–Uhlenbeck type. *Probab. Theory Related Fields* **85**, no. 4, 425–447.

[26] Siegmund, D. (1976) The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probab.* **4**, no. 6, 914–924.

[27] Steutel, F. W. and van Harn, K. (2004) *Infinite Divisibility of Probability Distributions on the Real Line.* Marcel Dekker, Inc., New York.

[28] Whittaker, E. T. and Watson, G. N. (1996) *A Course of Modern Analysis.* Cambridge University Press, Cambridge.

Article III

Scaling limits for a class of regular Ξ-coalescents

Möhle, M. and Vetter, B. (2023) Stochastic Process. Appl. 162, 387–422.

Abstract.

Let $N_t^{(n)}$ denote the number of blocks in a Ξ-coalescent restricted to a sample of size $n \in \mathbb{N}$ after time $t \geq 0$. Under the assumption of a certain curvature condition on a function well known from the literature, we prove the existence of sequences $(v(n,t))_{n\in\mathbb{N}}$ for which $(\log N_t^{(n)} - \log v(n,t))_{t\geq 0}$ converges to an Ornstein–Uhlenbeck type process as $n \to \infty$. The curvature condition is intrinsically related to the behavior of Ξ near the origin. The method of proof is to show the uniform convergence of the associated generators. Via Siegmund duality an analogous result for the fixation line is proven. Several examples are studied.

Keywords: Block counting process; fixation line; Ornstein–Uhlenbeck type process; regular coalescent; simultaneous multiple collisions; time-inhomogeneous process; weak convergence

2020 Mathematics Subject Classification: Primary 60J90 Secondary 60J27

III.1 Introduction

Exchangeable coalescents are continuous-time Markov processes taking values in the space $\mathcal{P}$ of partitions of $\mathbb{N} := \{1,2,\ldots\}$, where blocks merge over time. Their distribution is determined by a finite measure Ξ on the infinite simplex $\Delta := \{(u_1,u_2,\ldots) : u_1 \geq u_2 \geq \cdots \geq 0, \sum_{i\geq 1} u_i \leq 1\}$. Coalescents can be constructed from appropriate Poisson point processes (Schweinsberg [31]), which allows to identify the class of exchangeable coalescents with the class of finite measures Ξ on Δ. In the Cannings model [6, 7], a discrete-time haploid population model with non-overlapping generations and finite, constant population size, individuals of the same generation follow an exchangeable reproduction law, independently of the other generations. Start with a sample of individuals in one generation and put members into the same block when they have a common parent one generation in the past. We obtain a discrete-time partition-valued ancestral process by merging individuals who share a common ancestor when going backwards further in time, and the coagulation of ancestral lineages corresponds to the merging of blocks. Under suitable conditions exchangeable coalescents then arise as the weak limit of these ancestral processes, properly time-scaled, as the total population size tends to infinity, since a certain form of consistency relation holds for Cannings models, see [26].

Most coalescents treated in the literature belong to one of the following subclasses. The coalescent $(\Pi_t)_{t\geq 0}$, starting from an infinite number of blocks, is said to come down from infinity if the number of blocks is finite at all times $t > 0$ almost surely, and it is said to stay infinite if the number of blocks is infinite at all times $t > 0$ almost surely. For coalescents with dust the number of original blocks that have not been involved in any merger up to time $t > 0$ is infinite with positive probability. Schweinsberg [31] determined conditions to decide on the schemes.

Let Λ be a finite measure on the unit interval $[0,1]$. The Λ-coalescent, which allows only for multiple but not for simultaneous multiple mergers of ancestral lineages, is the particular Ξ-coalescent, where the measure Ξ on Δ is concentrated on $[0,1]\times\{0\}\times\{0\}\cdots$ with $\Xi(B\times\{0\}\times\{0\}\times\cdots) := \Lambda(B)$ for all Borel sets $B\subseteq[0,1]$.

Suppose that $(\Pi_t)_{t\geq 0}$ is standard, i.e., Π_0 is the partition of $\mathbb{N}$ into singletons. For $t\geq 0$ and $n\in\mathbb{N}$ the restriction $\Pi_t^{(n)} := \{B\cap[n] : B\in\Pi_t, B\cap[n]\neq\emptyset\}$ of Π_t to $[n] := \{1,\ldots,n\}$ has values in the space $\mathcal{P}_n$ of partitions of $[n]$. Suppose that $\Pi^{(n)} := (\Pi_t^{(n)})_{t\geq 0}$ is in a state with $k\in[n]$ blocks. For $j\geq 1$, $k_1\geq\cdots\geq k_j$ with $k_1+\cdots+k_j = k$ and $k_1\geq 2$ we speak of a $(k_1,\ldots,k_j)$-collision when $\Pi^{(n)}$ jumps to a state with j blocks and $k_1,\ldots,k_j$ blocks merge into single blocks, respectively. Next we introduce some standard notation. Define $|u| := \sum_{i\geq 1} u_i$ and $(u,u) := \sum_{i\geq 1} u_i^2$ for $u\in\Delta$, $0 := (0,0,\ldots)\in\Delta$, $a := \Xi(\{0\})$, and the measures Ξ_0 and ν via $\Xi = a\varepsilon_0 + \Xi_0$ and $\nu(\mathrm{d}u) := \Xi_0(\mathrm{d}u)/(u,u)$. A $(k_1,\ldots,k_j)$-collision, $j\in\mathbb{N}$, $k_1\geq\cdots\geq k_j$ with $k_1\geq 2$, occurs at the rate (Schweinsberg [31])

$$\begin{aligned}\phi_j(k_1,\ldots,k_j) &= a1_{\{r=1,k_1=2\}}\\ &\quad + \int_\Delta \sum_{l=0}^{s}\binom{s}{l}(1-|u|)^{s-l}\sum_{i_1\neq\cdots\neq i_{r+l}} u_{i_1}^{k_1}\cdots u_{i_{r+l}}^{k_{r+l}}\,\nu(\mathrm{d}u),\end{aligned}$$

where $s := |\{i\in[j] : k_i = 1\}|$ and $r := j-s$.

The aim of this work is to analyze the block counting process $N^{(n)} := (N_t^{(n)})_{t\geq 0} := (|\Pi_t^{(n)}|)_{t\geq 0}$ for large initial state, more precisely, to determine scaling functions $v(n,t)$ for which $N_t^{(n)}/v(n,t)$ converges in distribution as $n\to\infty$. For coalescents with dust it is proven in [14] and [25] with different methods that $(N_t^{(n)}/n)_{t\geq 0}$ converges in the space $D_{[0,1]}[0,\infty)$ of càdlàg paths endowed with the Skorohod topology to the so-called frequency of singletons process as $n\to\infty$. The Bolthausen–Sznitman coalescent in which the driving measure Λ is the uniform distribution on $[0,1]$ has been thoroughly studied in the literature and is an example of a dust-free Λ-coalescent that stays infinite. Goldschmidt and Martin [16] and Baur and Bertoin [1] proved for every $t\geq 0$ the almost sure convergence of $N_t^{(n)}/n^{e^{-t}}$ as $n\to\infty$. This almost sure convergence follows from the construction of the Bolthausen–Sznitman coalescent as clusters of path-connected vertices in a random recursive tree by removing edges at random as time evolves. In [24], it is shown via exact moment calculations that $(N_t^{(n)}/n^{e^{-t}})_{t\geq 0}$ converges in $D_{[0,\infty)}[0,\infty)$ as $n\to\infty$. In [27], the authors obtain the convergence of the scaled block counting process in the Skorohod space for a more general class of Λ-coalescents, where Λ is essentially a beta distribution with parameters 1 and $b>0$.

We extend the results of [27] not only to a larger class of Λ-coalescents but even to a large class of Ξ-coalescents. Our key assumption (III.5) covers the class of Λ-coalescents treated in [27], as shown in Section II.4. The coalescents treated in this paper stay infinite, most coalescents with dust are included but many dust-free coalescents are covered as well. The key assumption (III.5) involves a certain rate function γ known from the literature, which roughly speaking describes the expected size of a jump of the block counting process. The main result (Theorem III.2.7) states that, for a properly chosen scaling $v(n,t)$, the process $(\log N_t^{(n)} - \log v(n,t))_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n\to\infty$ to an Ornstein–Uhlenbeck type process. For information on Ornstein–Uhlenbeck type processes we refer the reader exemplary to [30].

The work of Limic [21] is concerned with the small-time behavior of the block counting process $(N_t)_{t\geq 0} := (|\Pi_t|)_{t\geq 0}$ of Ξ-coalescents $(\Pi_t)_{t\geq 0}$ that come down from infinity. See also [2] and [22] for Λ-coalescents. Under the regularity condition (cf. [21, Eq. (R)])

$$\int_\Delta |u|^2\,\nu(\mathrm{d}u) \;<\; \infty, \tag{III.1}$$

a speed $v(t)$ of coming down of infinity is defined for which $N_t/v(t)$ converges almost surely as $t \to 0+$. The scaling $v(n,t)$ in our main convergence result (Theorem III.2.7) is defined similarly to the speed $v(t)$.

The fixation line $(L_t)_{t\geq 0}$ has been introduced for Λ-coalescents by Hénard [18] and further studied in [14] for Ξ-coalescents. It can be characterized as the Siegmund dual [32] of the block counting process satisfying ([14, Theorem 2.9])

$$\mathbb{P}(L_t^{(m)} \geq n) \;=\; \mathbb{P}(N_t^{(n)} \leq m), \qquad m,n \in \mathbb{N}, t \geq 0, \tag{III.2}$$

where the upper indices denote the initial states $L_0^{(m)} = m$ and $N_0^{(n)} = n$, respectively. Theorem III.2.10 states the convergence of the fixation line in the Skorohod space after suitable scaling.

The paper is organized as follows. The results are presented in Section III.2. In Subsection III.2.1 the function γ and the key assumption (III.5) are treated. The scaling $v(n,t)$ is defined in Subsection III.2.2 and certain properties of the scaling are collected. In Subsection III.2.3 the block counting process is revisited and the main convergence result is stated. Subsection III.2.4 provides the analogous convergence result for the fixation line. Subsection III.2.5 summarizes the obtained convergence and duality results in non-logarithmic form. Several illustrating examples are provided in Section II.4, including an example which clarifies the relation to the results in [27] for a class of Λ-coalescents and including examples of Ξ-coalescents with discrete measure Ξ. The proofs are provided in Section III.4 in the order of appearance of the respective results. The approach to prove the main convergence result is to show the uniform convergence of the associated infinitesimal generators.

III.2 Results

III.2.1 The rate function γ

The following function γ has been proven to be of great significance to the study of coalescents, see [19] and, although in different form, [21] for Ξ-coalescents, and [2, 9, 10, 22] for Λ-coalescents. Define $\gamma : [0,\infty) \to \mathbb{R}$ via

$$\gamma(x) \;:=\; a\binom{x}{2} \;+\; \int_\Delta \sum_{i\geq 1} \big((1-u_i)^x - 1 + xu_i\big)\nu(\mathrm{d}u), \tag{III.3}$$

for $x \geq 0$. The main reason why the function γ is so important to the study of exchangeable coalescent processes is the fact that, if the coalescent is in a state with $k \in \mathbb{N}$ blocks, then (see the forthcoming Eq. (III.14)) $\gamma(k)$ is the expected rate of decrease of the number of blocks. The properties of γ collected in the following lemma are essentially known from the (above cited) literature.

Lemma III.2.1. *Let γ be defined by* (III.3)*. Then $\gamma(0) = \gamma(1) = 0$. Moreover, $\gamma(x) > 0$ for $x > 1$, $\gamma(x) \leq x(x-1)(a/2 + \Xi_0(\Delta))$ for $x \geq 2$, and $\gamma \in C_\infty((0,\infty))$ with derivative*

$$\gamma'(x) \;=\; a\Big(x-\frac{1}{2}\Big) + \int_\Delta \sum_{i\geq 1} \big((1-u_i)^x \log(1-u_i)+u_i\big)\nu(\mathrm{d}u), \quad x > 0,$$

and higher derivatives

$$\gamma^{(k)}(x) \;=\; a\delta_{k2} \;+\; \int_\Delta \sum_{i\geq 1}(1-u_i)^x\big(\log(1-u_i)\big)^k\nu(\mathrm{d}u),$$

$x > 0, k \in \mathbb{N}\setminus\{1\}$, where δ_{kl} denotes the Kronecker symbol. The map $x \mapsto \gamma(x)/x$ is strictly increasing on $[1,\infty)$. In particular, the map γ is strictly increasing on $[1,\infty)$.

We now introduce a parameter which will turn out to be of fundamental interest for our purposes. Define

$$\kappa \;:=\; \lim_{x\to\infty} x\gamma''(x) \;\in\; [0,\infty] \tag{III.4}$$

whenever this limit exists in $[0,\infty]$. In this case we call κ the *asymptotic curvature* of γ or simply the *curvature parameter* of the underlying Ξ-coalescent. Proposition III.2.3 shows that (III.4) is intrinsically related to the behavior of the measure Ξ near $0 \in \Delta$. In Section III.3 the curvature parameter κ is computed for several examples.

Let us briefly comment on the coming down from infinity (cdi) property of the coalescent. Some important coalescents, for example all beta coalescents (see Example III.3.1) and all NLG-coalescents (see Example III.3.2), come down from infinity if and only if $\kappa = \infty$. Note however that, in general, neither $\kappa = \infty$ implies cdi (see Example III.3.5) nor cdi implies $\kappa = \infty$ (see Example 6.1 b) of [19]).

Lemma III.2.1 implies that, up to multiplicative constants, $\gamma(x)$ lies for all sufficiently large x in between x and $x(x-1)$. The key assumption (III.5) of our convergence theorem (Theorem III.2.7) is a more precise condition for the growth of $\gamma(x)$, see (III.6), and can be compactly stated in terms of the curvature of γ as follows.

$$\text{The limit } \kappa \text{ in (III.4) exists and is finite.} \tag{III.5}$$

Using Lemma III.2.1 it is easily seen that (III.5) implies that $a := \Xi(\{0\}) = 0$. In particular, (III.5) excludes the Kingman coalescent. We will see in Section III.2.3 that the assumptions of Theorem III.2.7 exclude all coalescents that come down from infinity and only covers coalescents that stay infinite. If Assumption A of [27] holds with $\kappa := b$, then (III.5) holds, showing that all convergence results of [27] are covered by the following convergence theorems.

The following Proposition III.2.2 provides several conditions, each being equivalent to the key assumption (III.5). The proof shows that Proposition III.2.2 holds for any function $\gamma \in C_2((0,\infty))$ such that γ'' is non-negative and ultimately non-increasing.

Proposition III.2.2. *The following five conditions are equivalent.*

(i) Assumption (III.5) *holds, i.e., the limit* $\kappa := \lim_{x\to\infty} x\gamma''(x)$ *exists and is finite.*

(ii) $\lim_{x\to\infty}(\gamma'(x) - \gamma(x)/x) = \kappa$.

(iii) There exists a function $L : (0,\infty) \to (0,\infty)$ *being slowly varying at* ∞ *such that*

$$\frac{\gamma(x)}{x} \;=\; \kappa \log x \;+\; \log L(x), \qquad x > 0. \tag{III.6}$$

(iv) For all $y > 0$ *the limit* $d(y) := \lim_{x\to\infty}(\gamma(yx)/(yx) - \gamma(x)/x)$ *exists and* $d(y) = \kappa \log y$.

(v) $\lim_{x\to\infty}(\gamma'(yx) - \gamma'(x)) = \kappa \log y$ *for all* $y > 0$.

Remarks.

1. Assume that the coalescent has dust. Equivalently, $\lim_{x\to\infty} \gamma(x)/x = \int_\Delta |u|\,\nu(\mathrm{d}u) =: \mu < \infty$. Thus, (III.6) and, hence, all conditions of Proposition III.2.2 hold with $\kappa = 0$ and a slowly varying function L satisfying $\lim_{x\to\infty} L(x) = e^{\mu} < \infty$. Note however that there exist dust-free coalescents (even Λ-coalescents) which satisfy $\kappa := \lim_{x\to\infty} x\gamma''(x) = 0$. We refer the reader to Examples III.3.2 and III.3.3 in Section III.3.

2. The characterization theorem for regularly varying functions [4, Theorem 1.4.1] implies that the limit $d(y)$ in Proposition III.2.2

(iv) is necessarily of the form $d(y) = \kappa \log y$, $y > 0$, for some $\kappa \in \mathbb{R}$, if it exists, and due to $\lim_{x\to\infty} \gamma(x)/x = \int_\Delta |u|\,\nu(\mathrm{d}u) \in [0,\infty]$, only $\kappa \geq 0$ can occur.

3. In the terminology of [4, Section 3], the function γ' is a de Haan function with 1-index κ.

Proposition III.2.2 provides conditions being equivalent to the key assumption (III.5). However, all these conditions involve the rate function γ. Proposition III.2.3 below provides two additional equivalent conditions of assumption (III.5), which do not involve the rate function γ anymore and are instead more directly stated in terms of the measure Ξ of the coalescent and hence more intuitive to understand. Proposition III.2.3 essentially shows how (III.5) is related to the behavior of the measure Ξ near the point $0 \in \Delta$. In order to state the result, let us introduce the functions $F, F_1, F_2, \ldots : [0,1) \to [0,\infty)$ and $G, G_1, G_2, \ldots : [0,1] \to [0,\infty)$ via

$$\begin{aligned}
F_i(t) &:= \int_\Delta 1_{[0,t]}(u_i)(\log(1-u_i))^2\,\nu(\mathrm{d}u), \qquad i \in \mathbb{N}, t \in [0,1),\\
F(t) &:= \sum_{i\geq 1} F_i(t) = \int_\Delta \sum_{i\geq 1} 1_{[0,t]}(u_i)(\log(1-u_i))^2\,\nu(\mathrm{d}u),\ t \in [0,1),\\
G_i(t) &:= \int_\Delta 1_{[0,t]}(u_i)u_i^2\,\nu(\mathrm{d}u), \qquad i \in \mathbb{N}, t \in [0,1],\\
G(t) &:= \sum_{i\geq 1} G_i(t) = \int_\Delta \sum_{i\geq 1} 1_{[0,t]}(u_i)u_i^2\,\nu(\mathrm{d}u), \qquad t \in [0,1].
\end{aligned} \tag{III.7}$$

Note that $F_i(0) = G_i(0) = 0$ for all $i \in \mathbb{N}$ and, hence, $F(0) = G(0) = 0$. For every $t \in (0,1)$ there exists a constant $C_t \in (0,\infty)$ (choose, for example, $C_t := (-\log(1-t))/t$) such that $-\log(1-x) \leq C_t x$ for all $x \in [0,t]$. Applying this inequality with $x := u_i \leq t$ yields

$$\begin{aligned}
F_i(t) \leq F(t) &= \int_\Delta \sum_{i\geq 1} 1_{[0,t]}(u_i)(-\log(1-u_i))^2\,\nu(\mathrm{d}u)\\
&\leq C_t^2 \int_\Delta \sum_{i\geq 1} 1_{[0,t]}(u_i)u_i^2\,\nu(\mathrm{d}u)\\
&\leq C_t^2 \int_\Delta \sum_{i\geq 1} u_i^2\,\nu(\mathrm{d}u) = C_t^2\Xi(\Delta) < \infty.
\end{aligned}$$

Obviously, $G_i(t) \leq G(t) \leq \int_\Delta \sum_{i\geq 1} u_i^2\,\nu(\mathrm{d}u) = \Xi(\Delta) < \infty$. From $u_i \leq -\log(1-u_i)$ we conclude that $G_i(t) \leq F_i(t)$ for all $i \in \mathbb{N}$ and $t \in [0,1)$ and, hence, $G(t) \leq F(t)$ for all $t \in [0,1)$. Moreover, the functions $F, G, F_1, G_1, F_2, G_2, \ldots$ are non-decreasing, hence Riemann integrable.

Proposition III.2.3. *Let Ξ be a finite measure on Δ and let κ be some constant in $[0,\infty)$. Then the following three conditions are equivalent.*

(i) Assumption (III.5) *holds, i.e., the limit $\kappa = \lim_{x\to\infty} x\gamma''(x)$ exists and is finite.*

(ii) $\lim_{t\to 0+} t^{-1}F(t) = \kappa$. *(iii)* $\lim_{t\to 0+} t^{-1}G(t) = \kappa$.

In particular, for Λ-coalescents, (III.5) *is equivalent to*

$$\lim_{t\to 0+} \frac{\Lambda([0,t])}{t} \;=\; \kappa. \tag{III.8}$$

Relation (III.8) already appears in Lemma 9.1 of [27], but its importance was not (fully) discovered there.

III.2.2 The scaling function

Define $v : [1,\infty) \times [0,\infty) \to [1,\infty)$ (implicitly) via

$$v(1,t) \;:=\; 1 \quad \text{and} \quad \int_{v(x,t)}^{x} \frac{\mathrm{d}u}{\gamma(u)} \;=\; t, \qquad x > 1, t \geq 0. \tag{III.9}$$

The following two propositions clarify the existence of v and provide basic properties of v with an emphasis on coalescents with dust, coalescents that come down from infinity and coalescents that satisfy the key assumption (III.5).

Proposition III.2.4. *For each $x > 1$ and $t \geq 0$, the solution $v(x,t) \in (1,x]$ to the integral equation in* (III.9) *exists and is unique. Moreover, $v \in C_1((1,\infty)\times[0,\infty))$ with*

$$\frac{\mathrm{d}}{\mathrm{d}t}v(x,t) \;=\; -\gamma(v(x,t)), \quad \frac{\mathrm{d}}{\mathrm{d}x}v(x,t) \;=\; \frac{\gamma(v(x,t))}{\gamma(x)}, \tag{III.10}$$

$x > 1, t \geq 0$. For every $x \geq 1$ the map $t \mapsto v(x,t)$, $t \geq 0$, is non-increasing and for every $t \geq 0$ the map $x \mapsto v(x,t)$, $x \geq 1$, is non-decreasing.

Remark. If the coalescent is in a state with $k \in \mathbb{N}$ blocks, then $\gamma(k)$ is the expected rate of decrease of the block counting process. The choice of the scaling $v(x,t)$ then becomes plausible as, for each $x \geq 1$, it is the solution to the initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}v(x,t) \;=\; -\gamma(v(x,t)), \qquad t \geq 0, \qquad v(x,0) \;=\; x.$$

Proposition III.2.5. *Let γ be defined by* (III.3) *and let v be defined by* (III.9).

(i) If the coalescent has dust, i.e., $a := \Xi(\{0\}) = 0$ and $\mu := \int_\Delta |u|\nu(\mathrm{d}u) < \infty$, then $v(x,t) \sim xe^{-\mu t}$ as $x \to \infty$ for every $t \geq 0$.

(ii) Suppose that $\int_c^\infty (\gamma(u))^{-1}\,\mathrm{d}u < \infty$ for some (and hence all) $c > 1$. Then, for every $t > 0$, the solution $v(t) \in (1,\infty)$ to the equation

$$\int_{v(t)}^\infty \frac{\mathrm{d}u}{\gamma(u)} \;=\; t \tag{III.11}$$

exists and $\lim_{x\to\infty} v(x,t) = v(t)$.

(iii) Suppose that (III.5) *holds. Then, for every $t \geq 0$, there exists a slowly varying function $L_t : [1,\infty) \to (0,\infty)$ such that $v(x,t) = x^{e^{-\kappa t}} L_t(x)$ for all $x \geq 1$.*

Remarks.

1. If the coalescent has dust, then, as $n \to \infty$, $N^{(n)}/n$ converges in $D_{[0,1]}[0,\infty)$ to the so-called frequency of singletons process [14], so $(v(n,t))_{n\in\mathbb{N}}$ as in (i) is a reasonable scaling sequence for the block counting process.

2. For regular Ξ-coalescents that come down from infinity the integral $\int_c^\infty (\gamma(u))^{-1}\,\mathrm{d}u$ is finite for all $c > 1$. The function $v(t)$ defined by (III.11) is the "speed of coming down from infinity" as defined in [21], although with a slightly different function γ. See also [2] and [22].

3. The finiteness of the integral $\int_c^\infty (\gamma(u))^{-1}\,\mathrm{d}u$ can be viewed as a Grey's condition for the Ξ-coalescent. Grey's condition originally stems (see [17, 33, 34]) from the study of continuous-state branching processes with the function γ replaced by the branching mechanism of the considered branching process. For Λ-coalescents the rate function γ itself is the branching mechanism of a continuous-state branching process.

As seen in Lemma III.2.1, the asymptotic growth as $x \to \infty$ of $\gamma(x)$ is at least of order x. Part (i) of the following proposition shows that altering γ additively by a function of asymptotic order smaller than x asymptotically does essentially not change the scaling function $v(.,t)$. In Part (ii), the slowly varying function L_t of the scaling function is asymptotically calculated for a special case.

Proposition III.2.6. *Let γ and $v(x,t)$ be defined by* (III.3) *and* (III.9)*, respectively, and suppose that* (III.5) *holds.*

(i) Assume that there exists a continuous function $\gamma_1 : (1,\infty) \to (0,\infty)$ such that $(\gamma(x) - \gamma_1(x))/x \to 0$ as $x \to \infty$. Then, for each $t \geq 0$, there exists $x_0(t) > 1$ such that the scaling $v_1(x,t)$, defined by the integral equation in (III.9) *with γ_1 in place of γ, exists for all $x \geq x_0(t)$. Moreover, $v(x,t) \sim v_1(x,t)$ as $x \to \infty$.*

(ii) According to Proposition III.2.2, γ satisfies (III.6) *with $\kappa \in [0,\infty)$ and a slowly varying function $L : (0,\infty) \to (0,\infty)$. If $L(x) \to C$ as $x \to \infty$ for some constant $C > 0$, then, for each $t \geq 0$, $v(x,t) \sim x^{e^{-\kappa t}} C^{-\kappa^{-1}(1-e^{-\kappa t})}$ as $x \to \infty$ if $\kappa > 0$ and $v(x,t) \sim xC^{-t}$ as $x \to \infty$ if $\kappa = 0$.*

Remark. Assume that the coalescent has dust or, equivalently, that $\mu := \lim_{x\to\infty} \gamma(x)/x < \infty$. Then γ satisfies (III.6) with $\kappa = 0$. Thus, $L(x) = e^{\gamma(x)/x}$ and, hence, $\lim_{x\to\infty} L(x) = e^{\mu} \in [1,\infty)$. By Proposition III.2.6 (ii), $v(x,t) \sim xe^{-\mu t}$ as $x \to \infty$ for all $t \geq 0$, which also proves Proposition III.2.5 (i). For dust-free coalescents, $\lim_{x\to\infty} v(x,t)/x = 0$ for all $t > 0$.

III.2.3 Results concerning the block counting process

Let $n \in \mathbb{N}$. The block counting process $(N_t^{(n)})_{t\geq 0}$ with initial state $N_0^{(n)} = n$ jumps from state $k \in \{2,\ldots,n\}$ to state $j \in \{1,\ldots,k-1\}$ at the rate (see [13, Eq. (1.3)] or [14, Proposition 2.1])

$$q_{k,j} \;=\; a\binom{k}{2} 1_{\{j=k-1\}} \;+\; \int_\Delta \sum_{i=1}^{j} f_{kji}(u)\,\nu(\mathrm{d}u),$$

where

$$f_{kji}(u) \;:=\; \sum_{\substack{k_1,\ldots,k_i\in\mathbb{N}\\ k_1+\cdots+k_i=k-j+i}} \frac{k!}{(j-i)!k_1!\cdots k_i!}(1-|u|)^{j-i} \sum_{\substack{l_1,\ldots,l_i\in\mathbb{N}\\ l_1<\cdots<l_i}} u_{l_1}^{k_1}\cdots u_{l_i}^{k_i}$$

for $i \in \{1,\ldots,j\}$ and $u \in \Delta$.

From the Poisson point process construction of the coalescent it follows that the jump rates of the block counting process can be described in terms of an urn model as in [23] as follows. Fix $u \in \Delta$ and partition the interval $[0,1)$ into "urns" $J_0, J_1, \ldots$ of lengths $u_0 := 1-|u|, u_1, u_2, \ldots$, i.e., $J_0 := [0,u_0), J_1 := [u_0, u_0+u_1), J_2 :=$

$[u_0+u_1, u_0+u_1+u_2)$ and so on. The "balls" $Z_1, Z_2, \ldots$ are i.i.d. random variables, where Z_1 has an uniform distribution on $[0,1)$. Let $X_i(k,u) := \sum_{j=1}^k 1_{\{Z_j=i\}}$ denote the number of balls in urn $i \in \mathbb{N}_0 := \{0,1,\ldots\}$ after $k \in \mathbb{N}_0$ throws. Note that

$$Y(k,u) \;:=\; X_0(k,u) \;+\; \sum_{i\geq 1} 1_{\{X_i(k,u)>0\}}, \tag{III.12}$$

$k \in \mathbb{N}, u \in \Delta$, is the sum of the number of balls in urn J_0 and the number of all other occupied urns. Then

$$q_{k,j} \;=\; a\binom{k}{2}1_{\{j=k-1\}} \;+\; \int_\Delta \mathbb{P}(Y(k,u)=j)\,\nu(\mathrm{d}u), \tag{III.13}$$

$j,k \in \mathbb{N}, j<k$. This representation of the jump rates will turn out to be crucial to the proof of the main convergence result (Theorem III.2.7). Relation (III.13) also provides further insight into the function γ. For example, by (III.13), for all $k \in \mathbb{N}$,

$$\sum_{j=1}^{k-1}(k-j)q_{k,j} = a\binom{k}{2} + \int_\Delta \mathbb{E}(k-Y(k,u))\,\nu(\mathrm{d}u) \tag{III.14}$$

$$= a\binom{k}{2} + \int_\Delta \Big(k - k(1-|u|) - \sum_{i\geq 1}(1-(1-u_i)^k)\Big)\nu(\mathrm{d}u) = \gamma(k).$$

Thus, if the block counting process is in state $k \in \mathbb{N}$, then $\gamma(k) = \sum_{j=1}^{k-1}(k-j)q_{k,j}$ is the expected rate of decrease of the block counting process. Lemma III.5.2 provided in the appendix shows that $Y(k,u)/k \to 1-|u| =: u_0$ almost surely as $k\to\infty$ for every $u \in \Delta$.

Define $\Delta^* := \{u \in \Delta : |u|=1\}$. Assume that $\nu(\Delta^*)=0$ and that the regularity condition (III.1) holds. Define the function $\psi : \mathbb{R} \to \mathbb{C}$ via

$$\psi(x) \;:=\; \int_\Delta \big((1-|u|)^{ix} - 1 + ix|u|\big)\nu(\mathrm{d}u), \qquad x \in \mathbb{R}. \tag{III.15}$$

Note that (III.1) ensures that $\psi(x) \in \mathbb{C}$. Define the transformation $g : \Delta\setminus\Delta^* \to (-\infty,0)$ via $g(u) := \log(1-|u|)$ for all $u \in \Delta\setminus\Delta^*$ and let $\varrho := \nu_g$ denote the image measure of ν under g. Then,

$$\begin{aligned}\psi(x) \;&=\; \int_{(-\infty,0)} \big(e^{ixt} - 1 + ix(1-e^t)\big)\varrho(\mathrm{d}t)\\ &=\; ix\int_{(-\infty,0)}\Big(1-e^t+\frac{t}{1+t^2}\Big)\varrho(\mathrm{d}t)\\ &\quad+ \int_{(-\infty,0)}\Big(e^{ixt}-1-ix\frac{t}{1+t^2}\Big)\varrho(\mathrm{d}t), \quad x \in \mathbb{R}.\end{aligned}$$

Thus, ψ is the characteristic exponent of an infinitely divisible distribution. The regularity condition (III.1) is required for $\varrho := \nu_g$ to be a Lévy measure. In the following, for $t \geq 0$, S_t denotes a random variable with characteristic function ϕ_t, given by

$$\phi_t(x) \;:=\; \exp\biggl(\int_0^t \psi(e^{-\kappa s}x)\,\mathrm{d}s\biggr), \qquad x \in \mathbb{R}, t \geq 0, \qquad \text{(III.16)}$$

where ψ is defined by (III.15) and $\kappa \in [0, \infty)$. Note that ψ can be recovered from the values $\phi_t(x)$, $x \in \mathbb{R}$, $t \geq 0$, namely, for any $x \in \mathbb{R}$, $\psi(x) = t^{-1} \log \phi_t(x)$ if $\kappa = 0$ and $\psi(x) = -\lim_{t\to\infty} \frac{\mathrm{d}}{\mathrm{d}t} \log \phi_t(x)$ if $0 < \kappa < \infty$.

The limiting process X arising in the main convergence result (Theorem III.2.7 below), whose distribution is determined through its semigroup $(T_t^X)_{t\geq 0}$ via (III.16) and (III.17), belongs to the class of Ornstein–Uhlenbeck type processes [30]. The semigroup $(T_t^X)_{t\geq 0}$ belongs to the class of generalized Mehler semigroups (see [5]), since $\phi_{t+s}(x) = \phi_t(e^{-\kappa s}x)\phi_s(x)$ for $x \in \mathbb{R}$ and $s, t \geq 0$. Clearly, $(T_t^X)_{t\geq 0}$ is a Feller semigroup on $\widehat{C}(\mathbb{R})$, the space of continuous functions $f : \mathbb{R} \to \mathbb{R}$ vanishing at infinity.

Theorem III.2.7. *Suppose that* Ξ *satisfies* (III.1) *and* $\Xi(\Delta^*) = 0$. *Let* γ *be defined by* (III.3) *and suppose that* (III.5) *holds, i.e., the limit* $\kappa := \lim_{x\to\infty} x\gamma''(x) \in [0, \infty)$ *exists. Moreover, let the scaling* $v(n, t)$ *be defined by* (III.9). *Then the logarithmically scaled block counting process* $(\log N_t^{(n)} - \log v(n, t))_{t\geq 0}$ *converges in* $D_{\mathbb{R}}[0, \infty)$ *to* X *as* $n \to \infty$, *where* $X = (X_t)_{t\geq 0}$ *is an Ornstein–Uhlenbeck type process with state space* $\mathbb{R}$, *initial value* $X_0 = 0$ *and Mehler semigroup* $(T_t^X)_{t\geq 0}$ *given by*

$$T_t^X f(x) \;:=\; \mathbb{E}(f(X_{s+t})|X_s = x) \;=\; \mathbb{E}(f(e^{-\kappa t}x + S_t)), \qquad \text{(III.17)}$$

$k \in \mathbb{N}, u \in \Delta$, *with the distribution of* S_t *defined via its characteristic function* (III.16).

Remark. For results on the generator A^X of the limiting process X arising in Theorem III.2.7 we refer the reader to (III.33).

Remark. The limiting process X in Theorem III.2.7 determines the distribution of S_t ($\stackrel{d}{=} X_t$) and, hence, the values (III.16) of its characteristic function ϕ_t, which in turn determines the function ψ defined via (III.15). In general it is however impossible to recover from ψ the original measure Ξ of the underlying coalescent.

Remark. (stationary distribution) Under the assumptions of Theorem III.2.7 an application of [4, Theorems 4.1 and 4.2] yields the following. If

$$\int_{\{u\in\Delta:|u|>\varepsilon\}} \log\log(1-|u|)^{-1}\,\nu(\mathrm{d}u) < \infty, \tag{III.18}$$

for some $\varepsilon \in (1-e^{-1},1)$, then X_t converges in distribution as $t\to\infty$ to the unique stationary distribution μ of X, where μ is self-decomposable with characteristic function ϕ given by $\phi(x) := \exp(\int_0^\infty \psi(e^{-\kappa s}x)\,\mathrm{d}s)$, $x\in\mathbb{R}$. If (III.18) does not hold, then the process X has no stationary distribution. For Λ-coalescents, this remark remains valid if in Eq. (III.18) the measure ν is replaced by Λ.

Remark. If (III.5) and, hence, (III.6) holds, then $\int_c^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$ for some (and hence all) $c\in(1,\infty)$. For $\varepsilon\in(0,1)$ define $\Delta^\epsilon := \{u\in\Delta : |u|\le 1-\varepsilon\}$. Further, define $\Delta_f := \{u\in\Delta : u_1+\ldots+u_n = 1 \text{ for some } n\in\mathbb{N}\}$. Under the regularity condition (III.1) it holds that $\nu(\Delta\backslash\Delta^\varepsilon)<\infty$ for all $\varepsilon\in(0,1)$ (see [21, p. 229]). Due to $\Xi(\Delta_f)=0$, the coalescents covered by Theorem III.2.7 hence stay infinite [31, Proposition 33]. In other words the coalescents covered by Theorem III.2.7 either have dust or they have no dust and are not coming down from infinity. A schematic representation of the space $\mathcal{M}(\Delta)$ of all finite measures Ξ on $(\Delta,\mathcal{B}(\Delta))$ is provided in Figure III.1. In this representation, $\mathcal{M}(\Delta)$ is equipped with the topology of weak convergence, i.e., $\Xi_n\to\Xi$ as $n\to\infty$ if and only if $\lim_{n\to\infty}\int_\Delta f\,\mathrm{d}\Xi_n = \int_\Delta f\,\mathrm{d}\Xi$ for all continuous functions $f:\Delta\to\mathbb{R}$. Note that, since Δ is compact, all continuous functions $f:\Delta\to\mathbb{R}$ are bounded and uniformly continuous. The space $\mathcal{M}(\Delta)$ is metrizable, for example via the metric

$$d(\Xi_1,\Xi_2) \ := \ \sum_{i\ge1} 2^{-i}\frac{1}{1+\|f_i\|}\left|\int f_i\,\mathrm{d}\Xi_1 - \int f_i\,\mathrm{d}\Xi_2\right|,$$

where $\{f_1,f_2,\ldots\}$ is a dense set of real-valued continuous functions on Δ. The results in Parthasarathy [28, Chapter 6] imply that, with this metric, $\mathcal{M}(\Delta)$ is a compact Polish (separable complete metric) space.

III.2.4 Results concerning the fixation line

The fixation line has been introduced by Hénard [18] for Λ-coalescents and further studied in [14] for general Ξ-coalescents. The fixation line $(L_t^{(n)})_{t\ge0}$ with initial state $L_0^{(n)}=n$ is a Markov process

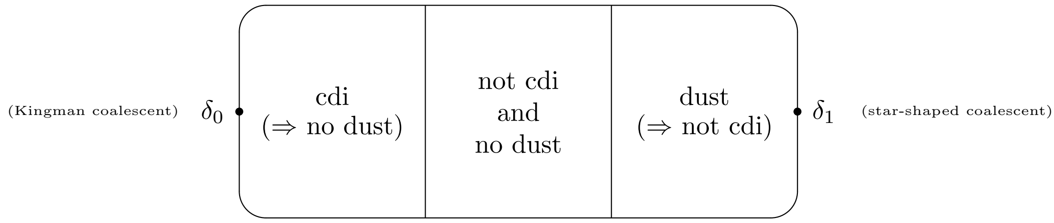

Figure III.1: A schematic representation of the space of all exchangeable coalescents (Ξ-coalescents). Each point in the oval region corresponds to a finite measure Ξ on $(\Delta, \mathcal{B}(\Delta))$. The compact Polish space $\mathcal{M}(\Delta)$ is divided into three regions of exchangeable coalescents, those coming down from infinity (cdi) to the left, the ones not coming down from infinity and having no dust in the middle and those having dust to the right.

which moves from state $i \in \{n, n+1, \ldots\}$ to state $j \in \mathbb{N}$ with $j > i$ at the rate (see [14, Proposition 2.5])

$$\gamma_{i,j} \;=\; a\binom{j}{2}\delta_{j,i+1} \;+\; \int_\Delta \mathbb{P}\big(Y(j,u) = i, Y(j+1,u) = i+1\big)\nu(\mathrm{d}u),$$

where $a := \Xi(\{0\})$ and $Y(.,u)$ is defined via (III.12). The fixation line does not explode if and only if the coalescents stays infinite [14, Remark 2.11]. Recall that (see (III.2)) the block counting process is Siegmund dual to the fixation line. In this subsection we will see that this duality property transfers the convergence result for the block counting process (Theorem III.2.7) into an analogous convergence result (Theorem III.2.10) for the fixation line. The arguments are similar as for the fixation line. We start as follows. Assume that $\int_2^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$. Define the function $w : [1,\infty) \times [0,\infty) \to [1,\infty)$ via

$$w(1,t) \;:=\; 1 \quad \text{and} \quad \int_x^{w(x,t)} \frac{\mathrm{d}u}{\gamma(u)} \;=\; t, \qquad x > 1, t \geq 0. \tag{III.19}$$

Proposition III.2.8. *Assume that* $\int_2^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$*. Then, for each* $x > 1$ *and* $t \geq 0$*, the solution* $w(x,t) \in [x,\infty)$ *to the integral equation in* (III.19) *exists and is unique. Furthermore,* $w \in C_1((1,\infty) \times [0,\infty))$ *with*

$$\frac{\mathrm{d}}{\mathrm{d}t} w(x,t) \;=\; \gamma(w(x,t)), \quad \frac{\mathrm{d}}{\mathrm{d}x} w(x,t) \;=\; \frac{\gamma(w(x,t))}{\gamma(x)}, \tag{III.20}$$

$x > 1, t \geq 0$*. The maps* $x \mapsto w(x,t)$*,* $x \geq 1$*, and* $t \mapsto w(x,t)$*,* $t \geq 0$*, are strictly increasing.*

It is readily seen from (III.9) and (III.19) that $v(w(x,t),t) = x = w(v(x,t),t)$ for all $x \geq 1$ and $t \geq 0$. Thus, for fixed $t \geq 0$, $w(.,t)$ is the inverse of $v(.,t)$. This aspect is utilized in the proof of Proposition

III.2.9 below, whose statements are variants of Propositions III.2.5 and III.2.6 for the scaling of the fixation line. Note that, under the key assumption (III.5), in particular when the coalescent has dust, it holds that $\int_c^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$ for every $c > 1$, so $w(x,t)$ is well-defined.

Proposition III.2.9. *Let γ be defined by* (III.3) *and w be defined by* (III.19)*.*

(i) If $a := \Xi(\{0\}) = 0$ and $\mu := \int_\Delta |u|\,\nu(\mathrm{d}u) < \infty$, then $w(x,t) \sim xe^{\mu t}$ as $x \to \infty$ for each $t \geq 0$.

(ii) Suppose that γ satisfies (III.5) *with $\kappa \geq 0$. Then, for every $t \geq 0$, there exists a slowly varying function $L_t^{\#} : [1,\infty) \to (0,\infty)$ such that $w(x,t) = x^{e^{\kappa t}} L_t^{\#}(x)$ for all $x \geq 1$.*

(iii) Suppose that γ satisfies (III.5) *with $\kappa \geq 0$. Assume that there exists a continuous function $\gamma_1 : (1,\infty) \to (0,\infty)$ such that $(\gamma(x) - \gamma_1(x))/x \to 0$ as $x \to \infty$. Then the scaling $w_1(x,t)$, defined by* (III.19) *with γ_1 in place of γ, exists for all $t \geq 0$ and $x \geq 1$. Moreover, assume that the map $x \mapsto \gamma_1(x)/x$, $x > 1$, is non-decreasing if $\kappa = 0$. Then $w(x,t) \sim w_1(x,t)$ as $x \to \infty$.*

Remark. The regular variation of $w(.,t)$ under the key assumption (III.5) is a consequence of the regular variation of $v(.,t)$ (see Proposition III.2.5) and the fact that $w(.,t)$ and $v(.,t)$ are inverse. The slowly varying part $L_t^{\#}$ can be retrieved from L_t with the use of the de Bruijn conjugate, see [4, Theorem 1.5.13 and Proposition 1.5.15] and the proof of Proposition III.2.9 in Section II.8 for further details.

The following theorem is the analog of Theorem III.2.7 for the fixation line.

Theorem III.2.10. *Suppose that Ξ satisfies* (III.1) *and $\Xi(\Delta^*) = 0$. Let γ be defined by* (III.3) *and suppose that* (III.5) *holds with $0 \leq \kappa < \infty$. Let the scaling $w(n,t)$ be defined by* (III.19)*. Then $(\log L_t^{(n)} - \log w(n,t))_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ to Y as $n \to \infty$, where $Y = (Y_t)_{t\geq 0}$ is an Ornstein–Uhlenbeck type process with state space $\mathbb{R}$, initial value $Y_0 = 0$ and Mehler semigroup $(T_t^Y)_{t\geq 0}$ given by*

$$T_t^Y f(y) := \mathbb{E}(f(Y_{s+t})|Y_s = y) = \mathbb{E}(f(e^{\kappa t}y - e^{\kappa t}S_t)), \qquad \text{(III.21)}$$

for $y \in \mathbb{R}, f \in B(\mathbb{R})$ and $s,t \geq 0$ with the distribution of S_t defined via its characteristic function (III.16)*.*

III.2.5 Siegmund duality and summary of results

Let $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ be the limiting processes arising in Theorems III.2.7 and III.2.10, respectively. For $t \geq 0$ define $\widetilde{X}_t := e^{X_t}$ and $\widetilde{Y}_t := e^{Y_t}$. Consider the "exponential" Markov processes $\widetilde{X} := (\widetilde{X}_t)_{t\geq 0}$ and $\widetilde{Y} := (\widetilde{Y}_t)_{t\geq 0}$ both having state space $E := (0,\infty)$. From (III.17) and (III.21) it follows that the semigroups $(T_t^{\widetilde{X}})_{t\geq 0}$ and $(T_t^{\widetilde{Y}})_{t\geq 0}$ of $\widetilde{X}$ and $\widetilde{Y}$ are given by

$$T_t^{\widetilde{X}} f(x) \;=\; \mathbb{E}(f(x^{e^{-\kappa t}} e^{S_t})), \qquad t \geq 0, f \in B(E), x \in E, \quad \text{(III.22)}$$

and

$$T_t^{\widetilde{Y}} g(y) \;=\; \mathbb{E}(g(y^{e^{\kappa t}} e^{-e^{\kappa t} S_t})), \qquad t \geq 0, g \in B(E), y \in E, \quad \text{(III.23)}$$

where S_t has characteristic function (III.16).

Fix $t \geq 0$, define $\alpha := e^{-\kappa t}$ and let $H : E \times E \to \{0,1\}$ denote the Siegmund duality kernel, i.e., $H(x,y) := 1$ for $x \leq y$ and $H(x,y) := 0$ otherwise. For $x, y \in E$, by (III.22), $T_t^{\widetilde{X}} H(.,y)(x) = \mathbb{E}(H(x^\alpha e^{S_t}, y)) = \mathbb{P}(x^\alpha e^{S_t} \leq y)$. Similarly, by (III.23), $T_t^{\widetilde{Y}} H(x,.)(y) = \mathbb{P}(H(x, y^{1/\alpha} e^{-S_t/\alpha})) = \mathbb{P}(x \leq y^{1/\alpha} e^{-S_t/\alpha}) = \mathbb{P}(x^\alpha e^{S_t} \leq y)$. Thus, $T_t^{\widetilde{X}} H(.,y)(x) = T_t^{\widetilde{Y}} H(x,.)(y)$ for all $t \geq 0$ and $x, y \in E$, showing that $\widetilde{X}$ is Siegmund dual to $\widetilde{Y}$.

Since the map $D_{\mathbb{R}}[0,\infty) \ni x = (x_t)_{t\geq 0} \mapsto (e^{x_t})_{t\geq 0} \in D_E[0,\infty)$ is continuous, an application of the continuous mapping theorem shows that Theorems III.2.7 and III.2.10 can be summarized as follows.

Theorem III.2.11. *Under the conditions of Theorem III.2.7 the following two assertions hold.*

(i) As $n \to \infty$, the scaled block counting process $(N_t^{(n)}/v(n,t))_{t\geq 0}$ converges in $D_E[0,\infty)$ to the Markov process $\widetilde{X}$ with $\widetilde{X}_0 = 1$ and semigroup (III.22)*.*

(ii) As $n \to \infty$, the scaled fixation line $(L_t^{(n)}/w(n,t))_{t\geq 0}$ converges in $D_E[0,\infty)$ to the Markov process $\widetilde{Y}$ with $\widetilde{Y}_0 = 1$ and semigroup (III.23)*.*

Thus, under the assumptions of Theorem III.2.7, the commutative diagram in Figure 2 holds.

We close the result section by providing formulas for the infinitesimal generators of the processes $\widetilde{X}$ and $\widetilde{Y}$. Applying the generator formulas $A^{\widetilde{X}} f(x) = A^X (f \circ \exp)(\log x)$ and $A^{\widetilde{Y}} g(y) = A^Y (g \circ \exp)(\log y)$

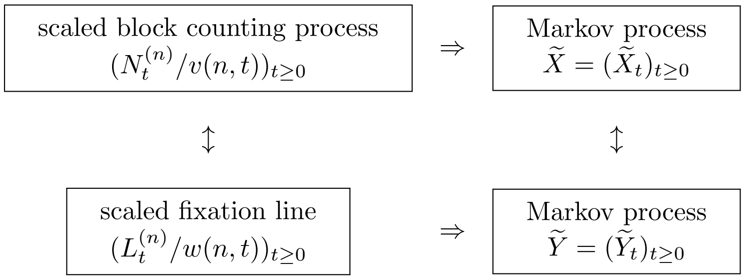

Figure III.2: Commutative diagram summarizing the convergence and duality results. In the diagram, "$\Rightarrow$" stands for convergence in $D_E[0,\infty)$ and "$\updownarrow$" for Siegmund duality $\mathbb{P}(N_t^{(n)} \leq m) = \mathbb{P}(L_t^{(m)} \geq n)$, $n, m \in \mathbb{N}$, and $\mathbb{P}(\widetilde{X}_t^{(x)} \leq y) = \mathbb{P}(\widetilde{Y}_t^{(y)} \geq x)$, $x, y \in E$, $t \geq 0$, respectively, where the upper indices indicate the initial states of the corresponding processes.

yield that the generators $A^{\widetilde{X}}$ and $A^{\widetilde{Y}}$ of $\widetilde{X}$ and $\widetilde{Y}$ satisfy

$$\begin{aligned} A^{\widetilde{X}} f(x) &= -\kappa x (\log x) f'(x) \\ &\quad + \int_\Delta \big(f(x(1-|u|)) - f(x) + |u| x f'(x)\big)\nu(\mathrm{d}u) \end{aligned} \tag{III.24}$$

for $x > 0$ and $f \in \widetilde{D}$ and

$$\begin{aligned} A^{\widetilde{Y}} g(y) &= \kappa y (\log y) g'(y) \\ &\quad + \int_\Delta \big(g(y/(1-|u|)) - g(y) - |u| y g'(y)\big)\nu(\mathrm{d}u) \end{aligned} \tag{III.25}$$

for $y > 0$ and $g \in \widetilde{D}$, where $\widetilde{D}$ denotes the space of all functions $f : E \to \mathbb{R}$ such that the maps f, $x \mapsto x f'(x)$, $x \mapsto x^2 f''(x)$ and $x \mapsto x(\log x) f'(x)$ belong to $\widehat{C}(E)$. Note that $\widetilde{D}$ is a core for both generators, $A^{\widetilde{X}}$ and $A^{\widetilde{Y}}$.

III.3 Examples

In this section several illustrating examples are provided. For most of the examples Theorem III.2.7 and Theorem III.2.10 are applicable. Example III.3.1 treats the Λ-coalescent, where $\Lambda = \beta(a, b)$ is a beta distribution with parameters $a, b > 0$. For the $\beta(1, b)$-coalescent scaling limits have already been obtained in [27], and we clarify beforehand the relation between [27] and this work. Example III.3.2 studies the Λ-coalescent introduced in [25], where the measure Λ is a negative logarithmic gamma distribution (NLG-coalescent). Example III.3.3 provides a simple dust-free Λ-coalescent which nevertheless satisfies $\kappa = 0$. Example III.3.4 presents a true Ξ-coalescent for which Theorem III.2.7 and Theorem III.2.10 are applicable.

We start with putting the results of [27] in the context of our work. Let $b > 0$. In [27] it is shown that $(\log N_t^{(n)} - e^{-bt} \log n)_{t\geq 0}$

converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to an Ornstein–Uhlenbeck type process provided that the coalescent's driving measure Λ satisfies

$$\Lambda(\{0\}) = \Lambda(\{1\}) = 0 \quad \text{and} \quad c := \int_{[0,1]} u^{-1}\,(\Lambda - b\lambda)(\mathrm{d}u) < \infty. \tag{III.26}$$

Here λ denotes Lebesgue measure. Condition (III.26) essentially forces the coalescent to behave similarly to the Bolthausen–Sznitman coalescent (BS-coalescent), which is the Λ-coalescent where Λ is the uniform distribution on $[0,1]$. For $x \geq 0$ define $\gamma_{\mathrm{BS}}(x) := \int_0^1 \big((1-u)^x - 1 + ux\big)u^{-2}\,\mathrm{d}u$. Let $\Psi := (\log \Gamma)' = \Gamma'/\Gamma$ denote the logarithmic derivative of the gamma function (digamma function). It is easily checked that $\gamma_{\mathrm{BS}}(x) = x(\Psi(x+1) - \Psi(1) - 1) = x\log x - (\Psi(1)+1)x + O(1)$ as $x \to \infty$. If (III.26) holds, then

$$\begin{aligned}\gamma(x) &= b\gamma_{\mathrm{BS}}(x) + \int_{[0,1]} \big((1-u)^x - 1 + xu\big)(\Lambda - b\lambda)(\mathrm{d}u)\\ &= bx\log x + x(-b(1+\Psi(1)) + c + o(1))\end{aligned}$$

for $x > 0$ such that (III.6) and, hence, (III.5) are satisfied with $\kappa := b$ and the slowly varying function L in (III.6) satisfies $L(x) \to \exp(-b(1+\Psi(1)) + c)$ as $x \to \infty$.

Example III.3.1. (beta coalescent) Let $\Lambda = \beta(a,b)$ be the beta distribution with parameters $a, b > 0$. For the corresponding Λ-coalescent, the function γ, defined via (III.3), can be calculated explicitly. For $a \notin \{1,2\}$, a technical but straightforward calculation shows that

$$\begin{aligned}\gamma(x) &= \frac{(x+a+b-1)(x+a+b-2)}{(a-1)(a-2)}\frac{\mathrm{B}(a,x+b)}{\mathrm{B}(a,b)}\\ &\quad + \frac{a+b-1}{a-1}x - \frac{(a+b-1)(a+b-2)}{(a-1)(a-2)}\end{aligned} \tag{III.27}$$

for all $x \geq 0$, where $\mathrm{B}(.,.)$ denotes the beta function. The boundary cases $a=1$ and $a=2$ need to be treated separately. For $a=1$ one obtains

$$\gamma(x) = b(x+b-1)\big(\Psi(x+b) - \Psi(b)\big) - bx, \qquad x \geq 0, \tag{III.28}$$

where Ψ denotes the digamma function. For $a=2$ it follows that

$$\gamma(x) = (b+1)x - b(b+1)\big(\Psi(x+b) - \Psi(b)\big), \qquad x \geq 0. \tag{III.29}$$

Since $\Psi(x+b) = \log x + O(x^{-1})$ as $x \to \infty$ it follows from (III.27), (III.28) and (III.29) that

$$\frac{\gamma(x)}{x}\begin{cases} = \dfrac{\Gamma(a+b)}{\Gamma(b)(1-a)(2-a)}x^{1-a}+O(1) & \text{for } a<1,\\ = b\log x - b(\Psi(b)+1)+O\Big(\dfrac{\log x}{x}\Big) & \text{for } a=1,\\ \to \dfrac{a+b-1}{a-1} & \text{for } a>1.\end{cases}$$

For all $y\in(0,\infty)$ the limit $d(y)$, defined in Proposition III.2.2 (iv), is thus given by

$$d(y) \;=\; \begin{cases} -\infty 1_{(0,1)}(y)+\infty 1_{(1,\infty)}(y) & \text{for } a<1 \text{ (cdi)},\\ b\log y & \text{for } a=1 \text{ (not cdi, no dust)},\\ 0 & \text{for } a>1 \text{ (dust)},\end{cases}$$

and the curvature parameter κ is given by

$$\kappa \;:=\; \lim_{x\to\infty} x\gamma''(x) \;=\; \begin{cases} \infty & \text{for } a<1,\\ b & \text{for } a=1,\\ 0 & \text{for } a>1.\end{cases}$$

For beta coalescents the curvature parameter κ thus characterizes both, the dust property and the cdi property. Theorems III.2.7 and III.2.10 are hence applicable for the $\beta(a,b)$-coalescent with $a\ge 1$. Let us distinguish two cases.

Case 1: If $a>1$ (dust case), then $\kappa=0$, $v(x,t)\sim e^{-\mu t}x$ and $w(x,t)\sim e^{\mu t}x$ as $x\to\infty$ with $\mu:=\int u^{-1}\Lambda(\mathrm{d}u)=(a+b-1)/(a-1)$. In this case, Theorems III.2.7 and III.2.10 are in essence logarithmic versions of [14, Theorem 2.1].

Case 2: Now assume that $a=1$, i.e., that $\Lambda=\beta(1,b)$ is the beta distribution with parameters 1 and $b>0$ having density $u\mapsto b(1-u)^{b-1}$, $u\in(0,1)$, with respect to Lebesgue measure on $(0,1)$. Then, $\kappa=b>0$.

From the discussion above (see also [27, Example 2 or Proposition 11]) it follows that

$$\gamma(x) \;=\; bx\log x \;+\; x(\log C_b+o(1)), \qquad x>0,$$

where $C_b:=\exp(-b(\Psi(b)+1))$. Independently one can verify that

$$x\gamma''(x) \;=\; bx\int_0^1 \frac{(1-u)^{x+b-1}(\log(1-u))^2}{u^2}\,\mathrm{d}u \;\to\; b, \qquad x\to\infty.$$

Let the scaling sequence $v(n,t)$ be defined by (III.9) for $n\ge 2$. By Proposition III.2.6, $v(x,t)\sim x^{e^{-bt}}C_b^{b^{-1}(e^{-bt}-1)}=x^{e^{-bt}}e^{(\Psi(b)+1)(1-e^{-bt})}$ as

$x \to \infty$. Similarly, $w(x,t) \sim x^{e^{bt}} e^{-(\Psi(b)+1)(1-e^{-bt})}$ as $x \to \infty$. By Theorems III.2.7 and III.2.10, both processes $(\log N_t^{(n)} - \log v(n,t))_{t\geq 0}$ and $(\log L_t^{(n)} - \log w(n,t))_{t\geq 0}$ converge in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$. As $n \to \infty$, the process $(\log N_t^{(n)} - e^{-bt}\log n)_{t\geq 0}$ converges as well due to the specifics of the scaling sequence, and the limiting processes generator A^X, which can be determined using Lemma III.4.1, is given by

$$\begin{aligned} A^X f(x) &= bf'(x)(1+\Psi(b)-x) \\ &\quad + \int_{[0,1]} \big(f(x+\log(1-u)) - f(x) + uf'(x)\big)u^{-2}\,\Lambda(\mathrm{d}u) \end{aligned}$$

for $x \in \mathbb{R}$ and f belonging to a core D, in agreement with the results of [27].

Further examples are now provided for which Theorems III.2.7 and III.2.10 are applicable.

Example III.3.2. (NLG-coalescent) Fix $\alpha, \varrho > 0$. Assume that Λ is the negative logarithmic gamma (NLG) distribution having density $u \mapsto \alpha^\varrho u^{\alpha-1}(-\log u)^{\varrho-1}/\Gamma(\varrho)$, $u \in (0,1)$, with respect to Lebesgue measure on $(0,1)$. The corresponding Λ-coalescent was introduced in [25, Example 3.2]. The asymptotics of $\gamma(x)$ as $x \to \infty$ is obtained as follows. For all $n \in \mathbb{N}$, $\gamma(n) = \sum_{j=0}^{n-1} a_j$, where

$$\begin{aligned} a_j &:= \int \frac{1-(1-u)^j}{u}\Lambda(\mathrm{d}u) \\ &\sim \begin{cases} \dfrac{\alpha^\varrho}{\Gamma(\varrho)}\dfrac{\Gamma(\alpha)}{1-\alpha} j^{1-\alpha}(\log j)^{\varrho-1} & \text{if } 0<\alpha<1, \\ \dfrac{(\log j)^\varrho}{\Gamma(\varrho+1)} & \text{if } \alpha = 1, \\ \left(\dfrac{\alpha}{\alpha-1}\right)^\varrho & \text{if } 1<\alpha<\infty, \end{cases} \end{aligned}$$

as $j \to \infty$ by [25, Lemma 7.3], applied with $a := \alpha$, $b := j$ and $c := \varrho$. Thus, as $n \to \infty$, the arithmetic mean $\gamma(n)/n = n^{-1}\sum_{j=0}^{n-1} a_j$ of the sequence $(a_j)_{j\in\mathbb{N}_0}$ satisfies

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

The coalescent has dust if and only if $\gamma(n)/n$ is bounded, so if and only if $\alpha > 1$. This coalescent comes down from infinity if and only if $\sum_{n=2}^{\infty} 1/\gamma(n) < \infty$, so if and only if $\alpha < 1$ or $\alpha = 1$ and $\varrho > 1$. It is easily seen that

$$\begin{aligned} \kappa \;&:=\; \lim_{\varepsilon\to 0+} \frac{\Lambda([0,\varepsilon])}{\varepsilon} \\ &=\; \begin{cases} \infty & \text{if } 0<\alpha<1 \text{ or if } \alpha=1 \text{ and } 1<\varrho<\infty, \\ 1 & \text{if } \alpha=\varrho=1 \text{ (Bolthausen–Sznitman coalescent)}, \\ 0 & \text{if } 1<\alpha<\infty \text{ or if } \alpha=1 \text{ and } 0<\varrho<1. \end{cases} \end{aligned}$$

In particular, $\kappa = \infty$ if and only if the coalescent comes down from infinity. Theorems III.2.7 and III.2.10 are applicable if and only if $\kappa < \infty$, so if and only if $\alpha > 1$ or $\alpha = 1$ and $0 < \varrho \le 1$. In the following three cases are distinguished.

Case 1: For $1 < \alpha < \infty$ the coalescent has dust. Hence, $\kappa = 0$, $v(x,t) \sim e^{-\mu t}x$ and $w(x,t) \sim e^{\mu t}x$ as $x \to \infty$ with $\mu := \int u^{-1}\,\Lambda(\mathrm{d}u) = \lim_{x\to\infty}\gamma(x)/x = (\alpha/(\alpha-1))^{\varrho}$. Theorems III.2.7 and III.2.10 are applicable and in essence logarithmic versions of [14, Theorem 2.1].

Case 2: For $\alpha = \varrho = 1$ we obtain the Bolthausen–Sznitman coalescent already studied in Example III.3.1.

Case 3: Assume that $\alpha = 1$ and $0 < \varrho < 1$. Then $\kappa = 0$ but nevertheless the coalescent is dust-free. Let $x > 1$. By the definition (III.3) of the rate function γ,

$$\begin{aligned} \gamma(x) \;&=\; \frac{1}{\Gamma(\varrho)} \int_0^1 \frac{(1-u)^x - 1 + xu}{u} \frac{(-\log u)^{\varrho-1}}{u}\,\mathrm{d}u \\ &=\; \frac{1}{\Gamma(\varrho+1)} \int_0^1 \frac{1-(1-u)^x - xu(1-u)^{x-1}}{u^2} (-\log u)^{\varrho}\,\mathrm{d}u, \end{aligned}$$

where the last equality holds by partial integration. The substitution $t = xu$ yields

$$\frac{\gamma(x)}{x} \;=\; \frac{(\log x)^{\varrho}}{\Gamma(\varrho+1)} \int_0^x \frac{1-(1-\frac{t}{x})^x - t(1-\frac{t}{x})^{x-1}}{t^2} \left(1 - \frac{\log t}{\log x}\right)^{\varrho} \mathrm{d}t.$$

A careful analysis shows that, as $x \to \infty$, the latter integral is asymptotically equal to $\int_0^{\infty} (1 - e^{-t} - te^{-t})/t^2\,\mathrm{d}t + O(1/\log x) = [(e^{-t}-1)/t]_0^{\infty} + O(1/\log x) = 1 + O(1/\log x)$, which implies that

$$\frac{\gamma(x)}{x} - \frac{(\log x)^{\varrho}}{\Gamma(\varrho+1)} \;=\; O((\log x)^{\rho-1}) \;\to\; 0, \qquad x \to \infty.$$

Proposition III.2.6 (i), applied with $\gamma_1(x) := x(\log x)^\varrho/\Gamma(\varrho+1)$, shows that the scaling $v(x,t)$ in Theorem III.2.7 satisfies $v(x,t) \sim v_1(x,t)$ as $x \to \infty$, where $v_1(x,t)$ is the solution to the equation

$$\begin{aligned} t &= \int_{v_1(x,t)}^x \frac{\mathrm{d}u}{\gamma_1(u)} = \Gamma(\varrho+1) \int_{v_1(x,t)}^x \frac{\mathrm{d}u}{u(\log u)^\varrho} \\ &= \frac{\Gamma(\varrho+1)}{1-\varrho}\big((\log x)^{1-\varrho} - (\log v_1(x,t))^{1-\varrho}\big), \end{aligned}$$

whenever it exists and $v_1(x,t) := 1$ otherwise. Define $C_\varrho := (1-\varrho)/\Gamma(1+\varrho)$. Solving for $v_1(x,t)$ yields

$$v_1(x,t) = \exp\Big(\Big((\log x)^{1-\varrho} - C_\varrho t\Big)^{\frac{1}{1-\varrho}}\Big), \qquad x > \exp\Big((C_\varrho t)^{(1-\varrho)^{-1}}\Big).$$

By Theorem III.2.7, the process $(\log N_t^{(n)} - \log v_1(n,t))_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to an Ornstein–Uhlenbeck type process X, whose generator A^X satisfies (see (III.33))

$$A^X f(x) = \frac{1}{\Gamma(\varrho)} \int_0^1 \big(f(x+\log(1-u)) - f(x) + uf'(x)\big) \frac{(-\log u)^{\varrho-1}}{u^2}\,\mathrm{d}u,$$

$f \in D, x \in \mathbb{R}$ where D, the space of all 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 A^X. Clearly, Theorem III.2.10 is applicable as well. Similar arguments as for the block counting process show that $w(x,t) \sim w_1(x,t)$ as $x \to \infty$, where

$$w_1(x,t) := \exp\Big(\Big(\frac{(1-\varrho)t}{\Gamma(\varrho+1)} + (\log x)^{1-\varrho}\Big)^{\frac{1}{1-\varrho}}\Big).$$

Thus, the logarithmically scaled fixation line $(\log L_t^{(n)} - \log w_1(n,t))_{t\geq 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ as $n \to \infty$ to an Ornstein–Uhlenbeck type process Y. The generator A^Y of the limiting process Y satisfies

$$A^Y g(y) = \frac{1}{\Gamma(\varrho)} \int_0^1 \big(g(y-\log(1-u)) - g(y) - ug'(y)\big) \frac{(-\log u)^{\varrho-1}}{u^2}\,\mathrm{d}u,$$

$g \in D, y \in \mathbb{R}$. Note that Assumption A of [27] is not satisfied in the situation of Case 3, so both convergence results, for the block counting process and the fixation line, cannot be derived from the results provided in [27].

We provide another example of a dust-free Λ-coalescent which nevertheless satisfies $\kappa = 0$.

Example III.3.3. Assume that the measure Λ has density $u \mapsto 1/(1-\log u)$, $u \in (0,1)$, with respect to Lebesgue measure on $(0,1)$. Then $\Lambda([0,\varepsilon]) = \int_0^\varepsilon 1/(1-\log u)\,\mathrm{d}u \sim e\varepsilon/(-\log\varepsilon)$ as $\varepsilon \to 0+$, and, hence, $\kappa = \lim_{\varepsilon\to 0+}\varepsilon^{-1}\Lambda([0,\varepsilon]) = 0$. Nevertheless, the corresponding Λ-coalescent is dust-free, since $\int u^{-1}\,\Lambda(\mathrm{d}u) = [-\log(1-\log u)]_0^1 = \infty$. The function L in Proposition III.2.2 (iii) satisfies $L(x) = e^{\gamma(x)/x}$ and, hence, $L(x) \sim \log x$ as $x \to \infty$. Theorem III.2.7 is applicable. By Proposition III.2.6 (i), applied with $\gamma_1(x) := x\log\log x$, the scaling $v(x,t)$ can be chosen as the solution to the integral equation

$$\begin{aligned} t \;&=\; \int_{v(x,t)}^x \frac{1}{u\log\log u}\,\mathrm{d}u \;=\; [-\mathrm{Ei}(1,-\log\log u)]_{v(x,t)}^x \\ &=\; \mathrm{Ei}(1,-\log\log v(x,t)) - \mathrm{Ei}(1,-\log\log x), \end{aligned}$$

where $\mathrm{Ei}(x) := \int_1^\infty t^{-1}e^{-xt}\,\mathrm{d}t$ denotes the exponential integral. By Lemma III.4.1, the generator A^X of the limiting process X in Theorem III.2.7 satisfies

$$A^X f(x) \;=\; \int_0^1 \frac{f(x+\log(1-u)) - f(x) + uf'(x)}{u^2(1-\log u)}\,\mathrm{d}u,$$

$x \in \mathbb{R}, f \in D$, where D denotes 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 $\widehat{C}(\mathbb{R})$. We leave the formulation of the analogous results for the fixation line to the interested reader.

In the following an example with simultaneous multiple collisions is provided. The basic idea is to choose the measure Ξ such that the corresponding Ξ-coalescent is dust-free, regular and stays infinite. We slightly modify the example studied in [19] as follows.

Example III.3.4. Let $p_1, p_2, \ldots \in (0,1)$ with $\sum_{m=1}^\infty p_m < \infty$ and let $k_1, k_2, \ldots \in \mathbb{N}$ such that $k_m p_m < 1$ for all $m \in \mathbb{N}$ and $\sum_{m=1}^\infty k_m p_m < \infty$. Suppose that Ξ assigns for each $m \in \mathbb{N}$ mass p_m to the point $x^{(m)} \in \Delta$ whose first k_m coordinates are equal to p_m and all other coordinates are equal to 0. Note that $\Xi(\Delta) = \sum_{m=1}^\infty p_m < \infty$. Moreover, $|x^{(m)}| = k_m p_m < 1$ for all $m \in \mathbb{N}$ and, hence, $\Xi(\Delta_f) = 0$ and $\Xi(\Delta^*) = 0$. The corresponding Ξ-coalescent is dust-free, since

$$\begin{aligned} \int_\Delta |u|\,\nu(\mathrm{d}u) \;&=\; \sum_{m=1}^\infty |x^{(m)}| \frac{p_m}{(x^{(m)}, x^{(m)})} \\ &=\; \sum_{m=1}^\infty k_m p_m \frac{p_m}{k_m p_m^2} \;=\; \sum_{m=1}^\infty 1 \;=\; \infty, \end{aligned}$$

and regular, since $\int_\Delta |u|^2\,\nu(\mathrm{d}u) = \sum_{m=1}^\infty (k_m p_m)^2 p_m/(k_m p_m^2) = \sum_{m=1}^\infty k_m p_m < \infty$. Note that (see [19, Proposition 1]) all regular Ξ-coalescents are non-critical, i.e., $\nu(\Delta \setminus \Delta^\varepsilon) < \infty$ for some $\varepsilon \in (0,1)$. For all $x \geq 0$,

$$\begin{aligned}
\gamma(x) &= \int_\Delta \sum_{i\geq 1} \big((1-u_i)^x - 1 + xu_i\big)\nu(\mathrm{d}u) \\
&= \sum_{m=1}^\infty k_m\big((1-p_m)^x - 1 + xp_m\big)\frac{p_m}{k_m p_m^2} \\
&= \sum_{m=1}^\infty \frac{(1-p_m)^x - 1 + xp_m}{p_m}.
\end{aligned}$$

Note that $\gamma(x)$ does not depend on the sequence $(k_m)_{m\in\mathbb{N}}$ and is hence solely determined by the sequence $(p_m)_{m\in\mathbb{N}}$.

For example, if $p_m = p^m$, $m \in \mathbb{N}$, for some $p \in (0, 1/2]$, then [19, Example 6.1 b)],

$$\gamma(x) \sim \kappa_p x \log x, \qquad x \to \infty,$$

with constant $\kappa_p := -1/\log p$. Thus, $\sum_{n=2}^\infty 1/\gamma(n) = \infty$. By Schweinsberg's criterion [31, Proposition 33] for non-critical coalescents, the Ξ-coalescent stays infinite.

We have hence constructed a class of dust-free and regular Ξ-coalescents that stay infinite. For all $x > 0$,

$$\begin{aligned}
\gamma''(x) &= \sum_{m=1}^\infty \frac{(1-p^m)^x(\log(1-p^m))^2}{p^m} \\
&\sim \int_0^\infty \frac{(1-p^t)^x(\log(1-p^t))^2}{p^t}\,\mathrm{d}t \\
&= \kappa_p \int_0^1 \frac{(1-u)^x(\log(1-u))^2}{u^2}\,\mathrm{d}u \\
&\sim \kappa_p \int_0^1 (1-u)^x\,\mathrm{d}u = \frac{\kappa_p}{x+1},
\end{aligned}$$

which shows that $x\gamma''(x) \to \kappa_p$ as $x \to \infty$. Thus, Theorems III.2.7 and III.2.10 are applicable.

For other choices of the sequence $(p_m)_{m\in\mathbb{N}}$ one obtains further examples with different behavior. Intuitively, $\gamma(x)/x$ grows very slowly if p_m tends to 0 extremely fast. One such choice is $p_m := p^{e^m}$ in which case we have $\gamma(x) \sim x \log\log x$ as $x \to \infty$, see also [19, Example 6.1 c)]. In this case $\kappa := \lim_{x\to\infty} x\gamma''(x) = 0$. Nevertheless the coalescent is dust-free. Theorems III.2.7 and III.2.10 are applicable with scalings $v(x,t) = xL_t(x)$ and $w(x,t) = xL_t^{\#}(x)$, where L_t and $L_t^{\#}$ are the slowly varying functions from Propositions III.2.5 and III.2.9, respectively.

We end this section by providing a simple example of a Λ-coalescent that does not come down from infinity but nevertheless has curvature parameter $\kappa = \infty$.

Example III.3.5. Let $\alpha > 0$. Consider a Λ-coalescent such that $\gamma(x) \sim x(\log x)(\log\log x)^\alpha$. Such a Λ-coalescent can be easily constructed. By Cauchy's condensation test, the series $\sum_{n=2}^\infty 1/\gamma(n)$ converges if and only if $\alpha > 1$. By Schweinsberg's criterion, this coalescent therefore comes down from infinity if and only if $\alpha > 1$. However, $\kappa = \infty$, no matter how $\alpha > 0$ is chosen. For $\alpha \le 1$, this coalescent does not come down from infinity but nevertheless satisfies $\kappa = \infty$.

III.4 Proofs

III.4.1 The function γ

Proof. (of Lemma III.2.1) First assume that $a = 0$. Clearly, $\gamma(0) = \gamma(1) = 0$. From Bernoulli's inequality (and $\nu(\{(1,0,\ldots)\}) = 0$) it follows that $\gamma(x) > 0$ for $x > 1$. By the mean value theorem, there exist $\xi_i^{(1)} \in (0,u_i)$ and $\xi_i^{(2)} \in (0,\xi_i^{(1)})$ for $x \ge 2, i \in \mathbb{N}$ and $u \in \Delta$ such that $\sum_{i\ge1}(x^{-1}((1-u_i)^x - 1) + u_i) = \sum_{i\ge1} u_i(1-(1-\xi_i^{(1)})^{x-1}) = (x-1)\sum_{i\ge1} u_i\xi_i^{(1)}(1-\xi_i^{(2)})^{x-2} \le (x-1)(u,u)$. Hence, $\gamma(x)/x = \int_\Delta \sum_{i\ge1}(x^{-1}((1-u_i)^x-1)+u_i)\,\nu(\mathrm{d}u) \le (x-1)\Xi_0(\Delta)$.

Let $u \in \Delta$. By [25, Lemma 4.1], the map $\Phi(x) := \sum_{i\ge1}(1-(1-u_i)^x)$, $x \ge 0$, is infinitely often differentiable on $(0,\infty)$ with derivatives

$$\Phi^{(k)}(x) \;=\; -\sum_{i\ge1}(1-u_i)^x(\log(1-u_i))^k, \qquad x>0, k\in\mathbb{N}, u\in\Delta.$$

Thus,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\sum_{i\ge1}\big((1-u_i)^x - 1 + xu_i\big)\\
&\quad= \frac{\mathrm{d}}{\mathrm{d}x}\big(x|u| - \Phi(x)\big) \;=\; |u| \;+\; \sum_{i\ge1}(1-u_i)^x\log(1-u_i)\\
&\quad= \sum_{i\ge1}\big((1-u_i)^x\log(1-u_i) + u_i\big)
\end{aligned}$$

and

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}\sum_{i\geq 1}\big((1-u_i)^x-1+xu_i\big) \;=\; -\Phi^{(k)}(x)$$
$$=\; \sum_{i\geq 1}(1-u_i)^x(\log(1-u_i))^k,$$

$k\in\mathbb{N}\backslash\{1\}$. Note that, for every $k\in\mathbb{N}$, the k-th derivative is bounded by $C_k(u,u)$ for some $C_k>0$. Hence, it is allowed to differentiate (with respect to x) below the integral such that $\gamma\in C_\infty((0,\infty))$ with derivatives as stated in the lemma.

Since $\gamma'(x)>0$ for $x>1$, the function γ is strictly increasing on $[1,\infty)$. Since $x\mapsto x^{-1}((1-u_i)^x-1)$, $x\geq 1$, is strictly increasing for every $u\in\Delta$ and $i\in\mathbb{N}$, the map $x\mapsto\gamma(x)/x=\int_\Delta\sum_{i\geq 1}(x^{-1}((1-u_i)^x-1)+u_i)\,\nu(\mathrm{d}u)$, $x\geq 1$, is strictly increasing as well.

For $a>0$ the value $a\binom{x}{2}$ is added, which shows that the results remain valid for $a>0$ as stated in the lemma. □

Proof. (of Proposition III.2.2) We prove this proposition by verifying the implications "(i) ⇒ (ii) ⇔ (iii) ⇔ (iv) ⇒ (v)" and "(v) ⇒ (i)". Define the functions $g,L:(0,\infty)\to(0,\infty)$ via $g(x):=\exp(\gamma(x)/x)=x^\kappa L(x)$, $x>0$. From $g'(x)=g(x)(\gamma'(x)/x-\gamma(x)/x^2)$ and $\frac{\mathrm{d}}{\mathrm{d}x}(x\gamma'(x)-\gamma(x))=x\gamma''(x)$, $x>0$, it follows that

$$\frac{xg'(x)}{g(x)} \;=\; \gamma'(x)-\frac{\gamma(x)}{x} \;=\; \frac{1}{x}\bigg(\int_1^x u\gamma''(u)\,\mathrm{d}u \;+\; \gamma'(1)\bigg), \qquad \text{(III.30)}$$

$x>0$. Due to the second equality of (III.30), (i) implies (ii). The map $x\mapsto x^2g'(x)=g(x)(x\gamma'(x)-\gamma(x))$, $x>0$, is non-decreasing, since

$$\frac{\mathrm{d}}{\mathrm{d}x}g(x)(x\gamma'(x)-\gamma(x)) \;=\; g(x)\bigg(\Big(\gamma'(x)-\frac{\gamma(x)}{x}\Big)^2 \;+\; x\gamma''(x)\bigg) \;\geq\; 0,$$

$x>0$. Applying [20, Theorem 2] to the function $x\mapsto g(x^{-1})$, $x>0$, hence shows that (ii) is equivalent to the regular variation of g of index κ, i.e., equivalent to the slow variation of L. Thus, (ii) and (iii) are equivalent. Conditions (iii) and (iv) are equivalent by the definition of slow variation. Suppose that (iv) holds. It is already proven that (iv) implies (ii) such that

$$\lim_{x\to\infty}\big(\gamma'(yx)-\gamma'(x)\big) \overset{\text{(ii)}}{=} \lim_{x\to\infty}\bigg(\frac{\gamma(yx)}{yx}-\frac{\gamma(x)}{x}\bigg) \overset{\text{(iv)}}{=} \kappa\log y, \qquad y>0,$$

and (v) holds. Finally, it is shown that (v) implies (i). By the mean value theorem, there exists $\xi = \xi(x, y)$ between y and 1 such that $\gamma'(yx) - \gamma'(x) = \gamma''(x\xi)x(y-1)$ for all $x, y > 0$. Given (v), $\lim_{x\to\infty} x\gamma''(x\xi) = (\kappa \log y)/(y-1)$ for $y > 0, y \neq 1$. Since γ'' is non-negative and ultimately non-increasing, $\limsup_{x\to\infty} x\gamma''(x) \leq \lim_{x\to\infty} x\gamma''(x\xi(x,y)) = (\kappa \log y)/(y-1)$ for all $y \in (0,1)$ and $\liminf_{x\to\infty} x\gamma''(x) \geq \lim_{x\to\infty} x\gamma''(x\xi(x,y)) = (\kappa \log y)/(y-1)$ for all $y > 1$. Letting $y \to 1$ establishes (i), since $\lim_{y\to 1}(\log y)/(y-1) = 1$. □

Proof. (of Proposition III.2.3) We prove the equivalence of (i) and (ii) and afterwards the equivalence of (ii) and (iii).

(i) ⇔ (ii): Define $U(t) := 0$ for $t < 0$ and $U(t) := F(1-e^{-t})$ for $t \geq 0$, with F defined as in (III.7). Let $x > 0$. By monotone convergence and Fubini's theorem,

$$\begin{aligned}
\gamma''(x) &= \sum_{i\geq 1} \int_\Delta (1-u_i)^x (\log(1-u_i))^2 \,\nu(\mathrm{d}u) \\
&= \sum_{i\geq 1} \int_\Delta \int_{[u_i,1)} x(1-y)^{x-1} \,\lambda(\mathrm{d}y)(\log(1-u_i))^2 \,\nu(\mathrm{d}u) \\
&= \sum_{i\geq 1} x \int_{[0,1)} (1-y)^{x-1} \int_\Delta 1_{[0,y]}(u_i)(\log(1-u_i))^2 \,\nu(\mathrm{d}u)\lambda(\mathrm{d}y) \\
&= \sum_{i\geq 1} x \int_0^1 (1-y)^{x-1} F_i(y)\,\mathrm{d}y \;=\; x\int_0^1 (1-y)^{x-1}F(y)\,\mathrm{d}y \\
&= x\int_0^\infty e^{-xt}F(1-e^{-t})\,\mathrm{d}t \;=\; x\int_0^\infty e^{-xt}U(t)\,\mathrm{d}t \;=:\; \widehat{U}(x).
\end{aligned}$$

Thus, the map $\gamma'' = \widehat{U}$ is the Laplace–Stieltjes transform of U. By Karamata's Tauberian theorem [4, Theorem 1.7.1'], applied with $\ell \equiv 1$, $c := \kappa \in [0,\infty)$ and $\rho := 1$, the condition $x\widehat{U}(x) = x\gamma''(x) \to \kappa$ as $x \to \infty$ is equivalent to $t^{-1}U(t) \to \kappa$ as $t \to 0+$, which shows that (i) and (ii) are equivalent.

(ii) ⇔ (iii): Let $\varepsilon > 0$. Choose $\delta = \delta(\varepsilon) \in (0,1)$ sufficiently small such that $(\log(1-x))^2 \leq (1+\varepsilon)x^2$ for all $x \in [0,\delta]$. For all $i \in \mathbb{N}$ and $t \in [0,\delta]$ it follows that

$$\begin{aligned}
F_i(t) &= \int_\Delta 1_{[0,t]}(u_i)(\log(1-u_i))^2 \,\nu(\mathrm{d}u) \\
&\leq (1+\varepsilon)\int_\Delta 1_{[0,t]}(u_i)u_i^2 \,\nu(\mathrm{d}u) \;=\; (1+\varepsilon)G_i(t).
\end{aligned}$$

Summation over all $i \in \mathbb{N}$ yields $F(t) \leq (1+\varepsilon)G(t)$ for all $t \in [0,\delta]$. Thus, $\liminf_{t\to 0+} t^{-1}F(t) \leq (1+\varepsilon)\liminf_{t\to 0+} t^{-1}G(t)$ and $\limsup_{t\to 0+} t^{-1}F(t) \leq (1+\varepsilon)\limsup_{t\to 0+} t^{-1}G(t)$. Since $\varepsilon > 0$ can be chosen arbitrarily small it follows that $\liminf_{t\to 0+} t^{-1}F(t) \leq \liminf_{t\to 0+} t^{-1}G(t)$ and $\limsup_{t\to 0+} t^{-1}F(t) \leq \limsup_{t\to 0+} t^{-1}G(t)$. The converse two inequalities are obviously satisfied, since $G(t) \leq F(t)$ for all $t \in [0,1)$. The equivalence of (ii) and (iii) now follows immediately.

For Λ-coalescents it is easily seen that $G(t) = \Lambda([0,t])$, $t \in [0,1]$, showing that (iii) reduces to (III.8). □

III.4.2 The normalizing function v

Recall that $v(x,t)$ is the solution to $\int_{v(x,t)}^{x} (\gamma(u))^{-1}\,\mathrm{d}u = t$ for $x > 1$ and $t \geq 0$.

Proof. (of Proposition III.2.4) In order to see that v is well-defined fix $x > 1$ and define $F_x : (1,x] \to [0,\infty)$ via $F_x(y) := \int_y^x (\gamma(u))^{-1}\,\mathrm{d}u$, $y \in (1,x]$. Then $F_x(y) > 0$ for $y \in (1,x]$, since $\gamma(u) > 0$ for $u > 1$, and $F_x \in C_1((1,x])$ with $F_x'(y) = -(\gamma(y))^{-1}$, $y \in (1,x]$, since γ is continuous. In particular, F_x is strictly decreasing. Clearly, $F_x(x) = 0$. There exists $C \in \mathbb{R}$ such that $\gamma(u) \leq u(u-1)C$ for $u > 1$. Then

$$F_x(y) \;\geq\; \frac{1}{C}\int_y^x \frac{\mathrm{d}u}{u(u-1)} \;=\; \frac{1}{C}\log\frac{1-x^{-1}}{1-y^{-1}}, \qquad y \in (1,x],$$

such that $\lim_{y\to 1+} F_x(y) = \infty$. By the intermediate value theorem, the solution $v(x,t) \in (1,x]$ to the equation $F_x(v(x,t)) = t$ exists and is unique for every $t \geq 0$. Since $F_x'(y) < 0$ for $y \in (1,x]$, the function F_x is injective and the inverse function $F_x^{-1} : [0,\infty) \to (1,x]$ exists and is differentiable with $(F_x^{-1})'(t) = -\gamma(F_x^{-1}(t))$, $t \geq 0$. Hence, $t \mapsto v(x,t) = F_x^{-1}(t)$, $t \geq 0$, is differentiable and

$$\frac{\mathrm{d}}{\mathrm{d}t}v(x,t) \;=\; -\gamma(v(x,t)), \qquad t \geq 0.$$

Differentiating both sides of the integral equation in (III.9) with respect to x leads to $(\gamma(x))^{-1} - \frac{\mathrm{d}}{\mathrm{d}x}v(x,t)(\gamma(v(x,t)))^{-1} = 0$. Equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}x}v(x,t) \;=\; \frac{\gamma(v(x,t))}{\gamma(x)}, \qquad x > 1, t \geq 0.$$

The two monotonicity statements follow from the formulas for the derivatives (and can also be deduced directly from Eq. (III.9)). □

Proof. (of Proposition III.2.5) (i) Suppose that $\Xi(\{0\}) = 0$ and $\mu := \int_\Delta |u|\,\nu(\mathrm{d}u) < \infty$. Fix $t > 0$ and let $\varepsilon > 0$ be arbitrary. Due to $\lim_{x\to\infty} \gamma(x)/x = \mu$, there exists $x_0 > 1$ such that $(\mu u)/\gamma(u) \in (1-\varepsilon, 1+\varepsilon)$ for all $u \in (v(x,t), x)$ as long as $x \geq x_0$. Thus,

$$\frac{1-\varepsilon}{\mu}\int_{v(x,t)}^{x}\frac{\mathrm{d}u}{u} \;\leq\; \int_{v(x,t)}^{x}\frac{\mathrm{d}u}{\gamma(u)} \;\leq\; \frac{1+\varepsilon}{\mu}\int_{v(x,t)}^{x}\frac{\mathrm{d}u}{u},$$

such that, by Eq. (III.9), $\exp(-\mu t/(1-\varepsilon)) \leq v(x,t)/x \leq \exp(-\mu t/(1+\varepsilon))$ for $x \geq x_0$. It follows that $v(x,t) \sim xe^{-\mu t}$ as $x \to \infty$, since $\varepsilon > 0$ can be chosen arbitrarily small. Clearly, $v(x,0) = x$, so the statement also holds for $t = 0$.

(ii) Define $F : (1,\infty) \to (0,\infty)$ via $F(y) = \int_y^\infty (\gamma(u))^{-1}\,\mathrm{d}u$, $y > 1$, and suppose that $F(y) < \infty$ for some (and hence all) $y > 1$. Similarly to the proof of (i), it follows that $\lim_{y\to 1+} F(y) = \infty$, $\lim_{y\to\infty} F(y) = 0$, $F \in C_1((1,\infty))$ and $F'(y) = -(\gamma(y))^{-1}$, $y > 1$. Thus, the solution $v(t)$ to the equation $F(v(t)) = t$ exists and is unique for every $t \geq 0$. The limit $c(t) := \lim_{x\to\infty} v(x,t)$ exists for every $t > 0$, since $x \mapsto v(x,t)$, $x \geq 1$, is non-decreasing, and $c(t) < \infty$ due to $\lim_{y\to\infty} F(y) = 0$. From

$$\int_{v(x,t)}^{v(t)}\frac{\mathrm{d}u}{\gamma(u)} \;=\; \int_{v(x,t)}^{x}\frac{\mathrm{d}u}{\gamma(u)} - \int_{v(t)}^{\infty}\frac{\mathrm{d}u}{\gamma(u)} + \int_{x}^{\infty}\frac{\mathrm{d}u}{\gamma(u)} \;=\; \int_{x}^{\infty}\frac{\mathrm{d}u}{\gamma(u)},$$

$x \geq 1, t \geq 0$, we obtain that $F(c(t)) - F(v(t)) = \lim_{x\to\infty}(F(v(x,t)) - F(v(t))) = \lim_{x\to\infty} F(x) = 0$. Since F is injective, $c(t) = v(t)$ for each $t > 0$. The proof of (ii) is complete.

(iii) Due to $v(x,0) = x$, the claim is true for $t = 0$ with $L_0(x) = 1$ for $x \geq 1$. Fix $t > 0$. By assumption and Proposition III.2.2, there exists a slowly varying function $L : (0,\infty) \to (0,\infty)$ such that $\gamma(x) = \kappa x \log x + x \log L(x)$ for $x > 0$. The fact that $L(x) = o(x^\varepsilon)$ and $L(x) = \omega(x^{-\varepsilon})$ for every $\varepsilon > 0$ is repeatedly used in this proof.

First suppose that $\kappa > 0$ and let $0 < \varepsilon < \kappa$ be arbitrary. Recall that $\lim_{x\to\infty} v(x,t) = \infty$. There exists $x_0 > 1$ such that $(\kappa - \varepsilon)u \log u \leq \gamma(u) \leq (\kappa + \varepsilon)u \log u$ for every $u \in (v(x,t), x)$ and $x \geq x_0$. Thus,

$$\frac{1}{\kappa+\varepsilon}\int_{v(x,t)}^{x}\frac{\mathrm{d}u}{u\log u} \;\leq\; \int_{v(x,t)}^{x}\frac{\mathrm{d}u}{\gamma(u)} \;\leq\; \frac{1}{\kappa-\varepsilon}\int_{v(x,t)}^{x}\frac{\mathrm{d}u}{u\log u},$$

$x \geq x_0$. Computing the integrals on both sides and using Eq. (III.9) yields

$$\frac{1}{\kappa+\varepsilon}\log\left(\frac{\log x}{\log v(x,t)}\right) \;\leq\; t \;\leq\; \frac{1}{\kappa-\varepsilon}\log\left(\frac{\log x}{\log v(x,t)}\right)$$

or, equivalently, $x^{e^{-(\kappa+\varepsilon)t}} \leq v(x,t) \leq x^{e^{-(\kappa-\varepsilon)t}}$ for all $x \geq x_0$. From (III.10) it follows that

$$\frac{\frac{\mathrm{d}}{\mathrm{d}x}v(x,t)x}{v(x,t)} = \frac{\gamma(v(x,t))}{v(x,t)}\frac{x}{\gamma(x)} = \frac{\kappa\log v(x,t)+\log L(v(x,t))}{\kappa\log x+\log L(x)},$$

$x > 1$, such that

$$\begin{aligned} e^{-(\kappa+\varepsilon)t} &\leq \liminf_{x\to\infty}\frac{\frac{\mathrm{d}}{\mathrm{d}x}v(x,t)x}{v(x,t)} \\ &\leq \limsup_{x\to\infty}\frac{\frac{\mathrm{d}}{\mathrm{d}x}v(x,t)x}{v(x,t)} \leq e^{-(\kappa-\varepsilon)t}. \end{aligned} \tag{III.31}$$

We are going to show similar inequalities for $\kappa = 0$. Suppose that $\gamma(x) = x\log L(x)$, $x > 0$, for some slowly varying function L. By Proposition III.2.2, $(xL'(x))/L(x) = \gamma'(x) - \gamma(x)/x \to \kappa = 0$ as $x \to \infty$. From (III.10) it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\log L(v(x,t)) = \frac{-L'(v(x,t))v(x,t)\log L(v(x,t))}{L(v(x,t))},$$

$x > 1$. Note that $L(x) = \exp(\gamma(x)/x)$ is non-decreasing on $[1,\infty)$ and, by definition, $v(x,t) \leq x$ for all $x \geq 1$. For $\varepsilon > 0$ there exists $x_1 > 1$ such that $|\frac{\mathrm{d}}{\mathrm{d}s}\log L(v(x,s))| \leq \varepsilon\log L(x)$ for all $s \in [0,t]$ and $x \geq x_1$. We hence obtain

$$\begin{aligned} \big|\log L(v(x,t)) - \log L(x)\big| &\leq \int_0^t \left|\frac{\mathrm{d}}{\mathrm{d}s}\log L(v(x,s))\right|\mathrm{d}s \\ &\leq \varepsilon t\log L(x) \end{aligned}$$

such that

$$\frac{\frac{\mathrm{d}}{\mathrm{d}x}v(x,t)x}{v(x,t)} = \frac{\log L(v(x,t))}{\log L(x)} \in [1-\varepsilon t, 1+\varepsilon] \tag{III.32}$$

for all $x \geq x_1$. Letting $\varepsilon \to 0+$ in (III.31) and (III.32) yields $\lim_{x\to\infty}\frac{\frac{\mathrm{d}}{\mathrm{d}x}v(x,t)x}{v(x,t)} = e^{-\kappa t}$ for $\kappa \geq 0$. A “variant at infinity” of [20, Theorem 2] completes the proof. □

Proof. (of Proposition III.2.6) (i) Define the function $L_1 : (1,\infty) \to (0,\infty)$ via $\gamma_1(x) = \kappa x\log x + x\log L_1(x)$, $x > 1$. By assumption,

$$r(x) := \frac{\gamma(x)-\gamma_1(x)}{x} = \log\left(\frac{L(x)}{L_1(x)}\right) \to 0, \qquad x \to \infty.$$

In particular, $L(x) \sim L_1(x)$ as $x \to \infty$. Hence, L_1 is slowly varying and, as a consequence, γ_1 satisfies (III.6). Unfortunately, the scaling $v_1(x,t)$, defined by the integral equation in (III.9) with γ_1 in place of γ, does not exist globally. The reason is that the condition $\lim_{y\to 1+} F_x(y) = \infty$ from the proof of Proposition III.2.4 cannot be guaranteed. However, γ_1 is continuous and positive on $(1,\infty)$ and $\int_c^\infty (\gamma_1(u))^{-1} \mathrm{d}u = \infty$ for each $c > 1$. Carefully reading the proof of Proposition III.2.4 shows that the statements of Proposition III.2.4 remain true with the restriction that, for each $t \geq 0$, $x \geq x_0(t)$ for some $x_0(t) > 1$. Moreover, we can choose $x_0(t)$ in such a way that $x_0(s) \leq x_0(t)$ for $s \leq t$. In particular, the scaling $v_1(x,t)$ exists for $x \geq x_0(t)$ and $t \geq 0$. Now fix $t \geq 0$. From (III.10) it follows that

$$\begin{aligned} \log v(x,t) - \log x &= -\int_0^t \frac{\gamma(v(x,s))}{v(x,s)}\,\mathrm{d}s \\ &= -\int_0^t \big(\kappa \log v(x,s) + \log L(v(x,s)\big)\mathrm{d}s \end{aligned}$$

for $x > 1$. The same equalities hold when $x \geq x_0(t)$ and $v(x,t)$ and L are replaced by $v_1(x,t)$ and L_1, respectively. Then, for $x > x_0(t)$,

$$\begin{aligned} \left|\log \frac{v(x,t)}{v_1(x,t)}\right| &\leq \kappa \int_0^t \left|\log \frac{v(x,s)}{v_1(x,s)}\right| \mathrm{d}s \\ &\qquad + t \sup_{y \geq v(x,t)} |r(y)| + \int_0^t \left|\log \frac{L_1(v(x,s))}{L_1(v_1(x,s))}\right| \mathrm{d}s. \end{aligned}$$

Let $c_1, c_2 > 0$ be arbitrary. The representation theorem for slowly varying functions [4, Theorem 1.3.1] states the existence of functions $\varepsilon, \delta : (0,\infty) \to \mathbb{R}$ with $\lim_{x\to\infty} \varepsilon(x) = 0$ and $\lim_{x\to\infty} \delta(x) = d \in \mathbb{R}$ such that $\log L_1(x) = \delta(x) + \int_1^x (\varepsilon(u)/u)\,\mathrm{d}u$, $x > 0$. Hence,

$$\begin{aligned} \left|\log \frac{L_1(v(x,s))}{L_1(v_1(x,s))}\right| &\leq \big|\delta(v(x,s)) - \delta(v_1(x,s))\big| + \int_{v_1(x,s)}^{v(x,s)} \frac{|\varepsilon(u)|}{u}\,\mathrm{d}u \\ &\leq c_2 + c_1 \left|\log \frac{v(x,s)}{v_1(x,s)}\right| \end{aligned}$$

for sufficiently large x and $s \in [0,t]$, where in the last inequality it is used that $\inf_{s\in[0,t]} v(x,s) = v(x,t) \to \infty$ and $\inf_{s\in[0,t]} v_1(x,s) \to \infty$ as $x \to \infty$. Thus,

$$\left|\log \frac{v(x,t)}{v_1(x,t)}\right| \leq t\Big(\sup_{y\geq v(x,t)} |r(y)| + c_2\Big) + (\kappa + c_1) \int_0^t \left|\log \frac{v(x,s)}{v_1(x,s)}\right| \mathrm{d}s.$$

By Gronwall's inequality,

$$\left|\log \frac{v(x,t)}{v_1(x,t)}\right| \leq t\Big(\sup_{y\geq v_1(x,t)} |r(y)| + c_2\Big) \exp(t(\kappa + c_1)).$$

We conclude that $\lim_{x\to\infty} |\log v(x,t) - \log v_1(x,t)| = 0$, which completes the proof of (i), because $c_2 > 0$ is arbitrarily small.

(ii) First assume that $\gamma : (0,\infty) \to (0,\infty)$ is a function of the form (III.6) with $L \equiv C$ for some constant $C > 0$. Then the integral in (III.9) can be calculated explicitly and it is easily seen that $v(x,t)$, defined via $v(x,t) := x^{e^{-\kappa t}} C^{\kappa^{-1}(e^{-\kappa t}-1)}$ if $\kappa > 0$ and $v(x,t) := xC^{-t}$ if $\kappa = 0$, solves (III.9) for every $x > 1$ and $t \geq 0$. If L only satisfies $L(x) \to C$ as $x \to \infty$, then the same formulas for $v(x,t)$ hold, but with equality replaced by asymptotic equality as $x \to \infty$, as shown in (i). □

III.4.3 Proof of Theorem III.2.7

The scaled block counting process and the scaled fixation line are in general time-inhomogeneous Markov processes. We therefore add a further "time variable" and consider the associated time-space processes, which are time-homogeneous. We want to show the uniform convergence of the generators. First a well known result ([30, Theorem 3.1]) concerning generators of Ornstein–Uhlenbeck type processes on $\mathbb{R}^d$ is applied. The short proof is an adaption of the proof of [27, Lemma 6] to the Ξ-coalescent setting.

Lemma III.4.1. *Suppose that* Ξ *satisfies* (III.1) *and* $\Xi(\Delta^*) = 0$. *Fix* $\kappa \in [0,\infty)$ *and let the family of operators* $(T_t^X)_{t\geq 0}$ *be defined by* (III.17). *Then* $(T_t^X)_{t\geq 0}$ *is a Feller semigroup on* $\widehat{C}(\mathbb{R})$. *Let* D *denote the space of all 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* $\widehat{C}(\mathbb{R})$. *Then* D *is a core for the generator* A^X *corresponding to* $(T_t^X)_{t\geq 0}$ *and*

$$\begin{aligned} A^X f(x) &= -\kappa x f'(x) && \text{(III.33)} \\ &\quad + \int_\Delta \big(f(x + \log(1-|u|)) - f(x) + |u| f'(x)\big)\nu(\mathrm{d}u), \end{aligned}$$

$x \in \mathbb{R}, f \in D$.

Proof. (of Lemma III.4.1) Substituting $g : \Delta \setminus \Delta^* \to \mathbb{R}$, $g(u) := \log(1-|u|)$, $u \in \Delta \setminus \Delta^*$, shows that (III.33) is an integro-differential operator of the form (1.1) of Sato and Yamazato [30] with dimension $d = 1$. In [30], 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 [30]), but Step 3 of the proof of [30, Theorem 3.1] shows that (III.33) even holds for functions $f \in D$ ($\supset C_c^2$). Note that the space D is denoted by F_1 in [30].

The fact that D is a core for A^X is only a different phrasing of the claim in Step 5 of the proof of [30, Theorem 3.1]. $\square$

When writing semigroups or generators in the remainder of the proof section, we mostly omit the upper index that identifies the corresponding process. We only use the symbol tilde to indicate the time-space process.

The time-space process $\widetilde{X} := (t, X_t)_{t\geq 0}$ is a time-homogeneous Markov process with state space $\widetilde{E} := [0,\infty) \times \mathbb{R}$ and semigroup $\widetilde{T} := (\widetilde{T}_t)_{t\geq 0}$, given by

$$\widetilde{T}_t f(s,x) \ := \ \mathbb{E}(f(s+t, e^{-\kappa t}x + S_t)),$$

$(s,x) \in \widetilde{E}, f \in B(\widetilde{E}), t \geq 0$. For $f \in \widehat{C}(\widetilde{E})$ and $s \geq 0$, let the map $x \mapsto f(s,x)$, $x \in \mathbb{R}$, be denoted by $\pi f(s,x)$. Let $\widetilde{D}$ denote the space of functions $f \in \widehat{C}(\widetilde{E})$ of the form $f(s,x) = \sum_{i=1}^{l} g_i(s)h_i(x)$, $(s,x) \in \widetilde{E}$, 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 $i \in \{1,\ldots,l\}$. By [27, Proposition 10], $\widetilde{T}$ is a Feller semigroup, $\widetilde{D}$ is a core for the generator $\widetilde{A}$ corresponding to $\widetilde{T}$ and

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

For $n \in \mathbb{N}$ the logarithmically scaled block counting process $X^{(n)} := (X_t^{(n)})_{t\geq 0} := (\log N_t^{(n)} - \log v(n,t))_{t\geq 0}$ is a time-inhomogeneous Markov process. The random variable $X_s^{(n)}$ takes values in $E_{n,s} := \{x \in \mathbb{R} : e^x v(n,s) \in [n]\}$. The "generator" $(A_s^{(n)})_{s\geq 0}$ of $X^{(n)}$ is given by

$$\begin{aligned}
A_s^{(n)} f(x) \ &= \ f'(x)\frac{-\frac{\mathrm{d}}{\mathrm{d}s}v(n,s)}{v(n,s)} \\
&\quad + \sum_{j=1}^{xv(n,s)-1} \big(f(\log j - \log v(n,s)) - f(x)\big) q_{xv(n,s),j} \\
&= \ f'(x)\frac{\gamma(v(n,s))}{v(n,s)} \\
&\quad + \sum_{j=1}^{xv(n,s)-1} \big(f(\log j - \log v(n,s)) - f(x)\big) q_{xv(n,s),j}
\end{aligned}$$

for $x \in E_{n,s}$ and $s \geq 0$. Here $f \in C_1(\mathbb{R})$ such that $f, f' \in \widehat{C}(\mathbb{R})$. The time-space process $\widetilde{X}^{(n)} := (t, X_t^{(n)})_{t\geq 0}$ is a time-homogeneous Markov process with state space $\widetilde{E}_n := \{(s,x) \in [0,\infty) \times \mathbb{R} :$

$e^x v(n,s) \in [n]\}$ and semigroup $\widetilde{T}^{(n)} := (\widetilde{T}_t^{(n)})_{t\ge 0}$, given by

$$\widetilde{T}_t^{(n)} f(s,x) \;:=\; \mathbb{E}\big(f(s+t, \log N_t^{(e^x v(n,s))} - \log v(n,s+t))\big),$$

$(s,x) \in \widetilde{E}_n, f \in B(\widetilde{E}), t \ge 0$. For $f \in \widetilde{D}$ (restricted to $\widetilde{E}_n \subset \widetilde{E}$) the corresponding generator $\widetilde{A}^{(n)}$ is given by

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

Proof. (of Theorem III.2.7) Write $k := k(s,x,n) := e^x v(n,s)$ for $(s,x) \in \widetilde{E}_n$ and $n \in \mathbb{N}$. Let $h \in D$. Define $R(k,x) := \frac{\gamma(ke^{-x})}{ke^{-x}} - \gamma(k)/k$ and

$$S(k,x) \;:=\; \sum_{j=1}^{k-1} \big(h(x + \log \tfrac{j}{k}) - h(x) + (1 - \tfrac{j}{k}) h'(x)\big) q_{k,j},$$

$k \in \mathbb{N}, x \in \mathbb{R}$, such that

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

Define the continuous function $I : \mathbb{R} \times [0,1] \to \mathbb{R}$ via $I(x,y) := h(x + \log(1-y)) - h(x) + y h'(x)$, $y \in [0,1)$, and $I(x,1) := -h(x) + h'(x)$ for $x \in \mathbb{R}$. From Eq. (III.13) and the definition of I it follows that

$$S(k,x) \;=\; \int_\Delta \mathbb{E}(I(x, 1 - Y(k,u)/k))\, \nu(\mathrm{d}u), \qquad k \in \mathbb{N}, x \in \mathbb{R}.$$

Also,

$$A^X h(x) \;=\; -\kappa x h'(x) \;+\; \int_\Delta I(x, |u|)\, \nu(\mathrm{d}u), \qquad x \in \mathbb{R}.$$

Part 1 of the proof treats the convergence of $R(k,x)$ and Part 2 the convergence of $S(k,x)$.

Part 1. By assumption and Proposition III.2.2, there exist $\kappa \ge 0$ and a slowly varying function $L : (0,\infty) \to (0,\infty)$ such that $\gamma(x) = \kappa x \log x + x \log L(x)$, $x > 0$. Then $R(k,x) + \kappa x = \log(L(ke^{-x})/L(k))$ for $k \in \mathbb{N}$ and $x \in \mathbb{R}$. Applying [4, Theorem 1.5.6 (ii)], a boundary for the growth of slowly varying functions, yields the existence of $C > 0$ such that $|R(k,x) + \kappa x)| \le C + |x|$ for $k \in \mathbb{N}$ and $-\infty < x \le \log k$. For $c > 0$ there exist $-\infty < K_1 < K_2 < \infty$ such that

$$|h'(x)(R(k,x) + \kappa x)| \le C|h'(x)| + |xh'(x)| \le c, \qquad \text{(III.34)}$$

$x \in \mathbb{R} \setminus [K_1, K_2], x \le \log k, k \in \mathbb{N}$, since h' and the map $x \mapsto xh'(x)$, $x \in \mathbb{R}$, vanish as $|x| \to \infty$. The present restriction $x \le \log k(s,x,n)$

is met for $(s,x) \in \widetilde{E}_n$ and $n \in \mathbb{N}$. Let $T > 0$ be arbitrary. By the uniform convergence theorem for slowly varying functions ([4, Theorem 1.5.2]) and $\lim_{n\to\infty} \inf_{s\in[0,T]} v(n,s) = \infty$,

$$\begin{aligned} &\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T], x\in[K_1,K_2]} |R(k,x)+\kappa x| \qquad \text{(III.35)} \\ &\quad = \lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T], x\in[K_1,K_2]} |\log(L(v(n,s))/L(e^x v(n,s)))| \; = \; 0. \end{aligned}$$

From (III.34), (III.35) and arbitrariness of c it follows that

$$\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} |h'(x)(R(k,x)+\kappa x)| \; = \; 0. \qquad \text{(III.36)}$$

Part 2. Note that, as $n \to \infty$, $k = e^x v(n,s) \to \infty$ or $x \to -\infty$. For example, either $k \geq \sqrt{v(n,T)}$ or $x < -\frac{1}{2}\log v(n,T)$ for each $(s,x) \in \widetilde{E}_n$ with $s \in [0,T]$ and $n \in \mathbb{N}$. In order to prove that

$$\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} \left|\mathbb{E}\big(I(x,1-Y(k,u)/k) - I(x,|u|)\big)\right| \; = \; 0, \quad \text{(III.37)}$$

$u \in \Delta \setminus (\Delta^* \cup \{0\})$, it therefore suffices to show that $\lim_{x\to-\infty} I(x,|u|) = 0$, $\lim_{x\to-\infty} \mathbb{E}(I(x,Y(k,u)/k)) = 0$ for any $k \in \mathbb{N}$ and $\lim_{k\to\infty} \sup_{x\in\mathbb{R}} |\mathbb{E}(I(x,1-Y(k,u)/k) - I(x,|u|))| = 0$ for each $u \in \Delta \setminus (\Delta^* \cup \{0\})$.

Clearly, $\sup_{x\in\mathbb{R},y\in[0,1]} |I(x,y)| \leq 2||h|| + ||h'|| < \infty$. In particular, the family of functions $\mathcal{I} := \{I(x,.) : x \in \mathbb{R}\}$ is uniformly bounded. The family $\mathcal{I}$ is equicontinuous on any interval $[0,c]$ with $c < 1$, since h is uniformly continuous and h' is bounded. In view of [27, Lemma 9], the almost sure convergence of $1 - Y(k,u)/k$ to $|u|$ as $k \to \infty$ implies that $\lim_{k\to\infty} \sup_{x\in\mathbb{R}} |\mathbb{E}(I(x,1-Y(k,u)/k)) - I(x,|u|)| = 0$ for any $u \in \Delta \setminus (\Delta^* \cup \{0\})$. The cited lemma does not allow the limiting "random" variable $|u|$ to assume the values 0 and 1 with positive probability, hence we impose the restriction of u to $\Delta \setminus (\Delta^* \cup \{0\})$. For any $y \in [0,1]$, $\lim_{x\to-\infty} I(x,y) = 0$, since $\lim_{|x|\to\infty} h'(x) = \lim_{|x|\to\infty} h(x) = 0$. Thus, $\lim_{x\to-\infty} I(x,|u|) = 0$ and, by dominated convergence, $\lim_{x\to-\infty} \mathbb{E}(I(x,1-Y(k,u)/k))$ for $k \in \mathbb{N}$ and $u \in \Delta$, which completes the proof of (III.37).

Taylor's theorem applied to $y \mapsto h(x+\log(1-y))$, $y < 1$, evaluated at $y = 0$ with mean value remainder states the existence of $\xi \in (0,y)$ such that

$$I(x,y) \; = \; (1-\xi)^{-2}(h''(x+\log(1-\xi)) - h'(x+\log(1-\xi)))(y-\xi)y,$$

$x \in \mathbb{R}, y \in (0,1)$. In particular, $\sup_{x\in\mathbb{R}} |I(x,y)| \le (1-c)^{-2}(||h''|| - ||h'||)y^2 < \infty$ for $0 \le y \le c$ and any $c < 1$. Thus, there exists $C \in \mathbb{R}$ such that $\sup_{x\in\mathbb{R}} |I(x,y)| \le Cy^2$ for every $y \in [0,1]$. From Lemma III.5.3 it follows that $\sup_{n\in\mathbb{N}} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} |\mathbb{E}(I(x, 1-Y(k,u)/k)) - I(x,|u|)| \le \sup_{k\in\mathbb{N}, x\in\mathbb{R}} |\mathbb{E}(I(x, 1-Y(k,u)/k))| + \sup_{x\in\mathbb{R}} |I(x,|u|)| \le (D_2+1)C|u|^2$ for any $u \in \Delta$. Due to (III.37) and (III.1), the dominated convergence theorem is applicable such that

$$\begin{aligned}
\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} \bigg| \int_\Delta \mathbb{E}(I(x, 1-Y(k,u)/k))\, \nu(\mathrm{d}u) \qquad\qquad & \qquad \text{(III.38)} \\
- \int_\Delta I(x,|u|)\, \nu(\mathrm{d}u) \bigg| & \\
\le \lim_{n\to\infty} \int_\Delta \sup_{(s,x)\in\widetilde{E}_n, s\in[0,T]} \big| \mathbb{E}(I(x, 1-Y(k,u)/k)) & \\
- I(x,|u|) \big| \nu(\mathrm{d}u) \;=\; 0. &
\end{aligned}$$

Here we made use of $\nu(\Delta^* \cup \{0\}) = 0$.

Eqs. (III.36) and (III.38) imply

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

Hence,

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

From [11, 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. Assumption (III.5) is only used in Part 1 of the proof of Theorem III.2.7, whereas Part 2 remains correct for every measure Ξ satisfying $\Xi(\Delta^* \cup \{0\}) = 0$ and $\int_2^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$.

III.4.4 Proofs concerning the fixation line

Propositions III.2.8 and III.2.9 treat the normalizing function $w(x,t)$ for the fixation line, implicitly defined via $\int_x^{w(x,t)} (\gamma(u))^{-1}\,\mathrm{d}u = t$. Proposition III.2.8 verifies the existence of w.

Proof. (of Proposition III.2.8) Suppose that $\int_2^\infty (\gamma(u))^{-1}\,\mathrm{d}u = \infty$. Fix $x > 1$. The function $F_x : [x,\infty) \to \mathbb{R}$, defined by $F_x(y) := \int_x^y (\gamma(u))^{-1}\,\mathrm{d}u$, $y \in [x,\infty)$, is continuous, strictly increasing and satisfies $F_x(x) = 0$ and $\lim_{y\to\infty} F_x(y) = \infty$. Thus, the solution $w(x,t)$

to the equation $t = F_x(w(x,t)) = \int_x^{w(x,t)} (\gamma(u))^{-1}\,\mathrm{d}u$, exists, lies in the interval $[x,\infty)$ and is unique for every $t \geq 0$. The function F_x is differentiable and $F_x'(y) = (\gamma(y))^{-1} > 0$, $y \in [x,\infty)$, and, as a consequence, the inverse $F_x^{-1} : [0,\infty) \to [x,\infty)$ exists, is differentiable and $(F_x^{-1})'(t) = \gamma(F_x^{-1}(t))$, $t \geq 0$. Clearly, $w(x,t) = F_x^{-1}(t)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t} w(x,t) \;=\; \gamma(w(x,t)), \qquad t \geq 0, x > 1.$$

The formula for $\frac{\mathrm{d}}{\mathrm{d}x} w(x,t)$ follows from differentiation of both sides of the integral equation in (III.19) with respect to x. □

The proof of Proposition III.2.9 could be copied from the respective one for the block counting process, the proof given instead uses the fact that $v(.,t)$ and $w(.,t)$ are inverse.

Proof. (of Proposition III.2.9) We first prove (ii), and then (i) and (iii). Note that the situation of (i) is a special case of (ii) with $\kappa = 0$. Fix $t \geq 0$. According to Proposition III.2.5 there exists a slowly varying function $L_t : [1,\infty) \to (0,\infty)$ such that $v(x,t) = x^{e^{-\kappa t}} L_t(x)$, $x \geq 1$. As the function $w(.,t)$ is the inverse of $v(.,t)$, it is regularly varying with index $e^{\kappa t}$. More precisely, it follows from [4, Proposition 1.5.15], applied with $f(x) := v(x,t)$, $a := e^{-\kappa t}$, $b := 1$ and $l(x) := L_t(x)$, $x \geq 1$, that $w(x,t) \sim x^{e^{\kappa t}} L_t^{\#,0}(x^{e^{\kappa t}})$ as $x \to \infty$, where $L_t^{\#,0}$ is the de Bruijn conjugate of the slowly varying function $x \mapsto (L_t(x))^{e^{\kappa t}}$, $x \geq 1$, i.e., a slowly varying function satisfying $\lim_{x\to\infty} L_t^{\#,0}(x(L_t(x))^{e^{\kappa t}})(L_t(x))^{e^{\kappa t}} = 1$. See, e.g., [4, Theorem 1.5.13] for a definition of the de Bruijn conjugate of slowly varying functions. The function $L_t^{\#}$, defined via $w(x,t) = x^{e^{\kappa t}} L_t^{\#}(x)$, $x \geq 1$, is asymptotically equal to the slowly varying function $L_t^{\#,0}(x^{e^{\kappa t}})$, thus slowly varying itself, which completes the proof of (ii).

(i) Assume that $\Xi(\{0\}) = 0$ and $\mu = \int_\Delta |u|\,\nu(\mathrm{d}u) < \infty$, and recall that $\kappa = 0$. Proposition III.2.5 states that $\lim_{x\to\infty} L_t(x) = e^{-\mu t}$. We can thus choose $L_t^{\#,0}(x) := e^{\mu t}$, $x \geq 1$. From $L_t^{\#}(x) \sim L_t^{\#,0}(x) = e^{\mu t}$ it follows that $w(x,t) \sim x e^{\mu t}$ as $x \to \infty$.

(iii) As seen in the proof of Proposition III.2.6, there exists a slowly varying function $L_1 : (1,\infty) \to (0,\infty)$ such that $\gamma_1(x) = \kappa x \log x + x \log L_1(x)$, $x > 1$. The function γ_1 is continuous and positive on $(1,\infty)$ and $\int_2^\infty (\gamma_1(u))^{-1}\,\mathrm{d}u = \infty$. The proof of Proposition III.2.8 shows that the scaling $w_1(x,t)$, defined by (III.19) with γ_1 in place of γ, exists for $x \geq 1$ and $t \geq 0$. Fix $t \geq 0$. According to Proposition III.2.6 the scaling $v_1(x,t)$, defined by the integral equation in (III.9) with γ_1 in place of γ, exists for $x \geq x_0(t)$, where $x_0(t) > 1$. The proof of Proposition III.2.5 shows the existence of a slowly varying function $L_{t,1}$ such that $v_1(x,t) = x^{e^{-\kappa t}} L_{t,1}(x)$ for $x \geq x_0(t)$. Here we used that the map $x \mapsto \gamma_1(x)/x$, $x > 1$, and, hence, the function L_1 are non-decreasing if $\kappa = 0$. The scalings $v_1(.,t)$ and $w_1(.,t)$ are obviously inverse (on suitable domains). From Part (ii) it follows that there exists a slowly varying function $L_{t,1}^{\#}$ such that $w_1(x,t) = x^{e^{\kappa t}} L_{t,1}^{\#}(x)$. From $v(x,t) \sim v_1(x,t)$ it follows that $L_t(x) \sim L_{t,1}(x)$ and $x = w(v(x,t)) \sim w(v_1(x,t)) = x(L_t(x))^{e^{\kappa t}} L_t^{\#}(v_1(x,t))$ as $x \to \infty$ and $x = w_1(v_1(x,t),t) = x(L_{t,1}(x))^{e^{\kappa t}} L_{t,1}^{\#}(v_1(x,t))$ for $x \geq x_0(t)$. Hence, $L_{t,1}^{\#}(v_2(x,t)) \sim L_{t,2}^{\#}(v_2(x,t))$, consequently $L_{t,1}^{\#}(x) \sim L_{t,2}^{\#}(x)$ and we finally have $w_1(x,t) \sim w_2(x,t)$ as $x \to \infty$. □

We proceed to prove the convergence of the scaled fixation line. The involved state spaces and semigroups are denoted by the same symbols as for the block counting process.

Proof. (of Theorem III.2.10) Define $Y_t^{(n)} := \log L_t^{(n)} - \log w(n,t)$ for $n \in \mathbb{N}$ and $t \geq 0$. We start by proving the convergence of the one-dimensional distributions. Fix $t \geq 0$, $x \in \mathbb{R}$ and write $k := \lceil e^x w(n,t) \rceil \in \mathbb{N}$. Note that $Y_t \stackrel{\mathrm{d}}{=} -e^{\kappa t} S_t$. By duality (Eq. (III.2)),

$$\begin{aligned} \mathbb{P}(Y_t^{(n)} \geq x) &= \mathbb{P}(L_t^{(n)} \geq k) = \mathbb{P}(N_t^{(k)} \leq n) \\ &= \mathbb{P}\big(\log N_t^{(k)} - \log v(k,t) \leq \log(n/v(k,t))\big). \end{aligned}$$

By Proposition III.2.5, the function $v(.,t)$ varies regularly with index $e^{-\kappa t}$. From $\lim_{n\to\infty} w(n,t) = \infty$ it hence follows that $n/v(k,t) = v(w(n,t),t)/v(\lceil e^x w(n,t) \rceil, t) \to e^{-xe^{-\kappa t}}$ as $n \to \infty$. Theorem III.2.7 implies that

$$\lim_{n\to\infty} \mathbb{P}(Y_t^{(n)} \geq x) = \mathbb{P}(S_t \leq -xe^{-\kappa t}) = \mathbb{P}(Y_t \geq x) \qquad \text{(III.39)}$$

for $-xe^{-\kappa t}$ in the set C_{S_t} of continuity points of S_t. From (III.39) we obtain the weak convergence of $Y_t^{(n)}$ to Y_t as $n \to \infty$ for each $t \geq 0$, since $-xe^{-\kappa t} \in C_{S_t}$ if and only if $x \in C_{Y_t}$.

The time-space processes $\widetilde{Y}^{(n)} := (t, Y_t^{(n)})_{t\geq 0}$, $n \in \mathbb{N}$, and $\widetilde{Y} := (t, X_t)_{t\geq 0}$ are time-homogeneous Markov processes with state spaces $\widetilde{E}_n = \{(s,x) : s \geq 0, e^x w(n,s) \in \{n, n+1, \ldots\}\}$ and $\widetilde{E} = [0,\infty) \times \mathbb{R}$. Set $k := k(s,x,n) := e^x w(n,s) \in \{n, n+1, \ldots\}$ for $(s,x) \in \widetilde{E}_n$ and $n \in \mathbb{N}$. The semigroups $(\widetilde{T}_t^{(n)})_{t\geq 0}$ and $(\widetilde{T}_t)_{t\geq 0}$ of $\widetilde{Y}^{(n)}$ and $\widetilde{Y}$ are given by

$$\begin{aligned}\widetilde{T}_t^{(n)} f(s,x) &:= \mathbb{E}(f(s+t, Y_{s+t}^{(n)}) | Y_s^{(n)} = x)\\ &= \mathbb{E}(f(s+t, \log L_t^{(k)} - \log w(n, s+t))\\ &= \mathbb{E}(f(s+t, \log(w(k,t)/w(n,s+t)) + Y_t^{(k)})),\end{aligned}$$

$(s,y) \in \widetilde{E}_n$, and

$$\widetilde{T}_t f(s,x) := \mathbb{E}(f(s+t, Y_{s+t}) | Y_s = x) = \mathbb{E}(f(s+t, e^{\kappa t}x + Y_t)),$$

$(s,x) \in \widetilde{E}$, for $f \in B(\widetilde{E}), t \geq 0$ and $n \in \mathbb{N}$. Fix $t > 0$ and first let $f \in B(\widetilde{E})$ be of the form $f(s,x) = g(s)h(x)$, $(s,x) \in \widetilde{E}$, where $g \in B([0,\infty))$ and $h \in C_c(\mathbb{R})$. Clearly, $\widetilde{T}_t^{(n)} f(s,x) = g(s+t)\mathbb{E}(h(\log(w(k,t)/w(n,s+t)) + Y_t^{(k)}))$, $(s,x) \in \widetilde{E}_n, n \in \mathbb{N}$, and $\widetilde{T}_t f(s,x) = g(s+t)\mathbb{E}(h(e^{\kappa t}x + Y_t))$, $(s,x) \in \widetilde{E}$. If we are able to show that

$$\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n} \big|\mathbb{E}\big(h(\log(w(k,t)/w(n,s+t)) + Y_t^{(k)})\big) - \mathbb{E}(h(e^{bt}x + Y_t))\big| = 0, \tag{III.40}$$

then

$$\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n} |\widetilde{T}_t^{(n)} f(s,x) - \widetilde{T}_t f(s,x)| = 0. \tag{III.41}$$

The algebra of functions $f \in B(\widetilde{E})$ of the form $f(s,x) = \sum_{i=1}^{l} g_i(s)h_i(x)$, $(s,x) \in \widetilde{E}$, where $l \in \mathbb{N}, g_i \in B([0,\infty))$ and $h_i \in C_c(\mathbb{R})$, separates points and vanishes nowhere. According to the Stone–Weierstrass theorem for locally compact spaces (see e.g. [8]) it is a dense subset of $B(\widetilde{E})$ such that (III.41) holds for $f \in B(\widetilde{E})$. [11, 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 $n \to \infty$. It remains to verify (III.40).

From

$$s+t = \int_x^{w(x,s)} \frac{\mathrm{d}u}{\gamma(u)} + \int_{w(x,s)}^{w(w(x,s),t)} \frac{\mathrm{d}u}{\gamma(u)} = \int_x^{w(w(x,s),t)} \frac{\mathrm{d}u}{\gamma(u)}$$

it follows that $w(x,s+t) = w(w(x,s),t)$ for $(s,x) \in \widetilde{E}$. By Proposition III.2.8, there exists a slowly varying function $L_t^{\#} : [1,\infty) \to (0,\infty)$ such that $w(x,t) = x^{e^{\kappa t}} L_t^{\#}(x)$ for $x \geq 1$. Applying Proposition III.5.4 to the right-hand side of

$$\frac{w(k,t)}{w(n,s+t)} = \frac{w(e^x w(n,s),t)}{w(w(n,s),t)} = e^{e^{\kappa t}x} \frac{L_t^{\#}(e^x w(n,s))}{L_t^{\#}(w(n,s))}, \qquad (s,x) \in \widetilde{E}_n, n \in \mathbb{N},$$

provides

$$\lim_{x\to\infty} \inf_{s:(s,x)\in\widetilde{E}_n, n\in\mathbb{N}} \log \frac{w(k,t)}{w(n,s+t)} \; = \; \infty$$

and

$$\lim_{x\to-\infty} \sup_{s:(s,x)\in\widetilde{E}_n, n\in\mathbb{N}} \log \frac{w(k,t)}{w(n,s+t)} \; = \; -\infty.$$

The family $\{Y_t^{(k)} : k \in \mathbb{N}\}$ is tight due to the convergence $Y_t^{(k)} \to Y_t$ in distribution as $k \to \infty$ and Prokhorov's theorem. By dominated convergence and since h has compact support,

$$\begin{aligned} &\lim_{x\to\infty} \sup_{s:(s,x)\in\widetilde{E}_n, n\in\mathbb{N}} \left|\mathbb{E}\big(h(\log(w(k,t)/w(n,s+t)) + Y_t^{(k)})\big)\right| \qquad \text{(III.42)} \\ &\quad = \lim_{x\to-\infty} \sup_{s:(s,x)\in\widetilde{E}_n, n\in\mathbb{N}} \left|\mathbb{E}\big(h(\log(w(k,t)/w(n,s+t)) + Y_t^{(k)})\big)\right| = 0, \end{aligned}$$

such as

$$\lim_{x\to\infty} \mathbb{E}(h(e^{\kappa t}x + Y_t)) \; = \; \lim_{x\to-\infty} \mathbb{E}(h(e^{\kappa t}x + Y_t)) \; = \; 0. \qquad \text{(III.43)}$$

For any compact interval $K \subset \mathbb{R}$ we have that, by the uniform convergence theorem for slowly varying functions [4, Theorem 1.5.2],

$$\begin{aligned} &\lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, x\in K} \left| \log \frac{w(k,t)}{w(n,s+t)} - e^{\kappa t}x \right| \\ &\quad = \lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, x\in K} \left| \log \frac{L_t^{\#}(e^x w(n,s))}{L_t^{\#}(w(n,s))} \right| \; = \; 0. \end{aligned}$$

The function h is uniformly continuous. Note that $\lim_{n\to\infty} \inf_{(s,x)\in\widetilde{E}_n, x\in K} k(s,x,n) = \infty$. From the convergence $Y_t^{(k)} \to Y_t$ in distribution as $k \to \infty$ it hence follows that

$$\begin{aligned} \lim_{n\to\infty} \sup_{(s,x)\in\widetilde{E}_n, x\in K} &\Big|\mathbb{E}\big(h(\log(w(k,t)/w(n,s+t)) + Y_t^{(k)})\big) \\ &- \; \mathbb{E}(h(e^{\kappa t}x + Y_t))\Big| \; = \; 0. \qquad \text{(III.44)} \end{aligned}$$

Finally, Eqs. (III.42), (III.43) and (III.44) imply (III.40). The proof is complete. □

III.5 Appendix

We collect some fundamental results concerning the model described in Section III.2.3 involving an infinite number of urns. Let $u \in \Delta$. Recall that $X_i(n,u)$ denotes the number of balls in urn $J_i \in \mathbb{N}_0$ after

n balls have been allocated. Let $K(n,u) := \sum_{i\geq 1} 1_{\{X_i(n,u)>0\}}$ denote the number of occupied urns (disregarding urn J_0). The following law of large numbers result holds.

Lemma III.5.1. *For all $u \in \Delta$, $K(n,u)/\mathbb{E}(K(n,u)) \to 1$ almost surely as $n \to \infty$.*

Proof. We proceed as in the proof of [15, Theorem 1]. Fix $u \in \Delta$ and write $K(n) := K(n,u)$ for convenience. Define $\Phi : [0,\infty) \to [0,\infty)$ via $\Phi(x) := \sum_{i\geq 1}(1-(1-u_i)^x)$, $x \geq 0$. Note that $\Phi(0) = 0$, $\Phi(1) = |u| \leq 1$ and $\mathbb{E}(K(n)) = \Phi(n)$, $n \in \mathbb{N}$. The function Φ is non-decreasing, concave and differentiable on $(0,\infty)$ with derivative $\Phi'(x) = \sum_{i\geq 1}(1-u_i)^x(-\log(1-u_i))$. In particular, for all $x \geq 1$, $\Phi'(x) \leq \Phi'(1) = \sum_{i\geq 1}(1-u_i)(-\log(1-u_i)) \leq \sum_{i\geq 1} u_i \leq 1$. Thus, for each $m \in \mathbb{N}$, there exists $n_m \in \mathbb{N}$ such that $m^2 \leq \Phi(n_m) \leq m^2 + 1$. Tschebyscheff's inequality together with $\mathrm{Var}(K(n)) \leq \Phi(2n) - \Phi(n) \leq \Phi(n)$ yields

$$\mathbb{P}\left(\left|\frac{K(n_m)}{\Phi(n_m)} - 1\right| \geq \varepsilon\right) \leq \frac{\mathrm{Var}(K(n_m))}{\varepsilon^2(\Phi(n_m))^2} \leq \frac{1}{\varepsilon^2\Phi(n_m)} \leq \frac{1}{\varepsilon^2 m^2}$$

for all $m \in \mathbb{N}$ and $\varepsilon > 0$. Thus, $\sum_{m\geq 1}\mathbb{P}(|K(n_m)/\Phi(n_m) - 1| \geq \varepsilon) < \infty$ for all $\varepsilon > 0$. By the Borel–Cantelli lemma it follows that $K(n_m)/\Phi(n_m) \to 1$ almost surely as $m \to \infty$.

For $n \in \mathbb{N}$ with $n_m \leq n \leq n_{m+1}$ the monotonicity inequalities $K(n_m) \leq K(n) \leq K(n_{m+1})$ and $\Phi(n_m) \leq \Phi(n) \leq \Phi(n_{m+1})$ hold, which allows to sandwich the fraction $K(n)/\Phi(n)$ via

$$\frac{K(n_m)}{\Phi(n_{m+1})} \leq \frac{K(n)}{\Phi(n)} \leq \frac{K(n_{m+1})}{\Phi(n_m)},$$

where both sides converge to 1 almost surely, since $\Phi(n_m)/\Phi(n_{m+1}) \to 1$. $\square$

The following two results deal with the random variables $Y(n,u) := X_0(n,u) + K(n,u)$ defined in (III.12). Lemma III.5.2 concerns the limiting behavior of $Y(n,u)/n$ as $n \to \infty$.

Lemma III.5.2. *For all $u = (u_1, u_2, \ldots) \in \Delta$, $Y(n,u)/n \to u_0$ almost surely as $n \to \infty$, where $u_0 := 1 - |u| := 1 - \sum_{i\geq 1} u_i$.*

Proof. Fix $u \in \Delta$. We have $Y(n,u) = X_0(n,u) + K(n,u)$, $n \in \mathbb{N}$. Clearly, $X_0(n,u)/n \to u_0$ almost surely as $n \to \infty$, since $X_0(n,u)$ has a binomial distribution with parameters n and u_0. By Lemma

III.5.1, $K(n,u)/\mathbb{E}(K(n,u)) \to 1$ almost surely as $n \to \infty$. Moreover,

$$\frac{\mathbb{E}(K(n,u))}{n} \;=\; \sum_{i\geq 1} \frac{1-(1-u_i)^n}{n} \;\to\; 0$$

as $n \to \infty$ by dominated convergence, since $(1-(1-u_i)^n)/n \leq 1/n \to 0$ and $(1-(1-u_i)^n)/n \leq u_i$, where the dominating map $i \mapsto u_i$ is integrable with respect to the counting measure on $\mathbb{N}$. Thus,

$$\frac{K(n,u)}{n} \;=\; \frac{K(n,u)}{\mathbb{E}(K(n,u))}\frac{\mathbb{E}(K(n,u))}{n} \;\to\; 1\cdot 0 \;=\; 0$$

almost surely as $n \to \infty$. Therefore, $Y(n,u)/n \to u_0$ almost surely as $n \to \infty$. □

The following result (Lemma III.5.3) is used in the proof of the main convergence theorem (Theorem III.2.7). It presents bounds for particular moments of the random variable $Y(n,u)$ defined in (III.12). Lemma 18 of [21] provides similar bounds.

Lemma III.5.3. *There exist constants $D_1, D_2 \in \mathbb{R}$ such that, for all $u \in \Delta$,*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left(\left(\frac{Y(n,u)}{n} - (1-|u|)\right)^2\right) \;\leq\; D_1|u|^2$$

and

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left(\left(\frac{Y(n,u)}{n} - 1\right)^2\right) \;\leq\; D_2|u|^2.$$

Proof. Fix $n \in \mathbb{N}$ and $u \in \Delta$. Define $u_0 := 1-|u|$. We omit the parameter (n,u) and write (III.12) as $Y = X_0 + K$, where $K := K(n,u) := \sum_{i\geq 1} 1_{\{X_i(n,u)>0\}}$ denotes the number of occupied urns (disregarding urn J_0). Furthermore, define $\widetilde{Y} := Y/n - u_0 = X_0/n - u_0 + K/n$. Calculations that are similar to the following (but come from a different motivation) are carried out in the proof of [25, Lemma 6.1]. The formulas for $\mathbb{E}(K)$ and $\mathbb{E}(X_0K)$ can be found there. We have

$$\widetilde{Y}^2 = \left(\frac{X_0}{n} - u_0 + \frac{K}{n}\right)^2 = \left(\frac{X_0}{n} - u_0\right)^2 + \frac{2X_0K}{n^2} - \frac{2u_0K}{n} + \frac{K^2}{n^2}.$$

Recall that X_0 has a binomial distribution with parameters n and u_0. In particular, $\mathbb{E}(X_0) = nu_0$ and $\mathbb{E}((X_0/n - u_0)^2) = n^{-2}\mathrm{Var}(X_0) =$

$|u|(1-|u|)/n$. Together with $K^2 = K + \sum_{i\neq j} 1_{\{X_iX_j>0\}}$ it follows that

$$\begin{aligned}\mathbb{E}(\widetilde{Y}^2) \;&=\; \frac{u_0(1-u_0)}{n} \;+\; \frac{2\mathbb{E}(X_0K)}{n^2}\\ &\quad-\; \frac{2u_0\mathbb{E}(K)}{n} \;+\; \frac{\mathbb{E}(K)}{n^2} \;+\; \frac{1}{n^2}\sum_{i\neq j}\mathbb{P}(X_iX_j>0).\end{aligned}$$

Adding and subtracting $u_0(1-u_0)/n = |u|(1-|u|)/n$ leads to

$$\begin{aligned}\mathbb{E}(\widetilde{Y}^2) \;&=\; 2\bigg(\frac{\mathbb{E}(X_0K)}{n^2} - \frac{u_0\mathbb{E}(K)}{n} + \frac{u_0(1-u_0)}{n}\bigg)\\ &\quad+\; \frac{1}{n}\bigg(\frac{\mathbb{E}(K)}{n} - |u|\bigg) \;+\; \frac{|u|^2}{n} \;+\; \frac{1}{n^2}\sum_{i\neq j}\mathbb{P}(X_iX_j>0).\end{aligned}$$

We have $\mathbb{E}(K) = \sum_{i\geq 1}(1-(1-u_i)^n)$ and $\mathbb{E}(X_0K) = nu_0\sum_{i\geq 1}(1-(1-u_i)^{n-1})$. Moreover, by Bernoulli's inequality, $1-(1-u_i)^{n-1} \leq (n-1)u_i$ for $i\in\mathbb{N}$. We conclude that

$$\begin{aligned}\frac{\mathbb{E}(X_0K)}{n^2} - \frac{u_0\mathbb{E}(K)}{n} + \frac{u_0(1-u_0)}{n} \;&=\; \frac{u_0}{n}\sum_{i\geq 1}u_i(1-(1-u_i)^{n-1})\\ &\leq\; \frac{(n-1)u_0}{n}(u,u) \;\leq\; |u|^2.\end{aligned}$$

Also, $n^{-1}\mathbb{E}(K) - |u| = n^{-1}\sum_{i\geq 1}(1-(1-u_i)^n - nu_i) \leq 0$. From the generalized Bernoulli inequality $1-(1-u_i)^n-(1-u_j)^n+(1-u_i-u_j)^n \leq n(n-1)u_iu_j$, $i,j\in\mathbb{N}$, it follows that

$$\begin{aligned}\frac{1}{n^2}\sum_{i\neq j}\mathbb{P}(X_iX_j>0) \;&=\; \frac{1}{n^2}\sum_{i\neq j}(1-(1-u_i)^n\\ &\qquad -\; (1-u_j)^n \;+\; (1-u_i-u_j)^n)\\ &\leq\; \sum_{i,j\geq 1}u_iu_j \;=\; |u|^2.\end{aligned}$$

Collecting all bounds yields that $\mathbb{E}(\widetilde{Y}^2)$ is bounded by $4|u|^2$, which shows that the first claim holds with $D_1 := 4$. Concerning the second claim, note that

$$\begin{aligned}0 \;\leq\; \mathbb{E}((Y/n-1)^2) \;&=\; \mathbb{E}((\widetilde{Y}-|u|)^2)\\ &=\; \mathbb{E}(\widetilde{Y}^2) \;-\; 2|u|\mathbb{E}(\widetilde{Y}) \;+\; |u|^2 \;\leq\; \mathbb{E}(\widetilde{Y}^2) \;+\; |u|^2,\end{aligned}$$

since $\mathbb{E}(\widetilde{Y}) = n^{-1}\mathbb{E}(K) \geq 0$, showing that we can choose $D_2 := D_1 + 1 = 5$. □

The following result is needed in the proof of Theorem III.2.10.

Proposition III.5.4. *Let $\alpha > 0$ and the function $L : [1,\infty) \to (0,\infty)$ be slowly varying with $0 < \inf_{y\in[1,K]} L(y) \le \sup_{y\in[1,K]} L(y) < \infty$ for any $K > 1$. Then $\lim_{x\to\infty} \inf_{y\ge x^{-1}\vee 1} x^\alpha L(xy)/L(y) = \infty$ and $\lim_{x\to 0} \sup_{y\ge x^{-1}\vee 1} x^\alpha L(xy)/L(y) = 0$.*

Proof. By the representation theorem for slowly varying functions [4, Theorem 1.3.1], there exist functions $\delta : [1,\infty) \to (0,\infty)$ and $\varepsilon : [1,\infty) \to \mathbb{R}$ with $\lim_{x\to\infty} \delta(x) =: d \in (0,\infty)$ and $\lim_{x\to\infty} \varepsilon(x) = 0$ such that $L(x) = \delta(x) \exp(\int_1^x \frac{\varepsilon(u)}{u}\,\mathrm{d}u)$, $x \ge 1$. Furthermore we can choose ε such that $\|\varepsilon\| \le \alpha/2$, if δ is adapted accordingly. By the additional boundary assumption for L and the convergence of δ to $d \in (0,\infty)$, $0 < \inf_{y\ge 1} \delta(y) \le \sup_{y\ge 1} \delta(y) < \infty$. Thus, $0 < \inf_{y\ge x^{-1}\vee 1} \delta(xy)/\delta(y) \le \sup_{y\ge x^{-1}\vee 1} \delta(xy)/\delta(y) < \infty$. From

$$\begin{aligned} x^\alpha \frac{L(xy)}{L(y)} &= \frac{\delta(xy)}{\delta(y)} \exp\left(\int_1^x \frac{\alpha}{u}\,\mathrm{d}u + \int_y^{xy} \frac{\varepsilon(u)}{u}\,\mathrm{d}u \right) \\ &= \frac{\delta(xy)}{\delta(y)} \exp\left(\int_1^x \frac{\alpha + \varepsilon(uy)}{u}\,\mathrm{d}u \right) \end{aligned}$$

it follows that

$$\begin{aligned} \inf_{y\ge x^{-1}\vee 1} x^\alpha \frac{L(xy)}{L(y)} &\ge \inf_{y\ge x^{-1}\vee 1} \frac{\delta(xy)}{\delta(y)} \exp\left(\int_1^x \frac{\alpha}{2u}\,\mathrm{d}u \right) \\ &= x^{\alpha/2} \inf_{y\ge x^{-1}\vee 1} \frac{\delta(xy)}{\delta(y)} \to \infty \end{aligned}$$

as $x \to \infty$ and

$$\sup_{y\ge x^{-1}\vee 1} x^\alpha \frac{L(xy)}{L(y)} \le x^{\alpha/2} \sup_{y\ge x^{-1}\vee 1} \frac{\delta(xy)}{\delta(y)} \to 0$$

as $x \to 0+$. □

Remark. The function $L_t^\# : [1,\infty) \to (0,\infty)$, defined via $L_t^\#(x) := w(x,t)/x^{e^{\kappa t}}$, $x \ge 1$, is slowly varying. Due to $w(x,t) \ge x$, it holds that $L_t^\#(x) \ge x^{1-e^{\kappa t}}$ for $x \ge 1$ and $t \ge 0$, and since $L_t^\#$ is continuous, Proposition III.5.4 applies.

Bibliography

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

[2] Berestycki, J., Berestycki, N. and Limic, V. (2010) The Λ-coalescent speed of coming down from infinity. *Ann. Probab.* **38**, no. 1, 207–233.

[3] Berestycki, J., Berestycki, N. and Limic, V. (2014) A small-time coupling between Λ-coalescents and branching processes. *Ann. Appl. Probab.* **24**, no. 2, 449–475.

[4] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1989) *Regular Variation.* Cambridge University Press.

[5] Bogachev, V. I., Röckner, M. and Schmuland, B. (1996) Generalized Mehler semigroups and applications. *Probab. Theory Related Fields* **105**, no. 2, 193–225.

[6] Cannings, C. (1974) The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Adv. Appl. Probab.* **6**, no. 2, 260–290.

[7] Cannings, C. (1975) The latent roots of certain Markov chains arising in genetics: a new approach, II. Further haploid models. *Adv. Appl. Probab.* **7**, no. 2, 264–282.

[8] de Branges, L. (1959) The Stone–Weierstrass theorem. *Proc. Amer. Math. Soc.* **10**, no. 5, 822–824.

[9] Diehl, C. and Kersting, G. (2019) Tree lengths for general Λ-coalescents and the asymptotic site frequency spectrum around the Bolthausen–Sznitman coalescent. *Ann. Appl. Probab.* **29**, no. 5, 2700–2743.

[10] Diehl, C. and Kersting, G. (2019) External branch lengths of Λ-coalescents without a dust component. *Electron. J. Probab.* **24**, no. 134, 36pp.

[11] Ethier, S. N. and Kurtz, T. G. (1986) *Markov Processes, Characterization and Convergence.* Wiley.

[12] Foucart, C., Ma, C. and Mallein, B. (2019) Coalescences in continuous-state branching processes. *Electron. J. Probab.* **24**, no. 103, 52 pp.

[13] Freund, F. and Möhle, M. (2009) On the number of allelic types for samples taken from exchangeable coalescents with mutation. *Adv. in Appl. Probab.* **41**, no. 4, 1082–1101.

[14] Gaiser, F. and Möhle, M. (2016) On the block counting process and the fixation line of exchangeable coalescents. *ALEA Lat. Am. J. Probab. Math. Stat.* **13**, no. 2, 809–833.

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

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

[17] Grey, D. R. (1974) Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probab.* **11**, no. 4, 669–677.

[18] Hénard, O. (2015) The fixation line in the Λ-coalescent. *Ann. Appl. Probab.* **25**, no. 5, 3007–3032.

[19] Herriger, P. and Möhle, M. (2012) Conditions for exchangeable coalescents to come down from infinity. *ALEA Lat. Am. J. Probab. Math. Stat.* **9**, no. 2, 637–665.

[20] Lamperti, J. (1958) An occupation time theorem for a class of stochastic processes. *Trans. Amer. Math. Soc.* **88**, 380–387.

[21] Limic, V. (2010) On the speed of coming down from infinity for Ξ-coalescent processes. *Electron. J. Probab.* **15**, no. 8, 217–240.

[22] Limic, V. and Talarczyk, A. (2015) Second order asymptotics for the block counting process in a class of regularly varying Λ-coalescents. *Ann. Probab.* **43**, no. 3, 1419–1455.

[23] Möhle, M. (2010) Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson–Dirichlet coalescent. *Stochastic Process. Appl.* **120**, no. 11, 2159–2173.

[24] Möhle, M. (2015) 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**, no. 1, 35–53.

[25] Möhle, M. (2021) The rate of convergence of the block counting process of exchangeable coalescents with dust. *ALEA Lat. Am. J. Probab. Math. Stat.* **18**, no. 2, 1195–1220.

[26] Möhle, M. and Sagitov, S. (2001) A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, no. 4, 1547–1462.

[27] Möhle, M. and Vetter, B. (2022) Scaling limits for the block counting process and the fixation line for a class of Λ-coalescents. *ALEA Lat. Am. J. Probab. Math. Stat.* **19**, no. 1, 641–664.

[28] Parthasarathy, K. R. (1967) *Probability Measures on Metric Spaces.* Academic Press. New York.

[29] Ranga Rao, R. (1962) Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33**, no. 2, 659–680.

[30] Sato, K. and Yamazato, M. (1984) Operator-selfdecomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type. *Stochastic Process. Appl.* **17**, no. 1, 73–100.

[31] Schweinsberg, J. (2000) Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5**, no. 12, 1–50.

[32] Siegmund, D. (1976) The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probab.* **4**, no. 6, 914–924.

[33] Silverstein, M. L. (1968) A new approach to local times. *J. Math. Mech.* **17**, no. 11, 1023–1054.

[34] Silverstein, M. L. (1969) Continuous state branching semigroups. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **14**, 96–112.