A small-time coupling between Λ-coalescents and branching processes

Julien Berestycki[1], Nathanaël Berestycki[2], and Vlada Limic[3]

Draft version — January 20, 2012

Abstract

We describe a new general connection between Λ-coalescents and genealogies of continuous-state branching processes. This connection is based on the construction of an explicit coupling using a particle representation inspired by the lookdown process of Donnelly and Kurtz. This coupling has the property that the coalescent comes down from infinity if and only if the branching process becomes extinct, thereby answering a question of Bertoin and Le Gall. The coupling also offers new perspective on the speed of coming down from infinity, and allows us to relate power-law behavior for $N^{\Lambda}(t)$ to the classical upper and lower indices arising in the study of pathwise properties of Lévy processes.

AMS 2000 Subject Classification. 60J25, 60F99, 92D25

Key words and phrases. Λ-coalescents, continuous-state branching processes, particle system representation, Lévy processes, lookdown construction, Fleming-Viot processes

1. Université Pierre et Marie Curie - Paris VI. Supported in part by ANR A3 and ANR MADCOF.
2. University of Cambridge. Supported in part by EPSRC grant EP/GO55068/1 and EP/I03372X/1.
3. Université de Provence. Research supported in part by NSERC Discovery Grant, by Alfred P. Sloan Research Fellowship, and by ANR MAEV research grant.

1 Introduction and main results

Coalescents with multiple collisions, also known as Λ-*coalescents* are Markovian models of coagulation. Introduced and first studied independently by Pitman [24] and by Sagitov [27] (also considered in a contemporaneous work of Donnelly and Kurtz [14]), these processes have been intensely researched in the last decade. The research is mostly motivated by the fact that Λ-coalescents arise naturally as scaling limits for the genealogy in certain exchangeable population dynamics models. We refer to [6, 5] for an introduction and a survey of the relevant literature.

The *standard* Λ-coalescent starts with infinitely many microscopic particles that coalesce into larger clusters as time runs. Our interest in this paper concerns the small-time behavior of (standard) Λ-coalescents, in particular the phenomenon of *coming down from infinity* (a precise definition will be given below). Our main goal is to answer a question which arose from work of Bertoin and Le Gall [10]. They observed that the Schweinsberg condition [28] for coming down from infinity for Λ-coalescents is equivalent to the condition for extinction of related continuous-state branching processes (CSBPs), and asked if a deeper connection exists between these two classes of processes. In this paper, we construct an explicit coupling between a given Λ-coalescent and a certain associated CSBP, and therefore answer the above question of Bertoin and Le Gall.

This coupling makes use of a particle system representation based on a lookdown process in the spirit of Donnelly and Kurtz [13, 14]. Apart from its interest from a purely theoretical point of view, our coupling gives a new understanding of the asymptotic form of the "speed of coming down from infinity" (as discussed by the authors in [2]), and leads to precise quantitative results for the corresponding Λ-coalescent observed at small times. In particular, the power-law exponents for the number of blocks in a particular Λ-coalescent are shown to coincide with the classical notion of upper and lower indices of the Lévy measure of the associated CSBP.

The methodology in this paper has several points in common with [3, 4], where an analogous link between Beta-coalescents and α-stable continuous-state branching processes was used. However, in these papers the central tool was an explicit embedding of the lookdown process into the (stable) Continuous Random Tree, which allowed for many explicit computations. Here, we show that the correct way to generalize this picture for an arbitrary Λ-coalescents is directly via the particle system approach of the lookdown process.

In the rest of the paper, we denote by $\stackrel{d}{=}$ the equivalence in distribution. We also use the standard Bachmann-Landau notation $\sim, O(\cdot), o(\cdot), \asymp$ for comparing asymptotic behavior of deterministic and stochastic functions and sequences.

1.1 Coalescents and CSBPs

Let Λ be an arbitrary finite measure on $[0,1]$, and let $(\Pi_t, t \geq 0)$ denote the associated Λ-coalescent. The Markov jump process $(\Pi_t, t \geq 0)$ takes values in the set of partitions of $\{1, 2, \ldots\}$. Its law is specified by the requirement that, for any $n \in \mathbb{N}$, the restriction Π^n of Π to $\{1, \ldots, n\}$ is a continuous-time Markov chain with transition rates given as follows: whenever Π^n has $b \in [2, n]$ blocks, any given k-tuple of blocks coalesces at rate $\lambda_{b,k} := \int_{(0,1]} r^{k-2}(1-r)^{b-k}\Lambda(dr)$.

We will always assume that $\Pi(0)$ is the trivial partition $\{\{i\} : i \in \mathbb{N}\}$. Let us call $N^\Lambda(t)$ the number of blocks of $\Pi(t)$ the coalescent at time t. The first question one may ask about these processes is whether the number of blocks ever becomes finite. In his seminal paper [24] Pitman noted that (provided $\Lambda(\{1\}) = 0$) as a consequence of the strong Markov property the following striking dichotomy holds: either $\mathbb{P}(N^\Lambda(t) = \infty, \forall t \geq 0) = 1$ or $\mathbb{P}(N^\Lambda(t) < \infty, \forall t > 0) = 1$. In the latter case the coalescent is said to *come down from infinity.* Finding a necessary and sufficient condition for this phenomenon was naturally one of the first problems to be studied. As part of his thesis work, Schweinsberg [28] derived the following criterion: the Λ-coalescent comes down from infinity if and only if

$$\sum_{b=2}^{\infty} \left(\sum_{k=2}^{b} (k-1) \binom{b}{k} \lambda_{b,k} \right)^{-1} < \infty. \tag{1}$$

Over the subsequent years, a series of remarkable links were discovered between Λ-coalescents and continuous-state branching processes (CSBP), for some special cases of Λ. The case of Kingman's coalescent ($\Lambda = \delta_0$) was analyzed by Perkins [23] in 1991, though he used a somewhat different language. Bertoin and Le Gall [7] did the case of the Bolthausen-Sznitman coalescent (where $\Lambda(dx) = dx$ is the uniform measure on [0,1]), and then Birkner et al. [11] did all the Beta-coalescents cases (where Λ is the Beta$(2-\alpha, \alpha)$ distribution, and $\alpha \in (0, 2)$).

While seeking a way to understand the above results as special cases of a general theorem, Bertoin and Le Gall [10] made the following observation. Consider the function

$$\psi(q) := \int_0^1 (e^{-qx} - 1 + qx) x^{-2} \Lambda(dx), \quad q \geq 0. \tag{2}$$

Then ψ is the Laplace exponent of a spectrally positive Lévy process and is thus the *branching mechanism* of a CSBP $(Z_t, t \geq 0)$. (Definitions and elementary properties of CSBPs may be found for instance in [19, 21] and [5] and Chapter 6 of [1]). In particular Grey [18] showed that a ψ-CSBP becomes extinct almost surely in finite time, if and only if

$$\int_1^\infty \frac{dq}{\psi(q)} < \infty. \tag{3}$$

Theorem 1 (Bertoin and Le Gall, [10])**.** *Conditions* (1) *and* (3) *are equivalent. In other words, a particular (standard) Λ-coalescent comes down from infinity if and only if the corresponding CSBP becomes extinct.*

The proof of Bertoin and Le Gall (see the end of Section 4 in [10]) is direct and analytical. However, Theorem 1 strongly suggests that a general probabilistic connection exists, and this prompted Bertoin and Le Gall to ask for a probabilistic proof of their result.

The main goal of the present work is to provide an explicit coupling that makes Theorem 1 "obvious". In fact, the coupling yields much more information, including a quantitative estimate on $N^\Lambda(t)$ for small times t. This estimate matches the "speed of coming down from infinity" obtained by the authors in [2] with a martingale method. In fact, the present coupling construction suggested that completely general result in the first place.

Organization and contents of the paper. Our coupling is based on a particle system representation for Λ-coalescents and a connection to a version of Donnelly and Kurtz's *lookdown process.* Both for the sake of completeness and of explaining the differences between our construction and that of [14], we will start by defining the lookdown process. More precisely, we will show that this construction is feasible whenever its driving point process $\pi = \sum_i \delta_{(t_i,p_i)}$, given on $(0,\infty)\times(0,1)$ satisfies $\sum_{t_i\le t} p_i^2 < \infty$ for all $t \ge 0$. This result, which we believe is of independent interest, is stated in Proposition 3.

We will then apply this construction to two distinct point processes, one arising from the Λ-coalescent and the other from the associated CSBP. This is done in Section 2. We then use these representations to obtain a coupling between the two processes. This allows us to conclude that the genealogy of the CSBP is, at small times, "close" to the Λ-coalescent. On the other hand, the CSBP gets extinct in finite time if and only if the number of individuals with descendants alive at a future time $t > 0$ is finite (Proposition 9). This directly yields Theorem 1 and its stronger quantitative version, Theorem 15.

We next use these results together with certain pathwise properties of Lévy processes and CSBPs to discuss the regularity of $N^\Lambda(t)$ as $t \to 0$. Our main result there (Proposition 20) shows that the power-law behavior for $N^\Lambda(t)$ is intimately related to the classical *upper and lower indices* of the Lévy measure of ψ, following Blumenthal and Getoor [12] and Pruitt [26]. The appendix contains an example of a measure Λ that is not "well-behaved", in the sense that the corresponding Λ-coalescent comes down from infinity but the lower and the upper indices are different. We show how this leads to truly oscillatory behavior for $N^\Lambda(t)$, which highlights potential difficulties in the analysis of small-time behavior of general Λ-coalescents.

2 Preliminaries

In this section we describe a general procedure known as the *lookdown construction*, enabling one to construct measure valued processes from point processes on $[0,1]\times\mathbb{R}_+$. The material discussed in this section is mostly well-known, but we prefer to give a brief account of the theory to set the ground for the construction of the coupling in section 3. Unless stated otherwise, we henceforth assume that $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$.

2.1 Lookdown construction

The lookdown construction was first introduced by Donnelly and Kurtz in 1996 [13]. Their goal was to give a construction of the Fleming-Viot superprocess that provides an explicit description of the genealogy of the individuals in the population (see [17] for a reader-friendly introduction to these notions). Donnelly and Kurtz subsequently modified their construction in [14] to include more general measure-valued processes (such as the Dawson-Watanabe superprocesses). It is this version that we use here, and that we will apply to the generalized Fleming-Viot superprocesses (which are dual to Λ-coalescents) as well as to the ratio processes associated to CSBPs. Our approach here shares common points with that of [11].

For a given (infinite size) population evolving in continuous time, let the genetic types of individuals be encoded as numbers in $[0,1]$. More precisely, for each $i \ge 1$ and $t \ge 0$, let $\xi_i(t) \in [0,1]$ be the genetic type of the individual i (or level i) at time t. As will be seen soon, for our models, the infinite particles system $((\xi_1(t),\xi_2(t),\ldots),\, t\ge 0)$ is such that the limiting empirical measure

$$\Xi_t(\cdot) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \delta_{\xi_i(t)}(\cdot)$$

exists simultaneously for all t, almost surely. The process $(\Xi_t(\cdot),\, t\ge 0)$ is a convenient way to track the evolution of the genetic composition of the population.

We first offer an informal description followed by a formal one in Definition 2. The evolution of $(\xi_i(t))_{i\geq 1}$ is driven by a point process (i.e. a countable collection of random points) $\pi = (p_i, t_i)_{i\in\mathbb{N}}$ in $[0,1]\times\mathbb{R}_+$, and a family of i.i.d. coin tosses. Each atom of π corresponds to a *birth* (or resampling) event. Changes in $(\xi_i(t), t\geq 0)_{i\geq 1}$ occur only at birth event times. Let $(p,t)\in\pi$. Then at time t, for each level $i\geq 1$, a coin is tossed, where the probability of head equals p, independently over levels. Those levels for which the coin comes up heads (let us denote this set by $I_{p,t}$) modify their label to $\xi_{\min I_{p,t}}(t-)$. In words, each level in $I_{p,t}$ immediately adopts the type of the smallest level participating in this birth event. For the remaining levels reassign the types so that their relative order immediately prior to this birth event is preserved. More precisely, for each $i\notin I_{p,t}$, let $\xi_i(t-) = \xi_{\phi(i)}(t)$ where ϕ is the unique increasing bijection from $\mathbb{N}\setminus\{\min I_{p,t}\}$ onto $\mathbb{N}\setminus I_{p,t}$.

A more formal description follows. Fix $(U_{i,j})_{i,j\geq 1}$, a collection of i.i.d. uniform variables on $[0,1]$. Let $\pi = \{(p_i,t_i) : i\in\mathbb{N}\}$ be a *fixed* point process on $[0,1]\times\mathbb{R}_+$ such that for any $0\leq t<\infty$

$$\sum_{i:\, t_i\leq t} p_i^2 < \infty. \tag{4}$$

(When we apply this construction later, π will be random and we will work conditionally given π. Condition (4) will then hold almost surely). For each $n\geq 1$, construct the label process associated with π as follows. We fix an infinite sequence of exchangeable random variables $(\xi_i(0))_{i\geq 1}$. Set $\xi_i^n(0) = \xi_i(0)$, $i = 1,\ldots,n$. For each $j\geq 1$ and $i\in\{1,\ldots,n\}$ define

$$A_i(t_j,p_j) \equiv A_j(i) := \{U_{i,j}\leq p_j\} \;\text{ and }\; i_1(j) := \min\{i\geq 1 :\ A_j(i) \text{ occurs}\}. \tag{5}$$

For $i\leq n$, let

$$m_j(i) := \sum_{l=1}^{i} \mathbf{1}_{A_l(j)},\; i\geq 1, \tag{6}$$

be the number of levels smaller or equal to i that participate in the birth event (p_j,t_j). Denote by J the set of atom indices $\{j\geq 1 : m_j(n)\geq 2\}$ for which two or more levels in $\{1,\ldots,n\}$ participate in the corresponding birth event. Order the collection of indices in J so that $t_{j_1} < t_{j_2} < \ldots$ (this is almost surely possible due to (4), see Proposition 3 below). Define $(\xi_i^n(t))_{1\leq i\leq n}$ to be constant over $[t_{j_k}, t_{j_{k+1}})$. Moreover, if $j\in J$, modify the labels at time t_j as follows: for each $1\leq i\leq n$ declare

$$\xi_i^n(t_j) = \xi^n_{i-(m_j(i)-1)_+}(t_j-)\mathbf{1}_{A_j(i)^c} + \xi^n_{i_1(j)}(t_j-)\mathbf{1}_{A_j(i)}, \tag{7}$$

where $m_j(i)$ is defined in (6).

Finally, observe a crucial property of the above construction: if $1\leq m<n$, then the restriction of ξ^n to the first m levels yields ξ^m, in symbols:

$$((\xi_1^n(t),\ldots,\xi_m^n(t)),\, t\geq 0) \equiv ((\xi_1^m(t),\ldots,\xi_m^m(t)),\, t\geq 0). \tag{8}$$

This fact is a simple consequence of the (lookdown) updating rule (7) that makes the type at level i depend only on the previous types at levels up to (and including) i. Therefore, one can unambiguously define the label process $(\xi_i, i = 1,2,\ldots)$ simultaneously for all i, as

$$\xi_i(t) := \xi_i^i(t) \equiv \lim_{n\to\infty} \xi_i^n(t),\; \forall t\geq 0,\; \forall i\geq 1. \tag{9}$$

Definition 2. We call $\xi := (\xi_i(t), t \geq 0)_{i\geq 1}$ the *label process* associated to π. We may write ξ^π for ξ in order to indicate this association. Unless otherwise specified we always assume that the $(\xi_i(0))_{i\geq 1}$ are i.i.d. uniformly distributed on $[0, 1]$.

In the sequel we will often focus on $(N^\pi(t),\, t \geq 0)$, the number of (distinct) types in the population process, defined by

$$N^\pi(t) := \#\{\xi_1(t), \xi_2(t), \ldots\},\ t \geq 0. \tag{10}$$

Note that $N^\pi(t) \in \{1, 2, \ldots\} \cup \{\infty\}$ and $N^\pi(0) = \infty$, due to our assumptions on $\xi(0)$.

The next proposition justifies the above definition of ξ, and ensures that the corresponding limiting empirical measure exists (as a càdlàg Markov process when the process π is a Poisson point process). These facts will be used in the construction of the coupling without further reference in the sequel.

Proposition 3. *Let π be a point process satisfying* (4) *and let $(\xi_i)_{i\geq 1}$ be its label process. Then the limit $\Xi_t = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \delta_{\xi_i(t)}$ exists simultaneously for all t almost surely and is càdlàg with respect to the weak topology.*

Moreover, if π is random and satisfies (4) *almost surely,* $(\Xi_t,\, t \geq 0)$ *is a Markov process in its own filtration provided $U(t) = \sum_{i:t_i\leq t} p_i^2$ has independent increments.*

Definition 4. The process $(\Xi_t, t \geq 0)$ is the *lookdown (measure-valued) process* associated to π. We may write Ξ_t^π instead of Ξ_t to make explicit the dependence on the point process π.

Proof of Proposition 3. The proof can essentially be found in [14], up to a few modifications due to the difference in point of views. We explain how to adapt their arguments to our setting. Recall the notation of Definition 2. To show that ξ^n is well defined, note that, almost surely,

$$\#\{j \geq 1 :\, t_j \in [0, t] \text{ and } m_j(n) \geq 2\} < \infty,\ \forall t \geq 0. \tag{11}$$

Indeed, for each j the indicator $\mathbf{1}_{\{m_j(n)\geq 2\}}$ has expectation $1 - (1 - p_j)^n - np_j(1 - p_j)^{n-1} \leq \binom{n}{2} p_j^2$, and assumption (4) together with Borel–Cantelli lemma ensures (11). Thus the dynamic (inductive) update (7) is feasible, and the label process ξ associated to π is well-defined. A crucial feature of ξ is that for each fixed $t > 0$, the sequence $(\xi_i(t), i = 1, 2, \ldots)$ is exchangeable. Indeed, $(\xi_i(0), i = 1, 2, \ldots)$ is an exchangeable family, and the transitions preserve the exchangeability. An application of de Finetti's theorem now yields the existence of the limit

$$\Xi_t = \lim_{n\to\infty} \Xi_t^n,\ \text{where } \Xi_t^n := \frac{1}{n}\sum_{i=1}^n \delta_{\xi_i(t)}, \tag{12}$$

for any fixed time t, and hence for all $t \in \mathbb{Q}$ simultaneously, almost surely.

To see that the limit Ξ_t actually exists simultaneously for all t with probability one is more delicate and is proved by Donnelly and Kurtz in [14]. Essentially one can adapt the proof of their Lemma 3.4 to see that for each fixed $T > 0, \epsilon > 0$ and each Borel bounded function $f : [0, 1] \mapsto \mathbb{R}$ there exists a positive sequence $(\delta_l)_{l>0}$ such that $\sum_{l\geq 0} \delta_l < \infty$ and such that

$$\mathbb{P}\left(\sup_{t\leq T} \left| \int_0^1 f(x)\Xi_t^m(dx) - \int_0^1 f(x)\Xi_t^l(dx) \right| > \epsilon\right) \leq \delta_l. \tag{13}$$

This implies that the sequence $\int_0^1 f(x)\Xi_t^m(dx)$ is almost surely Cauchy. Since the space of bounded measurable functions is separable (see e.g. Lemma 1.2 in [14]) this is enough to guarantee existence $(\Xi_t, t \geq 0)$ as a process with values in the set of Borel measures. Moreover Ξ_t^n converges for all $t \leq T$ simultaneously almost surely.

Now assume that π is a random point process satisfying (4) and that $(U(t), t \geq 0)$ has independent increments. Then it is easy to check that the label process $(\xi_i(t), t \geq 0)$ is Markov (in its own filtration). The Markov property for Ξ then follows directly from the fact that exchangeable laws on $[0,1]^\infty$ are by De Finetti's theorem in one-to-one correspondence with the law of their empirical measure on $[0,1]$. (Note that, however, $(\Xi_t, t \geq 0)$ is **not** Markov with respect to the strictly greater filtration of the label process, since the type of individual 1 will tend to take over the population as time evolves). □

Remark 5. Donnelly and Kurtz [14] work under a different set of assumptions. Their setup is more general in the sense that they do not assume the consistency of the finite-n label processes $(\xi_i^n(t), t \geq 0)_{1 \leq i \leq n}$. (Furthermore, note that they also include a Markov mutation diffusion operator that drives the motion of labels in between reproduction events.) In fact, the total number of particles is allowed to vary in their setting. For this reason, their construction does not make sense conditionally given the (limiting) point process π, which is an important feature of our construction. The main novelty in our setting is the observation that the assumption (4) is in fact all that is needed to guarantee existence of the measure-valued process $(\Xi_t, t \geq 0)$ (it is also clear that this condition is necessary for the very construction of the label process). In the notation of Donnelly and Kurtz, this amounts to checking that the process $(U^n(t), t \geq 0)$ converges in distribution to $(U(t), t \geq 0)$.

2.2 Ancestral partitions, Fleming-Viot processes and Λ-coalescents

We next apply Proposition 3 in two different settings, corresponding to the Fleming-Viot process and to the CSBP, respectively. The upshot of this construction is a convenient way to track the respective genealogies. This is achieved through the *ancestral partition process*, associated to the process ξ constructed in Proposition 3.

Let π be a point process satisfying (4), and ξ^π its associated label process. Note that for each $s > 0$, the shifted point process $\pi^{-s} := \{(p, t-s) : (p,t) \in \pi, t \geq s\}$ also satisfies (4), and that, due to the updating rule (7), the label updates of the associated label process $\{\xi^{\pi^{-s}}(t), t \geq 0\}$ are the same as those of $\{\xi^\pi(t), t \geq s\}$. The difference between the two processes is manifested through their initial states, since for $i \neq j$ we have $\xi_i^{\pi^{-s}}(0) \neq \xi_j^{\pi^{-s}}(0)$, almost surely, while it is possible that $\xi_i^\pi(s) = \xi_j^\pi(s)$. Now fix some $T > 0$.

Definition 6. The ancestral partition process $(\mathcal{R}^T(t), 0 \leq t \leq T)$ takes values in the space of level partitions (or partitions of $\mathbb{N}$). For each $t \leq T$, $\mathcal{R}^T(t)$ is defined by the equivalence relation: $i \sim j$ in $\mathcal{R}^T(t)$ if and only if $\xi_i(T)$ and $\xi_j(T)$ descend from the same level at time t, or equivalently, if $\xi_i^{\pi^{-t}}(T-t) = \xi_j^{\pi^{-t}}(T-t)$ (see also equation (2.3) in [11]).

Note that $\mathcal{R}^T(T)$ is the trivial partition $\{\{i\} : i \in \mathbb{N}\}$ and that $\mathcal{R}^T(t_1)$ is a coarser partition than $\mathcal{R}^T(t_2)$, whenever $0 \leq t_1 \leq t_2 \leq T$.

We now briefly recall the definition of generalized Fleming-Viot processes as well as their link to Λ-coalescents. A generalized Λ-Fleming-Viot process (in the sense of Bertoin and Le Gall [9]) $(\rho_t, t \geq 0)$ is a Markov process taking values in the space $\mathcal{M}$ of probability measures on $[0,1]$. Its generator L is defined as follows: given a finite measure Λ on $[0,1]$,

$$LF(\mu) = \int_{(0,1]} y^{-2}\Lambda(dy) \int_{[0,1]} \mu(dx) \left(F((1-y)\mu + y\delta_x) - F(\mu)\right), \tag{14}$$

where $F : \mathcal{M} \to \mathbb{R}$ is a bounded continuous function. In words, a number y between 0 and 1 is sampled at rate $y^{-2}\Lambda(dy)$. A type x is sampled from ρ_{t-}. Then ρ_t is obtained from ρ_{t-} by scaling down ρ_{t-} by $(1-y)$, and adding to the result an atom at x of mass y.

Theorem 7. *Let Λ be a finite measure on $[0,1]$. Let π be a Poisson point process on $[0,1]\times\mathbb{R}$ with intensity $x^{-2}\Lambda(dx) \otimes dt$. Then the lookdown process Ξ^π (cf. Definition 4) is a Λ-generalized Fleming-Viot process with generator* (14)*, started from the uniform measure on* $[0,1]$*. Furthermore, the ancestral partition process $(\mathcal{R}^T(T-t), 0 \le t \le T)$ is the Λ-coalescent, run for time T.*

Proof. A careful proof of this fact can be found in Lemma 3.6 of [11], that is directly based on the work of Donnelly and Kurtz [14]. We include a simpler proof which relies instead on the duality introduced by Bertoin and Le Gall [8]. We start by the claim that the ancestral partition process $(\mathcal{R}^T(T-t), 0 \le t \le T)$ is the $\Lambda-$coalescent. This follows simply from the following observation: let π' be the point process obtained from π by applying the transformation $(p,t) \mapsto (p, T-t)$. Then π' has same law as π restricted to $[0,T]$ and is thus a Poisson point process on $[0,1]\times[0,T]$ with intensity $x^{-2}\Lambda(dx)\otimes dt\mathbf{1}_{[0,T]}(t)$. Now, the updating rule (7) can be rephrased as follows: at each atom (x,t) of π' one flips a coin for each active ancestral lineage with probability of heads equal to p and the lineages that come up heads merge. This is precisely the Poisson process construction of Λ-coalescents (see, e.g., Theorem 3.2 in [5]).

Let $(\rho_t, t \ge 0)$ be a Fleming-Viot process, and let

$$F_t(x) = \rho_t([0,x]), \quad 0 \le x \le 1$$

be the associated *bridge* process. Denote by F_t^{-1} the càdlàg inverse of the map $x \mapsto F_t(x)$. Let $V_1, V_2, \ldots,$ be i.i.d. uniform random variables in $[0,1]$, independent of $(\rho_t, t \ge 0)$. By the Glivenko-Cantelli theorem (see, e.g., (7.4) in Chapter 1 of Durrett [16]), noting that $(F_t^{-1}(V_i), i \ge 1)$ are i.i.d. samples from the random measure ρ_t, we have for each fixed $t \ge 0$

$$\rho_t = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \delta_{F_t^{-1}(V_i)}, \text{ almost surely,} \tag{15}$$

where the limit is taken in the sense of the weak topology on probability measures.

Bertoin and Le Gall [9] proved that the Λ-coalescent $(\Pi_t, t \ge 0)$ is dual to the generalized Fleming-Viot process corresponding to Λ in the following sense: if $n \ge 1$ and f is any continuous function on $[0,1]^n$, then

$$\mathbb{E}(f(F_t^{-1}(V_1), \ldots, F_t^{-1}(V_n))) = \mathbb{E}(f(Y(\Pi^n(t), V_1', \ldots, V_n'))), \tag{16}$$

where $\Pi^n(t)$ denotes the restriction of Π_t to $[n]$, the random variables $(V_1', \ldots, V_n')$ are i.i.d. uniform on $[0,1]$, and independent of $(\Pi_t, t \ge 0)$, and where the map Y is defined as follows:

for $\pi \in \mathcal{P}_n$ and $(x_1, \ldots, x_n) \in [0,1]^n$ let
$Y(\pi, x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ with $y_j = x_i$ for $i = \min\{k : k \sim_\pi j\}$.

Note that the duality relation (16) has the form of a generalized functional duality in the context of interacting particle systems (see [22]), and should not be confused with the notion of duality between coagulation and fragmentation processes of [25].

We next verify that, for each $t > 0$,

$$(\xi_1(t), \dots, \xi_n(t)) \overset{d}{=} Y(\Pi^n(t), V_1', \dots, V_n'). \tag{17}$$

This fact is an immediate consequence of our construction. Indeed, at time t two levels i and j have the same type $\xi_i(t) = \xi_j(t)$ if and only if they descend from the same level at time 0 (since all the $\xi_i(0)$ are almost surely distinct). Hence $\xi_i(t) = \xi_j(t)$ if and only if i and j belong to the same block of $\mathcal{R}^t(0)$. Therefore

$$(\xi_1(t), \dots, \xi_n(t)) = Y'(\mathcal{R}^t(0), \xi_1(0), \dots, \xi_n(0)),$$

where for $\pi = (B_1, B_2, \dots) \in \mathcal{P}_n$ and $(x_1, \dots, x_n) \in [0,1]^n$ we let

$$Y'(\pi, x_1, \dots, x_n) = (y_1, \dots, y_n) \text{ with } y_j = x_i \text{ for } j \in B_i.$$

Clearly, as long as the random variables $\Pi \in \mathbb{P}_n$ and $(X_1, \dots, X_n)$ (i.i.d. uniform on $[0,1]$) are independent one has

$$Y(\Pi, X_1, \dots, X_n) \overset{d}{=} Y'(\Pi, X_1, \dots, X_n),$$

and since the $\xi_i(0)$ are i.i.d. uniform on $[0,1]$ and $\mathcal{R}^t(0) \overset{d}{=} \Pi(t)$, this proves the claim (17). Due to (16), one concludes that $(F_t^{-1}(V_1), \dots, F_t^{-1}(V_n))_{t\geq 0}$ and $(\xi_1(t), \dots, \xi_n(t))_{t\geq 0}$ have the same one-dimensional marginals. This implies that

$$\forall t \geq 0,\ \Xi_t^n \overset{d}{=} \frac{1}{n}\sum_{i=1}^n \delta_{F_t^{-1}(V_i)},\ n \geq 1, \text{ and hence that } \Xi_t \overset{d}{=} \rho_t.$$

Our argument was carried out under the assumption that the initial state is the uniform law on $[0,1]$. However, it would equally apply if the $\xi_i(0)$ were drawn independently from any other law on $[0,1]$. Since Ξ and ρ are both càdlàg Markov processes, they must be equal in distribution. $\square$

2.3 Lookdown process of a CSBP

Recall ψ from (2) and consider a CSBP $(Z(t), t \geq 0)$ with branching mechanism ψ (see, e.g., [1] or [5, Chapter 4.2] for an elementary introduction). In the sequel, we often refer to any such process as ψ-CSBP. In this section assume that Z is started from $Z(0) = 1$. Following Bertoin and Le Gall [7], recall existence of a two parameter branching family $(Z_t(x), t \geq 0, x \in [0,1])$, such that for each fixed $x \in [0,1]$, $(Z_t(x),\ t \geq 0)$ is a ψ-CSBP started from $Z_0(x) = x$, independent from the ψ-CSBP $(Z_t(1) - Z_t(x),\ t \geq 0)$. In particular $(Z_t(1),\ t \geq 0) \overset{d}{=} (Z(t),\ t \geq 0)$. The quantity $Z_t(x)$ can be interpreted as the population size at time t, descended from the initial fraction x of the population at time 0. Furthermore, the branching property also implies that, for any $t > 0$, $(Z_t(x),\ x \in [0,1])$ is a subordinator.

We briefly recall the setting of [11]. For each fixed $t \geq 0$, define $M_t([x_1, x_2]) := Z_t(x_2) - Z_t(x_1)$, for all $0 \leq x_1 \leq x_2 \leq 1$. Then M_t extends to a random measure on $[0,1]$. The process

$M = (M_t,\, t \geq 0)$ is easily seen to be Markov, with a generator given by (see, (1.15) in [11] for the general case formula)

$$\mathcal{L}F(\mu) = \int_0^1 \mu(da) \int_{[0,1]} \nu(dh)(F(\mu + h\delta_a) - F(\mu) - hF'(\mu; a)),$$

where $\nu(dh) = \Lambda(dh)/h^2$ and $F'(\mu; a)$ denotes the Fréchet derivative of F at μ in the direction δ_a (see, e.g., (1.4) in [11]). The process M encodes the genealogy of the CSBP $(Z_t(1), t \geq 0)$ (this is a continuous time/space analogue to the relation between a Galton–Watson process and the associated tree). The composition of the population is then well-described by the *ratio process* $R = (R_t,\, t \geq 0)$ defined by $R_t = \frac{1}{Z_t(1)} M_t$, taking values in the space of probability measures. Now define

$$\pi^Z = \{(\Delta Z(t)/Z(t), t) : t \geq 0\} \tag{18}$$

to be the point process of normalized jump sizes of Z. Here and below (without further mention), we will account in π^Z only the points $(\Delta Z(t)/Z(t), t) \in (0, \infty) \times [0, \infty)$, which represent the true jumps of the process.

Lemma 8. *The condition* (4) *holds for* π^Z*, and the associated lookdown process* $(\Theta_t(.), t \geq 0)$ *is equal in law to the ratio process* $(R(t), t \geq 0)$*.*

Proof. A detailed proof is given in the "Proof of (2.4)" in [11] pp. 313-315, although the idea goes back at least to Theorem 3.2 in [14]. □

2.3.1 Evolution of the number of types

Let Z be a CSBP with branching mechanism ψ started from $Z_0 = 1$, and assume that Grey's condition (3) is satisfied. Denote by $\zeta_Z = \inf\{t \geq 0 : Z(t) = 0\}$ its (almost surely finite) extinction time. Let π^Z be the associated point process of rescaled jump sizes (18), and note that π^Z has no points in $[0, 1] \times (\zeta_Z, \infty)$. Recall definition (10), and define $N^Z(t) = N^{\pi^Z}(t)$, $t < \zeta_Z$, and $N^Z(t) = 0$, $t \geq \zeta_Z$.

Let us define $v(t) := \inf\{z > 0 : \int_z^\infty \psi(q)^{-1} dq < t\}$ with the convention that $\inf \emptyset = \infty$ or equivalently let $v(t)$ be the solution of

$$\int_{v(t)}^\infty \frac{dq}{\psi(q)} = t. \tag{19}$$

Recall from Duquesne and Le Gall [15] that the function v describes the evolution of the number of alive families at time t in a ψ-CSBP. More precisely,

Proposition 9. *If (3) is satisfied then* $N^Z(t) < \infty$*, for all* $t > 0$*, almost surely, and moreover*

$$(N^Z(t), t \geq 0) \overset{d}{=} (Q(v(t)), t \geq 0). \tag{20}$$

where $t \mapsto Q(t)$ *is a standard Poisson counting process, and where* $v(t)$ *is defined in (19). In particular,*

$$\lim_{t \to 0} \frac{N^Z(t)}{v(t)} = 1, \text{ almost surely.} \tag{21}$$

If (3) is not satisfied then both v *and* N^Z *are infinite for all* $t > 0$ *almost surely.*

Proof. This essentially follows from Theorem 12 in [4] and Corollary 1.4.2 (ii) in [15] (see also Corollary 4.1 in [5] for an elementary sketch of proof). Indeed, when Grey's condition is satisfied, we may use the construction of [4] for the Donnelly-Kurtz lookdown process, where the labeling process $(\theta_1(t), \theta_2(t), \ldots)$ is directly defined in terms of the excursions of a Continuous Random Tree (CRT) with branching mechanism ψ (see [15] or [4] for the basic terminology and properties of these objects, to which we will refer in this proof). Let $(H_s, 0 \le s \le T_1)$ be the height process associated with $(Z_s, s \ge 0)$, where $T_1 := \inf\{u > 0 : L_u^0 > 1\}$ and where $(L_u^0, u \ge 0)$ is the local time process at level 0 of $(H_s, s \ge 0)$. It follows from the construction in [4] that one can embed the lookdown construction in the CRT so that for any $t > 0$, $N^Z(t)$ is exactly the number of excursions of $(H_s, 0 \le s \le T_1)$ that reach level t. It follows directly (by excursion theory for $(H_s, 0 \le s \le T_1)$) that $(N_Z(t), t > 0)$ has the law of $(Q_{\tilde{v}(t)}, t \ge 0)$, where by definition:

$$\tilde{v}(t) = \mathbf{N}(\sup_{s \ge 0} H_s > t).$$

Here, $\mathbf{N}(\cdot)$ denotes the excursion measure of H. By Corollary 1.4.2 (ii) of [15], $\tilde{v}(t) = v(t) < \infty$, which proves the result. □

Remark 10. For each fixed $t > 0$, due to the exchangeability of the sequence $(\xi_i(t), i = 1, 2, \ldots)$, the number of types $N^Z(t)$ is almost surely equal to the number of atoms of the purely atomic measure $\Xi_t = \lim_{n\to\infty} n^{-1} \sum_{i=1}^n \delta_{\xi_i(t)}$.

Remark 11. The property $\mathbb{P}(N^Z(t) < \infty) = 1$ may seem counter-intuitive in view of the fact that types are not destroyed in any particular application of the updating rule (7). However, an accumulation of many densely placed small lookdown jumps "pushes off" to infinity all but finitely many types in any positive amount of time, whenever Grey's condition (3) is fulfilled.

3 The coupling

3.1 Coupling construction

We can now explain the coupling between Λ-coalescents and CSBP. The key idea is to use the following result due to Lamperti, which expresses any CSBP as a time-change of a Lévy process.

Consider a Lévy process $(X_t,\, t \ge 0)$ with Laplace exponent ψ given in (2), and assume $X_0 = x \in (0, 1]$. Define

$$U^{-1}(t) := \inf\left\{ s > 0 : \int_0^s \frac{du}{X_u} > t \right\}, \tag{22}$$

and

$$Z_t = X_{U^{-1}(t)}, \quad t \ge 0. \tag{23}$$

Theorem 12. (Lamperti [19, 21].) *The process $(Z_t,\, t \ge 0)$ is a ψ-CSBP started from $Z_0 = x$.*

Construction. We now describe the coupling between the genealogies of a CSBP and Fleming-Viot processes. Assume that the Lévy process X and its corresponding CSBP Z (Lamperti time-changed as above) satisfy $X_0 = Z_0 = 1$. As before, denote by π^Z the point process of the rescaled jump sizes of Z. Call $\xi = (\xi_i(t), t \ge 0)_{i \ge 1}$ the label process of π^Z

obtained from the lookdown construction applied to Z.
Consider simultaneously the point process $\pi^X = (\Delta X(t_i), t_i)$ of (unscaled) jump sizes of X, and its associated label process $\theta = (\theta_i(t),\, t \geq 0)_{i\geq 1}$, as well as the lookdown measure $\Theta = (\Theta_t,\, t \geq 0)$. Then Θ is a Λ-Fleming-Viot process, and hence (due to Theorem 7) has a genealogy given by a Λ-coalescent. Indeed, since X is a Lévy process, due to the Lévy-Itô decomposition, the point process of jumps $\pi^X = (\Delta X(t_i), t_i)$ is a Poisson point process with intensity $\nu(dx) \otimes dt$, where $\nu(dx) = x^{-2}\Lambda(dx)$ is the Lévy measure of X.

Heuristics. For a small $t > 0$, the two point processes π^X and π^Z, restricted to $[0, t]$, are "close to each other". Indeed, each point $(p, t) \in \pi^X$ also corresponds to a point $(\tilde{p}, \tilde{t}) \in \pi^Z$, where $\tilde{t} = U^{-1}(t)$, and $\tilde{p} = p/Z(\tilde{t})$. Now, since $(X_t, t \geq 0)$ is almost surely continuous at $t = 0$, the time-change U^{-1} is almost surely differentiable at $t = 0$ with derivative close to 1. Therefore, $U^{-1}(t) \sim t$ as $t \to 0$, and one deduces that for small t, $\tilde{t} \approx t$. Likewise, invoking the continuity of Z and the fact $Z_0 = 1$, we have $Z(\tilde{t}) \approx 1$, hence $(\tilde{p}, \tilde{t}) \approx (p, t)$.
It is therefore reasonable to believe that for small t, $N^X(t) \approx N^Z(t)$, where $N^X(t)$ (resp. $N^Z(t)$) is the number of types in the lookdown process associated to π^X (resp. π^Z) at time t. At the same time, by Proposition 9 we also know $N^Z(t) \sim v(t)$ almost surely as $t \to 0$, and all of the above strongly suggests that the same is true for N^X in place of N^Z.
Finally, due to Theorem 7, we have

$$N^X(t) \stackrel{d}{=} N^\Lambda(t), \text{ for each fixed } t \geq 0, \tag{24}$$

where $N^\Lambda(t)$ is (as usual) the number of blocks in the corresponding Λ-coalescent at time t. The reader can easily check this property by restricting attention to the first n levels, and using the updating rule (7), as well as the fact that $(\pi^X(t),\, t \in [0, T])$ and $(\pi^X(T - t),\, t \in [0, T])$ have the same distribution. Therefore, we obtain $N^\Lambda(t) \sim v(t)$ in probability, as $t \to 0$.

We will now turn these heuristic observations into a rigorous argument for Proposition 15), starting with a monotonicity lemma.

Definition 13. Given two point processes π and π^+ on $[0, 1] \times \mathbb{R}_+$ on the same probability space, and a random time $T \geq 0$, measurable with respect to the filtration generated by π and π^+, we write $\pi \lhd_{|[0,T]} \pi^+$ (or $\pi \lhd \pi^+$ on $[0, T]$) if there exists an increasing càdlàg process $r : [0, T] \mapsto \mathbb{R}^+$ such that, almost surely, $r(0) = 0$ and

$$\pi = \{(p_i, t_i) : i \geq 1\} \text{ and } \pi^+ = \{(q_i, r(t_i)) : i \geq 1\},$$

where $p_i \leq q_i$, for each $i \geq 1$ such that $t_i \leq T$.

In words, $\pi \lhd \pi^+$ on $[0, T]$, if the atoms of π^+ are those of π, time-changed by r, and multiplied in size by a (possibly non-constant and random) quantity not smaller than 1. Observe that r preserves the order of the atoms, almost surely. In our main applications, the form of r will be rather simple. Furthermore, the processes π and π^+ of interest will both have (countably) infinitely many atoms in any interval of positive length, almost surely, ensuring that $\{r(T) < \infty\} = \{T < \infty\}$, almost surely.

Consider now π and π^+ such that $\pi \lhd \pi^+$ on $[0, T]$ for some finite random time T, and both

$$\sum_{i:\, t_i \leq t} p_i^2 < \infty,\ \sum_{i:\, t_i \leq t} q_i^2 < \infty,\ \forall t \geq 0, \text{ almost surely.}$$

One can then construct a coupling of Ξ^π (with its label processes $\xi = (\xi_i(t), t \geq 0)_{i\geq 1})$ and Ξ^{π^+} (with its label processes $\xi^+ = (\xi_i^+(t), t \geq 0)_{i\geq 1})$, by using the same collection $\{U_{i,j}\}_{i,j\in\mathbb{N}}$ of i.i.d. uniform random variables to specify the levels participating in the resampling events in Definition 2. Due to $\pi \lhd \pi^+$ on $[0,T]$, the following result is obvious by construction:

Lemma 14. *If $\pi \lhd \pi^+$ on $[0,T]$, then*

$$\mathbb{P}(N^{\pi^+}(r(s)) \leq N^\pi(s),\ \forall s \in [0,T]) = 1.$$

3.2 Proof of Theorem 1 and the asymptotic for the number of blocks

To prove Theorem 1 it suffices to show that $N^\Lambda(t)$ is infinite for all $t > 0$ whenever $v(t) = \infty, \forall t > 0$ and is finite for all $t > 0$ in the converse case. This is now a consequence of the above coupling, used to show the following proposition.

Proposition 15. *For each $\varepsilon \in (0,1)$,*

$$\mathbb{P}\left(\liminf_{t\to 0} \frac{N^X(t)}{v\left(\frac{1+\varepsilon}{1-\varepsilon}\,t\right)} \geq \frac{1}{1+\varepsilon}\ ,\ \limsup_{t\to 0} \frac{N^X(t)}{v\left(\frac{1-\varepsilon}{1+\varepsilon}\,t\right)} \leq \frac{1}{1-\varepsilon}\right) = 1, \tag{25}$$

and therefore

$$\lim_{t\to 0} \mathbb{P}\left[\frac{1}{(1+\varepsilon)^2}\cdot v\left(\frac{1+\varepsilon}{1-\varepsilon}\,t\right) \leq N^\Lambda(t) \leq \frac{1}{(1-\varepsilon)^2}\cdot v\left(\frac{1-\varepsilon}{1+\varepsilon}\,t\right)\right] = 1. \tag{26}$$

Remark 16. Observe that N^X and N^Λ have only the same one-dimensional marginal distributions, but they are not equal in distribution as processes. For instance while the first one only decreases by jumps of size 1 (this is known at least in the stable case, see [20]), the second one can decrease by jumps of arbitrary integral length. Thus one cannot obtain more than (26) from (25). This result is clearly weaker than Theorem 1 in [2]

$$\lim_{t\to 0} \frac{N^\Lambda(t)}{v(t)} = 1, \text{ almost surely.} \tag{27}$$

As mentioned in the introduction, it is the use of a sophisticated martingale technique which yields this stronger result there. However, it was the knowledge of the coupling described bellow that initiated [2] and suggested the form of the asymptotics in the first place.

Proof of Proposition 15. We start by showing (25) for ε sufficiently small. The conclusion (26) will then readily follow. Let us assume for the moment that $\mathrm{supp}(\Lambda) \subset [0,\eta]$ where $\eta < 1$, and fix some $\varepsilon \in (0, 1/\eta - 1)$. Consider again the Lévy process X with Laplace exponent ψ such that $X_0 = 1$, and let

$$\pi = \pi^X = \{(\Delta X_t, t) : t > 0\}$$

be the corresponding Poisson point process. Let π_ε^- (resp. π_ε^+) be the image of π under the map $(p,t) \mapsto (p(1-\varepsilon), t)$ (resp. $(p,t) \mapsto (p(1+\varepsilon), t)$). Due to our assumptions on $\mathrm{supp}(\Lambda)$ and the choice of ε, we have that for each atom (p,t) of π, $p(1+\varepsilon) < 1$ almost surely. Therefore, both π_ε^+ and π_ε^- are Poisson point processes on $(0,1)\times\mathbb{R}_+$. Let $\nu_\varepsilon^+ \otimes dt$ (resp. $\nu_\varepsilon^- \otimes dt$) be the intensity measure corresponding to π_ε^+ (resp. π_ε^-). If f is a Borel function on $[0,1]$, then ν_ε^+ is obtained by the formula

$$\int_{[0,1]} f(x)\nu_\varepsilon^+(dx) = \int_{[0,1]} f(x(1+\varepsilon))\nu(dx),$$

and ν_ε^- is obtained by an analogous formula with $1-\varepsilon$ in place of $1+\varepsilon$. For $\lambda > 0$, let

$$\psi_\varepsilon^\pm(\lambda) := \int_{(0,1)} (e^{-\lambda x} - 1 + \lambda x)\nu_\varepsilon^\pm(dx).$$

By the above observation we see that, for each $\lambda > 0$,

$$\psi_\varepsilon^+(\lambda) = \psi(\lambda(1+\varepsilon)),\ \psi_\varepsilon^-(\lambda) = \psi(\lambda(1-\varepsilon)). \tag{28}$$

Therefore, if we let $u_\varepsilon^\pm(t) := \int_t^\infty d\lambda/\psi_\varepsilon^\pm(\lambda)$ and $v_\varepsilon^\pm(t) = (u_\varepsilon^\pm)^{-1}(t)$ the càdlàg inverse of $u_\varepsilon^\pm$, we have

$$u_\varepsilon^+(s) = \frac{1}{1+\varepsilon}u(s(1+\varepsilon)) \text{ and } u_\varepsilon^-(s) = \frac{1}{1-\varepsilon}u(s(1-\varepsilon)),$$

hence

$$v_\varepsilon^+(t) = \frac{1}{1+\varepsilon}v(t(1+\varepsilon)) \quad\text{and}\quad v_\varepsilon^-(t) = \frac{1}{1-\varepsilon}v(t(1-\varepsilon)). \tag{29}$$

Recall that $X_0 = 1$, and define

$$X_t^+ = (1+\varepsilon)X_t - \varepsilon, \quad\text{and}\quad X_t^- = (1-\varepsilon)X_t + \varepsilon,\ t > 0.$$

Then it easy to see that both $(X_t^+, t \geq 0)$ and $(X_t^-, t \geq 0)$ are Lévy processes such that $X_0^+ = X_0^- = 1$. Moreover, the Laplace exponent of X^+ (resp. X^-) is ψ_ε^+ (resp. ψ_ε^-).

Define $T_\varepsilon^+ = \inf\{s : |X^+(s) - 1| > \varepsilon\}$ and $T_\varepsilon^- = \inf\{s : |X^-(s) - 1| > \varepsilon\}$. Then, almost surely we have, for all $t \geq 0$

$$\frac{\Delta X^-(t)}{X^-(t)} \leq \frac{\Delta X^-(t)}{1-\varepsilon} = \Delta X(t) = \frac{\Delta X^+(t)}{1+\varepsilon} \leq \frac{\Delta X^+(t)}{X^+(t)} \text{ on } \{t \leq T_\varepsilon^+ \wedge T_\varepsilon^-\}. \tag{30}$$

Using the Lamperti transform, now define two continuous-state branching processes with branching mechanism ψ_ε^+ and ψ_ε^-, respectively, by setting $U_\pm(t) := \int_0^t \frac{1}{X_u^\pm} du$,

$$U_\pm^{-1}(t) := \inf\{s \geq 0 : U_\pm(s) > t\} \text{ and } Z_t^+ := X^+_{U_+^{-1}(t)},\ Z_t^- := X^-_{U_-^{-1}(t)},\ t \geq 0.$$

Finally define $\pi^{Z^+} := \{(\Delta Z_s^+/Z_s^+, s) :\ s \geq 0\}$ and $\pi^{Z^-} := \{(\Delta Z_s^-/Z_s^-, s) :\ s \geq 0\}$. Due to (30), we have that almost surely

$$\pi^{Z^-} \lhd_{|[0,U_-(T_\varepsilon^+ \wedge T_\varepsilon^-)]} \pi \text{ (with } r = U_-^{-1}) \text{ and } \pi \lhd_{|[0,T_\varepsilon^+ \wedge T_\varepsilon^-]} \pi^{Z^+} \text{ (with } r = U_+),$$

where $\lhd$ is as in Definition 13. Both $T_\varepsilon^+ \wedge T_\varepsilon^-$ and $U_-(T_\varepsilon^+ \wedge T_\varepsilon^-)$ are clearly strictly positive and finite, almost surely. Hence, Lemma 14 gives that almost surely, for all $t \geq 0$

$$N^\pi(t) \leq N^{\pi^{Z^-}}(U_-(t)) \text{ and } N^\pi(t) \geq N^{\pi^{Z^+}}(U_+(t)), \quad\text{on } \{t \leq T_\varepsilon^+ \wedge T_\varepsilon^-\}.$$

Observe that this is already enough to prove Theorem 1 since $v_\varepsilon^\pm$ is finite if and only if v is finite, and thus $N^{\pi^{Z^+}}(U_+(t)) = \infty$ for all $t > 0$ if $v(t) = \infty$ for all $t > 0$ and likewise $N^{\pi^{Z^-}}(U_-(t)) < \infty$ for all $t > 0$ if $v(t) < \infty$ for all $t > 0$.

Proposition 9 implies that

$$\lim_{t\to 0} \frac{N^{\pi^{Z^-}}(t)}{v_\varepsilon^-(t)} = \lim_{t\to 0} \frac{N^{\pi^{Z^+}}(t)}{v_\varepsilon^+(t)} = 1, \text{ almost surely.}$$

This together with $\mathbb{P}(T_\varepsilon^+ \wedge T_\varepsilon^- > 0) = 1$ and the discussion above yields

$$\limsup_{t\to 0} \frac{N^\pi(t)}{v_\varepsilon^-(U_-(t))} \le 1, \text{ and } \liminf_{t\to 0} \frac{N^\pi(t)}{v_\varepsilon^+(U_+(t))} \ge 1, \tag{31}$$

almost surely. Moreover, it is easy to check that almost surely, for all $t \ge 0$

$$t/(1+\varepsilon) \le U_\pm(t) \le t/(1-\varepsilon), \quad \text{on } \{t \le T_\varepsilon^+ \wedge T_\varepsilon^-\}. \tag{32}$$

Due to monotonicity of $v_\varepsilon^\pm$ and (32), we have that again almost surely, for all $t \ge 0$

$$\frac{N^\pi(t)}{v_\varepsilon^-(U_-(t))} \ge \frac{N^\pi(t)}{v_\varepsilon^-(t/(1+\varepsilon))} \text{ and } \frac{N^\pi(t)}{v_\varepsilon^+(U_+(t))} \le \frac{N^\pi(t)}{v_\varepsilon^+(t/(1-\varepsilon))}, \text{ on } \{t \le T_\varepsilon^+ \wedge T_\varepsilon^-\}. \tag{33}$$

Combining (29), (31) and (33), and recalling $\mathbb{P}(T_\varepsilon^+ \wedge T_\varepsilon^- > 0) = 1$, we can now conclude that

$$\limsup_{t\to 0} \frac{N^\pi(t)}{v\left(t\frac{1-\varepsilon}{1+\varepsilon}\right)} \le \frac{1}{1-\varepsilon} \text{ and } \liminf_{t\to 0} \frac{N^\pi(t)}{v\left(t\frac{1+\varepsilon}{1-\varepsilon}\right)} \ge \frac{1}{1+\varepsilon}, \text{ almost surely.} \tag{34}$$

Since $N^X = N^\pi$ by definition, this gives (25), under the hypothesis that Λ does not give positive mass to a neighborhood of 1. Otherwise, we modify the above argument in the following way. For a fixed $\eta \in (0,1)$, since $x^{-2}\Lambda(dx)$ assigns a finite mass to $(1-\eta, 1]$, the first time T_η when X makes a jump of size strictly greater than η has an exponential random variable law (with finite rate), hence it is strictly positive with probability one. The analysis (30)–(34) clearly works if $T_\epsilon^+ \wedge T_\epsilon^-$ is everywhere replaced by $T_\epsilon^+ \wedge T_\epsilon^- \wedge T_\eta$, yielding (25).

In particular, almost surely, for all t sufficiently small,

$$\frac{1}{(1+\varepsilon)^2} \cdot v\left(t\frac{1+\varepsilon}{1-\varepsilon}\right) \le N^X(t) \le \frac{1}{(1-\varepsilon)^2} \cdot v\left(t\frac{1-\varepsilon}{1+\varepsilon}\right).$$

The limit (26) is easily deduced from (24) and this final estimate. □

The asymptotics (27) in the sense of convergence in probability can be obtained from Proposition 15 under additional assumptions on v (that is, on Λ) as the following result shows.

Proposition 17. *Assume $\Lambda(\{0\}) = 0$. Then, the convergence*

$$N^\Lambda(t)/v(t) \to 1 \quad \textit{in probability}$$

holds at least if

$$\lim_{\varepsilon\to 0} \limsup_{t\to 0} \frac{v(t(1-\varepsilon))}{v(t)} = 1, \ \lim_{\varepsilon\to 0} \liminf_{t\to 0} \frac{v(t(1+\varepsilon))}{v(t)} = 1, \tag{35}$$

and, in particular, if

$$\psi(v(t)) = O(v(t)/t), \ \textit{as } t \to 0. \tag{36}$$

Proof of Proposition 17. The first claim follows by simple calculus manipulations from (26). To see why (36) implies (35), we note that $\psi : [0, \infty) \to \mathbb{R}^+$ of (2) is a (strictly) increasing and convex function on $[0, \infty)$. Furthermore, $v'_\psi(s) = -\psi(v_\psi(s))$, so that v_ψ is decreasing with its derivative decreasing in absolute value. Therefore, for $\varepsilon > 0$ small enough,

$$|v(t(1+\varepsilon)) - v(t)| = \int_t^{t(1+\varepsilon)} |v'(s)| ds \leq |v'(t)|\varepsilon t = \psi(v(t)) t\varepsilon.$$

Similarly,

$$\begin{aligned} |v(t(1-\varepsilon)) - v(t)| &= \int_{t(1-\varepsilon)}^{t} |v'(s)| ds \leq |v'(t(1-\varepsilon))| t\varepsilon \\ &= \psi(v(t(1-\varepsilon))) t(1-\varepsilon) \frac{\varepsilon}{1-\varepsilon}. \end{aligned}$$

Hence (35) will hold provided $\psi(v(t))t = O(v(t))$. □

4 Regularity indices and consequences

In this section we use the quantitative estimates obtained above (Proposition 15) to get concrete information on the small-time behavior of $N^\Lambda(t)$. We are particularly concerned with power-law behavior, which as we show below turns out to be intimately related to the notion of upper and lower indices, which arose in seminal papers by Blumenthal and Getoor [12] and Pruitt [26] on pathwise properties of Lévy process.

Let $X = (X_t, t \geq 0)$ be a Lévy process with Laplace exponent ψ given by (2). We call $\nu(dx) = x^{-2}\Lambda(dx)$, and recall that we assume that $\Lambda(\{0\}) = 0$ to avoid a Kingman component. As discussed above, we may also assume that $\text{supp}(\Lambda) \subset [0, 1/2)$. The following definitions and properties of the upper-index β and of the lower-index δ of X can be found in [12] and [26].

Definition 18. The *upper index* is defined by

$$\beta := \inf \left\{ \alpha > 0 : \int_{|x| \leq 1} |x|^\alpha \nu(dx) < \infty \right\} \in [0, 2]. \tag{37}$$

To define the lower-index, following Pruitt [26], we introduce the function $h(x) = G(x) + K(x) + M(x)$, where (since in our setting $\text{supp}(\nu) \subset \mathbb{R}^+$ and moreover the drift is 0)

$$G(x) = \nu(y : y > x), \ K(x) = x^{-2} \int_{y \leq x} y^2 \nu(dy),$$

and

$$M(x) = x^{-1} \left| \int_{y \leq x} \frac{y^3}{1+y^2} \nu(dy) - \int_{y > x} \frac{y}{1+y^2} \nu(dy) \right|.$$

Definition 19. The *lower index* is defined by

$$\delta := \inf\{\alpha : \liminf_{x \to 0} x^\alpha h(x) = 0\}. \tag{38}$$

Note that the upper index β of (37) is similarly given by

$$\beta = \inf\{\alpha : \limsup_{x\to 0} x^\alpha h(x) = 0\}.$$

Therefore, it must be

$$0 \leq \delta \leq \beta \leq 2.$$

The constants β and δ characterize the asymptotic behavior of X near 0 (see (3.4) in Pruitt [26], and Figure 3). More precisely, if $M_t := \sup_{0\leq s\leq t} |X_s|$, then

$$\limsup_{t\to 0} M_t/t^\kappa = \begin{cases} 0 \text{ if } \kappa < 1/\beta \\ \infty \text{ if } \kappa > 1/\beta \end{cases}, \qquad \liminf_{t\to 0} M_t/t^\kappa = \begin{cases} 0 \text{ if } \kappa < 1/\delta \\ \infty \text{ if } \kappa > 1/\delta \end{cases}.$$

In this section we show the following result:

Proposition 20. *If the lower-index δ is strictly greater than 1, then for any $\varepsilon > 0$,*

$$\frac{N^\Lambda(t)}{t^{-1/(\beta+\varepsilon-1)}} \to \infty, \quad \text{in probability,}$$

and, for any $\varepsilon \in (0, \delta - 1)$

$$\frac{N^\Lambda(t)}{t^{-1/(\delta-\varepsilon-1)}} \to 0, \quad \text{in probability.}$$

Remark 21. When Grey's condition for extinction holds, we know (Lemma 22) that $\beta \geq 1$. However, by modifying the construction in Section 5, it is possible to find examples of Lévy measures ν such that both $\beta > 1$ and Grey's condition does not hold (that is, the corresponding coalescent does not come down from infinity). See the second to last paragraph of Section 5.

Informally speaking, the following lemma states that as $t \to 0$ the function $q \mapsto \psi(q)$ is of order at most q^β and at least q^δ.

Lemma 22. *For each $\varepsilon > 0$ small enough, there exist finite constants $c_{\varepsilon,\beta}$ and $c_{\varepsilon,\delta}$ such that for all v large enough $c_{\varepsilon,\delta} v^{\delta-\varepsilon} \leq \psi(v) \leq c_{\varepsilon,\beta} v^{\beta+\varepsilon}$. Hence if Λ is such that the Λ-coalescent comes down from infinity, then $\beta \geq 1$.*

Proof. Observe that for large q,

$$\psi(q) \asymp q^2 \int_{[0,1/q]} x^2 \nu(dx) + q \int_{[1/q,1]} x\, \nu(dx), \ q \to \infty, \tag{39}$$

where $f(q) \asymp g(q)$ means that both $f = O(g)$ and $g = O(f)$. Indeed, for $x \leq 1/q$ one can use Taylor's approximation to get $e^{-qx} - 1 + qx \in [q^2x^2/6, q^2x^2/2]$ while for $x \geq 1/q$ an easy computation shows $e^{-qx} - 1 + qx \in [qx/e, qx]$.

By definition (37), we have that $\int_{[0,1]} x^{\beta+\varepsilon}\nu(dx) < \infty$. Therefore

$$\sum_{n=0}^\infty e^{-(n+1)(\beta+\varepsilon)} \nu([e^{-n-1}, e^{-n}]) \ \leq \ \sum_{n=0}^\infty \int_{e^{-n-1}}^{e^{-n}} x^{\beta+\varepsilon}\nu(dx) = \int_{[0,1]} x^{\beta+\varepsilon}\nu(dx) < \infty.$$

In particular, there exists a constant $c > 0$ such that for all $n \geq 1$,

$$\nu([e^{-n-1}, e^{-n}]) \leq c\, e^{(n+1)(\beta+\varepsilon)}. \tag{40}$$

As a consequence, for $\epsilon < 2 - \beta$

$$\int_0^{1/q} x^2 \nu(dx) \quad \leq \quad \sum_{n=\lfloor \log q \rfloor}^{\infty} \int_{e^{-n-1}}^{e^{-n}} x^2 \nu(dx) \leq c \sum_{n=\lfloor \log q \rfloor}^{\infty} e^{(n+1)(\beta+\varepsilon)} e^{-2n} \leq c\, q^{\beta-2+\varepsilon},$$

where the finite positive constant c may change from one inequality to the next one. Similarly, one estimates

$$\int_{1/q}^{1} x\nu(dx) \leq c\, q^{\beta-1+\varepsilon}.$$

Together with (39), this yields the upper bound $\psi(q) = O(q^{\beta+\varepsilon})$.

For the lower bound, recall the definition (38) and related notation. Observe that

$$\int_{y>x} \frac{y}{1+y^2} \nu(dy) \asymp \int_{y>x} y\nu(dy), \;\; x \in (0,1).$$

The first integral in the definition of $M(x)$ is of order $\int_{y \leq x} y d\Lambda(y) = O(x)$, so it is negligible, in comparison. Also, note that as $x \to 0$

$$G(x) = \int_{y>x} \nu(dy) \leq x^{-1} \int_{y>x} y\nu(dy) \asymp M(x).$$

Combining this with the definition of $K(x)$ and (39) one gets

$$h(x) \asymp \psi(1/x), \;\; \text{as } x \to 0. \tag{41}$$

Due to (38), we have $h(x) \geq Cx^{-\delta+\varepsilon}$ for all x sufficiently small and for some $C > 0$, and the lower bound for ψ now easily follows.

Finally, assume that a given Λ-coalescent comes down from infinity. Then by Theorem 1, Grey's condition (3) is satisfied for the corresponding measure Λ. Since for each $\varepsilon > 0$, $\psi(q) \leq cq^{\beta+\varepsilon}$, we deduce that $\beta \geq 1$. □

Remark 23. Note that (41) also implies the stated upper bound on $\psi(v)$.

The asymptotic behavior of $\psi(q)$ as $q \to \infty$ induces the asymptotic behavior of $v(t)$ as $t \to 0$:

Corollary 24. *Assume that the Λ-coalescent comes down from infinity.*

(i) If $\beta \geq 1$, we have $\liminf_{t\to 0} t^{1/(\beta-1+\varepsilon)} v(t) = +\infty$ for any $\varepsilon > 0$.

(ii) If $\delta > 1$, then $\limsup_{t\to 0} t^{1/(\delta-\varepsilon-1)} v(t) = 0$ for any $\varepsilon \in (0, \delta - 1)$.

Proof. Recall(19) and let $c_{\varepsilon,\beta}$ and $c_{\varepsilon,\delta}$ be as in Lemma 22. It follows immediately that

$$\frac{v^{-(\beta+\varepsilon)+1}}{c_{\varepsilon,\beta}(\beta+\varepsilon-1)} \leq \int_v^{\infty} \frac{dq}{\psi(q)} \leq \frac{v^{-(\delta-\varepsilon)+1}}{c_{\varepsilon,\delta}(\delta-\varepsilon-1)}.$$

Note that since $\beta \geq 1$ we are able to integrate the lower bound for $1/\psi$, but we use the additional constraint $\delta > 1$ in order to be able to integrate the upper bound for $1/\psi$, and thus derive the right-hand-side of the above inequality. Setting the middle term to t, after rearranging, we obtain

$$\left(\frac{c_{\varepsilon,\beta}}{\beta+\varepsilon-1}\right)^{1/(\beta-1+\varepsilon)} \cdot t^{-1/(\beta-1+\varepsilon)} \leq v(t) \leq \left(\frac{c_{\varepsilon,\delta}}{\delta-\varepsilon-1}\right)^{1/(\delta-1-\varepsilon)} \cdot t^{-1/(\delta-1-\varepsilon)},$$

implying both statements. □

Proposition 15 and Corollary 24 together yield Proposition 20 (using Theorem 1 in [2] instead of Proposition 15 yields the same result in the stronger almost sure sense).

5 Appendix: an instructive example

In this section we discuss a class of examples that illustrate potential difficulties in analyzing functions ψ and v directly. In particular, we show that for some $\beta \neq \delta$, one can choose the measure Λ in such a way that $\psi(q)$ oscillates between q^δ and q^β, resulting in analogous oscillations for $v(t)$ between $t^{-1/(\beta-1)}$ and $t^{-1/(\delta-1)}$. This shows that the upper and lower bounds of Proposition 20 are sharp in general. As a bonus we provide examples of Λ-coalescents with $\delta = 1$ that come down from infinity. Let $\beta \in (1,2)$ be fixed. Set $a_n = e^{-n}$, $n \geq 0$ and for each $n \geq 0$ define the interval J_n as

$$J_n = (a_{n+1}, a_n].$$

For a subsequence $(a_{n_k})_{k\geq 0}$ of $(a_n)_{n\geq 0}$ define the measure

$$\nu(dx;(a_{n_k})_{k\geq 0}) \equiv \nu(dx) = \sum_k \mathbf{1}_{J_{n_k}}(x)\frac{1}{x^{\beta+1}}\,dx.$$

Then it is easy to check that for any choice of such a subsequence, the corresponding measure ν has the upper index β. It is moreover easily seen that if $n_k = k$ then (recalling (2) and (19)) $\psi(q) \asymp q^\beta$, $u(t) := \int_t^\infty dq/\psi(q) \asymp (1/t)^{\beta-1}$, as $t \to \infty$, and as a result $v(t) \asymp t^{-1/(\beta-1)}$, as $t \to 0$. The remaining calculations however confirm that if one chooses the intervals J_{n_k} sparse enough as $k \to \infty$, the asymptotic behavior of the functions ψ , u, and v can become quite irregular.

By (39), estimating $u(t)$ as $t \to \infty$ (up to constants) amounts to estimating

$$\int_t^\infty \frac{1}{q^2\int_{[0,1/q]} x^2\nu(dx) + q\int_{(1/q,1]} x\nu(dx)}\,dq.$$

Define $k^* = k^*(q) = \max\{k : a_{n_k} \geq 1/q\}$ and set $\beta_1 := \beta - 1$ and $\beta_2 := 2 - \beta$ so that

$\beta_1, \beta_2 > 0$. First compute

$$\begin{aligned}
\int_{[0,1/q]} x^2\, \nu(dx) &= \sum_k \int_{J_{n_k}\cap[0,1/q]} x^{1-\beta}\, dx \\
&= \frac{1}{\beta_2} \sum_{k: a_{n_k}<1/q} (a_{n_k}^{2-\beta} - a_{n_k+1}^{2-\beta}) + \sum_{k: a_{n_k}\geq 1/q > a_{n_k+1}} \int_{a_{n_k+1}}^{1/q} x^{1-\beta}\, dx \\
&= \frac{1}{\beta_2} \sum_{k=k^*+1}^{\infty} e^{-n_k(2-\beta)}(1-e^{-(2-\beta)}) + \int_{\exp(-n_{k^*}-1)\wedge 1/q}^{1/q} x^{1-\beta}\, dx \\
&= \frac{1}{\beta_2}\left((1-e^{-\beta_2}) \sum_{k=k^*+1}^{\infty} e^{-n_k\beta_2} + \frac{1}{q^{\beta_2}} - \frac{1}{(e^{n_{k^*}+1}\vee q)^{\beta_2}} \right),
\end{aligned} \tag{42}$$

and similarly

$$\int_{(1/q,1]} x\nu(dx) = \frac{1}{\beta_1}\left(\sum_{l=1}^{k^*-1} e^{n_l\beta_1}(e^{\beta_1}-1) + [(q\wedge e^{n_{k^*}+1})^{\beta_1} - e^{n_{k^*}\beta_1}] \right). \tag{43}$$

From now on assume $q \geq 1$ and let $k^* = k^*(q)$ be as defined above. Note that if $1/q \in J_{n_{k^*}}$ (meaning $n_{k^*} \leq \log(q) < n_{k^*}+1$) then

$$\begin{aligned}
&q^2 \cdot \left[\frac{1}{q^{\beta_2}} - \frac{1}{e^{(n_{k^*}+1)\beta_2}} \right] + q \cdot [q^{\beta_1} - e^{n_{k^*}\beta_1}] = \\
&q^2 \cdot \left[\frac{1}{q^{2-\beta}} - \frac{1}{e^{(n_{k^*}+1)(2-\beta)}} \right] + q \cdot [q^{\beta-1} - e^{n_{k^*}(\beta-1)}] \asymp q^\beta,
\end{aligned}$$

where for the last estimate it is best to consider separately the two cases $\log(q) \in [n_{k^*}, n_{k^*}+1/2)$ and $\log(q) \in [n_{k^*}+1/2, n_{k^*}+1)$. One can check similarly that (still assuming $1/q \in J_{n_{k^*}}$) the initial terms, corresponding to the non-negative series from (42) and (43), are of the order at most $q^{\beta-2}$ and $q^{\beta-1}$, respectively. Hence,

$$\psi(q) \asymp q^\beta,\ \tfrac{1}{q} \in \cup_k J_{n_k}, \tag{44}$$

which agrees well with the "regular" setting where $n_k = k$. If on the contrary, $1/q \notin \cup_k J_{n_k}$, then $n_{l-1}+1 \leq \log(q) < n_l$ for $l = k^*(q)+1$, so that computations (42) and (43) imply

$$\psi(q) \asymp c_1(\beta) q^2 \sum_{k=l}^{\infty} e^{-n_k(2-\beta)} + c_2(\beta) q \sum_{k=1}^{l-1} e^{n_k(\beta-1)},$$

where $c_i(\beta) \in (0,\infty)$, $i = 1,2$ are constants depending on β only. Due to the properties of the exponential function we then have

$$\psi(q) \asymp q^2 e^{-n_l(2-\beta)} + q e^{n_{l-1}(\beta-1)},\ \tfrac{1}{q} \in (a_{n_l}, a_{n_{l-1}+1}]. \tag{45}$$

Therefore, we need to estimate up to constants, for large t,

$$\sum_k \int_{[t,\infty)\cap[e^{n_k}, e^{n_k+1})} \frac{1}{q^\beta}\, dq + \sum_l \int_{[t,\infty)\cap[e^{n_{l-1}+1}, e^{n_l})} \frac{1}{q^2 e^{-n_l(2-\beta)} + q e^{n_{l-1}(\beta-1)}}\, dq. \tag{46}$$

The first series of integrals above can easily be evaluated as being of order

$$\sum_{n_k \geq \log t} e^{-n_k(\beta-1)} \asymp e^{-n_{k^*(t)}(\beta-1)}. \tag{47}$$

Using the formula

$$\int_a^b \frac{dx}{Bx + Cx^2} = \left[\frac{1}{B}\log\left|\frac{Cx}{Cx+B}\right|\right]_a^b,$$

for each l such that $\log t \leq n_{l-1} + 1$, the lth summand in the second series in (46) equals

$$\frac{1}{e^{n_{l-1}(\beta-1)}} \log\left|\frac{e^{n_l - n_{l-1}-1}(e^{n_l(\beta-2)}e^{n_{l-1}+1} + e^{n_{l-1}(\beta-1)})}{e^{n_l(\beta-1)} + e^{n_{l-1}(\beta-1)}}\right| \asymp \frac{(2-\beta)(n_l - n_{l-1})}{e^{n_{l-1}(\beta-1)}}, \tag{48}$$

since $\beta < 2$ and $n_{l-1} \leq n_l$.

Consider the following class of examples: for some $\varepsilon \geq 0$, define inductively $m_0 := 1$, $m_{r+1} := m_r + e^{\varepsilon m_r}$, for $r \geq 0$, and let

$$n_{j+1} := n_j + 1, \text{ whenever } n_j \in [m_{2r}, m_{2r+1}) \text{ for some } r \in \mathbb{N},$$

and otherwise (here it must be $n_j = m_{2r+1}$ for some $r \in \mathbb{N}$) define $n_{j+1} = m_{2r+2} = n_j + e^{\varepsilon n_j}$. In words, the strictly increasing sequence $(n_k)_{k\geq 1}$ looks like the simplest arithmetic progression over a long interval, then it makes a jump (if $\varepsilon > 0$ its size is huge in comparison to the current value of the sequence), and immediately after the sequence continues its slow increase by 1 unit at a time, until the next even larger jump, etc.

Now fix some $\varepsilon \in (0, \beta-1)$. Due to (46)–(48), the corresponding Λ-coalescent comes down from infinity, since

$$\sum_{l \geq 1} \frac{n_l - n_{l-1}}{e^{n_{l-1}(\beta-1)}} < \infty. \tag{49}$$

Consider first the case $1/t \in \cup_k J_{n_k}$, and more precisely let $1/t = e^{-n_j} = e^{-m_{2r}}$ for some $r \in \mathbb{N}$ (or equivalently, $\log t$ is just at the beginning of the rth long interval where n increases by increments of 1). Then $k^*(t) = j$ and so the expressions in (46) is of order

$$e^{-n_j(\beta-1)} + \sum_{l \geq j+1} \frac{(2-\beta)(n_l - n_{l-1})}{e^{n_{l-1}(\beta-1)}} \asymp \left(\frac{1}{t}\right)^{\beta-1} + \frac{(2-\beta)(m_{2r+2} - m_{2r+1})}{e^{m_{2r+1}(\beta-1)}} \asymp \left(\frac{1}{t}\right)^{\beta-1}.$$

The middle asymptotic was obtained by splitting the sum in l into two sums, one over the indices l satisfying $n_{l-1} \in \cup_s [m_{2s}, m_{2s+1})$ and the other over the indices l satisfying $n_{l-1} \in \{m_{2s+1} : s \in \mathbb{N}\}$. The first sum is easily seen to contribute another term of order $(1/t)^{\beta-1}$, while for the second sum the dominant term is given by the l for which $n_{l-1} = m_{2r+1}$. The final asymptotic is obtained by noting that due to the definition of the sequence $(m_s)_{s\geq 1}$, we have $(m_{2r+2} - m_{2r+1})/e^{m_{2r+1}(\beta-1)} = 1/e^{m_{2r+1}(\beta-1-\varepsilon)} = 1/e^{(m_{2r}+e^{\varepsilon m_{2r}})(\beta-1-\varepsilon)}$, and rewriting this last expression in terms of t as $\left(\frac{1}{te^{t^\varepsilon}}\right)^{\beta-1-\varepsilon)} = o\left(\frac{1}{t}\right)^{\beta-1}$, we conclude that for t of the form $t = e^{m_{2r}}$ we have

$$u(t) = \int_t^\infty \frac{1}{\psi(q)}\, dq \asymp \left(\frac{1}{t}\right)^{\beta-1},$$

as would be true for all t in the regularly varying case $\varepsilon = 0$.

We now focus on the opposite case $1/t \notin \cup_k J_{n_k}$, and in particular let us consider $t = e^{m_{2r+1}+1}$ for some $r \in \mathbb{N}$. Suppose that n_j is such that $n_j = m_{2r+1}$ and $n_{j+1} = m_{2r+1}+e^{\varepsilon m_{2r+1}}$. Then we have $k^*(t) = j$ and so it can be easily checked that the contribution of (47) to $u(t)$ is again of order $(1/t)^{\beta-1}$. However, the contribution of (48) to $u(t)$ is of order

$$\sum_{l\geq j+1} \frac{(2-\beta)(n_l - n_{l-1})}{e^{n_{l-1}(\beta-1)}} \asymp \frac{(2-\beta)(m_{2r+2} - m_{2r+1})}{e^{m_{2r+1}(\beta-1)}} = \frac{(2-\beta)}{e^{m_{2r+1}(\beta-1-\varepsilon)}} \asymp \left(\frac{1}{t}\right)^{\beta-1-\varepsilon} >> \left(\frac{1}{t}\right)^{\beta-1}.$$

So for t of the form $t = e^{m_{2r+1}+1}$we have $u(t) \asymp (1/t)^{\beta-1+\varepsilon}$.

The above class of examples can be generalized in the following way: instead of a fixed $\varepsilon \in (0, \beta-1)$, on could introduce a non-negative sequence $(\varepsilon_r)_{r\geq 1}$, redefine $m_0 := 1$, $m_{r+1} := m_r + e^{\varepsilon_r m_r}$, for $r \geq 0$, and keep the old definition of $(n_j)_{j\geq 1}$ in terms of $(m_r)_{r\geq 1}$.

Now if $\varepsilon_r = \beta - 1$ identically for all r, the corresponding coalescent does not come down from infinity, while we noted at the beginning of the section that the corresponding upper index is $\beta > 1$.

Similarly, if $\limsup_r \varepsilon_r = \beta - 1$ (and $\varepsilon_r < \beta - 1$, $\forall r$) where the terms close to $\beta - 1$ in the sequence $(\varepsilon_r)_r$ are sufficiently sparse so that (49) holds, then the corresponding coalescent comes down from infinity. However, (38), (41) and (44)–(45) imply that the lower index δ equals to 1, while the upper index is still $\beta > 1$.

References

[1] K.B. Athreya and P. Ney (1972). *Branching processes.* Springer, New York.

[2] J. Berestycki, N. Berestycki and V. Limic (2010). The Λ-coalescent speed of coming down from infinity. *Ann. Probab.*, 38, 207-233.

[3] J. Berestycki, N. Berestycki and J. Schweinsberg (2008). Small-time behavior of beta-coalescents. *Ann. Inst. H. Poincaré (B): Probab. Stat.* 44(2), 214–238.

[4] J. Berestycki, N. Berestycki and J. Schweinsberg (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.*, 35, 1835–1887.

[5] N. Berestycki (2009). *Recent progress in coalescent theory.* Ensaios Matematicos, Vol. 16.

[6] J. Bertoin (2006). *Random Fragmentation and Coagulation Processes.* Cambridge Studies in Advanced Mathematics.

[7] J. Bertoin and J.-F. Le Gall (2000). The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Theory Related Fields* **117**, 249–266.

[8] J. Bertoin and J.-F. Le Gall (2003). Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields* **126**, 261–288.

[9] J. Bertoin and J.-F. Le Gall (2005). Stochastic flows associated to coalescent processes II: Stochastic differential equations. *Ann. Inst. H. Poincaré Probabilités et Statistiques*, 41, 307–333.

[10] J. Bertoin and J.-F. Le Gall (2006). Stochastic flows associated to coalescent processes III: Limit theorems. *Illinois J. Math*, 50, 147–181.

[11] M. Birkner, J. Blath, M. Capaldo, A. Etheridge, M. Möhle, J. Schweinsberg, and A. Wakolbinger (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.*, 10, 303–325.

[12] R. M. Blumenthal and R. K. Getoor (1961). Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.*, 10, 493–516.

[13] P. Donnelly, and T. Kurtz (1996). A Countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.* **24**, 698–742.

[14] P. Donnelly, and T. Kurtz (1999). Particle Representations for Measure-Valued Population Models. *Ann. Probab.* **27**, 166–205.

[15] T. Duquesne and J.-F. Le Gall (2002). *Random Trees, Lévy Processes, and Spatial Branching Processes. Astérisque* **281**.

[16] R. Durrett (2004). *Probability: theory and examples.* Duxbury advanced series, 3$^{\text{rd}}$ edition.

[17] A.M. Etheridge (2000). *An introduction to superprocesses.* American mathematical society, University Lecture series (Vol. 20).

[18] D.R. Grey (1974). Asymptotic behavior of continuous-time continuous-state branching processes. *J. Appl. Prob.*, 11:669–677.

[19] J. Lamperti (1967). The limit of a sequence of branching processes. *Z. Wahrsch. Verw. Gebiete* **7**, 271–288.

[20] C. Labbé (2011). From flows of Lambda Fleming-Viot processes to lookdown processes via flows of partitions. ArXiv preprint 1107.3419v1.

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

[22] T.M. Liggett (1985). *Interacting particle systems.* Springer, New York.

[23] E. Perkins (1991). Conditional Dawson-Watanabe processes and Fleming-Viot processes, *Seminar in Stochastic Processes,* Birkhauser, 142–155.

[24] J. Pitman (1999). Coalescents with multiple collisions. *Ann Probab.* 27, 1870–1902.

[25] J. Pitman (2006). Combinatorial stochastic processes. *Lecture notes in mathematics, Ecole d'Eté de probabilités de Saint-Flour* XXXII-2002. Lecture Notes Math. **1875**, Springer.

[26] W. Pruitt (1981). The growth of random walks and Lévy processes. *Ann. Probab.*, 9(6):948–956.

[27] S. Sagitov (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl Prob.* 36:4, 1116–1125.

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