The Annals of Probability
2010, Vol. 38, No. 1, 207–233
DOI: 10.1214/09-AOP475

THE Λ-COALESCENT SPEED OF COMING DOWN FROM INFINITY

BY JULIEN BERESTYCKI, NATHANAËL BERESTYCKI AND VLADA LIMIC[1]

Université Pierre et Marie Curie-ParisVI, Cambridge University and C.N.R.S.

Consider a Λ-coalescent that comes down from infinity (meaning that it starts from a configuration containing infinitely many blocks at time 0, yet it has a finite number N_t of blocks at any positive time $t > 0$). We exhibit a deterministic function $v : (0, \infty) \to (0, \infty)$ such that $N_t/v(t) \to 1$, almost surely, and in L^p for any $p \geq 1$, as $t \to 0$. Our approach relies on a novel martingale technique.

1. Introduction. Various natural population genetics models lead to a representation of the genealogical tree by a process called Kingman's coalescent [16, 17]. Kingman's coalescent is a Markov process which can be informally described as follows: in a fixed sample of n individuals from the population, each pair of ancestral lineages coalesces at rate 1.

In population genetics, one uses the above process to quantify polymorphism in a homogeneously mixing population under neutral evolution. However, there is some evidence that for modeling evolution of marine populations (see, e.g., [19]), the use of coalescent processes which allow *multiple collisions* is more appropriate than that of Kingman's coalescent where only pairs of blocks can merge at any given time. Similarly, multiple collisions are natural for modeling evolution of viral populations, where natural selection plays a very strong role. They also emerge in the fine-scale mapping of disease loci [21].

A suitable family of mathematical models has been introduced and studied by Pitman [22] and Sagitov [26] under the name Λ-*coalescents* or *coalescents with multiple collisions*. We postpone the precise definitions of these processes until the next section.

Let $N^{\Lambda} \equiv N := (N_t, t \geq 0)$ be the number of blocks process corresponding to a particular Λ-coalescent process. In view of applications, we concentrate on Λ-coalescents such that $P(N_t < \infty, t > 0) = 1$ and $\lim_{t \to 0+} N_t = \infty$ (here N is really an entrance law). This property is typically referred to as coming down from infinity (see Section 2.2 for a formal definition). It is important to understand the

Received July 2008; revised February 2009.

[1]Supported in part by NSERC Discovery grant and by Alfred P. Sloan Research Fellowship.

AMS 2000 subject classifications. 60J25, 60F99, 92D25.

Key words and phrases. Exchangeable coalescents, small-time asymptotics, coming down from infinity, martingale techniques, fluid limits.

nature of divergence of N_t as t decreases to 0. In the current paper, our goal is to exhibit a function $v:(0,\infty)\to(0,\infty)$ such that

$$\lim_{t\to 0}\frac{N_t}{v(t)}=1 \qquad \text{almost surely.}$$

We call any such v the *speed* of coming down from infinity (speed of CDI) for the corresponding Λ-coalescent. Note that the limit above is in fact the limit as $t\to 0+$; from now on we always write $t\to 0$. The exact form of the function v is implicit and somewhat technical (see Theorem 1 for the precise statement). However, in many situations of interest, one can find a simpler function $g(t)$, often a power in t, such that $g(t)/v(t)\to 1$, and therefore $N_t/g(t)\to 1$, as $t\to 0$. Then we also refer to g as the speed of CDI for the corresponding coalescent. As mentioned above, Kingman's coalescent is the simplest Λ-coalescent. In particular, one can quickly find its speed of CDI by considering the "time-reversed" process. Analogous time-reversals for general Λ-coalescents seem to be difficult to grasp. The speed of CDI was recently determined for Beta-coalescents and their "perturbations" in Berestycki, Berestycki and Schweinsberg [4] and [5] and Bertoin and Le Gall [6] (where convergence is established in probability). See also the comment following the statement of Theorem 1 below.

With the above biology motivation in mind, there is a strong interest in understanding (see, e.g., [10, 13, 20]) analogues of Ewens' sampling formula for Λ-coalescents. It seems that only Kingman's coalescent allows for an exact solution (see, e.g., [12] or [14]) while in the general case, one should aim for good approximations. The only previous detailed analysis of this kind was carried out in [5] and [4] for the special case of Beta-coalescents. The above result can be viewed as the first step towards analogous understanding of the general Λ-coalescent case.

In a parallel work [3] we discuss the consequence of our main results to the problem of quantifying polymorphism in a population whose genealogy is driven by a coalescent with multiple collisions. In the same paper, we will describe a general connection between the small-time asymptotics of Λ-coalescents and continuous random trees and their associated continuous-state branching processes as well as generalized Fleming–Viot processes. These connections enable one to guess the form of function $v(t)$, and they imply the convergence in probability of the quantity $N_t/v(t)$ which is of interest under certain technical conditions. They can also be useful in determining the power law order of growth of v as $t\to 0$.

To the best of our knowledge, the martingale analysis in the current context is novel. We believe that it is of independent interest. Although similar in spirit, our setting is different from the general setting of Darling and Norris [8]. For their technique to apply, it is necessary to start with good bounds on the accumulated absolute difference of the "drifts" of the Markov chain and the solution to the corresponding differential "fluid-limit" equation. Here it seems difficult to obtain such bounds. However, it is possible to work directly [cf. the local martingale M'_z

from (22)] with the accumulated (nonabsolute) difference of the drifts in order to obtain sufficiently good asymptotic estimates.

The rest of the paper is organized as follows. Section 2 contains definitions and notations. The main results are stated in Section 3 and are proved in Section 4, with some technical estimates postponed until the Appendix.

2. Definitions and preliminaries.

2.1. *Notation.* We recall some standard notation, and introduce additional notation to simplify the exposition.

Denote the set of real (resp. rational) numbers by $\mathbb{R}$ (resp. $\mathbb{Q}$) and set $\mathbb{R}_+ = (0, \infty)$. For $a, b \in \mathbb{R}$, denote by $a \wedge b$ (resp. $a \vee b$) the minimum (resp. maximum) of the two numbers.
Let $\mathbb{N} := \{1, 2, \ldots\}$ and let $\mathcal{P}$ be the set of partitions of $\mathbb{N}$. Furthermore, for $n \in \mathbb{N}$ denote by $\mathcal{P}_n$ the set of partitions of $[n] := \{1, \ldots, n\}$.
If f is a function, defined in a left-neighborhood $(s - \varepsilon, s)$ of a point s, we denote by $f(s-)$ the left limit of f at s.
Given two functions $f, g : \mathbb{R}_+ \to \mathbb{R}_+$, write $f = O(g)$ if $\limsup f(x)/g(x) < \infty$, $f = o(g)$ if $\limsup f(x)/g(x) = 0$, and $f \sim g$ if $\lim f(x)/g(x) = 1$. The point at which the limits are taken might vary, depending on the context.
If X and Y are two random objects, we write $X \stackrel{d}{=} Y$ to indicate their equivalence in distribution. As usual, convergence in distribution will be denoted by $\Rightarrow$ symbol.
If $\mathcal{F} = (\mathcal{F}_t, t \geq 0)$ is a filtration, and T is a stopping time relative to $\mathcal{F}$, denote by $\mathcal{F}_T$ the standard filtration generated by T (see, e.g., [11], page 389).
For ν a finite or σ-finite measure, denote the support of ν by $\mathrm{supp}(\nu)$.

2.2. Λ-*coalescents.* Let Λ be a finite measure on $[0, 1]$. The Λ-coalescent is a Markov process $(\Pi_t, t \geq 0)$ with values in $\mathcal{P}$ (the set of partitions of $\mathbb{N}$), characterized as follows. If $n \in \mathbb{N}$, then the restriction $(\Pi_t^{(n)}, t \geq 0)$ of $(\Pi_t, t \geq 0)$ to $[n]$ is a Markov chain, taking values in $\mathcal{P}_n$, with a following dynamics: whenever $\Pi_t^{(n)}$ is a partition consisting of b blocks, the rate at which a given k-tuple of its blocks merges is

$$\lambda_{b,k} = \int_{[0,1]} x^{k-2}(1 - x)^{b-k} \Lambda(dx). \tag{1}$$

Note that mergers of several blocks into one block are possible, but multiple mergers do not occur simultaneously. For a generalization of Λ-coalescents where multiple mergers are possible, see Schweinsberg [29]. For a generalization of Λ-coalescents to spatial (not a mean-field) setting, see Limic and Sturm [18].

We will quote here several basic properties of the Λ-coalescent, and refer the reader to Pitman [22] for details and additional analysis. When $\Lambda(\{0\}) = 0$, the corresponding Λ-coalescent can be constructed via a Poisson point process in the following way. Let

$$\pi(\cdot) = \sum_{i \in \mathbb{N}} \delta_{t_i, x_i}(\cdot) \tag{2}$$

be a Poisson point process on $\mathbb{R}_+ \times (0, 1)$ with the intensity measure $dt \otimes \nu(dx)$ where $\nu(dx) = x^{-2}\Lambda(dx)$. Each atom (t, x) of π influences the evolution of the process Π as follows: for each block of $\Pi(t-)$, flip a coin with probability of heads equal to x; all the blocks corresponding to coins that come up "head" then merge immediately into one single block while all other blocks remain unchanged. Note that in order to make this construction rigorous, one first considers the restrictions $(\Pi^{(n)}(t), t \geq 0)$, since the measure $\nu(dx) = x^{-2}\Lambda(dx)$ may have infinite total mass.

We next recall a remarkable property of Λ-coalescents. Let E be the event that for all $t > 0$ there are infinitely many blocks, and let F be the event that for all $t > 0$ there are only finitely many blocks. Pitman [22] showed that, if $\Lambda(\{1\}) = 0$, only the following two types of behavior are possible, depending on the measure Λ: either $P(E) = 1$ or $P(F) = 1$. When $P(F) = 1$, the process Π is said to *come down from infinity*. For instance, Kingman's coalescent comes down from infinity, while if $\Lambda(dx) = dx$ is the uniform measure on $(0, 1)$, then the corresponding Λ-coalescent does not come down from infinity. This particular Λ-coalescent was discovered by Bolthausen and Sznitman [7] in connection with spin glasses.

A necessary and sufficient condition for a Λ-coalescent to come down from infinity was given by Schweinsberg [28]: define

$$\gamma_b = \sum_{k=2}^{b} (k-1) \binom{b}{k} \lambda_{b,k},$$

then the Λ-coalescent comes down from infinity if and only if $\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty$.

Recently, Bertoin and Le Gall [6] observed that this condition is equivalent to the following requirement: define

$$\psi_\Lambda(q) \equiv \psi(q) := \int_{[0,1]} (e^{-qx} - 1 + qx)\nu(dx), \tag{3}$$

where $\nu(dx) = x^{-2}\Lambda(dx)$, then

$$\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty \quad \text{if and only if} \quad \int_a^{\infty} \frac{dq}{\psi(q)} < \infty, \tag{4}$$

where the right-hand side is finite for some (and then automatically for all) $a > 0$. Somewhat remarkably, the divergence rate function v is given [cf. definition (8) in the next section] in terms of the right-hand side in (4). The condition (4) is well known in the Lévy processes literature as the Grey's criterion for extinction of the underlying continuous-state branching process. We refer the reader to [3] for further explanation of the above connections.

3. Main results. Let Λ be a finite measure on $[0, 1]$, and let $(\Pi_t, t \geq 0)$ be a Λ-coalescent. Without loss of generality, we may, and will, henceforth assume that Λ is a probability measure, that is,

$$\Lambda[0,1] = 1. \tag{5}$$

Indeed, a scaling of the total mass of Λ by a constant factor will induce the scaling of the speed of evolution (and therefore, that of coming down from infinity) by the same factor, and the speed of CDI v from (8) below will scale in the same way.

To each such measure Λ we associate a function ψ defined in (3). Moreover, for a probability measure $\tilde{\Lambda}$ of the form $\tilde{\Lambda} = (1-c)\Lambda + c\delta_0$, where Λ has no atom at 0, we may rewrite as

$$\psi_{\tilde{\Lambda}}(q) = \frac{c}{2}q^2 + (1-c)\int_{[0,1]} (e^{-qx} - 1 + qx)\nu(dx). \tag{6}$$

Note that if $c = 1$ we retrieve the Kingman coalescent, whose small-time behavior is well understood. Henceforth we assume that $c < 1$.

When ψ is such that the integral in (4) is finite, or equivalently, when the corresponding Λ-coalescent comes down from infinity, we can define

$$u_\psi(t) \equiv u(t) := \int_t^\infty \frac{dq}{\psi(q)} \in (0, \infty), \qquad t > 0, \tag{7}$$

and its càdlàg inverse

$$v_\psi(t) \equiv v(t) := \inf\left\{s > 0 : \int_s^\infty \frac{1}{\psi(q)}\,dq < t\right\}, \qquad t > 0. \tag{8}$$

Denote by $(N^\Lambda(t), t \geq 0) = (N_t^\Lambda, t \geq 0)$ the number of blocks process for the Λ-coalescent $(\Pi(t), t \geq 0)$. The first main result of this paper is following theorem.

THEOREM 1.

$$\lim_{t \to 0} \frac{N^\Lambda(t)}{v_\psi(t)} = 1 \qquad \textit{almost surely}. \tag{9}$$

Note that if Π does not come from infinity, both $N_t^\Lambda = N^\Lambda(t) = \infty$, for all $t \geq 0$, almost surely, and the formal definition (8) yields $v_\psi \equiv \infty$, so (9) extends trivially if $\infty/\infty = 1$.

We next comment on some special cases of Theorem 1. When $\Lambda = \delta_0$, we have $v(t) = 2/t$, and we recover the well-known result that for Kingman's coalescent, the number of blocks is almost surely asymptotic to $2/t$. Another interesting case occurs when Λ has the Beta$(2-\alpha, \alpha)$ distribution for some $1 < \alpha < 2$. That is,

$$\Lambda(dx) = \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} x^{1-\alpha}(1-x)^{\alpha-1}\,dx. \tag{10}$$

Here it is not hard to see that $\psi(q) \sim c_1 q^\alpha$ as $q \to \infty$, and thus that

$$v(t) \sim c_2 t^{-1/(\alpha-1)} \qquad \text{as } t \to 0,$$

where $c_1 = (\Gamma(\alpha)\alpha(\alpha-1))^{-1}$ and $c_2 = (\alpha\Gamma(\alpha))^{-1/(\alpha-1)}$. In fact these calculations can easily be generalized to the case where Λ is regularly varying near 0 with index $1 < \alpha < 2$. In this case, Theorem 1 strengthens Lemma 3 in [6].

However, we emphasize that the most delicate case of the above theorem occurs when the measure Λ "wildly oscillates" in any neighborhood of 0. An example of such a measure is constructed in the appendix of [3]. It illustrates potential difficulties in the analysis of functions ψ, u or v directly.

With a bit more work, we obtain as the second main result an analogue to Theorem 1 in terms of convergence of moments.

THEOREM 2. *For any $d \in [1, \infty)$,*

$$\lim_{s \to 0} E\left(\sup_{t \in [0,s]} \left| \frac{N^\Lambda(t)}{v_\psi(t)} - 1 \right|^d\right) = 0. \tag{11}$$

The following consequence of Theorem 1 says that, among all the Λ-coalescents such that $\Lambda[0,1] = 1$, Kingman's coalescent is extremal for the speed of coming down from infinity.

COROLLARY 3. *Assume* (5). *Then with probability* 1, *for any $\varepsilon > 0$, and for all t sufficiently small*,

$$N^\Lambda(t) \geq \frac{2}{t}(1 - \varepsilon).$$

PROOF. Without loss of generality assume that the Λ-coalescent comes down from infinity. To see how the corollary follows from Theorem 1, observe that since $e^{-qx} \leq 1 - qx + q^2x^2/2$ for $x > 0$,

$$\psi(q) \leq \frac{q^2}{2} \int_{[0,1]} x^2 \nu(dx) \leq \frac{q^2}{2} \qquad \text{[due to (5)]}. \tag{12}$$

Hence

$$u_\psi(s) \geq \int_s^\infty \frac{2}{q^2}\, dq = \frac{2}{s} \quad \text{and} \quad v_\psi(t) \geq \frac{2}{t}. \tag{13}$$

Due to Theorem 1, $N^\Lambda(t) \sim v_\psi(t)$ as $t \to 0$, implying that $N^\Lambda(t) \geq 2(1-\varepsilon)/t$ with high probability for all t small. $\square$

REMARK 4. It is interesting to compare the last result with the following fact shown in Angel et al. [1]:

$$\int_0^1 N^\Lambda(t)\, dt = \infty, \tag{14}$$

regardless of the choice of the finite measure Λ. Corollary 3 may be used to give an alternative proof of (14).

The following result is interesting from the perspective of applications in population genetics. More specifically, the total length of the coalescent tree is relevant for predicting the number of mutations in a large but finite sample. Assume (4), so that the coalescent comes down from infinity. Let $N^{\Lambda,n}$ denote the number of blocks process of the restriction $\Pi^{(n)}$ with initial state $\Pi_0^{(n)} = \{\{1\},\dots,\{n\}\}$ as defined at the beginning of Section 2.2. Let $\tau_n := \inf\{s > 0 : N^\Lambda(s) \le n\}$, and let $H_n := \{N^\Lambda(\tau_n) = n\}$ be the event that the (unrestricted) Λ-coalescent ever attains a configuration with exactly n blocks. Then, due to the strong Markov property, the conditional law of $(N^\Lambda(s+\tau_n), s \ge 0)$ given $\mathcal{F}_{\tau_n}$ on the event H_n, equals the law of $N^{\Lambda,n}$. Let $t_n = u_\psi(n)$ so that $v_\psi(t_n) = n$.

THEOREM 5. *For each $s > 0$ we have*

$$\lim_{n\to\infty} \frac{\int_0^s N^{\Lambda,n}(t)\,dt}{\int_0^s v_\psi(t_n+t)\,dt} = \lim_{n\to\infty} \frac{\int_0^s N^{\Lambda,n}(t)\,dt}{\int_0^s E(N^{\Lambda,n}(t))\,dt} = 1 \qquad \textit{in probability}.$$

For Kingman and Beta coalescents [*i.e.*, *when Λ is of the form $\Lambda = \delta_0$ or* (10) *with* $1 < \alpha < 2$], *the above convergence holds almost surely.*

Let $\tau_1^n = \inf\{t \ge 0 : N^{\Lambda,n}(t) = 1\}$, so that $\int_0^{\tau_1^n} N^{\Lambda,n}(t)\,dt$ equals the total length of the (Λ-)coalescent tree with n leaves. Moreover, for any fixed $s > 0$,

$$\int_s^{\tau_1^n} N^{\Lambda,n}(t)\,dt \to \int_s^{\tau_1} N^\Lambda(t)\,dt \qquad \text{almost surely}$$

(see Section 4.4) where the limit is a finite random variable. Hence the above theorem yields the asymptotics for the total length of the coalescent (genealogical) tree. Some more detailed analysis is postponed until [3].

Whereas Theorem 1 is a law of large numbers-type result for N^Λ, Theorem 5 is a law of large numbers-type result for $\int_0^{\tau_1^n} N^{\Lambda,n}(t)\,dt$. A central limit theorem for lengths of partial coalescent trees is obtained by Delmas, Dhersin and Siri-Jegousse [9] (see also [27]) for the Beta-coalescent case, similar questions for general Λ-coalescents remain open.

4. Martingale based arguments.

We now proceed toward the proof of Theorem 1. The following easy-to-check facts will be used in our analysis.

LEMMA 6. *The function $\psi : [0,\infty) \to \mathbb{R}_+$ of* (3) *is* (*strictly*) *increasing on* $[0,\infty)$, *and convex on* $(0,\infty)$. *Furthermore, for v_ψ, as in* (8), *we have $v'_\psi(s) = -\psi(v_\psi(s))$, so that v_ψ is decreasing with its derivative decreasing in absolute value.*

Due to Lemma 14, postponed until the next section, we can, and will, suppose without loss of generality that $\mathrm{supp}(\Lambda) \subset [0, 1/4]$. This assumption simplifies some technical estimates.

In this section we write N instead of N^{Λ} whenever not in risk of confusion, and we also abbreviate $v = v_{\psi}$. We start by observing that the function v is the unique solution of the following integral equation:

$$\log(v(t)) - \log(v(z)) + \int_z^t \frac{\psi(v(r))}{v(r)}\,dr = 0 \qquad \forall 0 < z < t, \tag{15}$$

with the "initial condition" $v(0+) = \infty$ [see Lemma 9 for properties of $\psi(q)/q$]. It is then natural to consider, for each fixed $z > 0$, the process

$$M(t) := \log(N(t)) - \log(N(z)) + \int_z^t \frac{\psi(N(r))}{N(r)}\,dr, \qquad t \ge z. \tag{16}$$

Let $n_0 \ge 1$ be fixed. Define

$$\tau_{n_0} := \inf\{s > 0 : N(s) \le n_0\}. \tag{17}$$

The following proposition tells us that $M(t \wedge \tau_{n_0})$ is "almost" (up to a bounded drift correction, and integrability condition) a martingale, with respect to the natural filtration $(\mathcal{F}_t, t \ge 0)$ generated by the underlying Λ-coalescent process. Its proof uses some general facts about binomial distributions, with precise statements and arguments postponed until the Appendix. In particular, in the rest of this section the parameter n_0 is taken to be the integer n_0 from Lemma 19.

As usual, $E[dX_s|\mathcal{F}_s]$ denotes the infinitesimal drift of a continuous-time process $(X_s, s \ge 0)$ with respect to the filtration $\mathcal{F}$ at time s. Similarly, we denote by $E[(dX_s)^2|\mathcal{F}_s]$ the corresponding infinitesimal second moment. That is,

$$\frac{E[dX_s|\mathcal{F}_s]}{ds} := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E[X_{s+\varepsilon} - X_s|\mathcal{F}_s]$$

and

$$\frac{E[(dX_s)^2|\mathcal{F}_s]}{ds} := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E[(X_{s+\varepsilon} - X_s)^2|\mathcal{F}_s].$$

PROPOSITION 7. *There exists some deterministic $C < \infty$ such that*

$$E[d\log(N(s))|\mathcal{F}_s] = \left(-\frac{\psi(N(s))}{N(s)} + h(s)\right) ds, \tag{18}$$

where $(h(s), s \ge z)$ is an $\mathcal{F}$-adapted process such that $\sup_{s \in [z, z \wedge \tau_{n_0}]} |h(s)| \le C$, and

$$E[[d\log(N(s))]^2|\mathcal{F}_s]\mathbf{1}_{\{s \le \tau_{n_0}\}} \le C\,ds \qquad \textit{almost surely.}$$

Both estimates are valid uniformly over $z > 0$.

PROOF. To prove the proposition, it suffices to show that for each $s > 0$, we have on $\{N(s) \geq n_0\}$,

$$\left|\frac{E(d\log(N(s))|\mathcal{F}_s)}{ds} + \frac{\psi(N(s))}{N(s)}\right| = |h(s)| = O\left(\int_{[0,1/4]} p^2 \nu(dp)\right) \tag{19}$$

and

$$E([d\log(N(s))]^2|\mathcal{F}_s) = O\left(\int_{[0,1/4]} p^2 \nu(dp)\right) ds, \tag{20}$$

where $O(\cdot)$ can be taken uniformly in s. Note that the finite integrals above are in fact taken over [0, 1], since $\nu(dx) = \nu(dx)\mathbf{1}_{\{x\in[0,1/4]\}}$ by assumption.

Recall the Poisson point process construction of Section 2.2 and fix $n \geq n_0$. If $\Lambda(\{0\}) = 0$, then on the event $\{N(s) = n\}$ an atom of size p arrives at rate $\nu(dp)\,ds$, and given that, $\log N(s) = \log n$ jumps to $\log(B_{n,p} + \mathbf{1}_{\{B_{n,p}<n\}})$ where $B_{n,p}$ has Binomial$(n, 1-p)$ distribution. Hence we have

$$E(d\log(N(s))|\mathcal{F}_s) = \int_{[0,1]} E\left[\log\frac{B_{n,p} + \mathbf{1}_{\{B_{n,p}<n\}}}{n}\right]\nu(dp)\,ds.$$

In the general case where $\Lambda(\{0\}) = c \in (0, 1)$, we have on the same event

$$\begin{aligned}\frac{E(d\log(N(s))|\mathcal{F}_s)}{ds} &= (1-c)\int_{[0,1]} E\left[\log\frac{B_{n,p} + \mathbf{1}_{\{B_{n,p}<n\}}}{n}\right]\nu(dp)\\ &\quad + c\binom{n}{2}\log\frac{n-1}{n}.\end{aligned}$$

Let $\psi_0(q) = q^2/2$ be the function ψ corresponding to the atomic $\Lambda(dx) = \delta_0(dx)$. Note that

$$c\binom{n}{2}\log\frac{n-1}{n} = -\frac{c}{2}n + \frac{c}{4} + O(1/n) = -c\frac{\psi_0(n)}{n} + \frac{c}{4} + O(1/n).$$

In view of (6), the estimate (19) will follow by Lemma 19 in the Appendix provided that

$$|np - 1 + (1-p)^n - (e^{-np} - 1 + np)| = |(1-p)^n - e^{-np}| \leq Cnp^2 \tag{21}$$

for all $n \geq n_0$, $p \leq 1/4$ and for some $C < \infty$. Note that $e^{-np} > (1-p)^n$, and, in fact,

$$e^{-np} - (1-p)^n = e^{-np}\left(1 - \exp\left\{-n\left(\frac{p^2}{2} + \frac{p^3}{3} + \cdots\right)\right\}\right).$$

Therefore, for $p \leq 1/4$ we have

$$\begin{aligned}1 - \exp\left\{-n\left(\frac{p^2}{2} + \frac{p^3}{3} + \cdots\right)\right\} &\leq 1 - \exp\left\{-\frac{n}{2}(p^2 + p^3 + \cdots)\right\}\\ &\leq 1 - \exp\left(-\frac{2}{3}np^2\right)\\ &\leq \frac{2}{3}np^2;\end{aligned}$$

hence both (21) and (19) hold.

To bound the infinitesimal variance on the event $\{N(s) = n\}$, use the second estimate in Lemma 19, together with the fact

$$\frac{E([d\log(N(s))]^2|\mathcal{F}_s)}{ds} \le (1-c)\int_{[0,1]} E\left[\log^2 \frac{B_{n,p} + \mathbf{1}_{\{B_{n,p}<n\}}}{n}\right]\nu(dp) + O\left(\frac{1}{n^2}\right)\binom{n}{2}.$$

Finally, note that both bounds (19) and (20) are uniform in the choice of z. $\square$

4.1. *Proof of Theorem* 1. Recall the process M from (16) and define

$$M'_z(t) \equiv M'(t) := M(t \wedge \tau_{n_0}) + \int_z^{t\wedge\tau_{n_0}} h(r)\,dr, \qquad t \ge z, \tag{22}$$

so that M'_z has martingale increments due to Proposition 7. A general property of the Doob–Meyer martingale correction (that one can check easily) implies that $E([dM'_z(t)]^2|\mathcal{F}_t) \le E[[d\log(N(t))]^2|\mathcal{F}_t]$, so that

$$E\big(M'_z(s) - M'_z(z)\big)^2 \le C(s-z) \qquad \forall 0 < z < s, \tag{23}$$

where C is the constant from Proposition 7.

Define a family of deterministic functions $(v_x, x \in \mathbb{R})$ by

$$v_x(t) = v(t+x), \qquad t \ge -x,$$

and note that each v_x satisfies an appropriate analogue of (15) on its entire domain, namely, $v_x(-x+) = \infty$ and

$$\log(v_x(t)) - \log(v_x(z)) + \int_z^t \frac{\psi(v_x(r))}{v_x(r)}\,dr = 0 \qquad \forall -x < z < t. \tag{24}$$

For each fixed $z > 0$ and each $x > -z$, define

$$M_{z,x}(t) := \log\frac{N(t)}{v_x(t)} - \log\frac{N(z)}{v_x(z)} + \int_z^t \left[\frac{\psi(N(r))}{N(r)} - \frac{\psi(v_x(r))}{v_x(r)} + h(r)\right]dr, \qquad t \ge z,$$

where h is given in (18).

Moreover, given $X \in \mathcal{F}_z$ such that $P(X > -z) = 1$, we can consider the process $M_{z,X}$. The advantage of this approach will be apparent soon.

For fixed $z > 0$, the processes M'_z, $M_{z,x}$ and $M_{z,X}$ are all adapted to the filtration $(\mathcal{F}_r, r \ge z)$.

REMARK 8. Strictly speaking, the processes $M'_z, M_{z,x}$ and $M_{z,X}$ defined above are *local* martingales (see [23] Chapter II or [24], Chapters VI, 31–34 for definition and first properties) since we do not know a priori whether $\log(N(t))$ has finite expectation. However, the optional stopping and Doob moment estimates that we apply below still hold in this more general setting.

LEMMA 9. *The function $q \mapsto \psi(q)/q$ is increasing.*

PROOF. Note that $q \mapsto \psi(q)/q$ is smooth, and that its derivative at q equals

$$\frac{\psi'(q)q - \psi(q)}{q^2} = \frac{\int(1-(xq+1)e^{-qx})\nu(dx)}{q^2}$$
$$= \frac{\int(1-(xq+1)e^{-qx})/x^2\Lambda(dx)}{q^2}.$$

It is a simple matter to check that the integrand in the numerator is positive for all $x > 0$, and that its limit as $x \to 0$ is $q^2/2$, so again it is positive.

The reader is invited to verify in a similar manner that $\lim_{q\to\infty}(\psi(q)/q)'' = -\int_{[0,1]} e^{-qx}x\Lambda(dx)$ which implies that $q \mapsto \psi(q)/q$ is asymptotically concave. Our argument does not make use of this fact. □

The following deterministic lemma is a crucial step in our analysis. It overcomes the need for a priori estimates necessary for the method of [8] to apply, as discussed in the Introduction.

LEMMA 10. *Suppose $f, g : [a,b] \mapsto \mathbb{R}$ are deterministic càdlàg functions such that*

$$\sup_{x\in[a,b]}\left|f(x) + \int_a^x g(u)\,du\right| \le c \tag{25}$$

for some $c < \infty$. If, in addition, $f(x)g(x) > 0$, $x \in [a,b]$ whenever $f(x) \ne 0$, then both

$$\sup_{x\in[a,b]}\left|\int_a^x g(u)\,du\right| \le c \quad \text{and} \quad \sup_{x\in[a,b]}|f(x)| \le 2c.$$

PROOF. Due to the assumptions, we know that at any point x where $f(x)$ is positive (resp. negative) $h(x) := \int_a^x g(u)\,du$ is increasing (resp. decreasing) from the right. Define $t_1 := \min\{x \in [a,b] : |h(x)| > c\}$, with the convention that $t_1 = b$ if this set is empty. Suppose $t_1 < b$. By continuity of h, it must be that $|h(t_1)| = c$. Without loss of generality, assume

$$h(t_1) = c \quad \text{and hence} \quad h(t_1 + \varepsilon) > c \tag{26}$$

for all small enough $\varepsilon > 0$. Having $f(t) < 0$ for all $t \in (t_1, t_1 + \varepsilon)$ would imply that h is decreasing on that same interval, contradicting (26). Therefore, there exists $t \in (t_1, t_1 + \varepsilon)$ such that $f(t) \geq 0$. But since $h(t) > c$ by (26), this would in turn contradict (25). Hence it must be $t_1 = b$, so that the uniform bound on $|h|$ holds, which together with (25) implies the uniform bound on $|f|$. $\square$

Since $N(t) \to \infty$ as $t \to 0$, almost surely, we have

$$P(\tau_{n_0} > 0) = 1 \quad \text{or equivalently} \quad \lim_{s \to 0} P(\tau_{n_0} \leq s) = 0.$$

Therefore, for any family $(Y_s, s > 0)$ of random variables, we have $\lim_{s \to 0, s \leq \tau_{n_0}} Y_s = \lim_{s \to 0} Y_s$, almost surely (in the sense that whenever one of the limits exists so does the other). Without loss of generality we will henceforth write $M_{z,x}(t)$ instead of $M_{z,x}(t \wedge \tau_{n_0})$, $t \in [z, s]$ instead of $t \in [z, s \wedge \tau_{n_0}]$, etc.

Fix $\alpha^* \in (0, 1/2)$. By Doob's L^2-inequality for martingales and (23) we have

$$\begin{aligned} P\Big(\sup_{t \in [z,s]} |M'_z(t) - M'_z(z)| > s^{\alpha^*}\Big) &\leq s^{-2\alpha^*} \sup_{t \in [z,s]} E\big[\big(M'_z(t) - M'_z(z)\big)^2\big] \\ &\leq s^{-2\alpha^*} C(s - z) = O(s^{1-2\alpha^*}). \end{aligned} \tag{27}$$

Denote by

$$A'_z(s) \equiv A'_z := \Big\{\sup_{t \in [z,s]} |M'_z(t) - M'_z(z)| \leq s^{\alpha^*}\Big\}$$

the complement of the above event. Henceforth we assume that $s < (1/C)^{1/(1-\alpha^*)}$. Note that then $\int_z^s h(r)\,dr \leq \int_z^s C\,dr \leq Cs \leq s^{\alpha^*}$. So we obtain that on A'_z [hence with probability greater than $1 - O(s^{1-2\alpha^*})$],

$$\sup_{t \in [z,s]} \left|\log N(t) - \log N(z) + \int_z^t \frac{\psi(N(r))}{N(r)}\,dr\right| \leq 2s^{\alpha^*}.$$

We conclude that $A'_z \subset A_z$, where

$$\begin{aligned} &A_z(s) \equiv A_z \\ &\quad := \left\{\sup_{t_1, t_2 \in [z,s]} \left|\log N(t_2) - \log N(t_1) + \int_{t_1}^{t_2} \frac{\psi(N(r))}{N(r)}\,dr\right| \leq 4s^{\alpha^*}\right\}. \end{aligned} \tag{28}$$

The advantage of the new definition is that $A_{z_1} \subset A_{z_2}$ whenever $z_1 \leq z_2 \leq s$. Moreover, the bound in (27) is uniform in $z \in (0, s)$, hence the decreasing property of probability measures implies

$$\begin{aligned} P\Big(\bigcap_{z \in (0,s)} A_z\Big) &= P\left(\sup_{t_1, t_2 \in (0,s]} \left|\log N(t_2) - \log N(t_1) + \int_{t_1}^{t_2} \frac{\psi(N(r))}{N(r)}\,dr\right| \leq 4s^{\alpha^*}\right) \\ &= 1 - O(s^{1-2\alpha^*}). \end{aligned}$$

Let X_z be the random variable defined by

$$N(z) = v(X_z + z) = v_{X_z}(z). \tag{29}$$

LEMMA 11. *We have* $\lim_{z\to 0} X_z = 0$, *almost surely.*

PROOF. Since N is nonincreasing and v is (strictly) decreasing, it is easy to see that $(X_z + z, z > 0)$ is also a nondecreasing process, almost surely. Therefore $\lim_{z\to 0} X_z + z \geq 0$ exists, almost surely. Moreover, the above limit equals 0 with probability 1, since $X_z + z = u(N(z))$, and since $P(N(0+) = \infty) = 1$ and $\lim_{x\to\infty} u(x) = 0$. □

Due to (24) and (28), we have, in particular, that

$$A_z = \left\{ \sup_{t_1,t_2\in[z,s]} \left| \log \frac{N(t_2)}{v_{X_z}(t_2)} - \log \frac{N(t_1)}{v_{X_z}(t_1)} \right.\right.$$
$$\left.\left. + \int_{t_1}^{t_2} \left[\frac{\psi(N(r))}{N(r)} - \frac{\psi(v_{X_z}(r))}{v_{X_z}(r)} \right] dr \right| \leq 4s^{\alpha^*} \right\}.$$

After plugging in $t_1 = z$, we obtain

$$A_z \subset \left\{ \sup_{t\in[z,s]} \left| \log \frac{N(t)}{v_{X_z}(t)} + \int_z^t \left[\frac{\psi(N(r))}{N(r)} - \frac{\psi(v_{X_z}(r))}{v_{X_z}(r)} \right] dr \right| \leq 4s^{\alpha^*} \right\}.$$

Lemma 9 implies the hypotheses of Lemma 10 omega-by-omega (with $a = z$, $b = s$ and the obvious choice of f and g), therefore

$$A_z(s) = A_z \subset \left\{ \sup_{t\in[z,s]} \left| \log \frac{N(t)}{v_{X_z}(t)} \right| \leq 8s^{\alpha^*} \right\}. \tag{30}$$

By fixing $t < s$ and varying $z \in (0, t]$ [note that $\log \frac{v_{X_z}(t)}{v_{X_{z'}}(t)} = \log \frac{N(t)}{v_{X_{z'}}(t)} - \log \frac{N(t)}{v_{X_z}(t)}$] we obtain

$$\bigcap_{z\in(0,s)} A_z(s) \subset \left\{ \sup_{z,z'\in(0,s),t\in[z\vee z',s]} \left| \log \frac{v_{X_z}(t)}{v_{X_{z'}}(t)} \right| \leq 16s^{\alpha^*} \right\},$$

which together with (30) implies

$$\bigcap_{z\in(0,s)} A_z(s) \subset \left\{ \sup_{t\in(0,s]} \left| \log \frac{N(t)}{\lim_n v_{X_{z_n}}(t)} \right| \leq 24s^{\alpha^*} \right\},$$

where $(z_n)_{n\geq 1}$ is any given deterministic sequence of strictly positive numbers converging to 0. Due to Lemma 11, the continuity of v implies $\lim_{n\to\infty} v_{X_{z_n}}(t) = v(t)$, $\forall t \in (0, s]$, almost surely. To summarize, we have just proved:

PROPOSITION 12. *If* $\operatorname{supp}(\Lambda) \subset [0, 1/4]$, *then*

$$P\left(\sup_{t\in(0,s\wedge\tau_{n_0}]}\left|\log\frac{N(t)}{v(t)}\right| \le 24s^{\alpha^*}\right) \ge P\left(\bigcap_{z\in(0,s)} A_z(s)\right) = 1 - O(s^{1-2\alpha^*}).$$

Theorem 1 now follows due to the Borel–Cantelli lemma, after choosing a deterministic sequence $(s_m)_{m\ge1}$ of strictly positive numbers converging to 0 sufficiently fast so that $\sum_m (s_m)^{1-2\alpha^*} < \infty$.

REMARK 13. The fixed scale assumption $\Lambda[0,1](=\Lambda[0,1/4]) = 1$ has not been used in the above argument.

4.2. *Relaxing assumptions on* $\operatorname{supp}(\Lambda)$. Given a probability measure Λ on $[0, 1]$ and a positive $\eta \le 1$, define its restriction Λ_η by

$$\Lambda_\eta(dx) = \Lambda(dx)\mathbf{1}_{[0,\eta]}(dx).$$

For each $\eta \in (0, 1]$, denote by ψ_η the function ψ_{Λ_η} that corresponds to Λ_η [cf. (3)], and by v_η the corresponding rate function from (8).

LEMMA 14. *All the* Λ_η*-coalescents, where* $\eta \in (0, 1]$, *have the same speed of CDI. Moreover, for any fixed* $\eta \in (0, 1)$,

$$\lim_{t\to0}\frac{v(t)}{v_\eta(t)} = 1, \tag{31}$$

so it suffices to prove Theorem 1 *for one* $\eta \in (0, 1)$ *in order to prove it for all* $\eta \in (0, 1]$.

PROOF. Fix $\eta \in (0, 1)$. Assume first that $\Lambda(\{0\}) = 0$. Then it is easy to see that one can find a coupling of the two coalescent processes defined by Λ and by Λ_η, respectively, such that the corresponding coalescent block counting processes N^Λ and N^{Λ_η} coincide for all $t \in (0, T_\eta)$ where $P(T_\eta > 0) = 1$. Namely, recall the PPP construction of Section 2.2 and set $T_\eta := \min\{t > 0 : (t, p)$ is an atom of π and $p > \eta\}$.

If $\Lambda(\{0\}) > 0$, let $\Lambda'(dx) = \Lambda(dx)\mathbf{1}_{(0,1)}(x)$, and note that the PPP-based construction of Λ'-coalescent can be enriched by superimposing pairwise coalescent events at rate $\Lambda(\{0\})$ thus yielding a construction of Λ-coalescent. Again, one can couple such constructions of Λ-coalescent and Λ_η-coalescent so that the two processes agree until T_η as discussed above.

To prove the lemma, it now suffices to show (31) for any fixed $\eta \in (0, 1)$. Note that we trivially have $v(t) \le v_\eta(t)$ for all $t > 0$, since $\psi_\eta(q) \le \psi(q)$ for all $q > 0$. Moreover,

$$\psi_\eta(q) = \psi(q) - a_\eta q + b_\eta + O(e^{-q\eta}),$$

where $a_\eta := \int_{(\eta,1]}(1/x)\Lambda(dx)$ and $b_\eta := \int_{(\eta,1]}(1/x^2)\Lambda(dx)$. Therefore, for any $0 \le z \le t$,

$$\log\frac{v(t)}{v_\eta(t)} - \log\frac{v(z)}{v_\eta(z)} + \int_z^t \left[\frac{\psi(v(r))}{v(r)} - \frac{\psi(v_\eta(r))}{v_\eta(r)} + h_z(r)\right] dr = 0,$$

where $h_z(r)$ is now a deterministic function, bounded by a fixed constant C, uniformly over z. The rest of the argument is a deterministic (and easier) analogue of the argument given in Section 4.1. We leave it to an interested reader. □

If $\Lambda(\{0\}) > 0$, then the size of the atom at 0 determines the speed of CDI. More precisely, we have:

COROLLARY 15. *If* $\Lambda(\{0\}) = c > 0$, *then for all* $\eta \in (0, 1]$,

$$v_\eta(t) \sim \frac{2}{ct}, \qquad t \to 0.$$

PROOF. Denote by v_0 the above function $2/(ct)$ and note that it corresponds to $\Lambda(dx) = c\delta_0(dx)$ and $\psi_0(q) = \frac{cq^2}{2}$, in terms of (8). Next note that if $\eta \in (0, 1]$, then

$$\psi_\eta(q) = \frac{cq^2}{2} + f(q) = \psi_0(q) + f(q),$$

where $f(q) = o(q^2)$ is a nonnegative function. In particular, $v_\eta(t) \le v_0(t)$, $t > 0$. Moreover, since for any ε, we can find $q(\varepsilon) < \infty$, such that

$$\psi_\eta(q) \le \frac{c(1+\varepsilon)q^2}{2} \qquad \text{for all } q \ge q(\varepsilon).$$

We have by the same reasoning, $v_\eta(t) \ge v_0(t)/(1+\varepsilon)$ for all sufficiently small t. Letting $\varepsilon \to 0$ implies the statement. □

4.3. *Proof of Theorem* 2. Assume that the parameter n_0 is the maximum of the corresponding quantities from Lemmas 19 and 20. Assume initially that $\mathrm{supp}(\Lambda) \subset [0, 1/4]$ and fix $z > 0$. With the notation of Section 4.1 in mind, let $M_{z,X_z} \equiv M$ be the process given by

$$\begin{aligned} M_t &:= \log\frac{N(t\wedge\tau_{n_0})}{v(X_z + t\wedge\tau_{n_0})} \\ &\quad + \int_z^{t\wedge\tau_{n_0}} \left(\frac{\psi(N(r))}{N(r)} - \frac{\psi(v(X_z+r))}{v(X_z+r)} + h(r)\right) dr, \qquad t \ge z. \end{aligned} \tag{32}$$

Then $M_z = 0$, and due to Proposition 7, M is a martingale (in the sense that M_t is an integrable random variable, $t \ge z$). Note that here we use M as abbreviation; the above process should not be confounded with M from (16).

We next obtain better estimates on the tails of the distribution of M_t, via an analogue of Hoeffding's inequality [15] for discrete martingale sums. Since M has only downward jumps, a simple case of a general result of Barlow, Jacka and Yor ([2], Proposition 4.2.1; see also [25]) implies that for any $c > 0$,

$$S^{(c)} := \left(\exp\left\{cM_t - \frac{c^2 C(t-z)}{2}\right\}, t \geq z\right),$$

is a supermartingale started from $S_z^{(c)} = 1$, with respect to the usual filtration $\mathcal{F}$. Note that D_t in [2, 25] corresponds to $E[(dM_t)^2|\mathcal{F}_t]$ in our notation, and that C is the uniform upper bound from Proposition 7.

Fix some $x \in \mathbb{R}_+$. Let $c = x/(C(s-z))$, and $y = \exp\{cx/2\} = \exp\{cx - c^2 C(s-z)/2)\}$, and let $T_y = \inf\{t \geq z : S_t^{(c)} > y\}$. Since $S^{(c)}$ only has downward jumps, it must be $S_{T_y}^{(c)} = y$ on $\{T_y < \infty\}$. Since $S^{(c)}$ is supermartingale, using optional stopping at $T_y \wedge s$, we have

$$\begin{aligned} 1 = E(S_z^{(c)}) &\geq E(S_{T_y \wedge s}^{(c)}) \\ &= yP(T_y \leq s) + E(S_s^{(c)} \mathbf{1}_{T_y > s}) \\ &\geq yP(T_y \leq s). \end{aligned}$$

It follows that

$$\begin{aligned} P\Big(\sup_{t\in[z,s]} M_t > x\Big) &\leq P\Big(\sup_{t\in[z,s]} S_t^{(c)} > e^{cx - c^2 C(s-z)/2}\Big) \\ &\leq P(T_y \leq s) \\ &\leq \frac{1}{y} = \exp\left\{-\frac{x^2}{2C(s-z)}\right\}. \end{aligned}$$

In order to obtain the "left tails" we use [2] Proposition 4.2.1 in a less trivial sense. If $c > 0$, then

$$S^{(-c)} := \left(\exp\left\{-cM_t - \frac{c^2 C(t-z)}{2} - \frac{c^2}{2}\sum_{s\leq t}(\Delta_s M)^2\right\}, t \geq z\right),$$

is a supermartingale where $\Delta_s M = M(s) - M(s-) = \Delta_s \log N(s \wedge \tau_{n_0})$. Define

$$\begin{aligned} E^{(c)}(t) &:= \exp\left\{c \sum_{t\in[z,s]} (\Delta_t M)^2 - e^{9c/4} K_0(t-z)\right\} \\ &= \exp\left\{c \sum_{t\in[z,s]} (\Delta_t \log N(s\wedge \tau_{n_0}))^2 - e^{9c/4} K_0(t-z)\right\}, \end{aligned}$$

where K_0 is the constant from Lemma 20. Due to Lemma 20, we have that for each $c > 0$, the process $(E^{(c)}(t), t \geq z)$ is a nonnegative super-martingale started from

$E^{(c)}(z) = 1$. Indeed, it is easy to verify in the sense of calculations of Proposition 7 that

$$\begin{aligned}
&E\big(dE^{(c)}(t)|\mathcal{F}_t\big)\\
&\quad = E^{(c)}(t)\cdot E[\exp\{c(\Delta_t M)^2\} - 1|\mathcal{F}_t] - e^{9c/4}K_0 \cdot E^{(c)}(t)\,dt\\
&\quad \le E^{(c)}(t)\cdot\left[\sum_{n\ge n_0}\mathbf{1}_{\{N(t)=n\}}\int_{[0,1/4]}(e^{9c/4}K_0p^2)/p^2\Lambda(dp) - e^{9c/4}K_0\right]dt\\
&\quad = 0,
\end{aligned}$$

almost surely. To include the case $\Lambda(\{0\}) > 0$ in the above calculation, note that by a standard estimate (50) and Taylor's series expansion,

$$\binom{n}{2}\big(\exp\{c\log^2\big((n-1)/n\big)\} - 1\big) = \frac{c}{2} + O\left(\frac{c}{n} + \frac{e^c}{n^2}\right).$$

Without loss of generality one can assume that both $K_0 \ge 1$ and $c/2 + O(c/n + e^c/n^2) \le e^{9c/4}$ for $n \ge n_0$ and all $c > 0$.

Then for $x > 0$, we have

$$\begin{aligned}
&P\Big(\inf_{t\in[z,s]} M_t < -x\Big)\\
&\quad \le P\left(\inf_{t\in[z,s]} M_t < -x, c^2\sum_{t\in[z,s]}(\Delta_s M)^2 \le cx\right)\\
&\qquad + P\left(\sum_{t\in[z,s]}(\Delta_t M)^2 > x/c\right)\\
&\quad \le P\Big(\sup_{t\in[z,s]} S_t^{(-c)} > e^{cx/2 - c^2C(s-z)/2}\Big)\\
&\qquad + P\Big(\sup_{t\in[z,s]} E^{(c^2)}(t) > e^{xc - e^{9c^2/4}K_0(s-z)}\Big)\\
&\quad \le e^{-cx/2 + c^2C(s-z)/2} + e^{-xc + e^{9c^2/4}K_0(s-z)}.
\end{aligned}$$

We plug in $c = \frac{2}{3}\sqrt{\log[x/(K_0(s-z))]}$ [here we assume that $x > 2K_0(s-z)$]. Since in each exponent the second term is negligible when compared to the first, we get the sub-exponential estimate

$$P\Big(\inf_{t\in[z,s]} M_t < -x\Big) = O\big(r(x; s-z)\big),$$

where

$$r(x; s) := \exp\big\{-x\sqrt{\log[x/(K_0 s)]}/4\big\}.$$

Now another omega-by-omega application of Lemmas 9 and 10 yields

$$1 - O\big(r(x; s - z)\big) \le P\Big(\sup_{t\in[z,s]} |M_t| \le x\Big)$$
$$\le P\Big(\sup_{t\in[z,s]}\Big|\log \frac{N(t \wedge \tau_{n_0})}{v(X_z + t \wedge \tau_{n_0})}\Big| \le 2(x + Cs)\Big).$$

Since $\lim_{z\to 0} v(X_z + t) = v(t)$ as argued before, in the limit we obtain

$$P\Big(\sup_{t\in[0,s]}\Big|\log \frac{N(t \wedge \tau_{n_0})}{v(t \wedge \tau_{n_0})}\Big| \le 2(x + Cs)\Big) \ge 1 - O(r(x; s)). \tag{33}$$

Note that since N is an integer-valued process and v is a decreasing function, $\inf_{t\in[0,s]} \log(N(t)/v(t)) \ge \inf_{t\in[0,s\wedge\tau_{n_0}]} \log(N(t)/v(t)) - \log n_0$, almost surely. Now (33) together with the observation $N(t) \le N(t \wedge \tau_{n_0})$ implies that the random variable

$$\Xi_s := \sup_{t\in[0,s]}\Big|\log \frac{N(t)}{v(t)}\Big| = \log\Big(\sup_{t\in[0,s]}\Big|\frac{N(t)}{v(t)}\Big| \vee \sup_{t\in[0,s]}\Big|\frac{v(t)}{N(t)}\Big|\Big)$$

satisfies $P(\Xi_s > x) = O(r(x; s))$, hence

$$P\Big(\sup_{t\in[0,s]}\Big|\frac{N(t)}{v(t)}\Big| \ge y\Big) \le O\Big(\frac{1}{y^{\sqrt{\log\log(y)-\log(K_0 s)}/4}}\Big) \qquad \text{as } y \to \infty.$$

In particular, for any $d \ge 1$, we can find a constant $D(d) < \infty$ such that

$$E\Big(\sup_{t\in[0,s]}\Big|\frac{N(t)}{v(t)}\Big|^d\Big) < D(d), \tag{34}$$

hence (for a possibly different constant) $E(\sup_{t\in[0,s]}|N(t)/v(t) - 1|^d) < D(d)$. Now the almost sure convergence of Theorem 1 combined with an application of dominated convergence theorem completes the argument.

For the case of general $\mathrm{supp}(\Lambda)$, recall the notation of Section 4.2. In addition, denote by $N_{1/4}(t)$ the number of blocks process corresponding to $\Lambda_{1/4}$. Due to the coupling construction used in the argument of Lemma 14, we have

$$N_{1/4}(t) \ge N(t), \qquad t \ge 0,$$

and moreover,

$$\sup_{t\in[0,s]} \frac{v_{1/4}(t)}{v(t)} < \infty.$$

Therefore estimate (34), established for the $\Lambda_{1/4}$-coalescent, will imply the same estimate [with possibly different constant $D(d)$] for the Λ-coalescent.

4.4. *Proof of Theorem* 5. Recall the notation $t_n = u_\psi(n) = u(n)$ introduced before the statement of Theorem 5. It suffices to show that any subsequence $(n_k)_{k\geq 1}$ contains a further subsequence $(n_{k(j)})_{j\geq 1}$ such that

$$\lim_{j\to\infty} \frac{\int_0^s N^{\Lambda,n_{k(j)}}(t)\,dt}{\int_0^s v(t_{n_{k(j)}}+t)\,dt} = 1 = \lim_{j\to\infty} \frac{\int_0^s N^{\Lambda,n_{k(j)}}(t)\,dt}{\int_0^s E(N^{\Lambda,n_{k(j)}}(t))\,dt} \quad \text{almost surely.} \tag{35}$$

For $t \geq 0$, define

$$\begin{aligned} M_t^n &:= \log \frac{N^{\Lambda,n}(t\wedge \tau_{n_0}^n)}{v(t_n + t\wedge \tau_{n_0}^n)} \\ &\quad + \int_0^{t\wedge\tau_{n_0}^n} \left(\frac{\psi(N^{\Lambda,n}(r))}{N^{\Lambda,n}(r)} - \frac{\psi(v(t_n+r))}{v(t_n+r)} + h^n(r) \right) dr, \end{aligned} \tag{36}$$

where h^n is the drift compensator of $\log(N^{\Lambda,n})$ with respect to the filtration generated by the underlying Λ-coalescent and where

$$\tau_{n_0}^n := \inf\{s > 0 : N^{\Lambda,n}(s) \leq n_0\}.$$

Then M^n in (36) is a direct analogue of martingale (32). In particular, note that by definition of t_n, $M_0^n = 0$, and as in (23),

$$E((M_t^n)^2) \leq Ct.$$

Recall τ_{n_0} defined in (17), and note that with probability 1, $\tau_{n_0}^n$ increases to τ_{n_0} as $n\to\infty$. The arguments leading to Proposition 12 apply in the current setting to yield for a fixed $\alpha^* < 1/2$, and for all n (for $n \leq n_0$ the result holds trivially),

$$P\left(\sup_{t\in[0,s]} \left|\log \frac{N^{\Lambda,n}(t\wedge\tau_{n_0}^n)}{v(t_n+t\wedge\tau_{n_0}^n)}\right| \leq 24 s^{\alpha^*}\right) \geq 1 - O(s^{1-2\alpha^*}) \quad \text{and} \tag{37}$$

$$P\left(\sup_{t\in[0,s]} \left|\log \frac{N^{\Lambda,n}(t\wedge\tau_{n_0}^n)}{v(t_n+t\wedge\tau_{n_0}^n)}\right| \leq 2(x+Cs)\right) \geq 1 - O(r(x;s)). \tag{38}$$

Fix some subsequence $(n_k)_{k\geq 1}$. We now show the first convergence statement in (35). Choose any sequence s_j of positive numbers decreasing to 0 so that

$$\sum_j s_j^{1-2\alpha^*} < \infty. \tag{39}$$

Next choose a further subsequence of $(n_k)_{k\geq 1}$, denoted again by $(n_j)_{j\geq 1}$ to simplify notation, so that

$$\begin{aligned} &\lim_{j\to\infty} \int_0^{s_j} v(t_{n_j}+t)\,dt = \infty, \\ &\lim_{j\to\infty} \frac{\int_{s_j}^s v(t_{n_j}+t)\,dt}{\int_0^{s_j} v(t_{n_j}+t)\,dt} = \lim_{j\to\infty} \frac{\int_{s_j}^s N^{\Lambda,n_j}(t)\,dt}{\int_0^{s_j} N^{\Lambda,n_j}(t)\,dt} = 0, \end{aligned} \tag{40}$$

where the last limit is taken almost surely. Note that here we use observations (13) and (14) and the following straightforward facts: for any fixed $0 \le a < b \le s$, $\int_a^b v(t_{n_j} + t)\,dt \uparrow \int_a^b v(t)\,dt$ and $\int_a^b N^{\Lambda,n_j}(t)\,dt \uparrow \int_a^b N^{\Lambda}(t)\,dt$. Due to (37), (39) and the Borel–Cantelli lemma, we have

$$(41)\qquad \lim_{j\to\infty} \sup_{t\in[0,s_j]} \left| \frac{N^{\Lambda,n_j}(t)}{v(t_{n_j}+t)} - 1 \right| = 0 \qquad \text{almost surely.}$$

The first statement in (35) now follows by a simple calculus fact: if $(f_n)_{n\ge1}$, $(g_n)_{n\ge1}$, $f_n, g_n : [0,s] \to [0,\infty)$, are two sequences of integrable functions such that for some positive sequence $\delta_n \to 0$ it is true that

$$\lim_{n\to\infty} \int_0^{\delta_n} f_n(t)\,dt = \infty, \qquad \lim_{n\to\infty} \frac{\int_{\delta_n}^s f_n(t)\,dt}{\int_0^{\delta_n} f_n(t)\,dt} = \lim_{n\to\infty} \frac{\int_{\delta_n}^s g_n(t)\,dt}{\int_0^{\delta_n} g_n(t)\,dt} = 0,$$

and

$$\lim_{n\to\infty} \sup_{t\in[0,\delta_n]} \left| \frac{f_n(t)}{g_n(t)} - 1 \right| = 0,$$

then

$$\lim_{n\to\infty} \frac{\int_0^s f_n(t)\,dt}{\int_0^s g_n(t)\,dt} = 1.$$

For the second convergence statement in (35), note that (similar to the argument for Theorem 2), almost sure convergence (41) together with estimate (38) and the dominated convergence theorem, yield

$$(42)\qquad \lim_{j\to\infty} \sup_{t\in[0,s_j]} \left| \frac{EN^{\Lambda,n_j}(t)}{v(t_{n_j}+t)} - 1 \right| = 0 \qquad \text{almost surely.}$$

Note that without loss of generality we may assume that

$$(43)\qquad \lim_{j\to\infty} \frac{\int_{s_j}^s EN^{\Lambda,n_j}(t)\,dt}{\int_0^{s_j} EN^{\Lambda,n_j}(t)\,dt} = 0.$$

The previous argument applies.

The final statement of Theorem 5 will follow from Corollary 16, which is stated and proved in next subsection.

4.4.1. *Discussion on almost sure convergence*. It is an open question whether the convergence of Theorem 5 holds almost surely. Our technique seems too crude to verify it in general, yet we offer below a partial result in this direction. One standard approach would be to use the monotonicity

$$\int_0^s N^{\Lambda,n}(t)\,dt \le \int_0^s N^{\Lambda,n+1}(t)\,dt \quad \text{and} \quad \int_0^s v(t_n+t)\,dt \le \int_0^s v(t_{n+1}+t)\,dt.$$

It would suffice to find a subsequence n_j along which convergence holds in the almost sure sense, and in addition, such that

$$\lim_{j\to\infty} \frac{\int_0^s v(t_{n_j}+t)\,dt}{\int_0^s v(t_{n_{j+1}}+t)\,dt} = 1. \tag{44}$$

COROLLARY 16. *Assume that $\alpha^* < 1/2$ is fixed, and that two sequences $(s_j)_{j\geq 1}$ and $(n_j)_{j\geq 1}$ are given where n_j is nondecreasing. If in addition to (39) and (44), we have*

$$\lim_{j\to\infty} \int_0^{s_j} v(t_{n_j}+t)\,dt = \infty, \qquad \lim_{j\to\infty} \frac{\int_{s_j}^s v(t_{n_j}+t)\,dt}{\int_0^{s_j} v(t_{n_j}+t)\,dt} = 0 \quad \textit{and} \tag{45}$$

$$\lim_{j\to\infty} \frac{\int_{s_j}^s v(t)\,dt}{\int_{s_j}^s v(t_{n_j}+t)\,dt} < \infty, \tag{46}$$

then the convergence of Theorem 5 holds almost surely.

PROOF. As discussed above, due to (44) and monotonicity, it suffices to show convergence as stated in Theorem 5 along the sequence $(n_j)_{j\geq 1}$. Due to the Borel–Cantelli lemma, (37), (39), (42) and the fact

$$P\Big(\limsup_j \{\tau_{n_0}^{n_j} < s_j\}\Big) = 0,$$

we have, as for Theorems 1 and 2, that

$$\lim_{j\to\infty} \frac{\int_0^{s_j} N^{\Lambda,n_j}(t)\,dt}{\int_0^{s_j} v(t_{n_j}+t)\,dt} = 1 \qquad \text{almost surely}$$

and

$$\lim_{j\to\infty} \frac{\int_0^{s_j} EN^{\Lambda,n_j}(t)\,dt}{\int_0^{s_j} v(t_{n_j}+t)\,dt} = 1.$$

Due to (45), we have

$$\liminf_{j\to\infty} \frac{\int_0^s N^{\Lambda,n_j}(t)\,dt}{\int_0^s v(t_{n_j}+t)\,dt} \geq 1 \qquad \text{almost surely}$$

and

$$\liminf_{j\to\infty} \frac{\int_0^s EN^{\Lambda,n_j}(t)\,dt}{\int_0^s v(t_{n_j}+t)\,dt} \geq 1.$$

For the corresponding upper bound on the lim sup, note that due to Theorem 1 (resp. Theorem 2) there exists a positive finite random variable C_0 (resp. positive constant C_0) such that

$$\frac{\int_{s_j}^s N^\Lambda(t)\,dt}{\int_{s_j}^s v(t)\,dt} \le 1 + C_0 \qquad \text{a.s.,} \left(\text{resp. } \frac{\int_{s_j}^s EN^\Lambda(t)\,dt}{\int_{s_j}^s v(t)\,dt} \le 1 + C_0\right), \text{ for all } j \ge 1.$$

Due to (45) and (46) and monotonicity $N^{\Lambda,n_j}(t) \le N^\Lambda(t)$ (with probability 1), we now have both

$$\lim_j \frac{\int_{s_j}^s N^{\Lambda,n_j}(t)\,dt}{\int_0^{s_j} v(t_{n_j}+t)\,dt} = 0 \qquad \text{almost surely,} \quad \text{and} \quad \lim_j \frac{\int_{s_j}^s EN^{\Lambda,n_j}(t)\,dt}{\int_0^{s_j} v(t_{n_j}+t)\,dt} = 0,$$

which completes the argument. □

Taking for example $\alpha^* = 1/4$, $s_j = 1/j^3$, and $n_j = \exp(\log^2 j)$ (resp. $n_j = j^\eta$ with $\eta > 3(\alpha - 1)$) in the case of Kingman (resp. Beta) coalescent, one can verify (left to the reader) the hypotheses of the last corollary, implying the final statement of Theorem 5.

APPENDIX: BINOMIAL CALCULATIONS

LEMMA 17. *If X has* Binomial(n, p) *distribution and if* $Y = X - \mathbf{1}_{\{X>0\}}$, *then*:

$$\begin{aligned} &\text{(i)} \quad EY = np - 1 + (1-p)^n; \\ \text{(47)} \quad &\text{(ii)} \quad \operatorname{var}(Y) = np(1-p) + (1-p)^n\bigl(1-(1-p)^n\bigr) - 2np(1-p)^n; \\ &\text{(iii)} \quad EY^2 = -np - np^2 + n^2p^2 + 1 - (1-p)^n. \end{aligned}$$

PROOF. Property (i) is trivial, (ii) follows easily from the fact that

$$\operatorname{cov}\bigl(X, \mathbf{1}_{\{X>0\}}\bigr) = np(1-p)^n$$

and (iii) is implied by (i) and (ii). □

COROLLARY 18. *If X has* Binomial$(n, 1-p)$ *distribution and if* $Y = X + \mathbf{1}_{\{X<n\}}$, *then*

$$E\left[\left(\frac{n-Y}{n}\right)^2\right] = O(p^2).$$

PROOF. Note that $n - Y$ has the distribution of the variable Y from Lemma 17. Hence its second moment is given in (47). Since for $p < 1/n$ we have

$$(1-p)^n = 1 - np + O(n^2p^2),$$

the claim of the corollary is true in this case. Now if $p \ge 1/n$ then $np = O(n^2p^2)$ therefore the largest term in (47) is again of order n^2p^2. □

LEMMA 19. *There exists* $n_0 \in \mathbb{N}$ *and* $C_0 < \infty$ *such that for all* $n \geq n_0$ *and all* $p \leq 1/4$, *if* X *has* Binomial$(n, 1-p)$ *distribution*, *then*

$$\left| E\big[\log(X + \mathbf{1}_{\{X<n\}}) - \log n\big] + \frac{np - 1 + (1-p)^n}{n} \right| \leq C_0 p^2$$

and

$$E\big[(\log(X + \mathbf{1}_{\{X<n\}}) - \log n)^2\big] \leq C_0 p^2.$$

PROOF. Let $Y = n - X$ as before, and abbreviate

$$T \equiv T_n := \log(X + \mathbf{1}_{\{X<n\}}) - \log n = \log\left(1 - \frac{Y - \mathbf{1}_{\{Y>0\}}}{n}\right). \tag{48}$$

We split the computation according to the event

$$A_n = \{Y \leq n/2\},$$

whose complement due to a large deviation bound has probability bounded by

$$\exp\left\{-n\left(\frac{1}{2}\log\frac{1}{2p} + \frac{1}{2}\log\frac{1}{2(1-p)}\right)\right\} = 2^n p^{n/2}(1-p)^{n/2}, \tag{49}$$

uniformly in $p \leq 1/4$ and n. On A_n^c we have $|T| \leq \log n$, and on A_n we apply a calculus fact,

$$|\log(1-x) + x| \leq \frac{x^2}{2(1-x)} \leq x^2, \qquad x \in [0, 1/2], \tag{50}$$

to obtain

$$\left| E[T] + E\left[\frac{Y - \mathbf{1}_{\{Y>0\}}}{n}\mathbf{1}_{A_n}\right]\right| \leq (\log n)P(A_n^c) + E\left[\frac{(Y - \mathbf{1}_{\{Y>0\}})^2}{n^2}\mathbf{1}_{A_n}\right].$$

Furthermore, since $(Y - \mathbf{1}_{\{Y>0\}})/n \leq 1$, we conclude

$$\left| E[T] + E\left[\frac{Y - \mathbf{1}_{\{Y>0\}}}{n}\right]\right| \leq (\log n + 1)P(A_n^c) + E\left[\frac{(Y - \mathbf{1}_{\{Y>0\}})^2}{n^2}\right]. \tag{51}$$

Note that by Corollary 18 and Lemma 17(i), in order to prove the first estimate of the lemma, it remains to show

$$(\log n)P(A_n^c) \leq (\log n)2^n p^{n/2}(1-p)^{n/2} \leq Cp^2 \tag{52}$$

for some $C < \infty$, all $p \in [0, 1/4]$, and all n large. Now consider $f : p \mapsto (p(1-p))^{n/2}/p^2$. Its derivative at p equals $g(p)(n(1-2p)/2 - 2(1-p))$ where $g(p)$ is a positive function. It is easy to check that if $p \leq 1/4$, then $n(1-2p)/2 - 2p^2(1-p) > 0$ for all $n \geq 6$. Therefore f is an increasing function of p, so in order to verify (52) for all $p \leq 1/4$, it suffices to check it for $p = 1/4$. This corresponds to having $(\log n)2^n(3/16)^{n/2} \leq C/16$, that will hold for all large $n = n(C)$ given a $C > 0$.

For the second estimate, again use the partitioning according to A_n and (50) to obtain

$$ET^2 \le E[T^2 \mathbf{1}_{A_n}] + \log^2 n P(A_n^c) \le \left(\frac{3}{2}\right)^2 E\left[\frac{(Y - \mathbf{1}_{\{Y>0\}})^2}{n^2}\right] + \log^2 n P(A_n^c),$$

which differs from (51) only by an extra factor of order $\log n$ multiplying $P(A_n^c)$, so the previous argument carries over. □

LEMMA 20. *There exists $n_0 \in \mathbb{N}$ and $K_0 < \infty$ such that for all $n \ge n_0$, $p \le 1/4$ and $c > 0$, if X has* Binomial$(n, 1-p)$ *distribution, then*

$$E\left[\exp\{c[\log(X + \mathbf{1}_{\{X<n\}}) - \log n]^2\} - 1\right] \le e^{9c/4} K_0 p^2.$$

PROOF. The strategy is the same as that used for the second estimate in the previous lemma, some details are left to the reader.

Recall that $Y = n - X$ and observe that

$$\begin{aligned}
E[e^{cT^2} - 1] &\le n^{c \log n} P(A_n^c) + E[(e^{cT^2} - 1)\mathbf{1}_{A_n}] \\
&\le n^{c \log n} P(A_n^c) + E\left[\left(\exp\left\{c \frac{9(Y - \mathbf{1}_{\{Y>0\}})^2}{4n^2}\right\} - 1\right)\mathbf{1}_{A_n}\right] \\
&\le n^{c \log n} P(A_n^c) + E\left[\exp\left\{c \frac{9(Y - \mathbf{1}_{\{Y>0\}})^2}{4n^2}\right\} - 1\right].
\end{aligned}$$

Hence it suffices to show that for some K_0, all $c > 0$ and all n, p as specified above, we have

$$E\left[\exp\left\{c \frac{(Y - \mathbf{1}_{\{Y>0\}})^2}{n^2}\right\} - 1\right] \le e^c K_0 p^2. \tag{53}$$

Without loss of generality, one can assume that $c > 1$.

The left-hand side above

$$\sum_{k=1}^{n} \binom{n}{k} p^k (1-p)^{n-k} \left(e^{c(k-1)^2/n^2} - 1\right)$$

can be bounded, using Taylor's expansion, by

$$\begin{aligned}
&\sum_{k=1}^{n} \binom{n}{k} p^k (1-p)^{n-k} \left\{c \frac{(k-1)^2}{n^2} + \frac{e^c}{2} \frac{(k-1)^4}{n^4}\right\} \\
&\quad = \frac{c}{n^2}\left(E(Y-1)^2 - P(Y=0)\right) \\
&\qquad + \frac{e^c}{2n^4}\left(E(Y-1)^4 - P(Y=0)\right).
\end{aligned} \tag{54}$$

Next compute

$$
\begin{aligned}
&E(Y-1)^2 - P(Y=0)\\
&\quad = \mathrm{Var}(Y-1) + \bigl(E(Y-1)\bigr)^2 - P(Y=0)\\
&\quad = np(1-p) + (np-1)^2 - (1-p)^n \qquad \text{[recall (21)]} \qquad (55)\\
&\quad \le np(1-p) + (np-1)^2 - e^{-np} + 2np^2/3\\
&\quad \le (np)^2 + O(np^2),
\end{aligned}
$$

where, for the last inequality, we recall that $e^{-x} - 1 + x > 0$ for $x > 0$. Similarly, using the fact

$$(y-1)^4 = y(y-1)(y-2)(y-3) + 2y(y-1)(y-2) + (y-1)^2$$

as well as the expressions for Binomial factorial moments, we have

$$
\begin{aligned}
&E(Y-1)^4 - P(Y=0) \qquad (56)\\
&\quad = n(n-1)(n-2)(n-3)p^4\\
&\qquad + 2n(n-1)(n-2)p^3 + E(Y-1)^2 - P(Y=0)\\
&\quad \le n^4p^4 + 2n^3p^3 + (np)^2 + O(np^2)\\
&\quad \le 4n^4p^2 + O(np^2). \qquad (57)
\end{aligned}
$$

Now (54)–(57) yield (53), and therefore the statement of the lemma, with appropriately chosen n_0. □

Acknowledgments. We are grateful to Etienne Pardoux for his careful reading of a preliminary draft, and to anonymous referees for their valuable comments and suggestions. The research project on which both the current work and [3] report, started at the time when J. Berestycki was a junior faculty at Université de Provence, while N. Berestycki was a postdoctoral fellow and V. Limic a junior faculty at the University of British Columbia.

REFERENCES

[1] ANGEL, O., BERESTYCKI, N., HAMMOND, A. and LIMIC, V. (2009). Global divergence of spatial coalescents. In preparation.

[2] BARLOW, M. T., JACKA, S. D. and YOR, M. (1986). Inequalities for a pair of processes stopped at a random time. *Proc. London Math. Soc.* **52** 142–172. MR812449

[3] BERESTYCKI, J., BERESTYCKI, N. and LIMIC, V. (2008). Interpreting Λ-coalescent speed of coming down from infinity via particle representation of super-processes. In preparation.

[4] BERESTYCKI, J., BERESTYCKI, N. and SCHWEINSBERG, J. (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.* **35** 1835–1887. MR2349577

[5] BERESTYCKI, J., BERESTYCKI, N. and SCHWEINSBERG, J. (2008). Small-time behavior of beta coalescents. *Ann. Inst. H. Poincaré Probab. Statist.* **44** 214–238. MR2446321

[6] BERTOIN, J. and LE GALL, J.-F. (2006). Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math.* **50** 147–181. MR2247827

[7] BOLTHAUSEN, E. and SZNITMAN, A.-S. (1998). On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197** 247–276. MR1652734

[8] DARLING, R. W. R. and NORRIS, J. R. (2008). Differential equation approximations for Markov chains. *Probab. Surv.* **5** 37–79. MR2395153

[9] DELMAS, J.-F., DHERSIN, J.-S. and SIRI-JEGOUSSE, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.* **18** 997–1025. MR2418236

[10] DONG, R., GNEDIN, A. and PITMAN, J. (2007). Exchangeable partitions derived from Markovian coalescents. *Ann. Appl. Probab.* **17** 1172–1201. MR2344303

[11] DURRETT, R. (2004). *Probability: Theory and Examples*, 3rd ed. Duxbury Press, Belmont, CA. MR1609153

[12] DURRETT, R. (2002). *Probability Models for DNA Sequence Evolution*. Springer, New York. MR1903526

[13] ELDON, B. and WAKELEY, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* **172** 2621–2633.

[14] EWENS, W. J. (2004). *Mathematical Population Genetics*. Springer, New York. MR554616

[15] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. MR0144363

[16] KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248. MR671034

[17] KINGMAN, J. F. C. (1982). On the genealogy of large populations. *J. Appl. Probab.* **19A** 27–43. MR633178

[18] LIMIC, V. and STURM, A. (2006). The spatial Λ-coalescent. *Electron. J. Probab.* **11** 363–393. MR2223040

[19] LI, G. and HEDGECOCK, D. (1998). Genetic heterogeneity, detected by PCR SSCP, among samples of larval Pacific oysters (Crassostrea gigas) supports the hypothesis of large variance in reproductive success. *Canad. J. Fish. Aquat. Sci.* **55** 1025–1033.

[20] MÖHLE, M. (2006). On sampling distributions for coalescent processes with simultaneous multiple collisions. *Bernoulli* **12** 35–53. MR2202319

[21] MORRIS, A. P., WHITTAKER, J. C. and BALDING, D. J. (2002). Fine-scale mapping of disease loci via shattered coalescent modeling of genealogies. *Amer. J. Hum. Genet.* **70** 686–707.

[22] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. MR1742892

[23] REVUZ, D. and YOR, M. (1994). *Continuous Martingales and Brownian Motion*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1303781

[24] ROGERS, L. C. G. and WILLIAMS, D. (1987). *Diffusions, Markov Processes and Martingales*, 2nd ed. Cambridge Univ. Press, Cambridge. MR1780932

[25] DE LA PEÑA, V. H. (1999). A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27** 537–564. MR1681153

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

[27] SIRI-JEGOUSSE, A. (2008). *Etude de Processus Appliqués Aux Modéles de Population*. Ph.D. thesis, Paris V.

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

[29] SCHWEINSBERG, J. (2000). Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5** 50. MR1781024

J. BERESTYCKI
LABORATOIRE DE PROBABILITÉS
ET MODÈLES ALÉATOIRES / UMR 7599
UNIVERSITÉ PIERRE ET MARIE CURIE-PARIS VI
4 PLACE JUSSIEU
F-75252 PARIS CEDEX 05
FRANCE

N. BERESTYCKI
STATISTICAL LABORATORY
UNIVERSITY OF CAMBRIDGE
WILBERFORCE RD.
CAMBRIDGE CB3 0WB
UNITED KINGDOM
E-MAIL: N.Berestycki@statslab.cam.ac.uk

V. LIMIC
UNIVERSITÉ DE PROVENCE
TECHNOPÔLE DE CHÂTEAU-GOMBERT
UMR 6632, LATP, CMI
39, RUE F. JOLIOT CURIE
13453 MARSEILLE, CEDEX 13
FRANCE