Global divergence of spatial coalescents

Omer Angel[*] Nathanaël Berestycki[†] Vlada Limic[‡]

September 2009

Abstract

A class of processes called spatial Λ-coalescents was recently introduced by Limic and Sturm (2006). In these models particles perform independent random walks on some underlying graph G. In addition, particles on the same site merge randomly according to some given coalescing mechanism. The goal of the current work is to obtain several asymptotic results for these processes. If $G = \mathbb{Z}^d$, and the coalescing mechanism is Kingman's coalescent, then starting with N particles at the origin, the number of particles is of order $(\log^* N)^d$ at any fixed time (where $\log^*$ is the inverse tower function). At sufficiently large times this number is of order $(\log^* N)^{d-2}$. Beta-coalescents behave similarly, with $\log\log N$ in place of $\log^* N$. Moreover, it is shown that on any graph and for general Λ-coalescent, starting with infinitely many particles at a single site, the total number of particles will remain infinite at all times, almost surely.

[*]University of British Columbia; supported in part by NSERC.
[†]Cambridge University; supported in part by EPSRC grant EP/G055068/1.
[‡]CNRS, UMR 6632; supported in part by Alfred P. Sloan Research Fellowship.

Contents

1 Introduction

1.1 Motivation and main results

A number of important mathematical population genetics models lead to a description of the genealogical tree of a population by a coalescent process known as Kingman's coalescent, introduced by Kingman in 1982 [23, 24]. The discovery of this process was a big step forward in theoretical population genetics. From a quantitative point of view, it serves as the basic model for evaluating genetic diversity in a well-mixed population at a neutral locus, and when individuals do not typically have a large number of offspring.

There has been recent interest in analogous models where an individual may give birth to a positive fraction of the population. For instance, in the case of strong natural selection (e.g., in viral populations), or when there is a large variation in the offspring distribution of a typical individual (e.g., marine species), there is convincing empirical evidence that the use of coalescents which allow for *multiple collisions* is more appropriate than that of Kingman's coalescent. The corresponding stochastic processes, called Λ-coalescents, were introduced by Pitman [29] and Sagitov [31]. In the next section we give the precise definitions and some background on these processes. These general models are much more intricate to analyse than Kingman's coalescent, for example, the analogue of Ewens's sampling formula is currently available (in an asymptotic form) only for a subclass of Λ-coalescents, see Berestycki et al. [5, 4].

In this paper, we study spatially structured generalisations of Λ-coalescent processes, which describe the genealogy of a population while taking into account various geographic factors such as migration, isolation, etc. They were introduced and named the spatial Λ-coalescents by Limic and Sturm [27]. These processes can be informally described as follows: given a locally finite graph $G = (V, E)$, each vertex $v \in V$ represents a the site of a colony, and each edge represents potential migratory routes between two adjacent colonies. Initially, we are given n particles on G. Particles which are on the same site, coalesce according to either Kingman's coalescent, or a more general Λ-coalescent. Simultaneously, each particle performs an independent random walk in continuous time, with some given rate $\theta > 0$.

In the language of theoretical population biology, a sample of n individuals is selected from the population at the present time, and their ancestral lineages are followed in reversed time. These above two transition rules reflect the idea that individuals typically reproduce within their own colony (so that only particles on the same site may coalesce), and occasionally there is a rare migration event, which corresponds to the random walk transitions. In the case where the coalescence mechanism is simply Kingman's coalescent, we note that this model may be viewed as the ancestral partition process associated with with Kimura's celebrated *stepping-stone model* [21, 22].

Our main results provide information about the limiting behavior of these processes as the size of the sample n tends to infinity, at both small and large time-scales. We consider the case where all the particles are located at the origin o of G at time $t = 0$. The only general assumptions we make on G are that it is connected and has bounded degree $\max_{v \in V} \text{degree}(v) < \infty$, though some of our more precise results on the asymptotic behaviour are restricted to the setting where G is the d-dimensional lattice $\mathbb{Z}^d$.

Let $\text{Vol}\, B(o, r)$ be the volume of the ball of radius r around o. Define the function $\log^* n$ as the inverse $\log^* n := \inf\{m \geq 1 : \text{Tow}(m) \geq n\}$ of the tower function:

$$\text{Tow}(n) = e^{\text{Tow}(n-1)} := \underbrace{e^{e^{\cdot^{\cdot^{\cdot^{e}}}}}}_{n\text{iterations}} . \tag{1}$$

Theorem 1.1. *Consider the spatial Kingman coalescent on a graph G with bounded degree such that $\text{Vol}\, B(o, k) \sim ck^d$ for some c, d. Start with n particles at $o \in G$, and let $N^n(t)$ be the total number of particles at time t. There are constants $C, c > 0$ depending only on t and the degree bound such that*

$$\mathbb{P}\left(c \leq \frac{N^n(t)}{\text{Vol}\, B(o, \log^* n)} \leq C \right) \xrightarrow[n \to \infty]{} 1.$$

Remark. The function $\log^* n$ tends to infinity with n, but at a very slow rate. For instance, when $n = 10^{78}$ (the total number of atoms in the universe), $\log^* n = 4$. For all practical purposes, $\log^* n$ is a constant equal to 3.

Remark. The behaviour in Theorem 1.1 contrasts with the non-spatial case, where $N^n(t)$ converges (without renormalisation) to a finite random variable $N(t)$ for all $t > 0$. In the lattice case $G = \mathbb{Z}^d$, we see that $N^n(t)$ diverges as $(\log^* n)^d$, i.e. extremely slowly. Even on a regular tree, where the balls have maximal volume given the degree, $N^n(t)$ diverges only as $e^{c \log^* n}$.

Our next result considers another biologically relevant case where the coalescence mechanism is given by a Beta-coalescent with parameter $1 < \alpha < 2$. (This process is defined in the next section). Our result is easier to state when the graph G is simply $\mathbb{Z}^d$. Here the number of particles that survive up to a fixed constant time is of order $(\log\log n)^d$, rather than $(\log^* n)^d$ as in the case of Kingman's coalescent.

Theorem 1.2. *Fix $1 < \alpha < 2$, and consider the spatial Beta($2-\alpha, \alpha$) coalescent on a graph G with bounded degree. Start with n particles at $o \in G$, and let $N^n(t)$ be the total number of particles at time t. There are constants $C, c > 0$ depending only on t, α and the degree bound such that*

$$\mathbb{P}\left(c \leq \frac{N^n(t)}{(\log\log n)^d} \leq C \right) \xrightarrow[n\to\infty]{} 1.$$

The next result concerns the case of more general Λ-coalescents (defined in Section 1.3). It states that regardless of the geometry of the underlying graph G, spatial coalescents are always globally divergent. Let $N^n(t)$ be the number of particles at time $t > 0$ when there are n particles located initially at a given vertex o of G.

Theorem 1.3. *For any measure Λ and any infinite G, consider the spatial Λ coalescent on G started with n particles at $o \in G$. Let $N^n(t)$ be the total number of particles at time t, then $N^n(t) \to \infty$ almost surely, as $n \to \infty$.*

This should be compared with the behaviour in the non-spatial case. There, depending on Λ, either for all $t > 0$ $N^n(t)$ converges almost surely to a finite random variable or it diverges for all $t > 0$. In the spatial case the divergence is universal. If Λ is such that the non-spatial coalescent diverges, then the result is trivial (see the discussion of the coming down from infinity property in Section 1.3).

While the above theorems consider the state of the system at a fixed time t, we also provide estimates for the number of particles that survive for a long time. Here the diffusion of particles plays a more important role, and the results depend in a fundamental way on the underlying graph. We focus on the case $G = \mathbb{Z}^d$. When $d = 1, 2$ or 3, this case has special biological relevance. For instance, when $d = 1$, one might think of coastal marine species, since the coast may be modeled as a one-dimensional object.

Theorem 1.4. *Assume that the coalescence mechanism is Kingman's coalescent. Let $G = \mathbb{Z}^d$, let $m = \log^* n$, and fix $\delta > 0$. Then there exist some constants $c > 0$ and $C > 0$ (depending only on d, δ) such that, if $d > 2$,*

$$\mathbb{P}\left(cm^{d-2} < N^n(\delta m^2) < Cm^{d-2} \right) \xrightarrow[n\to\infty]{} 1,$$

while, if $d = 2$, then

$$\mathbb{P}\left(c\log m < N^n(\delta m^2) < C\log m \right) \xrightarrow[n\to\infty]{} 1.$$

If the coalescent mechanism is a Beta distribution with parameters $(2-\alpha, \alpha)$ *and* $1 < \alpha < 2$*, then the same statement holds with* $m = \log \log n$*.*

For the study of the long-term behaviour of the system of particles, it is in fact natural, as suggested by Theorem 1.4, to rescale the particle system in time and in space, running the system with a clock that ticks m^2 times as quickly as the original one, while squeezing space by a factor of m^{-1}. Theorem 1.4 points to the rescaled system exhibiting a Boltzmann-Grad limiting behaviour, in which a typical particle meets and coalesces with a number of particles that is bounded away from zero and from ∞, as $n \to \infty$. The behaviour of the rescaled system mirrors the system of Brownian coagulating particles studied in [18] and [17], in which the Smoluchowski PDE $\partial_t u = \Delta u - cu^2$ is derived as the governing macroscopic behaviour. It is worth pointing out that in the current case, the discrete structure of the lattice remains important in determining the frequency of coalescence events even after space and time have been rescaled, which would alter the formula fixing the reaction coefficient in the limiting PDE.

The presence of this rescaling shows how the random system we consider may be thought of as a microscopic description of the small-time evolution of a solution of the Smoluchowski PDE that starts from a singular initial condition, such as a Dirac delta measure at a given spatial location.

Remark. Kesten [19, 20] studied the number of allelic types in a Wright-Fisher model with small mutation probability, where the allelic types have values in $\mathbb{Z}$ and mutations take the form of independent random walk steps. Specifically, the allelic type of an offspring is generally identical to that of its parent, but may change with a small probability (depending on the total population size). When it does change, the type jumps by some distribution on $\mathbb{Z}$ with finite support. The number of types (and, in fact, their relative positions in space) then has an equilibrium distribution. Kesten showed that, at equilibrium the number of observed types is of order $\log^* n$, where n is the sample size.

At a glance, this model appears related to the one-dimensional spatial Kingman coalescent, by considering the merging of ancestral lineages of samples from the population at equilibrium. Indeed, as the total population size $N \to \infty$, and after rescaling time, ancestral lineages of particles with the same type coalesce at rate 1, and each line mutates according to some random walk. The resulting (Moran) model is not quite the spatial Kingman coalescent: the interactions between the coalescent and the random walks is different. To shed light on the difference, let us explain Kesten's main result in these terms. Kingman's coalescent started with n particles gives rise to a tree. Now consider branching random walks in continuous time, with branching given by the tree (i.e., a tree-indexed random walk indexed by the coalescent tree). Then the number of occupied sites is of order $\log^* n$. This is true for similar reasons, and indeed the can be proved along the lines we use in Section 3

1.2 Heuristics and proof ideas

It is evident from Theorems 1.1 and 1.2 that the long term behaviour of the number of particles in the spatial coalescent depends delicately on the precise nature of the coalescent. We now describe the approximate behaviour of the spatial coalescent started with a large number of particles, all located at o. The proofs are mostly a detailed treatment of the following heuristic observations.

To understand the finite initial condition, we turn to the infinite one. Consider some Λ-coalescent which comes down from infinity (see below). Let N_t be the number of particles in the (non-spatial) coalescent started with $N_0 = \infty$. For Kingman's coalescent it is the case that $N_t \sim 2/t$, whereas for Beta-coalescents with parameters $(2-\alpha, \alpha)$ with $1 < \alpha < 2$, we have $N_t \sim c_\alpha t^{1-\alpha}$ [5, 6]. The rough description that follows applies to both of these, as well as more general coalescents. In general, one would expect N_t to be concentrated (for small t) around some function $g(t)$ (such a function is computed in [3]). The coalescent started with N particles is similar to the infinite coalescent observed from time $g^{-1}(N)$ onward.

Consider now the non-spatial coalescent with emigration, where each particle also disappears at some rate ρ. In fact, the parameter ρ may depend on the size of the population, as long as the model is still attractive, i.e. $n\rho(n)$ is non-decreasing. It turns out that for coalescents that come down from infinity, the emigration does not influence N_t so much, and N_t is still close to $g(t)$. The total number of particles that emigrate when starting with N particles is then close to a Poisson variable with mean

$$f(N) := \int_{g^{-1}(N)} g(t)\rho(g(t))dt. \tag{2}$$

(The upper bound of integration is some arbitrary constant.)

Now comes the key observation: if N is large, the number of particles migrating back into o is negligible (under a technical condition that holds for most spatial coalescents), and in fact, an overwhelming proportion of those particles that emigrate will have emigrated by time $g^{-1}(f(N))$. Thus we find that at this time, the number of particles at o and each of its neighbors is of order $f(N)$. A second observation is that the resulting populations can be approximated by independent spatial coalescents, when observed from time $g^{-1}(f(N))$ onward. In particular, at time $g^{-1}(f \circ f(N))$ there are of the order of $f \circ f(N)$ particles at each vertex in $B(o, 2)$. This "cascading onto neighbors" continues until step m, where m is such that $f \circ \cdots \circ f(N)$ (m repeated iterations of f) is of order 1. Note that in these m steps a ball of radius m has been roughly filled.

Applying this heuristics to the case of Kingman's coalescent and the Beta-coalescents with parameters $(2-\alpha, \alpha)$ and $1 < \alpha < 2$, gives the following. For Kingman's coalescent and constant ρ, we have $f(n) \sim 2\rho \log n$, and for Beta-coalescents we have $f(n) \sim C_\alpha \rho n^{2-\alpha}$, for some constant $C_\alpha > 0$. Thus in the first case, $m = \log^* N$. In the second case, we find $m \sim c \log\log N$. In general, this gives $m \sim f^*(n)$, where

$$f^*(n) = \inf\Big\{m \geq 1 : \underbrace{f \circ \cdots \circ f}_{m \text{ iterations}}(n) \leq 1\Big\}. \tag{3}$$

Note that if $\rho(n)$ decreases fast enough so that $f(n)$ is bounded, then it follows from this heuristic analysis that the spatial coalescent will come down from infinity globally. However, when ρ is constant, it can be proved that f is always unbounded, which in turn implies the result about global divergence of any spatial Λ-coalescent.

Turning to long time asymptotics, by the above reasoning we may start from a configuration consisting of a tight number of particles at each site of the ball of radius m around the origin. Since the number of particles per site is tight, the type of coalescent influences the evolution less than the diffusion. In particular, the structure of the underlying graph has a profound effect on the asymptotics. For simplicity, let us restrict ourselves to d-dimensional Euclidean lattices with $d \geq 3$. Let $\rho(t)$ denote the average number of particles per site in the ball of radius m at time t. Then at time $t_0 = 1$ we have $\rho(t_0) \asymp 1$ and $\lim_{t\to\infty} \rho(t) = 0$. Each particle present in the configuration at time t coalesces with another particle at an average rate approximately $\rho(t)$, so that $\frac{d}{dt}N(t) = -N(t)\rho(t)/2$. Dividing by the volume of the ball, one arrives to the ODE

$$\frac{d}{dt}\rho(t) = -\frac{1}{2}\rho(t)^2, \tag{4}$$

whose solution is given by $\rho(t) = 2/(t+c)$ for some $c > 0$.

The approximation (4) should be valid as long as the diffusion of particles away from the initial region (i.e., $B(o,m)$) is negligible. The influence of diffusion should start to be visible at times of order m^2. In particular, at time m^2, the density $\rho(m^2)$ is of order m^{-2}, so the total number of remaining particles is of order m^{d-2}. Assuming the plausible claim that the remaining particles are approximately uniformly distributed over a ball of radius order m, a simple calculation (using hitting probabilities for random walks) now implies that each of them has a positive probability of never meeting any other particle again, and so the number of particles that survive indefinitely is of order m^{d-2}.

We point out that van den Berg and Kesten [8, 9] have shown a density decay similar to (4) for a related model of coalescing random walks. However their results differ in two ways. On the one hand, the coalescence mechanism which they analyse is different. On the other hand, and more importantly, their initial condition is initially homogeneous in space, and not restricted to a large ball. This restriction is the cause of much of the difficulty in our case – see Section 7 for more details.

1.3 Definitions and background on spatial coalescents

Kingman's coalescent. Suppose we are given an integer $n \geq 1$. *Kingman's n-coalescent* is the Markov process $(\Pi^n_t, t \geq 0)$, with values in the set $\mathcal{P}_n$ of partitions of $[n] := \{1, \dots, n\}$, such that $\Pi^n_0 = \{\{1\}, \{2\}, \dots, \{n\}\}$, and such that each pair of blocks merges at rate 1, and these are the only transitions of the process. Blocks of the partition Π^n_t may be viewed as indistinguishable particles, and we often refer to the number of blocks of Π^n_t as the number of particles alive at time t. A simple but essential property of Kingman's n-coalescent is the so-called *sampling consistency* property: the restriction of $(\Pi^{n+1}_t, t \geq 0)$ to $[n]$ has the same distribution as an n-coalescent. This enables one to construct a Markov process $(\Pi_t, t \geq 0)$ with state space $\mathcal{P}$, the set of partitions of $\mathbb{N}$, such that the law of Π when restricted to $[n]$ equals the law of Π^n. In particular, the initial state of the process is the trivial partition $\Pi_0 = \{\{1\}, \{2\}, \dots\}$. The process Π is called *Kingman's coalescent.* For background reading, see for instance [14, 30, 7].

Λ-coalescents. Let Λ be a finite measure on $[0, 1]$. A *coalescent with multiple collisions*, or *Λ-coalescent*, is a Markov process $(\Pi_t, t \geq 0)$ with values in the set of partitions of $\mathbb{N}$ characterized by the following properties. If $n \in \mathbb{N}$, then the restriction of $(\Pi_t, t \geq 0)$ to $[n]$ is a Markov chain $(\Pi^{(n)}_t, t \geq 0)$, where $\Pi^n_0 = \{\{1\}, \{2\}, \dots, \{n\}\}$, and where the only possible transitions are mergers of blocks (it is possible to merge several blocks simultaneously into one block, but no two mergers of this kind can occur simultaneously) so that whenever the current configuration consists of b blocks, any given k-tuple of blocks merges at rate

$$\lambda_{b,k} = \int_{[0,1]} x^{k-2}(1-x)^{b-k}\Lambda(dx). \tag{5}$$

Note that 0^0 is interpreted as 1, so that an atom of Λ at 0 causes each pair of particles to coalesce at a finite positive rate $\Lambda(\{0\})$. In this way any Λ-coalescent can be thought of as a superposition of a "pure" coalescent with multiple collisions driven by measure $\Lambda(dx)\mathbf{1}_{\{(0,1]\}}(x)$, and a time-changed Kingman's coalescent. An atom of Λ at 1 causes all the particles to coalesce at some positive fixed rate. Such Λ-coalescent may be viewed as a killed Λ'-coalescent where $\Lambda'(dx) = \Lambda(dx)\mathbf{1}_{\{[0,1)\}}(x)$. Kingman's coalescent is a particular Λ-coalescent, obtained when the measure Λ equals δ_0, the unit Dirac mass at 0. Any Λ-coalescent Π is sampling consistent, that is, if $m < n$ then the restriction of Π^n to $[m]$ is equal in law to Π^m. It is this observation that allows one to construct an infinite version of the process. It is interesting to note the following fact shown by Pitman [29]: Λ-coalescents are the only exchangeable Markov coalescent processes without simultaneous collisions. We refer the reader to [29] for definitions and further properties.

If $\Lambda(dx) = dx$, the corresponding Λ-coalescent is usually referred to as the Bolthausen-Sznitman coalescent, that arises in the context of spin glasses. Another important case is when Λ is the Beta$(2-\alpha,\alpha)$ distribution for some parameter $1 < \alpha < 2$, that is,

$$\Lambda(dx) = \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} x^{1-\alpha}(1-x)^{\alpha-1}\,dx. \tag{6}$$

Such Λ-coalescents are called *Beta-coalescents*. Their relevance is apparent in view of a result of Schweinsberg [32]: the Beta-coalescent corresponding to a fixed parameter α arises in the scaling limit of population models where the offspring distribution of a typical individual is in the domain of attraction of a stable law with index α. Apart from the Kingman's coalescent, this is the best-understood class of Λ-coalescents (see, e.g., [11, 6, 5]).

Spatial coalescents. As informally described above, spatial coalescents are processes which combine spatial motion of individual particles with coalescence of particles located on the same site of a given graph of bounded degree. Formally, let Λ be a given finite measure on $[0,1]$. A spatial Λ-coalescent, as defined in [27], is a Markov processes $(\Pi^\ell_t, t \geq 0)$ with values in the space $\mathcal{P}^\ell = \mathcal{P} \times V^{\{1,2,\ldots\}}$ of partitions of $\{1,2,\ldots\}$ indexed by spatial locations. That is, an element $x = (\pi,\ell) \in \mathcal{P}^\ell$ consists of a partition $\pi = \{A_1, A_2, \ldots\}$, and a sequence $\ell = (\ell_1, \ell_2, \ldots)$, where ℓ_i specifies the location of the block A_i. There are only two types of transitions possible for $\Pi^\ell_t = (\Pi_t, \ell_t)$: (i) provided there are b blocks at a location $v \in V$, then any given k-tuple of them will merge at rate $\lambda_{b,k}$ given by (5), independently over v; and (ii) independently of the coalescent mechanism, each block A_k of π migrates at rate θ. This means that if the block is at v, then some vertex w is chosen according to the distribution $p(v,\cdot)$, where $p(v,w)$ is a given Markov kernel. When this happens, ℓ_k is changed from v to w. To simplify the discussion, we will assume unless otherwise specified, that $p(x,y)$ is the transition kernel for the simple random walk on the underlying graph.

If π is a partition let $i \sim_\pi j$ mean that the particles labelled i and j belong to the same block of π. For $(\pi, \ell) \in \mathcal{P}^\ell$ and $v \in V$, denote by $\#_v(\pi, \ell)$ the number of blocks in π with label (location) v.

Spatial Λ-coalescents inherit the sampling consistency directly from Λ-coalescents. Namely, if we consider a spatial coalescent started from $n+1$ particles (that is, blocks) and consider its restriction to the first n particles, the new process has the law of a spatial coalescent started from n particles. This simple property will be used on several occasions. In particular, it implies that if (π^1, ℓ^1) and (π^2, ℓ^2) are such that $\#_v(\pi^1, \ell^1) \le \#_v(\pi^2, \ell^2)$, for all v, then there exists a coupling of two spatial coalescents $((\Pi_t^1, \ell_t^1), (\Pi_t^2, \ell_t^2)), t \ge 0)$ such that $(\Pi_0^i, \ell_0^i) = (\pi^i, \ell^i)$, $i = 1, 2$ and $\#_v(\Pi_t^1, \ell_t^1) \le \#_v(\Pi_t^2, \ell_t^2)$ for all v, almost surely. The same property guarantees the existence of spatial coalescents started with infinitely many particles on an infinite graph (see Theorem 1 in [27] for a particular construction).

Spatial Λ-coalescents may be started from configurations containing countably infinitely many particles at each site of G, see [27]. However, our results concern spatial Λ-coalescents started from the following initial condition:

$$\Pi_0^\ell = (\{\{1\}, \{2\}, \ldots\}, (o, o, \ldots)), \tag{7}$$

where o is some given reference vertex called the *origin* of G. In words, all the infinitely many particles are initially located at the origin o.

From now on we abbreviate

$$X_v(t) = \#_v(\Pi_t, \ell_t) \qquad \text{and} \qquad X_v^n(t) = \#_v(\Pi_t^n, \ell_t). \tag{8}$$

We denote the *total number of blocks* by $N^*(t) = \sum_{v \in V} X_v(t)$ (resp. $N^n(t) = \sum_{v \in V} X_v^n(t)$). When not in risk of confusion, we will drop the superscript n to simplify notations. It is clear from the definitions that both $(\{X_v(t)\}_{v \in G}, t \ge 0)$ and $(N^*(t), t \ge 0)$ have Markovian transitions, with respect to the filtration generated by the coalescent process Π. They carry only partial information about the evolution of the corresponding spatial coalescent, in particular, they do not determine the evolving partition structure.

Coming down from infinity. Let $(\Pi_t, t \ge 0)$ be Kingman's coalescent. As already mentioned, Kingman [23, 24] showed that while Π starts with an infinite number of blocks at $t = 0$, its number of blocks becomes finite for all $t > 0$, almost surely. For coalescents with multiple collisions, such a phenomenon may or may not happen, depending on the measure Λ. More precisely, Pitman [29] showed that there are only two possibilities: let E (resp. F) denote the event that for all $t > 0$ there are infinitely (resp. finitely) many blocks. Then, if $\Lambda(\{1\}) = 0$, either $P(E) = 1$ or $P(F) = 1$. When $P(F) = 1$, the process Π is said to *come down from infinity*. For instance, a Beta-coalescent comes down from infinity if and only if $1 < \alpha < 2$, henceforth we make this an assumption whenever working with Beta-coalescents.

In the context of spatial coalescents, assuming that $\Lambda(\{1\}) = 0$, Proposition 11 in [27] implies that when the initial number of particles is infinite, then $X_v(t)$ becomes finite for all $v \in V$ and $t > 0$ with probability 1, if and only if the underlying measure Λ is such that the mean-field (i.e., non-spatial) Λ-coalescent comes down from infinity. In this situation, we may say that the spatial coalescent comes down from infinity *locally*. Naturally, this stays true if the initial condition is (7).

Other notations. Unless specified otherwise, c, C (and variations $c_1, C_2, \ldots$) will henceforth denote positive constants that depend only on the dimension d, and that may change from line to line. Typically, $c, c_1, \ldots$ denote sufficiently small, whereas $C, C_2, \ldots$ denote sufficiently large constants. We also use the symbols $a_n \sim b_n$ and $a_n \asymp b_n$ to denote respectively that $a_n/b_n \to 1$ as $n \to \infty$, and a_n/b_n is bounded away from 0 and ∞.

Organization of the paper. The rest of the paper is organized as follows. Section 2 starts with some preliminary remarks and observations, mainly dealing on the one hand with certain large deviation estimates for Kingman's coalescent and Ewens' sampling formula, and on the other hand, about various couplings between spatial Λ-coalescents and mean-field (i.e., non-spatial) Λ-coalescents, which will be used throughout the paper. Section 3 contains a proof of Theorem 1.1 on the behaviour of the spatial Kingman coalescent in finite time. As many of the subsequent results in the paper build on this, we recommend that the reader starts there. Section 4 contains the proof of Theorem 1.2 which deals with the finite-time behaviour of the spatial Beta-coalescents. Section 5 returns to the general case of spatial Λ-coalescents and arbitrary graphs with bounded degree, and contains the proof of global divergence (Theorem 1.3). The final sections 6 and 7 deal respectively with the lower bound and the upper bound for long term behaviour of Kingman's coalescent (Theorem 1.4). The lower bound is true for general Λ-coalescents (section 6.2), but we provide an alternate shorter proof (section 6.3) for the special case of Kingman's coalescent, which also gives tighter bounds. The proof of the upper bound (7) turns out to be the most technical part, requiring a delicate multi-scale analysis.

Sections 4, 5 and 6–7 may be read in different orders depending on the interest of the reader.

2 Preliminary lemmas

2.1 Some large deviation estimates

We begin with an easy large deviation result for a sum of exponential random variables, which we prefer to state in an abstract form now so as to refer to it on several occasions later. In our applications, $\mathbb{E}S$ will typically be small.

Lemma 2.1. *For an index set I, let $\mu_i > 0$ for each $i \in I$, and let $\{E_i\}_{i\in I}$ be independent exponential random variables with $\mathbb{E}E_i = \mu_i$. Let $S = \sum_{i\in I} E_i$. Then for any $0 < \varepsilon < 1$*

$$\mathbb{P}\big(S < (1-\varepsilon)\mathbb{E}S\big) \le \exp\left(-\frac{\varepsilon^2(\mathbb{E}S)^2}{4\operatorname{Var} S}\right).$$

Additionally, for $0 < \varepsilon < \frac{2\operatorname{Var} S}{\mathbb{E}S \sup\{\mu_i\}}$,

$$\mathbb{P}\big(S > (1+\varepsilon)\mathbb{E}S\big) \le \exp\left(-\frac{\varepsilon^2(\mathbb{E}S)^2}{4\operatorname{Var} S}\right).$$

Remark. If $I = \{n, n+1, \ldots\}$ and $\mu_i \sim ci^{-\alpha}$ for some $\alpha > 1$, then as $n \to \infty$, $\frac{2\operatorname{Var} S}{\mathbb{E}S \sup\{\mu_i\}}$ is bounded away from 0, hence the upper-bound holds for all $\varepsilon > 0$ small enough.

Proof. Using Markov's inequality, for any $0 < \lambda \le \frac{1}{2}\inf\{\mu_i^{-1}\}$

$$\begin{aligned}\mathbb{P}\big(S > (1+\varepsilon)\mathbb{E}S\big) &\le e^{-\lambda(1+\varepsilon)\mathbb{E}S}\mathbb{E}e^{\lambda S}\\ &= e^{-\lambda(1+\varepsilon)\mathbb{E}S}\prod \frac{1}{1-\lambda\mu_i}\\ &< e^{-\lambda(1+\varepsilon)\mathbb{E}S}\exp\Big(\sum \lambda\mu_i + \lambda^2\mu_i^2\Big)\\ &= e^{-\lambda\varepsilon\mathbb{E}S+\lambda^2\operatorname{Var}S},\end{aligned}$$

where we have used that for $x \in (0,1/2)$ we have $-\ln(1-x) < x + x^2$. Taking $\lambda = \frac{\varepsilon\mathbb{E}S}{2\operatorname{Var}S}$, which is allowed since $\varepsilon < \frac{2\operatorname{Var}S}{\mathbb{E}S\sup\{\mu_i\}}$, yields the upper bound.

The lower bound follows from a similar argument with $\lambda = -\frac{\varepsilon\mathbb{E}S}{2\operatorname{Var}S}$. □

We now apply this to get a large deviation estimate for Kingman's coalescent. This uses a simple idea which can already be found in Aldous [1], who used it to prove a central limit theorem for the number of particles at time t. Denote by $\mathbb{P}^n$ the law of the (non-spatial) Kingman coalescent started with n blocks. Let $N(t)$ be the number of blocks at time t.

Lemma 2.2. *Let $t = t(n) \to 0$ in such a way that $t(n)^{-1} = o(n)$. For any $0 < \varepsilon < 1/2$, for n large enough,*

$$\mathbb{P}^n\left(1-\varepsilon < \frac{N(t)}{2/t} < 1+\varepsilon\right) > 1 - \exp\left(-\frac{\varepsilon^2}{t}\right).$$

Proof. For the upper bound, let $m = \lceil (1+\varepsilon)2/t \rceil$. The time it takes the process to get from n to m particles is a sum of independent exponential random variables with means $\binom{k}{2}^{-1}$ for $k = m+1, \dots, n$. Call this sum S. If $N(t) > m$ then $S > t$. We have

$$\mathbb{E}S = \sum_{k=m+1}^{n} \binom{k}{2}^{-1} \sim 2m^{-1} \sim t/(1+\varepsilon)$$

provided $m = o(n)$. Similarly,

$$\operatorname{Var}S = \sum_{k=m+1}^{n} \binom{k}{2}^{-2} \sim (4/3)m^{-3}.$$

Thus, for $\varepsilon < 2/3 + o(1)$,

$$\mathbb{P}(S > t) < \exp\left(-\frac{(3+o(1))\varepsilon^2}{2t}\right),$$

by Lemma 2.1. The lower bound is similar using the upper bound on S. □

We now consider Kingman's coalescent with spatial migration. Let $\mathbb{P}^n$ be the law of a simplified process where n particles initially located at a single site o coalesce according to Kingman's dynamics, while each particle (or block of particles) migrates at rate ρ, and any block that migrates away from o is ignored from that time onwards. Denote by Z_n the total number of blocks that ever migrate away from o.

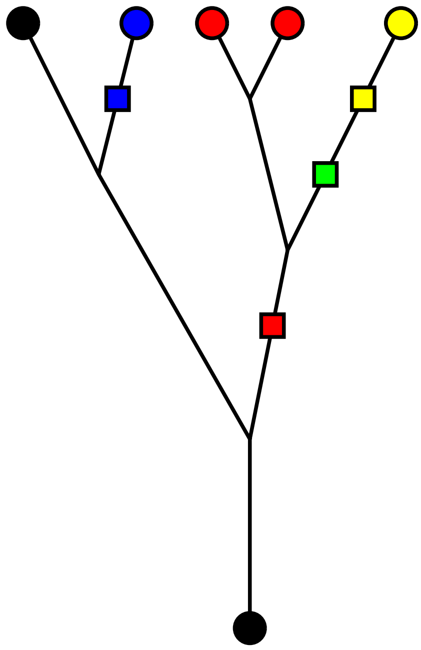

Figure 1: The random partition generated by mutations (squares). Here $Z_5 = 4$.

We can think of each migration event as of a "unique mutation on the genealogical tree", for example by thinking of its occurrence time as its label. Since migrations happen at rate ρ for each block present in the configuration at site o, one quickly realizes that Z_n is a realization from a well-known distribution arising in mathematical population genetics. Namely, set $\theta = 2\rho$, and suppose that on the (non-spatial) Kingman coalescent tree mutation marks occur at a Poisson rate of $\theta/2$ per unit length. Using the language of mathematical population genetics, assume the infinite alleles models (all mutations create a different allele, and so different individuals in the original sample of n are in the same family if and only if they descend from the same mutation and there has been no other mutation between this common ancestor and the present individuals). The marks of the mutation process on define a random partition Π_θ on the leaves of the tree by declaring that i and j are in the same block of Π_θ if and only if there is no mutation mark on the shortest path that connects i and j. In Figure 1 different blocks of this partition are represented by different colors. Then it is easy to see that Z_n has the law of the number of blocks in Π_θ. It is well-known (see, e.g., (3.24) in Pitman [30]) that Z_n is of order $\theta \log n$ for large n. The following large deviation estimate is part of the folklore, but we could not find a precise reference for it in the literature.

Lemma 2.3. *Fix $\varepsilon > 0$. There are $c, C > 0$ such that*

$$\mathbb{P}^n\left(\left|\frac{Z_n}{\log n} - \theta\right| > \varepsilon\right) < Cn^{-c}.$$

Furthermore, for any U,

$$\mathbb{P}^n(Z_n > U) < Cn^C e^{-U}.$$

The proof is based on the Chinese restaurant process representation of Ewens's sampling formula. Let $K_{n,i}$ be the number of blocks of size i in Π_θ, where $i = 1, \dots, n$. Then the distribution of $(K_{n,1}, \dots, K_{n,n})$ is given by Ewens's sampling formula $ESF(\theta)$:

$$P(K_{n,i} = a_i,\, i = 1, \dots, n) = \frac{n!}{\theta(\theta+1)\cdots(\theta+n-1} \prod_{i=1}^{n} \frac{\theta^{a_i}}{i^{a_i} a_i!}, \tag{9}$$

for any given collection $a_1, \dots, a_n$ of non-negative integers such that $\sum_{i=1}^n i a_i = n$.

The Chinese restaurant process representation of (9) (see [30, §3.1]), states that the number of blocks in Π_θ satisfies

$$Z_n \stackrel{d}{=} \sum_{i=1}^{n} \zeta_i \tag{10}$$

where ζ_i are independent Bernoulli random variables with mean

$$P(\zeta_i = 1) = \frac{\theta}{i+\theta}.$$

Proof of Lemma 2.3. By (10) we have for $\lambda > 0$

$$\begin{aligned}
\mathbb{E}e^{-\lambda Z_n} &= \prod_{i\le n} \left(1 - \frac{\theta}{i+\theta}(1-e^{-\lambda})\right) \\
&< \exp\left(\sum_{i\le n} -\frac{\theta}{i+\theta}(1-e^{-\lambda})\right) \\
&< \exp\left(-\theta(1-e^{-\lambda})(C+\log n)\right).
\end{aligned}$$

By Markov's inequality

$$\begin{aligned}
\mathbb{P}(Z_n < (1-\varepsilon)\theta\log n) &< \exp\left(\lambda(1-\varepsilon)\theta\log n - \theta(1-e^{-\lambda})(C+\log n)\right) \\
&< \exp\left((-\lambda\varepsilon\theta + O(\lambda^2))\log n + C\right)
\end{aligned}$$

For small positive λ the coefficient of $\log n$ is strictly negative.

Similarly, for $\lambda > 0$

$$\begin{aligned}
\mathbb{E}e^{\lambda Z_n} &= \prod_{i=1}^{n} \left(1 + \frac{\theta}{i+\theta}(e^{\lambda}-1)\right) \\
&< \exp\left(\sum_{i=1}^{n} \frac{\theta}{i}(e^{\lambda}-1)\right) \\
&< \exp\Big(\theta(e^{\lambda}-1)(C+\log n)\Big).
\end{aligned}$$

Thus, by Markov's inequality,

$$\mathbb{P}\left(Z_n > U\right) < \exp\Big(\theta(e^{\lambda}-1)(C+\log n) - \lambda U\Big) \tag{11}$$

Taking $U = (1+\varepsilon)\theta\log n$ and $\lambda = \lambda(\varepsilon,\theta)$ small enough gives the first upper bound. Taking $\lambda = 1$ gives the second claim. □

A similar computation can be found in Greven et al. [16, Lemma 3.3].

2.2 Coupling and comparison

The final general tool we need is a way to couple our process with simpler coalescing processes such as non-spatial ones. While it is traditional in coalescent theory to keep track of the entire partition structure as time evolves, we are here only interested in the number of particles still present in the system at any given time, for which this structure can be quite cumbersome to manipulate. Instead, we note that the process X is equivalent to one where there is at each time t only a subset of the original n particles, and the other have been killed. Label the n initial particles $1, \ldots, n$, and let $x_1, \ldots, x_n$ be the initial locations of these particles. Let $S^1, \ldots, S^n$ be n i.i.d. simple random walks on G in continuous time with rate ρ started at $x_1, \ldots, x_n$. Any particular set of k particles on the same vertex v may coalesce at rat $\lambda_{b,k}$ if there are b particles in total on that site. When such an event occurs, involving particles with labels $i_1 < \ldots < i_k$, let the newly formed particle follow the trajectory of S^{i_1}. Thus we only keep the lowest particle and kill all the other particles.

The *Poissonian construction* (see below) enables us to decide ahead, for each vertex $v \in V$, which labels would coalesce if they were on v and at what time. This construction is relatively heavy but useful and we will recall it in greater detail in the proof of the next lemma, which is a result of monotonicity.

Suppose our n initial particles are partitioned into classes, represented by a partition $\pi = (B_1, \ldots, B_r)$ of $\{1, \ldots, n\}$ for some $r \geq 1$. We wish to compare the system $X = \{X_v(t), t \geq 0\}_{v \in V}$ to one which consists only of the particles belonging to a particular block B of π. For any block B of π, denote by X^B a process which has the law of a spatial coalescent restricted to particles from the class B. Let $N(t)$ denote the total number of particles of $X(t)$ and let $N^B(t)$ denote the total number of particles of $X^B(t)$.

Lemma 2.4. *There is a coupling $(X, X^{B_1}, \ldots, X^{B_r})$ such that, almost surely,*

$$\forall v \in V \quad X_v^B(t) \leq X_v(t) \leq \sum_{i=1}^r X_v^{B_i}(t),$$

for all block B of π, and hence

$$N^B(t) \leq N(t) \leq \sum_{i=1}^r N^{B_i}(t).$$

Note that in the coupling we construct the processes X^{B_i} are not independent. In fact, under weak assumptions on the coalescent and when all blocks are small, there is a coupling where they are independent, see Lemma 7.2.

Proof. The idea is simple. Fix a realization of the process $(X(t), t \geq 0)$ as described above. In addition, we start by colouring red every particle from B. If k particles coalesce at some point and at least one of them is red, then colour the newly formed particle red as well. The reader may easily check that, by the consistency of Λ-coalescents, the process restricted to the red particles has exactly the same distribution as $(X^B(t), t \geq 0)$. Therefore, in this coupling we see that $X_v^B(t) \leq X_v(t)$ since X_v may contain both red and black particles. Moreover, in this construction it is also the case that $X_v(t) \leq \sum_{i=1}^r X_v^{B_i}(t)$, since any particle with label i

counted in $X_v(t)$ will also be counted in $X_v^B(t)$, where B is the block that contains i. Thus the right hand side inequality in the first estimate follows, and the second estimate is a trivial consequence of the first. $\square$

A second type of coupling we will need is between a spatial Λ-coalescent and a mean-field (i.e., non-spatial) one. Fix a vertex $u \in V$ of the graph, and consider a spatial Λ-coalescent $\{X_v(t), t \geq 0\}_{v \in V}$ started with a finite number of particles and such that initially $X_u(0) = n$. Let $M(t)$ denote the number of particles on u at time t that have always stayed at u, and let $Z(t)$ denote the number of particles that jumped out of u prior to time t. In parallel, let $(N(t), t \geq 0)$ denote the number of particles at time t in a mean-field Λ-coalescent started with n particles.

Lemma 2.5. *There exists a coupling of X and N such that:*

$$M(t) \leq N(t) \leq M(t) + Z(t), \qquad \text{a.s. for all } t \geq 0, \tag{12}$$

and

$$N(t) - Z(t) \leq X_u(t) \leq N(t) + Z(t), \text{ a.s. for all } t \geq 0. \tag{13}$$

Proof. The process $M(t)$ may be realized as a mean-field coalescent where, in addition, particles are killed at rate ρ. In that case, if we let $Z(t)$ denote the total number of particles that have been killed, we see immediately that on the one hand, $M(t) \leq N(t)$, and on the other hand, $N(t) = M(t) + \bar{Z}(t)$ where $\bar{Z}(t) \leq Z(t)$. Indeed, $M(t) + Z(t)$ counts the number of particles if we freeze particle instead of killing them. However, in $N(t)$ these particles keep coalescing, and so the difference $\bar{Z}(t) = N(t) - M(t) \leq Z(t)$. This proves (12). For (13), note first that $X_u(t) \geq M(t) = N(t) - \bar{Z}(t) \geq N(t) - Z(t)$. Finally, the last inequality in (13) is obtained by observing that $X_u(t)$ is made of particles that never jumped out of u (there are $M(t)$ such particles) and of particles that have jumped out of u and have come back at some time later, potentially coalescing in the meantime. There can never be more than $Z(t)$ such particles, since this is the total number of particles that jump out of u. $\square$

In fact, one can be slightly more precise than the above estimate. We shall need the following observation. Define two processes

$$S(t) = Z(t) - \int_0^t \rho M(s) ds \qquad\qquad V(t) = S(t)^2 - \int_0^t \rho M(s) ds. \tag{14}$$

It is a standard (and easy) fact that both are continuous time martingales under the law $\mathbb{P}^n$, with respect to the filtration $\mathcal{F}$ generated by the above coupling process. In fact, if we define $\mathcal{G} = \sigma\{N(u), u \geq 0\}$ to be the σ-algebra generated by N, and let $\mathcal{F}_t^* = \sigma\{\mathcal{G}, \mathcal{F}_t\}$, then the processes $S(\cdot)$ and $V(\cdot)$ are continuous-time martingales with respect to the filtration $\mathcal{F}^*$.

Lemma 2.6. *For any time interval $[a, b]$, we have the stochastic domination*

$$\mathbb{P}(Z(b) - Z(a) \geq x | \mathcal{G}) \leq \mathbb{P}\left(\text{Poisson}\left(\rho \int_a^b N(s)\, ds\right) \geq x \,\middle|\, \mathcal{G}\right).$$

Proof. Given $\mathcal{G}$, Z is a pure jumps process with jumps of size 1 that arrive at rate $\rho M(t) \leq \rho N(t)$ at time t, almost surely. $\square$

Finally, a more global comparison with mean-field coalescents can be obtained as follows in the case of the spatial Kingman coalescent. Let S be an arbitrary subset of vertices and consider the restriction of X to S.

Lemma 2.7. *Fix a time $\tau \leq 2$, and vertex set S, and assume that all particles are in S at time 0. Let $Z = Z(\tau)$ be the number of distinct particles that exit S by time τ, and let $N_S(t)$ be the number of particles in S at time t. Then, for some $c = c(\epsilon) > 0$,*

$$\mathbb{P}\left(N_S(\tau) > Z + \frac{(4+\varepsilon)|S|}{\tau}\right) < e^{-c|S|/\tau}.$$

Note that the bound is independent of the starting configuration. This Lemma is a precursor to Lemma 7.5.

Proof. Let Q_t be the number of particles in S that have survived until time t but have not left S. We have then that $N_S^n(t) \leq Z + Q_t$. The rate of coalescence inside S at time t is

$$\sum_{v\in S} \binom{X_v(t)}{2} \geq |S| \binom{Q_t/|S|}{2}$$

(by Jensen's inequality for $\binom{x}{2}$.) If $Q_t < 2|S|$ for some $t \leq \tau$ then we are done (since $\tau \leq 2$.) Otherwise, $|S| \cdot \binom{Q_t/|S|}{2} \geq \frac{1}{2|S|}\binom{Q_t}{2}$, and so Q_t is dominated by a Kingman coalescent slowed down by a factor of $2|S|$. Lemma 2.2 completes the proof. □

3 Finite time behaviour of the spatial Kingman coalescent

3.1 Lower bound

Recall that $\theta = 2\rho$, where ρ is the jump rate of particles.

Lemma 3.1. *Fix constants $a_0, a_1, \varepsilon > 0$. Consider the coalescent started with n particles at $u \in G$ and none elsewhere: $X_v(0) = n\delta_{u,v}$. Let $\tau = a(\log n)^{-3}$ for some $a \in [a_0, a_1]$, and define the event*

$$A = \{\forall v, X_v(\tau) \in [(1-\varepsilon)Q_v, (1+\varepsilon)Q_v]\},$$

where

$$Q_v = \begin{cases} 2/\tau & v = u \\ (\rho/d_u)\log n & |v-u| = 1 \\ 0 & |v-u| > 1 \end{cases}.$$

Then there are constants γ, C depending only on $\varepsilon, a_0, a_1, d_u$ such that

$$\mathbb{P}^n(A^c) < Ce^{-\gamma\sqrt[3]{\log n}}$$

Proof. Let $Z(t)$ be the number of distinct labels corresponding to particles that exit u during $[0,t]$ (each label is counted at most once). Let $N(t)$ denote the total number of particles in the coupling with the mean-field coalescent of Lemma 2.5. Thus we have:

$$N(t) - Z(t) \leq X_u(t) \leq N(t) + Z(t), \text{ almost surely.} \tag{15}$$

We thus need to estimate $N(t)$ and $Z(t)$. For any fixed ε, by Lemma 2.2 we have

$$\mathbb{P}^n(|\tau N(\tau)/2 - 1| > \varepsilon) < Ce^{-c/\tau} < Cn^{-1}. \tag{16}$$

By Lemma 2.3:

$$\mathbb{P}^n\left(\left|\frac{Z(\infty)}{\log n} - \theta\right| > \varepsilon\right) < Cn^{-c}. \tag{17}$$

However, $Z(\infty) - Z(\tau)$ is the number of particles that exit after time τ. By (16) and Lemma 2.3

$$\begin{aligned}\mathbb{P}^n(Z(\infty) - Z(\tau) > \varepsilon \log n) &< \mathbb{P}^n(N(\tau) > 3/\tau) + \\ &\qquad \mathbb{P}^n\left(Z(\infty) - Z(\tau) > \frac{\varepsilon}{3a_1} N(\tau)^{1/3} \middle| N(\tau)\right) \leq 3/\tau)) \\ &< Cn^{-1} + C(\log n)^C e^{-c\sqrt[3]{\log n}} \\ &< Ce^{-\gamma \sqrt[3]{\log n}}.\end{aligned}$$

Thus with probability larger than $1 - Ce^{-\gamma\sqrt[3]{\log n}}$, we have $|Z(\tau) - Z(\infty)| < \varepsilon Z(\infty)$, and so with this probability, $|N(\tau) - 2/\tau| < 2\varepsilon/\tau$ while $Z(\tau) \leq 2\log n = o(N(\tau))$. In light of (15) this gives the required bounds on $X_u(\tau)$.

Conditioned on $Z(\tau)$, the probability that there is a particle that makes more than a single jump by time τ is at most $\rho\tau Z(\tau)$. Thus, given that $Z(\tau) \leq (1+\varepsilon)\log n$, the probability that any particle moves twice is at most $C(\log n)^{-2}$. However, if no particle jumps twice then there is no particle at distance greater than 1.

Similarly, given $Z(\tau)$, the probability that there is a coalescence event outside of u before time τ is at most $\tau\binom{Z(\tau)}{2}$. Given the bound on $Z(\tau)$, this is at most $C(\log n)^{-1}$. So with high probability (at least $1 - C(\log n)^{-1}$) there is no coalescence outside of u before time τ.

Finally, since $Z(\tau)$ is concentrated, and since each particle that moves selects a random neighbor of u, on the event that there are no further moves or coalescence events involving these particles, $X_v(\tau)$ is concentrated near $Z(\tau)/d_u$ for any $v \sim u$. The number of particles moving to any particular neighbor has variance of order $\log n$, and using a normal approximation to binomial random variables, the probability of deviating by $\varepsilon \log n$ from the mean is no more than Cn^{-c}. This completes the proof of the lemma. □

To get a hold on the global behavior of the spatial coalescent, a key idea is to iterate the estimates obtained in Lemma 3.1. We define recursively a sequence of times $(t_k)_{k\geq 1}$ at which we observe the process. Let $t_0 = 0$ and inductively let t_k be defined by $t_{k+1} = t_k + (\log^{(k)} n)^{-3}$, where $\log^{(k)}$ denotes the k^{th} iterate of the logarithm:

$$\log^{(k)} n = \underbrace{\log \ldots \log}_{k \text{ applications}} n.$$

Note that each increment in the sequence is much greater than all previous ones.

Let $m = \log^* n$, as defined in the introduction, and define m' by

$$m' = \min\left\{k : \exp\left(\gamma\sqrt[3]{\log^{(k+1)} n}\right) < \operatorname{Vol} B(o,m)^4\right\}, \tag{18}$$

where γ is as in Lemma 3.1. Note that m' is quite close to m (indeed, note first that $m' < m$ for large enough n, and moreover if D is the maximal degree of the graph, then with the trivial bound of $\operatorname{Vol} B(o,m) < D^m$ we get $m - m' < \log^*(Cm^3)$.)

Lemma 3.2. *Define the events*

$$B_k = \left\{\forall v, \frac{X_v(t_k)}{Q_k(v)} \in [\tfrac{1}{D} - \varepsilon, 1+\varepsilon]\right\},$$

where D is the maximal degree in G, and

$$Q_k(v) = \begin{cases} (\log^{(k)} n)^3 & |v| < k \\ \rho \log^{(k)} n & |v| = k \\ 0 & |v| > k. \end{cases}$$

Then with m' defined by (18)

$$\mathbb{P}\left(\bigcap_{k=1}^{m'} B_k\right) \xrightarrow[n\to\infty]{} 1.$$

Proof. Note that Lemma 3.1 gives us a bound on $\mathbb{P}^n(B_1^c)$. We proceed by induction, and use Lemmas 2.4 and 3.1 for each step. Assume B_k holds, then at time t_k all particles are in $B(o,k)$ and the number of particles at each site is between $\frac{\log^{(k)} n}{2D}$ and $2(\log^{(k)} n)^3$ (we may assume ε is small enough that B_k implies these bounds.)

With Lemma 2.4 in mind, consider the evolution over the time interval $t_{k+1} - t_k = (\log^{(k)} n)^{-3}$ of particles from a single site $u \in B(o,k)$. By Lemma 3.1 at time t_{k+1} the number of remaining particles is at most $(1+\varepsilon)(\log^{(k+1)} n)^3$, and the number of particles moving to each neighbor is $(1\pm\varepsilon)/d_u \log^{(k+1)} n$. Call a vertex where this fails to hold a *failure*. Note that if B_k holds and there is no failure at step k then B_{k+1} holds as well.

If the number of particles at the beginning of a stage at v is q then the probability of a failure at v at that stage is at most $C\exp\left(-\gamma\sqrt[3]{\log q}\right)$. Given the lower bound on q we have for large enough n,

$$\sqrt[3]{\log q} \geq \tfrac{1}{2}\sqrt[3]{\log^{(k+1)} n}.$$

It follows that the probability of failure for some $u \in B(o,k)$ at the k'th step is at most

$$C \operatorname{Vol} B(o,k) \exp\left(-\tfrac{\gamma}{2}\sqrt[3]{\log^{(k+1)} n}\right) < C \operatorname{Vol} B(o,m) \exp\left(-\tfrac{\gamma}{2}\sqrt[3]{\log^{(k+1)} n}\right).$$

Because the sequence $\log^{(k+1)} n$ decreases very quickly, the probability of failure in some step $k \leq m'-1$ (and using the definition of m') is at most

$$2C \operatorname{Vol} B(o,m) \exp\left(-\tfrac{\gamma}{2}\sqrt[3]{\log^{(m')} n}\right).$$

By the choice of m', this is at most $2C/\operatorname{Vol} B(o,m)$, which tends to 0 as $n \to \infty$. □

The previous lemma gave us a fairly accurate description of the spatial coalescent up to time $t_{m'}$ which is still $o(1)$, even though it decreases very slowly as $m \to \infty$. In order to push this analysis up to a constant time t, we need some additional estimates. We begin with the lower bound, since it is simpler. Henceforth, we let $t > 0$ be a fixed time.

Lemma 3.3. *Let m' be given by (18) and fix t. The collection $X_t(v)$ for $v \in B(o,m')$ stochastically dominates with high probability independent Bernoulli random variables ζ_v with mean $p > 0$ depending only on t. That is, the coalescent can be coupled with the independent variables ζ_v so that*

$$\mathbb{P}\big(\forall v \in B(o,m'), X_v(t) \geq \zeta_v\big) \xrightarrow[n\to\infty]{} 1.$$

Proof. By Lemma 3.2 there is some time $t' < t$ (for n large enough) so that with probability tending to 1, each site in $B(o,m')$ is not empty at time t'. On this event, fix one particle at each $v \in B(o,m')$, and color it red. Consider the evolution with coloring, so that if a red particle coalesces with another particle the newly formed particle retains the red color. Now, it is obvious that between time t' and t, each red particle has probability $e^{-\rho(t-t')} > e^{-\rho t}$ of not migrating, independently of all other red particles, so the claim holds. □

3.2 Upper bound

We now turn to the upper bound of Theorem 1.1. After step m', the detailed description of Lemma 3.2 will begin to break for some vertices, since as the number of particles per vertex decreases, the probability of failures becomes large. We overcome this by combining the second part of Lemma 2.3 with Lemma 2.7.

Lemma 3.4. *Fix $t > 0$ and assume $\operatorname{Vol} B(o,m) < Cm^d$ for some C, d. Then with high probability there is no particle outside $B(o, m + (\log m)^4)$ at or before time t.*

Proof. By Lemma 3.2, with high probability at time $t_{m'}$ the number of particles inside $B(o,m')$ is at most $3 \operatorname{Vol} B(o,m')(\log^{(m')} n)^3$, while there are no particles elsewhere. The bound on ball volumes together with some algebraic manipulation of (18) yields

$$(\log^{(m')} n)^3 < Ce^{C(\log m)^3},$$

for some C.

Disregard any coalescence after time $t_{m'}$, so that each particle performs a simple random walk independently of all other particles. The probability that any given particle makes more than $(\log m)^4$ steps by time t is at most $e^{-c(\log m)^4}$ (for some c depending on t.) Therefore the probability that any particle is outside $B(o, m + (\log m)^4)$ at or before time t is at most

$$3 \operatorname{Vol} B(o,m')(\log^{(m')} n)^3 e^{-c(\log m)^4} < C \exp\big(-c(\log m)^4 + C \log m^3 + C \log m\big)$$

which tends to 0 as $m \to \infty$. □

Proof of Theorem 1.1. The lower bound is a corollary of Lemma 3.3 which implies that

$$\mathbb{P}\big(N^n(t) < c \operatorname{Vol} B(o,m')\big) \xrightarrow[n\to\infty]{} 0.$$

Given the nice asymptotics for the volume of balls and that $m' \sim m$ we find $\operatorname{Vol} B(o,m) \sim \operatorname{Vol} B(o,m')$.

The upper bound is a corollary of Lemmas 2.7 and 3.4. By Lemma 3.4, no particles exit $B(o, m + (\log m)^4)$ before time t. By Lemma 2.7, the number of particles within the ball is with high probability at most a constant times its volume. □

Remark. The lower bound $N^n(t) > c \operatorname{Vol} B(o,m')$ holds for any bounded degree graph. However, in general this can be much smaller then $V(o,m)$.

For bounded degree graphs, the method above gives an upper bound of $c \operatorname{Vol} B(o,cm)$ for some c. This is because the number of particles at time $t_{m'}$ is exponential in m. It is possible to show that the bound $C \operatorname{Vol} B(o,m)$ also holds in this generality, though we do not include the proof. One way of achieving this is by considering the evolution of the total number of particles in $B(o,k)$ for $k > m'$, similarly to Section 6.

4 Results for spatial Beta-coalescents

We now turn to the proof of Theorem 1.2. In fact we are going to prove a slightly more general result. Assume that Λ has a density: $\Lambda(dx) = g(x)dx$, and that g is sufficiently regular near 0: there exists $B > 0$ and $1 < \alpha < 2$ such that

$$g(x) \sim Bx^{1-\alpha}, \qquad x \to 0 \tag{19}$$

Thus the case where Λ is the Beta$(2-\alpha, \alpha)$ is included. A consequence of (19) is the following estimate for the rate of coalescence events when there are n particles remaining:

Lemma 4.1. *λ_n is increasing in n. Furthermore, there exists $c > 0$ which depends only on B such that $\lambda_n \sim cn^\alpha$.*

Proof. The monotonicity of λ_n in n is a consequence of the natural consistency of Λ-coalescents. The second part of the statement is a consequence of (19) and Tauberian theorems. See, e.g., [10, Lemma 4] for more details. □

4.1 Lower bound

Define the following parameters

$$\beta = \frac{\alpha - 1}{2} \qquad \tau = an^{-\beta} \text{ for some } a \in [a_0, a_1] \qquad \gamma = \min\{1 - \alpha/2, \beta/2, 1/8\},$$

and observe that γ satisfies $\alpha - 2 + \gamma \leq -\gamma < 0$. We also consider the quantity

$$Y_n = \int_0^\tau N(s)\, ds.$$

Lemma 4.2. *Assume that Λ satisfies* (19)*. Then for some c, C depending only on Λ,*

$$\mathbb{P}(Y_n \geq n^{2-\alpha+\gamma}) \leq Cn^{-\gamma}, \tag{20}$$

and

$$\mathbb{P}(Y_n \leq cn^\gamma) \leq Cn^{-\gamma}. \tag{21}$$

Proof. The key fact is that if the process $N(t)$ attains some value k, then it stays at k for an exponentially distributed time with mean $1/\lambda_k$. Since the probability of hitting k is at most 1,

$$\mathbb{E}Y_n \le \sum_{k\le n} \frac{k}{\lambda_k} \le cn^{2-\alpha}$$

by Lemma 4.1. The upper bound (20) follows by Markov's inequality.

The lower bound is more delicate. We argue that with high probability the first $M = n^{\alpha-1+\gamma}$ jumps all occur before time τ and that throughout these jumps $N(t)$ remains above $n/2$. Summing only these first jumps will give the lower bound.

Let B_m be the number of particles lost in the next coalescence when there are m particles present. It is known [6, Lemma 7.1] that there exists $C > 0$ such that

$$\mathbb{P}(B_m > k) \le Ck^{-\alpha} \text{ for all } m, k \ge 1. \tag{22}$$

In particular, $\mathbb{E}B_m < c$ for some constant depending only on Λ. Thus the total size of the first M jumps has expectation at most cM. Let t_k be the time of the kth jump in $N(t)$, then by Markov's inequality

$$\mathbb{P}^n(N(t_M) < n/2) < \frac{cM}{n-n/2} < cn^{\alpha-2+\gamma} < cn^{-\gamma}. \tag{23}$$

On the event that $N(t_M) \ge n/2$, the rate of each of the first M jumps is at least $\lambda_{n/2}$. Thus, by Markov's inequality, and by monotonicity of λ_m,

$$\mathbb{P}(t_M > \tau, N(t_M) \ge n/2) \le \frac{M/\lambda_{n/2}}{\tau} \le cn^{-1+\gamma+\beta} < cn^{-\gamma}. \tag{24}$$

Thus, combining (24) with (23), $\mathbb{P}(A^c) < cn^{-\gamma}$, where $A = \{t_M < \tau, N(t_M) \ge n/2\}$.

Note that, on the event A,

$$Y_n = \int_0^\tau N(t)dt \ge \int_0^{t_M} N(t)dt \ge (n/2)t_M.$$

It thus suffices to show that $\mathbb{P}^n(t_M \le cn^{\gamma-1}) \le Cn^{-\gamma}$. However, the rate of each jump is at most λ_n, and therefore

$$t_M \succeq \sum_{i=1}^M E_i$$

where E_i are i.i.d. exponentials with rate λ_n. Now, from Lemma 4.1 we know that

$$\mathbb{E}\sum_{i\le M} E_i \sim cn^{\gamma-1},$$

and by Lemma 2.1 with $\epsilon = 1/2$,

$$\mathbb{P}\left(\sum_{i\le M} E_i < cn^{\gamma-1}/2\right) < \exp\left(-\frac{1}{16}n^{\alpha-1+\gamma}\right) < Cn^{-\gamma}$$

as desired. This completes the proof of Lemma 4.2. □

The next result gives a lower bound on the number of particles that exit the origin. This complements the upper bound of Lemma 2.6. The idea is that as long as Z is small, the true behaviour is close to the upper bound.

Lemma 4.3. *Let A be the event $\{Z_\tau < n^\gamma\}$. Then $\mathbb{P}(A) = O(n^{-\gamma})$.*

Proof. We introduce the random time T_a defined for any $0 < a < 1$ by $T_a = \inf\{t > 0 : Z(t) \geq aN(t)\}$. Define

$$A_1 = \{Z_{\tau \wedge T_a} \leq n^\gamma\} \qquad\qquad A_2 = A \cap \{\tau > T_a\}.$$

Note that $A \subset A_1 \cup A_2$ so it suffices to prove that $P(A_i) = O(n^{-\gamma})$, for $i = 1, 2$.

Consider A_1 first. Recall the notations introduced in Lemma 2.6, and note that T_a is a stopping time with respect to the filtration $\mathcal{F}^*$. Since $N(t)$ is non-increasing with limit 1 and since $Z(t)$ is non-decreasing and non-negative integer valued, T_a is finite if and only if at least one particle leaves o. This must eventually happen, so T_a is a.s. finite. Denote by $\tilde{\mathbb{P}}_n$ the law

$$\tilde{\mathbb{P}}_n(\cdot) = \mathbb{P}_n(\cdot|\mathcal{G}),$$

of all processes, conditioned on the entire evolution of N.

Consider the martingale S_t stopped at time T_a. By Doob's inequality, we find that for any $\delta > 0$

$$\begin{aligned}\tilde{\mathbb{P}}_n\left(\sup_{s\leq T_a}|S_s| \geq \delta\int_0^{T_a} N(s)ds\right) &\leq \frac{4\rho\tilde{\mathbb{E}}_n\left(\int_0^{T_a} M_u du\right)}{\delta^2\left(\int_0^{T_a} N(s)ds\right)^2} \wedge 1 \\ &\leq \frac{4\rho}{\delta^2\int_0^{T_a} N(s)ds} \wedge 1. \end{aligned} \tag{25}$$

The last inequality follows from the first bound of (12), which implies that $\tilde{\mathbb{E}}_n(\int_0^{T_a} M(u)du) \leq \int_0^{T_a} N(u)du$. Define the event

$$A_s = \left\{1 - a - \delta < \frac{Z_s}{\rho\int_0^s N(u)\,du} < 1 + \delta\right\}.$$

Until time T_a we have $M(t) \geq (1-a)N(t)$, and so (12) and (25) imply

$$\tilde{\mathbb{P}}_n(A_s^c) \leq \frac{4\rho}{\delta^2\int_0^{T_a} N(s)\,ds}.$$

We fix a and δ such that $1 - a - \delta > 1/2$. After taking the expectation, we obtain, using (20):

$$\mathbb{P}^n(A_1) \leq O(n^{-\gamma}) + n^{\alpha-2-\gamma} = O(n^{-\gamma})$$

Turning to A_2, note that

$$A_2 \subset \{aN(\tau) \leq n^\gamma\}$$

We claim that

$$\mathbb{P}^n(aN(\tau) \leq n^\gamma) \leq Cn^{-\gamma}. \tag{26}$$

To see this, we use the following very rough estimate. Note that by (22), there is a probability at least $1 - Cn^{-\gamma\alpha}$ that $N(s) \in [n^\gamma + 1, 2n^\gamma]$ for some s. In this case, the process will wait an amount of time greater than an exponential Y with rate λ_{2n^γ} before the next jump. It follows that (since $\gamma \leq \beta/2$ and $\alpha < 2$),

$$\begin{aligned}\mathbb{P}^n(N(\tau) \leq n^\gamma) &\leq \big(1 - Cn^{-\gamma\alpha}\big)\mathbb{P}(Y \leq \tau)\\ &\leq 1 - Cn^{-\gamma\alpha} - \exp(-c\tau n^{\alpha\gamma})\\ &\leq 1 - Cn^{-\gamma\alpha} - \exp(-cn^{-\gamma})\\ &< cn^{-\gamma}.\end{aligned}$$

This completes the proof of Lemma 4.3. □

We are now ready to start proving the lower-bound of Theorem 1.2. Let $t_0 > 0$ be a fixed time.

Lemma 4.4. *Fix constants $1 < a_0 < a_1$. Consider the coalescent started with n particles at $u \in G$ and none elsewhere: $X_v(0) = n\delta_{u,v}$. Let $\tau = an^{-\beta}$ for some $a \in [a_0, a_1]$, and define the event A by*

$$X_v(\tau) \geq n^\gamma/(4d_u) \qquad \textit{for } v = u \textit{ and any } v \sim u.$$

Then there are constants c, C depending only on a_0, a_1, d_u such that $\mathbb{P}(A^c) < Cn^{-c}$.

Proof. The statement corresponding to $v = u$ is simply (26).

For $v \sim u$, Lemma 4.3 gives a bound on the probability that not many particles leave the origin. It is overwhelmingly improbable that v does not get a constant times its fair share of these. It remains to estimate the number of particles that move to v and subsequently coalesce.

Indeed, even if all the particles migrate to v at time 0, so that they have strictly more opportunities to coalesce, the remaining number would still be large enough. Indeed, if all of them jump at time 0 onto v, it takes an exponential amount of time Y with parameter $\lambda_{n^\gamma/(4d)}$ before the first coalescence occurs. Since $\lambda_m \leq cm^\alpha$ for all $m \geq 1$, we deduce that $\mathbb{E}(Y) \geq cn^{-\gamma\alpha}$. However, since $\gamma \leq \beta/2$ and $\alpha < 2$, this is much larger than $\tau = an^{-\beta}$. It follows that $\mathbb{P}(Y < \tau) \leq cn^{\alpha\gamma-\beta}$.

Note also that by Lemma 4.1, the total jump rate of n^γ particles is smaller than the total coalescence rate (because $\alpha > 1$), so that the probability any of these particles makes an extra jump is even smaller than the above probability. It follows that at time τ, there are at least $n^\gamma/(4d)$ particles at v with probability at least $1 - Cn^{-c}$. □

Proof of Theorem 1.2(lower bound). Let $f_k(n) = f \circ f \ldots \circ f(n)$ (k times) where $f(n) = n^\gamma/4d$. We define the sequence of times $(\tau_k)_{k=1}^\infty$

$$\tau_k = \tau_{k-1} + af_{k-1}(n)^{-\beta}$$

It is easy to check that if we take $k = k(n) = \log\log n/(-2\log\gamma)$, then

$$f_k(n) \geq c\exp(\sqrt{\log n})$$

Let A' be the event that at each site within radius k there are at least $f_k(n)$ particles at time τ_k. On A', reasoning as in Lemma 3.4, (at each site of this ball at least one particle may remain with positive probability until time t_0), we see that $N^n(\tau) \geq \operatorname{Vol} B(o,k) \geq c \operatorname{Vol} B(o, c \log\log n)$ for some $c > 0$. Thus to finish the lower bound of Theorem 1.2, it suffices to compute the cumulative error probability in the iterated application of Lemma 4.4. However, it is trivial to check that:

$$\mathbb{P}(A'^c) \leq \sum_{i=1}^{k} C \operatorname{Vol} B(o,i) f_i(n)^{-\gamma} \leq Ck \operatorname{Vol} B(o,k) f_k(n)^{-\gamma}.$$

Since $\operatorname{Vol} B(o,k) < D^k$, where D is the maximal degree of the graph, this converges to 0 as $n \to \infty$. □

4.2 Upper bound

Let us now turn to the proof of the upper-bound, which takes only a few more estimates than the lower-bound.

Lemma 4.5. *In the spatial coalescent on any graph G, take some set $A \subset V$ and let Q_t be the number of particles that are present in A throughout the time interval $[0,t]$. Then there are constants $c, C > 0$ which depend only on Λ, such that:*

$$\mathbb{P}(Q_{t_0} > C t_0^{-1/(\alpha-1)} |A|) < \exp(-c|A|).$$

Proof. If particles exit A we ignore them from that time even if they return, hence we may assume that any particle leaving A is immediately killed. Even ignoring emigration, the main reason Q_t is small is coalescence. The rate of a coalescence event at a site v holding X_v particles is $\lambda_{X_v} \sim c(X_v)^\alpha$. At each such event at least one particle disappears, and therefore the total rate of decrease of Q_t is at least

$$\sum_{v \in A} (cX_v(t))^\alpha \geq c|A|^{1-\alpha} Q_t^\alpha$$

which is a consequence of Jensen's inequality, since $\alpha > 1$. (This is similar to [27, Theorem 12], but the above inequality is stronger). Thus Q_t is dominated by a pure death chain where the rate of decrease from i to $i-1$ is $c|A|^{1-\alpha} i^\alpha$, and these are the only transitions.

It is easy to finish using Lemma 2.1. Let E_k be independent exponential random variables with mean $\mu_k = c|A|^{\alpha-1} k^{-\alpha}$, and define $S_K = \sum_{k>K} E_k$. Then we have

$$\mathbb{P}(Q_t > K) < \mathbb{P}(S_K > t).$$

To apply Lemma 2.1 to S_K we need to estimate $\mathbb{E} S_K$ and $\operatorname{Var} S_K$: note that for suitable constants, as $K \to \infty$,

$$\mathbb{E} S_K = \sum_{k>K} \mu_k^{-1} \sim c_1 |A|^{\alpha-1} K^{1-\alpha} \tag{27}$$

and

$$\operatorname{Var} S_K = \sum_{k>K} \mu_k^{-2} \sim c_2 |A|^{2\alpha-2} K^{1-2\alpha}. \tag{28}$$

In particular $\frac{\operatorname{Var} S_K}{\mathbb{E}S_k \mu_K}$ is asymptotically constant and we may apply Lemma 2.1 with some constant ε. Thus for some $c_3 > 0$,

$$\begin{aligned}\mathbb{P}(S_K > 2\mathbb{E}S_K) &\leq \exp\left(-c\frac{(\mathbb{E}S_K)^2}{\operatorname{Var} S_K}\right) \\ &< e^{-c_3 K}.\end{aligned}$$

Now, if K is such that $\mathbb{E}S_K < t_0/2$ we may conclude that

$$\mathbb{P}(Q_{t_0} > K) < e^{-c_3 K}.$$

From (27) we see that $K = Ct_0^{-1/(\alpha-1)}|A|$ works for C large enough. □

Lemma 4.6. *Fix constants $a_0, a_1, \varepsilon > 0$. Consider the coalescent started with n particles at $u \in G$ and none elsewhere: $X_v(0) = n\delta_{u,v}$. Let $\tau = an^{-\beta}$ for some $a \in [a_0, a_1]$, and define the event A by*

$$X_v(\tau) \leq C_1 Q_v$$

with

$$Q_v = \begin{cases} n^{3/4} & \text{if } v = u \\ n^{2-\alpha+\gamma} & \text{if } |v-u| \leq r := \lceil 4/(\alpha-1) \rceil \\ 0 & \text{else} \end{cases}$$

Then there are constants C, C_1 depending only on Λ, a_0, a_1 such that $P(A^c) < Cn^{-\gamma}$.

Proof. With sufficiently high probability at most $n^{2-\alpha+\gamma}$ particles leave the origin by time τ (by Lemma 2.6 and (20)). This implies the bound for $0 < |v-u| \leq r$.

Some of the n^γ particles leaving u may coalesce before time τ (thus reducing further the number of particles), but we claim that with polynomially high probability, none of these particles may make more than r jumps by time τ. Indeed, the probability that by time τ, a given particle has jumped more than r times is smaller than $(\rho n^{-\beta})^r$ and there are at most n^γ particles which jump out of the origin. If r is such that $1 - r\beta < -\gamma$, the probability of any particle reaching distance r is indeed smaller than $C < n^{-\gamma}$, settling the case $|v-u| \geq r$. For the case $v = u$ we invoke Lemma 4.5 with an arbitrary set A containing u of size $c\log n$. If c is large enough then with probability at most $n^{-\gamma}$ we have $Q_\tau < Cn^{1/2}|A| < n^{3/4}$. However, $X_u(\tau) < Q_\tau + Z_\tau$, and so by Lemma 2.6 and (20) again, $X_u(\tau)$ is also small. □

Proof of Theorem 1.2: upper bound. Note that for any α we have $\gamma \leq \beta/2 < 1/2$. Let $c = \max(\gamma, 3/4)$, and note that $c < 1$. Let $C_2 = C_1 \times \operatorname{Vol} B(o,r)$, where C_1 and r are the constants in Lemma 4.6.

Let $f(n) = C_2 n^{3/4}$, and as before set $f_k(n) = f \circ \cdots \circ f(n)$ (k times). Also set $\tau_1 = \tau = n^{-\beta}$, and

$$\tau_k = \tau_{k-1} + af_{k-1}(n)^{-\beta}.$$

Let A_i be the event that at time τ_i there are no particles outside $B(o, ir)$ and that for all $|v| \leq ir$:

$$X_v(\tau_i) \leq f_i(n).$$

Choose $k = k(n)$ to be the maximal k so that $f_k(n) > \log n$. It is clear that $f_k(n) < (\log n)^2$. It is also straightforward to check that $k \sim c \log\log n$, and that $\tau_k = o(1)$.

Applying Lemma 4.6 iteratively, we see that

$$\mathbb{P}(A_k^c) \leq \sum_{i<k} C f_i(n)^{-\gamma} \operatorname{Vol}(B(o, ir)) \leq C \operatorname{Vol}(B(o, kr)) f_k(n)^{-\gamma} \xrightarrow[n\to\infty]{} 0. \tag{29}$$

Consequently, with high probability at time τ_k the total number of remaining particles is at most $C f_k(n) \operatorname{Vol}(B(o, kr))$, and all are inside the ball of radius kr.

Consider now the set $A = B(o, M \log\log n)$ for some large M to be specified soon. In order for any particle to exit A by time t it must survive to time τ_k and jump at least $M \log\log n - kr$ times by time t. Thus the expected number of particles that exit A by time t is at most

$$C f_k(n) \operatorname{Vol}(B(o, kr)) e^{-c(M \log\log n - kr)} < C(\log n)^2 (\log\log n)^d e^{-(cM - c') \log\log n}.$$

Fix M large enough that this tends to 0.

Finally, if no particle leaves A then $\sum_v X_v(t) = Q_t$. By Lemma 4.5, with high probability the number of particles that remain in A throughout $[0, t]$ is at most $O((\log\log n)^d)$. □

5 Global divergence of spatial Λ-coalescents

5.1 Infinite tree length for Λ-coalescents

Fix an arbitrary probability measure Λ on $[0, 1]$. Consider the corresponding mean-field Λ-coalescent that starts from a configuration consisting of infinitely many blocks, and let $(K^n(s), s \geq 0)$ be the number of blocks process of its restriction to the first n particles. Define:

$$X_n(t) \equiv X_n = \int_0^t (K^n(s) - 1) ds. \tag{30}$$

We are interested in the quantity X_n due to the following observation: if K^n is a good approximation for the number of blocks at the origin of the spatial Λ-coalescent at small times s, then for t small, ρX_n approximates well the number of particles that emigrate from the origin up to time t (see, for instance, Lemma 2.6). The key ingredient in the proof of Theorem 1.3 is the following result.

Lemma 5.1. *For any fixed $t > 0$ we have $X_n \xrightarrow[n\to\infty]{} \infty$ almost surely.*

Proof. Denote by $\sim^t$ the equivalence relation on the labels generated by the coalescent blocks at time t. For $n \geq 2$ let

$$\tau_n := \min\{t > 0 : \exists j < n \text{ s.t. } n \sim^t j\}$$

be the first time that the particle labelled n coalesces with any of the particles with smaller labels. We have that

$$K^n(s) = K^{n-1}(s) + \mathbf{1}_{\{s<\tau_n\}},$$

and therefore

$$X_n = X_{n-1} + (\tau_n \wedge t),$$

i.e. the contribution to X_n of particle n is $\tau_n \wedge t$.

Define $\mathcal{F}_n$ to be the σ-algebra generated by $\{K_s^j\}_{j\le n, s>0}$. Conditioned on $\mathcal{F}_{n-1}$, the infinitesimal rate of coalescence of particle n with particles with smaller labels at time s is given by

$$\int_{[0,1]} \frac{1}{x^2} \cdot x \cdot (1-(1-x)^{K^{n-1}(s)})\, d\Lambda(x).$$

Applying $(1-x)^k \ge 1-kx$ (for $x \in [0,1]$) we find that the rate of coalescence of particle n is at most $K^{n-1}(s)$ (with equality if and only if Λ is the point mass at 0, in which case the coalescent is Kingman's coalescent). Thus

$$\begin{aligned}
\mathbb{E}(\tau_n \wedge t | \mathcal{F}_{n-1}) &= \int_0^t \mathbb{P}(\tau_n > s | \mathcal{F}_{n-1})\, ds \\
&\ge \int_0^t \exp\left(-\int_0^s K^{n-1}(u)\, du\right) ds \\
&\ge \int_0^t \exp\left(-s - \int_0^t (K^{n-1}(u) - 1)\, du\right) ds \\
&= e^{-X_{n-1}} \int_0^t e^{-s} ds = (1-e^{-t}) e^{-X_{n-1}}.
\end{aligned}$$

Note that X_n is increasing and consider the martingale

$$M_n = X_n - \sum_{k=2}^n \mathbb{E}(\tau_k \wedge t | \mathcal{F}_{k-1}).$$

On the event that X_n is bounded, the last calculation implies that $\mathbb{E}(\tau_k \wedge t|\mathcal{F}_{k-1})$ is bounded from below, hence $M_n \to -\infty$. Since M is a martingale, the last event has probability 0. □

Note that a different proof of Lemma 5.1 follows from Corollary 3 in [3], although the arguments there are significantly more involved.

5.2 Proof of Theorem 1.3

We now consider the spatial coalescent corresponding to some fixed Λ as in the previous section, on an arbitrary locally finite graph G. As usual, let n denote the initial size of the population, with all particles initially located at o, a fixed vertex of G. Recall the definitions of the processes M and Z in Lemma 2.5. Both processes M and Z depend implicitly on n, omitted from the notation. We consider the usual coupling of coalescents that correspond to different n.

Lemma 5.2. *For any $t > 0$ we have that $Z(t) \to \infty$ almost surely as $n \to \infty$.*

Proof. We follow the argument of Lemma 4.3, except that we are only interested in showing that Z diverges, which simplifies the argument. Since $Z(t)$ is non-decreasing in n it suffices to show that for any fixed m we have $\mathbb{P}(Z(t) < m) \xrightarrow[n\to\infty]{} 0$.

Recall the martingales (14). On the event $\{Z(t) \le m\}$, we have for all $s \le t$ that $M_s \ge N(s) - Z_s \ge N(s) - m$, and therefore $S_t \le m + \rho m t - \int_0^t \rho N(t)dt$. Due to Lemma 5.1, for any fixed m, t and any sufficiently large n, on the event $\{Z(t) \le m\}$ (this event also depends on n)

$$S_t \le -\frac{1}{2}\int_0^t \rho N(t).$$

As in (25), Doob's maximal inequality yields that for large enough n

$$\mathbb{P}(Z_t < m|\mathcal{G}) \le \mathbb{P}\left(\sup_{s\le t}|S_s| \ge \frac{1}{2}\int_0^t \rho N_s ds \,\middle|\, \mathcal{G}\right) \le \frac{16}{\rho\int_0^t N_s ds}.$$

By Lemma 5.1 the right-hand side tends to 0 almost surely, so $\mathbb{P}(Z(t) < m) \xrightarrow[n\to\infty]{} 0$. ☐

Fix $\varepsilon > 0$ and a vertex v of the graph, and let $E_{m,\varepsilon,v} = \{\sup_{t\in[0,\varepsilon]} X_v^n(t) \ge m\}$ be the event that at some time $t < \varepsilon$ there are at least m particles located at site v.

Lemma 5.3. *We have* $\mathbb{P}^n(E_{m,\varepsilon,v}) \xrightarrow[n\to\infty]{} 1$.

Proof. Note that the claim is trivially true if $v = o$. We prove it first for v a neighbor of o. Take

$$t_0 = \eta \min\{\varepsilon, \lambda_m^{-1}, (\rho m)^{-1}\},$$

where η is an arbitrarily small number. Now, choose $n_0 = n_0(t_0)$ large enough that $Z_{t_0} > 2dm$ with probability at least $1 - \eta$, where $d = \deg(o)$. By the weak law of large numbers, one can choose $n_0 = n_0(t_0, h)$ large enough that on the event $\{Z_{t_0} > 2dm\}$, v receives at least m particles from o with probability at least $1 - \eta$. We concentrate on this event of high probability, and on these m particles, ignoring any further particles that might visit v.

Jumps from v occur at rate ρ per particle, so the probability that any of the m particles above leave v before time t_0 is at most η. Since a coalescent event involving any k-tuple of particles occurs at total rate λ_k (increasing in k), and since at a given time $s < t_0$ there are up to m of the above particles located at v, the probability of a coalescent event before time t_0 in which two or more of the m above particles participate is at most η. It follows that $\mathbb{P}(X_v^n(t) \ge m) \ge 1 - 4\eta$, for all $n > n_0$. Since η can be made arbitrarily small, this proves our claim for v a neighbor of o.

For other v we use induction in the distance $|v|$ to o. Indeed, such v has a neighbor u satisfying $|u| < |v|$. For any fixed m', η, and n sufficiently large, we have $\mathbb{P}^n(E_{m',\varepsilon,u}) \ge 1 - \eta$. Given this, and using the strong Markov property, one can repeat the previous argument with m' sufficiently large to conclude that with probability at least $1 - 2\eta$ there will be at least m particles at v (arriving from u) at some time $t < 2\varepsilon$. ☐

Proof of Theorem 1.3. Again, due to monotonicity in n and t, it suffices to show that for any $m < \infty$ and any $t > 0$, we have $\lim_{n\to\infty} \mathbb{P}(N_t^n > m) = 1$.

Let $\eta, \varepsilon > 0$ be small numbers. Fix $m < \infty$, and choose a subgraph $G^m \subset G$ of size m such that the distance between any two vertices of G^m is larger than $1/\eta$. By Lemma 5.3 we have that

$$\mathbb{P}^n(E_{1,\varepsilon,v}) \xrightarrow[n\to\infty]{} 1, \ \forall v \in G^m.$$

Moreover, if $A_{v,\varepsilon}$ is the event that the first (if any) particle that enters v before time ε stays at v up to time ε (while it may possibly coalesce with other particles), note that $\mathbb{P}(A_{v,\varepsilon}|E_{1,\varepsilon,v}) \geq e^{-\rho\varepsilon}$. By choosing ε sufficiently small we arrive to

$$\lim_{n\to\infty} \mathbb{P}(\cap_{v\in G^m} A_{v,\varepsilon}) \geq 1 - \eta.$$

However, given $\cap_{v\in G^m} A_{v,\varepsilon}$, the probability that any pair of the above particles (located at mutual distance greater than $1/\eta$ at time ε) will coalesce before time t tends to 0 as $\eta \to 0$. □

6 Lower bound for the long time asymptotics

We now turn to the large time asymptotic behaviour of spatial coalescents. The underlying measure Λ does not play an important role here as it did for the behaviour at constant times. The reason for this is that, as explained in the introduction (Section 1.2), at the beginning of this phase, say at constant time $t > 0$, the number of particles at each site is tight with respect to n. When the number of particles at a site is small, the coalescents corresponding to different choices of Λ behave similarly. In fact, the density of particles quickly decays, and once it is small enough, it rarely happens that more than two particles are at the same location. With at most two particles at each site, any spatial Λ-coalescent is equivalent to spatial Kingman's coalescent.

An important quantity in this setting is the radius m of the region (ball) which is initially "filled" with particles. As we have seen, for Kingman's coalescent the radius of this ball is $m = \log^* n$, while in the case of Beta-coalescents it is approximately $m = \log\log n$ up to constants. In the general case, the radius m should be a certain function of both n and Λ, namely $m = f^*(n)$ where f^* is defined in (3). This was rigorously established only for Kingman's coalescent and those with "regular variation" (i.e., satisfying (19)). However, the results which we present in this section and the next one, are valid for essentially arbitrary coalescence mechanisms (subject to (49) for the upper bound in Section 7), and assume that the spatial Λ-coalescent starts with a possibly random but tight number of particles per site in a large ball of radius m. See Theorem 6.4 for the full statement. Note that in this result as in the rest of the paper, we will be taking limits as m tends to ∞, recalling that m is itself a function of n when applying these results to get Theorem 1.4.

Let us first present some further heuristic arguments for the lower bound in Theorem 1.4. Consider for the moment the case $d > 2$, so that the random walk migration process is transient. The first heuristic comes from the first moment calculation and simple Green function estimates: label the particles in an arbitrary way and let S_i be the total number of particles that ever coalesce with the ith particle. Observing that a typical particle is at initial distance of order k away from an order k^{d-1} particles, where k ranges from 1 up to a number of order m, gives for a typical i

$$\mathbb{E}(S_i) \asymp \sum_{k=1}^{m} k^{d-1} \frac{1}{k^{d-2}} \asymp m^2,$$

where we use the fact that the probability that two particles ever coalesce is proportional to the probability that their corresponding walks intersect (visit the same site *at the same time*) (the constant comes from the delayed coalescence dynamics). The fact that this probability is approximately k^{2-d} is a well-known Green function estimate. Since $N = \sum_{i=1}^{M} 1/S_i$ gives the total number of clusters that survive forever (with M being the initial number of particles, of order m^d), and since $\mathbb{E}(1/S_i) \geq 1/\mathbb{E}(S_i)$ we arrive at

$$\mathbb{E}(N) \geq cm^{d-2}. \tag{31}$$

While Jensen's inequality may seem crude, this does give the correct exponents because the distribution of S_i is sufficiently concentrated. The next section contains results confirming this heuristic.

6.1 Technical random walk lemmas

We begin with technical results concerning random walks. Most of these are standard yet difficult to "pinpoint" in the random walk literature. Let $(S_n, n \geq 0)$ be simple symmetric random walk on $\mathbb{Z}^2$, started from a point $X_0 \in B(o, 2m)$ which will later be chosen in a certain random fashion (very roughly speaking, close to uniform) and recall that $m \to \infty$. Let $(X_t, t \geq 0)$ be a continuous time random walk on $\mathbb{Z}^2$ obtained as $X_t := S_{N_t^*}$ where $(N_t^*, t \geq 0)$ is a Poisson process with rate 1, independent of X. Let S' be a lazy version of S, with $S'_0 = S_0$ and step distribution given by $\mathbb{P}(S'_{n+1} - S'_n = 0) = 1/2$ and $\mathbb{P}(S'_{n+1} - S'_n = \pm \mathbf{e}_i) = 1/8$, where $\mathbf{e}_1, \mathbf{e}_2$ are the coordinate vectors, then $(S'_{N_{2t}^*}, t \geq 0)$ has the same law as X. We write $\mathbb{P}_x$ for the corresponding probability measures when $X_0 = x$.

Define $\tau'_x := \inf\{n \geq 0 : S'_n = x\}$, $\tau_x := \inf\{s > 0 : X_s = x\}$ to be the hitting times of x. Similarly, let ${\tau'}_x^+ := \inf\{n \geq 1 : S'_n = x\}$ denote the *positive* hitting time of x. We abbreviate $\tau' = \tau'_0$, $\tau = \tau_0$ etc.

The next result is a variation of an Erdös-Taylor formula [15] (see also [13], p. 354). We assume that as $m \to \infty$,

$$\mathbb{E}\left(\frac{1}{\|X_0\|_+^2}\right) = O\left(\frac{1}{\log m}\right), \qquad \mathbb{E}\left(\log \frac{m}{\|X_0\|_+}\right) = O(1), \tag{32}$$

where for any $y \in \mathbb{Z}$ we abbreviate $\|y\|_+ := \|y\| \vee 1$.

Lemma 6.1. *Assume $d = 2$ and fix $s > 4$, assume a random $\|X_0\| \leq 2m$ satisfies* (32)*. Then*

$$\mathbb{P}(\tau < sm^2) \asymp \frac{\log s}{\log m + \log s}, \tag{33}$$

where the constants implicit in the $\asymp$ notation depend only on those implicit in (32) *(and not on s or m).*

It is easy to check that X_0 drawn from a uniform on $B(o,2m)$ or from a difference of two independent uniforms on $B(o,m)$ will satisfy the hypotheses of Lemma 6.1 and therefore (33) with universal constants (not depending on m) for any fixed $s>4$. Also note that if $\mathbb{P}(X_0 \notin B(o,2m)) = 1$, under no further restriction on the distribution of X_0, the upper bound on the probabilities $\mathbb{P}(\tau < sm^2)$ holds with the same constant as in Lemma 6.1. Indeed, in order for τ to happen, the walk needs first to enter $B(o,2m)$ at a location close to its boundary, for which the argument below gives the required estimate.

Proof. We estimate the above probability for any given $x \in B(o,2m)$, and then integrate over the law of X_0. Without loss of generality, assume that sm^2 is an integer. Use the "last-exit decomposition":

$$\mathbf{1}_{\{\tau'<sm^2\}} = \sum_{k=1}^{sm^2-1} \mathbf{1}_{\{S'_k=0\}} \prod_{j=k+1}^{sm^2-1} \mathbf{1}_{\{S'_j\neq 0\}}$$

together with the Markov property, to obtain

$$\mathbb{P}_x(\tau' < sm^2) = \sum_{k=1}^{sm^2-1} \mathbb{P}_x(S'_k = 0)\mathbb{P}_0(\tau'^+ > sm^2 - k - 1).$$

We now apply a local central limit theorem and an estimate on the distribution of excursion length [33], statement E1 on p. 167, (or [25, Prop. 4.2.4]) and [33] statement P10 on p. 79 (or [25, Theorem 2.1.1]). We find that for some universal sequence $e_n \xrightarrow[n\to\infty]{} 0$

$$\begin{aligned}\mathbb{P}_x(\tau' < sm^2) &= \sum_{k=1}^{sm^2-1} \frac{1}{k} e^{-\frac{2\|x\|^2}{k}} \frac{1+e_{sm^2-k+1}}{\log(sm^2-k+1)} + O\left(\sum_{k=1}^{sm^2-1} \frac{1}{k\|x\|_+^2} \frac{1}{\log(sm^2-k+1)}\right)\\ &= \sum_{k=1}^{sm^2-1} \frac{1}{k} e^{-\frac{2\|x\|^2}{k}} \frac{1+e_{sm^2-k+1}}{\log(sm^2-k+1)} + O\left(\frac{1}{\|x\|_+^2}\right).\end{aligned}$$

Split this sum in three: For $k \le \|x\|^2$ use $e^{-x} < x^{-1}$ to get a total contribution of $O(1/\log(sm^2))$. For $k > sm^2 - \sqrt{sm^2}$ each term is at most C/k so the total contribution is $O(1/(\sqrt{s}\cdot m))$. Finally, for the intermediate k's each term is $\asymp 1/k \log sm^2$, so the total contribution is $\asymp \log(sm^2/\|x\|^2)/\log(sm^2)$. Thus

$$\mathbb{P}_x(\tau' < sm^2) \asymp \frac{\log s + 2\log(m/\|x\|_+) + O(1)}{\log(sm^2)} + O\left(\frac{1}{\|x\|_+^2}\right),$$

uniformly over $x \in B(o, sm/2)$. Taking expectation with respect to X_0 while using (32) and $s>2$, yields $\mathbb{P}(\tau' < sm^2) \asymp \frac{\log s}{\log(sm^2)}$ as $m \to \infty$.

Going back to the continuous time random walk, we have $\mathbb{P}(|N^*_{2t} - 2t| > \varepsilon t) \le e^{-c(\varepsilon)t}, t \ge 0$ for some $c(\varepsilon) > 0$, accounting for an additional error of $O(e^{-c(\varepsilon)sm^2}) = o(1/\log(sm))$ in the corresponding estimate for τ. ☐

We will also need later a simpler result which goes along the same lines.

Lemma 6.2. *Assume $d=2$ and $X_0 = x$ is such that $\|x\| = m$. For all $c_1 > 0$, there exists $c_2 > 0$ which depends only on c_1 such that $\mathbb{P}_x(\tau_0 < c_1 m^2) \ge c_2/\log m$.*

Proof. First we note that by easy large deviations on Poisson random variables, it suffices to prove the same inequality with τ_0 replaced by the discrete time τ_0'. By the strong Markov property, note that if $K(t)$ counts the number of hits of 0 by time t, then for all $c > 0$,

$$\mathbb{E}_x\big(K(cm^2)\big) \leq \mathbb{P}_x(\tau_0' \leq cm^2)\mathbb{E}_0(K(cm^2)). \tag{34}$$

By the local central limit theorem,

$$\mathbb{E}_0(K(cm^2)) = \sum_{k=0}^{cm^2} \mathbb{P}_0(S_k' = 0) \asymp \sum_{k=1}^{cm^2} \frac{1}{k} \sim 2c \log m. \tag{35}$$

Also,

$$\mathbb{E}_x(K(cm^2)) = \sum_{k=0}^{cm^2} \mathbb{P}_0(S_k' = x) \asymp \sum_{k=1}^{cm^2} \frac{e^{-c'\|x\|^2/(2k)}}{k} \geq \sum_{k=cm^2/2}^{cm^2} \frac{e^{-c'm^2/(2k)}}{k} \geq c''. \tag{36}$$

Combining (34)–(36), we complete the proof. □

If $d \geq 3$, we denote by G_X the Green function of a d-dimensional walk X. It is well-known (see e.g. [33]) that

$$G_X(x) \sim c\|x\|^{2-d}, \text{ as } \|x\| \to \infty, \tag{37}$$

for some constant c that depends on d (here $\|x\|$ denotes the Euclidean norm in $\mathbb{Z}^d$).

Let $X(\cdot), Y(\cdot)$ be independent continuous random walks in $\mathbb{Z}^d$ $d \geq 2$ with jump rate 1 and with starting points uniform in $B(o, m)$. Denote by $\sigma_t\{X\} := \sigma\{X(s), 0 \leq s \leq t\}$ the natural filtration of X and let $\sigma\{X\} := \sigma_\infty\{X\} = \sigma\{X(s), 0 \leq s < \infty\}$. Define the stopping time $\tau := \inf\{t : X(t) = Y(t)\}$. Define the collision events by

$$H_s \equiv H_s(X, Y) := \{\tau \leq s\}, \qquad H \equiv H(X, Y) := \{\tau < \infty\}.$$

Lemma 6.3. *Let X, Y be independent continuous time random walks in $\mathbb{Z}^d$ starting at uniform points at $B(o, m)$. For any $d > 2$ we have*

$$\mathbb{P}(H) \asymp m^{2-d}, \tag{38}$$

$$\operatorname{Var}\left(\mathbb{P}(H|\sigma\{X\})\right) \leq Cm^{2(2-d)}, \tag{39}$$

while if $d = 2$, for any $t > 4$ we have

$$\mathbb{P}\big(H_{tm^2}\big) \asymp \frac{\log t}{\log m + \log t}, \tag{40}$$

$$\operatorname{Var}\left(\mathbb{P}(H_{tm^2}|\sigma\{X\})\right) \leq C\left(\frac{\log t}{\log m + \log t}\right)^2, \tag{41}$$

where C and the constants in $\asymp$ relation depend only on d.

Proof. Assume first that $d \geq 3$. Note that the difference $X(t) - Y(t)$ is also a continuous time simple random walk (with a doubled rate of jumps), and abbreviate $G_{X-Y} = G$. It is well-known and easy to check that

$$\mathbb{P}(H | X_0 = x_0, Y_0 = y_0) = \frac{G(x_0 - y_0) - \mathbf{1}_{\{x_0 = y_0\}}}{G(0)}. \tag{42}$$

Since $x \mapsto \|x\|^{2-d}$ is integrable near 0 as a function on $\mathbb{R}^d$, then (37) implies that

$$\frac{1}{\operatorname{Vol} B(o,m)} \sum_{y \in B(o,m)} G(x-y) \leq C m^{2-d}, \qquad \text{for any } x \in \mathbb{Z}^d. \tag{43}$$

If $x \in B(o,m)$, then a corresponding lower bound holds since a positive fraction of points in $B(o,m)$ is at distance order m from x. Hence, for any $x \in B(o,m)$,

$$\frac{1}{\operatorname{Vol} B(o,m)} \sum_{y \in B(o,m)} G(x-y) \asymp m^{2-d},$$

where the constants implicit in $\asymp$ depend only on d. Due to (42), averaging over $x \in B(o,m)$ gives that $\mathbb{P}(H) \asymp m^{2-d}$ as claimed. (It is not hard to show similarly that $\mathbb{P}(H) \sim cm^{2-d}$ for some c.)

In order to show (39), introduce a third random walk Y' independent from, and identically distributed as, X and Y. In analogy to H define $H' = \{\exists t, X(t) = Y'(t)\}$. Given $\sigma\{X\}$, the events H, H' are independent and have the same probability. Thus

$$\begin{aligned}
\operatorname{Var} \mathbb{P}(H|\sigma\{X\}) &\leq \mathbb{E}\big[\mathbb{P}(H|\sigma\{X\})^2\big] \\
&= \mathbb{E}\big[\mathbb{P}(H|\sigma\{X\})\mathbb{P}(H'|\sigma\{X\})\big] \\
&= \mathbb{E}\big[\mathbb{P}(\exists t, s : X(t) = Y(t), X(s) = Y'(s)|\sigma\{X\})\big] \\
&\leq 2\mathbb{P}(\exists t, s : t \leq s, X(t) = Y(t), X(s) = Y'(s)),
\end{aligned}$$

where for the last inequality we use the symmetry between Y and Y'. Denote by $\mathcal{F}_\tau$ the standard σ-field generated by processes X and Y up to time τ. On the event $\{\tau < \infty\}$, due to the strong Markov property and (42),

$$\mathbb{P}\big(\exists s \geq \tau : X(s) = Y'(s) \mid \mathcal{F}_\tau\big) \leq c\mathbb{E}\big[G(X(\tau) - Y'(\tau)) \mid \mathcal{F}_\tau\big].$$

Let $Z = X(\tau) - (Y'(\tau) - Y'(0))$. Noting that $Y'(0)$ is independent from both $\mathcal{F}_\tau$ and Z, we have $\mathbb{E}(G(Z - Y'(0))|\mathcal{F}_\tau, Z) \leq Cm^{2-d}$, almost surely, and therefore

$$\mathbb{E}\big(G(X(\tau) - Y'(\tau))|\mathcal{F}_\tau\big) = \mathbb{E}[\mathbb{E}\big(G(Z - Y'(0))|\mathcal{F}_\tau, Z\big)|\mathcal{F}_\tau] \leq Cm^{2-d}.$$

In view of the discussion above this yields a uniform bound on $\operatorname{Var} \mathbb{P}(H|\sigma\{X\})$.

If $d = 2$, we proceed similarly, with H replaced by H_{tm^2}. In particular, Lemma 6.1 gives the asymptotics of $\mathbb{P}(H_{tm^2})$. For the conditional variance estimate, one obtains as above

$$\operatorname{Var} \mathbb{P}(H_{tm^2}(X,Y)|\sigma\{X\}) \leq 2\mathbb{P}[\mathbf{1}_{\{\tau<t\}}\mathbb{P}(H_{tm^2}(X'', Y'')|\mathcal{F}_\tau)],$$

where X'', Y'' are independent random walks started from $X(\tau)$ and $Y'(\tau)$, respectively, and otherwise independent of $\mathcal{F}_\tau$. The result follows as before, since by Lemma 6.1, $\mathbb{P}(H_{tm^2}(X'', Y'')|\mathcal{F}_\tau) \asymp \frac{\log t}{\log m + \log t}$. □

6.2 Proof of the lower bound

We return to the spatial coalescent. Let Λ be an arbitrary finite measure on $(0,1)$. Consider a spatial coalescent with initial configuration $X(0)$ that stochastically dominates i.i.d. Bernoulli random variables with mean $p > 0$ in $B(o,m)$ (we make no assumptions on the initial configuration outside of $B(o,m)$). With a slight abuse of notation, we write $N^m(t)$ in this section for the total number of particles at time t, and we define $N \equiv N^m = \lim_{t\to\infty} N^m(t)$ be the number of particles that survive to time ∞.

Theorem 6.4. *Consider the spatial coalescent with initial state dominating Bernoulli variables in $B(o,m)$. If $d > 2$, then there exist a constant $a > 0$ such that*

$$\mathbb{P}(N > am^{d-2}) \xrightarrow[m\to\infty]{} 1.$$

If $d = 2$, then there exists a constant $a > 0$ such that, for any $t > 4$,

$$\mathbb{P}\left(N^m(tm^2) > a\frac{\log m}{\log t}\right) \xrightarrow[m\to\infty]{} 1.$$

Note that, since the total number of particles is non-increasing, the lower bound in the $d = 2$ case holds for any $t > 1$ with modified constant a (or with $\log(2+t)$ in place of $\log t$ for any positive t).

We begin with a lemma stating a similar result for a simpler initial condition and with an "instantaneous" coalescent mechanism, where two particles coalesce as soon as they visit the same site. This model is called *coalescing random walks* (CRW). Afterwards we couple the two models to obtain Theorem 6.4.

Lemma 6.5. *Consider a system of s coalescing random walks, such that their initial positions are i.i.d. uniform points in $B(o,m)$, where*

$$s \equiv s(a) = \begin{cases} am^{d-2}, & d \geq 3, \\ a\log m, & d = 2. \end{cases}$$

Let $Z(t)$ denote the total number of particles at time t and let $Z = \lim_{t\to\infty} Z(t)$. If $d > 2$, then for some $a > 0$ we have $\mathbb{P}(Z > am^{d-2}/4) \xrightarrow[m\to\infty]{} 1$.
If $d = 2$, then for some a and all $t > 4$, we have $\mathbb{P}\left(Z(tm^2) > a\frac{\log m + \log t}{4\log t}\right) \xrightarrow[m\to\infty]{} 1$.

Proof. We use the following explicit construction of the CRW model with the given initial condition: Let $(X_i(t), t \geq 0), i = 0, 1, \ldots, s-1$ be a family of i.i.d. (non-coalescing) random walks, such that for each i, $X_i(0)$ is uniform in $B(o,m)$. At time 0, each block contains a single particle that is assigned a unique label in $\{0, 1, \ldots, s-1\}$. While present in the system, the particle (or block of particles) carrying label i follows the trajectory of X_i. If the trajectories of blocks labeled i and j ever intersect, they instantaneously merge into a new block that inherits the smaller label $i \wedge j$.

Consider first the case $d > 2$. For each pair i, j let $A_{i,j} := \{\forall u \geq 0 : X_i(u) \neq X_j(u)\} = H(X_i, X_j)^c$. Then on $A_{i,j}$ the blocks carrying labels i and j cannot merge as a consequence of a single coalescence event, but might merge due to a collection of coalescence events involving lower indexed particles. However, on the event

$$A_k := \bigcap_{i<k} A_{k,i}, \tag{44}$$

the block carrying label k stays in the system indefinitely.

Consider the filtration $\mathcal{F}_k = \sigma\{X_i(\cdot), i \leq k\}$. Define $p_k = \mathbb{P}(A_k|\mathcal{F}_{k-1})$, and note that $p_0 = 1$. The random variables $\{p_k\}$, are a non-increasing sequence of random variables. To see this we use the fact that the random walks are independent and so

$$p_k \leq \mathbb{P}\left(\bigcap_{i=0}^{k-2} A_{k,i}|\mathcal{F}_{k-1}\right) = \mathbb{P}\left(\bigcap_{i=0}^{k-2} A_{k-1,i}|\mathcal{F}_{k-2}\right) = p_{k-1}, \text{ almost surely.}$$

Next, define events

$$B_k = A_k \cup \{p_k < 1/2\},$$

and note that

$$\mathbb{P}(B_k|\mathcal{F}_{k-1}) = \begin{cases} 1, & p_k < 1/2, \\ p_k, & p_k \geq 1/2. \end{cases}$$

Consider the martingale

$$M_k = \sum_{i=0}^{k} \mathbf{1}_{B_i} - \mathbb{P}(B_i|\mathcal{F}_{i-1}).$$

Note that M_k has increments with variance bounded (crudely) by 1. Thus $\operatorname{Var} M_s < s$ (here s is the initial total number of blocks) and, by Markov's inequality,

$$\mathbb{P}(|M_s| > s/4) \leq \frac{s}{(s/4)^2} = \frac{16}{s}.$$

However, $\mathbb{P}(B_k|\mathcal{F}_{k-1}) \geq 1/2$, so by the definition of M, we find

$$\mathbb{P}\left(\sum_{i<s} \mathbf{1}_{B_i} < s/4\right) \leq \frac{16}{s} \xrightarrow[m\to\infty]{} 0. \tag{45}$$

Since p_k is non-increasing and since on the event $\{p_k \geq 1/2\}$ the events A_k and B_k coincide, we realize that on the event $\{p_s \geq 1/2\}$

$$\sum_{i<s} \mathbf{1}_{A_i} = \sum_{i<s} \mathbf{1}_{B_i}.$$

Thus if we prove that

$$\mathbb{P}(p_s < 1/2) \xrightarrow[m\to\infty]{} 0, \tag{46}$$

then (45) would imply the lemma. To this end we show that p_s is bounded below by a random quantity that is concentrated above 1/2, via second moment estimates. Specifically, from the definition (44) we have

$$1 - p_s \leq \sum_{i<s} \mathbb{P}(A_{s,i}^c|\mathcal{F}_{s-1}) = \sum_{i<s} \mathbb{P}(A_{s,i}^c|\sigma\{X_i\}),$$

where the last identity is due to independence of $\sigma\{X_i\}$ for different i's. Moreover, $\{\mathbb{P}(A^c_{s,i}|\sigma\{X_i\}), i = 0, \dots, s-1\}$ is an i.i.d. family of random variables. Using (38),

$$\mathbb{E}\left(\sum_{i<s} \mathbb{P}(A^c_{s,i}|\sigma\{X_i\})\right) < s \cdot Cm^{2-d} \leq Ca.$$

We choose $a = 1/(4C)$ so that this expectation is at most $1/4$. Due to (39),

$$\mathrm{Var}\left(\sum_{i<s} \mathbb{P}(A^c_{s,i}|\sigma\{X_i\})\right) \leq s \cdot Cm^{2(2-d)} \to 0.$$

so the sum is concentrated near its mean, and (46) follows.

In the case $d = 2$, the proof is almost identical. We take $s = a \log m$ and $a < 1/(4C \log t)$, where C is the constant that appears in (40). The event $A_{i,j}$ is accordingly redefined as $A_{i,j} := H_{tm^2}(X_i, X_j)^c$. Otherwise, the argument proceeds exactly as above, with (40), (41) used in place of (38), (39). □

Proof of Theorem 6.4. The idea is to couple the spatial coalescent X with a system of coalescing random walks, denoted $X^\diamond$, with an initial state of s particles at i.i.d. sites, uniform in $B(o, m)$. We first argue that it is possible to couple the initial states so that w.h.p. $X^\diamond_v(0) \leq X_v(0)$ (at every vertex). Indeed, in $X^\diamond$, there are $N^\diamond(0) \leq s$ occupied sites (since there may be repetitions) and given S, these sites are uniformly sampled from the ball $B(0, m)$ without replacement. On the other hand, $X(0)$ dominates a Bernoulli configuration on $B(o, m)$, hence $X(0)$ has at least $\mathrm{Bin}(\#B(o, m), p)$ particles sampled without replacement. Since $\mathbb{P}(\mathrm{Bin}(\#B(o, m), p) > s) \to 1$, this holds.

The second step of the proof is that if the initial configurations satisfy $X^\diamond_v(0) \leq X_v(0)$ for all v, then there is a coupling of the processes so that

$$X_v(t) \geq X^\diamond_v(t), \quad t \geq 0, v \in V. \tag{47}$$

To see this, observe that by the consistency property of spatial Λ-coalescent it suffices to prove the result assuming that $X_v(0) = X^\diamond_v(0)$ for all $v \in V$. In this case, (47) follows easily by induction on the number of particles: Just apply the consistency property of spatial Λ-coalescents, after the first time that two particles occupy the same site. (This idea is further exploited in Lemma 7.2.)

Finally, Theorem 6.4 follows by Lemma 6.5. □

6.3 Concentration of the number of particles

The main result of this section is a concentration result for the number of particles alive at a certain time. This provides a soft alternate route for the lower-bound on the long-time behaviour of the spatial coalescent, as we briefly explain.

Theorem 6.6. *Fix $t > 0$, and consider a spatial Kingman coalescent started from some arbitrary configuration containing a finite number of particles. Then we have*

$$\mathrm{Var}(N(t)) \leq \mathbb{E}N(t).$$

Proof. The tool used here again is a comparison to the coalescing random walk model, where particles coalesce immediately upon meeting. We denote by $(X^\diamond(t), t \geq 0)$ a system of instantaneously coalescing random walks started from a certain set of vertices A in a graph $G = (V, E)$, to be chosen suitably later, and let $N^\diamond(t)$ denote the total number of particles at time t. The proof is based on Arratia's correlation inequality [2, Lemma 1], which states that

$$\mathbb{E}X^\diamond_x(t)X^\diamond_y(t) \leq \mathbb{E}X^\diamond_x(t) \cdot \mathbb{E}X^\diamond_y(t). \tag{48}$$

Thus at any time, any two sites are negatively correlated. This inequality holds not just for the process on $\mathbb{Z}^d$, but on any edge weighted graph.

We now remark that the spatial Kingman coalescent on $\mathbb{Z}^d$ can be approximated by a system of instantly coalescing random walks on a larger graph. For any integer N such that $N > n$ (the initial number of particles), consider the graph $G_N = (V, E)$ with vertices $V = \mathbb{Z}^d \times \{1, \dots, N\}$. The edges of G_N are of two types. If $x \sim y$ in $\mathbb{Z}^d$ then there is an edge between (x, i) and (y, j) with weight ρ/N. Additionally, there is an edge with weight $1/2$ between (x, i) and (x, j) for any x, i, j. Call the set $x \times \{1, \dots, N\}$ a cluster. Clusters correspond to vertices of $\mathbb{Z}^d$ in a natural way. The $\mathbb{Z}^d$ coordinate of a continuous time random walk on G_N is a continuous time random walk on $\mathbb{Z}^d$ with jump rate ρ. However, two walks may be present in the same cluster and not meet. It is clear that as long as two random walks are in the same cluster they will meet at rate one (since each may jump into the vertex occupied by the other).

The probability of two random walks meeting when one jumps from one cluster to another is of order $1/N$. Thus as long as the number of particles is negligible compared to N, the projection onto $\mathbb{Z}^d$ of the coalescing random walks $X^\diamond_N$ on G_N is close to the spatial Kingman coalescent on $\mathbb{Z}^d$, and so as $N \to \infty$, the projection of $X^\diamond_N(t)$ converges to $X(t)$ (in the sense of vague convergence, identifying X_v and the projection of $X^\diamond_N$ to point measures on $\mathbb{Z}^d$). More precisely, for an initial configuration $X(0)$ of particles on $\mathbb{Z}^d$, we define a set $A \subset V_N$ by choosing for each v (arbitrarily) $X_v(0)$ particles from the cluster of v. Let $X^\diamond_N(t)$ be the process of coalescing random walks on G_N started with this configuration. Then if $M^\diamond_N(t)$ denote the total number of particles of $X^\diamond_N(t)$,

$$\begin{aligned}
\mathbb{E}(M^\diamond_N(t)^2) &= \sum_{x \in V_N} \mathbb{E}X^\diamond_{N,t}(x) + \sum_{x \neq y \in V_N} \mathbb{E}X^\diamond_{N,t}(x)X^\diamond_{N,t}(y) \\
&\leq \mathbb{E}M^\diamond_N(t) + \sum_{x \neq y \in V_N} \mathbb{E}X^\diamond_{N,t}(x)\mathbb{E}X^\diamond_{N,t}(y) \\
&\leq \mathbb{E}M^\diamond_N(t) + (\mathbb{E}M^\diamond_N(t))^2.
\end{aligned}$$

Thus for any N we have $\operatorname{Var} M^\diamond_N(t) \leq \mathbb{E}M^\diamond_N(t)$. By dominated convergence (since all processes have at most n particles) we see that

$$\lim_{N\to\infty} \mathbb{E}M^\diamond_N(t) = \mathbb{E}N(t) \qquad \lim_{N\to\infty} \mathbb{E}M^\diamond_N(t)^2 = \mathbb{E}N(t)^2,$$

and the result follows. □

As a simple corollary of this result, we obtain an alternate proof of Theorem 6.4. We have already seen in (31) that $\mathbb{E}(N(\infty)) \geq cm^{d-2}$ for some $c > 0$ if $d \geq 3$ (this argument is a simple Green function estimate, and is easy to adapt to the case $d = 2$). Applying Theorem 6.6 concludes the proof.

It would be also possible to derive a lower-bound on the expected number of particles in a system of instantaneously coalescing random walks at time tm^2, starting from a set A which dominates i.i.d. Bernoulli random variables with mean $p > 0$, using technology from coalescing random walks. We briefly outline the steps needed to do this. First, starting from a configuration where there is a particle at every site of $\mathbb{Z}^d$, and using a famous result of Bramson and Griffeath [12] on the asymptotic density of particles, we conclude that about cm^{d-2} such particles are in a region of volume Cm^d for some large $C > 0$ to be chosen suitably. If we treat the particles that started outside of A as ghosts, we are then led to estimate the number of ghost particles among those cm^{d-2}. For this, one can use the duality with the voter model (see [26]) and [28, Lemma 4], which gives good control on the probability that the voter model escapes a ball of radius $\sqrt{t}$, for large t.

7 Upper bound for the number of survivors

Assume that Λ is a finite measure on $[0, 1]$ such that for some $a_0 > 0$, we have

$$\lambda_n \geq a_0 n \quad \text{for all } n \geq 2, \tag{49}$$

where $\lambda_n = \sum_{k=2}^n \lambda_{b,k}$ is the total merger rate when there are n particles. Note that most coalescents which come down from infinity satisfy (49), in particular, if $\Lambda = \delta_{\{0\}}$ (the Kingman case) then (49) holds since $\lambda_n = \binom{n}{2}$, and if Λ has the regular variation property of (19), then (49) holds by Lemma 4.1.

Our goal here is to prove the following result.

Theorem 7.1. *Fix $C_0 \in (0, \infty)$ and $\delta > 0$, and consider the spatial Λ-coalescent in $\mathbb{Z}^d$ satisfying (49), started from a configuration of at most $C_0 m^d$ particles located in $B(o, m)$, and no particles in $\mathbb{Z}^d \setminus B(o, m)$. There exists $C = C(\delta, C_0)$, such that if $d > 2$ then*

$$\mathbb{P}\left(N^*(\delta m^2) < Cm^{d-2}\right) \xrightarrow[m\to\infty]{} 1,$$

while, if $d = 2$,

$$\mathbb{P}\left(N^*(\delta m^2) < C \ln m\right) \xrightarrow[m\to\infty]{} 1.$$

Note that when $d > 2$ this order of magnitude bound is sharp, since Theorem 6.4 shows $N^*(\infty) \geq cm^{d-2}$. For $d = 2$, due to recurrence, $N^*(\infty) = 1$, almost surely.

The idea behind the proof is a comparison of the spatial system to a mean field approximation. The actual argument is based on a somewhat technical construction so we start with a non-technical overview. Recall the comparison with ODE described in (4): if at time t the density of particles averaged over some ball is $\rho(t)$ (typically small), then we approximate the spatial coalescent with the mean-field model where the coalescence rate per particle is $\rho(t)$ at time t, leading to the differential equation

$$\frac{d}{dt}\rho(t) = -\rho^2(t)/2, \quad t \geq s.$$

Hence $\rho(t)^{-1} = c + (t-s)/2$ and therefore $\rho(t) = \frac{2}{t-s+2\rho(s)^{-1}}$, $t \geq s$. Provided that all the particles in the spatial coalescent configuration are located in the ball of radius m during the whole interval $[s,t]$ (and that the above approximation is valid) then their total number is approximately $Cm^d\rho(t)$. In turn, this approximation remains valid as long as the particles remain inside a ball centered at the origin with radius of order m, i.e. up to time of order m^2. At times of order m^2, the number of remaining particles is of order m^{d-2}.

A key difficulty of the approach outlined above comes from the fact that some particles diffuse away from the densest regions relatively early in the evolution, which might enable them to survive longer. To account for such "runaways", we adopt a *multi-scale approach*, bounding at each stage the number of particles that "escape". This is done in Lemma 7.8. Lemma 7.5 provides the estimates on the number of non-escaping particles at each stage.

To justify the comparison of the spatial process with the mean field process we average over small time intervals (cf. Lemma 7.4 below). This is necessary since at any given time it is possible that no vertex contains more than a single particle, in which case the immediate rate of coalescence is 0. However, the system is unlikely to stay in such states long enough to hinder the approximation. Indeed, Lemma 7.4 implies that the average rate of coalescence is (up to constants) as predicted by the mean field heuristic. The multiplicative constants are inherent to the spatial structure, and it seems difficult to compute them.

7.1 Preparatory lemmas

Our first step is a comparison lemma between the spatial Λ-coalescent X and a slower spatial coalescent. We then consider a possibly more general spatial coalescent process $\{(\bar{X}_v(t), t \geq 0)\}_{v\in V}$. If the process consists initially of n particles labeled by $[n] = \{1, \dots, n\}$, a configuration consists as usual of labeled partitions of $[n]$, where the label of a block corresponds to its location on V. Equivalently, a configuration $\bar{x} = (\bar{x}_v)_{v\in V}$ may be thought of as giving the list of blocks (referred to as particles) present at each particular site $v \in V$. We will also sometimes abuse notation and denote by $X_v(t)$ the number of particles (i.e., blocks) present at time t and at position v. We assume that particles perform independent continuous-time simple random walks with jump rate ρ, and that there exists a family of real numbers $\bar{\lambda}_{\bar{x},S}$ such that for all configuration $\bar{x} = (\bar{x}_v)_{v\in V}$, all $v \in V$, any particular subset S of all blocks present at $v \in V$ coalesces at an instantaneous rate $\bar{\lambda}_{\bar{x},S}$, if the current configuration is $\bar{x}$. Moreover, coalescence events at different sites occur independently of one another, and are independent of the migration. We now make the following assumption on the family of rates $\bar{\lambda}_{\bar{x},S}$: if $v \in V$ and $\bar{x}_v$ contains $n \geq 2$ particles, then for every $2 \leq k \leq n$, we have:

$$\sum_{S:|S|\geq k} \bar{\lambda}_{\bar{x},S} \leq \sum_{\ell \geq k} \binom{n}{\ell} \lambda_{n,\ell}, \tag{50}$$

where $\lambda_{n,k}$ is the coalescence rate of any particular subset of size k in a Λ-coalescent. The idea behind (50) is that if X and $\bar{X}$ have the same number of particles at time t, then $X(t+\varepsilon)$ is stochastically dominated by $\bar{X}(t+\varepsilon)$.

Lemma 7.2. *Consider a Λ-coalescent X and a coalescent process $\bar{X}$ such that (50) holds, and $X_v(0) \leq \bar{X}_v(0)$ for all v. Then there is a coupling of the processes X and $\bar{X}$ such that $X_v(t) \leq \bar{X}_v(t)$ holds for all $v \in V$ and $t \geq 0$.*

Proof. By the consistency of spatial Λ-coalescents, it suffices to prove the result when $X_v(0) = \bar{X}_v(0)$ for all $v \in V$. We associate each particle of X with a particle of $\bar{X}$ and let them perform the same random walks as long as there are no coalescence events. A consequence of (50) is that it is possible to couple the processes so that if $X_v = \bar{X}_v$ then the coalescence events of X dominate those of $\bar{X}$, that is, any coalescence event in $\bar{X}$ occurs at the same time as an event in X involving at least as many particles.

The proof now proceeds by induction on the total number of particles, which are allowed to be distributed arbitrarily. By the above remark, we may couple the processes X and $\bar{X}$ so that the domination holds up to and including the first time t_0 of a coalescence event, which involves particles from X and possibly from $\bar{X}$. Assume that $\bar{X}$ also experiences a coagulation event at this time. (Else, we can artificially retain particles in X that were supposed to coagulate at time t_0. By the consistency property, this may only increase the process X stochastically.)

We now use the induction hypotheses to construct processes $(X'(t), t \geq t_0)$ and $(\bar{X}'(t), t \geq t_0)$ with initial configuration $X'(t_0) = \bar{X}'(t_0) = \bar{X}(t_0)$ such that $X'_u(t) \leq \bar{X}'_u(t)$ for all $t \geq t_0$. We can define $\bar{X}(t) = \bar{X}'(t)$ for $t > t_0$, and by consistency of the spatial Λ-coalescents, we extend the coupling to X for $t > t_0$ so that $X_u(t) \leq X'_u(t)$ for all $u \in V$, which proves the claim. $\square$

Remark. This lemma holds for more general spatial coalescents: e.g., the instantaneous coalescence rates $\lambda_{\bar{x},S}$ could be allowed to be arbitrary path-dependent (i.e., $\mathcal{F}_t$-measurable at time t), almost surely nonnegative and finite random variables. The only crucial assumption is that (50) holds uniformly.

We now apply Lemma 7.2 to the situation which is particularly useful in our setting. Recall that we are considering a spatial Λ-coalescent for which (49) holds. Assume that initially there are N particles, and let $\{X_v(t), t \geq 0\}_{v\in V}$ denote the number of particles of this process as a function of time and space.

Let π be a partition of $\{1, \dots, N\}$. We refer to the blocks of π as *classes*. Let $\{\bar{X}_v(t), t \geq 0\}_{v\in V}$ denote a process where classes evolve independently of one another, and particles within each class evolve according to a spatial $(\bar{\Lambda})$-coalescent, where $\bar{\Lambda}$ will be specified soon. That is, particles move as continuous-time simple random walks with rate ρ and coalesce when they are on the same site and from the same class according to a $\bar{\Lambda}$-coalescent.

Lemma 7.3. *Assume that the blocks of π are all of size 1 or 2, and that $\bar{\Lambda} = (a_0/\lambda_2)\Lambda$, where a_0 is the constant of (49) and $\lambda_2 = \lambda_{2,2}$ is the pairwise coalescence rate. Assume also that $X_v(0) \leq \bar{X}_v(0)$ for all v. Then there is a coupling of the processes X and $\bar{X}$ such that $X_v(t) \leq \bar{X}_v(t)$ holds for all $v \in V$ and $t \geq 0$.*

Proof. Observe first that our process $\bar{X}$ is of the type described above Lemma 7.2, so that it suffices to establish (50). Note however that if a configuration $\bar{x}$ contains n particles at site v, and S is a subset of particles with $|S| = k$ and $2 \leq k \leq n$, we have $\bar{\lambda}_{\bar{x},S} = 0$ for $k \geq 3$, while if $k = 2$, $\lambda_{\bar{x},S} = 0$ when the particles of S are not of the same class, and if they are of the same class, $\lambda_{\bar{x},S} = (a_0/\lambda_2)\lambda_2 = a_0$. Since there are at most n subsets of particles that are allowed to coalesce, we have

$$\sum_{S:|S|\geq 2} \bar{\lambda}_{\bar{x},S} \leq na_0 \leq \lambda_n = \sum_{k=2}^{n} \binom{n}{k} \lambda_{n,k},$$

which proves (50), and completes the proof. □

Lemma 7.4. *Fix c_0, C_0, and consider a spatial Λ-coalescent satisfying (49) with $N^*(0)$ particles all inside $B(o,R)$. Let $\rho(t) = \frac{N^*(t)}{R^d}$ be the inverse density, and assume $\rho(0) \in (c_0 R^{-2}, C_0)$. Denote $\tau = \rho(0)^{-2/d}$. Then for $d > 2$ we have*

$$\mathbb{P}\Big(\rho(\tau)^{-1} < \rho(0)^{-1} + c_1\tau\Big) < \exp\left(-cR^{(d-2)^2/d}\right),$$

where c_1 depends only on d, c_0, C_0, a_0.

If $d = 2$ we have

$$\mathbb{P}\Big(\rho(\tau)^{-1} < \rho(0)^{-1} + \frac{c_1\tau}{\log\tau}\Big) < \exp\left(-\frac{cR^2}{\tau\log\tau}\right).$$

Proof. We first argue that for some $C = C(d)$, it is possible to find at least $N^*(0)/4$ disjoint pairs in the set of initial particles, so that for each pair the initial distance between its particles is at most $C\rho(0)^{-1/d}$ (for large $\rho(0)$, the particles forming such a pair are initially located at the same site). To achieve this, cover $B(o,R)$ with $N^*(0)/2$ (disjoint) boxes of diameter $CRN^*(0)^{-1/d} = C'\rho^{-1/d}$ (this is possible for some C). Within each box match as many pairs as possible in an arbitrary manner. This leaves at most one unpaired particle in each ball, so at least $N^*(0)/2$ are matched, with all distances bounded as claimed. Refer to two particles forming a pair as "partners".

Consider the coupling from Lemma 7.3, where π corresponds to the partitioned formed by identifying particles with their partners (which therefore contains only singletons or doubletons). Let Z' be the total number of coalescence events in the process Π' where coalescence events involving members of different classes are not allowed and occur at rate a_0 when they are. Lemma 7.3 implies that $Z' \preceq Z$, in the sense of stochastic domination. Hence, it suffices to prove the claimed bounds for Z'. The advantage of considering Π' instead of Π is that different pairs of partners evolve independently.

From this point on, the arguments for the cases $d = 2$ and $d > 2$ differ. In dimensions $d > 2$, by our assumptions, $\tau > c_0$ for some c_0. The probability that random walkers started at distance at most $\rho^{-1/d}$ meet before time $\tau/2 = \rho^{-2/d}/2$ is at least $c\rho^{(d-2)/d}$. On this event, there is probability bounded from 0 that they coalesce before time τ. Thus the number of partners that coalesce by time τ dominates a $\mathrm{Bin}(N^*(0)/4, c\rho^{(d-2)/d})$ random variable. This random variable has expectation $cN^*(0)\rho^{(d-2)/d} \geq cR^{(d-2)^2/d}$. The bound in the lemma is the probability that this random variable is less than half its expectation.

Finally, if the number of coalesce events is at least $cN^*(0)\rho(0)^{(d-2)/d} = cN^*(0)\rho(0)\tau$ then

$$\rho(t)^{-1} = \frac{R^d}{N^*(t)} \geq \frac{R^d}{N^*(0)(1 - c\rho(0)\tau)} \geq \frac{R^d}{N^*(0)}(1 + c\rho(0)\tau) = \rho(0)^{-1} + c\tau.$$

In the case $d = 2$, each pair coalesces with probability at least $c/\log \tau$ (along the same lines) by Lemma 6.2. As above, the number of coalesce events is at least $cR^2/\tau \log \tau$ except with probability $e^{-cR^2/\tau \log \tau}$. On this event, a similar computation gives

$$\rho_t^{-1} \geq \rho_0^{-1} + \frac{c}{\tau \log \tau}. \qquad \square$$

Lemma 7.5. *Let S_t denote the number of particles (in the spatial coalescent) that remain in $B(o, R)$ during the whole interval $[0, t]$. In particular, $S_0 = \sum_{v \in B(o,R)} X_v(0)$. Fix $C_0 > 0$, and assume that $S_0 = n < C_0 R^d$ and that $2 < t < R^2$. Then for some C depending only on d,*

$$\mathbb{P}\left(\frac{S_t}{R^d} > \frac{C}{t}\right) < cR^4 \exp\big(- cR^{(d-2)^2/d}\big) \qquad \text{if } d > 2,$$

and

$$\mathbb{P}\left(\frac{S_t}{R^2} > \frac{C \log t}{t}\right) < (\log t)^{-1} \qquad \text{if } d = 2.$$

Proof. As usual, the case $d > 2$ is considered first. The previous lemma can be formulated as follows: The process $\{\rho_t^{-1}\}$ is unlikely to spend more than $u^{-2/d}$ units of time in the interval $[u, u + c_1 u^{-2/d}]$. Note that S_t can only decrease faster than $N^*(t)$, so this will also hold for the modified density $\rho(t) = \frac{S_t}{R^d}$.

We apply this to the following sequence of intervals. Let $u_0 = \rho(0)$ and $u_{k+1} = u_k + c_1 u_k^{-2/d}$. Let K be minimal with $u_K > t/c_1$. As long as $u_j < t/c_1$ the increment is at least $ct^{-2/d}$. It follows that $K < Ct^{1+2/d} < R^4$. If the process does not spend more than $u_k^{-2/d}$ time in $[u_k, u_{k+1}]$ then the time before ρ^{-1} exceeds t/c_1 is at most t. The probability that this fails to hold is at most $R^2 \exp\big(- cR^{(d-2)^2/d}\big)$.

This works provided $K > 1$, or equivalently $t \geq \rho(0)^{-2/d}$. If $t < \rho(0)^{-2/d}$ then we have

$$S_t \leq S_0 = \rho(0) R^d \leq \frac{R^d}{t^{d/2}} < \frac{R^d}{t}.$$

In the case $d = 2$, we instead have $u_{k+1} = u_k + \frac{c_1 u_k}{\log u_k}$. It is not hard to see that

$$u_k \asymp e^{\sqrt{2c_1 k + c}} \tag{51}$$

for some c depending on u_0. To this end, note that u is increasing and hence is dominated by the solution of the ODE $f' = c_1 f/\log f$ (at least once u is large enough that $u/\log u$ is increasing). This ODE is solved by $f = e^{\sqrt{2c_1 x + x}}$, giving the upper bound on u. For the other direction, note that once u_k is large u_{k+1}/u_k is close to 1. This implies that u dominates a solution of $f' = (c_1 - \varepsilon) f/\log f$.

Lemma 7.4 tells us that ρ_t^{-1} is unlikely to spend more than u_k units of time in $[u_k, u_{k+1}]$, and the probability of this unlikely event is at most

$$p_k = e^{-cR^2/u_k \log u_k}.$$

Let K be such that $u_K > \alpha t/\log t$, with α small to be determined soon. Note that now the failure probability for the last intervals is of order 1, so a union bound does not work.

However, Lemma 7.4 tells us more. If the process ρ_t^{-1} fails to exceed u_{k+1} in the next u_k units of time, then by the Markov property and Lemma 7.4 again, it gets a fresh chance to do so in the next u_k units of time. Therefore, the number of attempts is smaller than a geometric random variable with success probability $1-p_k$. It follows that the total time spent in $[u_0, u_K]$ is stochastically dominated by $Q := \sum_{k=1}^{K-1} u_k \operatorname{Geom}(1-p_k)$, where the geometric random variables are independent. The probability we wish to bound is therefore at most $\mathbb{P}(Q > t)$. By making α small we can guarantee $p_k < 1/2$ for all k, so that the geometric variables are typically small.

More precisely, from (51) it follows that $K \asymp \log^2 t$ and therefore that

$$\sum_{k<K} u_k \asymp \int_0^K e^{\sqrt{2c_1 t}} dt \asymp \sqrt{K} u_K \asymp \alpha t,$$

hence for small enough α we have $\mathbb{E}Q \leq 2\sum_{k=1}^{K-1} u_k \leq t/2$. Similarly, we can compute

$$\operatorname{Var} Q \leq C \sum_{k<K} u_k^2 < \frac{C\alpha^2 t^2}{\log t}.$$

The lemma now follows from Chebyshev's inequality and choosing α small enough that $4\alpha^2 C \leq 1$. ☐

The following is a fairly standard fact which follows easily from the optional stopping theorem and Doob's inequality:

Lemma 7.6. *If X_t is a continuous time random walk on $\mathbb{Z}^d$ then for all $x \in \mathbb{Z}^d$,*

$$\mathbb{P}\left(\sup_{s\leq t} \|X_s\|_2 \geq x\right) \leq C e^{-cx^2/t},$$

where c, C depend only on d.

Lemma 7.7. *If $W = \operatorname{Bin}(n,p)$, and $\Delta > 0$, then*

$$\mathbb{P}(W > 2np + \Delta) < \Delta^{-1/2}. \tag{52}$$

Proof. If $\Delta > (np)^2$, then Markov's inequality gives $\mathbb{P}(W > \Delta) \leq np/\Delta < \Delta^{-1/2}$. If $\Delta \geq (np)^2$, one can use Chebyshev's inequality to obtain $\mathbb{P}(W > 2np + \Delta) \leq \mathbb{P}(W - \mathbb{E}W > np) \leq (1-p)/(np) < \Delta^{-1/2}$. ☐

7.2 Completing the proof

Lemma 7.5 is almost sufficient to deduce Theorem 7.1. The missing piece is to account for the particles that "escape" from the ball under observation. We accomplish this by partitioning the time interval $[0, m^2]$ into several segments and applying Lemma 7.5 to each segment. More precisely, let $K = K(m) = \lfloor \log\log m \rfloor$, and consider the process at a particular sequence of times given by

$$t_k = \begin{cases} 0 & k = 0, \\ e^{k-K} m^2 & k = 1, \dots, K. \end{cases}$$

Thus $t_1 \asymp m^2/\ln m$, and the sequence increases geometrically up to $t_K = m^2$. At each time t_k, we will consider the behavior of the process with respect to the ball $B(o, R_k)$, where the radii R_k are defined by:

$$R_k = \begin{cases} 0 & k = 0, \\ \gamma(m + \sqrt{t_k(K+1-k)}) & k > 0, \end{cases}$$

where $\gamma > 1$ is some constant to be determined during the proof of Lemma 7.8. Note that R_k is increasing in k, and that $R_K = 2\gamma m$.

With the above notations in mind, let X_k (resp. Y_k) be the number of particles inside (resp. outside) $B(o, R_k)$ at time t_k. Let $A_m = A_m(\alpha, \beta, \gamma)$ be the event

$$A_m = \left\{ X_k < \frac{\alpha R_k^d}{t_k} \quad \text{and} \quad Y_k < \frac{\beta R_k^d}{t_k} \quad \text{for all } k \le K \right\} \qquad \text{when } d > 2.$$

and

$$A_m = \left\{ X_k < \frac{\alpha R_k^2 \log m}{t_k} \quad \text{and} \quad Y_k < \frac{\beta R_k^d \log m}{t_k} \quad \text{for all } k \le K \right\} \qquad \text{when } d = 2.$$

Lemma 7.8. *Assume the initial conditions of Theorem 7.1. Then for some choice of α, β, γ we have $\mathbb{P}(A_m) \xrightarrow[m \to \infty]{} 1$.*

Proof. The idea is to inductively bound X_{k+1}, Y_{k+1} in terms of X_k, Y_k. The bound on X_{k+1} is mostly an application of Lemma 7.5. However, to bound X_{k+1} we must also account for particles that are outside the ball $B(o, R_k)$ at time t_k, or particles that exit the ball $B(o, R_{k+1})$ at some time before time t_{k+1} and re-enter it. These quantities can be bounded in terms of Y_k and X_k as well as auxiliary quantities introduced soon. A delicate point in the proof is that the number of steps of the induction is not fixed ($\ln \ln m$), so we make sure that constants do not grow with m. Thus all constants below depend only on d.

At time t_0 our assumptions are that $Y_0 = 0$ and $X_0 \le C_0 m^d$. For the induction step we define two additional quantities: S_k and Z_k. Let S_k be the number of particles that remain in $B(o, R_k)$ throughout the time interval $[t_{k-1}, t_k]$. We wish to apply Lemma 7.5 to S_k. The conditions are clearly satisfied (recall $S_k \le C_0 m^d < C_0 R_k^d$, since $\gamma > 1$). Since $t_k - t_{k-1} \ge c t_k$ this will imply that with high probability

$$S_k < C_1 \frac{R_k^d}{t_k} \qquad \text{for all } k \le K. \tag{53}$$

(The probability of failure at each of $\log \log m$ steps is exponentially small.)

Let Z_k be the number of particles located inside $B(o, R_k)$ at time t_k that exit $B(o, R_{k+1})$ before time t_{k+1}. Lemma 7.6, bounds the escape probability for each of X_k particles inside $B(o, R_k)$. Coalescence can only reduce the number of escaping particles, so given X_k,

$$Z_k \preceq \text{Bin}\left(X_k, C_2 \exp\left(-c_2 \frac{(R_{k+1} - R_k)^2}{t_{k+1} - t_k} \right) \right).$$

Here c_2, C_2 depend only on d, and "$\preceq$" denotes stochastic domination.

If $k = 0$ this implies

$$\mathbb{E}Z_0 \le C_0 m^d C_2 e^{-c_2\gamma^2 m^2/t_1}.$$

Since $m^2/t_1 \asymp \ln m$, by making γ large enough we obtain $\mathbb{E}Z_0 = o(1)$. Then in particular $Z_0 = 0$ with probability tending to 1. For $k \ge 1$ we use Lemma 7.7, with $\Delta = K^4$ to find

$$\mathbb{P}\left(\bigcup_{k=1}^{K} \{Z_k \ge K^4 + 2\mathbb{E}Z_k\}\right) \le K/\sqrt{K^4} = 1/K.$$

Thus with probability at least $1 - K^{-1} \to 1$ (as $m \to \infty$), we have

$$Z_k < K^4 + 2C_2 X_k \exp\left(-c_2 \frac{(R_{k+1} - R_k)^2}{t_{k+1} - t_k}\right) \qquad \text{for all } k.$$

Using $t_{k+1} = et_k$, $k \ge 1$, together with $\sqrt{t_{k+1}(K-k)} \ge \sqrt{e/2}\sqrt{t_k(K+1-k)}$, we conclude that with high probability, $Z_0 = 0$ and

$$Z_k < K^4 + 2C_2 X_k e^{-c_3\gamma^2(K-k)} \qquad \text{for all } k \le K. \tag{54}$$

With these preparations in place, we are ready for the induction. Assume from here on that $Z_0 = 0$ and (53), (54) hold. We have, for each k, the deterministic bounds

$$\begin{aligned} X_k &\le S_k + Y_{k-1} + Z_{k-1}, \\ Y_k &\le Y_{k-1} + Z_{k-1}. \end{aligned}$$

To see this, note that particles in $B(o, R_k)$ either stayed inside (S), started outside (Y), or exited and returned (Z). The bound on Y is similar.

We now carry out an induction over k to bound X_k, Y_k for all $k = 1, \dots, K$. Suppose that the bound (from the event A_m) on X_j and Y_j hold for all $j < k$. It follows from (54) and the inductive hypothesis that

$$\begin{aligned} Y_k \le \sum_{j<k} Z_j &< kK^4 + 2C_2\alpha \sum_{1\le j<k} \frac{R_j^d}{t_j} e^{-c_3\gamma^2(K-j)} \\ &< kK^4 + 2C_2\alpha \frac{R_k^d}{t_k} \sum_{j<k} \frac{t_k}{t_j} e^{-c_3\gamma^2(K-j)} \\ &= kK^4 + 2C_2\alpha \frac{R_k^d}{t_k} e^{-c_3\gamma^2(K-k)} \sum_{j<k} e^{(1-c_3\gamma^2)(k-j)}. \end{aligned}$$

We require γ to be large enough that $c_3\gamma^2 > 2$. Then the last sum is at most $1/(e-1) < 1$ and so

$$Y_k < kK^4 + 2C_2\alpha \frac{R_k^d}{t_k} e^{-c_3\gamma^2(K-k)}. \tag{55}$$

This proves the induction step for Y_k with any choice of $\beta > 2C_2\alpha$ (since $kK^4 \le K^5 \ll R_k^d/t_k$), on the event from (54), for all sufficiently large m.

It remains to bound X_k, for which we will use the bounds on S_k, Y_{k-1} and Z_{k-1}. We already have

$$Y_{k-1} < (k-1)K^4 + 2C_2\alpha\frac{R_{k-1}^d}{t_{k-1}}e^{-c_3\gamma^2(K+1-k)} < (k-1)K^4 + 2eC_2\alpha\frac{R_k^d}{t_k}e^{-c_3\gamma^2}.$$

Using the induction hypothesis and (54) (with $k-1$ replacing k) one finds

$$Z_{k-1} < K^4 + 2C_2\alpha\frac{R_{k-1}^d}{t_{k-1}}e^{-c_3\gamma^2(K+1-k)} < K^4 + 2eC_2\alpha\frac{R_k^d}{t_k}e^{-c_3\gamma^2}.$$

Thus we have

$$\begin{aligned} X_k \le S_k + Y_{k-1} + Z_{k-1} &< C_1\frac{R_k^d}{t_k} + kK^4 + 4eC_2\alpha\frac{R_k^d}{t_k}e^{-c_3\gamma^2} \\ &= kK^4 + \left(C_1 + 4eC_2\alpha e^{-c_3\gamma^2}\right)\frac{R_k^d}{t_k}. \end{aligned} \tag{56}$$

To finish the proof it remains to select α and γ (and $\beta > 2C_2\alpha$) so that

$$C_1 + 4eC_2\alpha e^{-c_3\gamma^2} < \alpha.$$

This is done by requiring γ to satisfy $e^{c_3\gamma^2} > 4eC_2$ and taking any sufficiently large α.

Turning to the case $d = 2$, we proceed along the same lines. Lemma 7.5 gives with probability with high probability

$$S_k < C_1\frac{R_k^2\log m}{t_k} \qquad \text{for all } k \le K. \tag{57}$$

The failure probability is $\log^{-1} m$ at each of $\log\log m$ steps. Note that $\log t_i \sim \log m$ for all i, so we are not giving much away here. Furthermore, w.h.p. $Z_0 = 0$ and (54) holds (the proof of these facts does not depend on d.

We now repeat the induction. Given (54) and the induction hypothesis bounds on X we get (as above, with an extra $\log m$ factor)

$$Y_k < kK^4 + 2C_2\alpha\frac{R_k^2\log m}{t_k}e^{-c_3\gamma^2(K-k)}.$$

Since the bounds for S_k, Y_{k-1} and Z_{k-1} differ from the general case only by a $\log m$ on the R_k^2/t_k term, the bound for X_k gets the same factor as well. □

Acknowledgements. We wish to thank Alan Hammond for useful discussions and careful reading of a few preliminary drafts. This project was started when all authors were at the University of British Columbia. Part of it was also carried as the first author was at the University of Toronto. Further progress was made during visits to Marseilles through the support of V.L.'s Alfred P. Sloan fellowship, and to the University of Toronto. The hospitality and support of these departments is warmly acknowledged.

References

[1] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[2] R. Arratia. Limiting point processes for rescalings of coalescing and annihilating random walks on $\mathbf{Z}^d$. *Ann. Probab.*, 9(6):909–936, 1981.

[3] J. Berestycki, N. Berestycki, and V. Limic. The Λ-coalescent speed of coming down from infinity. *Ann. Probab.*, to appear.

[4] J. Berestycki, N. Berestycki, and V. Limic. Interpreting Λ-coalescent speed of coming down from infinity via particle representation of super-processes. In preparation.

[5] J. Berestycki, N. Berestycki, and J. Schweinsberg. Beta-coalescents and continuous stable random trees. *Ann. Probab.*, 35(5):1835–1887, 2007.

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

[7] N. Berestycki. Recent progress in coalescent theory. To appear in Ensaios Matematicos.

[8] J. v. d. Berg and H. Kesten. Asymptotic density in a coalescing random walk model. *Ann. Probab.*, 28(1):353–352, 2000.

[9] J. v. d. Berg and H. Kesten. Randomly coalescing random walks in dimension ≥ 3. In *In and out of equilibrium*, volume 51 of *Progr. Probab.*, pages 1–45. Birkhauser, Boston, 2002.

[10] J. Bertoin and J.-F. Le Gall. Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math.*, 50(1-4):147–181 (electronic), 2006.

[11] M. Birkner, J. Blath, M. Capaldo, A. Etheridge, M. Möhle, J. Schweinsberg, and A. Wakolbinger. Alpha-stable branching and beta-coalescents. *Electron. J. Probab.*, 10:no. 9, 303–325 (electronic), 2005.

[12] M. Bramson and D. Griffeath. Asymptotics for interacting particle systems on $\mathbf{Z}^d$. *Z. Wahrsch. Verw. Gebiete*, 53(2):183–196, 1980.

[13] J. T. Cox and D. Griffeath. Diffusive clustering in the two dimensional voter model. *Ann. Probab.*, 14:347–370, 1986.

[14] R. Durrett. *Probability models for DNA sequence evolution.* Springer, second edition, 2008.

[15] P. Erdös and S. J. Taylor. Some problems concerning the structure of random walk paths. *Acta Math. Acad. Sci. Hungar.*, 11:137–162, 1960.

[16] A. Greven, V. Limic, and A. Winter. Coalescent processes arising in the study of diffusive clustering. In preparation.

[17] A. Hammond and F. Rezakhanlou. Kinetic limit for a system of coagulating planar Brownian particles. *J. Stat. Phys.*, 124(2-4):997–1040, 2006.

[18] A. Hammond and F. Rezakhanlou. The kinetic limit of a system of coagulating Brownian particles. *Arch. Ration. Mech. Anal.*, 185(1):1–67, 2007.

[19] H. Kesten. The number of alleles in electrophoretic experiments. *Theoret. Pop. Biol.*, 18(2):290–294, 1980.

[20] H. Kesten. The number of distinguishable alleles according to the ohta-kimura model of neutral mutation. *J. Math. Biol.*, 10:167–187, 1980.

[21] M. Kimura. "Stepping stone" model of population genetics. *Ann. Rept. Nat. Inst. Genetics, Japan*, 3:62–63, 1953.

[22] M. Kimura and G. Weiss. The stepping stone model of population structure and the decrease of genetic correlations with distance. *Genetics*, 49:561–576, 1964.

[23] J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982.

[24] J. F. C. Kingman. On the genealogy of large populations. *J. Appl. Probab.*, Special Vol. 19A:27–43, 1982. Essays in statistical science.

[25] G. F. Lawler and V. Limic. Symmetric random walk. Book in preparation, http://www.math.uchicago.edu/ lawler/books.html.

[26] T. M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.

[27] V. Limic and A. Sturm. The spatial Λ-coalescent. *Electron. J. Probab.*, 11:no. 15, 363–393 (electronic), 2006.

[28] M. Merle. Hitting probability of a distant point for the voter model started with a single one. *Ann. Probab.*, 36(3):807–861, 2008.

[29] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.

[30] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

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

[32] J. Schweinsberg. Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106(1):107–139, 2003.

[33] F. Spitzer. *Principles of Random Walk*, volume 34 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 1976.

Omer Angel: angel@math.ubc.ca
Department of Mathematics, University of British Columbia,
Vancouver, BC, V6T 1Z2, Canada

Nathanael Berestycki: n.berestycki@statslab.cam.ac.uk
Statistical Laboratory, DPMMS
University of Cambridge
Wilberforce Rd. Cambridge CB3 0WB, United Kingdom

Vlada Limic: vlada@cmi.univ-mrs.fr
CNRS - Université de Provence,
Technopôle de Château-Gombert
UMR 6632, LATP, CMI
39, rue F. Joliot Curie, 13453 Marseille, cedex 13, FRANCE