The Annals of Probability
2018, Vol. 46, No. 4, 2000–2037
https://doi.org/10.1214/17-AOP1218

RANDOM PARTITIONS OF THE PLANE VIA POISSONIAN COLORING AND A SELF-SIMILAR PROCESS OF COALESCING PLANAR PARTITIONS

BY DAVID ALDOUS[1]

University of California, Berkeley

Plant differently colored points in the plane; then let random points ("Poisson rain") fall, and give each new point the color of the nearest existing point. Previous investigation and simulations strongly suggest that the colored regions converge (in some sense) to a random partition of the plane. We prove a weak version of this, showing that normalized empirical measures converge to Lebesgue measures on a random partition into measurable sets. Topological properties remain an open problem. In the course of the proof, which heavily exploits time-reversals, we encounter a novel self-similar process of coalescing planar partitions. In this process, sets $A(z)$ in the partition are associated with Poisson random points z, and the dynamics are as follows. Points are deleted randomly at rate 1; when z is deleted, its set $A(z)$ is adjoined to the set $A(z')$ of the nearest other point z'.

1. Introduction. The work in this paper has several motivations. We focus below on the most concrete motivation; more broadly, as indicated in Sections 1.2 and 5.2, we will encounter a kind of spatial analog of well-studied nonspatial models of stochastic fragmentation (in forward time) or stochastic coalescence (in reversed time). A minor variant of the process below has been considered independently by several researchers (see Section 1.2), but without any published results.

As the "elementary" variant,[2] choose $k \geq 2$ distinct points $z_1, \dots, z_k$ in the unit square, and assign to point z_i the color i from a palette of k colors. Take i.i.d. uniform random points $U_{k+1}, U_{k+2}, \dots$ in the unit square, and inductively, for $j \geq k+1$,

give point U_j the color of the closest point to U_j among $U_1, \dots, U_{j-1}$

where we interpret $U_i = z_i$, $1 \leq i \leq k$ (there is a unique closest point a.s.; throughout the paper we omit the "a.s." qualifier where no subtlety is involved). This defines a process $S_n = (S_n(i), 1 \leq i \leq k)$, where $S_n(i)$ is the set of color-i points among $(U_j, 1 \leq j \leq n)$. Simulations (see Figure 1) and intuition strongly suggest that there is (in some sense) an $n \to \infty$ limit, which is a random partition

Received January 2017; revised July 2017.
[1]Supported by NSF Grant DMS-15-04802.
MSC2010 subject classifications. Primary 60D05; secondary 60G57.
Key words and phrases. Random tessellation, Poisson point process, spatial tree, stochastic coalescence.

[2]We mean that the model *definition* is elementary.

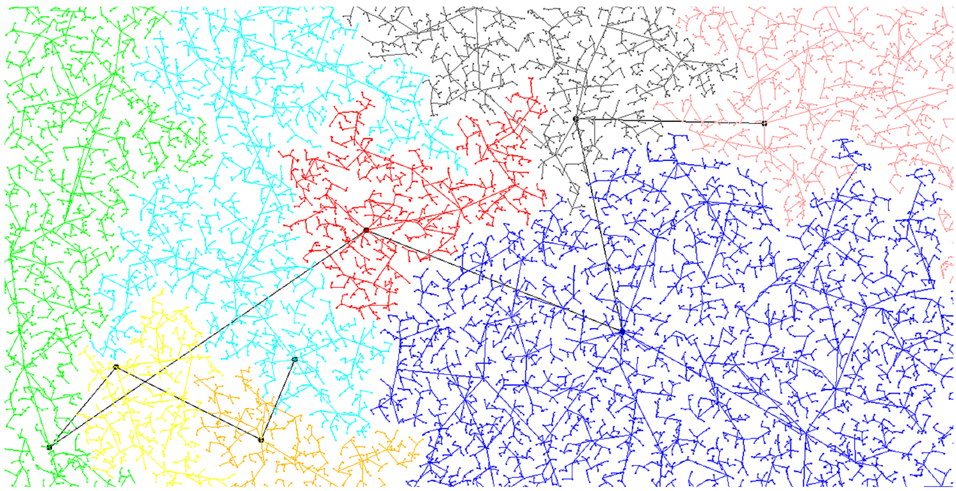

FIG. 1. *A realization within part of the unit square. Line segments indicate parent-child relation.*

of the square into k colored regions. Simulations (see Figure 2) also suggest that the boundaries between these (presumed) limit regions should be fractal, in some sense, though intuition is less clear here (see Section 5.3).

What can we actually prove? For rigorous study, it is more convenient to consider a slightly more sophisticated model. On the infinite plane $\mathbb{R}^2$ and the infinite time interval $-\infty < t < \infty$, there is a space–time Poisson point process (PPP), which we will envisage as the times and positions of arriving particles, such that the set of particles which arrive before time t forms a spatial PPP on $\mathbb{R}^2$ with in-

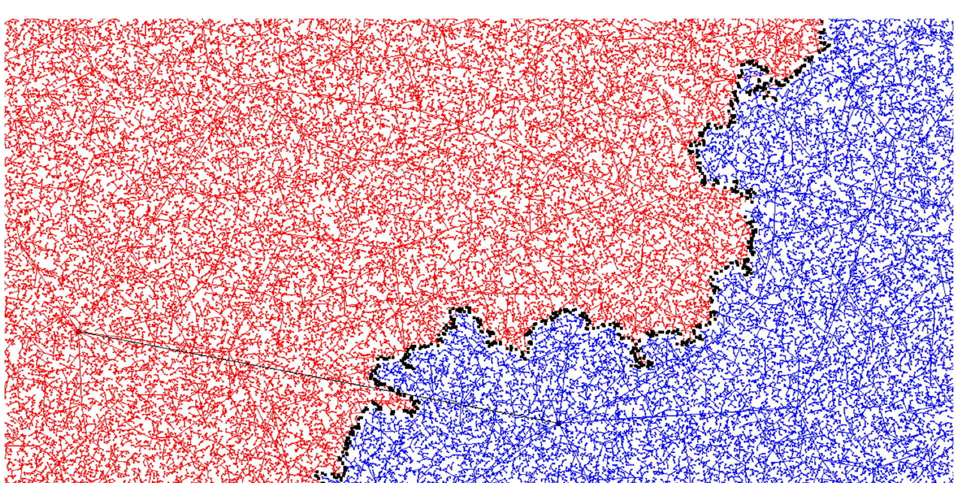

FIG. 2. *A close-up of the boundary between partition components suggests the boundary is fractal.*

tensity e^t per unit area. Within this process (more details and notation for what follows in this section will be given next in Section 1.1), imagine assigning a different color to each particle present at time t_1, and then as t increases suppose we color each newly arriving particle by the previous rule, that is, by copying the color of the nearest existing particle. Intuitively, what we see in the unit square within this model, at large times t, must be similar (up to boundary effects) as in the elementary model with a Poisson(e^{t_1}) number of initial particles and with $n \approx e^t$ total particles.

The advantage of this more sophisticated model is that we can exploit the exact self-similarity property of the underlying space–time PPP. In particular, by reversing time the "line of descent" by which a particle acquires its color from previous particles can be studied. Moreover, suppose the first intuitive suggestion is true. That is, after assigning different colors to particles at positions z at time t_1, suppose there is a $t \to \infty$ limit random partition $\mathcal{A}(t_1)$ of $\mathbb{R}^2$ into regions $A(t_1, z)$ occupied by particles with the color of the particle at z at time t_1. This (supposed) partition valued process $(\mathcal{A}(t_1), \infty > t_1 > -\infty)$ has a simple intuitive description in reversed time. Given $\mathcal{A}(t_1)$ and the time-t_1 particle positions z, we obtain $\mathcal{A}(t_1 - dt)$ by the rule

> delete each particle with probability dt; for each deleted particle, at position z say, let z' be the nearest other particle position, and replace $A(t_1, z')$ by $A(t_1, z') \cup A(t_1, z)$.

The purpose of this paper is to prove two intertwined results; that the random partition $\mathcal{A}(t_1)$ does exist as a certain type of limit of the coloring process (Theorem 2); and that the resulting reversed-time process $(\mathcal{A}(t_1) : \infty > t_1 > -\infty)$ is a self-similar version of the process defined by the rule above (Theorem 1).

1.1. *Notation and more detailed outline.* Write $\mathbb{R} \times \mathbb{R}^2$ for the set with elements (t, z), interpreted as "time" $t \in \mathbb{R}$ and "position" $z \in \mathbb{R}^2$. Write Ξ for the Poisson point process on $\mathbb{R} \times \mathbb{R}^2$ with mean measure $e^t\,dt\,dz$. All the random objects considered in this paper will be constructed from Ξ. We write a typical "point" of Ξ as $\xi = (t_\xi, z_\xi)$ or $\zeta = (t_\zeta, z_\zeta)$. We consider ξ as the label for an immortal particle with arrival or "birth" time t_ξ at position z_ξ, and so

$$\Xi_{\le t} := \{\xi \in \Xi : t_\xi \le t\}$$

denotes the set of particles which are alive at time t. Define $\Xi_{<t}$ analogously. Write

$$\mathcal{Z}_{\le t} = \{z_\xi : \xi \in \Xi_{\le t}\}$$

for the *positions* of the particles at time t. Of course, $\mathcal{Z}_{\le t}$ and $\mathcal{Z}_{<t}$ are Poisson point processes on $\mathbb{R}^2$ with rate e^t, that is mean measure $e^t\,dz$, because $\int_{-\infty}^t e^s\,ds = e^t$. The self-similarity properties of the PPP—that $\mathcal{Z}_{\le t_1}$ is distributed as a spatial rescaling of $\mathcal{Z}_{\le t_2}$—will extend to self-similarity for the process $(\mathcal{A}(t_1) : \infty > t_1 > -\infty)$ outlined in the previous section.

To each particle ξ, let us assign a *parent* particle $\zeta = \text{parent}(\xi)$, defined as the particle in $\Xi_{t_\xi -}$ for which the Euclidean distance $\|z_\zeta - z_\xi\|$ is minimized. This defines a (genealogical) *tree process*. So for each particle ξ there is an *ancestral sequence* of particles, written $(\text{parent}[i,\xi], i \geq 0)$, defined by $\text{parent}[0,\xi] = \xi$ and then recursively by

$$\text{parent}[i+1,\xi] = \text{parent}\big(\text{parent}[i,\xi]\big), \qquad i \geq 0.$$

The associated *line of descent* indicates the ancestor of ξ at each time $t < t_\xi$, that is,

$$\begin{aligned} \text{ancestor}(t,\xi) &= \text{parent}[i,\xi] \\ &\qquad \text{for } i \geq 1 \text{ such that } t_{\text{parent}[i,\xi]} \leq t < t_{\text{parent}[i-1,\xi]}, \end{aligned} \tag{1}$$

where for completeness we define

$$\text{ancestor}(t,\xi) = \xi \qquad \text{for } t \geq t_\xi.$$

The first part of the proof (Proposition 4 in Section 2) shows that for a typical particle ξ present at time 0, the distance to $\text{ancestor}(-t,\xi)$, the ancestor at time $-t$, is of order $e^{t/2}$, which is the same order as the distance to the *nearest* particle present at time $-t$. In the second part of the proof (Section 3) we first consider two particles present at time 0 and distance r apart. Their lines of descent merge at some random past time $-T_r$, and we need an upper bound (Proposition 13) on the tail of the distribution of T_r. The methods in these sections are very concrete—calculations and bounds involving Euclidean geometry and spatial Poisson processes—though rather intricate in detail.

The limit result we seek involves descendants (rather than ancestors) of typical particles, and we set up notation as follows.[3] For $t_1 \leq t_2$ and $\zeta \in \Xi_{\leq t_1}$, define

$$\text{Descend}(t_1,t_2,\zeta) := \big\{\xi \in \Xi_{\leq t_2} : \text{ancestor}(t_1,\xi) = \zeta\big\}. \tag{2}$$

This is the set of particles born before t_2 whose time-t_1 ancestor in the line of descent was ζ. In the coloring story, this is "the set of particles at time t_2 which have inherited the same color as ζ, if we gave all the particles at t_1 different colors." Then, still for $t_1 \leq t_2$ and $\zeta \in \Xi_{\leq t_1}$, define

$$\begin{aligned} &\mu_{t_1,t_2,\zeta} \text{ is the measure } \mu \text{ putting weight } e^{-t_2} \\ &\qquad \text{on the position of each particle in } \text{Descend}(t_1,t_2,\zeta). \end{aligned} \tag{3}$$

So $\mu_{t_1,t_2,\zeta}$ is a random element of the space $\mathcal{M}(\mathbb{R}^2)$ of finite measures on $\mathbb{R}^2$, equipped with the usual topology of weak convergence. To obtain the limit theorem, we first show (Proposition 24 in Section 4.2) that there exist $t_2 \to \infty$ limits

[3]For ancestor-descendant pairs, we systematically write ζ for the ancestor and ξ for the descendant.

in probability [as $\mathcal{M}(\mathbb{R}^2)$-valued random variables]; that is, there exist random measures $\{\mu_{t_1,\infty,\zeta} : \zeta \in \Xi_{\le t_1}\}$ such that

$$\mu_{t_1,t_2,\zeta} \to \mu_{t_1,\infty,\zeta} \qquad \text{in probability as } t_2 \to \infty\ (\forall \zeta \in \Xi_{\le t_1}). \tag{4}$$

The proof essentially relies on Proposition 4 and self-similarity. We then use Proposition 13 to show that a limit $\mu_{t_1,\infty,\zeta}$ is in fact Lebesgue measure restricted to some random set $A(t_1,\zeta)$, implying that the collection $\{A(t_1,\zeta) : \zeta \in \Xi_{\le t_1}\}$ is necessarily a partition of $\mathbb{R}^2$.

For fixed t, we can regard

$$\mathbf{Z}^{(t)} = \bigl\{\bigl(z_\xi, A(t,\xi)\bigr) : \xi \in \Xi_{\le t}\bigr\}$$

as a marked point process. As t increases, the process $(\mathbf{Z}^{(t)}, -\infty < t < \infty)$ evolves in a way one can describe qualitatively:

> new points arrive randomly at rate e^t per unit area per unit time; when a point ξ arrives at time t, the region $A(t,\zeta)$ associated with the closest existing point ζ is split into two regions $A(t+dt,\zeta)$ and $A(t+dt,\xi)$.

But the probability distribution over possible splits depends on $\mathbf{Z}^{(t)}$ in some complicated way which we are unable to describe explicitly.

However, the key feature of this process is that as t *decreases* the regions merge according to the simple rule stated earlier. To summarize (see discussion below for the formalization of *random measurable set*), we have the following.

THEOREM 1. *The space–time PPP* $\{(t_\xi, z_\xi) : \xi \in \Xi\}$ *can be extended to a process* $\{(t_\xi, z_\xi, A(t,\xi), t \ge t_\xi) : \xi \in \Xi\}$ *with the following properties*:

(a) *For each* $-\infty < t < \infty$, *the collection* $\{A(t,\xi) : \xi \in \Xi_{\le t}\}$ *is a random partition of* $\mathbb{R}^2$ *into measurable sets*.

(b) *The distribution of the entire time-varying process* $\{(t_\xi, z_\xi, A(t,\xi), t \ge t_\xi) : \xi \in \Xi\}$ *is invariant under the action of the Euclidean group on* $\mathbb{R}^2$.

(c) *The process whose state at time* t *is* $\{(z_\xi, A(t,\xi)) : \xi \in \Xi_{\le t}\}$ *evolves in reversed time according to the rule*:

> *during* $[t, t-dt]$, *for each* $\xi \in \Xi_{\le t}$ *delete* ξ (*i.e., remove the entry* $(z_\xi, A(t,\xi))$] *with probability* dt; *for each deleted particle* ξ, *let* ζ *be the nearest other particle, and set* $A(t-dt,\zeta) = A(t,\zeta) \cup A(t,\xi)$.

(d) *The action of the scaling map* $z \to e^{-t/2}z$ *on* $\mathbb{R}^2$ *that takes the distribution of* $\mathcal{Z}_{\le 0}$ *to the distribution of* $\mathcal{Z}_{\le t}$ *also takes the distribution of* $\{(z_\xi, A(0,\xi)) : \xi \in \Xi_{\le 0}\}$ *to the distribution of* $\{(z_\xi, A(t,\xi)) : \xi \in \Xi_{\le t}\}$.

The earlier statement (4) can now be rephrased as follows, where we define $\mu_{t,t_2,\xi}$ as at (3) and consider it as an $\mathcal{M}(\mathbb{R}^2)$-valued random variable. Write Λ_A for Lebesgue measure restricted to $A \subset \mathbb{R}^2$.

THEOREM 2. *For each $\xi \in \Xi_{\le t}$, we have*

$$\mu_{t,t_2,\xi} \to \Lambda_{A(t,\xi)} \qquad \textit{in probability as } t_2 \to \infty,$$

where the limit random sets $\{A(t,\xi) : \xi \in \Xi_{\le t}\}$ comprise a process with the properties stated in Theorem 1.

Note this implies that the limit random sets here and in Theorem 1 are $\sigma(\Xi)$-measurable.

Theorem 2 is a formalization of the "limit colored regions exist" result described in the opening section, but this particular formalization is mathematically weak in two senses. Our formalization via weak convergence of empirical measures means, in the original "elementary" version, that we are ignoring positions of $o(n)$ size subsets of the n particles. Second, our proof gives no information about topological properties of the regions $A(t,\xi)$, only that they are measurable. In fact, as mentioned below (4) a random region $A(t,\xi)$ is formalized as a random measure for which the density function (with respect to Lebesgue measure on the plane) is a.e. 0 or 1, and so $A(t,\xi)$ is well defined only up to Lebesgue-null sets of $\mathbb{R}^2$. So, for instance, the natural question "is ξ an element of $A(t,\xi)$" is not well-posed. But it is natural to guess that the following is true.

CONJECTURE 3. *For each t, one can identify the regions $\{A(t,\xi) : \xi \in \Xi_{\le t}\}$ so that the topological boundary of each region has Lebesgue measure zero.*

If true, we could rephrase the question above as the well-posed question "is each ξ in the interior of $A(t,\xi)$," and we conjecture the answer is "yes." More interestingly, assuming Conjecture 3 is true, it is natural to conjecture that the boundaries have some (suitably defined) nonrandom fractal dimension $1 \le d < 2$, and Section 5.3 contains heuristic discussion. Further related remarks are in the next section. One might expect the regions to be *connected* sets, but this seems incorrect; see Section 5.3.2. Finally, as a referee noted, our results do not even imply that the number of regions intersecting the unit square is a.s. finite, so this constitutes another conjecture.

1.2. *Background and analogous models*. To quote the unpublished notes [11]:

> The [elementary] model is described in ([13], Section 7.6.8, pages 270–271), although we are not sure of its origins: [we] probably first learned of the problem from Mathew Penrose in about 2003, while Ben Hambly [personal communication] recalls that the same problem arose elsewhere at about the same time.

The context of that line of work was on-line algorithms in computational and stochastic geometry. Separately, the present author learned [personal communication] that the elementary model has been considered by Ohad Feldheim as a spatial analog of the Pólya urn process.

The approach in [11] to the elementary model is to identify colored regions in the unit square as Voronoi regions, that is, the set of points for which the nearest particle has a given color. Then via the Hausdorff metric on closed sets, it makes sense to ask whether our notion of convergence of empirical measures can be strengthened to include convergence of Voronoi regions. In our language and model, this could only be true if Conjecture 3 is true. Arguments in [11] focus on the length ℓ_n of the boundary between the two regions (for two colors and n particles in the unit square). Using arguments with a more geometric flavor than ours, they raise and discuss the question of whether $\ell_n = O(n^{d/2})$ for some $d < 2$. This mirrors our "fractal dimension" question, and indeed would imply that Conjecture 3 is true. The arguments in this paper make surprisingly little use of the "local geometry" of the PPP, so one can hope that our results might be combined with more geometric arguments to make further progress.

Note also that, intuitively, the area of the Voronoi region of a given color should behave almost as a martingale, because a new particle near the boundary seems equally likely to make the area larger or smaller. If one could bound the martingale approximation well enough to establish a.s. convergence of such areas, the results of this paper would follow rather trivially. But doing so seems to require detailed knowledge of the geometry of the boundary.

The author's own interest in the model arose in the context of a *scale-invariant random spatial network* (SIRSN) [1, 2], studied as abstractions of road networks. A general conjecture is that any network built dynamically from randomly-arriving Poisson points by means of edges (now line segments in the plane) being created to attach an arriving point to the existing network by a "scale-invariant rule" (i.e., a rule which uses only relative distances, not absolute distances) should in the limit define a SIRSN. Of course, the rule in our model "creates an edge from the newly-arrived point to the closest existing point" is about the simplest scale-invariant rule one can imagine.[4] The fact that this "simplest case" is hard to analyze suggests that the general conjecture is very challenging.

There is extensive literature on stochastic fragmentation and coalescence models in the nongeometric "mean-field" setting [3, 6]. There is also substantial literature (see, e.g., [7] Chapter 9) concerning random partitions of the plane (tessellations, tilings, etc.). But the combination of these themes, that is, Markovian processes of refining or coarsening partitions in the plane, have been considered only in special refining models [8] and in variants of the STIT model [9, 15]. The coalescing partitions process in Theorem 1 is perhaps the only known self-similar Markovian process of pairwise merging partitions of $\mathbb{R}^2$ with explicit rates. Obversely, the topic of Markovian models of coarsening partitions seems little investigated—see Section 5.2 for further brief comments.

[4]Unfortunately, the tree-like structure of this model implies it does not satisfy the requirement of a SIRSN that mean route lengths be finite.

2. A bound on ancestor displacement.

2.1. *Compactness for the marked point process.* Our first objective is to obtain a concrete bound, Proposition 4, on the distance between the position z_ξ of a particle ξ (present at time 0) and the position $z_{\text{ancestor}(-t,\xi)}$ of its ancestor at time $-t$.

Some notation:

- $\mathbf{0}$ is the origin in $\mathbb{R}^2$.
- $\|x-y\|$ denotes Euclidean distance in $\mathbb{R}^2$.
- $\text{disc}(z,r)$ is the closed disc with center z and radius r.
- For measurable $B \subset \mathbb{R}^2$, write $\text{area}(B)$ for its area (2-dimensional Lebesgue measure) and $\text{diam}(B) = \sup_{x,y\in B} \|x-y\|$ for its diameter.

PROPOSITION 4. *There exists a function $G(r) \downarrow 0$ as $r \uparrow \infty$ such that, for all $z \in \mathbb{R}^2$ and all $t > 0$, conditional on $\Xi_{\le 0}$ having a particle ξ with $z_\xi = z$ and $t_\xi > -t$, we have*

$$G_t(r) := \mathbb{P}\big(\|z_{\text{ancestor}(-t,\xi)} - z_\xi\| > re^{t/2}\big) \le G(r), \qquad 0 < r < \infty.$$

Moreover, $\int_0^\infty rG(r)\,dr < \infty$.

The rest of Section 2 is devoted to the proof of Proposition 4 and a variant (Proposition 9). As mentioned earlier, the conceptual point of Proposition 4 is that the distance to the time t (in the past) ancestor is the same order of magnitude as the distance to the closest particle at that time, that is, order $e^{t/2}$. An expression for $G(r)$ is given at (17).

The elementary "thinning" property of Poisson processes leads to a corresponding property of our space–time Poisson point process Ξ. As t runs backwards over $\infty > t > -\infty$, the processes $\Xi_{\le t}$ evolve according to the rule:

each particle is deleted at stochastic rate 1.

This *Poisson thinning process* representation is the foundation for much of our analysis, as are the related *self-similarity* properties of our derived processes, discussed in Section 4.1.

To be pedantic, in forward time, we work with the filtration $\mathcal{F}_t = \sigma(\Xi_{\le t})$. In reversed time, we work with the filtration

$$\overleftarrow{\mathcal{F}}_t = \sigma\big(\big(\max(t_\xi, t), z_\xi\big) : \xi \in \Xi\big). \tag{5}$$

So $\overleftarrow{\mathcal{F}}_t$ tells us the positions of all particles, and the arrival times of particles born after time t, and the following "thinning process" property holds.

LEMMA 5. *Conditional on $\overleftarrow{\mathcal{F}}_t$, the previous lifetimes $\{t - t_\xi : z_\xi \in \mathcal{Z}_{\le t}\}$ of the particles alive at time t are i.i.d. with* Exponential(1) *distribution.*

2.2. *Derivation of an EA process.* We study lines of descent in the genealogical tree process. Consider a particle ξ present at time 0 at position z_0. From the thinning process representation, its arrival time $T_0 < 0$ is such that $-T_0$ has Exponential(1) distribution. For $i \geq 1$, write (T_i, Z_i) for the arrival time and position of its ith generation ancestor, that is, parent$[i, \xi]$. We will show how to represent this process in terms of a certain Markov process we will call the *excluded area* (EA) process.

Conditional on $\{T_0 = t_0\}$ the particles present at t_0 are distributed as the PPP $\Xi_{<t_0}$, and so parent$[1, \xi]$ is the closest such point to z_0, at position Z_1 say. Conditional also on $\{Z_1 = z_1\}$, we know there are no points of $\Xi_{<t_0}$ in the interior of $C_1 := \mathrm{disc}(z_0, \|z_1 - z_0\|)$. The arrival time T_1 of Z_1 has density function $\propto e^t$ on $-\infty < t < t_0$, implying that $t_0 - T_1$ has Exponential(1) distribution.

Now given $T_0 = t_0$ and $(T_1, Z_1) = (t_1, z_1)$, the information we have about $\Xi_{<t_1}$ is precisely the fact that it has no points in C_1. So Z_2 is the closest point to z_1 in a PPP of rate e^{t_1} on $\mathbb{R}^2 \setminus C_1$. And as before, $t_1 - T_2$ has Exponential(1) distribution.

Now given $T_0 = t_0$ and $(T_1, Z_1) = (t_1, z_1)$ and $(T_2, Z_2) = (t_2, z_2)$, we have built an "excluded region" $C_2 := C_1 \cup \mathrm{disc}(z_1, \|z_2 - z_1\|)$. The information we have about $\Xi_{<t_2}$ is precisely that it is a PPP of rate e^{t_2} with no points in C_2, and we can continue inductively to describe the entire process $((T_i, Z_i), i \geq 0)$.

2.3. *Definition of the EA process.* Here, we re-specify the process above in intrinsic terms. Working with time $\downarrow -\infty$ is rather counter-intuitive, so in the definition below it seems helpful to reverse the direction of time.

Consider the space $\mathbb{C}$ of triples $\mathbf{c} = (C, z, \tau)$ such that

$$C \text{ is a compact set in } \mathbb{R}^2; \qquad z \in C; \qquad 0 \leq \tau < \infty.$$

Given an element $\mathbf{c} = (C, z, \tau) \in \mathbb{C}$, we can define a probability distribution $\mu_{\mathbf{c}}$ on $\mathbb{C}$ as follows. Take a PPP $\widetilde{\Xi}$ of rate $e^{-\tau}$ on $\mathbb{R}^2 \setminus C$. Let ξ be the point of $\widetilde{\Xi}$ closest to z. Set

$$z' = \xi; \qquad C' = C \cup \mathrm{disc}\big(z, \|\xi - z\|\big); \qquad \tau' = \tau + \theta,$$

where θ has Exponential(1) distribution independent of $\widetilde{\Xi}$. Then let $\mu_{\mathbf{c}}$ be the distribution of (C', z', τ').

Define the *EA process* to be the $\mathbb{C}$-valued Markov chain $(\mathbf{C}_i = (C_i, Z_i, \tau_i), 0 \leq i < \infty)$ where, for each step i, the conditional distribution of $\mathbf{C}_{i+1}$ given $\mathbf{C}_i = \mathbf{c}$ is the distribution $\mu_{\mathbf{c}}$ specified above. Figure 3 provides an illustration. It is straightforward to formalize the argument in Section 2.2 to show the following.

LEMMA 6. *Condition on Ξ containing a particle ξ with $t_\xi \leq 0$ and $z_\xi = z_0$. The process $((t_{\mathrm{parent}[i,\xi]}, z_{\mathrm{parent}[i,\xi]}), 0 \leq i < \infty)$ of arrival times and positions of the ancestors of this ξ is distributed as the random process $((-\tau_i, Z_i), 0 \leq i < \infty)$ within the EA process $((C_i, Z_i, \tau_i), 0 \leq i < \infty)$ with initial state $(\{z_0\}, z_0, \tau_0)$, where τ_0 has* Exponential(1) *distribution.*

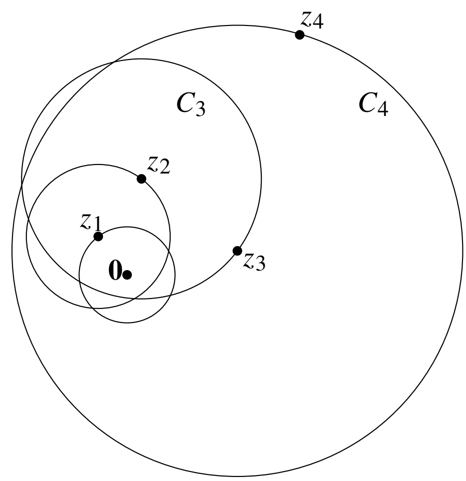

FIG. 3. *Illustration of the standard EA process.* C_i *is the union of the discs centered at* $\mathbf{0}, z_1, \dots, z_{i-1}$ *and* z_i *is on the boundary of* C_i.

Terminology. In what follows, we write *step* for the steps i of the EA chain, and *time* for the τ's.

2.4. *Geometric analysis of the EA process.* It is enough to study the *standard* EA process with initial state

$$(C_0, z_0, \tau_0) = (\{\mathbf{0}\}, \mathbf{0}, \tau_0),$$

where τ_0 has Exponential(1) distribution.[5] So in the context of Lemma 6 we will study ancestors of a particle present at position $\mathbf{0}$ at time 0. The starting observation, Lemma 7 below, is an expression for the growth of the area of C_i at each step. After that, we use geometric arguments to bound the diameter of C_i in terms of its area. Because Z_i is on the boundary of C_i this will be enough to prove Proposition 4.

LEMMA 7. *Conditional on* $\mathbf{C}_i = (C_i, z_i, \tau_i)$, *the increment* $\text{area}(C_{i+1}) - \text{area}(C_i)$ *has* $\text{Exponential}(e^{-\tau_i})$ *distribution*, *independent of* $\tau_{i+1} - \tau_i$.

PROOF. Writing a_r for the area of $\text{disc}(z_i, r) \setminus C_i$,

$$\begin{aligned}
&\mathbb{P}(\text{area}(C_{i+1}) - \text{area}(C_i) > a_r) \\
&\quad = \mathbb{P}(\text{ no point of a rate } e^{-\tau_i} \text{ Poisson process in } \text{disc}(z_i, r) \setminus C_i) \\
&\quad = \exp(-e^{-\tau_i} a_r).
\end{aligned}$$

[5]This is notationally more convenient than taking $\tau_0 = 0$, because of our convention that particles are labeled by position and arrival time.

The independence holds by construction. □

We can *lower* bound the diameter in terms of the area via the classical fact (called Bieberbach's inequality or the isodiametric inequality—see [12] for a short proof) that the disc is extremal:

$$\text{area}(C) \le \frac{\pi}{4}\big(\text{diam}(C)\big)^2, \qquad \text{all compact } C \subset \mathbb{R}^2. \tag{6}$$

We want a corresponding *upper* bound, to verify that C_j does not become long and thin. The bound will rely upon the following geometry lemma.

LEMMA 8. *Let C be a compact set in $\mathbb{R}^2$ and let D be a closed disc whose center is in C. Then*

$$\begin{aligned}&\text{diam}(C \cup D)\\ &\quad \le \max\left(\text{diam}(C) + \sqrt{\frac{2(\text{area}(C \cup D) - \text{area}(C))}{\pi}}, \sqrt{\frac{4\,\text{area}(C \cup D)}{\pi}}\right).\end{aligned}$$

PROOF. The right-hand side clearly bounds the distance between two points in C, and also between two points in D because

$$\sup_{z,z' \in D} \|z - z'\| = \text{diam}(D) = \sqrt{\frac{4}{\pi}\,\text{area}(D)} \le \sqrt{\frac{4}{\pi}\,\text{area}(C \cup D)}.$$

So it will suffice to prove the bound for one point in C and the other in D, that is to prove

$$\sup_{z \in C, z' \in D} \|z - z'\| - \text{diam}(C) \le \sqrt{\frac{2(\text{area}(C \cup D) - \text{area}(C))}{\pi}}. \tag{7}$$

Figure 4 illustrates the argument.

First, assume C is convex. If $D \subseteq C$, the result is trivial, so suppose not. Let y be a point on the boundary of D at maximal distance ($= r_0$, say) from C, and let w be a point in C with $\|y - w\| = r_0$. Then

$$\sup\{\|z - z'\| : z \in C, z' \in D\} \le \text{diam}(C) + r_0 \tag{8}$$

by applying the triangle inequality to the point in C closest to z'. Now consider the half-spaces defined by the line ℓ through w that is orthogonal to the line segment $\overline{wy}$. The convex set C must lie in the half-space not containing y, else by convexity some point in C would be closer to y. And the tangent line to the disc at y must be parallel to ℓ, otherwise some other point on the boundary would be farther from C than is y. But this implies that the line segment $\overline{wy}$ is part of the line segment $\overline{vy}$, where v is the center of the disc D. So $r_0 \le r :=$ radius of D, and

$$\text{area}(C \cup D) - \text{area}(C) \ge \text{area}(D \cap H),$$

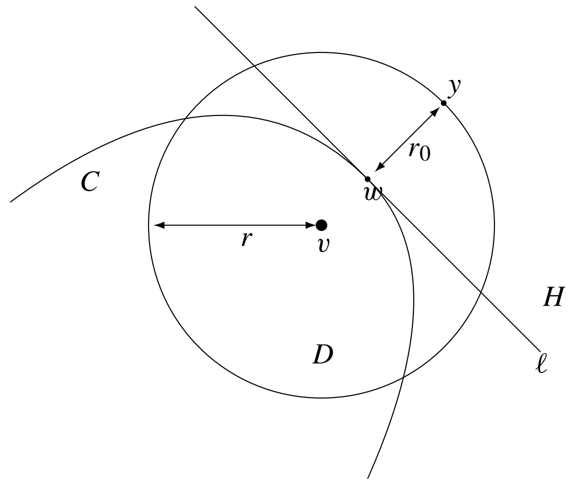

FIG. 4. *Illustration of proof of Lemma* 8.

where H is the half-space containing y. Now area$(D \cap H)$ is a certain function of r_0 and $r \geq r_0$, and clearly this function is, for fixed r_0, minimized at $r = r_0$, and there its value is $\frac{1}{2}\pi r_0^2$. So

$$\operatorname{area}(C \cup D) - \operatorname{area}(C) \geq \frac{1}{2}\pi r_0^2$$

and combining with (8) gives (7).

In proving (7), we assumed C was convex. For general C, we can apply (7) to its convex hull C^* and then, noting

$$\operatorname{diam}(C^*) = \operatorname{diam}(C), \qquad \operatorname{area}(C^* \cup D) - \operatorname{area}(C^*) \leq \operatorname{area}(C \cup D) - \operatorname{area}(C)$$

we see that (7) remains true for nonconvex C. $\square$

2.5. *Completing the proof of Proposition* 4. Returning to the standard EA process $\mathbf{C}_i = (C_i, Z_i, \tau_i)$, we now have sufficient tools to study τ_i and

$$A_i := \operatorname{area}(C_i), \qquad D_i := \operatorname{diam}(C_i).$$

From Lemma 7, we obtain a constructive representation of the distribution of $((A_i, \tau_i), 0 \leq i < \infty)$, as follows:

(9) The process $(\tau_i, i \geq 0)$ is a Poisson process of rate 1 on $(0, \infty)$.

(10)
$$A_i = \sum_{j=0}^{i-1} e^{\tau_j} \theta_j \qquad \text{where } (\theta_j, j \geq 0) \text{ are i.i.d. Exponential}(1),$$
independent of $(\tau_i, i \geq 0)$.

Then from Lemma 8 we get the inequality

$$D_{i+1} \le \max\left(D_i + \sqrt{\frac{2(A_{i+1} - A_i)}{\pi}}, \sqrt{\frac{4A_{i+1}}{\pi}}\right). \tag{11}$$

In this section, we use only the weaker inequality

$$D_{i+1} \le D_i + \sqrt{\frac{4A_{i+1}}{\pi}}. \tag{12}$$

Because $D_0 = 0$ this implies

$$D_k \le 2\pi^{-1/2} \sum_{i=1}^{k} A_i^{1/2}. \tag{13}$$

Because

$$A_i^{1/2} = \left(\sum_{j=0}^{i-1} e^{\tau_j}\theta_j\right)^{1/2} \le \sum_{j=0}^{i-1} (e^{\tau_j}\theta_j)^{1/2}$$

we find that

$$D_k \le \overline{D}_k := 2\pi^{-1/2} \sum_{j=0}^{k-1} (k-j) e^{\tau_j/2}\theta_j^{1/2}. \tag{14}$$

In Proposition 4, we seek to bound the probability of the event $\{\|z_{\text{ancestor}(-t,\xi)} - \mathbf{0}\| > re^{t/2}\}$ for a particle ξ at time 0 with position $z_\xi = \mathbf{0}$ (the case of general $z_\xi = z$ is the same, by translation-invariance). Fix t. Identifying the EA process with the "line of descent" process as in Lemma 6, the position $z_{\text{ancestor}(-t,\xi)}$ is by construction on the boundary of the region $C_{N(t)+1}$ for

$$N(t) = \max\{i : \tau_i < t\}.$$

Therefore, using (14),

$$\|z_{\text{ancestor}(-t,\xi)} - \mathbf{0}\| \le D_{N(t)+1} \le 2\pi^{-1/2} \sum_{j=0}^{N(t)} \big(N(t) + 1 - j\big) e^{\tau_j/2}\theta_j^{1/2}. \tag{15}$$

From properties of the rate-1 Poisson process $(\tau_j, j \ge 0)$ on $(0, \infty)$, the time-points $(t - \tau_{N(t)}, t - \tau_{N(t)-1}, t - \tau_{N(t)-1}, \dots, t - \tau_0)$ are distributed as an initial segment of a rate-1 Poisson process $(\sigma_j, j \ge 1)$ on $(0, \infty)$. So rewriting (15) in terms of the (σ_j) and $u = N(t) + 1 - j$ gives

$$\begin{aligned} &e^{-t/2}\|z_{\text{ancestor}(-t,\xi)} - \mathbf{0}\| \\ &\qquad \text{is stochastically dominated by } \chi := 2\pi^{-1/2} \sum_{u=1}^{\infty} u e^{-\sigma_u/2}\eta_u^{1/2}, \end{aligned} \tag{16}$$

where $(\eta_u, u \geq 1)$ are i.i.d. Exponential(1), independent of the Poisson process $(\sigma_u, u \geq 1)$. So Proposition 4 holds for

$$G(r) := \mathbb{P}(\chi > r), \qquad 0 < r < \infty \tag{17}$$

and it is easy to check that $\int_0^\infty r G(r)\, dr < \infty$.

2.6. *A large deviation bound for occupation times*. The following technical bound will enable us to bound the distance required for two lines of descent to merge (Proposition 13 later).

PROPOSITION 9. *For the standard EA process, write*

$$B = \bigcup_{\{i : e^{-\tau_i/2} \operatorname{diam}(C_i) > b\}} [\tau_{i-1}, \tau_i).$$

Then for sufficiently large b, there exist $A < \infty$ and $\rho > 0$ such that

$$\mathbb{P}\big(\mathrm{Leb}\big(B \cap [0, T]\big) > T/3\big) \leq A \exp(-\rho T), \qquad 0 < T < \infty, \tag{18}$$

where Leb *denotes Lebesgue measure.*

We previously used inequality (12) to bound $D_i := \operatorname{diam}(C_i)$ in terms of the areas $A_i = \operatorname{area}(C_i)$. Here, we will use a slightly different bound.

LEMMA 10. $D_i \leq \sqrt{\frac{4A_i}{\pi}} + \sum_{j=0}^{i-1} \sqrt{\frac{2(A_{j+1} - A_j)}{\pi}}$.

PROOF. Setting $\widetilde{D}_i := D_i - \sqrt{\frac{4A_i}{\pi}}$, inequality (11) becomes

$$\widetilde{D}_{i+1} \leq \max\left(\widetilde{D}_i + \sqrt{\frac{2(A_{i+1} - A_i)}{\pi}} + \left(\sqrt{\frac{4A_i}{\pi}} - \sqrt{\frac{4A_{i+1}}{\pi}}\right), 0\right).$$

But the term $(\sqrt{\frac{4A_i}{\pi}} - \sqrt{\frac{4A_{i+1}}{\pi}})$ is negative, and $\widetilde{D}_0 = 0$, so we find

$$\widetilde{D}_i \leq \sum_{j=0}^{i-1} \sqrt{\frac{2(A_{j+1} - A_j)}{\pi}}$$

establishing the asserted bound. $\square$

Recall the notation from (9,10): $(\tau_j, j \geq 0)$ denotes a Poisson process of rate 1 on $(0, \infty)$, and $(\theta_j, j \geq 0)$ denotes i.i.d. Exponential(1) random variables independent of $(\tau_i, i \geq 0)$. From (10) and Lemma 10,

$$D_i \leq \sqrt{\frac{4A_i}{\pi}} + \sqrt{\frac{2}{\pi}} D_i^*, \tag{19}$$

where

$$A_i = \sum_{j=0}^{i-1} e^{\tau_j}\theta_j; \qquad D_i^* = \sum_{j=0}^{i-1} e^{\tau_j/2}\theta_j^{1/2}.$$

To prove Proposition 9, we will rephrase inequalities from Section 2.5 in terms of continuous-time processes. These are processes of *Ornstein–Uhlenbeck-type*, in the terminology of [14]. Set $V_0 = 0$ and define $(V_t, 0 \le t < \infty)$ to be the process which increments by θ_j at time τ_j, and otherwise decreases at exponential rate 1. In symbols, writing $N_t^j = 1_{\{t \ge \tau_j\}}$,

$$dV_t = -V_t\,dt + \sum_{j \ge 0} \theta_j\,dN_t^j.$$

At time τ_i- (just before the jump at τ_i), we have

$$V_{\tau_i -} = \sum_{j=0}^{i-1} \theta_j e^{\tau_j - \tau_i} = e^{-\tau_i} A_i.$$

Because V_t is decreasing on $[\tau_{i-1}, \tau_i)$, we have

$$\text{if } e^{-\tau_i} A_i > b_A \qquad \text{then } V_t > b_A \text{ on } [\tau_{i-1}, \tau_i). \tag{20}$$

Similarly, define $(W_t, 0 \le t < \infty)$ to be the process which increments by $\theta_j^{1/2}$ at time τ_j, and otherwise decreases at exponential rate $1/2$. Take $W_0 = 0$. In symbols,

$$dW_t = -\frac{1}{2} W_t\,dt + \sum_j \theta_j^{1/2}\,dN_t^j.$$

At time τ_i-, we have

$$W_{\tau_i -} = \sum_{j=0}^{i-1} \theta_j^{1/2} e^{(\tau_j - \tau_i)/2} = e^{-\tau_i/2} D_i^*.$$

So as at (20)

$$\text{if } e^{-\tau_i/2} D_i^* > b_D \qquad \text{then } W_t > b_D \text{ on } [\tau_{i-1}, \tau_i). \tag{21}$$

Combining (19) with (20), (21) for appropriate b_A, b_D defined in terms of b, we now see that the proof of Proposition 9 reduces to proofs of large deviation bounds for occupation measures of the processes (V_t) and (W_t). That is, it suffices to prove the following.

PROPOSITION 11. *For sufficiently large b, there exist $A < \infty$ and $\rho > 0$ such that*

$$\mathbb{P}\left(\int_0^T 1_{\{V_t > b\}}\,dt > T/6\right) \le A \exp(-\rho T), \qquad 0 < T < \infty, \tag{22}$$

$$\mathbb{P}\left(\int_0^T 1_{\{W_t > b\}}\,dt > T/6\right) \le A \exp(-\rho T), \qquad 0 < T < \infty. \tag{23}$$

We will give the proof for (V_t), and then note that essentially the same proof works for (W_t).

Fix a high level b. The process regenerates at each downcrossing of b. So starting from the first downcrossing, there is an i.i.d. sequence $((L_b(i), K_b(i)), i \geq 1)$ where L_b is the duration and K_b is the "occupation time above b" between successive downcrossings. We can decompose L_b as $L'_b + K_b$ where L'_b is the time until first upcrossing of b and K_b is the subsequent time until the next downcrossing of b. It is easy to see

$$L'_b \to_p \infty \qquad \text{as } b \to \infty. \tag{24}$$

It is also easy to see that $K_b \to_p 0$ as $b \to \infty$, though we need the stronger result

$$\lim_{b \to \infty} \mathbb{E} \exp(\theta K_b) = 1, \qquad 0 < \theta < \infty. \tag{25}$$

To prove this, note that during an excursion above b the process (V_t) is upper bounded by the process (V_t^*) in which the drift term is $-b\,dt$ instead of $-V_t\,dt$. But the process (V_t^*) describes the workload in a $M/M/1$ queue with arrival rate 1 and service rate b. So the distribution of K_b is stochastically smaller than the server's busy period in that queue, and from classical exact formulas for that busy period distribution (e.g., [5]) one can deduce (25).

Writing τ_n for the time of the nth regeneration, (25) and the classical large deviation theorem for i.i.d. sums imply that for b sufficiently large

$$\mathbb{P}\left(\int_0^{\tau_n} 1_{\{V_t > b\}}\,dt > n/6\right) = \mathbb{P}\left(\sum_{i=1}^n K_b(i) > n/6\right) \qquad \text{decreases exponentially}$$

as $n \to \infty$. Next, by (24) we can choose b so that $\mathbb{P}(L'_b \geq 2) \geq 3/4$. Then

$$\mathbb{P}(\tau_n < n) \leq \mathbb{P}\left(\sum_{i=1}^n L'_b < n\right) \leq \mathbb{P}\left(\sum_{i=1}^n 1_{(L'_b \geq 2)} \leq n/2\right).$$

By the lower-tail Binomial large deviation bound, it follows that the probabilities $\mathbb{P}(\tau_n < n)$ decrease exponentially as $n \to \infty$, and this establishes (22).

The argument for (23) is essentially the same, and this completes the proof of Proposition 9.

3. Coalescence of lines of descent. In this section, we continue the style of analysis in Section 2 by studying the lines of descent of two particles present at time 0. This involves a coupled EA process, whose dynamics are described in Section 3.1. Note that Proposition 4 implies that, for particles at distance $r \gg 1$ apart, the "coalesce" time (time backwards to their most recent common ancestor) must be at least $(2 - o(1)) \log r$. Our goal is to give an upper bound, Proposition 13, on the coalesce time distribution. The central idea is to use Proposition 9 to show that, if not coalesced already, the lines of descent at time $-t$ are typically only order $e^{t/2}$ apart (the same order as the distance to the nearest time $-t$ particle): this is Proposition 16. A geometric argument then shows (Lemma 15) that there is a nonvanishing chance to merge in the next generation backwards. These ingredients are combined in Section 3.4 to prove Proposition 13.

3.1. *The coupled EA process*. Fix $t_0 \ge 0$. For the rest of Section 3 we condition on the time-t_0 configuration $\mathcal{Z}_{\le t_0}$ containing a particle at position z_0^1 and the time-0 configuration $\mathcal{Z}_{\le 0}$ containing a particle at position z_0^2. The distribution of the line of descent for each particle is just a translated and scaled version of the distribution of the EA process in Lemma 6. So we anticipate that the joint distribution of the two lines of descent can be described in terms of some suitably coupled versions of the EA process.

Precisely, we will specify the *coupled EA process*

$$((\mathbf{C}_i^1; \mathbf{C}_i^2), i = 0, 1, 2, \ldots) = ((C_i^1, z_i^1, \tau_i^1); (C_i^2, z_i^2, \tau_i^2)), i = 0, 1, 2, \ldots)$$

with initial states $(\{z_0^1\}, z_0^1, \tau_0^1)$ and $(\{z_0^2\}, z_0^2, \tau_0^2)$ where τ_0^2 and $\tau_0^1 + t_0$ are independent Exponential(1). At each step (before the coalescence step I_{coal} below), only one of the components ($\mathbf{C}_i^1$ or $\mathbf{C}_i^2$) changes. There are notational issues in describing this coupled processes. We write $(\mathbf{C}_i^1; \mathbf{C}_i^2)$ for the configuration after the ith step of the coupled process. Because only one component changes in each step before coalescence, we need different notation for the configuration of a given component after j changes of that particular component, and we write $(\mathbf{C}_{(j)}^1)$ and $(\mathbf{C}_{(j)}^2)$ for these "jump processes" of each component. And it is these jump processes which individually are evolving as the ordinary EA process.

The evolution rule for the coupled process, which will be derived from the dynamics of the underlying tree process as was done in Lemma 6, is as follows:

> Write the configuration after i steps as $(\mathbf{C}_i^1; \mathbf{C}_i^2)$. Before step I_{coal}, we must have $\tau_i^1 \ne \tau_i^2$; suppose $\tau_i^1 < \tau_i^2$ (the other case is symmetric). Take a PPP of rate $e^{-\tau_i^1}$ on $\mathbb{R}^2 \setminus (C_i^1 \cup C_i^2)$, but augmented with an extra point planted at z_i^2. Let ξ be the point of the augmented PPP closest to z_i^1. If $\xi = z_i^2$, then we say that the process coalesces at position z_i^2 and time τ_i^2; write $I_{\text{coal}} = i + 1$ for the coalesce step, $Z_{\text{coal}} = z_i^2$ for the coalesce position, $T_{\text{coal}} = \tau_i^2$ for the coalesce time. Otherwise, set $C_{i+1}^1 = C_i^1 \cup \operatorname{disc}(z_i^1, \|\xi - z_i^1\|)$, set $z_{i+1}^1 = \xi$ and take $\tau_{i+1}^1 = \tau_i^1 + \theta$ where θ has Exponential(1) distribution independent of all previously constructed random variables. Set $\mathbf{C}_{i+1}^2 = \mathbf{C}_i^2$.

Note in particular that the configuration after i steps determines the value

$$\tau_i := \min(\tau_i^1, \tau_i^2); \tag{26}$$

the arg min determines which component will change on the next step and τ_i determines the rate $e^{-\tau_i}$ of the PPP used to construct the next step.

REMARK. For completeness, let us give the behavior of the coupled process after coalescence, though this is not directly relevant to our arguments. If the coalesce step is $i+1$ as above, then $z^1_{i+1} = z^2_{i+1}$ and $\tau^1_{i+1} = \tau^2_{i+1}$ but maybe $C^1_{i+1} \neq C^2_{i+1}$. In subsequent steps k, we use the same PPP outside $C^1_k \cup C^2_k$ for each component and, therefore, we have $(\tau^1_k, z^1_k) = (\tau^2_k, z^2_k)$ for all $k \geq i+1$. Each of the two component jump processes does evolve as the EA process, except that for the first component process there is the extra planted point at z^2_0. But this extra point only comes into play if it is the exact point of coalescence, and so does not affect our arguments for upper bounding the coalesce time.

A realization of six initial steps of the coupled process is illustrated in Figure 5. On the left are the successive positions $z^1_0, z^1_{(1)}, z^1_{(2)}, z^1_{(3)}, z^1_{(4)}$ in the first component process, and on the right are the positions in the second component process. The associated times for the first process are $-t_0 < \tau^1_0 = \tau^1_{(0)} < \tau^1_{(1)} < \tau^1_{(2)} < \tau^1_{(3)} < \tau^1_{(4)}$ and for the second process are $0 < \tau^2_0 = \tau^2_{(0)} < \tau^2_{(1)} < \tau^2_{(2)}$. Suppose that the times associated with these steps are ordered as

$$-t_0 < \tau^1_{(0)} < \tau^1_{(1)} < \tau^1_{(2)} < \tau^2_{(0)} < \tau^1_{(3)} < \tau^2_{(1)}.$$

In terms of steps i of the coupled process (indicated in the figure as a, b, c, d, e, f), we have

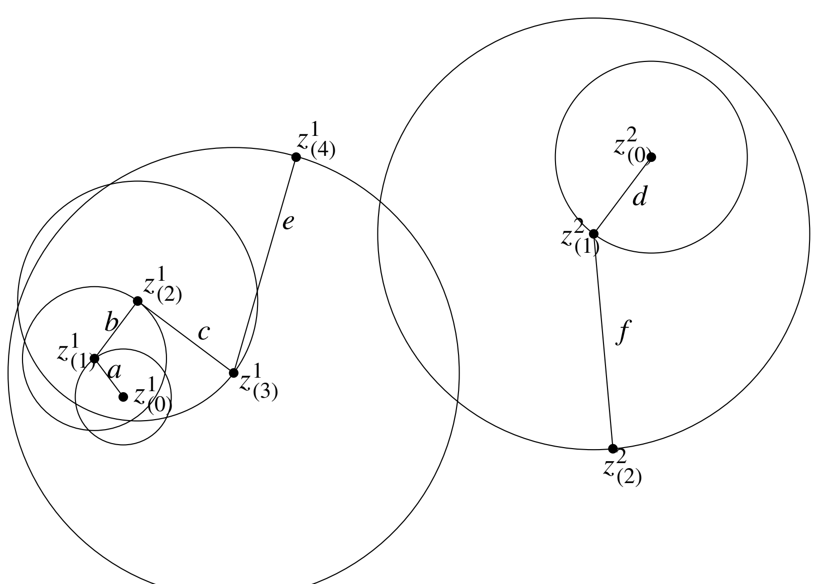

FIG. 5. *Initial steps of the coupled EA process.*

i	0	$a1$	$b2$	$c3$	$d4$	$e5$	$f6$
$z_i^1 =$	z_0^1	$z_{(1)}^1$	$z_{(2)}^1$	$z_{(3)}^1$	$z_{(3)}^1$	$z_{(4)}^1$	$z_{(4)}^1$
$z_i^2 =$	z_0^2	z_0^2	z_0^2	z_0^2	$z_{(1)}^2$	$z_{(1)}^2$	$z_{(2)}^2$
$\tau_i^1 =$	τ_0^1	$\tau_{(1)}^1$	$\tau_{(2)}^1$	$\tau_{(3)}^1$	$\tau_{(3)}^1$	$\tau_{(4)}^1$	$\tau_{(4)}^1$
$\tau_i^2 =$	τ_0^2	τ_0^2	τ_0^2	τ_0^2	$\tau_{(1)}^2$	$\tau_{(1)}^2$	$\tau_{(2)}^2$
$\tau_i =$	τ_0^1	$\tau_{(1)}^1$	$\tau_{(2)}^1$	τ_0^1	$\tau_{(3)}^1$	$\tau_{(1)}^1$	??

We will now relate how this description of the coupled EA process arises as the description of the two lines of descent within the tree process for the two given points of $\mathcal{Z}_{\le t_0}$ and $\mathcal{Z}_{\le 0}$ at positions z_0^1 and z_0^2. Consider Figure 5. Inductively, we have traced back the two lines of descent to $(-\tau_{(4)}^1, z_{(4)}^1)$ and $(-\tau_{(2)}^2, z_{(2)}^2)$, using 6 steps of the coupled process. What happens next depends on which of $\tau_{(4)}^1$ or $\tau_{(2)}^2$ (that is, which of τ_6^1 or τ_6^2) is smaller. Taking the case $\tau_{(4)}^1 < \tau_{(2)}^2$ (the other case is symmetric) then, to find the parent of $(-\tau_{(4)}^1, z_{(4)}^1)$ in the tree process, we need to search the region where vertices may have arrived before $-\tau_{(4)}^1$; this excludes both $C_{(4)}^1$ and the interior of $C_{(2)}^2$, because the latter contains no particles arriving before $-\tau_{(1)}^2$, and we must have $\tau_{(1)}^2 < \tau_{(4)}^1$ because of the rule that the component with smaller τ-value expanded first. However, the particle at $z_{(2)}^2$ arrived at time $-\tau_{(2)}^2$, which was before time $-\tau_{(4)}^1$, and so is eligible to be the parent of $(-\tau_{(4)}^1, z_{(4)}^1)$. So the parent of $(-\tau_{(4)}^1, z_{(4)}^1)$ is the closest particle to $z_{(4)}^1$ in the Poisson process $\mathcal{Z}_{\le -\tau_{(4)}^1}$, which has rate $e^{-\tau_{(4)}^1}$, on the complement of $C_{(4)}^1 \cup C_{(2)}^2$, or is $z_{(2)}^2$ if that particle is closer. In the latter case, the two lines of descent coalesce at $(-\tau_{(2)}^2, z_{(2)}^2)$.

In terms of steps i of the coupled process, $\tau_{(4)}^1 = \tau_6^1$ and $C_{(4)}^1 \cup C_{(2)}^2 = C_6^1 \cup C_6^2$ and $(-\tau_{(2)}^2, z_{(2)}^2) = (-\tau_6^2, z_6^2)$. This completes the derivation of the evolution rule stated in Section 3.1.

To summarize, we have the following.

LEMMA 12. *Condition on* Ξ *containing particles* ξ^1 *and* ξ^2 *with* $t_{\xi^1} \le t_0, t_{\xi^2} \le 0$ *and* $(z_{\xi^1}, z_{\xi^2}) = (z_0^1, z_0^2)$. *The joint process*

$$\big((t_{\text{parent}[i,\xi^1]}, z_{\text{parent}[i,\xi^1]}, t_{\text{parent}[i,\xi^2]}, z_{\text{parent}[i,\xi^2]}), 0 \le i < \infty\big)$$

of arrival times and positions of the ancestors of these particles is distributed as the random process $((-\tau_{(i)}^1, Z_{(i)}^1, -\tau_{(i)}^2, Z_{(i)}^2), 0 \le i < \infty)$ *within the coupled EA process* $((C_i^1, Z_i^1, \tau_i^1); (C_i^2, Z_i^2, \tau_i^2), 0 \le i < \infty)$ *with initial states* $(\{z_0^1\}, z_0^1, \tau_{(0)}^1)$ *and* $(\{z_0^2\}, z_0^2, \tau_{(0)}^2)$ *where* $\tau_{(0)}^1 + t_0$ *and* $\tau_{(0)}^2$ *are independent with* Exponential(1) *distribution.*

As mentioned before, we write I_{coal} for the first step I such that $z_I^1 = z_I^2$ (and call that point Z_{coal}), or equivalently for the first step I such that $\tau_I^1 = \tau_I^2$ (and call that time T_{coal}). We can now state the main result of Section 3.

PROPOSITION 13. *There exist constants $K, \beta < \infty$ and $\rho > 0$ such that, in the coupled EA process above, for any (z_0^1, z_0^2) and any $t_0 \geq 0$,*

$$\mathbb{P}(T_{\mathrm{coal}} > t) \leq K \exp(-\rho t) \qquad \text{for all } t > \beta \log^+ \|z_0^1 - z_0^2\|. \tag{27}$$

Heuristically, we expect $\|Z_{\mathrm{coal}} - z_0^1\| \asymp \exp(T_{\mathrm{coal}}/2)$ in the tails, and so the tail behavior of the form $\mathbb{P}(T_{\mathrm{coal}} > t) \asymp \exp(-\gamma t/2)$ would be equivalent to the tail behavior of the form $\mathbb{P}(\|Z_{\mathrm{coal}}\| > r) \asymp r^{-\gamma}$. We conjecture the latter is true; precisely, that the exponent

$$\gamma := -\lim_{r\to\infty} \frac{\log \mathbb{P}(\|Z_{\mathrm{coal}}\| > r)}{\log r} \tag{28}$$

exists and does not depend on $\|z_0^1 - z_0^2\|$ or t_0. This is closely related to the "are boundaries fractal" issue, as will be discussed in Section 5.3.

For ease of exposition, we will present the proof of Proposition 13 in the case $t_0 = 0$. The general case requires only minor modifications, noted below.

3.2. *The coalescence step.* Here, we will give conditions to ensure a non-vanishing probability of coalescing at the next step. This requires only a simple geometric lemma.

LEMMA 14. *Let $z_1, z_2 \in \mathbb{R}^2$, and let $r, \lambda > 0$. Let $\mathcal{Z}^\lambda$ be a Poisson point process of rate λ on $\mathbb{R}^2$. Then the event the nearest point $z \in \mathcal{Z}^\lambda$ to z_1 is also the nearest point to z_2, and $\min(\|z - z_1\|, \|z - z_2\|) > r$ has probability at least*

$$\exp\bigl(-\lambda\pi\bigl(r + \|z_2 - z_1\|\bigr)^2\bigr) - 2c_0\lambda^{1/2}\|z_2 - z_1\| \tag{29}$$

for a certain absolute constant c_0.

PROOF. Write $d = \|z_2 - z_1\|$. Consider the events:

(A): the distance from z_1 to the nearest point of $\mathcal{Z}^\lambda$ is at least $r + d$;

(B): the distances $D_1^{(\lambda)}$ and $D_2^{(\lambda)}$ from z_2 to the nearest two points of $\mathcal{Z}^\lambda$ are such that $D_2^{(\lambda)} - D_1^{(\lambda)} \leq 2d$.

We assert that, in order that the event in Lemma 14 occurs, it is sufficient that the event (A) occurs and the event (B) does not occur. To prove this assertion, let z_1' be the closest point of $\mathcal{Z}^\lambda$ to z_1, and let z_2' be the closest point of $\mathcal{Z}^\lambda$ to z_2. Suppose $z_2' \neq z_1'$. By the triangle inequality,

$$\|z_1 - z_2'\| \leq d + \|z_2 - z_2'\| \quad \text{and} \quad \|z_2 - z_1'\| \leq d + \|z_1 - z_1'\|.$$

Because z_1' is the closest point to z_1

$$\|z_1 - z_1'\| < \|z_1 - z_2'\|.$$

Combining these three inequalities leads to

$$\|z_2 - z_1'\| - \|z_2 - z_2'\| < 2d.$$

So if (B) fails, then then $z_1' = z_2' = z$ say. And so if (A) holds then, by the triangle inequality, $\min(\|z - z_1\|, \|z - z_2\|) > r$.

The probability $\mathbb{P}(A)$ equals the first term in (29). It is easy to check that $D_2^{(1)} - D_1^{(1)}$ has a density bounded by some c_0, so (by scaling) the density of $D_2^{(\lambda)} - D_1^{(\lambda)}$ is bounded by $c_0\lambda^{1/2}$. So $\mathbb{P}(B)$ is at most $2c_0\lambda^{1/2}d$. $\square$

We now apply this to the coalescence step.

LEMMA 15. *Consider a state* $(\mathbf{c}^1, \mathbf{c}^2) = ((C^1, z^1, \tau^1), (C^2, z^2, \tau^2))$ *of the coupled EA process started at* z_0^1 *and* z_0^2. *Suppose* $\tau^1 < \tau^2$. *Write* $\bar{\Delta} = \max(\operatorname{diam}(C^1), \operatorname{diam}(C^2)$. *Then the probability that the process coalesces at the next step is at least*

$$\exp\bigl(-4\pi e^{-\tau^1}(\|z_0^1 - z_0^2\| + 2\bar{\Delta})^2\bigr) - 2c_0\|z_0^1 - z_0^2\|e^{-\tau^1/2}. \tag{30}$$

PROOF. Because $z_0^1 \in C^1$ and $z_0^2 \in C^2$, we have

$$\|z^1 - z^2\| \le \|z_0^1 - z_0^2\| + 2\bar{\Delta} := r$$

and moreover, each $\operatorname{disc}(z^i, r)$ contains $C^1 \cup C^2$. We can now apply Lemma 14 with $\lambda = e^{-\tau^1}$; if the event in Lemma 14 occurs then the process coalesces at the next step. $\square$

3.3. *Diameters in the coupled process.* The key ingredient in proving Proposition 13 is the following extension of Proposition 9 to the coupled EA process, which will enable us to apply Lemma 15. This extension looks “obvious” but the proof is rather fussy.

Fix large b, and regard a step i of the coupled EA process as “good” if

$$e^{-\tau_i/2}\max\bigl(\operatorname{diam}(C_i^1), \operatorname{diam}(C_i^2)\bigr) \le b. \tag{31}$$

Let $N_b(T)$ be the number of “good” steps before[6] time T.

PROPOSITION 16. *In the coupled EA process, for sufficiently large* b *there exist constants* $a, \rho > 0$ *and* $K < \infty$ *such that*

$$\mathbb{P}\bigl(T_{\text{coal}} > T, N_b(T) \le aT\bigr) \le K\exp(-\rho T), \qquad 0 < T < \infty.$$

[6]In the general case $t_0 > 0$, we only count the number of steps with $\tau_i^1 > 0$.

The bound does not depend on the distance $\|z_0^1 - z_0^2\|$ between the particles.

In this Section 3.3, for an event indexed by T we say the event "has vanishing probability" if the probabilities are $O(\exp(-\rho T))$ as $T \to \infty$, for some $\rho > 0$. As in Proposition 9, for $j = 1, 2$ write

$$B^j = \bigcup_{\{i : e^{-\tau_i^j/2} \operatorname{diam}(C_i^j) > b\}} [\tau_{i-1}^j, \tau_i^j).$$

Note this is the same if use the indices $\tau_{(i)}^j$ of the jump processes.

Write

$$\widetilde{N}_b(T) := \operatorname{Leb}\big((B^1 \cup B^2)^c \cap [0, T]\big).$$

In words, this is the duration of time for which a "good" event similar to (31) is occurring. Applying Proposition 9 to both components of the coupled process, for sufficiently large b:

$$\text{the event } \{T_{\text{coal}} > T, \widetilde{N}_b(T) < T/3\} \qquad \text{has vanishing probability.} \tag{32}$$

This is almost what we are trying to prove as Proposition 16, except that we need to switch from "duration of time" $\widetilde{N}_b(T)$ to "number of steps" $N_b(T)$.

In the construction of the coupled EA process, we can start with two independent rate-1 PPPs on $(0, \infty)$ and use these as the values of $\tau_{(i)}^1$ and $\tau_{(i)}^2$ until the coalescence step. So on the event $\{T_{\text{coal}} > T\}$ these times, within $[0, T]$, coincide with the times of two independent rate-1 PPPs. Here, is a helpful way to record a consequence of this fact.[7]

LEMMA 17. *Let B_T be an event defined in terms of two independent rate-1 PPPs on $[0, T]$. Let B_T^* be the corresponding event defined in terms of the times $\tau_{(0)}^1 < \tau_{(1)}^1 < \tau_{(2)}^1 < \cdots < T$ and $\tau_{(0)}^2 < \tau_{(1)}^2 < \tau_{(2)}^2 < \cdots < T$ in the coupled EA process. Then*

$$\mathbb{P}\big(T_{\text{coal}} > T, B_T^*\big) \le \mathbb{P}(B_T).$$

We will apply this to events B_T which have vanishing probability, in which setting Lemma 17 says we can ignore such events for the purpose of proving Proposition 16. We state the following straightforward large deviation bounds for quantities associated with the PPP.

LEMMA 18. *Let (τ_i) be a rate-2 PPP on $(0, \infty)$, let $a > 0$ and let $H_T(a)$ be the sum of the lengths of the $\lfloor aT \rfloor$ longest intervals (τ_i, τ_{i+1}) with $\tau_i < T$. Then, for sufficiently small a,*

$$\textit{the event } \{H_T(a) \ge T/4\} \qquad \textit{has vanishing probability.} \tag{33}$$

[7] The fact is slightly subtle, in that the previous assertion is not true *conditional on the event* $\{T_{\text{coal}} > T\}$.

LEMMA 19. *Let (τ_i) be a rate-2 PPP on $(0,\infty)$, represented as the superposition of two independent rate-1 PPPs. Define τ_k^+ as the minimum value of τ_ℓ for $\ell \ge k+2$ such that the events at $\tau_{k+1}.\tau_{k+2},\dots,\tau_\ell$ include events from both component processes. Define*

$$G_T(a') = \sum\{\tau_k^+ - \tau_k : \tau_k \le T, \tau_k^+ - \tau_k > a'\}.$$

Then, for sufficiently large a',

$$\text{the event } \{G_T(a') \ge T/12\} \text{ has vanishing probability.} \tag{34}$$

Now choose a sufficiently small and a' sufficiently large that inequalities (33) and (34) hold. Write $H_T^*(a)$ and $G_T^*(a')$ for the random variables corresponding (as in Lemma 17) to $H_T(a)$ and $G_T(a')$ defined in terms of the times in the coupled EA process. Consider the event

$$\{T_{\text{coal}} > T, H_T^*(a) < T/4, G_T^*(a') < T/12, \widetilde{N}_b(T) \ge T/3\}. \tag{35}$$

On this event, take the "good" intervals comprising $\widetilde{N}_b(T)$ (with total length $\ge T/3$) and delete the intervals comprising $G_T^*(a')$ (with total length $< T/12$). There remain "good" intervals with total length $> T/4$, so there are at least $\lfloor aT \rfloor$ such intervals. In other words, on event (35) there are at least $\lfloor aT \rfloor$ steps i of the coupled process such that

$$[\tau_i, \tau_{i+1}) \qquad \text{is disjoint from } B^1 \cup B^2 \text{ and } \tau_i^+ \le \tau_i + a'. \tag{36}$$

For each such step i we have (see argument below)

$$e^{-\tau_i/2} \max(\operatorname{diam}(C_i^1), \operatorname{diam}(C_i^2)) \le be^{a'/2} := \beta, \qquad \text{say.} \tag{37}$$

In other words, on the event (35) we have $N_\beta(T) \ge \lfloor aT \rfloor$. So now,

$$\begin{aligned}
&\mathbb{P}(T_{\text{coal}} > T, N_\beta(T) \le aT) \\
&\quad \le \mathbb{P}(T_{\text{coal}} > T, H_T^*(a) > T/4) \\
&\qquad + \mathbb{P}(T_{\text{coal}} > T, G_T^*(a') > T/12) + \mathbb{P}(T_{\text{coal}} > T, \widetilde{N}_b(T) < T/3) \\
&\quad \le \mathbb{P}(H_T(a) > T/4) + \mathbb{P}(G_T^*(a') > T/12) + \mathbb{P}(T_{\text{coal}} > T, \widetilde{N}_b(T) < T/3)
\end{aligned}$$

using Lemma 17. Each term on the right has vanishing probability, by (32) and (33) and (34), and this establishes Proposition 16 (with β in place of b).

The argument that (36) implies (37) is illustrated by the case in Figure 6. Consider $\tau_k = \min(\tau_k^1, \tau_k^2) = \min(\tau_{(j)}^1, \tau_{(k-j)}^2)$ for some j, and suppose that (as in the figure) $\tau_{(j)}^1 > \tau_{(k-j)}^2$. Saying that

$$[\tau_k, \tau_{k+1}) \qquad \text{is disjoint from } B^1 \cup B^2$$

is saying that

$$[\tau_{(j-1)}^1, \tau_{(j)}^1) \text{ is not in } B^1, \quad \text{and} \quad [\tau_{(k-j)}^2, \tau_{(k-j+1)}^2) \text{ is not in } B^2$$

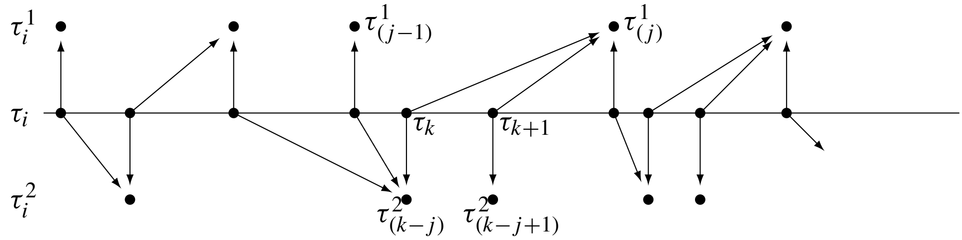

FIG. 6. *The times τ_i of steps of the coupled process are shown on the axis. The arrows point to the times $\tau_{(j)}^1$ and $\tau_{(i-j)}^2$ associated with the completed steps of the component processes.*

which is saying that

$$e^{-\tau_{(j)}^1/2}\operatorname{diam}(C_k^1) \le b \quad \text{and} \quad e^{-\tau_{(k-j+1)}^2/2}\operatorname{diam}(C_{k+1}^2) \le b.$$

Now consider the first time after τ_k that both components have expanded, that is,

$$\tau_k^+ := \min\{\tau_j : \tau_j^1 > \tau_k^1 \text{ and } \tau_j^2 > \tau_k^2\}.$$

Then the inequality above implies

$$e^{-\tau_k^+/2}\max(\operatorname{diam}(C_k^1), \operatorname{diam}(C_k^2)) \le b.$$

So when $\tau_k^+ \le \tau_k + a'$, we have (37).

3.4. *Proof of Proposition* 13. We need the following standard martingale-type bound.

LEMMA 20. *Let $S \ge 1$ be a stopping time for a filtration $(\mathcal{F}_n)$. For any $0 < p_0 < 1$ and $m \ge 1$,*

$$\mathbb{P}(S > n) \le (1 - p_0)^m + \mathbb{P}(L(n, p_0) < m, S > n),$$

where

$$L(n, p_0) = \big|\{i : 1 \le i \le n, \mathbb{P}(S = i | \mathcal{F}_{i-1}) \ge p_0\}\big|.$$

PROOF. The process (M_n) with $M_0 = 1$ and, for $n \ge 1$,

$$\begin{aligned} M_n &= 0 \qquad \text{on } \{S \le n\} \\ &= \frac{1}{\prod_{1 \le i \le n} \mathbb{P}(S > i | \mathcal{F}_{i-1})} \qquad \text{on } \{S > n\} \end{aligned}$$

is a martingale. On the event $\{L(n, p_0) \ge m, S > n\}$, we have $M_n \ge (1 - p_0)^{-m}$ and so

$$1 = \mathbb{E}M_n \ge \mathbb{E}M_n 1_{(L(n,p_0)\ge m, S>n)} \ge (1 - p_0)^{-m}\mathbb{P}(L(n, p_0) \ge m, S > n).$$

Because

$$\mathbb{P}(S > n) = \mathbb{P}\big(L(n, p_0) \geq m, S > n\big) + \mathbb{P}\big(L(n, p_0) < m, S > n\big)$$

the result follows. $\square$

We can now combine previous ingredients to prove Proposition 13. Suppose i is such that $I_{\text{coal}} > i$ and the configuration $(\mathbf{C}_i^1; \mathbf{C}_i^2)$ satisfies (31). Then by (30) the probability of coalescing on the next step is at least

$$\exp\big(-4\pi\big(\|z_0^1 - z_0^2\|e^{-\tau_i/2} + 2b\big)^2\big) - 2c_0\|z_0^1 - z_0^2\|e^{-\tau_i/2}.$$

So there exist constants $\alpha > 0$ and $p_0 > 0$ (determined by b and c_0) such that, for the natural filtration $(\mathcal{F}_i)$ of the coupled EA process,

$$\begin{aligned} &\mathbb{P}(I_{\text{coal}} = i + 1|\mathcal{F}_i) \geq p_0 \\ &\quad \text{on } \big\{I_{\text{coal}} > i, \|z_0^1 - z_0^2\|e^{-\tau_i/2} \leq \alpha,\ (\mathbf{C}_i^1; \mathbf{C}_i^2) \text{ satisfies (31)}\big\}. \end{aligned} \tag{38}$$

Appealing to Lemma 20,

$$\mathbb{P}(I_{\text{coal}} > n) \leq (1 - p_0)^m + \mathbb{P}(L_n < m, I_{\text{coal}} > n), \tag{39}$$

where

$$L_n = \big|\big\{i : 0 \leq i \leq n - 1, \mathbb{P}(I_{\text{coal}} = i + 1|\mathcal{F}_i) \geq p_0\big\}\big|.$$

Recall the definition of $N_b(t)$ in Proposition 16. Take $a > 0$ (to be specified later) and consider some $m < an$. If

$$\|z_0^1 - z_0^2\|e^{-\tau_j/2} \leq \alpha \qquad \text{for } j = \lfloor an\rfloor - m \tag{40}$$

then, on the event $\{I_{\text{coal}} > n, N_b(t) > an\}$, the events in (38) hold for at least m values of $i \leq n$, which implies $L_n \geq m$. So if (40) holds, then

$$\mathbb{P}(L_n < m, I_{\text{coal}} > n) \leq \mathbb{P}\big(I_{\text{coal}} > n, N_b(n) \leq an\big)$$

and then using (39) we have

$$\begin{aligned} &\mathbb{P}(I_{\text{coal}} > n) \\ &\quad \leq (1 - p_0)^m + \mathbb{P}\big(I_{\text{coal}} > n, N_b(n) \leq an\big) + \mathbb{P}\big(I_{\text{coal}} > n,\ \text{event (40) fails}\big). \end{aligned}$$

Proposition 16 implies that for sufficiently large b there exist constants $a, \rho > 0$ and $K < \infty$ such that

$$\mathbb{P}\big(T_{\text{coal}} > n/3, N_b(n) \leq an\big) \leq K\exp(-\rho n), \qquad n = 1, 2, 3, \ldots.$$

Using this choice of a above,

$$\begin{aligned} \mathbb{P}(I_{\text{coal}} > n) &\leq (1 - p_0)^m + O_{\exp}(n) \\ &\quad + \mathbb{P}(I_{\text{coal}} > n, T_{\text{coal}} \leq n/3) + \mathbb{P}\big(I_{\text{coal}} > n,\ \text{event (40) fails}\big), \end{aligned}$$

where $O_{\exp}(n)$ denotes a "vanishing probability" sequence which is $O(\rho^n)$ as $n \to \infty$ for some $\rho < 1$. Now note that elementary large deviation bounds for the rate-2 PPP (τ_i) show that

$$\mathbb{P}(I_{\text{coal}} > n, T_{\text{coal}} \le n/3), \qquad \mathbb{P}(I_{\text{coal}} \le n, T_{\text{coal}} > n), \qquad \mathbb{P}(\tau_n \le n/3) \quad \text{are all } O_{\exp}(n). \tag{41}$$

Choosing $m = \lfloor an/2 \rfloor$, we find

$$\mathbb{P}(I_{\text{coal}} > n) \le O_{\exp}(n) + \mathbb{P}\big(\text{event (40) fails}\big), \qquad n = 1, 2, 3, \dots,$$

where the $O_{\exp}(n)$ term does not depend on $\|z_0^1 - z_0^2\|$. From the definition of event (40) and the choice of m,

$$\mathbb{P}\big(\text{event (40) fails}\big) \le \mathbb{P}\left(\tau_j \le 2 \log \frac{\|z_0^1 - z_0^2\|}{\alpha}\right) \qquad \text{for } j = \lfloor an/2 \rfloor - 1. \tag{42}$$

From the final term in (41), there exists a constant β such that

$$\mathbb{P}\big(\text{event (40) fails}\big) = O_{\exp}(n) \qquad \text{for } n > \beta \log^+ \|z_0^1 - z_0^2\|.$$

So now,

$$\mathbb{P}(I_{\text{coal}} > n) \le O_{\exp}(n) \qquad \text{for } n > \beta \log^+ \|z_0^1 - z_0^2\|.$$

Because $\mathbb{P}(T_{\text{coal}} > n) \le \mathbb{P}(I_{\text{coal}} > n) + \mathbb{P}(I_{\text{coal}} \le n, T_{\text{coal}} > n)$, we have established Proposition 13.

4. Proof of the main theorems.

4.1. *Notation*. In this section, we use the preceding bounds to prove Theorems 1 and 2. Recall some definitions from Section 1.1. $\mathcal{M}(\mathbb{R}^2)$ denotes the space of finite measures on $\mathbb{R}^2$, equipped with the usual topology of weak convergence. For a particle ξ, the time-t ancestor is denoted $\text{ancestor}(t, \xi)$, and $\text{Descend}(t_1, t_2, \zeta)$ denotes the set of particles born before t_2 whose time-t_1 ancestor is ζ. And for $t_1 \le t_2$ and $\zeta \in \Xi_{\le t_1}$

$$\mu_{t_1,t_2,\zeta} \text{ is the measure } \mu \text{ putting weight } e^{-t_2} \text{ on the position of each particle in } \text{Descend}(t_1, t_2, \zeta). \tag{43}$$

Note that given $\Xi_{\le t_1}$, the "marks" $(\mu_{t_1,t_2,\zeta}, \zeta \in \Xi_{\le t})$ are still *random* elements of the space $\mathcal{M}(\mathbb{R}^2)$, whose distributions depend on all $\Xi_{\le t_1}$ and are dependent as ζ varies.

If we use $\Xi_{\le t}$ to define a translation-invariant marked PPP of the form $\{(\xi, m^+(\xi)), \xi \in \Xi_{\le t}\}$ with nonnegative real marks $m^+(\xi)$, then there is a spatial average rate of mark values, which we will write as

$$\text{ave}\big(m^+(\xi) : \xi \in \Xi_{\le t}\big)$$

defined as the value of a such that

$$\mathbb{E} \sum_{\xi \in \Xi_{\le t}, z_\xi \in B} m^+(\xi) = a \times \text{area}(B), \qquad B \subset \mathbb{R}^2.$$

For instance, we have, for $t_1 < t_2$,

$$\text{ave}(|\mu_{t_1,t_2,\xi}| : \xi \in \Xi_{\le t_1}) = 1, \tag{44}$$

where $|\mu|$ denotes the total mass of μ.

The self-similarity property of the underlying space–time PPP Ξ allows us to write down exact self-similarity properties for our marked point processes. In particular, the action of the scaling map $z \to e^{-t/2} z$ on $\mathbb{R}^2$, applied to the distribution of $\{(z_\xi, \mu_{t_1,t_2,\xi}), \xi \in \Xi_{\le t_1}\}$, gives a distribution which coincides with the distribution obtained from $\{(z_\xi, \mu_{t_1+t,t_2+t,\xi}), \xi \in \Xi_{\le t_1+t}\}$ under the action of rescaling weights $\mu \to e^{-t}\mu$. These self-similarity properties allow us to take previous results, which were stated in the context of time decreasing from 0 to $-t$, and rewrite them in the context of time decreasing from t to 0 and in the notation above. These rewritten results and simple consequences are recorded as Corollaries 21–23 below.

As a first example, Proposition 4 shows that for $t > 0$,

$$\text{ave}(\|z_{\text{ancestor}(-t,\xi)} - z_\xi\| : \xi \in \Xi_{\le 0}) \le K e^{t/2},$$

where $K = \int_0^\infty G(r)\, dr < \infty$. Using self-similarity, this implies the following.

COROLLARY 21. *For* $0 \le t_0 \le t$,

$$\text{ave}(\|z_{\text{ancestor}(t_0,\xi)} - z_\xi\| : \xi \in \Xi_{\le t}) \le K e^t e^{-t_0/2}.$$

Next, the fact that a set of cardinality k contains $k(k-1)$ distinct ordered pairs gives the first identity below, and the second follows from self-similarity. For $t > 0$,

$$\begin{aligned} &\text{ave}(|\mu_{0,t,\zeta}|(|\mu_{0,t,\zeta}| - e^{-t}) : \zeta \in \Xi_{\le 0}) \\ &\quad = e^{-2t}\, \text{ave}\Bigg(\sum_{\xi_2 \in \Xi_{\le t}} 1_{\{\text{ancestor}(0,\xi_2) = \text{ancestor}(0,\xi_1)\}} : \xi_1 \in \Xi_{\le t}\Bigg) \\ &\quad = e^{-t}\, \text{ave}\Bigg(\sum_{\xi_2 \in \Xi_{\le 0}} 1_{\{\text{ancestor}(-t,\xi_2) = \text{ancestor}(-t,\xi_1)\}} : \xi_1 \in \Xi_{\le 0}\Bigg) \\ &\quad = e^{-t} \int_{\square} \int_{\mathbb{R}^2} q(t; z_1, z_2)\, dz_2\, dz_1 = e^{-t} \int_{\mathbb{R}^2} q(t; \mathbf{0}, z)\, dz, \end{aligned} \tag{45}$$

where $\square$ denotes the unit square and $q(t; z_1, z_2)$ is the probability, given that $\Xi_{\le 0}$ has particles at z_1 and z_2, that they have the same ancestor at time $-t$. In the notation of Proposition 4, we have (using the triangle inequality) $q(t;\mathbf{0},z) \leq 2G_t(e^{-t/2}\|z\|/2)$. So from the conclusion of Proposition 4, we have

$$e^{-t}\int_{\mathbb{R}^2} q(t;\mathbf{0},z)\,dz \leq 2e^{-t}\int_{\mathbb{R}^2} G\bigl(e^{-t/2}\|z\|/2\bigr)\,dz = 2\int_{\mathbb{R}^2} G\bigl(\|z\|/2\bigr)\,dz < \infty.$$

Using (44) for the $-e^t$ term in (45), we have established the following.

COROLLARY 22. $\sup_{t>0} \operatorname{ave}(|\mu_{0,t,\zeta}|^2 : \zeta \in \Xi_{\leq 0}) < \infty$.

Finally, self-similarity allows us to rewrite Proposition 13 as follows.

COROLLARY 23. *Let* $s_0 \leq \bar{s} \leq s_1 \leq s_2$. *Let* $p(s_0, s_1, s_2; z_1, z_2)$ *be the probability, given that* $\Xi_{\leq s_1}$ *contains a particle at position* z_1 *and* $\Xi_{\leq s_2}$ *contains a particle at position* z_2, *that these particles have different time-*s_0 *ancestors. Then*

$$p(s_0, s_1, s_2; z_1, z_2) \leq K\exp\bigl(-\rho(\bar{s} - s_0)\bigr) \qquad \text{provided } \|z_1 - z_2\| \leq \sqrt{2}e^{-\bar{s}/2}$$

for the constants K, ρ *in Proposition* 13.

4.2. *Convergence of mark measures.* Here, we will prove the following.

PROPOSITION 24. *There exist* $\mathcal{M}(\mathbb{R}^2)$*-valued marks* $(\mu_{0,\infty,\zeta}, \zeta \in \Xi_{\leq 0})$ *such that* $\operatorname{ave}(|\mu_{0,\infty,\zeta}|, \zeta \in \Xi_{\leq 0}) = 1$ *and for all* $\zeta \in \Xi_{\leq 0}$,

$$\mu_{0,t,\zeta} \to \mu_{0,\infty,\zeta} \qquad \text{in probability as } t \to \infty.$$

The argument is slightly subtle. For large $t_1 < t_2$, a time-t_1 descendant ξ of ζ contributes mass e^{-t_1} to $\mu_{0,t_1,\zeta}$, whereas it contributes a *random* (mean e^{-t_1}) mass to $\mu_{0,t_2,\zeta}$ because the number of time-t_2 descendants of ξ is random. Even though these descendants are spatially close to ξ, this randomness means we can only deduce immediately that the induced non-uniform measure on time-t_2 descendants of ζ is close to $\mu_{0,t_1,\zeta}$; this is not the uniform measure $\mu_{0,t_2,\zeta}$. To handle this issue, we first prove convergence of total masses.

PROPOSITION 25. *There exist real-valued marks* $(m_{0,\zeta}, \zeta \in \Xi_{\leq 0})$ *such that* $\operatorname{ave}(m_{0,\zeta} : \zeta \in \Xi_{\leq 0}) = 1$ *and for all* $\zeta \in \Xi_{\leq 0}$,

$$|\mu_{0,t,\zeta}| \to m_{0,\zeta} \qquad \text{in probability as } t \to \infty.$$

PROOF. Write $\square$ for the unit square and $\square(t)$ for the scaled square of area e^{-t}. For $0 < t_0 < t_1 < t_2$ and $\varepsilon > 0$, write $A(t_0, t_1, t_2, \varepsilon)$ for the event:

> there exist at least $(1-\varepsilon)e^{t_1 - t_0}$ particles of $\Xi_{\leq t_1}$ in $\square(t_0)$, and at least $(1-\varepsilon)e^{t_2 - t_0}$ particles of $\Xi_{\leq t_2}$ in $\square(t_0)$, all with the same time-0 ancestor.

Define

$$\rho(t_0,\varepsilon) = \liminf_{t_1\to\infty} \liminf_{t_2\to\infty} \mathbb{P}\big(A(t_0,t_1,t_2,\varepsilon)\big). \tag{46}$$

We will show the following.

LEMMA 26. $\lim_{t_0\to\infty} \rho(t_0,\varepsilon) = 1$ *for each* $\varepsilon > 0$.

Granted that, consider

$$a(t_1,t_2) := \mathrm{ave}\big(\min\big(|\mu_{0,t_1,\zeta}|, |\mu_{0,t_2,\zeta}|\big), \zeta \in \Xi_{\le 0}\big) \le 1.$$

By averaging over area-t_0 squares in $\mathbb{R}^2$,

$$a(t_1,t_2) \ge (1-\varepsilon)\mathbb{P}\big(A(t_0,t_1,t_2,\varepsilon)\big).$$

So Lemma 26 implies

$$\liminf_{t_1\to\infty} \liminf_{t_2\to\infty} a(t_1,t_2) = 1. \tag{47}$$

By the triangle inequality and (44),

$$\mathrm{ave}\big(\big||\mu_{0,t_1,\zeta}| - |\mu_{0,t_2,\zeta}|\big| : \zeta \in \Xi_{\le 0}\big) \le 2\big(1 - a(t_1,t_2)\big).$$

Then (47) and the Cauchy criterion imply there exist limits $m_{0,\zeta}$ for which

$$\lim_{t\to\infty} \mathrm{ave}\big(\big||\mu_{0,t,\zeta}| - m_{0,\zeta}\big| : \zeta \in \Xi_{\le 0}\big) = 0.$$

Finally, Corollary 22 provides the "uniform integrability" condition needed to pass (44) to the limit to obtain $\mathrm{ave}(m_{0,\zeta} : \zeta \in \Xi_{\le 0}) = 1$. This establishes Proposition 25. $\square$

PROOF OF LEMMA 26. Write $\Xi_{\le t}^{\square(t_0)}$ for the restriction of $\Xi_{\le t}$ to particles within $\square(t_0)$. We can upper bound the mean number of pairs (ξ_1,ξ_2) in $\square(t_0)$ with $\xi_i \in \Xi_{\le t_i}$ and with different time-0 ancestors, as follows. Write

$$M(t_0,t_1,t_2) := \sum_{\xi_1 \in \Xi_{\le t_1}^{\square(t_0)}} \sum_{\xi_2 \in \Xi_{\le t_2}^{\square(t_0)}} 1_{\{\mathrm{ancestor}(0,\xi_1) \ne \mathrm{ancestor}(0,\xi_2)\}}.$$

Then

$$\mathbb{E}M(t_0,t_1,t_2) = e^{t_1}e^{t_2} \int_{\square(t_0)} \int_{\square(t_0)} p(0,t_1,t_2,z_1,z_2)\, dz_1\, dz_2,$$

where $p(0,t_1,t_2,z_1,z_2)$ is the probability, given that $\Xi_{\le t_1}$ has a particle at z_1 and $\Xi_{\le t_2}$ has a particle at z_2, that these two particles have different time-0 ancestors. But Corollary 23 shows that when z_1 and z_2 are in $\square(t_0)$ we have $p(0,t_1,t_2,z_1,z_2) \le K\exp(-\rho t_0)$. So

$$\mathbb{E}M(t_0,t_1,t_2) \le e^{t_1+t_2-2t_0} K \exp(-\rho t_0). \tag{48}$$

We now quote an elementary lemma.

LEMMA 27. *Let $I \subset J$ be finite sets, let $\sim$ be an equivalence relation on J and let B be a maximal-cardinality set in the corresponding partition of J. Let*

$$\rho = \frac{|\{(i,j) \in I \times (J \setminus I) : i \nsim j\}|}{|I| \cdot |J \setminus I|}.$$

Then $|B \cap I| \geq (1-\rho)|I|$ and $|B \cap (J \setminus I)| \geq (1-\rho)|J \setminus I|$.

We will apply the lemma with I and J being $\Xi_{\leq t_1}^{\square(t_0)}$ and $\Xi_{\leq t_2}^{\square(t_0)}$, so that $|I|$ and $|J|$ have Poisson distributions with means $e^{t_1-t_0}$ and $e^{t_2-t_0}$, and to the equivalence relation "same time-0 ancestor." Choose $\delta > 0$ such that $(1-\delta)(1-\delta(1-\delta)^{-2}) < \varepsilon$. On the event

$$(49)\qquad \begin{aligned} &|I| \geq (1-\delta)e^{t_1-t_0} \quad \text{and} \quad |J| \geq (1-\delta)e^{t_2-t_0} \quad \text{and} \\ &M(t_0,t_1,t_2) \leq \delta e^{t_1+t_2-2t_0}, \end{aligned}$$

Lemma 27 implies that event $A(t_0,t_1,t_2,\varepsilon)$ holds. The first two events in (49) have probabilities $\to 1$ as $t_1, t_2 \to \infty$, and so by (48) and Markov's inequality the limit $\rho(t_0,\varepsilon)$ at (46) satisfies

$$\rho(t_0,\varepsilon) \geq 1 - \delta^{-1} K \exp(-\rho t_0),$$

establishing Lemma 26. $\square$

Proof of Proposition 24. Take $0 < t_0 < t$. By self-similarity, Proposition 25 remains true if time 0 is replaced by an arbitrary time t_0: there exist real-valued marks $(m_{t_0,\xi}, \xi \in \Xi_{\leq t_0})$ such that for all $\xi \in \Xi_{\leq t_0}$

$$(50)\qquad |\mu_{t_0,t,\xi}| \to m_{t_0,\xi} \qquad \text{in probability as } t \to \infty.$$

For $\zeta \in \Xi_{\leq 0}$, define $\nu_{0,t_0,\zeta}$ to be the measure that puts weight $m_{t_0,\xi}$ on the position z_ξ of each particle $\xi \in \text{Descend}(0,t_0,\zeta)$. And define $\nu_{0,t_0,t,\zeta}$ to be the measure that puts weight $|\mu_{t_0,t,\xi}|$ on the position z_ξ of each particle $\xi \in \text{Descend}(0,t_0,\zeta)$. We will need to show that, for large t_0, the measures $\nu_{0,t_0,t,\zeta}$ and $\mu_{0,t,\zeta}$ are close.

We exploit the dual bounded Lipschitz metric on $\mathcal{M}(\mathbb{R}^2)$:

$$d(\nu,\nu') = \sup\left\{\left|\int f\,d\nu - \int f\,d\nu'\right| : \|f\|_{\mathrm{BL}} \leq 1\right\},$$

$$\|f\|_{\mathrm{BL}} := \max\left(\sup_z |f(z)|, \sup_{z_1 \neq z_2} \frac{|f(z_2) - f(z_1)|}{\|z_1 - z_2\|}\right).$$

This metric has the property

$$(51)\qquad d\left(c\sum_i \delta_{z_i}, c\sum_i \delta_{z_i'}\right) \leq c\sum_i \|z_i - z_i'\|.$$

Consider $0 < t_1 < t_2 < t$. The relationship between $\nu_{0,t_2,t,\zeta}$ and $\nu_{0,t_1,t,\zeta}$ is that for each $\xi \in \text{Descend}(0,t,\zeta)$ the weight e^{-t} moved from the position of ancestor(t_2, ξ) to the position of ancestor(t_1, ξ). Taking spatial averages and using (51), we find

$$\begin{aligned}
&\operatorname{ave}\bigl(d(\nu_{0,t_1,t,\zeta}, \nu_{0,t_2,t,\zeta}), \zeta \in \Xi_{\le 0}\bigr) \\
&\qquad \le e^{-t} \operatorname{ave}\bigl(\|z_{\mathrm{ancestor}(t_1,\xi)} - z_{\mathrm{ancestor}(t_2,\xi)}\|, \xi \in \Xi_{\le t}\bigr) \\
&\qquad \le 2Ke^{-t_1/2} \qquad \text{by Corollary 21.}
\end{aligned}$$

This and (50) are sufficient to imply that the ν's have a limit: for all $\zeta \in \Xi_{\le 0}$

$$\nu_{0,t,\zeta} \to \mu_{0,\infty,\zeta} \text{ (say), in probability as } t \to \infty. \tag{52}$$

Now by (50), we can write the definition of $\nu_{0,t_0,\zeta}$ as

$$\nu_{0,t_0,\zeta} = \sum \Bigl\{\Bigl(\lim_{u\to\infty} |\mu_{t_0,u,\xi}|\Bigr)\delta_{z_\xi} : \xi \in \mathrm{Descend}(0, t_0, \zeta)\Bigr\},$$

whereas (by definition) for $t > t_0$

$$\mu_{0,t,\zeta} = \sum \{\mu_{t_0,t,\xi} : \xi \in \mathrm{Descend}(0, t_0, \zeta)\}.$$

So now we have

$$\begin{aligned}
&d(\mu_{0,t,\zeta}, \mu_{0,\infty,\zeta}) \\
&\quad \le d(\nu_{0,t_0,\zeta}, \mu_{0,\infty,\zeta}) + d(\mu_{0,t,\zeta}, \nu_{0,t_0,\zeta}) \\
&\quad \le d(\nu_{0,t_0,\zeta}, \mu_{0,\infty,\zeta}) \\
&\qquad + \sum \Bigl\{d\Bigl(\lim_{u\to\infty} |\mu_{t_0,u,\xi}|\delta_{z_\xi}, \mu_{t_0,t,\xi}\Bigr) : \xi \in \mathrm{Descend}(0, t_0, \zeta)\Bigr\} \\
&\quad \le d(\nu_{0,t_0,\zeta}, \mu_{0,\infty,\zeta}) \\
&\qquad + \sum \Bigl\{\Bigl|\lim_{u\to\infty} |\mu_{t_0,u,\xi}| - |\mu_{t_0,t,\xi}|\Bigr| : \xi \in \mathrm{Descend}(0, t_0, \zeta)\Bigr\} \\
&\qquad + \sum \bigl\{d\bigl(\mu_{t_0,t,\xi}, |\mu_{t_0,t,\xi}|\delta_{z_\xi}\bigr) : \xi \in \mathrm{Descend}(0, t_0, \zeta)\bigr\}.
\end{aligned}$$

Taking the spatial average over $\zeta \in \Xi_{\le 0}$ of sums over all time-t_0 descendants of ζ is the same as taking the spatial average over all time-t_0 particles. So taking averages in the inequality above gives

$$\operatorname{ave}\bigl(d(\mu_{0,t,\zeta}, \mu_{0,\infty,\zeta}) : \zeta \in \Xi_{\le 0}\bigr) \le b_1(t_0) + b_2(t_0, t) + b_3(t_0, t), \tag{53}$$

where

$$\begin{aligned}
b_1(t_0) &= \operatorname{ave}\bigl(d(\nu_{0,t_0,\zeta}, \mu_{0,\infty,\zeta}) : \zeta \in \Xi_{\le 0}\bigr), \\
b_2(t_0, t) &= \operatorname{ave}\Bigl(\Bigl|\lim_{u\to\infty} |\mu_{t_0,u,\xi}| - |\mu_{t_0,t,\xi}|\Bigr| : \xi \in \Xi_{\le t_0}\Bigr), \\
b_3(t_0, t) &= \operatorname{ave}\bigl(d\bigl(\mu_{t_0,t,\xi}, |\mu_{t_0,t,\xi}|\delta_{z_\xi}\bigr) : \xi \in \Xi_{\le t_0}\bigr).
\end{aligned}$$

To prove Proposition 24, it is enough to prove

$$\operatorname{ave}\bigl(d(\mu_{0,t,\zeta}, \mu_{0,\infty,\zeta}) : \zeta \in \Xi_{\le 0}\bigr) \to 0 \qquad \text{as } t \to \infty. \tag{54}$$

We know $b_1(t_0) \to 0$ as $t_0 \to \infty$ by (52). And $b_2(0,t) \to 0$ as $t \to \infty$ by Proposition 25, and then by self-similarity $b_2(t_0,t) \to 0$ as $t \to \infty$ for all t_0. Finally,

$$\begin{aligned} d(\mu_{t_0,t,\xi}, |\mu_{t_0,t,\xi}|\delta_{z_\xi}) &\le \int \|z_\xi - z\| \mu_{t_0,t,\xi}(dz) \\ &= e^{-t} \sum \{ \|z_{\mathrm{ancestor}(-t_0,\chi)} - z_\chi\| : \chi \in \mathrm{Descend}(t_0,t,\xi) \} \end{aligned}$$

and so

$$\begin{aligned} b_3(t_0,t) &\le e^{-t} \operatorname{ave}(\|z_{\mathrm{ancestor}(-t_0,\chi)} - z_\chi\| : \chi \in \Xi_{\le t}) \\ &\le K e^{-t_0/2} \qquad \text{by Corollary 21.} \end{aligned}$$

Now taking limits in the inequality (53) establishes (54) and then Proposition 24.

4.3. *The random partition.* We will now show that a limit random measure $\mu_{0,\infty,\xi}$ in Proposition 24 is in fact Lebesgue measure Λ restricted to some random set. The fact that the $t \to \infty$ limit normalized empirical measure on $\mathcal{Z}_t$ is Λ implies that

$$\sum \{\mu_{0,\infty,\zeta} : \zeta \in \Xi_{\le 0}\} = \Lambda \qquad \text{a.s.}$$

So the random measures $\mu_{0,\infty,\zeta}$ have random densities $f_\zeta(z)$, $z \in \mathbb{R}^2$ such that

$$\sum \{ f_\zeta(z) : \zeta \in \Xi_{\le 0} \} = 1 \qquad \forall z \text{ a.s.}$$

As $t \to \infty$, we have

$$\begin{aligned} \mathbb{E} \Big\{ \sum_{z_{\xi_1} \in \square} \sum_{z_{\xi_2} \in \square} e^{-2t} 1_{\{\|z_{\xi_1} - z_{\xi_2}\| \le \delta\}} : \xi_1 \in \Xi_{\le t}, \xi_2 \in \Xi_{\le t} \Big\} \\ \to \int_\square \int_\square 1_{\{\|z_1 - z_2\| \le \delta\}} \, dz_1 \, dz_2. \end{aligned}$$

Now consider whether a pair (ξ_1, ξ_2) have different time-0 ancestors; precisely, consider

$$\begin{aligned} \lim_{t \to \infty} \mathbb{E} \Big\{ \sum_{z_{\xi_1} \in \square} \sum_{z_{\xi_2} \in \square} e^{-2t} 1_{\{\|z_{\xi_1} - z_{\xi_2}\| \le \delta\}} 1_{\{\mathrm{ancestor}(0,\xi_1) \ne \mathrm{ancestor}(0,\xi_2)\}} : \\ \xi_1 \in \Xi_{\le t}, \xi_2 \in \Xi_{\le t} \Big\}. \end{aligned} \tag{55}$$

From the fact $\mu_{0,t,\zeta} \to \mu_{0,\infty,\zeta}$, the limit in (55) equals

$$\mathbb{E} \int_\square \int_\square 1_{\{\|z_1 - z_2\| \le \delta\}} \Big(1 - \sum_{\zeta \in \Xi_{\le 0}} f_\zeta(z_1) f_\zeta(z_2) \Big) \, dz_1 \, dz_2. \tag{56}$$

But consider the probability $p(0,t,t,z_1,z_2)$, given that $\Xi_{\le t}$ has particles at z_1 and z_2, that they have different time-0 ancestors. By Corollary 23 with $\bar{s}$ defined by $\delta/\sqrt{2} = \exp(-\bar{s}/2)$,

$$\text{if } \|z_2 - z_1\| \le \delta \qquad \text{then } p(0,t,t,z_1,z_2) \le K\delta^{2\rho} \text{ for } t \ge \bar{s}.$$

So the limit in (55) also equals

$$\begin{aligned} &\int_\square \int_\square 1_{\{\|z_1-z_2\|\le\delta\}} \lim_t p(0,t,t,z_1,z_2)\,dz_1\,dz_2 \\ &\qquad \le K\delta^{2\rho} \int_\square \int_\square 1_{\{\|z_1-z_2\|\le\delta\}}\,dz_1\,dz_2. \end{aligned} \tag{57}$$

For probability distributions (a_i) and (b_i), we have $1 - \sum_i a_i b_i \ge 1 - \max_i a_i$. Applying this to (56) and using inequality (57),

$$\frac{\mathbb{E}\int_\square \int_\square 1_{\{\|z_1-z_2\|\le\delta\}}(1 - \max_{\zeta\in\Xi_{\le 0}} f_\zeta(z_1))\,dz_1\,dz_2}{\int_\square \int_\square 1_{\{\|z_1-z_2\|\le\delta\}}\,dz_1\,dz_2} \le K\delta^{2\rho}.$$

Letting $\delta \downarrow 0$ we deduce that a.s.

$$\max_{\zeta\in\Xi_{\le 0}} f_\zeta(z) = 1 \qquad \text{a.e.}$$

So defining

$$A(0,\zeta) = \{z : f_\zeta(z) = 1\}$$

and modifying on null sets, the random sets $\{A(0,\zeta) : \zeta \in \Xi_{\le 0}\}$ form a partition of $\mathbb{R}^2$, and $\mu_{0,\infty,\zeta}$ is Lebesgue measure restricted to $A(0,\zeta)$. So writing Λ_A for Lebesgue measure restricted to A, we can rewrite Proposition 24 as follows, using self-similarity to extend from the time-0 case to the general time t case.

PROPOSITION 28. *For each $-\infty < t < \infty$, there exists a random partition $\{A(t,\zeta) : \zeta \in \Xi_{\le t}\}$ of $\mathbb{R}^2$ into measurable sets such that for all $\zeta \in \Xi_{\le t}$*

$$\mu_{t,u,\zeta} \to \Lambda_{A(t,\zeta)} \qquad \textit{in probability as } u \to \infty.$$

4.4. *Completing the proofs.* Proposition 28 is essentially enough to prove Theorems 2 and 1. As noted in the Introduction, for fixed t we can regard

$$\mathbf{Z}^{(t)} = \{(z_\xi, A(t,\xi)) : \xi \in \Xi_{\le t}\}$$

as a marked point process. The fact that the evolution of the coloring process after time t, given $\mathcal{Z}_t$, does not depend on the arrival times of the particles in $\Xi_{\le t}$, means that $\mathbf{Z}^{(t)}$ is measurable with respect to the time-reversed filtration $\overleftarrow{\mathcal{F}}_t$ at (5). The statement in Theorem 1 was that the process $\mathbf{Z}^{(t)}$ evolves in reversed time according to the rule:

during $[t, t-dt]$, for each $\xi \in \Xi_{\le t}$ with probability dt delete ξ [that is, remove the entry $(z_\xi, A(t, \xi))$]; for each deleted particle ξ, let ζ be the nearest other particle, and set $A(t-dt, \zeta) = A(t, \zeta) \cup A(t, \xi)$.

To see how this arises, fix large T and for $-\infty < t \le T$ consider $\{(z_\xi, \mu_{t,T,\xi}) : \xi \in \Xi_{\le t}\}$ as a marked point process. From Lemma 5 (the thinning property of the PPP), in reversed time t this evolves precisely as a "coalescing measures process":

during $[t, t-dt]$, for each $\xi \in \Xi_{\le t}$ delete ξ [i.e., remove the entry $(z_\xi, \mu_{t,T,\xi})$] with probability dt; for each deleted particle ξ, let ζ be the nearest other particle, and set $\mu_{t-dt,T,\zeta} = \mu_{t,T,\zeta} + \mu_{t,T,\xi}$.

Taking the $T \to \infty$ limit given in Proposition 28, we obtain the former rule for the dynamics of $\mathbf{Z}^{(t)}$.

The other assertions of Theorem 1 hold by translation-invariance and self-similarity of the underlying space–time PPP Ξ.

5. Discussion.

5.1. *In what sense is this a tree process?* We have used the language of ancestors and descendants, but otherwise have not really exploited the implicit tree structure of the colored point process construction. If we draw the process as a random tree in the plane, with edges drawn as line segments, it is clear from Figure 1 that edges sometimes cross, so we do not get a "tree" in the usual sense. This suggests that, in the opening "k colors in the unit square" model, in the limit partition into k colored regions, these regions are not necessarily connected. Figure 7 illustrates how this could happen. Simulations strongly indicate that in fact a typical region is not connected but that only a very small proportion of its area is outside its largest connected component.

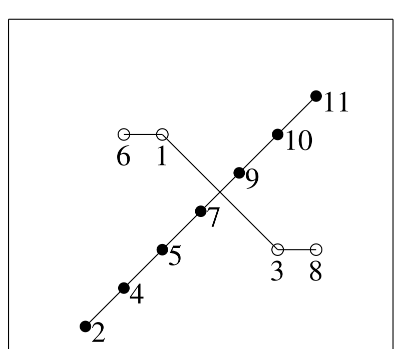

FIG. 7. *A possible realization of the tree in the unit square on the first $n = 11$ arriving points. The edge from 1 to 2 is omitted, to show the $k = 2$ subtrees associated with the first two vertices. At this stage, the unit square is Voronoi-partitioned into 2 components according to whether the nearest vertex is $\circ$ or $\bullet$, and the $\circ$ component is not connected. We expect this disconnection to persist in the $n \to \infty$ limit.*

5.2. *Other models of coalescing partitions*. There has been very little study of partition-valued processes in the plane which evolve by merging of adjacent components. One such process can be obtained by thinning a Poisson *line* process, but we are thinking of pairwise mergers. A well-studied implicit example is provided by bond percolation. As illustrated in Figure 8, to a percolation cluster of "open" edges (A) one can associate (this is *planar duality*) the region consisting of the union of the unit squares centered at the the vertices in the cluster (B), and then delete the open edges (C) and vertices to obtain a partition of the plane (D). The length of the boundary between two adjacent regions in this partition equals the number of "closed" edges between the original percolation clusters. So in the bond percolation model where edges become open at Exponential(1) times, the associated partition-valued process is such that the merger rate of adjacent regions equals the length of their common boundary. For this model, classical percolation theory [10] implies that infinite regions appear after time $\log 2$. More general models in which two adjacent components merge into one at some stochastic rate

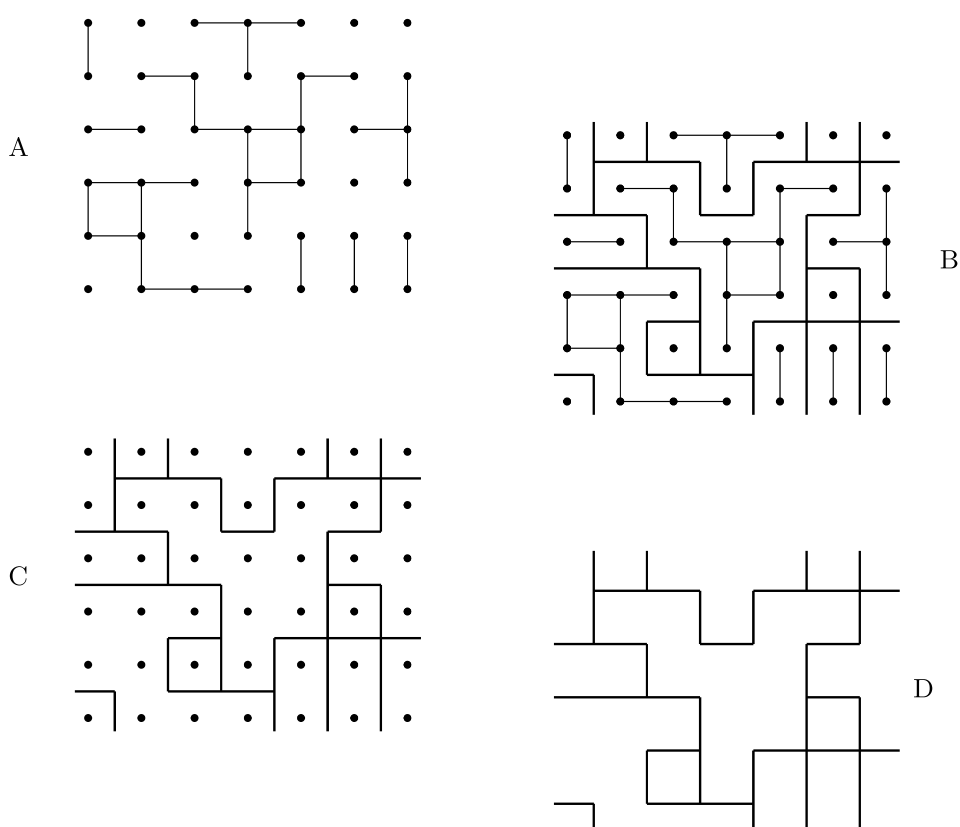

FIG. 8. *Bond percolation clusters as a partition of* $\mathbb{R}^2$.

determined by their geometry are discussed in [4], where it is conjectured that if large components do not grow too quickly (relative to small components), then there should be some self-similar asymptotics, but no such result is proved. The coalescing partitions process in this paper is perhaps the only known self-similar Markovian process of pairwise merging partitions of $\mathbb{R}^2$. In one dimension, the thinning process of Poisson points defines a self-similar process of merging adjacent intervals, which has an interpretation as intermediate-size asymptotics in the Kingman coalescent ([3] Section 3.1).

5.3. *Heuristic arguments*. Arguments in this Section 5.3 are heuristics, only parts of which seem easily formalized. We conjectured at (28) that the tail behavior of the "meeting distance" random variable $\|Z_{\mathrm{coal}}\|$ is of the form

$$\mathbb{P}\big(\|Z_{\mathrm{coal}}\| > r\big) \asymp r^{-\gamma} \qquad \text{as } r \to \infty \tag{58}$$

for some exponent γ. This is heuristically related to the issue of fractal dimension of the boundaries of the regions $(A(0,\xi), \xi \in \Xi_{\leq 0})$, as follows. Consider the boundaries within the unit square. Saying this has fractal dimension d is saying that for small $x > 0$ we need order x^{-d} radius-x discs to cover these boundaries. Consider a uniform random point z_1 in the square and another random point z_2 uniform on disc(z_1, x). The chance that z_1 and z_2 are in different components is the same order as the chance they are in the same covering disc, which chance is order x^{2-d}. But the former chance is the same order as the chance that the meeting distance M_x between their lines of descent is at least 1, that is, $\mathbb{P}(M_x > 1)$. So we expect $\mathbb{P}(M_x > 1) \asymp x^{2-d}$ as $x \downarrow 0$. Now by self-similarity $M_x =_d xM_1$. So setting $r = 1/x$,

$$\mathbb{P}(M_1 > r) \asymp r^{-(2-d)} \tag{59}$$

and we heuristically identify the fractal dimension as

$$d = 2 - \gamma.$$

5.3.1. *Heuristic* 1: *The boundary has fractal dimension* 1. For large t_0, consider the Voronoi regions associated with the different sets of particles Descend$(0, t_0, \zeta)$ as ζ varies. The particles in $\Xi_{\leq t_0}$ are separated by distance of order $e^{-t_0/2}$. Consider three points (z_1, z_2, z_3) on the boundary at time t_0 at distances of (say) $5e^{-t_0/2}$ apart. As t increases, particles arriving nearby move the boundary near these three points. The first such move is over a distance of order $e^{-t_0/2}$ and subsequent moves decrease geometrically. So in the $t \to \infty$ limit, the positions of the boundaries near (z_1, z_2, z_3) should become $(z_1 + D_1e^{-t_0/2}, z_2 + D_2e^{-t_0/2}, z_3 + D_3e^{-t_0/2})$ for some (D_1, D_2, D_3) with a nonvanishing limit as $t_0 \to \infty$. But this is saying that in the limit partition $(A(0,\zeta), \zeta \in \Xi_{\leq 0})$, on every scale $\sigma = \|y_1 - y_3\|$, for y_1 and y_3 on the boundary, the distance from the midpoint $(y_1 + y_3)/2$ to the boundary is of the form $D\sigma$ for some random $D > 0$. This is a hallmark of "fractal dimension = 1."

5.3.2. *Heuristic* 2: *The boundary has fractal dimension* 1. The genealogical tree defines a "line of descent" for each particle in Ξ; these particle positions are dense in $\mathbb{R}^2$, so let us suppose that in the continuum limit there is such a "line of descent" from almost all points z of $\mathbb{R}^2$ to infinity. Draw the tree via line segments in $\mathbb{R}^2$. Consider a point $(x, 0)$ on the x-axis. The route from there to infinity first crosses the $y = 1$ line at some random point $(c(x), 1)$. Consider the random set $\mathcal{C}$ of all such values $c(x)$ as x varies. This is stationary (translation-invariant) and so has some intensity γ, which cannot be zero; moreover, we expect $\gamma < \infty$ because the intensity of line segments of length $> a$ is finite for each $a > 0$. Then suppose that for each $c \in \mathcal{C}$, the sets of originating points $\{x : c(x) = c\}$ form some[8] interval $(x_-(c), x_+(c))$.

Next consider, for $z > 0$, the random quantity defined as

> the infimum of $y > 0$ such that the routes from $(x_1, 0)$ to infinity and from $(x_1 + z, 0)$ to infinity first hit the line $\{(x, y) : -\infty < x < \infty\}$ at the same point.

This has a distribution, say D_z, which does not depend on x_1. By considering endpoints of the intervals $(x_-(c), x_+(c))$, we have

$$\gamma = \lim_{\delta \downarrow 0} \mathbb{P}(D_\delta > 1)/\delta.$$

But by scale-invariance, we have $D_\delta =_d \delta D_1$, and so

$$\mathbb{P}(D_1 > d) \sim \gamma/d \qquad \text{as } d \to \infty.$$

But D_1 should have the same tail behavior as M_1 at (59). This suggests the scaling exponent is $\gamma = 1$, and hence the fractal dimension of the boundaries $= 1$.

5.3.3. *Heuristic* 3: *The boundary has fractal dimension* $\neq 1$. The heuristics at (58), (59) are essentially saying that the fractal dimension d is determined via the limit

$$\lim_{r \to \infty} \frac{\mathbb{P}(\|Z_{\text{coal}}\| > 2r)}{\mathbb{P}(\|Z_{\text{coal}}\| > r)} = 2^{d-2}.$$

But the limit is determined by the asymptotics of the coupled EA process, conditioned on not coalescing for a long time. As we saw in in Section 3.1, the dynamics of the coupled EA process involve the complicated geometry of excluded regions, and there seems no reason why that limit should turn out to be exactly $1/2$.

5.3.4. *Regarding Conjecture* 3. One might imagine that Conjecture 3 would follow easily from Proposition 13 via some general result of the form

> If $\{A, A^c\}$ is a partition of the unit square such that $\rho(r) \to 1$ as $r \to 0$, where $\rho(r)$ is the probability that two random points at distance r apart are in the same subset, then (after modifying A on a set of measure zero) the topological boundary of A has measure zero.

[8]The argument is unchanged if instead it is a union of a finite-mean number of intervals.

But this assertion is not true in general, by considering an example of the form $A = \bigcup_i \mathrm{disc}(z_i, r_i)$ for dense (z_i) and $r_i \downarrow 0$ very fast. Proving Conjecture 3 seems to require some new argument.

Acknowledgments. I thank Weijian Han for Figures 1 and 2, Andrew Wade for sharing his results and a referee for helpful suggestions. A dynamic simulation by Shirley Danlei Zhu can be seen at https://www.stat.berkeley.edu/~aldous/Research/Ugrad/40to2_Coloring_fast.gif.

REFERENCES

[1] ALDOUS, D. (2014). Scale-invariant random spatial networks. *Electron. J. Probab.* **19** no. 15, 41. MR3164768

[2] ALDOUS, D. and GANESAN, K. (2013). True scale-invariant random spatial networks. *Proc. Natl. Acad. Sci. USA* **110** 8782–8785. MR3082274

[3] ALDOUS, D. J. (1999). Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli* **5** 3–48. MR1673235

[4] ALDOUS, D. J., ONG, J. R. and ZHOU, W. (2010). Empires and percolation: Stochastic merging of adjacent regions. *J. Phys. A* **43** 025001, 10. MR2575204

[5] ASMUSSEN, S. (2003). *Applied Probability and Queues: Stochastic Modelling and Applied Probability*, 2nd ed. *Applications of Mathematics (New York)* **51**. Springer, New York. MR1978607

[6] BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes*. Cambridge Univ. Press, Cambridge. MR2253162

[7] CHIU, S. N., STOYAN, D., KENDALL, W. S. and MECKE, J. (2013). *Stochastic Geometry and Its Applications*, 3rd ed. Wiley, Chichester. MR3236788

[8] COWAN, R. (2010). New classes of random tessellations arising from iterative division of cells. *Adv. in Appl. Probab.* **42** 26–47. MR2666917

[9] GEORGII, H.-O., SCHREIBER, T. and THÄLE, C. (2015). Branching random tessellations with interaction: A thermodynamic view. *Ann. Probab.* **43** 1892–1943. MR3353818

[10] GRIMMETT, G. (1989). *Percolation*. Springer, New York. MR0995460

[11] JORDAN, J. and WADE, A. (2012). Evolution of random territories via on-line nearest-neighbours. Unpublished notes.

[12] LITTLEWOOD, J. E. (1986). *Littlewood's Miscellany*. Cambridge Univ. Press, Cambridge. MR0872858

[13] PENROSE, M. D. and WADE, A. R. (2010). Random directed and on-line networks. In *New Perspectives in Stochastic Geometry* 248–274. Oxford Univ. Press, Oxford. MR2654680

[14] SATO, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Univ. Press, Cambridge. MR1739520

[15] SCHREIBER, T. and THÄLE, C. (2013). Shape-driven nested Markov tessellations. *Stochastics* **85** 510–531. MR3176472

DEPARTMENT OF STATISTICS
367 EVANS HALL # 3860
UNIVERSITY OF CALIFORNIA, BERKELEY
BERKELEY, CALIFORNIA 94720
USA
E-MAIL: aldous@stat.berkeley.edu