

THE SEGREGATED Λ -COALESCENT

BY NIC FREEMAN¹

Oxford University

We construct an extension of the Λ -coalescent to a spatial continuum and analyse its behaviour. Like the Λ -coalescent, the individuals in our model can be separated into (i) a dust component and (ii) large blocks of coalesced individuals. We identify a five phase system, where our phases are defined according to changes in the qualitative behaviour of the dust and large blocks. We completely classify the phase behaviour, including necessary and sufficient conditions for the model to come down from infinity.

We believe that two of our phases are new to Λ -coalescent theory and directly reflect the incorporation of space into our model. Firstly, our semicritical phase sees a null but nonempty set of dust. In this phase the dust becomes a random fractal, of a type which is closely related to iterated function systems. Secondly, our model has a critical phase in which the coalescent comes down from infinity gradually during a bounded, deterministic time interval.

1. Introduction. Coalescent processes are stochastic models in which a collection of particles start out separated and come together over time. Modern coalescent theory began with the coalescent of [Kingman \(1982\)](#), which was introduced to describe the family trees of individuals sampled from large haploid populations. Kingman's coalescent was generalized, independently but in the same spirit, by [Donnelly and Kurtz \(1999\)](#), [Pitman \(1999\)](#) and [Sagitov \(1999\)](#). The resulting process became known as the Λ -coalescent.

We begin with a heuristic description of the Λ -coalescent. At time 0 the Λ -coalescent starts with a countable infinity of particles, with each particle representing an individual from the population. It is usual to label these initial particles with elements of \mathbb{N} . Then, at a countable set of random times during $(0, \infty)$, a subset of the currently present particles are selected, and these particles come together to form a coalesced block of particles. This coalesced block is thought of as a single new particle and may subsequently be coalesced into even larger blocks of particles.

The Λ -coalescent has been studied intensively over the past decade, and its behaviour is now well understood. See [Berestycki \(2009\)](#) for an introduction to the Λ -coalescent and its connections to other parts of probability.

At any time $t > 0$, we can divide the particles within the Λ -coalescent into two types, firstly, particles that have not been affected by a coalescence event during

Received November 2012; revised May 2013.

¹Supported by an EPSRC Studentship.

MSC2010 subject classifications. Primary 60G99; secondary 60J70, 60J85.

Key words and phrases. Lambda coalescent, coalescent, segregated.

$[0, t]$. These particles are singletons at time t and, collectively, make up the *dust component* of the coalescent. Secondly, we have large blocks of particles that were coalesced together during $[0, t]$. Each such block contains a nontrivial proportion of the countable infinity of initial particles (and is therefore infinite itself).

It is possible for the particles within the Λ -coalescent to come together so fast that the dust vanishes instantaneously after time 0, leaving only finitely many non-singleton blocks. In this case the Λ -coalescent is said to come down from infinity.

The Λ -coalescent is exchangeable, which means that its distribution does not change when the labels of the initial particles are permuted. This implies that the Λ -coalescent is a nonspatial model, since it means that the random forces which cause groups of particles to coalesce do not depend on the labels of the particles involved.

In reality, children begin life close to their parent and only travel so far in a single lifetime. Therefore, it is natural to ask if the geographical space in which the population lives has a noticeable effect on the genealogy of the population. This is believed to be the case; see, for example, [Etheridge \(2008\)](#).

In this article we construct a spatial extension of the Λ -coalescent in which the ancestral lines of individuals are more likely to coalesce if those individuals lived nearby. Our model behaves similarly to the Λ -coalescent but sees additional behaviour, notably an extra phase transition that is directly related to the introduction of space. The corresponding extra phase (known as the critical phase) contains behaviour that we believe is new to Λ -coalescent theory: in this phase our model comes down from infinity gradually over a deterministic, bounded interval of time.

We define our model in Sections 1.1 and 1.2 before stating our main results in Section 1.3. We compare our model and its behaviour to other Λ -coalescent type models in Sections 1.4–1.6. Our results are proved in Sections 2–5 and a brief outline of the proofs can be found in Section 1.7.

NOTATION. All the spaces we consider will be metric spaces, and we equip them with the corresponding topology and Borel σ -field. For sets A and B , we write $A \uplus B$ for the disjoint union of A and B (i.e., $A \uplus B = A \cup B$ with the implication that A and B are disjoint). We write $\bigcup A = \{a; \exists b \in A, a \in b\}$. If A is a finite set, then we write $|A|$ for the cardinality of A , with $|A| = \infty$ if A is infinite. We write $\mathbb{1}\{\mathcal{P}\}$ for the function which is 1 if the property \mathcal{P} holds and 0 if it does not. We set $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

1.1. Segregated spaces. The geographical space of our model, which we call a *segregated space*, is equipped with a tree structure, as follows. This structure will play a central role in the definition of the Segregated Λ -coalescent.

Let $S \in \{2, 3, 4, \dots\}$, and set $S = \{1, 2, \dots, S\}$. Let W_n be the set of words $w = w_1 w_2 \cdots w_n$ of length n with letters $w_i \in S$. Set $W_* = \bigcup_{n=0}^{\infty} W_n$ be the regular S -ary tree, where $W_0 = \{\emptyset\}$ and \emptyset is the empty word.² For each $w = w_1 w_2 \cdots w_n \in$

²We also use the symbol \emptyset for the empty set.

W_n we write $|w| = n$, with $|\emptyset| = 0$. If $n = 0$, then we define $w_1 \cdots w_n = \emptyset$. If $w = w_1 \cdots w_n$ and $i \in S$, then we set $wi = w_1 \cdots w_n i \in W_{n+1}$.

DEFINITION 1.1. Let (K, D_K) be a complete metric space, equipped with a family of nonempty measurable subsets $(K_w)_{w \in W_*}$ and a probability measure λ . We say K is a segregated space if it satisfies:

- ($\mathcal{K}1$) $K = K_\emptyset$ and for all $w \in W_*$, $K_w = \uplus_{i \in S} K_{wi}$.
- ($\mathcal{K}2$) There exists a sequence $(L_n)_{n=0}^\infty \subseteq (0, \infty)$ such that $L_n \rightarrow 0$ and for all $w \in W_*$, $\max\{D_K(x, y); x, y \in K_w\} \leq L_{|w|}$.
- ($\mathcal{K}3$) For all $w \in W_*$, $\lambda(K_w) = S^{-|w|}$.
- ($\mathcal{K}4$) For all $w \in W_*$ there exists $i \in S$ such that $\overline{K_{wi}} \subseteq K_w$.

We say K_w is a complex of K with level $|w|$. If $K_v \subseteq K_w$, then we say K_v is a subcomplex of K_w .

EXAMPLE 1.2. The prototype example of a segregated space is the middle third Cantor set. This is the unique nonempty compact subset \mathcal{K} of $[0, 1]$ which satisfies $F_1(\mathcal{K}) \uplus F_2(\mathcal{K}) = \mathcal{K}$, where $F_1(x) = x/3$ and $F_2(x) = 2/3 + x/3$. We set $K_\emptyset = \mathcal{K}$ and define K_w iteratively by the relation $K_{wi} = F_i(K_w)$. The measure λ is the uniform Bernoulli measure on \mathcal{K} , with $S = 2$ and $\lambda(K_w) = 2^{-|w|}$.

In Example 1.2, \mathcal{K} is a totally disconnected set of Lebesgue measure zero. This is unnatural from the point of view of population modelling, where it is usual to use \mathbb{R}^d as a model for spatial continua.

EXAMPLE 1.3. Let $S = 2$, and set $K = K_\emptyset = [0, 1]$. Note that $[0, 1] = [0, 1/2] \uplus (1/2, 1] = [0, 1/4] \uplus (1/4, 1/2] \uplus (1/2, 3/4] \uplus (3/4, 1] = \dots$. Set $K_1 = [0, 1/2]$, $K_2 = (1/2, 1]$, $K_{11} = [0, 1/4]$, $K_{12} = (1/4, 1/2]$, $K_{21} = (1/2, 3/4]$ and so on. Take λ as Lebesgue measure on $[0, 1]$. Note that this example is easily adapted to higher dimensions and $S \geq 2$.

We will use ($\mathcal{K}1$) so frequently that it would be impractical to reference it on every application. However, we will not use the other conditions without explicitly saying so. The purpose of ($\mathcal{K}2$) is as follows: suppose (w_n) is a sequence in W_* such that $|w_n| \rightarrow \infty$ and $K_{w_{n+1}} \subseteq K_{w_n}$. Then ($\mathcal{K}2$) implies that $\bigcap_n K_{w_n}$ is either empty or equal to a single point. Condition ($\mathcal{K}4$) is designed to prevent pathological examples of the sample space and will be discussed further in Section 4.

Initially, each point of K will be the location of precisely one individual. In view of ($\mathcal{K}3$), we think of λ as a uniform measure on K . The measure λ is important to us because it tells us whether a nonempty set of individuals (i.e., a subset of K) comprises a null or positive proportion of the total population.

LEMMA 1.4. For all $w \in W_*$, $\lambda(K_w) > 0$. For all $x \in K$, $\lambda(\{x\}) = 0$.

PROOF. The first statement follows trivially from $(\mathcal{K}3)$. If $x \in K$, then by $(\mathcal{K}1)$ for all $n \in \mathbb{N}$ we have $x \in K_w$ for some $w \in W_n$. Since $\cup_{w \in W_n} K_w = K$, we have $1 = \lambda(K) = \mathcal{S}^n \lambda(K_w) \geq \mathcal{S}^n \lambda(\{x\})$. Since $n \in \mathbb{N}$ was arbitrary, we must have $\lambda(\{x\}) = 0$. \square

1.2. *The Segregated Λ -coalescent.* For the remainder of this article, let K be a segregated space. Let λ_w be the restriction of λ to K_w , defined by $\lambda_w(\cdot) = \frac{\lambda(K_w \cap \cdot)}{\lambda(K_w)}$.

In this section we define our model, which we formulate as a stochastic flow on K . The rate of coalescence in our model is controlled by a sequence $(r_n)_{n \in \mathbb{N}_0}$ (recall $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$) such that $r_n \geq 0$ for all n . To avoid degeneracy, we assume that $r_n > 0$ for some $n \in \mathbb{N}_0$.

Heuristic definition: For each $w \in W_*$ the complex K_w is equipped with an exponential clock that rings repeatedly and at rate r_n , where $n = |w|$. Informally, if the clock for K_w rings at time $t \in \mathbb{R}$, then all the particles which are in K_w at time t —are coalesced together and jump at time t into a location p that is sampled according to λ_w . When this occurs, we say that a *coalescence event* (t, w, p) has occurred in K_w at time t with parent point p . We say the points $x \in K_w$ are affected by the coalescence event.

We write the resulting flow of particles as $X_{s,t} : K \rightarrow K$, where $X_{s,t}(x)$ is the location at time t of that the particle which was at x at time $s < t$. A graphical representation of the above paragraph can be seen in Figure 1.

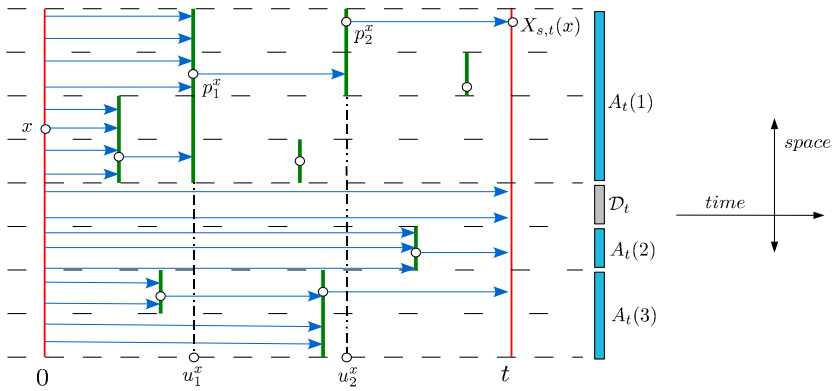


FIG. 1. The complexes of the geographical space with $S = 2$ are shown down to level 3, with dotted lines. Coalescence events are shown as thick vertical lines, with parents as circular dots. In this realization there are no coalescence events occurring in complexes of level 4 and above. The movement of some sample particles over $[0, t]$ is shown using arrows. The division of K into dust \mathcal{D}_t and nontrivial blocks $A_t = \{A_t(1), A_t(2), A_t(3)\}$ (see Section 1.3) over time $[0, t]$ is also shown. The sequences (u_m^x) and (p_m^x) corresponding to the point $x \in K$ are shown to illustrate Definition 2.1 (in this example, $N_1^x = 1$, $N_2^x = 2$ and $N_3^x = \infty$).

It is possible to sample the parent points p according to some measure other than λ_w , but in this article (for brevity) we restrict ourselves to that special case. See Freeman (2012) for a more general mechanism.

Of course, our heuristic definition only makes mathematical sense if the total number of coalescence events during $[s, t]$ is a.s. finite for all $s < t$ (equivalently, if the total rate $\sum S^n r_n$ of all the exponential clocks is finite). However, it does provide accurate intuition for the behaviour of our model in the general case. For arbitrary (r_n) , a mathematical definition which formalizes this intuition can be found in Section 2. We state our existence theorem below, with the understanding that it applies to the definition in Section 2. The proof appears in Section 2.1.

Let \mathbb{M} be the space of functions mapping K to itself, equipped with the metric $\|f, g\|_\infty = \sup\{D_K(f(x), g(x)); x \in K\}$.

THEOREM 1. *For each $s \leq t$, $X_{s,t}$ is an \mathbb{M} -valued random variable. The following properties hold:*

- *For all $s \leq t \leq u$, $X_{s,u} = X_{t,u} \circ X_{s,t}$ surely.*
- *For all $s \leq t < u \leq v$, $X_{u,v}$ and $X_{s,t}$ are independent.*
- *If $t_1 - s_1 = t_2 - s_2$, then X_{s_1,t_1} and X_{s_2,t_2} are identically distributed.*
- *For all $t \in \mathbb{R}$ and $x \in K$, $X_{t,t}(x) = x$ surely.*

The formula $X_{s,u} = X_{t,u} \circ X_{s,t}$ is known as the flow property and shows that the population which our model describes has a consistent genealogical structure.

The flow X is time homogeneous, and for most of this article we will be interested only in $(X_{0,t})_{t \geq 0}$. We think of each point $x \in K$ being home at time 0 to a single particle. The function $X_{0,t}$ specifies which particles are coalesced together during $[0, t]$ and where in space the resulting blocks of coalesced particles end up.

DEFINITION 1.5. If $X_{0,t}(K)$ is a finite set, then we say the Segregated Λ -coalescent has come down from infinity at time $t > 0$. If $X_{0,t}(K)$ is finite for all $t > 0$, then we say the Segregated Λ -coalescent comes down from infinity.

Loosely speaking, if our model is to come down from infinity, then a large number of coalescence events must occur (i.e., the rates r_n must be large). On a coalescence event, the coalesced particles jump through space to that events parent location. Thus if coalescence events occur fast, then our particles will also jump fast. The tree structure $(K_w)_{w \in W_*}$ on the segregated space K provides a method of controlling how far particles move when they jump and, as we see in Section 2, is the crucial ingredient that allows us to make sense of our model for any $(r_n) \subseteq [0, \infty)$.

1.3. *Phases transitions of the Segregated Λ -coalescent.* For each $t \geq 0$, we say (the individuals which began at) x and y are in the same *block* at time t if $X_{0,t}(x) = X_{0,t}(y)$, and in this case we write $x \sim_t y$. It is easily seen that \sim_t is an equivalence relation on K , and we write the equivalence class of $x \in K$ under \sim_t as $[x]_t$. We define

$$\begin{aligned} \mathcal{D}_t &= \{x \in K; [x]_t \text{ is a singleton}\}, \\ A_t &= \{[x]_t; x \in K \setminus \mathcal{D}_t\}. \end{aligned}$$

Thus, \mathcal{D}_t is the union of all the singleton blocks and is the *dust component* of our model at time t . The set A_t is the set of nonsingleton blocks at time t . It is easily seen that, for all $t \geq 0$,

$$(1.1) \quad K = \mathcal{D}_t \uplus \left(\bigcup A_t\right).$$

From the flow property (see Theorem 1) and the definition of \mathcal{D}_t , we have that

$$(1.2) \quad s < t \quad \Rightarrow \quad \mathcal{D}_t \subseteq \mathcal{D}_s.$$

As we vary the parameters (r_n) and \mathcal{S} , we say that a phase transition occurs within our model if we see a qualitative change in the behaviour of \mathcal{D}_t and A_t . In particular, we are interested in whether A_t is finite or infinite, and whether \mathcal{D}_t is empty or nonempty. When \mathcal{D}_t is nonempty we are also interested in whether \mathcal{D}_t is λ -null or has positive measure. Note that $A_0 = \emptyset$ and $\mathcal{D}_0 = K$ almost surely.

Let

$$(1.3) \quad \tau = \inf\{t \geq 0; \mathcal{D}_t = \emptyset\}$$

be the time at which the dust component has been entirely absorbed into the nonsingleton blocks. Since $r_N \neq 0$ for some $N \in \mathbb{N}$, at some time $\kappa < \infty$, the exponential clocks associated to all K_w with $|w| = N$ have rung, meaning that $\mathcal{D}_\kappa = \emptyset$; thus $\tau < \infty$ almost surely.

THEOREM 2. *Almost surely, if $t > \tau$, then $\mathcal{D}_t = \emptyset$ and A_t is finite. Further, if τ is not almost surely zero, then $\mathbb{P}[\tau > 0] = 1$.*

Therefore, we should classify our phases according to the behaviour seen during time $(0, \tau)$. It turns out that our model has five phases, which we now define. Our model is said to be:

- *lower subcritical* if:
 - for all $t \in (0, \infty)$, $\mathbb{P}[\tau > t] > 0$;
 - $\mathbb{P}[\forall t \in (0, \tau), \lambda(\mathcal{D}_t) > 0 \text{ and } |A_t| < \infty] = 1$.
- *upper subcritical* if:
 - for all $t \in (0, \infty)$, $\mathbb{P}[\tau > t] > 0$;
 - $\mathbb{P}[\forall t \in (0, \tau), \lambda(\mathcal{D}_t) > 0 \text{ and } |A_t| = \infty] = 1$.

- *semicritical* if:
 - for all $t \in (0, \infty)$, $\mathbb{P}[\tau > t] > 0$;
 - $\mathbb{P}[\forall t \in (0, \tau), \lambda(\mathcal{D}_t) = 0 \text{ and } |A_t| = \infty] = 1$.
- *critical* if there is some (deterministic) $t_0 \in (0, \infty)$ such that:
 - $\mathbb{P}[\tau < t_0] = 1$ and for all $t \in (0, t_0)$, $\mathbb{P}[t < \tau] \in (0, 1)$;
 - $\mathbb{P}[\forall t \in (0, \tau), \lambda(\mathcal{D}_t) = 0 \text{ and } |A_t| = \infty] = 1$.
- *supercritical* if $\mathbb{P}[\tau = 0] = 1$.

The quantity t_0 is known as the critical time. We are able to completely classify the phases of our model, as follows.

THEOREM 3. *Dependent only upon \mathcal{S} and (r_n) , our model is in precisely one of the five phases. In fact, the Segregated Λ -coalescent X is:*

- *lower subcritical* if and only if $\sum_n \mathcal{S}^n r_n < \infty$;
- *upper subcritical* if and only if $\sum_n \mathcal{S}^n r_n = \infty$ and $\sum r_n < \infty$;
- *semicritical* if and only if $\sum r_n = \infty$ and $\limsup_n \frac{1}{n} \sum_1^n r_j = 0$;
- *critical* if and only if $\limsup_n \frac{1}{n} \sum_1^n r_j \in (0, \infty)$;
- *supercritical* if and only if $\limsup_n \frac{1}{n} \sum_1^n r_j = \infty$.

Further, in the critical phase

$$(1.4) \quad t_0 = \frac{\limsup_n (\sum_0^n r_j) / n}{\log \mathcal{S}}.$$

It follows immediately from Theorem 3 that the Segregated Λ -coalescent comes down from infinity at $t > 0$ if and only if (1) X is supercritical or (2) X is critical and $t \geq t_0$.

As we commented above, the quantity $\sum \mathcal{S}^n r_n$ is the total rate of all the exponential clocks involved in the definition of our model. Since each $x \in K$ has $x \in K_w$ for precisely one $w \in W_n$, the quantity $\sum r_n$ is the rate at which coalescence events affect a single point of K . It is natural that these two quantities characterise phase transitions. The other quantity which appears in Theorem 3, $\limsup_n \frac{1}{n} \sum_1^n r_j$, relates to coming down from infinity and will be discussed when we outline our proofs in Section 1.7.

As the phase of our model changes, the behaviour of the dust is as expected, in that increasing the intensity of reproduction events reduces the fraction of dust. The lack of monotonicity in the behaviour of the nonsingleton blocks is explained as follows. In the lower subcritical phase there are simply not enough events to make anything more than finitely many nonsingleton blocks. Then, as the rate increases, there is an intermediate period where we see a countable infinity of nontrivial blocks. Eventually there are so many reproduction events that they frequently overlap, and we need (a.s.) only finitely many of them to cover K .

In the supercritical phase $\tau = 0$ almost surely and the question of the behaviour of our model at time τ is trivial. In the other phases we have the following result.

COROLLARY 4. *In all but the supercritical phase, almost surely, $\mathcal{D}_\tau = \emptyset$ and A_τ is finite.*

Theorems 2, 3 and Corollary 4 describe qualitative properties of our model. In particular, they are concerned with behaviour taking place on the tree structure $(K_w)_{w \in W_*}$ and are essentially independent of the metric D_K . However, if we examine more detailed properties of our model, then D_K does play a role. In Section 5 we describe the behaviour of the Hausdorff dimension of \mathcal{D}_t . Under some quite strong regularity conditions, with $K \subseteq \mathbb{R}^d$ and D_K equal to the Euclidean metric, we obtain the following result. Let $\chi(t)$ be the Hausdorff dimension of \mathcal{D}_t , conditional on $\mathcal{D}_t \neq \emptyset$, whenever this is defined.

- In the lower/upper subcritical and semicritical phases, $\chi(t) = \dim_{\mathcal{H}}(K)$ for all $t \geq 0$.
- In the critical phase, χ decreases linearly over $[0, t_0]$ with $\chi(0) = \dim_{\mathcal{H}}(K)$ and $\chi(t_0) = 0$.

Proof of the results stated in Sections 1.2 and 1.3 can be found in Sections 2–5. An outline of the proofs can be found in Section 1.7, but first we will compare our model's behaviour to some well-known coalescent and population models.

1.4. *Comparison to the Λ -coalescent.* Let Λ be a finite measure on $[0, 1]$, and consider the corresponding Λ -coalescent; see, for example, Berestycki (2009). Our model does not feature a spatial analogue of Kingman's coalescent, so we remove the Kingman component of the Λ -coalescent by specifying that $\Lambda(\{0\}) = 0$.

If $\Lambda(\{1\}) > 0$, then the effect on the Λ -coalescent of the atom at 1 is as follows; independently of all other mergers, and at rate $\Lambda(\{1\})$, the Λ -coalescent sees a merger which coagulates the whole population into a single block. Thus, the atom at 1 serves only to obfuscate the behaviour of the Λ -coalescent, and is typically removed.

In view of the above paragraph, suppose from now on that $\Lambda(\{1\}) = 0$. Consider the equivalent of this for the Segregated Λ -coalescent: we could set $r_0 = 0$, but if each of the 1-complexes sees a coalescence event, then we face essentially the same issue in that a finite number of coalescence events has covered the whole population. Of course, this could happen with n -complexes for any $n \in \mathbb{N}_0$, so in our spatial setting there is no simple way to remove the chance that a finite number of coalescence events may cover the whole population.

The Λ -coalescent is said to come down from infinity at time $t > 0$ if Π_t is a finite set. It is shown in Pitman (1999) that if the Λ -coalescent comes down from infinity at some $t > 0$, then it does so for all $t > 0$; hence the Λ -coalescent has no equivalent of our critical phase. Pitman's proof of this fact uses a zero-one law and relies on the Λ -coalescent containing only countably many individuals; for this reason the same argument cannot be applied to the Segregated Λ -coalescent.

Let $\mu^n = \int_0^1 x^n \Lambda(dx)$ and consider the case $\mu^{-1} = \infty$. If the Λ -coalescent does not come down from infinity [e.g., the coalescent of Bolthausen and Sznitman (1998)], then the Λ -coalescent has empty dust and a countable infinity of nontrivial blocks. This behaviour does not occur in our model. Alternatively, if the Λ -coalescent does come down from infinity, then it has empty dust and finitely many atoms, corresponding to our supercritical phase.

Now consider the case $\mu^{-1} < \infty$. It is shown in Freeman (2012) that if $\mu^{-2} = \infty$, then the Λ -coalescent has a countably infinity of nontrivial blocks, and a positive fraction of the population contained within the dust. Similarly, if $\mu^{-2} < \infty$, then it is well known [e.g., Example 19 in Pitman (1999)] that the Λ -coalescent has only finitely many nontrivial blocks, and has a nonnull proportion of the population contained in the dust. Thus the Λ -coalescent has equivalents of both our upper and lower subcritical phases.

To summarise the previous few paragraphs, the behaviour seen by our model is that of the Λ -coalescent, with following modifications:

1. Playing the role of the cases where $\Lambda(\{1\}) > 0$, we have the (always positive) probability of having only finitely many nontrivial blocks and no dust.
2. There is no possibility in our model of having a countable infinity of nontrivial blocks and empty dust. This behaviour is replaced by our semicritical phase, in which we see a countably infinity of nontrivial blocks and nonempty null dust.
3. The critical phase appears in between the semi- and supercritical phases.

1.5. *Connections to spatial Λ -coalescents.* Our model is a spatial extension of the Λ -coalescent. A mean-field version of the Λ -coalescent has already appeared in Limic and Sturm (2006), building on the mean-field version of Kingman's coalescent from Greven, Limic and Winter (2005). The model from Limic and Sturm (2006) is referred to in the literature as “the spatial Λ -coalescent”.

We refer the reader to Limic and Sturm (2006) for a proper description of their model and restrict ourselves here to outlining some important aspects in which it differs from our own. The model of Limic and Sturm uses a finite graph G as its geographical space, whereas we use a spatial continuum. The points of its geographical space may be inhabited by more than one block at any time, whereas we permit at most one block of individuals to inhabit a single point of K . Further, the blocks of individuals in the model of Limic and Sturm wander freely around G and may only be coagulated with other blocks that happen to be at the same vertex at the same time. By contrast, blocks in our own model do not move in between mergers, but a single merger involves a nonnull proportion of our geographical space.

Thus, the two models are very different and it is natural to expect different behaviour. In fact, such differences are readily seen. For example, Angel, Berestycki and Limic (2010) show that model of Limic and Sturm, modified slightly so as G is countably infinite and of bounded degree, does not come down from infinity.

1.6. *Connections to the Spatial Λ -Fleming–Viot process.* We now introduce a family of population models that are closely related to the Λ -coalescent. The dual of the Λ -coalescent is the Λ -Fleming–Viot process, constructed in [Le Gall and Bertoin \(2003\)](#) [and also implicitly in [Donnelly and Kurtz \(1999\)](#)]. The Λ -Fleming–Viot process is a natural generalisation of the classical Fleming–Viot process, which is itself dual to Kingman’s coalescent; see, for example, [Etheridge \(2011\)](#). The Spatial Λ -Fleming–Viot process (SAFV) was introduced in [Etheridge \(2008\)](#) as a spatial extension of the Λ -Fleming–Viot process.

The SAFV process is an infinite system of interacting Λ -Fleming–Viot processes, with one such process at each site of \mathbb{R}^d . See [Barton, Etheridge and Véber \(2010\)](#) for a precise description of the SAFV process, and see [Barton, Etheridge and Véber \(2013\)](#) for a survey of recent results concerning (variants of) the SAFV process.

Since the dual of the Λ -Fleming–Viot process is the Λ -coalescent, and the SAFV is a spatial version of the Λ -Fleming–Viot process, the dual to the SAFV process behaves like a spatial version of the Λ -coalescent. However, as we see below the dual of the SAFV process does not display the full range of Λ -coalescent like behaviour. Our own model, by contrast, shows that coalescent behaviour can be greatly enriched by the introduction of space.

The ancestral lineages of (individuals sampled from a population whose genealogy is described by) the SAFV process are compound Poisson processes. Like our own model, these ancestral lineages are only coalesced together when they move through space. However, each such lineage is a compound Poisson process that jumps at finite rate.

Now, fix $t > 0$. Since the geographical space used by Barton, Etheridge and Véber is a spatial continuum, we can pick one individual at each point of space and integrate across space (using Fubini’s theorem) to see that with positive probability there is a nonnull subset of the geographical space, containing an infinite subset of our chosen individuals, all of whom have not been affected by a reproduction event during $[0, t]$. In the language of coalescents, at all times there is positive probability of the dust being nonempty. Therefore, the model of [Barton, Etheridge and Véber \(2010\)](#) does not come down from infinity.

In light of the above paragraph, the reader might wonder why Barton, Etheridge and Véber force their ancestral lineages to be compound Poisson processes rather than using more general Lévy processes. The difficulty stems from an apparent incompatibility between the compensation mechanism (in \mathbb{R}^d) of “true” Lévy processes and the mathematical machinery used to construct the SAFV process; we refer the reader to [Barton, Etheridge and Véber \(2010\)](#) for a proper discussion. It would not be fair to claim that we have overcome this difficulty in our model. Rather, we choose our state space and reproduction mechanism in such a way as our ancestral lineages can jump at infinite rate, but without the need for Lévy process style compensation.

Another stochastic model with similar features, a system of interacting Cannings processes on the hierarchical group, is investigated by Greven et al. (2012). They carry out a renormalization procedure corresponding to the hierarchical mean field limit, obtaining a detailed description of the limiting object and its behaviour in terms of clustering and coexistence.

1.7. *Outline of the proofs.* Our proof of Theorem 3 comes via several lines of enquiry. Firstly Fubini’s theorem produces some useful information and, secondly, the spatial structure of K provides some basic connections between \mathcal{D}_t and A_t . However, the most important contribution comes via a connection between our model and Galton–Watson processes in Varying Environments (GWVEs). A GWVE is a classical Galton–Watson process, with the modification that the offspring distribution of an individual may depend on its generation number. An introduction to GWVEs can be found in Fearn (1972).

Note that branching structures also play a pivotal role in the study of the Λ -coalescent, as can be seen in, for example, Bertoin and Le Gall (2000), Birkner et al. (2005) and Berestycki, Berestycki and Limic (2014).

For each $w \in W_*$, let \mathcal{E}_w be the first time $t > 0$ at which a coalescence event occurs in the complex K_w (to be clear, the event must occur in precisely K_w and not just inside one of its subcomplexes). We refer to \mathcal{E}_w as the exponential clock for K_w . The connection to our model is as follows. For each $t > 0$ and $n \in \mathbb{N}_0$ we define

$$\mathcal{B}_n^t = \{K_w; |w| = n \text{ and for all } u \in W_*, \text{ if } K_w \subseteq K_u \text{ then } \mathcal{E}_u > t\}$$

and write $B_n^t = |\mathcal{B}_n^t|$ for the number of elements of \mathcal{B}_n^t . Set $\mathcal{B}^t = \bigcup_{n \in \mathbb{N}_0} \mathcal{B}_n^t$.

It can be seen (in Lemma 3.4) that (B_n^t) is a GWVE with an n th stage offspring distribution that is Binomial($S, e^{-tr_{n+1}}$). Note that the case $r_n = c \in (0, \infty)$, where the GWVEs are classical Galton–Watson processes, is part of the critical phase.

It turns out that the behaviour of $B_n^t = |\mathcal{B}_n^t|$ as $n \rightarrow \infty$ is closely connected to the behaviour of \mathcal{D}_t . In fact, Lemma 3.5 (which appears in Section 3.3) says that

$$(1.5) \quad \mathcal{D}_t = \bigcap_{n \in \mathbb{N}} \bigcup_{w \in \mathcal{B}_n^t} K_w.$$

A GWVE (B_n) is said to be *degenerate* if $\mathbb{P}[\exists n \in \mathbb{N}, B_n = 0] = 1$. In view of (1.5), it is important for us to understand when (B_n^t) is degenerate, since in this case $\mathcal{D}_t = \emptyset$. Conditions equivalent to degeneracy of GWVEs are in general not known, but conditions covering cases including (B_n^t) have been known for some time, in fact since Agresti (1975) and Jiřina (1976). Further conditions were given by Lyons (1992). The conditions of Jiřina (1976) are best suited to our setting, and we state them in Lemma 3.6.

The quantity $\limsup_n \frac{1}{n} \sum_1^n r_j$ (which appeared in Theorem 3) plays a central role in characterizing degeneracy of (B_n^t) . Its precise role is subtle but a partial

explanation of the formula is the following. When reproduction events are occurring at a high rate, it becomes common for a larger reproduction event to overwrite the effect of some of the preceding smaller ones. This is borne out by the appearance of the \limsup ; from Theorem 3 we see that, when $\sum r_n = \infty$, only the n -level reproduction events for which r_n is large enough to contribute to $\limsup_n \frac{1}{n} \sum_1^n r_j$ actually take part in determining the phase.

Formulas similar to (1.5) can be found in the random fractals literature at least as far back as Falconer (1986), Mauldin and Williams (1986) and Graf (1987) (although these authors did not use branching processes explicitly). Such formulas provide what is now a well-known connection between various classes of random fractals and branching processes. In fact, in Section 5 we use a result of Durand (2009) to calculate the Hausdorff dimension of \mathcal{D}_t , conditional on $\mathcal{D}_t \neq \emptyset$.

In addition to using GWVEs, in order to understand the behaviour of τ we will use some techniques from the percolation literature. Many types of branching process, including GWVEs, can be reformulated as an inhomogeneous percolation process on some suitable tree. The relationship is displayed in great generality by, for example, Lyons (1992). In the case of our GWVEs we have the following.

Consider W_* as the nodes of a regular \mathcal{S} -ary tree, rooted at \emptyset with edge set $\mathbb{E} = \{(w, wi); w \in W_*, i \in \mathcal{S}\}$. Fix $t > 0$. We say that the node $w \in W_*$ is open if $\mathcal{E}_w > t$ (and closed otherwise) and note that \mathcal{B}^t is the set of points in W_* which are connected to the root node \emptyset via edges with only open nodes at their endpoints. In the language of percolation, \mathcal{B}^t is the open cluster connected to \emptyset . Note that each node $w \in W_*$ chooses independently whether it is open or closed. The distribution of \mathcal{E}_w varies with $|w|$, so in fact \mathcal{B}^t is an inhomogeneous percolation on the \mathcal{S} -ary tree W_* .

2. Existence of the model. In this section we prove the existence of the Segregated Λ -coalescent. We begin with the definition of our model, formalizing the heuristic description given in Section 1.2.

Let \mathcal{P} be the measure on $W_* \times K$ defined by $\mathcal{P}(\{w\} \times A) = r_{|w|} \times \lambda_w(A)$ for each $w \in W_*$ and measurable $A \subseteq K$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a Poisson point process M in $\mathbb{R} \times W_* \times K$, of intensity $dt \otimes \mathcal{P}(dw, dy)$, where dt denotes Lebesgue measure. For (measurable) $I \subseteq \mathbb{R}$, $V \subseteq W_*$ and $A \subseteq K$ define

$$M_I = \{(t, w, y) \in M; t \in I\}, \quad M_{I \times V} = \{(t, w, y) \in M; t \in I, w \in V\}.$$

Note that, in terms of M , $\mathcal{E}_w = \inf\{s > 0; M_{(0,s] \times \{w\}} \neq \emptyset\}$.

It will be to our advantage to have some almost sure properties of M as “sure” properties of M . In particular, by standard properties of Poisson point processes [see, e.g., Kingman (1993)], with probability one:

- (a) For all $k \in \mathbb{N}$ and $n \in \mathbb{N} \cup \{0\}$, $M_{[-k,k] \times W_n}$ is finite.
- (b) For all $u \in \mathbb{R}$, $M_{\{u\}}$ is finite.

With slight abuse of notation, we simply redefine M so as (a) and (b) hold for all $\omega \in \Omega$.

Let us examine Figure 1 and determine which coalescence events, according to our heuristic, actually influence the final position of the lineages. Consider an event (s, w, y) in a complex K_w of level $|w| = n$ at time $s \in (0, t)$. The event had no effect on the position at time t of any of the lineages if:

- There was an event (s', w', y') such that $s < s' < t$ and $K_w \subseteq K_{w'}$.
- Or, the final event (s', w', y') such that $0 < s' < s$ and $K_w \subset K_{w'}$ had $y' \notin K_w$.

Hence, to work out where $x \in K$ should be mapped to over time $[s, t]$ we need only consider the following sequence of events.

First, look for the final level 0 event during $(s, t]$ which affected the point x . If we find one, say (u_1, w_1, p_1) , we then look for the final level 1 event which was after time u_1 and affected p_1 , and so on. If at any point we do not find a level n event, we simply move up to level $n + 1$ and look there. In symbols:

DEFINITION 2.1. Fix (x, s, t) with $x \in K$ and $s \leq t$. Let $(u_0, w_0, p_0) = (s, \emptyset, x)$ and set $N_0 = -1$. For as long as $N_{m+1} < \infty$ define inductively a pair of (possibly finite) sequences by

$$N_{m+1} = \min\{n > N_m; \exists(u, w, p) \in M_{(u_m, t] \times W_n} \text{ such that } p_m \in K_w\},$$

$$E_{m+1} = (u_{m+1}, w_{m+1}, p_{m+1}),$$

where

$$u_{m+1} = \max\{u \in (u_m, t]; (u, w, p) \in M_{(u_m, t] \times W_{N_{m+1}}} \text{ and } p_m \in K_w\}.$$

Define $(E_m) = \{(u_i, w_i, p_i); i = 1, 2, \dots\}$, $(N_m) = \{N_i; i = 1, 2, \dots\}$ and note that we do not include the term $i = 0$. Define (u_m) , (w_m) , (p_m) similarly. A graphical demonstration of Definition 2.1 can be seen in Figure 1.

Since (a) and (b) hold for all $\omega \in \Omega$, (E_m) and (N_m) are well defined for all $\omega \in \Omega$. The (finite or infinite) sequence of coalescence events (E_m) contains the only events which affected the final position of the lineage that started from x and moved during time $(s, t]$.

NOTATION 2.2 (Continuation of Definition 2.1). The sequence (E_m) depends on x, s and t . When we need this distinction (which will be most of the time), we write

$$E_m^{x,s,t} = (u_m^{x,s,t}, w_m^{x,s,t}, p_m^{x,s,t}).$$

We write $K_{w_m^{x,s,t}} = K_{w_m}^{x,s,t}$. Occasionally, if s and t are both clear from the context then we may omit them as superscripts and write $E_m^x = (u_m^x, w_m^x, p_m^x)$, $K_{w_m}^x = K_{w_m^x}$.

We will shortly define $X_{s,t}$ using the language above, but first we need to note a technical point that concerns the following lemma.

LEMMA 2.3. *If the sequence $(E_m^{x,s,t})$ is infinite, then the sequence $(p_m^{x,s,t})$ is convergent.*

PROOF. Suppose $(E_m^{x,s,t})$ is infinite. By Definition 2.1, $|w_m^{x,s,t}|$ is \mathbb{N} valued and strictly increasing, so $|w_m^{x,s,t}| \rightarrow \infty$ as $m \rightarrow \infty$. Thus $K_{w_m}^{x,s,t}$ is a decreasing sequence of sets. Note that $p_m^{x,s,t} \in K_{w_m}$. By $(\mathcal{K}2)$, $(p_m^{x,s,t})$ is a Cauchy sequence so, by completeness of K , $(p_m^{x,s,t})$ is convergent. \square

Suppose for a moment that all complexes K_w of K are closed, and recall our heuristic definition from Section 1.2. Then it makes intuitive sense that reproduction events occurring in complexes $K_{w'} \subseteq K_w$ cannot move particles in the flow from within K_w into $K \setminus K_w$. However, if some K_w is not closed, then it might be the case that an infinite sequence (u'_m, w'_m, p'_m) of events, with $K_{w'_m} \subseteq K_w$, could have $\lim p'_m \notin K_w$, because it could be that $\lim p'_m \in \overline{K_w} \setminus K_w$. In this case our construction would run into a serious problem; the flow property $X_{s,v} = X_{t,v} \circ X_{s,t}$ would fail.

To address this issue, we introduce the set

$$\mathcal{O} = \bigcup_{n \in \mathbb{N}} \bigcup_{\substack{w, w' \in W_n \\ w \neq w'}} \overline{K_w} \cap K_{w'}.$$

If $\mathcal{O} = \emptyset$, then we say K is *completely segregated*. Recall Examples 1.2 and 1.3, and note that Example 1.2 is completely segregated, but Example 1.3 is not.

Until further notice, which means until Section 4, we will assume that K is completely segregated, that is, $\mathcal{O} = \emptyset$. Then, in Section 4 we will discuss the modifications which are necessary to make our arguments work in the case $\mathcal{O} \neq \emptyset$.

The Segregated Λ -coalescent is the process $(X_{s,t})_{s \leq t}$ defined as follows:

$$(2.1) \quad X_{s,t}(x) = \begin{cases} x, & \text{if } (E_m^{x,s,t}) \text{ is empty,} \\ p_{\mathcal{M}}^{x,s,t}, & \text{if } (E_m^{x,s,t}) = (E_m^{x,s,t})_{m=1}^{\mathcal{M}} \text{ for } \mathcal{M} \in \mathbb{N}, \\ \lim_{m \rightarrow \infty} p_m^{x,s,t}, & \text{if } (E_m^{x,s,t}) \text{ is infinite.} \end{cases}$$

In fact, in Section 4 we will see that when $\mathcal{O} \neq \emptyset$, there is a \mathbb{P} -null set on which it makes sense to define X differently to (2.1). Until then we use (2.1) for all $\omega \in \Omega$. We now record some results which use the fact that $\mathcal{O} = \emptyset$.

LEMMA 2.4. *Every subcomplex K_w is a closed subset of K .*

PROOF. Let $n \in \mathbb{N}_0$ and $w \in W_n$. Note that by $(\mathcal{K}1)$ we have $K = \biguplus_{w \in W_n} K_w$. Hence, if $x \in \overline{K_w}$, then we must have $x \in K_{w'}$ for some $w' \in W_n$. If $x \notin K_w$, then we would have $x \in \overline{K_w} \cap K_{w'}$, which contradicts $\mathcal{O} = \emptyset$. \square

REMARK 2.5. By Lemma 2.4, if $\mathcal{O} = \emptyset$, then $(\mathcal{N}4)$ holds automatically.

LEMMA 2.6. For all $s < t$, all $x \in K$ and all $n \in \mathbb{N}$, if $(E_m^{x,s,t})$ is infinite, then $\lim_{m \rightarrow \infty} p_m^{x,s,t} \in K_{w_n}^{x,s,t}$.

PROOF. We noted in the proof of Lemma 2.3 that $K_{w_m}^{x,s,t}$ is a decreasing sequence of sets. Hence $p_m^{x,s,t} \in K_{w_m}^{x,s,t} \subseteq K_{w_n}^{x,s,t}$ for all $m \geq n$. By Lemma 2.4, $K_{w_l}^{x,s,t}$ is closed for all $l \in \mathbb{N}$ and hence $\lim_{m \rightarrow \infty} p_m^{x,s,t} \in K_{w_n}^{x,s,t}$. \square

LEMMA 2.7. For all $x \in K \setminus \mathcal{D}_t$ and $t > 0$, $[x]_t = K_{w_1}^{x,0,t}$.

PROOF. If $x \notin \mathcal{D}_t$, then for some $y \in K$ we have $y \neq x$ and $X_{0,t}(x) = X_{0,t}(y)$. It follows from Definition 2.1, that $E^{x,s,t} \neq \emptyset$ in this case. Note that if $y \in K_{w_1}^{x,0,t}$, then $X_{0,t}(y) = X_{0,t}(x)$, so $K^{x,0,t} \subseteq [x]_t$. Now, suppose $z \in K \setminus K_{w_1}^{x,0,t}$. If $(K \setminus K_{w_1}^{x,0,t}) \cap (K \setminus K_{w_1}^{z,0,t}) \neq \emptyset$, then by $(\mathcal{N}1)$ either (1) $(K \setminus K_{w_1}^{x,0,t}) \subseteq (K \setminus K_{w_1}^{z,0,t})$ or (2) $(K \setminus K_{w_1}^{z,0,t}) \cap (K \setminus K_{w_1}^{x,0,t})$.

If (1) holds, then $x \in K_{w_1}^{z,0,t}$, so by Definition 2.1 we must have $N_1^{x,0,t} \geq N_1^{z,0,t}$, which means $|w_1^{x,0,t}| \geq |w_1^{z,0,t}|$. But combined with (1) and $(\mathcal{N}1)$ this implies that $K \setminus K_{w_1}^{x,0,t} = K \setminus K_{w_1}^{z,0,t}$, which contradicts the definition of z .

Similarly, if (2) holds, then $z \in K_{w_1}^{x,0,t}$, so by Definition 2.1 we must have $N_{w_1}^{z,0,t} \geq N_{w_1}^{x,0,t}$, which means $|w_1^{z,0,t}| \geq |w_1^{x,0,t}|$. Combined with (2) and $(\mathcal{N}1)$ this implies that $K \setminus K_{w_1}^{x,0,t} = K \setminus K_{w_1}^{z,0,t}$, which again contradicts the definition of z .

Thus, we must have $(K \setminus K_{w_1}^{x,0,t}) \cap (K \setminus K_{w_1}^{z,0,t}) = \emptyset$. From this, Lemma 2.6 implies that $X_{0,t}(x) \neq X_{0,t}(z)$ so as $z \notin [x]_t$. Since z was arbitrary, $[x]_t = K_{w_1}^{x,0,t}$. \square

The remainder of Section 2 is concerned with proving Theorem 1 and establishing some regularity results that we require in Section 3. Readers who are more interested in proving Theorems 2, 3 and Corollary 4 may wish to omit Sections 2.1 and 2.2 and move straight on to Section 3.

2.1. *Proof of Theorem 1.* The proof comes in three parts, which correspond to the bullet points in the statement of Theorem 1. The first part is a careful check of the flow property.

PART 1: Let $s \leq t \leq v$, and fix $x \in K$. Write $y = X_{s,t}(x)$. When necessary we will emphasise the dependence with $y = y^{x,s,t}$. We divide into three cases.

If $N_1^{x,s,t} = N_1^{y,t,v} = \infty$: then for all $x \in K$, $X_{s,t}(x) = x = y$ and $X_{t,v}(y) = y$. Since $x = y$, $N_1^{x,s,v} = \infty$ and $X_{s,t}(x) = X_{t,v}(x) = X_{s,v}(x)$, so $X_{t,v}(X_{s,t}(x)) = X_{s,v}(x)$.

If $N_1^{x,s,t} = \infty$ and $N_1^{x,t,v} < \infty$, then $X_{s,t}(x) = x = y$, and hence we must have $u_1^{x,s,t} \geq v$. Hence $(E_m^{x,t,v}) = (E_m^{x,s,v})$ and thus $X_{t,v}(X_{s,t}(x)) = X_{t,v}(x) = X_{s,t}(x)$.
 If $N_1^{x,s,t} < \infty$, then we have $N_1^{x,s,v} < \infty$. Let

$$\mathcal{C}^{s,t,v} = \{x \in K; \exists m, u_m^{x,s,v} \geq t\}.$$

If $x \notin \mathcal{C}^{s,t,v}$, then $u_m^{x,s,v} < t$ for all m , so from the definitions we have $(E_m^{x,s,t}) = (E_m^{x,s,v})$. Hence $X_{t,v}(x) = X_{s,t}(x) = y$. Suppose it was the case that $(u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v}) \in (E^{y,t,v})$. Note $y \in K_{w_1}^{x,s,v}$ so we must have $(u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v}) \in (E_m^{x,s,v})$, which is a contradiction since $u_1^{y,t,v} \geq t$. Hence $(E_m^{y,t,v})$ is empty, and $X_{t,v} = t$. Thus, $X_{t,v}(X_{s,t}(x)) = X_{s,v}(x)$.

If $x \in \mathcal{C}^{s,t,v}$, let

$$\mathcal{M} = \max\{m; u_m^{x,s,v} < t\}$$

[which is well defined since $(u_m^{x,s,t})$ is strictly increasing], and from the definitions note that $(E_m^{x,s,t})_{\mathcal{M}} = (E_m^{x,s,v})_{\mathcal{M}}$.

By definition of \mathcal{M} we have $u_{\mathcal{M}+1}^{x,s,v} \geq t$, and since $p_{\mathcal{M}}^{x,s,v} = p_{\mathcal{M}}^{x,s,t}$, it holds that $N_{\mathcal{M}+1}^{x,s,v} \leq N_{\mathcal{M}+1}^{x,s,t}$. Hence $K_{w_{\mathcal{M}+1}}^{x,s,t} \subseteq K_{w_{\mathcal{M}+1}}^{x,s,v}$. By definition, $p_{\mathcal{M}}^{x,s,t} \in K_{w_{\mathcal{M}+1}}^{x,s,t}$ and, we have also that $(K_{w_m}^{x,s,t})$ is decreasing. We have already commented that $K_{w_{\mathcal{M}+1}}^{x,s,t} \subseteq K_{w_{\mathcal{M}+1}}^{x,s,v}$, so it follows from Lemma 2.6 that $y^{x,s,t} \in K_{w_{\mathcal{M}+1}}^{x,s,v}$.

Since both y and $p_{\mathcal{M}}^{x,s,v}$ are elements of $K_{w_{\mathcal{M}+1}}^{x,s,v}$, there is no $(u, w, p) \in (E_m^{y,t,v})$ such that $|w| < N_{\mathcal{M}+1}^{x,s,v}$ —such a (u, w, p) would also have featured in $(E_m^{x,s,v})$, which contradicts the definition of \mathcal{M} . Also, there are no $(u, w, p) \in (E_m^{y,t,v})$ such that $u > u_{\mathcal{M}+1}^{x,s,v}$ and $y \in K_w$ —such a (u, w, p) would feature in $(E_m^{x,s,v})$, which contradicts the definition of $u_{\mathcal{M}+1}^{x,s,v}$.

Combining the results of the previous two sentences, $(u_{\mathcal{M}+1}^{x,s,v}, w_{\mathcal{M}+1}^{x,s,v}, p_{\mathcal{M}+1}^{x,s,v}) = (u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v})$. Hence $(E_m^{x,s,v})_{m \geq \mathcal{M}+1} = (E_k^{x,s,v})_{k \geq 1}$, which implies that $X_{t,v}(y) = X_{s,v}(x)$. This completes the third case.

Since x and ω were arbitrary, in all cases we have that for all $\omega \in \Omega$, $X_{s,v} = X_{t,v} \circ X_{s,t}$.

PART 2: Let $s_1 < t_1 \leq s_2 < t_s$. Since $M_{(s_1,t_1]}$ and $M_{(s_2,t_2]}$ are independent, and the construction of $X_{s,t}$ depended only on $M_{s,t}$, it follows immediately that X_{s_1,t_1} and X_{s_2,t_2} are independent.

PART 3: Let $s_1 < t_1$ and $s_2 < t_2$ with $t_1 - s_1 = t_2 - s_2$. Then $M_{(s_1,t_1]}$ and $\{(u - (t_2 - t_1), w, p); (u, w, p) \in M_{(s_2,t_2]}\}$ are identical in distribution, from which it follows that X_{s_1,t_1} and X_{s_2,t_2} are also identical in distribution.

PART 4: Let $t \in \mathbb{R}$. Note that $(t, t]$ is empty, so as by Definition 2.1 we have that $(E^{x,t,t})$ is empty for all $x \in K$. Thus $X_{t,t}$ is the identity function.

2.2. *Regularity.* Recall that our underlying Poisson point process Π is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout this section we denote the dependence on $\omega \in \Omega$ of X by writing $X_{0,t}(\cdot)(\omega)$. Let $\mathbb{B}(K)$ denote the Borel σ -algebra on K , and recall that D_K denotes the metric on K .

LEMMA 2.8. *K is separable.*

PROOF. By Lemma 1.4, each K_w is nonempty. For each K_w pick some point $x(w) \in K_w$ and define $\mathbb{D} = \{x(w); w \in W^*\}$. Note that \mathbb{D} is countable.

Let O be an open set of K . Since K is a metric space, for some $r > 0$ and $y \in K$, $B_r(y) \subseteq O$. For $y \in K$ and $n \in \mathbb{N}$ let $K_{(y,n)}$ be the unique complex K_w of K such that $|w| = n$ and $y \in K_w$. By (A2), for some $n \in \mathbb{N}$ we have $r_n < r/2$, so as $K_{(y,n)} \subseteq B_r(y) \subseteq O$. By definition of $x(w)$ there is $w \in W_*$ such that $x(w) \in K_{(y,n)} \subseteq O$. Hence \mathbb{D} is a countable dense subset of K . \square

LEMMA 2.9. *The Borel σ -algebra on K is generated by $(K_w)_{w \in W_*}$.*

PROOF. By Definition 1.1, each K_w is measurable, so it is clear that $\sigma(K_w; w \in W_*) \subseteq \mathbb{B}(K)$. We will now prove the reverse inclusion.

By Lemma 2.8, K is separable, hence any open subset of K can be written as a union of only countably many open balls of K . Hence $\mathbb{B}(K)$ is generated by the open balls of K . So the proof is complete if we can show that any open ball of K is contained in $\sigma(K_w; w \in W_*)$.

To this end, let $B_r(x) = \{y \in K; |y - x| < r\}$ be a fixed but arbitrary open ball in K . By (A1), for each $y \in K$ and $n \in \mathbb{N}$, let $K_{(y,n)}$ be the unique complex K_w of K such that $|w| = n$ and $y \in K_w$. Note that

$$(2.2) \quad B_r(x) \supseteq \bigcup \{K_w; w \in W_*, K_w \subseteq B_r(x)\}$$

is tautologically true, and, since W_* is countable, the union on the right is countable. Now, suppose that $y \in B_r(x)$. Since $B_r(x)$ is open, for some $\varepsilon > 0$ we have $B_\varepsilon(y) \subseteq B_r(x)$. By (A2), for some sufficiently large $n \in \mathbb{N}$ we have $K_{w(y,n)} \subseteq B_\varepsilon(y) \subseteq B_r(x)$. However, this implies that $K_{w(y,n)} \in \{K_w; w \in W_*, K_w \subseteq B_r(x)\}$, so as $y \in \bigcup \{K_w; w \in W_*, K_w \subseteq B_r(x)\}$. Hence, in fact (2.2) is an equality, and thus $B_r(x) \in \sigma(K_w; w \in W_*)$. \square

LEMMA 2.10. *For each $s < t$, $(x, \omega) \mapsto X_{s,t}(x)(\omega)$ is a measurable function from $K \times \Omega \rightarrow K$.*

PROOF. The definition of our model is translation invariant across time, so it suffices to consider the case $s = 0$. For $v \in W_*$ let $\mathcal{W}(v) = \{w' \in W_*; K_v \subseteq K_{w'}\}$.

Fix $w \in W_*$. We note

$$\begin{aligned} & \{(x, \omega) \in K \times \Omega; X_{0,t}(x)(\omega) \in K_w\} \\ &= \bigcup_{v \in W_*} K_v \times \{\omega \in \Omega; X_{0,t}(K_v) \subseteq K_w\} \\ &= \bigcup_{v \in W_*} K_v \times [\{\omega \in \Omega \setminus \mathcal{A}; K_u \subseteq K_v\} \\ & \quad \uplus (\mathcal{A} \cap \{\omega \in \Omega; \nexists(u', w', p') \in \Pi_{(0,t)}, \\ & \quad K_v \subseteq K_{w'}, p' \notin K_w, \Pi_{[u',t] \times \mathcal{W}(v)} = \emptyset\})]. \end{aligned}$$

Note that $\{\omega \in \Omega \setminus \mathcal{A}; K_u \subseteq K_v\}$ is either empty or equal to the measurable set $\Omega \setminus \mathcal{A}$. From the representation above, it follows that $\{(x, \omega) \in K \times \Omega; X_{0,t}(x)(\omega) \in K_w\}$ is an element of the product σ -algebra on $K \times \Omega$. Lemma 2.9 completes the proof. \square

REMARK 2.11 (On particle paths). For each $x \in K$, $t \mapsto X_{0,t}(x)$ is a càdlàg function, and $\omega \mapsto X_{0,\cdot}(x)$ is a random variable in the space of càdlàg K -valued paths (with the usual weak topology). Proof of this result is no more than a long exercise in manipulating the definitions and is not included in this article; see Freeman (2012).

3. Proof of the phase transitions. In this section we prove the results that were stated in Section 1.3, namely Theorems 2, 3 and Corollary 4. We begin with some results based on Fubini’s theorem and the spatial structure of K .

LEMMA 3.1. *If $\sum r_n = \infty$, then $\mathbb{P}[\lambda_t(\mathcal{D}_t) = 0] = 1$.*

PROOF. Fix $t > 0$. By Fubini’s theorem, which applies by Lemma 2.10,

$$(3.1) \quad \mathbb{E}[\lambda(\mathcal{D}_t)] = \mathbb{E}\left[\int_K \mathbb{1}\{x \in \mathcal{D}_t\} \lambda(dx)\right] = \int_K \mathbb{P}[x \in \mathcal{D}_t] \lambda(dx).$$

The rate at which each point $x \in K$ is affected by reproduction events is $\sum_{n \in \mathbb{N}_0} r_n$. Hence, if $\sum r_n = \infty$, then each $x \in K$ has almost surely been involved in some reproduction event during $[0, t]$ and thus $\mathbb{P}[x \in \mathcal{D}_t] = 0$. By (3.1), $\mathbb{P}[\lambda(\mathcal{D}_t) = 0] = 1$. \square

LEMMA 3.2. *Suppose A_t is finite. Then either $\mathcal{D}_t = \emptyset$ or, for some $w \in W_*$, $K_w \subseteq \mathcal{D}_t$. In the latter case, $\lambda_t(\mathcal{D}_t) > 0$.*

PROOF. Suppose that A_t is finite and \mathcal{D}_t is nonempty. Enumerate A_t as $\{[x_i]_t; i = 1, \dots, N\}$ where $x_i \in K$ are such that $[x_i]_t \neq [x_j]_t$ for $i \neq j$. By definition of $[x]_t$ we have $[x_i]_t \cap [x_j]_t = \emptyset$ for $i \neq j$ and by Lemma 2.7 we have $[x_i]_t = K_{w_i}^{x_i, 0,t}$ for all i .

Let $\mathcal{N} = \max\{|w_1^{x_i, 0, t}|; i = 1, \dots, N\}$, and note that by $(\mathcal{K}1)$ we have $K = \uplus_{w \in W_{\mathcal{N}}} K_w$. Since $\mathcal{D}_t \neq \emptyset$, by (1.1) there is some $x \in K \setminus (\cup A_t) = K \setminus (\cup_i K_{w_1^{x_i, 0, t}})$. Combined with $K = \uplus_{w \in W_{\mathcal{N}}} K_w$, this means that for some $w' \in W_{\mathcal{N}}$ we must have $K_{w'} \subseteq K \setminus (\cup_i K_{w_1^{x_i, 0, t}})$, and by (1.1) we must have $K_{w'} \subseteq \mathcal{D}_t$. It follows immediately from Lemma 1.4 that $\lambda(\mathcal{D}_t) > 0$. \square

LEMMA 3.3. *If $\sum S^n r_n = \infty$, then $\mathbb{P}[\exists t > 0, \exists w \in W_*, K_w \subseteq \mathcal{D}_t] = 0$.*

PROOF. Since W_* is countable, the lemma follows if we can show that $\mathbb{P}[\exists t > 0, K_w \subseteq \mathcal{D}_t] = 0$ for an arbitrary $w \in W_*$. So fix $w \in W_*$, and set $n = |w|$. The rate at which K_w is affected by reproduction events is

$$\begin{aligned} \int_{W^* \times K} \mathbb{1}\{K_w \cap K_{w'} \neq \emptyset\} \mathcal{P}(dw', dp) &= \sum_{w' \in W_*} r_{|w|} \mathbb{1}\{K_w \cap K_{w'} \neq \emptyset\} \\ &\geq \sum_{w' \in W_*} r_{|w|} \mathbb{1}\{K_{w'} \subseteq K_w\}. \end{aligned}$$

Now, by $(\mathcal{K}1)$, $K_{w'} \subseteq K_w$ if and only if $w' = wv$ for some $v \in W_*$. Hence,

$$\sum_{w' \in W_*} r_{|w|} \mathbb{1}\{K_{w'} \subseteq K_w\} = \sum_{v \in W_*} r_{|wv|} = \frac{1}{S^n} \sum_{m=n}^{\infty} S^m r_m = \infty.$$

It follows immediately that (with probability one) K_w is affected by a reproduction event during $(0, t]$ for any $t > 0$. Hence $\mathbb{P}[\exists t, K_w \subseteq \mathcal{D}_t] = 0$. \square

The following pair of lemmas, which play a crucial part in our arguments, begin the connection between \mathcal{D}_t , A_t and B^t .

LEMMA 3.4. *For each fixed $t > 0$, $(B_n^t)_{n \geq 0}$ is a GWVE. The initial state is $B_0^t = \mathbb{1}\{\mathcal{E}_{\emptyset} > t\}$, and the n th stage offspring distribution is Binomial with S trials and success probability $e^{-tr_{n+1}}$.*

PROOF. Note first that $e^{-r_{n^t}}$ is the probability that K_w , where $|w| = n$, does not see its clock ring during $(0, t]$. If $w \in \mathcal{B}_n^t$ and $|w| = n$, then the (conditional) probability that $w_i \in \mathcal{B}_{n+1}^t$ is just $e^{-r_{n+1^t}}$. The clocks corresponding to K_{w_i} and K_{w_j} are independent if $i \neq j$. Thus the offspring distribution of $w \in \mathcal{B}_n^t$ is Binomial with S trials and success probability $e^{-r_{n+1^t}}$. \square

LEMMA 3.5. *For each $t > 0$, $\mathcal{D}_t = \bigcap_{n \in \mathbb{N}_0} \bigcup_{w \in \mathcal{B}_n^t} K_w$.*

PROOF. Suppose first that $x \in \mathcal{D}_t$, so as $[x]_t$ is a singleton. If $\mathcal{E}_w \leq t$ for some $K_w \ni x$, then by definition of X the set $X_{0,t}(K_w)$ would be a single point. By

Lemma 1.4 the set K_w is infinite, and by Lemma 2.7 $K_w \ni x$ implies $K_w \subseteq [x]_t$, so in fact we must have $\mathcal{E}_w > t$ for all $K_w \ni x$. Thus $x \in \bigcap_n \bigcap_{w \in \mathcal{B}_n^t} K_w$.

Similarly, if $x \in \bigcap_n \bigcap_{w \in \mathcal{B}_n^t} K_w$, then $\mathcal{E}_w > t$ for all $K_w \ni x$. If $x \notin \mathcal{D}_t$, then by Lemma 2.7 we would have $X_{0,t}^{-1}(x) = K_{w_1}^{x,0,t}$, which implies that $\mathcal{E}_{w_1} \leq t$, in contradiction to the above. Hence we must have $x \in \mathcal{D}_t$. \square

3.1. *Degeneracy of GWVEs.* We make use of Lemma 3.5 through a result of Jiřina (1976) which characterises the extinction criteria of particular GWVEs. Define

$$(3.2) \quad m_n^t = \mathbb{E}[B_n^t] = S^n e^{-t \sum_0^n r_j},$$

$$(3.3) \quad g^t = \sum_{n=1}^{\infty} \frac{(1 - e^{-r_{n+1}t})^{\mathcal{S}} + \mathcal{S}e^{-r_{n+1}t} - 1}{\mathcal{S}e^{-r_{n+1}t} m_n^t}.$$

Note that $m_n^t \in (0, \infty)$. For all $x > 0$ and $n \in \mathbb{N}$, $n \geq 2$, it holds that $(1 - x)^n + nx - 1 > 0$, and hence $g^t \in (0, \infty]$.

LEMMA 3.6 [Jiřina (1976)]. $\mathbb{P}[\exists n \in \mathbb{N}, B_n^t = 0] < 1$ if and only if both $\inf m_n^t > 0$ and $g^t < \infty$.

PROOF. Lemma 1.1 of Jiřina (1976) tells us that if $\inf_n m_n^t = 0$, then $\mathbb{P}[\exists n \in \mathbb{N}, B_n^t = 0] = 1$. Our offspring distributions are binomial with \mathcal{S} trials, which is sufficient for Theorem 2.3 of Jiřina (1976) to apply. From this we obtain that if $\inf_n m_n^t > 0$, then $\mathbb{P}[\exists n \in \mathbb{N}, B_n^t = 0] = 1$ if and only if $g^t = \infty$. \square

LEMMA 3.7. Suppose that $v \in (0, \infty)$ is such that $\inf_n m_n^v > 0$. Then for all $u \in (0, v)$, $\inf_n m_n^u > 0$ and $g^u < \infty$.

PROOF. Let $\inf_n m_n^v > 0$. Suppose that $\varepsilon > 0$ and that for infinitely many n we have $\frac{1}{n} \sum_0^n r_j > \frac{\varepsilon + \log \mathcal{S}}{v}$. For such n ,

$$m_n^v = \left(\exp \left(\log \mathcal{S} - v \frac{1}{n} \sum_0^n r_j \right) \right)^n \leq \left(\exp \left(\log \mathcal{S} - v \frac{\varepsilon + \log \mathcal{S}}{v} \right) \right)^n = e^{-\varepsilon n}.$$

This may not occur for infinitely many n since $\inf m_n^v > 0$. So, we may assume that both $\inf_n m_n^v > 0$, and

$$(3.4) \quad \limsup_n \frac{1}{n} \sum_0^n r_j \leq \frac{\log \mathcal{S}}{v}.$$

Let $u \in (0, v)$. By the above, for some $\varepsilon > 0$ we have $0 < u \limsup_n \frac{1}{n} \sum_0^n r_j < \log \mathcal{S} - \varepsilon$. Hence, there exists $N \in \mathbb{N}$ (dependent on ε) such that for all $n > N$,

$0 < u \frac{1}{n} \sum_0^n r_j < \log \mathcal{S} - \varepsilon$. Thus

$$m_n^u = \left(\exp \left(\log \mathcal{S} - u \frac{1}{n} \sum_0^n r_j \right) \right)^n \geq (\exp(\varepsilon))^n,$$

so clearly $\inf m_n^u > 0$. Also,

$$g^u \leq \sum_1^\infty \frac{\mathcal{S}}{\mathcal{S} e^{-r_{n+1} u} m_n^u} = \mathcal{S} \sum_1^\infty \frac{1}{m_{n+1}^u} \leq \mathcal{S} \sum_1^\infty \frac{1}{(e^\varepsilon)^{n+1}} < \infty$$

as required. \square

LEMMA 3.8. *There exist $C_1, C_2 \in (0, \infty)$ (dependent only upon \mathcal{S}) such that for all $t > 0$,*

$$C_1 \sum_n \frac{e^{-r_{n+1} t}}{m_n^t} \leq g^t \leq C_2 \sum_n \frac{e^{-r_{n+1} t}}{m_n^t}.$$

PROOF. Let $f_n : (0, \infty) \rightarrow (0, \infty)$ be the function $f_n(x) = (1-x)^n + nx - 1$. It is elementary to show that there exists $C_1, C_2 \in (0, \infty)$ (dependent only upon \mathcal{S}) such that for all $x \in (0, 1]$, $C_1 x^2 \leq f_n(x) \leq C_2 x^2$. Since $g^t = \sum_n \frac{f_n(e^{-r_{n+1} t})}{\mathcal{S} e^{-r_{n+1} t} m_n^t}$, the stated result follows. \square

LEMMA 3.9. *For each (r_n) , precisely one of the following three cases occurs:*

1. $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_0^n r_j = \infty$ if and only if $\inf m_n^t = 0$ for all $t > 0$.
2. $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_0^n r_j = 0$ if and only if $\inf m_n^t > 0$ for all $t > 0$.
3. $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_0^n r_j \in (0, \infty)$ if and only if both (1) if $t < t_0$, then $\inf_n m_n^t > 0$ and (2) if $t \geq t_0$ and $\inf m^{t_0} > 0$ then $g^{t_0} = \infty$.

PROOF. Since $\mathcal{L} = \limsup_n \frac{1}{n} \sum_0^n r_j$ exists in $[0, \infty]$ precisely one of $\mathcal{L} = 0$, $\mathcal{L} = \infty$ and $\mathcal{L} \in (0, \infty)$ occurs. We give each case in turn.

CASE 1: Suppose that $\limsup_n \frac{1}{n} \sum_0^n r_j = \infty$. For any $t > 0$, we can pick a subsequence (r_{i_n}) of (r_n) such that for all n , $\frac{1}{i_n} \sum_0^{i_n} r_j > \frac{\log \mathcal{S} + 1}{t}$. Hence $m_{i_n}^t \leq (\exp(-1))^{i_n}$ for all n , and since $i_n \rightarrow \infty$ it follows that $\inf_n m_n^t = 0$.

Conversely, if $\limsup_n \frac{1}{n} \sum_0^n r_j = C < \infty$, then for $t = \frac{\log \mathcal{S}}{2C} > 0$ we note that

$$m_n^t = \left(\exp \left(\log \mathcal{S} - \frac{\log \mathcal{S}}{2} \frac{1}{C} \frac{1}{n} \sum_0^n r_j \right) \right)^n.$$

For sufficiently large n , $\frac{1}{n} \sum_0^n r_j \leq \frac{3}{2} C$, and hence for sufficiently large n , $m_n^t \geq (\exp(\frac{1}{3} \log \mathcal{S}))^n$. Hence $\inf_n m_n^t > 0$.

CASE 2: Suppose that $\limsup_n \frac{1}{n} \sum_0^n r_j = 0$, and let $t > 0$. Then, for all sufficiently large n we have $\frac{1}{n} \sum_0^n r_j \leq \frac{1}{2t}$. Hence, for all sufficiently large n , $m_n^t \geq (\exp(\frac{1}{2} \log \mathcal{S}))^n$. Thus $\inf_m m_n^t > 0$.

Conversely, suppose that $\inf_n m_n^t > 0$ for all t . Fixing t , and using the first step of the proof of Lemma 3.7, we obtain from (3.4) that $\limsup_n \frac{1}{n} \sum_0^n r_j \leq \frac{\log \mathcal{S}}{t}$. However, we have $\inf_n m_n^t > 0$ for all $t > 0$, so $\limsup_n \frac{1}{n} \sum_0^n r_j = 0$.

CASE 3: Suppose that $\limsup_n \frac{1}{n} \sum_0^n r_j \in (0, \infty)$. Recall from (1.4) that $t_0 = \frac{\log \mathcal{S}}{\mathcal{L}}$ where $\mathcal{L} = \limsup_n \frac{1}{n} \sum_0^n r_j$. Consider first when $s < t_0$. Then there exists $\varepsilon \in (0, \log |S|)$ such that $s \leq \frac{\log |S| - \varepsilon}{\mathcal{L}}$. Hence,

$$m_n^s \geq \left(\exp \left(\log |S| - \frac{\log |S| - \varepsilon}{\mathcal{L}} \frac{1}{n} \sum_1^n r_j \right) \right)^n.$$

There exists $N \in \mathbb{N}$ such that for all $n > N$, $\frac{1}{n} \sum_1^n r_j \leq \frac{\log |S| - \varepsilon/2}{\log |S| - \varepsilon} \mathcal{L}$. Hence, for all $n > N$,

$$m_n^s \geq \left(\exp \left(\log |S| - (\log |S| - \varepsilon) \frac{\log |S| - \varepsilon/2}{(\log |S| - \varepsilon)} \right) \right)^n = (\exp(\varepsilon/2))^n.$$

Thus $m_n^s \rightarrow \infty$, and hence $\inf m_n^s > 0$.

Now consider t_0 itself. Define $a_n \in \mathbb{R}$ by $r_n = \mathcal{L} + a_n$. In this notation $m_n^{t_0} = \exp(-t_0 \sum_1^n a_j)$. We now consider two cases.

Firstly, if $\limsup_n \sum_1^n a_j = \infty$, then it is immediate that $\inf_n m_n^{t_0} = 0$. Since m_n^t is a decreasing function of t (for each fixed $n \in \mathbb{N}_0$), this implies that $\inf_n m_n^t = 0$ for all $t \geq t_0$.

It remains only to consider the case $\limsup \sum_1^n a_j < \infty$, in which case $\inf m_n^{t_0} > 0$. By Lemma 3.8 there exists $C \in (0, \infty)$ such that $g^{t_0} \geq C (\inf_n m_n^{t_0})^{-1} \times \sum_1^n e^{-r_{n+1} t_0}$. Since $\limsup \frac{1}{n} \sum_1^n r_j \in (0, \infty)$, (r_n) has a subsequence (r_{i_n}) such that $\limsup_n r_{i_n} < \infty$. Hence $g^{t_0} = \infty$. By Lemma 3.7 this implies that $\inf_n m_n^t = 0$ for all $t > t_0$, which completes the proof. \square

3.2. *Continuity.* In this section we establish that various aspects of our model are, in some sense, continuous across time.

For $s < t$, define $\mathcal{F}_{s,t} = \sigma(M_{(s,t)})$. Recall that in Section 1.7 we showed that the GWVE B^t is equivalent to an inhomogeneous percolation on the tree W_* . For a possibly random \mathcal{F} -measurable time $t \in [0, \infty)$, let $\mathcal{E}^t = \inf\{s > t; M_{(t,s] \times \{w\}} \neq \emptyset\}$. In words, this is the first time after t at which K_w sees a coalescence event. For each $w \in W_*$, $s \in (0, \infty)$ and possibly random \mathcal{F} -measurable time t , define

$$Q_{w,t} = \{\exists \text{ a sequence } (i_n)_{n \in \mathbb{N}} \subseteq S \text{ such that } \forall m \in \mathbb{N}, \mathcal{E}_{wi_1 \dots i_m} > t\},$$

$$Q_{w,s}^t = \{\exists \text{ a sequence } (i_n)_{n \in \mathbb{N}} \subseteq S \text{ such that } \forall m \in \mathbb{N}, \mathcal{E}_{wi_1 \dots i_m}^t > s\},$$

$$R_w^t = \Omega \setminus \bigcup_{s>0} (Q_{w,s}^t).$$

When we wish to use a random time t in the above definition we will say so explicitly. For the remainder of this section, the symbol t is used only for deterministic times.

In the language of percolation, $Q_{w,s}$ is the event that $w \in W_*$ is connected to infinity at time t . The set $Q_{w,s}^t$ is the event that a connection between w and infinity that exists at time t will continue to exist until (at least) time $t + s$. The set R_w^t is the event that any connection between w and infinity which might exist at time t will be instantaneously disconnected immediately after time t .

LEMMA 3.10. *Let $t \in [0, \infty)$ and $w \in W_*$. Then $\mathbb{P}[R_w^t]$ is either 0 or 1.*

PROOF. If $s_1 \leq s_2$, then $Q_{w,s_2}^t \subseteq Q_{w,s_1}^t$. Thus, for all $N \in \mathbb{N}$ we have $R_w^t = \Omega \setminus \bigcup_{n \geq N} Q_{w,1/n}^t$. Noting that $Q_{w,s}^t$ is $\mathcal{F}_{(s,t+s]}$ measurable, we obtain that R_w^t is $\mathcal{F}_{t,t+1/N}$ measurable for all $N \in \mathbb{N}$. The stated result then follows from the Kolmogorov zero-one law. \square

LEMMA 3.11. *The function $t \mapsto \phi(t) = \mathbb{P}[\forall n \in \mathbb{N}_0, B_n^t \neq 0]$ is strictly monotone decreasing over $[0, \infty)$. Further, ϕ is left continuous over $t \in [0, \infty)$. If $\phi(s) > 0$, then ϕ is right continuous on $[0, s)$.*

PROOF. Note that $B_n^t \subseteq B_n^s$ for all $s \leq t$; it follows immediately that $\phi(t)$ is decreasing. The time at which clock \mathcal{E}_w ring has a continuous distribution on $[0, \infty)$, so for all $[a, b] \subseteq [0, \infty)$ there is positive probability of having $\mathcal{E}_w \in [a, b]$. It follows from this that $\phi(t)$ is strictly decreasing.

For continuity, note that $\phi(t) = \lim_n \mathbb{P}[B_n^t \neq 0]$, which is a decreasing limit as $n \rightarrow \infty$. Each \mathcal{E}_w has continuous distribution, so the definition of B^t implies that the function $t \mapsto \mathbb{P}[B_n^t \neq 0]$ is continuous in t . Thus $\phi(t)$ is upper semicontinuous. Since $\phi(t)$ is also decreasing, $\phi(t)$ is left continuous on $[0, \infty)$.

Let $0 \leq t < s$ and be such that $\phi(s) > 0$. In order to show that ϕ is right continuous at t , we must prove that the event

$$(3.5) \quad \{\forall n \in \mathbb{N}_0, B_n^t \neq 0 \text{ and } \forall u > t \exists n \in \mathbb{N}_0, B_n^u = 0\}$$

has probability zero. Note that

$$(3.5) = Q_{\emptyset,t} \cap \left(\bigcap_{u>t} (\Omega \setminus Q_{\emptyset,u}) \right) = Q_{\emptyset,t} \cap \left(\bigcap_{u>0} (\Omega \setminus Q_{\emptyset,t+u}) \right)$$

$$(3.6) \quad = Q_{\emptyset,t} \cap \left(\Omega \setminus \bigcup_{u>0} Q_{\emptyset,t+u} \right) \subseteq Q_{\emptyset,t} \cap \left(\Omega \setminus \bigcup_{u>0} Q_{\emptyset,u}^t \right)$$

$$= Q_{\emptyset,t} \cap R_{\emptyset}^t.$$

Suppose (3.5) has positive probability. Then by (3.6) we have $\mathbb{P}[R_{\emptyset}^t] > 0$, which by Lemma 3.10 implies that $\mathbb{P}[R_{\emptyset}^t] = 1$. By the time homogeneity of our model

this means that also $\mathbb{P}[R_\emptyset^0] = 1$. Hence $\mathbb{P}[Q_{\emptyset,u}^0] = \mathbb{P}[Q_{\emptyset,u}] = 0$ for all $u > 0$, which means that B^u is almost surely degenerate for all $u > 0$, in contradiction to our hypothesis that $\phi(s) > 0$. So in fact $\mathbb{P}[A] = 0$, which completes the proof. \square

LEMMA 3.12. *If $\limsup_n \frac{1}{n} \sum_0^n r_j \in (0, \infty)$, then $\mathbb{P}[\pi < t_0] = 1$.*

PROOF. Suppose that $\limsup_n \frac{1}{n} \sum_0^n r_j \in (0, \infty)$. By combining Lemmas 3.7 and 3.8 we have that $\mathbb{P}[\forall n \in \mathbb{N}_0, B_n^{t_0} \neq 0] = 0$. Hence, by the left continuity proved in Lemma 3.11, for all $\varepsilon > 0$ there is some $\delta > 0$ such that $\mathbb{P}[\forall n \in \mathbb{N}_0, B_n^{t_0-s} \neq 0] < \varepsilon$ for all $s \in [0, \delta)$. Hence $\pi < t_0$ almost surely. \square

Let (\mathcal{G}_t) denote the usual augmented filtration [see Section II.67 of Rogers and Williams (2000)] of $\mathcal{F}_t = \sigma(M_{[0,s]}; s \leq t)$. Let $\pi = \inf\{t > 0; \exists n \in \mathbb{N}, B_n^t = 0\}$, and note that, since (\mathcal{G}_t) is right continuous, π is a (\mathcal{G}_t) stopping time.

LEMMA 3.13. *If $\limsup_n \frac{1}{n} \sum_0^n r_j < \infty$, then $\mathbb{P}[\exists n \in \mathbb{N}_0, B_n^\pi = 0] = 1$.*

PROOF. Let $\limsup_n \frac{1}{n} \sum_0^n r_j < \infty$, and suppose (for a contradiction) that $\mathbb{P}[\forall n \in \mathbb{N}_0, B_n^\pi \neq 0] > 0$. By definition of π , almost surely for all $t > \pi$ there is some $n \in \mathbb{N}_0$ such that $B_n^t = 0$. Thus

$$(3.7) \quad \{\forall n \in \mathbb{N}_0, B_n^\pi \neq 0 \text{ and } \forall u > \pi \exists n \in \mathbb{N}_0 B_n^u = 0\}$$

has positive probability. The same rearrangement as was used in (3.6), with π in place of t , shows that $(3.7) \subseteq Q_{\emptyset,\pi} \cap R_\emptyset^\pi$. Hence $\mathbb{P}[R_\emptyset^\pi] > 0$.

By the strong Markov property of the time-homogeneous process M at the (\mathcal{G}_t) stopping time π , we have that R_\emptyset^π and R_\emptyset^0 have the same distribution, hence also $\mathbb{P}[R_\emptyset^\pi] > 0$. By Lemma 3.10 we thus have $\mathbb{P}[R_\emptyset^0] = 1$. Thus $\mathbb{P}[Q_{\emptyset,u}^0] = \mathbb{P}[Q_{\emptyset,u}] = 0$ for all $u > 0$, which means that B^u is almost surely degenerate for all $u > 0$.

However, since $\limsup_n \frac{1}{n} \sum_0^n r_j < \infty$, by Lemmas 3.7 and 3.9 there is almost surely some $\varepsilon > 0$ such that $\mathbb{P}[\forall n \in \mathbb{N}_0, B_n^\varepsilon \neq 0] > 0$, so we reach a contradiction and conclude that in fact, $\mathbb{P}[\forall n \in \mathbb{N}_0, B_n^\pi \neq 0] = 0$. \square

3.3. *Dust and GWVEs.* In this section we build on Lemma 3.5 and relate the behaviour of \mathcal{D}_t to the behaviour of B^t .

LEMMA 3.14. *Let $t > 0$. Then $\mathcal{D}_t = \emptyset$ if and only if $\exists n \in \mathbb{N}, B_n^t = 0$.*

PROOF. Fix $t > 0$. Suppose first that for some (random) $n \in \mathbb{N}$, $B_n^t = 0$. Then, by Lemma 3.5, $\mathcal{D}_t = \emptyset$. For the converse, If $B_n^t \neq 0$ for all n , then it is easily seen that there exists a sequence $(w(n))_{n \in \mathbb{N}}$ such that $w(n) \in \mathcal{B}_n^t$ and $K_{w(n)} \supseteq K_{w(n+1)}$. By Lemma 2.4, each subcomplex K_w is a closed subset of K . It follows from this and completeness of K that $\bigcap_{n \in \mathbb{N}} K_{w(n)}$ is nonempty. By Lemma 3.5, $\bigcap_{n \in \mathbb{N}} K_{w(n)} \subseteq \mathcal{D}_t$ so the proof is complete. \square

REMARK 3.15. The argument above has appeared many times in various guises in the random fractals literature; see, for example, Lemma 8 of Durand (2009). However, if K is not completely segregated, then (Lemma 2.6 does not apply so as) K_w might not be closed and $\bigcap_{n \in \mathbb{N}} K_{w(n)}$ could be empty. We address this issue in Section 4.

As a consequence of Lemma 3.14, $\tau = \pi$. The reason for making the distinction between τ and π will become clear in Section 4.1.

LEMMA 3.16. *Let $t > 0$. If $\mathcal{D}_t = \emptyset$, then A_t is finite.*

PROOF. If \mathcal{D}_t is empty, then by Lemma 3.14 there is some $n \in \mathbb{N}_0$ such that $B_n^t = 0$. Let $N = \min\{n \in \mathbb{N}; B_n^t = 0\}$. Since $\mathcal{D}_t = \emptyset$, by Lemma 2.7, for all $x \in K$ we have $[x]_t = K_{w_1^{x,s,t}}$.

Suppose that $|w_1^{x,s,t}| > N + 1$. Then there is some $w \in W_{N+1}$ such that $K_w \supseteq K_{w_1^{x,s,t}}$. Since $|w| > N$ we have $w \notin \mathcal{B}^t$ so there must be some $w' \in W_*$ such that $|w'| \leq |w| < |w_1^{x,s,t}|$ and $K_{w'} \supseteq K_w \supseteq K_{w_1^{x,s,t}}$ with $\mathcal{E}_{w'} \leq t$. Since then $x \in K_{w'}$ and $|w'| \leq |w_1^{x,s,t}|$, this contradicts the definition of $w_1^{x,s,t}$. Hence in fact $|w_1^{x,s,t}| \leq N + 1$.

Therefore, A_t is a subset of $\{K_w; w \in W_{N+1}\}$, which implies that A_t is finite. □

LEMMA 3.17. *Let $t > 0$. Then $\mathbb{P}[\lim_n B_n^t = \infty \text{ or } \exists n, B_n^t = 0] = 1$. Further, almost surely, $\mathcal{D}_t \neq \emptyset$ if and only if $\lim_{n \rightarrow \infty} B_n^t = \infty$.*

PROOF. To prove the first statement, we use the result of Theorem 1 in Jagers (1974), which is a restatement (with minor correction) of a result in Church (1967).

The probability of a individual at stage n in the process B^t having exactly one offspring is given by $p_{n1}^t = S e^{-r_n t} (1 - e^{-r_n t})^{S-1}$. Note that for $a \in [0, 1]$ and $n \geq 1$, $a(1 - a)^n \leq \frac{1}{n+1} (1 - \frac{1}{n+1})^n$. Since $S \geq 2$ we have $p_{n1}^t \leq (1 - 1/S)^{S-1} < 1$. Hence $\sum_n (1 - p_{n1}^t) = \infty$, and from Jagers (1974) we have $\mathbb{P}[\lim_n B_n^t = \infty \text{ or } \exists n \geq \mathcal{N}, B_n^t = 0] = 1$. It follows immediately from this and Lemma 3.14 that $\lim_n B_n^t = \infty$ is almost surely equivalent to $\mathcal{D}_t \neq \emptyset$. □

LEMMA 3.18. *Let $t > 0$. If $\sum r_n < \infty$, then $\mathbb{P}[\mathcal{D}_t = \emptyset \text{ or } \lambda(\mathcal{D}_t) > 0] = 1$.*

PROOF. The process $n \mapsto B_n^t / \mathbb{E}[B_n^t]$ is a discreet parameter, nonnegative martingale. By the martingale convergence theorem, there is some random variable L^t such that $\frac{B_n^t}{\mathbb{E}[B_n^t]} \rightarrow L^t$ almost surely.

Recall that in (3.2) we gave a formula for $\mathbb{E}[B_n^t]$. Since $\sum r_n < \infty$,

$$(3.8) \quad \mathbb{E}[B_n^t] \geq S^n \left(\mathbb{E}[B_0^t] \exp\left(-t \sum_0^n r_j\right) \right) = C S^n,$$

where $C = C(t) > 0$. In the language of D’Souza and Biggins (1992), (3.8) means that B^t is uniformly supercritical. Since the offspring distribution of B^t is uniformly bounded (by S), Theorem 2 of D’Souza and Biggins (1992) applies. In our notation this means that

$$(3.9) \quad \{B_n^t \rightarrow \infty\} = \{L^t > 0\}.$$

Now, suppose $\omega \in \mathcal{A}$ and that $\mathcal{D}_t \neq \emptyset$. By Lemma 3.5, for all $n \in \mathbb{N}$ we have $B_n^t = |\mathcal{B}_n^t| \geq 1$. By the first part of Lemma 3.17 it follows that (almost surely) $B_n^t \rightarrow \infty$ as $n \rightarrow \infty$. By (3.9), $\lim_n \frac{B_n^t}{\mathbb{E}[B_n^t]} > 0$. From this and (3.8), $\liminf_n \frac{B_n^t}{C S^n} > 0$, where $\liminf_n \frac{B_n^t}{C S^n}$ could potentially be infinite. In fact, though, $B_n^t \leq B_0^t S^n \leq S^n$ so $\liminf_n \frac{B_n^t}{C S^n}$ is finite. We write $l = \liminf_{n \rightarrow \infty} \frac{B_n^t}{C S^n}$ where $l \in (0, \infty)$ (note l is random). Then there exists $N \in \mathbb{N}$ such that for all $n > N$, $\frac{B_n^t}{C S^n} \geq l/2$. So for all $n > N$ we have $B_n^t \geq \frac{Cl}{2} S^n$.

Note that the sets $\bigcup_{w \in \mathcal{B}_n^t} K_w$ are decreasing as n increases. By Lemma 3.5, $\lambda(\mathcal{D}_t) = \lim_n \lambda(\bigcup_{w \in \mathcal{B}_n^t} K_w)$. Recall that $\lambda(K_w) > 0$ and $\lambda(K_w)$ depends only on $|w|$ [by Lemma 1.4 and $(\mathcal{K}3)$, resp.]. Hence $\lambda(K_w) = \frac{\lambda(K)}{S^{|w|}}$. By $(\mathcal{K}1)$, $\lambda(\bigcup_{w \in \mathcal{B}_n^t} K_w) = |\mathcal{B}_n^t| \lambda(K_w) = \frac{B_n^t}{S^n} \lambda(K)$. Thus, for $n > N$, $\lambda(\bigcup_{w \in \mathcal{B}_n^t} K_w) \geq Cl \lambda(K) > 0$. The result follows by Lemma 3.5. \square

LEMMA 3.19. *If $\limsup \frac{1}{n} \sum_0^n r_j = 0$, then for all $t \in \mathbb{R}$, $\mathbb{P}[\tau > t] > 0$.*

PROOF. If $\limsup \frac{1}{n} \sum_0^n r_j = 0$, then by Lemma 3.9 we have $\inf_n m_n^t > 0$ for all $t > 0$. By Lemma 3.7 we thus have $g^t < \infty$ for all $t > 0$, and by Lemma 3.7 we thus have $\mathbb{P}[\forall n \in \mathbb{N}, B_n^t \neq 0] > 0$. By Lemma 3.14 we thus have $\mathbb{P}[\mathcal{D}_t > 0] > 0$ for all $t \in \mathbb{R}$. Since $\mathbb{P}[\mathcal{D}_t > 0] > 0$ for all $t \in \mathbb{R}$ we also have $\mathbb{P}[\tau > t] > 0$ for all $t \in \mathbb{R}$. \square

3.4. *Proofs of Theorems 2, 3 and Corollary 4.* Let us begin by proving Theorem 3. Note that the criteria given for our five phases in terms of \mathcal{S} and (r_n) assign possible each choice of \mathcal{S} and (r_n) to precisely one phase. Therefore, it suffices to show that the criteria for each phase are sufficient. Let us begin by covering the supercritical phase.

PROOF OF THEOREM 3. *Supercritical.* Suppose that $\limsup \frac{1}{n} \sum_0^n r_j = \infty$. We need to show that $\tau = 0$ almost surely. By Lemma 3.9, $\inf_n m_n^s = 0$ for all $s > 0$, and thus from Lemma 3.6 we have that $\mathbb{P}[\exists n \in \mathbb{N}, B_n^s = 0] = 1$. By Lemma 3.14 $\mathbb{P}[\mathcal{D}_s = \emptyset] = 1$ for all $s > 0$ and by (1.2) we have $\mathbb{P}[\forall s > 0, \mathcal{D}_s = \emptyset] = 1$. Hence $\mathbb{P}[\tau = 0] = 1$.

We will now give the arguments for the four remaining phases. Note that, in the lower/upper subcritical and semicritical phases, Lemma 3.19 tells us that $\mathbb{P}[\tau >$

$0] = 1$ and $\mathbb{P}[\tau > t] > 0$ for all $t \in \mathbb{R}$. The behaviour of τ in the critical phase is discussed below.

Lower subcritical. Suppose that $\sum \mathcal{S}^n r_n < \infty$, and consider $t \in (0, \tau)$. Hence, by definition of M , for each (deterministic) $s < \infty$ only finitely many coalescence events occur during $[0, s]$. Since $t < \tau < \infty$, almost surely only finitely many coalescence events have occurred during $[0, t]$. By Lemma 2.7, A_t is a finite set, which combined with Lemma 3.2 implies that $\lambda(\mathcal{D}_t) > 0$.

Upper subcritical. Suppose that $\sum \mathcal{S}^n r_n = \infty$ and $\sum r_n < \infty$. Consider $t \in (0, \tau)$. If A_t was finite, then since $\mathcal{D}_t \neq \emptyset$, by Lemma 3.2 the set \mathcal{D}_t must contain a subcomplex of K ; but Lemma 3.3 implies that this is not the case. Hence in fact A_t must be infinite. By Lemma 3.18, $\lambda(\mathcal{D}_t) > 0$ almost surely.

Semicritical. Suppose that $\sum r_n = \infty$ and $\limsup \frac{1}{n} \sum_0^n r_j = 0$. Consider $t \in (0, \tau)$. The same argument applies here as given above in the upper subcritical case to show that A_t is infinite. However, in this case Lemma 3.1 tells us that $\mathbb{P}[\lambda(\mathcal{D}_s) = 0] = 1$ for fixed $s \in (0, \infty)$. By (1.2), in fact $\mathbb{P}[\forall s > 0, \lambda(\mathcal{D}_s) = 0]$.

Critical. Suppose that $\limsup \sum_0^n r_j \in (0, \infty)$ and write $\mathcal{L} = \limsup \sum_0^n r_j$. Consider first if $s < t_0$. Then by combining Lemmas 3.6 and 3.8 we have that $\mathbb{P}[\exists n \in \mathbb{N} B_n^s = 0] < 1$, so as by Lemma 3.17, $\mathbb{P}[\mathcal{D}_s \neq \emptyset] > 0$ for $s < t_0$. Hence $\mathbb{P}[\tau > s] \in (0, 1)$ for all $s < t_0$.

Similarly, by Lemma 3.6, $\mathbb{P}[\exists n \in \mathbb{N}, B_n^{t_0} = 0] = 1$ and hence by Lemma 3.17, $\mathbb{P}[\mathcal{D}_{t_0} = \emptyset] = 1$. By (1.2) we then have $\mathbb{P}[\mathcal{D}_s = \emptyset] = 1$ for all $s \geq t_0$.

By Lemma 3.14 we have $\mathbb{P}[\tau = \pi] = 1$, so by Lemma 3.12 we have $\mathbb{P}[\tau < t_0] = 1$. Now consider $t \in (0, \tau)$. The same argument as in the semicritical case tells us that A_t must be infinite and that $\mathbb{P}[\forall s > 0, \lambda(\mathcal{D}_s) = 0] = 1$. This completes the proof of Theorem 3. \square

We now prove Theorem 2.

PROOF OF THEOREM 2. The first part of the statement of Theorem 2 is a trivial consequence of Lemma 3.16; if $t > \tau$, then $\mathcal{D}_t = \emptyset$ almost surely and, by Lemma 3.16, A_t is almost surely finite.

For the second statement, by Theorem 3, we have that $\mathbb{P}[\tau = 0] = 1$ if and only if our model is supercritical. So, assume our model is not supercritical. By Theorem 3 there is some (deterministic) $\delta > 0$ such that $\mathbb{P}[\mathcal{D}_\delta \neq \emptyset] > 0$, so as by Lemmas 3.11 and 3.14 we have that $t \mapsto \mathbb{P}[\mathcal{D}_t \neq \emptyset]$ is continuous on $[0, \delta)$. Note that for all $\varepsilon > 0$,

$$\mathbb{P}[\tau = 0] \leq \mathbb{P}[\tau \leq \varepsilon] = \mathbb{P}[\mathcal{D}_\varepsilon = \emptyset] = 1 - \mathbb{P}[\mathcal{D}_\varepsilon \neq \emptyset],$$

which by continuity tends to $1 - \mathbb{P}[\mathcal{D}_0 \neq \emptyset] = 0$ as $\varepsilon \downarrow 0$. Hence $\tau > 0$ almost surely. \square

We now prove Corollary 4.

PROOF OF COROLLARY 4. By Lemma 3.14 we have $\tau = \pi$ almost surely, and combining this with Lemma 3.13, we obtain that $\mathcal{D}_\tau = 0$ almost surely, provided $\limsup_n \frac{1}{n} \sum_0^n r_j < \infty$. By Theorem 3, if our model is not supercritical, then $\limsup_n \frac{1}{n} \sum_0^n r_j < \infty$, which completes the proof. \square

4. The case $\mathcal{O} \neq \emptyset$. We now describe the modifications required to prove our results in the case where K is a segregated space but potentially not a completely segregated space (i.e., we allow $\mathcal{O} \neq \emptyset$). Essentially, the difference in this case is that Lemma 2.4 breaks down. We used Lemma 2.4 in precisely two places, namely the proofs of Lemmas 2.6 and 3.14. It is these two lemmas which we will “repair” to deal with the case $\mathcal{O} \neq \emptyset$. We will replace them, respectively, with the following.

LEMMA A. *Almost surely, for all $0 < s < t$, all $x \in K$ and all $n \in \mathbb{N}$, if $(E_m^{x,s,t})$ is infinite, then $\lim_{m \rightarrow \infty} p_m^{x,s,t} \in K_{w_n}^{x,s,t}$.*

LEMMA B. *Almost surely, for all $t \in [0, \infty)$, $\mathcal{D}_t = \emptyset$ if and only if $B_n^t = 0$ for some $n \in \mathbb{N}_0$.*

Lemmas A and B are the same (resp.) as Lemmas 2.6 and 3.14, except for the presence in both cases of the “almost surely”. Their proofs are given in Section 4.1.

Let \mathcal{N} be the null set of Ω on which the “almost surely” in Lemma A does not hold. This is the null set (that we mentioned in Section 2) on which we wish to define X differently in the case $\mathcal{O} \neq \emptyset$. So, fix some point $x^* \in K$ and from now, for all $s < t$ define

$$(4.1) \quad X_{s,t} = \begin{cases} \text{as in (2.1),} & \text{if } \omega \in \Omega \setminus \mathcal{N}, \\ x^*, & \text{if } \omega \in \mathcal{N}. \end{cases}$$

In words, when $\omega \in \mathcal{N}$ the flow instantaneously (and at all times) moves all the particles to the point x^* . Thus $\tau = 0$ on \mathcal{N} .

The arguments in the proof of Theorem 1, using Lemma A in place of Lemma 2.6, work as before so long as $\omega \in \Omega \setminus \mathcal{N}$. On the other hand, for $\omega \in \mathcal{N}$ it is readily seen that (4.1) trivially implies the conclusions of Theorem 1. Thus Theorem 1 remains true.

Essentially the same principle applies to using Lemma B in place of Lemma 3.14; the results in Section 3 that were stated for all $\omega \in \Omega$ and rely upon Lemmas 2.6 and/or 3.14 now hold only almost surely. It is a simple matter to check that this is sufficient to make the proof of Theorems 2, 3 and Corollary 4 go through as before.

4.1. *Proofs of Lemmas A and B.* Recall Remark 2.5, that $(\mathcal{H}4)$ was immediate if $\mathcal{O} = \emptyset$. In the case $\mathcal{O} \neq \emptyset$ it is $(\mathcal{H}4)$ that fills the gap, as the following arguments show.

We first prove Lemma A.

PROOF OF LEMMA A. Let $\mathcal{E} = \{(x, s, t) \in K \times (0, \infty)^2; |E^{x,s,t}| = \infty\}$ and define an equivalence relation on \mathcal{E} by $(x, s, t) \sim (x', s', t')$ if and only if $E^{x,s,t} = E^{x',s',t'}$. Let $[x, s, t]$ denote the equivalence class of (x, s, t) in \mathcal{E} under \sim . In view of this definition, we write $E[x, s, t] = E^{x,s,t}$ and similarly for $(u_m^{x,s,t}, w_m^{x,s,t}, p^{x,s,t})$. It follows from $(\mathcal{N}1)$ and Definition 2.1 that $E_1^{x,s,t} = E_1^{x',s',t'}$ if and only if $(x, s, t) \sim (x', s', t')$.

Let $(\hat{s}_k, \hat{t}_k)_{k \in \mathbb{N}}$ be a countable dense subset of $\{(s, t) \in \mathbb{R}^2; s \leq t\}$, and let $(\hat{x}_j)_{j \in \mathbb{N}}$ be such that for all $w \in W_*$ there is some $x_j \in K_w$. Note that since $E_1^{x,s,t} < E_2^{x,s,t}$ for all x, s, t , for each $(x, s, t) \in \mathcal{E}$ there is some k such that $E^{x,s,t} = E^{x,\hat{s}_k,\hat{t}_k}$. For fixed k , from Definition 2.1 we have $E_1^{x,\hat{s}_k,\hat{t}_k} = E_1^{x',\hat{s}_k,\hat{t}_k}$ if and only if $w_1^{x,\hat{s}_k,\hat{t}_k}$; hence for all $x \in K$ and all $k \in \mathbb{N}$ there is some $j \in \mathbb{N}$ such that $E^{x,s,t} = E^{\hat{x}_j,\hat{s}_k,\hat{t}_k}$. Thus

$$(4.2) \quad \left\{ \exists (x, s, t) \in \mathcal{E}, \exists n \in \mathbb{N}, \lim_{m \rightarrow \infty} p_m^{x,s,t} \notin K_{w_n}^{x,s,t} \right\}$$

$$(4.3) \quad = \bigcup_{j,k,n \in \mathbb{N}} \{ p_m^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \notin K_{w_n}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \}.$$

For fixed j, k and n , each $p_m^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]}$ is sampled (independently) from within $K_{w_m}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]}$ according to λ . So, by $(\mathcal{N}4)$, with probability $1/S$ we have

$$(4.4) \quad p_m^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \in K_{w_{m+1}}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \subseteq \overline{K}_{w_{m+1}}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \subseteq K_{w_m}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]}.$$

By the Borel–Cantelli lemma, (4.4) occurs for infinitely many $m \in \mathbb{N}$, with probability one. So, for each $n \in \mathbb{N}$ almost surely we can find some $m \geq n$ for which (4.4) holds, implying that $\lim_{r \rightarrow \infty} p_r^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \in \overline{K}_{w_{m+1}}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \subseteq K_{w_n}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]}$. Thus $\mathbb{P}[p_m^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]} \notin K_{w_n}^{[\hat{x}_j,\hat{s}_k,\hat{t}_k]}] = 0$. Combining this with (4.3) we have that (4.2) is a \mathbb{P} -null, which completes the proof. \square

PROOF OF LEMMA B. We give the proof of Lemma B in two parts, the first of which is the following lemma. Thanks to Lemma A, all the results stated in Sections 3–3.2 are available to us, with the caveat that results which previously held for all $\omega \in \Omega$ now hold only almost surely.

LEMMA 4.1. $\mathbb{P}[\forall t \in \mathbb{Q}, \text{ if } \forall n \in \mathbb{N} B_n^t \neq \emptyset \text{ then } \mathcal{D}_t \neq \emptyset] = 1$.

PROOF. Since \mathbb{Q} is countable, it suffices to prove the result for a single fixed $t \in \mathbb{Q}$. In fact, we can prove for any fixed $t \in \mathbb{R}$, as follows. Suppose that $Q = \{\forall n \in \mathbb{N}, B_n^t \neq \emptyset\}$ has positive probability (and note that if this has probability

zero, then there is nothing to prove). Let \mathbb{P}_Q denote the conditional measure of \mathbb{P} on the event Q .

By Lemma 4.10 of Lyons (1992), the distribution of B^t under \mathbb{P}_Q is that of a GWVE with n th stage offspring distribution given by

$$(4.5) \quad f_n^*(s) = \frac{f_n(q_n + (1 - q_n)s) - q_{n-1}}{1 - q_{n-1}},$$

where $q_n = \mathbb{P}[\exists m \geq n B_m^t = 0 | B_n^t = 1]$, and $f_n(s)$ is the generating function of the offspring distribution in stage n of B^t . Since $q_n = \lim_{N \rightarrow \infty} f_n \circ \dots \circ f_{n+N}(0)$, we have $f_n^*(0) = 0$ which in turn means that, under \mathbb{P}_Q , each individual in B^t has at least one child.

This tells us the behaviour of $|\mathcal{B}^t|$ under \mathbb{P}_Q , but we need a little extra work to describe \mathcal{B}^t itself. The clocks $\{\mathcal{E}_w; w \in W_n\}$ are all independent and identically distributed (under \mathbb{P}). Therefore, under \mathbb{P}_Q , the number k of elements of W_n which are in \mathcal{B}_n^t is given by (4.5), but precisely which such elements of W_n is given by the uniform distribution on the set of subsets of W_n of size k .

In view of this description, define a sequence $(w_n)_{n \in \mathbb{N}_0}$ as follows. Set $w_0 = \emptyset$ and note that $w_0 \in B_0^t$ \mathbb{P}_Q -almost surely. Now, if w_n is defined and $w_n \in B_n^t$ \mathbb{P}_Q -almost surely, enumerate the set of children of w_n in \mathcal{B}_{n+1}^t as $C_{n+1} \subseteq W_{n+1}$. Independently of all else, sample w_{n+1} uniformly from C_{n+1} .

By our description of B^t under \mathbb{P}_Q , using (K4) we have that with probability (at least) $1/S$, $\overline{K}_{w_{n+1}} \subseteq K_{w_n}$. The offspring distributions of each individual in \mathcal{B}_n^t are independent, hence by the Borel–Cantelli lemma there is almost surely an infinite subsequence (\tilde{w}_n) of (w_n) such that $\overline{K}_{\tilde{w}_{n+1}} \subseteq K_{\tilde{w}_n}$ for all n . Thus, $\bigcap_n K_w = \bigcap_n K_{\tilde{w}_n}$ is nonempty. It follows from this and the “almost sure” replacement of Lemma 3.5 (see our comments before Lemma 4.1) that, almost surely, $\mathcal{D}_t \supseteq \bigcap_n K_{\tilde{w}_n} \neq \emptyset$. \square

Our task now is to upgrade Lemma 4.1 into Lemma B. By Lemma 4.1, almost surely, for all $q \in Q$, $\tau < q$ if and only if $\pi < q$. Since Q is dense in \mathbb{R} it follows immediately that $\tau = Q$ almost surely. Recall that $B_n^t \leq B_n^s$ for $s \leq t$. It follows from this and Lemma 3.13 that, almost surely, for all $t \in \mathbb{R}$, $\exists n \in \mathbb{N}_0 B_n^t = 0$ if and only if $t \geq \pi$. Hence, by (the almost sure version of) Lemma 3.5 we have $\mathcal{D}_\pi = \emptyset$. Since $\tau = \pi$ almost surely we have $\mathcal{D}_\tau = \mathcal{D}_\pi$ almost surely and thus $\mathbb{P}[\mathcal{D}_\tau = \emptyset] = 1$. Therefore, using (1.2) we have that almost surely, $\mathcal{D}_t = 0 \Leftrightarrow t \geq \tau \Leftrightarrow t \geq \pi \Leftrightarrow \exists n \in \mathbb{N}_0, B_n^t = 0$. \square

5. The Hausdorff dimension of the dust. It is natural to ask further questions about the nonempty null dust in the semicritical and critical phases. According to Lemma 3.5 and our comments following (1.5), \mathcal{D}_t is a random fractal. In fact, \mathcal{D}_t belongs to the large class of random fractals which are, in some sense, stochastic generalizations of iterated function systems (IFSs). See Falconer (2003) for an introduction to fractal geometry and IFSs.

IFSs have been generalised in many directions, both deterministically and stochastically, and formulas for the Hausdorff dimension of the corresponding attractors have been obtained in increasing generality; see Durand (2009), Mörters (2010) and the references therein. Generality sufficient to cope with \mathcal{D}_t , at least in terms of Hausdorff dimension, seems to have been reached only recently and a result corresponding to the Hausdorff measure of \mathcal{D}_t does not seem to be known.

Let $\|\cdot\|$ denote the Euclidean norm on \mathbb{R}^d , and let \mathcal{L}^d denote d -dimensional Lebesgue measure. Let A° denote the (topological) interior of the set A , and let the diameter of A be given by $\text{diam}(A) = \sup\{\|x - y\|; x, y \in E\}$. Recall that a similarity f is a function between subsets of \mathbb{R}^d such that for some $\eta \in (0, \infty)$ and all x, y , $\|f(x) - f(y)\| = \eta\|x - y\|$. We write $\eta = \text{lip}(f)$. Recall also that $\text{dim}_{\mathcal{H}}(A)$ denotes the Hausdorff dimension of A (for $A \subseteq \mathbb{R}^d$ this is with respect to the metric $(x, y) \mapsto \|x - y\|$).

In order to link our results to those of Durand (2009), we must impose some extra assumptions on K .

DEFINITION 5.1. We say K is D -compatible if $K \subseteq \mathbb{R}^d$, $D_K = \|\cdot\|$, and:

1. For all $w \in W_*$, K_w is compact.
2. For all $w \in W_*$ and $i \in S$ there exists a similarity $f^{(w,i)} : K_w \rightarrow K_{wi}$.
3. There exists $\varepsilon, \varepsilon' \in (0, 1)$ and a sequence $(l_n) \subseteq [\varepsilon, \varepsilon']$ such that for all $w \in W$, $\text{lip}(f^{(w,i)}) = l_{|w|}$.
4. There exists $\kappa > 0$ such that for all $w \in W_*$, $\mathcal{L}^d(K_w^\circ) \geq \kappa \text{diam}(K_w)^d$.

THEOREM 5. Suppose that K is D -compatible and that $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$. Let $\mathcal{L} = \limsup_n \frac{1}{n} \sum_1^n r_j$ and $\mathcal{S} = \limsup_n \frac{1}{n} \sum_1^n (-\log l_n)$. Conditionally on $\{\mathcal{D}_t \neq \emptyset\}$,

$$\text{dim}_{\mathcal{H}}(\mathcal{D}_t) = \left(\frac{\log |S| - t\mathcal{L}}{\mathcal{S}} \right) \vee 0.$$

By setting $r_n = 0$, it follows that $\text{dim}_{\mathcal{H}}(K) = \frac{\log |S|}{\mathcal{S}}$.

PROOF. For each $s > 0$ and $n \in \mathbb{N}$ let $\alpha_{s,n}^t = |S|(l_n)^s e^{-r_{n+1}t}$ and $\rho^t(s) = \liminf_n \frac{1}{n} \sum_{j=1}^n \log \alpha_{s,j}^t$. Using (A1) and the D -compatibility conditions we apply Theorem 1 of Durand (2009), which yields that, if $\mathcal{D}_t \neq \emptyset$, then $\text{dim}_{\mathcal{H}}(\mathcal{D}_t) = \sup\{s \in [0, \infty); \rho^t(s) > 0\}$.

By Theorem 3, if $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$, then $0 \leq \limsup_n \frac{1}{n} \sum_1^n r_j < \infty$. Note also that by 3 of the D -compatibility conditions, $0 \leq -\log \varepsilon' \leq -\log(l_n) \leq -\log \varepsilon < \infty$. A short calculation shows that

$$\rho^t(s) = \log |S| - t \limsup_{n \rightarrow \infty} \left(\frac{1}{n} \sum_1^n r_j \right) - s \limsup_{n \rightarrow \infty} \left(\frac{1}{n} \sum_1^n (-\log l_n) \right).$$

The result follows. \square

Acknowledgment. I would like to thank the referee for several comments that greatly improved the presentation of this article.

REFERENCES

- AGRESTI, A. (1975). On the extinction times of varying and random environment branching processes. *J. Appl. Probab.* **12** 39–46. [MR0365733](#)
- ANGEL, O., BERESTYCKI, N. and LIMIC, V. (2010). Global divergence of spatial coalescents. *Probab. Theory Related Fields* **152** 1–45.
- BARTON, N. H., ETHERIDGE, A. M. and VÉBER, A. (2010). A new model for evolution in a spatial continuum. *Electron. J. Probab.* **15** 162–216. [MR2594876](#)
- BARTON, N. H., ETHERIDGE, A. M. and VÉBER, A. (2013). Modelling evolution in a spatial continuum. *J. Stat. Mech. Theory Exp.* **1** P01002, 38. [MR3036210](#)
- BERESTYCKI, N. (2009). *Recent Progress in Coalescent Theory. Ensaios Matemáticos [Mathematical Surveys]* **16**. Sociedade Brasileira de Matemática, Rio de Janeiro. [MR2574323](#)
- BERESTYCKI, J., BERESTYCKI, N. and LIMIC, V. (2014). A small-time coupling between Λ -coalescents and branching processes. *Ann. Appl. Probab.* **24** 449–475. [MR3178488](#)
- BERTOIN, J. and LE GALL, J.-F. (2000). The Bolthausen–Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Theory Related Fields* **117** 249–266. [MR1771663](#)
- LE GALL, J.-F. and BERTOIN, J. (2003). Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields* **126** 261–288. [MR1990057](#)
- D’SOUZA, J. C. and BIGGINS, J. D. (1992). The supercritical Galton–Watson process in varying environments. *Stochastic Process. Appl.* **42** 39–47. [MR1172506](#)
- BIRKNER, M., BLATH, J., CAPALDO, M., ETHERIDGE, A., MÖHLE, M., SCHWEINSBERG, J. and WAKOLBINGER, A. (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.* **10** 303–325 (electronic). [MR2120246](#)
- BOLTHAUSEN, E. and SZNITMAN, A. S. (1998). On Ruelle’s probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197** 247–276. [MR1652734](#)
- CHURCH, J. D. (1967). Composition limit theorems for probability generating functions. *MRC Techn. Summ. Rept.* **732** 1–41.
- DONNELLY, P. and KURTZ, T. G. (1999). Genealogical processes for Fleming–Viot models with selection and recombination. *Ann. Appl. Probab.* **9** 1091–1148. [MR1728556](#)
- DURAND, A. (2009). Random fractals and tree-indexed Markov chains. *Rev. Mat. Iberoam.* **25** 1089–1126. [MR2590694](#)
- ETHERIDGE, A. M. (2008). Drift, draft and structure: Some mathematical models of evolution. In *Stochastic Models in Biological Sciences. Banach Center Publications* **80** 121–144. Polish Acad. Sci., Warsaw. [MR2433141](#)
- ETHERIDGE, A. (2011). *Some Mathematical Models from Population Genetics. Lecture Notes in Math.* **2012**. Springer, Heidelberg. [MR2759587](#)
- FALCONER, K. J. (1986). Random fractals. *Math. Proc. Cambridge Philos. Soc.* **100** 559–582. [MR0857731](#)
- FALCONER, K. (2003). *Fractal Geometry: Mathematical Foundations and Applications*, 2nd ed. Wiley, Hoboken, NJ. [MR2118797](#)
- FEARN, D. H. (1972). Galton–Watson processes with generation dependence. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. IV: Biology and Health* 159–172. Univ. California Press, Berkeley, CA. [MR0400434](#)
- FREEMAN, N. (2012). The Segregated Λ -coalescent. Ph.D. thesis, Oxford Univ.
- GRAF, S. (1987). Statistically self-similar fractals. *Probab. Theory Related Fields* **74** 357–392. [MR0873885](#)

- GREVEN, A., LIMIC, V. and WINTER, A. (2005). Representation theorems for interacting Moran models, interacting Fisher–Wright diffusions and applications. *Electron. J. Probab.* **10** 1286–1356 (electronic). [MR2176385](#)
- GREVEN, A., DEN HOLLANDER, F., KLIEM, S. and KLIMOVSKY, A. (2012). Renormalisation of hierarchically interacting Cannings processes. Available at [arXiv:1209.1856v1](#).
- JAGERS, P. (1974). Galton–Watson processes in varying environments. *J. Appl. Probab.* **11** 174–178. [MR0368197](#)
- JIRINA, M. (1976). Extinction of non-homogeneous Galton–Watson processes. *J. Appl. Probab.* **13** 132–137. [MR0394912](#)
- KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248. [MR0671034](#)
- KINGMAN, J. F. C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. Oxford Univ. Press, New York. [MR1207584](#)
- LIMIC, V. and STURM, A. (2006). The spatial Λ -coalescent. *Electron. J. Probab.* **11** 363–393 (electronic). [MR2223040](#)
- LYONS, R. (1992). Random walks, capacity and percolation on trees. *Ann. Probab.* **20** 2043–2088. [MR1188053](#)
- MAULDIN, R. D. and WILLIAMS, S. C. (1986). Random recursive constructions: Asymptotic geometric and topological properties. *Trans. Amer. Math. Soc.* **295** 325–346. [MR0831202](#)
- MÖRTERS, P. (2010). *New Perspectives in Stochastic Geometry*. Oxford Univ. Press, Oxford.
- PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. [MR1742892](#)
- ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes and Martingales*. Cambridge Univ. Press, Cambridge.
- SAGITOV, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36** 1116–1125. [MR1742154](#)

SCHOOL OF MATHEMATICS
UNIVERSITY OF BRISTOL
UNIVERSITY WALK
BRISTOL
UNITED KINGDOM
E-MAIL: nicfreeman1209@gmail.com