

Survival, extinction and ergodicity in a spatially continuous population model

N. Berestycki, A.M. Etheridge and M. Hutzenthaler ^{*†}

May 31, 2009

Abstract

We consider a model recently introduced by Barton & Etheridge for a population evolving in a spatial continuum, in which a succession of catastrophic events of varying intensity allows for the possibility of large-scale extinction and recolonisation. These reproduction events are based on a Poisson process of spatial events (rather than individuals) and the potential number of offspring produced during such an event is Poisson with a certain intensity. We show that if this intensity is sufficiently large the population, when started from a translation invariant initial condition in \mathbb{R}^d , survives with probability one, whereas for low intensities the population dies out. Moreover we prove that ergodicity holds even in low dimensions. This contrasts sharply with the Dawson-Watanabe process and other traditional models in which reproduction of different individuals is uncorrelated.

1 Introduction

1.1 Background

A natural starting point in modelling a biological population is the Galton-Watson branching process in which each individual, independently, leaves behind a random number of offspring in the next generation. In order to extend this model to include the spatial structure of a population one can superpose a spatial motion of individuals. Thus, during their lifetime, individuals, again independently of one another, move around in space according to Brownian motion, say, and we arrive at branching Brownian motion. Of course this model is very special and so in the hope of obtaining a more robust model (that is one which approximates a wide range of individual based models that capture the same key features - independent reproduction and spatial motion of individuals) one passes to a *diffusion approximation*, corresponding to looking at the evolution of very large populations over very long timescales. Provided that the variance

^{*}Research supported by the DFG in the Dutch German Bilateral Research Group "Mathematics of Random Spatial Models from Physics and Biology" (FOR 498)

[†]Research supported by EPSRC Grant no GR/T19537/01

AMS 2000 subject classifications. 60K35, 92D25

Key words and phrases. Coalescence, ergodicity, survival, extinction, diffusion approximation

of the number of offspring of each individual in the population is finite, this leads us to the Dawson-Watanabe superprocess (also known as super-Brownian motion). The total size of the population then follows a Feller diffusion.

Sadly, although such models reveal a rich and beautiful mathematical structure, from the point of view of biological modelling they are inadequate. The difficulty is that they do not predict a stable population; the underlying branching structure dictates that the population will either die out in finite time or grow without bound. This would not be a fatal flaw if, by virtue of working in an unbounded spatial domain, the population still settled down to an equilibrium and indeed in at least three spatial dimensions this is the case, but many biological populations are evolving in one- or two-dimensional continua and there the branching Brownian motion and superprocess models do not converge to an equilibrium. In fact, worse still, they predict that if not extinct, at large times, the population will form ‘clumps’ of arbitrarily large density and extent (Felsenstein 1975, Sawyer & Fleischmann 1978).

Traditional work-arounds for these shortcomings have been to either (a) exogenously specify the total population size or (b) to suppose that the population is in fact subdivided into *demes* situated at the vertices of a discrete graph and specify the population size within each deme. These two approaches lead to the Fleming-Viot superprocess and to Kimura’s stepping stone model respectively. For a spatially extended population one expects the population size to be determined by local rules and so the Fleming-Viot approach is unsatisfactory. Kimura’s model is preferable in that one can specify the population sizes in each of the demes (although the model rapidly becomes unwieldy if they are not taken to be deterministic). However, the conditions of the model assume that the population within each deme is sufficiently large that reproduction there can be modelled through a (Wright-Fisher) diffusion approximation. And again one is faced with the task of justifying the choice of population sizes. (In fact most analyses assume the population size in each deme to be identically constant.)

The stepping stone model is particularly attractive from the perspective of population genetics as there is a simple and elegant description of the genealogical trees relating individuals in a sample from the population: as we trace backwards in time, lineages ancestral to the sample are described by a system of coalescing random walks. Thus lineages migrate between demes, reflecting the spatial motion of individuals in the forwards in time model, and, in addition, each pair of lineages in a given deme coalesces at a rate inversely proportional to the specified population size there. However, if one is modelling a population in a continuum one is forced to choose (rather arbitrarily) a subdivision of the population. Moreover, as discussed for example in Barton et al. (2002), in a spatial continuum ‘neighbourhood size’ (which dictates the number of neighbours with whom an individual can interact during reproduction) can be small,

violating the assumptions of the model. One ramification of this is that the genealogies may now exhibit not just pairwise coalescences of ancestral lineages, but also *multiple* mergers. In other words, locally it might be more appropriate not to consider just a Kingman coalescent, but instead one of the Λ -coalescents introduced by Pitman (1999) and Sagitov (1999), or more generally a Ξ -coalescent (see for example Schweinsberg 2000 or Möhle & Sagitov 2001) which allows for *simultaneous* multiple mergers. A coalescence model on a discrete lattice which allows for multiple mergers has been recently introduced by Limic & Sturm (2006), but, as discussed in Angel et al. (2006), even after taking a hydrodynamic rescaling of the system, the macroscopic properties of this process are highly sensitive to the microscopic structure of the underlying lattice.

Considerable effort has gone into finding more satisfactory models of populations evolving in spatial continua. There has been particular focus on modifications of the branching Brownian motion and superprocess models. A natural way to overcome the clumping of the population is to introduce local population regulation. Thus, because of competition for resources, individuals living in crowded regions will have fewer offspring than those living in sparsely populated regions. Models of this type have been studied by many authors, e.g. Bolker & Pacala (1997), Birkner & Depperschmidt (2007), Etheridge (2004), Hutzenthaler & Wakolbinger (2007), Law, Murrell & Dieckman (2003), Mueller & Tribe (1994). In particular it is known that such ‘locally regulated’ populations can survive provided, roughly, that individuals migrate fast enough that they are not killed through competition with their own close relatives.

The difficulty is that the existing models are extremely cumbersome. Moreover, if we are going to use our models as a basis for inference from genetic data, we need a description of the genealogical trees relating individuals in a sample from the population and this is not (yet at least) available (although see Barton et al. 2002 for some approximate results). What is really required are consistent and tractable (forwards in time) models for the evolution of a population and (backwards in time) descriptions of the corresponding genealogical trees.

In recent work Barton & Etheridge have introduced a new framework in which to model populations in a spatial continuum (see Etheridge 2008 for an outline). Their approach departs from those described above in two key ways. First, density dependent reproduction is achieved by basing reproduction events on a spatial Poisson process rather than on individuals (as in a branching process). Second, it allows for the explicit incorporation of large-scale extinction-recolonisation events. This captures the large scale fluctuations experienced by real populations in which the movement and reproductive success of many individuals can be correlated. For example, climate change has caused extreme extinction and recolonization events that dominate the demographic history of humans and other species (e.g. Eller et al. 2004). Backwards in time, the ge-

neological trees in this new framework are described by a system of coalescing compound Poisson jump processes with multiple mergers which can be thought of as a spatial Λ -coalescent or more generally Ξ -coalescent. (We note however that this is different from the spatial Λ -coalescent studied by Limic & Sturm 2006.)

In its simplest form, the model of Barton & Etheridge can be described as a spatial version of the Λ -Fleming-Viot process and, like a classical Fleming-Viot process, it is obtained as a limit of individual based models. It is these prelimiting models that concern us here.

1.2 The model

We shall suppose that the population to be modelled evolves in \mathbb{R}^d . Our results hold for any dimension, but from the biological point of view, the most interesting case is $d = 2$. The state of the population at time t will be represented by a purely atomic measure, X_t , with an atom at the spatial location of each of the individuals currently alive. We write \mathcal{N} for the space of purely atomic measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ (where as usual $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel subsets of \mathbb{R}^d) and $\mathcal{M}_1(\mathcal{N})$ for the probability measures on \mathcal{N} .

The model is parametrized by a (non-trivial) σ -finite measure

$$\xi(dr, du) = \mu(dr)\nu_r(du) \tag{1}$$

on $(0, \infty) \times (0, 1]$ and a real number $\lambda > 0$. We assume that

$$\int_0^\infty \int_0^1 ur^d (1 + r^d) \nu_r(du)\mu(dr) < \infty. \tag{2}$$

As we discuss briefly in §2, this condition is more than enough to guarantee existence of the model and, in particular, is stronger than that required for existence of the limiting spatial Λ -Fleming-Viot process of Barton & Etheridge. It will be used in several coupling arguments below and also in our proof of Theorem 1.5. To avoid trivialities we also assume that

$$\xi((0, \infty) \times (0, 1]) > 0. \tag{3}$$

The dynamics of $\{X_t\}_{t \geq 0}$ are described as follows.

- a. Let Π be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+ \times (0, 1]$ with rate $dt \otimes dx \otimes \xi(dr, du)$.
- b. If (t, x, r, u) is a point of Π , then at time t a reproduction event takes place in the ball $B_r(x)$ of radius r and centre $x \in \mathbb{R}^d$.
- c. If the ball is empty do nothing. If not:
 - (a) for each individual in $B_r(x)$, independently flip a coin which shows heads with probability u and kill all those individuals with a head;

- (b) throw down new individuals according to an independent Poisson point process with intensity $u\lambda\mathbf{1}_{B_r(x)}$.

If (t, x, r, u) is a point of Π , then we shall refer to a reproduction event based on the point x . Of course there are many variants on this theme (which is why we refer to Barton and Etheridge's work as providing a framework rather than a model), some of which are described in Etheridge (2008). Our results could be adapted without difficulty to many of these settings. In this paper we shall be concerned only with survival, extinction and ergodicity of the population. In the full model, the new particles generated in such a reproduction event all adopt the type of a single parent chosen uniformly at random from the individuals in the ball at time $t-$ (or more generally the type of one of a number of possible parents). This mechanism can then be thought of as regulating the reproductive success of individuals. If the ball $B_r(x)$ is crowded, then each individual has only a small chance of reproducing. On the other hand if the ball is only sparsely populated they have a significant chance of producing a Poisson number of offspring with mean $\lambda u C_d r^d$ where

$$C_d r^d = \frac{2r^d \pi^{d/2}}{d\Gamma(d/2)} \quad (4)$$

is the volume of a ball of radius r in \mathbb{R}^d . The limiting model is obtained by letting λ tend to infinity. Although at first sight this corresponds to taking the population density to infinity, in fact a signature of finite local population size is retained since we still see multiple mergers of ancestral lineages.

Of course, for $\lambda < \infty$ any reproductive event has positive probability of leaving no individuals in a given region of space, but, because neighbourhoods overlap, an empty neighbourhood can subsequently become recolonised. The question is whether this is enough to prevent the population from dying out. As we shall see, there is a critical value of the parameter λ below which extinction is certain, but above which the population, started from a Poisson random measure with intensity ρdx , survives.

Definition 1.1 ((λ, ξ) -process). *We shall call a process $\{X_t\}_{t \geq 0}$ governed by the dynamics above a (λ, ξ) -process.*

1.3 The main results

Let us state our results a little more formally. First we need some notation. For two measures $\phi, \psi \in \mathcal{N}$ we write $\phi \leq \psi$ if $\phi(A) \leq \psi(A)$ for all Borel sets $A \in \mathcal{B}(\mathbb{R}^d)$. This defines a partial order on \mathcal{N} with minimal element $\underline{0}$, the zero measure. For a random measure Φ , we write $\mathcal{L}(\Phi)$ for its law. We shall say that a random measure Φ is dominated by Ψ , written $\mathcal{L}(\Phi) \leq \mathcal{L}(\Psi)$, if there is a coupling $\mathcal{L}(\Phi, \Psi)$ such that $\Phi \leq \Psi$ almost surely. Finally, let us denote by \mathcal{T} the translation invariant probability measures on \mathcal{N} and by $\mathcal{T}_{>0}$ the subset of these that assign zero mass to the trivial measure $\underline{0}$.

Our main results are as follows.

Theorem 1.2. *For a measurable subset $A \subseteq \mathbb{R}^d$, denote by $N(A)$ the number of lattice points $x \in \mathbb{Z}^d$ such that the unit cube of \mathbb{R}^d centered at x is entirely contained in A . Let $\{X_t\}_{t \geq 0}$ denote a (λ, ξ) -process with X_0 a Poisson point process on \mathbb{R}^d with intensity λdx . For each fixed ξ , and for each $\gamma > 0$ there exists λ_0 such that for every $\lambda > \lambda_0$ and every $\varepsilon > 0$, there exists a nonrandom $N_0 = N_0(\varepsilon) \geq 1$ such that*

$$\liminf_{t \rightarrow \infty} \mathbb{P}[X_t(A) > \gamma N(A)] \geq 1 - \varepsilon \quad (5)$$

whenever $N(A) > N_0$.

Theorem 1.3. *Let $\{X_t\}_{t \geq 0}$ denote a (λ, ξ) -process. There exists a measure $\bar{\nu} \in \mathcal{M}_1(\mathcal{N})$ such that*

$$\mathcal{L}(X_t) \xrightarrow{w} \mathbb{P}[X_0 = \underline{0}] \cdot \delta_{\underline{0}} + \mathbb{P}[X_0 \neq \underline{0}] \cdot \bar{\nu} \quad \text{as } t \rightarrow \infty \quad (6)$$

whenever $\mathcal{L}(X_0) \in \mathcal{T}$.

Definition 1.4 (Upper invariant measure). *We call $\bar{\nu}$ the upper invariant measure for the (λ, ξ) -process.*

Theorem 1.5. *Let $\{X_t\}_{t \geq 0}$ be a (λ, ξ) -process. For each fixed ξ , there exists a $\lambda_0 > 0$ such that the upper invariant measure $\bar{\nu} = \delta_{\underline{0}}$ for $0 \leq \lambda < \lambda_0$.*

The proofs of survival (Theorem 1.2) and extinction (Theorem 1.5) rest on comparison with oriented percolation and subcritical branching respectively. Some care is needed in the proof of survival as we are comparing our continuum model to one in discrete space. This turns out to be rather delicate. It eventually leads us to edge percolation rather than the more usual site percolation, but beyond that, the proofs follow a standard pattern. They do highlight the importance of the Poisson structure of reproduction events which provides something akin to the graphical representation of lattice based interacting particle system models. The main novelty in the work lies in the model itself. Exhibiting the survival and ergodicity properties of populations within this framework is an essential step in a larger programme that develops our understanding of the models to a point where they can be properly justified with reference to biological data.

The rest of the paper is laid out as follows. In §2 we prove, via coupling, some preliminary results that establish, in particular, monotone dependence of the survival probability on both the parameter λ and the initial state of the process. As a by-product we prove existence of the process. We also check that asymptotically the population ‘forgets’ its initial state. The proofs of Theorem 1.2, Theorem 1.3 and Theorem 1.5 are provided in §3, §4 and §5 respectively.

2 Preliminary results

In this section we use simple coupling arguments to prove some elementary monotonicity properties of the process $\{X_t\}_{t \geq 0}$.

Lemma 2.1. *Let $\{X_t\}_{t \geq 0}$ be a (λ, ξ) -process. Suppose that $\mathcal{L}(X_0) \leq \mathcal{L}(Z)$ where Z is a Poisson point process with intensity λdx , then $\mathcal{L}(X_t) \leq \mathcal{L}(Z)$ for all $t \geq 0$.*

Proof. To explain the idea we proceed heuristically. Our aim is to construct, simultaneously, a copy of $\{X_t\}_{t \geq 0}$ and a process $\{Z_t\}_{t \geq 0}$ whose marginal distribution is $\mathcal{L}(Z)$ for every $t \geq 0$. The individuals in the process $\{X_t\}_{t \geq 0}$, which will be coloured red, evolve according to the dynamics described in §1.2. However, we augment this process with ‘ghost’ particles, to be coloured white. Thus, if at a point $(t, x, r, u) \in \Pi$ we find $B_r(x)$ empty at time $t-$, we introduce white particles according to a Poisson process with intensity $u\lambda \mathbf{1}_{B_r(x)}$. Once born, ghost particles evolve exactly like real particles, except that our colouring rules are as follows: if at the time of a reproduction event there is a red individual available in $B_r(x)$ to act as parent, then all newborns are red, otherwise they are white.

Now observe that for $(t, x, r, u) \in \Pi$, if the combined red and white population in $B_r(x)$ at time $t-$ (which may be zero) is distributed according to a Poisson point process with intensity $\lambda \mathbf{1}_{B_r(x)}$, then during the reproduction, first these individuals are ‘thinned’ according to u , thus reducing the intensity in the ball to $(1 - u)\lambda \mathbf{1}_{B_r(x)}$, and then we replenish with an independent Poisson distribution of intensity $u\lambda \mathbf{1}_{B_r(x)}$ resulting once again in a total population which is Poisson with intensity $\lambda \mathbf{1}_{B_r(x)}$. Thus the combined population at any time t has distribution $\mathcal{L}(Z)$ and since the red particles follow the dynamics of a (λ, ξ) -process the result of the lemma will follow.

Of course the difficulty with the argument above is that reproduction events fall on the population at a potentially unbounded rate and so some further justification is required. Let us denote the combined red and white particles at time t by $Z_t \in \mathcal{N}$. Consider the births and deaths in a bounded set $A \in \mathcal{B}(\mathbb{R}^d)$. Notice that once we ignore the labels, births and deaths in A do not depend on the *configuration* of particles outside A (although the *times* of changes within and outwith A will be correlated). Moreover, if there are currently n individuals in A , the rate at which the configuration in A changes is certainly bounded above by

$$\int_0^\infty \int_0^1 r^d \left(nu + |A|(1 - e^{-u\lambda|A|}) \right) \nu_r(du) \mu(dr),$$

which is finite by Assumption (2). Thus if A is a bounded set and T_M is the first time there are more than M particles in A , the *total* number of individuals in A up to time $t \wedge T_M$ is well defined. Moreover, note that the number of white and red particles in A

up to time T_M is a nonnegative martingale, and it is easy to conclude from this that $T_M \rightarrow \infty$ almost surely as $M \rightarrow \infty$. This, in turn, implies that the process of the total number of (uncoloured) particles in A is well-defined. We must also see that the colour of an individual alive at time t is well-defined. To do this, we trace backwards in time and consider all the ‘potential ancestors’ of that individual. Recall that if an individual is an offspring of a reproduction event, then it will be coloured red if *any* of the individuals living within the region affected by that event is red. Thus we must determine the labels of all such individuals and these are our ‘potential ancestors’. We now check that in order to determine the label of an individual at a finite time t we need only examine the labels of an almost surely finite number of potential ancestors at time zero. To do this we compare with a branching process. We exploit the reversibility of the process $\{Z_t\}_{t \geq 0}$. Consider then a single ancestral lineage. We trace back until the first time that it is an offspring of a reproduction event. We must then add *all* the individuals alive before the reproduction event to our pool of potential ancestors. If the reproduction event affects a ball of radius r , this will be a Poisson number of individuals with parameter $\lambda C_d r^d$. Different (potential) ancestral lineages do not evolve independently of one another, as two or more could be hit by the same reproduction event, but notice that by comparing with a branching process we are ‘double counting’. All that remains is to check that the branching process is non-explosive. The rate at which a given lineage branches (that is, is affected by an event at which branching may occur - note that it could also die then) can be computed as follows. Note that an annulus centered at the origin between radii s and $s + \delta s$ has $d^{-1} C_d s^{d-1} \delta s$ Lebesgue measure up to terms of size $o(\delta s)$. Thus, integrating over the distance s of the location of the possible centre of the event, we find by Fubini’s theorem that this rate is equal to

$$\begin{aligned} & \int_0^\infty \frac{C_d}{d} s^{d-1} ds \int_s^\infty \mu(dr) \int_0^1 u \nu_r(du) \\ &= \int_0^\infty \int_0^1 u C_d r^d \nu_r(du) \mu(dr) \end{aligned} \tag{7}$$

which is finite by Assumption (2).

Given that it does so, the mean number of descendants that it leaves is then

$$\frac{\int_0^\infty \int_0^1 \lambda C_d^2 u r^{2d} \nu_r(du) \mu(dr)}{\int_0^\infty \int_0^1 C_d u r^d \nu_r(du) \mu(dr)} \tag{8}$$

which is also finite (again by Assumption 2). Comparing with Theorem 1, § III.2 of Athreya & Ney (1972), we see that this is sufficient to guarantee that the branching process will not explode and so the number of potential ancestors at time zero is finite and the proof is complete. \square

Remark 2.2 (Existence of the process). *Notice that a by-product of our proof of Lemma 2.1 is a proof of the existence of the model. We don’t actually believe that*

the condition (2) is necessary, but it is interesting to note that one arrives at a similar condition on trying to mimic standard existence proofs for interacting particle systems as described for example in Liggett (1985), Chapter 1, Theorem 3.9. Liggett's condition (3.8) turns out to be the stronger requirement. The rather strong condition on the behaviour of the measure ξ for large values of r arises in that setting because the addition of a single individual in a large and otherwise unpopulated region can radically change the behaviour of the process. In the limit as $\lambda \rightarrow \infty$, such large unpopulated regions do not occur and the conditions for existence of the limiting model are much weaker:

$$\int_0^\infty \int_0^1 ur^d \nu_r(du) \mu(dr) < \infty \quad (9)$$

and

$$\int_{\mathbb{R}^d} (1 \wedge |x|^2) \left(\int_{(|x|/2, \infty)} \int_{[0,1]} \frac{L_r(x)}{\pi r^d} u \nu_r(du) \mu(dr) \right) dx < \infty, \quad (10)$$

where $L_r(x)$ is the area of $B_r(0) \cap B_r(x)$ and $B_r(x)$ denotes the ball of radius r centred on $x \in \mathbb{R}^d$. Essentially the first guarantees that the configuration of types at a single point does not change too rapidly and the second ensures that we can fit together the configurations at different points in a consistent way by ensuring that the ancestral lineages follow well-defined compound Poisson processes.

The following Lemma provides a version of Lemma 2.1 for arbitrary initial conditions. It could be restated as

$$\limsup_{t \rightarrow \infty} \mathcal{L}(X_t) \leq \mathcal{L}(Z) \quad (11)$$

where Z is a Poisson point process with intensity λdx .

Lemma 2.3. *Let $\{X_t\}_{t \geq 0}$ be a (λ, ξ) -process. There exist processes $\{\check{X}_t\}_{t \geq 0}$ and $\{Z_t\}_{t \geq 0}$ and a coupling $\mathcal{L}(X_t, \check{X}_t, Z_t)$ such that $X_t \leq Z_t + \check{X}_t$ for all $t \geq 0$ almost surely where, for each $t \geq 0$, Z_t is a Poisson point process with intensity λdx and for every bounded Borel set $A \subset \mathbb{R}^d$ there exists a random time τ such that $\check{X}_t(A) = 0$ for all $t \geq \tau$ almost surely.*

Proof. For this and the remaining proofs in this section we describe the coupling employed, but tacitly assume that everything is well-defined (which follows easily from the arguments in the proof of Lemma 2.1). Assume X_0 and Z_0 have respective laws $\mathcal{L}(X_0)$ and $\mathcal{L}(Z_0)$, and define \check{X}_0 to be the point measure made of particles in X_0 but not in Z_0 , which we note $\check{X}_0 = X_0 \setminus Z_0$. The coupling here is similar to that in the proof of Lemma 2.1 except that now we use three colours: red, white and blue. The X -process will be made up of red and blue particles, where blue particles are ancestral, that is they were already present at time zero and have not yet been killed. Red particles are

those individuals in X that are offspring of a reproduction event that has happened since time zero. The process \check{X}_t consists only of those blue particles at time t which were already in \check{X}_0 (since these particles are blue, they have never been affected by a reproductive event). A reproduction event for which $B_r(x)$ contains either blue or red particles at time $t-$ will result in red offspring, whereas a pure white patch or an empty patch will produce white offspring. At time zero we suppose that the white particles are distributed according to $\mathcal{L}(Z)$ and so as before the combined red and white population will have law $\mathcal{L}(Z)$ at every $t \geq 0$. It is plain from this construction that

$$X_t \leq \check{X}_t + Z_t$$

almost surely for all $t \geq 0$, since the right-hand side contains all blue, red and white particles, while the left-hand side contains only the red and blue particles. Furthermore, since all the blue particles in any bounded set will be killed in a finite random time, the result follows. \square

Lemma 2.4. *Let $\{X_t\}_{t \geq 0}$ and $\{Y_t\}_{t \geq 0}$ be two (λ, ξ) -processes. If $\mathcal{L}(X_0) \leq \mathcal{L}(Y_0)$ then*

$$\mathcal{L}(X_t) \leq \mathcal{L}(Y_t)$$

for all $t \geq 0$.

Proof. This time we first choose initial configurations X_0 and Y_0 with $X_0 \leq Y_0$. We construct the process $\{Y_t\}_{t \geq 0}$ according to the dynamics described in §1.2, but again we attach a colour to each individual in the population. At time zero, the individuals of X_0 are all red, while those in $Y_0 \setminus X_0$ are green. Once again the rules mirror those in the proof of Lemma 2.1. For $t > 0$, colours are determined as follows. For each point $(t, x, r, u) \in \Pi$, if, at time $t-$, $B_r(x)$ contains *at least one* red individual, then all *new* individuals created are also red. If, on the other hand, all individuals in $B_r(x)$ at time $t-$ are green, then all new individuals created during the reproduction event are green. Our colouring does not change the dynamics of the combined population (of reds and greens) which therefore provides a copy of $\{Y_t\}_{t \geq 0}$ and a moment's thought shows that the red particles alone provide us with a copy of $\{X_t\}_{t \geq 0}$ and the proof is complete. \square

Lemma 2.5. *Let $\{X_t^{\lambda_1}\}_{t \geq 0}$ be a (λ_1, ξ) -process and let $\{X_t^{\lambda_2}\}_{t \geq 0}$ be a (λ_2, ξ) -process where $\lambda_1 < \lambda_2$. If $\mathcal{L}(X_0^{\lambda_1}) \leq \mathcal{L}(X_0^{\lambda_2})$, then*

$$\mathcal{L}(X_t^{\lambda_1}) \leq \mathcal{L}(X_t^{\lambda_2})$$

for all $t \geq 0$.

Proof. Again we can use a straightforward coupling. We construct the X^{λ_1} population as a subset of the X^{λ_2} population, colouring the former individuals red and the ‘excess’

individuals of X^{λ_2} green. At the time of a reproduction event, if there are only green individuals in $B_r(x)$, then we produce green offspring with intensity $u\lambda_2\mathbf{1}_{B_r(x)}$, but if there are red individuals then we produce red offspring with intensity $u\lambda_1\mathbf{1}_{B_r(x)}$ and green offspring with intensity $u(\lambda_2 - \lambda_1)\mathbf{1}_{B_r(x)}$. It is elementary to check that the red population is a (λ_1, ξ) -process and the combined red and green populations yield a (λ_2, ξ) -process. \square

3 Survival

Lemma 2.3 immediately implies that if we restrict our population to live not on \mathbb{R}^d but on some bounded region, then eventually it will go extinct (although simulations on a large torus suggest that this can take a very long time). However, on an infinite space, Theorem 1.2 tells us that for suitable choices of parameters (and initial condition) the (λ, ξ) -process can survive forever. In this section we provide a detailed proof of this result through a comparison with an oriented percolation model.

Assumption:

Throughout this section we assume that

$$\xi \left([\sqrt{d+3}, d+1] \times (0, 1] \right) > 0. \quad (12)$$

Evidently under assumption (3) we can always rescale space to ensure that this is true.

Our first task is to construct a discrete lattice approximation of the system. We consider cubic cells of unit Lebesgue measure about each element of \mathbb{Z}^d , that is, subsets of \mathbb{R}^d of the form $\prod_{i=1}^d (x_i - 1/2, x_i + 1/2)$, where $(x_1, \dots, x_d) \in \mathbb{Z}^d$ is the centre of the cell. Note that any point in the cell is within distance $\sqrt{d}/2$ of the centre. This gives a regular tessellation of \mathbb{R}^d (minus countably many hyperplanes, of zero Lebesgue measure).

We will not use coordinates again and so from now on x_i will denote a point in \mathbb{R}^d rather than the i th component of x .

Let $Y(t)$ denote the random element of $\{0, 1\}^{\mathbb{Z}^d}$ defined by putting a 1 on a site of \mathbb{Z}^d if and only if the corresponding cell has at least one particle in it. That is, for $x \in \mathbb{Z}^d$,

$$Y(t, x) = 1 \text{ if and only if } \int_{C(x)} X(t, dx) > 0 \quad (13)$$

where $C(x)$ is the cell centered at x . We call Y the *cell-occupancy process* of X .

We start with the following simple geometric lemma:

Lemma 3.1. *Let C be a cell and $x \in C$ be a point within the cell. Let $y \in \mathbb{R}^d$ with $\|y - x\| = r > 0$. Then there exists a constant α depending solely on the dimension, such that for all $h \in (0, 1)$,*

$$\text{Vol}\{B_{r+h}(y) \cap C\} \geq \alpha h^d.$$

Proof. Assume first that $y \notin C$. Then there exists a point $z \in \partial C$ on the line segment $[x, y]$ between x and y . By the triangle inequality, we have:

$$\{B_{h/2}(z) \cap C\} \subset \{B_{r+h}(y) \cap C\}$$

so that, in particular, the volume of the right-hand side is at least that of the left-hand side. Now, since z is on the boundary of C and $h/2$ is at most half the sidelength of the cell, $B_{h/2}(z) \cap C$ contains (at least) a portion of \mathbb{R}^d isometric to a ball of radius $h/2$ intersected with the positive orthant $O^+ = \{x \in \mathbb{R}^d : x_1 > 0, \dots, x_d > 0\}$. But note that

$$\begin{aligned} \text{Vol}\{B_{h/2}(0) \cap O^+\} &= \frac{1}{2^d} \text{Vol}B_{h/2}(0) \\ &= \alpha h^d. \end{aligned}$$

where α depends only on the dimension. The case $y \in C$ is similar. This proves Lemma 3.1. \square

For our next lemma, we suppose that there are individuals alive at points $x_1, \dots, x_m \in \mathbb{R}^d$ in distinct cells and estimate the rate at which cells in the corresponding occupancy process become empty. The idea is that although the rate at which x_i is hit by a reproductive event may be high, if λ is sufficiently large, then with high probability even if the individual at x_i is killed, at least one offspring will be born in cell containing x_i . From the point of view of the occupancy process there is no change in that cell. The only difficulty is that if x_i happens to be near the boundary of the cell $C(x_i)$, the volume of the intersection $B_r(x) \cap C$ might be small, in which case the chance of a ‘replacement’ being born is also small. We use Lemma 3.1 to control the size of the set of radii which contribute to this effect.

We shall refer to an event in which a particle from the set $A = \{x_1, \dots, x_m\}$ dies without there being a birth of another particle in the same cell as a *net loss*, and call $\delta(A)$ the instantaneous rate at which this occurs.

Lemma 3.2. *Let x_1, x_2, \dots, x_m belong to distinct cells. Given $\varepsilon > 0$ there exists λ_0 (independent of m) such that for $\lambda > \lambda_0$, the instantaneous rate at which we see a net loss from the set $A = \{x_1, x_2, \dots, x_m\}$ is at most $2m\varepsilon$.*

Proof. Consider an event corresponding to a point $(t, x, r, u) \in \Pi$. Each individual x_j within $B_r(x)$ dies (and is not subsequently replaced in the same cell) with probability

$$u \exp(-\lambda u \text{Vol}\{B_r(x) \cap C(x_j)\}).$$

Thus the probability that at least one individual from A is killed and not replaced is at most

$$\sum_{j=1}^m \mathbf{1}_{\{\|x-x_j\| \leq r\}} \times u \exp(-\lambda u \text{Vol}\{B_r(x) \cap C(x_j)\}).$$

Let $h \in (0, 1)$ (so that Lemma 3.1 holds) and write $\delta(A)$ for the infinitesimal rate at which an individual from A is killed and not replaced by a particle in the same cell. Using Lemma 3.1 we have

$$\begin{aligned}
\delta(A) &\leq \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \sum_{j=1}^m \mathbf{1}_{\{\|x-x_j\| \leq r\}} u e^{-\lambda u \text{Vol}\{B_r(x) \cap C(x_j)\}} \nu_r(du) \mu(dr) dx \\
&\leq \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 u \sum_{j=1}^m \mathbf{1}_{\{\|x-x_j\| \leq r\}} \left(\mathbf{1}_{\{r \leq \|x-x_j\|+h\}} + e^{-\lambda u \alpha h^d} \right) \nu_r(du) \mu(dr) dx \\
&= \sum_{j=1}^m \int_{\mathbb{R}^d} \int_{\mathbb{R}_+} \int_0^1 u \mathbf{1}_{\{\|x-x_j\| \leq r \leq \|x-x_j\|+h\}} \nu_r(du) \mu(dr) dx \\
&\quad + \sum_{j=1}^m \int_{\mathbb{R}^d} \int_{\mathbb{R}_+} \int_0^1 \mathbf{1}_{\{\|x-x_j\| \leq r\}} u e^{-\lambda u \alpha h^d} \nu_r(du) \mu(dr) dx \\
&= S_1 + S_2,
\end{aligned}$$

say. We treat the two terms S_1 and S_2 one at a time. For S_1 , we use Fubini's Theorem to obtain

$$\begin{aligned}
S_1 &= \sum_{j=1}^m \int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbf{1}_{\{\|x-x_j\| \leq r \leq \|x-x_j\|+h\}} u \nu_r(du) \mu(dr) dx \\
&= \sum_{j=1}^m \int_0^\infty \int_0^1 \int_{\mathbb{R}^d} \mathbf{1}_{\{\|x-x_j\| \leq r \leq \|x-x_j\|+h\}} u dx \nu_r(du) \mu(dr) \\
&= m \int_0^\infty \int_0^1 \text{Vol}\{R(r, h)\} u \nu_r(du) \mu(dr) \\
&= m \varphi_1(h),
\end{aligned}$$

say, where

$$R(r, h) := B_{r+h}(0) \setminus B_r(0) \tag{14}$$

is the annulus between radii r and $r+h$. Note that by the Lebesgue Dominated Convergence Theorem and (2), $\varphi_1(h) \rightarrow 0$ as $h \rightarrow 0$. Thus, given $\varepsilon > 0$, we can choose h small enough that $\varphi_1(h) < \varepsilon$, in which case $S_1 \leq m\varepsilon$. Notice further that S_1 , and hence our choice of h , is independent of λ .

For S_2 , a similar use of Fubini's Theorem yields

$$S_2 = m \varphi_2(\lambda) := m \int_{\mathbb{R}_+} \int_0^1 \text{Vol}\{B_r(0)\} u e^{-u \lambda \alpha h^d} \nu_r(du) \mu(dr).$$

Now observe that with h fixed (as above) we also have $\varphi_2(\lambda) \rightarrow 0$ as $\lambda \rightarrow \infty$, again by the Lebesgue Dominated Convergence Theorem. Choosing λ_0 such that $S_2 \leq m\varepsilon$ whenever $\lambda > \lambda_0$, the proof is complete. \square

We now extend this to provide what will be a key estimate in our comparison with oriented percolation. First we need some notation. If $x \in \mathbb{R}^d$, we let $C(x) = C(\bar{x})$, where \bar{x} is the point of \mathbb{Z}^d closest to x . Now let $T > 0$, fix $m \geq 1$ and let $\{x_i\}_{i=1}^m$ and $\{y_i\}_{i=1}^m$ be points in \mathbb{R}^d such that for all $1 \leq i \leq m$, $\bar{x}_i \sim \bar{y}_i$ (that is \bar{x}_i and \bar{y}_i are neighbours in \mathbb{Z}^d). We write \mathbf{L} for a collection of m space-time paths going from x_i to y_i in time T , such that each path in \mathbf{L} jumps exactly once between x_i and y_i (at some unspecified time in $(0, T)$). We think of \mathbf{L} as a collection of lineages. Finally, we let $\pi_t(\mathbf{L})$ be the projection of \mathbf{L} onto \mathbb{R}^d at time t , that is, the collection of m points in \mathbb{R}^d which tell us the locations of the lineages at time $0 \leq t \leq T$. Thus, in particular, $\pi_t(\mathbf{L}) \subset \bigcup_{i=1}^m \{x_i, y_i\}$.

For $0 \leq t \leq T$, let $E_t(\mathbf{L})$ be the event that no lineage in \mathbf{L} suffers a net loss (by which, as before, we mean is killed by a reproduction event in which no offspring are born in the cell containing the current position of that lineage) during the interval $(0, t)$. On the event $E_t(\mathbf{L})$, there is a set of lineages \mathbf{L}_t which brings particles from locations x_i to points within $C(\bar{x}_i) \cup C(\bar{y}_i)$, with perhaps a finite number of jumps in between. We write $E(\mathbf{L})$ for the event $\bigcap_{0 \leq t \leq T} E_t(\mathbf{L})$. On the event $E(\mathbf{L})$ there is a set of lineages taking particles from points in $C(\bar{x}_i)$ to points in $C(\bar{y}_i)$ in $[0, T]$.

Lemma 3.3. *In the notation above, let x_1, x_2, \dots, x_m and y_1, y_2, \dots, y_m occupy distinct cells. For any $\varepsilon > 0$, and for any $T > 0$, there exists $\lambda_1 = \lambda_1(\varepsilon, T)$ such that if $\lambda > \lambda_1$, then*

$$\mathbb{P}[E(\mathbf{L})] \geq (1 - \varepsilon)^m.$$

Proof. Notice that our estimates of Lemma 3.2 are independent of the positions of the individuals of A within their respective cells. It follows that for every $\varepsilon > 0$, we can find a λ_0 such that for all $\lambda > \lambda_0$,

$$\mathbb{P}[E(\mathbf{L})] \geq \exp(-2m\varepsilon T).$$

The existence of the required λ_1 is now immediate. □

We can use this estimate to bound below the probability that m occupied cells have each ‘infected’ a neighbouring cell by time T . We need some more notation. Let $x, y \in \mathbb{R}^d$, and recall that \bar{x} (resp. \bar{y}) is the point of \mathbb{Z}^d closest to x (resp. y). Suppose that $\bar{x} \sim \bar{y}$. We shall say that there is a *potential* infection from \bar{x} to \bar{y} in time T , and denote this event by $J(\bar{x}, \bar{y})$, if there is a reproduction event in the time interval $(0, T)$ which affects both $C(\bar{x})$ and $C(\bar{y})$ and which, provided that $C(\bar{x})$ is non-empty, leaves at least one offspring in $C(\bar{y})$. In fact we shall consider only reproduction events based on points in $C(\bar{x})$ and affecting a ball of radius at most three in order to guarantee independence of the events $J(\bar{x}_i, \bar{y}_i)$ for sufficiently well-spaced points x_i . We say further that there is a *proper* infection if also the corresponding lineage carrying a particle from

its location in $C(\bar{x})$ to a location in $C(\bar{y})$ in $[0, T]$ does not suffer a net loss. We use $I(\bar{x}, \bar{y})$ to denote this event. Note that under $I(\bar{x}, \bar{y})$, if $C(\bar{x})$ is occupied at time zero, then $C(\bar{y})$ will be occupied at time T .

Lemma 3.4. *Let $\varepsilon > 0$. For all $T > 0$, there exists $\lambda_2 = \lambda_2(\varepsilon, T) > 0$ such that for all $\lambda > \lambda_2$ and $m \geq 1$, if $x_i, y_i \in \mathbb{R}^d, 1 \leq i \leq m$, with $\bar{x}_i \sim \bar{y}_i$ for $1 \leq i \leq m$, $\|\bar{x}_i - \bar{x}_j\| > L := 2d + 6$ for all $1 \leq i \neq j \leq m$, then*

$$\mathbb{P} \left[\bigcap_{i=1}^m I(\bar{x}_i, \bar{y}_i) \right] \geq (1 - \varepsilon)^m.$$

Proof. Let $T > 0$ be fixed, and consider first the case $m = 1$. We begin by estimating the probability of the event $J(\bar{x}_1, \bar{y}_1)$. We require a lower bound on the probability that there is an event affecting $C(\bar{x}_1)$ and $C(\bar{y}_1)$ in which (provided the region it affects is not empty immediately before the event) at least one offspring is born in $C(\bar{y}_1)$. We should like an estimate that will generalise easily to $m > 1$ and so we restrict our attention to reproduction events based on the points of $C(\bar{x}_1)$ which will *not* affect $C(\bar{x}_i)$ or $C(\bar{y}_i)$ for $i \neq 1$. For this we consider only those events for which $r_1 \leq r \leq r_2$, where $r_1 = \sqrt{d+3}$ and $r_2 = d+2$. The choice of r_1 and r_2 is such that $r_1 < r_2$ for all dimensions d , and r_1 is the maximal distance between two points in adjacent cells. Events satisfying these criteria which also result in the birth of offspring in $C(\bar{y}_1)$ appear at instantaneous rate

$$b(\lambda) \geq \bar{b}(\lambda) := \int_{\sqrt{d+3}}^{d+2} \mu(dr) \int_0^1 (1 - e^{-\lambda u}) \nu_r(du).$$

Thus

$$\mathbb{P}[J(\bar{x}_1, \bar{y}_1)] \geq 1 - e^{-T\bar{b}(\lambda)}.$$

Now, note that the event $J(\bar{x}_1, \bar{y}_1)$ depends only on the Poisson events based on points within $C(\bar{x})$, so that the events $J(\bar{x}_i, \bar{y}_i), 1 \leq i \leq m$, are independent of one another.

Hence

$$\mathbb{P} \left[\bigcap_{i=1}^m J(\bar{x}_i, \bar{y}_i) \right] \geq (1 - e^{-T\bar{b}(\lambda)})^m.$$

On the event $J = \bigcap J(\bar{x}_i, \bar{y}_i)$, we may choose a set of lineages \mathbf{L} which leads from locations x_i to points $y_i \in C(\bar{y}_i)$ by single jumps, and extend this definition of \mathbf{L} arbitrarily outside of the event J . Note also that

$$\bigcap_{i=1}^m I(\bar{x}_i, \bar{y}_i) \supseteq \bigcap_{i=1}^m J(\bar{x}_i, \bar{y}_i) \cap E(\mathbf{L}).$$

Although \mathbf{L} is a random set of lineages, we may still apply Lemma 3.3 since, conditional on J , the Poisson point process which governs the dynamics can be realised by adding to an unconditional realisation of the Poisson events, those that create the desired

infections $J(\bar{x}_i, \bar{y}_i)$ (this is the Palm distribution of the Poisson point process). Since $\|x_i - x_j\| > L = 2d + 6$, and the radius of all those events is smaller than $r_2 = d + 2$, it follows that these additional events can never kill $\pi_t(\mathbf{L}_t)$. Choose λ_2 sufficiently large that $\lambda_2 > \lambda_1$ (from Lemma 3.3) and $e^{-Tb(\lambda_2)} < \varepsilon$. For all $\lambda > \lambda_2$, noting that $\bar{b}(\lambda)$ is a non-increasing function of λ , we have, setting $I = \bigcap_{i=1}^m I(\bar{x}_i, \bar{y}_i)$:

$$\begin{aligned} \mathbb{P}[I] &\geq \mathbb{P}[J \cap E(\mathbf{L})] \\ &= \mathbb{P}[E(\mathbf{L})|J]\mathbb{P}[J] \\ &\geq (1 - \varepsilon)^{2m}. \end{aligned}$$

Noting that $(1 - \varepsilon)^2 \geq 1 - 2\varepsilon$, we conclude

$$\mathbb{P}[I] \geq (1 - 2\varepsilon)^m,$$

as desired. □

Proof of Theorem 1.2. As is often the case with survival results of this nature, the key to the proof of Theorem 1.2 is a comparison with oriented percolation. Some care is needed, because we are dealing with edge-oriented percolation (whereas results are typically given for site-oriented percolation) and there are some dependencies.

We start by recalling the definition of L -dependent oriented edge percolation. Let \mathbf{E}^d be the set of edges e connecting vertices of the form (x, t) and $(y, t + 1)$ where $x, y \in \mathbb{Z}^d$ and $x \sim y$ are neighbours in \mathbb{Z}^d , and where $t \in \mathbb{Z}_+$. We call x the base point of e . Let $(\omega(e))_{e \in \mathbf{E}^d}$ be a random element of $\{0, 1\}^{\mathbf{E}^d}$, where we think of $\omega(e) = 1$ as the edge being open, and $\omega(e) = 0$ as the edge being closed. Let $L \geq 1$. We say that ω is a realization of L -dependent oriented edge percolation with density $1 - \gamma$ if for every $k \geq 1$, and for every $e_1, \dots, e_k \in \mathbf{E}^d$ with respective base points $x_1, \dots, x_k \in \mathbb{Z}^d$ such that $\|x_i - x_j\|_\infty \geq L$ for every $1 \leq i \neq j \leq k$, then we have

$$P[\omega(e_1) = 1, \dots, \omega(e_k) = 1] \geq (1 - \gamma)^k.$$

Let W_0^p be a random set of vertices of \mathbb{Z}^d chosen by independent Bernoulli trials with parameter p . Let W_n^p be the set of vertices in $\{n\} \times \mathbb{Z}^d$ which can be reached from $\{0\} \times W_0^p$ using only open edges of the oriented percolation configuration ω . This determines the open cluster of the oriented percolation.

The next step is to show that our cell-occupancy process at time n , $Y(n, x)$ of equation (13), dominates W_n^p for $n \in \mathbb{Z}_+$ and a suitable choice of p . For this, fix $M = 2d + 6$ and let L be such that $\|x\|_\infty \geq L$ implies $\|x\| \geq M$. Fix $T > 0$ arbitrarily, and declare an edge between (x, n) and $(y, n + 1)$ to be open if $x \sim y$ and if $I(x, y)$ holds between times nT and $(n + 1)T$ (in the sense of Lemma 3.4). Then by construction $Y(n, x)$ stochastically dominates W_n^p where $p = 1 - \exp(-\rho)$.

To recover the exact formulation of Theorem 1.2, we now recall the following useful result, proved in a more general form by Liggett et al. (1997).

Theorem 3.5 (Grimmett (1999), Theorem 7.65). *Let $d, k \geq 1$. There exists a non-decreasing function $\pi : [0, 1] \rightarrow [0, 1]$ satisfying $\pi(\delta) \rightarrow 1$ as $\delta \rightarrow 1$ such that the following holds. If $Y = \{Y_x : x \in \mathbb{Z}^d\}$ is a k -dependent family of random variables satisfying*

$$\mathbb{P}[Y_x = 1] \geq \delta \quad \text{for all } x \in \mathbb{Z}^d,$$

then Y stochastically dominates $Z^{\pi(\delta)}$, a family of independent random variables satisfying

$$\mathbb{P}[Z_x^{\pi(\delta)} = 1] = 1 - \mathbb{P}[Z_x^{\pi(\delta)} = 0] = \pi(\delta).$$

This tells us, in particular, that our oriented L -dependent edge percolation stochastically dominates independent oriented edge percolation with a slightly lower edge density. Finally then we choose $\delta < 1$ sufficiently large that independent oriented edge percolation percolates (that such a δ exists is well known, see for example, Durrett 1988). Now choose ε sufficiently small that in the notation of Theorem 3.5, $\pi(1 - \varepsilon) > \delta$ and we have that for each $n \in \mathbb{Z}_+$, $Y(n, x)$ dominates the open sites at time n for *independent* oriented edge percolation. Now, for independent edge percolation the probabilities for different sites of being in the open cluster are positively correlated (by the FKG inequality) and our model is translation invariant and so the result now follows. \square

4 Ergodicity

In this section we prove Theorem 1.3 in three stages. Let us first remark that since $\underline{0}$ is a trap for the process

$$\mathcal{L}(X_t) = \mathbb{P}[X_0 = \underline{0}] \cdot \delta_{\underline{0}} + \mathbb{P}[X_0 \neq \underline{0}] \cdot \mathcal{L}(X_t | X_0 \neq \underline{0})$$

and so it suffices to prove the result for $\mathcal{L}(X_0) \in \mathcal{T}_{>0}$.

Our first lemma, which is an elementary corollary of the results of §2, deals with the case when the initial condition is a Poisson point process with intensity λdx .

Lemma 4.1. *Let $\{Y_t\}_{t \geq 0}$ be a (λ, ξ) -process with Y_0 distributed according to a Poisson point process with intensity λdx . Then $\mathcal{L}(Y_t)$ converges weakly to some measure $\bar{\nu} \in \mathcal{M}_1(\mathcal{N})$ as $t \rightarrow \infty$.*

Proof. Using Lemma 2.1, $\mathcal{L}(Y_s)$ is dominated by $\mathcal{L}(Y_0)$ for every $s \geq 0$. Since by Lemma 2.4, the dynamics of $\{Y_t\}_{t \geq 0}$ preserve the partial order of \mathcal{N} , combined with the Markov property this implies that

$$\mathcal{L}(Y_{t+s}) = \mathcal{L}^{\mathcal{L}(Y_s)}(Y_t) \leq \mathcal{L}^{\mathcal{L}(Y_0)}(Y_t) \tag{15}$$

for all $s, t \geq 0$. Thus the sequence $\{\mathcal{L}(Y_t)\}_{t \geq 0}$ is decreasing. It is also tight (since $\delta_0 \leq \mathcal{L}(Y_t) \leq \mathcal{L}(Y_0)$) and so $\mathcal{L}(Y_t)$ converges to some $\bar{\nu} \in \mathcal{M}_1(\mathcal{N})$ as $t \rightarrow \infty$ (see e.g. Kamae et al. 1977, Theorem 6). \square

In the second stage we will need two properties on the ancestral process, which are as follows: first, if the number of potential ancestors of a given particle doesn't die out, then it must tend to infinity. Secondly, these ancestors must be sufficiently spread out in space. This is to be expected since the location of particles is always dominated by a Poisson process. Let X be a (λ, ξ) -process and denote by $|X_t|$ the total number of particles in X_t .

Lemma 4.2. *Let $(X_t)_{t \geq 0}$ be a (λ, ξ) -process such that X_0 is almost surely a finite measure. Then the total mass of X_t converges to zero or to infinity almost surely as $t \rightarrow \infty$.*

Proof. If $|X_t|$ is less than or equal to $n \geq 1$, then irrespective of the path $(|X_s|)_{0 \leq s \leq t}$, the probability that $|X_{t+1}| = 0$ is uniformly bounded below (no matter how the particles of X_t are distributed in space) by a constant depending only on n . (To see this, consider killing them just one at a time.) The following result of Jagers (1992) then tells us that conditional on survival, $|X_t| \rightarrow \infty$ as $t \rightarrow \infty$. \square

Theorem 4.3 (Jagers 1992, Theorem 2). *Consider a sequence of random variables X_1, X_2, \dots defined in some probability space and taking values in $[0, \infty)$. Assume 0 is absorbing in the sense that $X_n = 0$ implies $X_{n+1} = 0$ and suppose that there is always a risk of extinction in the following way. For any x there is a $\delta > 0$ such that $\mathbb{P}[\exists n : X_n = 0 | X_1, \dots, X_k] \geq \delta$ on the event $\{X_k \leq x\}$. Then with probability 1 either*

- *there is an n such that all $X_k = 0$ for $k \geq n$ or*
- *$X_k \rightarrow \infty$ as $k \rightarrow \infty$.*

Now we show that the (λ, ξ) -process is asymptotically spread out in space on the event of survival. More formally we say for $k, l \in \mathbb{N}$ that a configuration $\phi \in \mathcal{N}$ is (k, l) -spread out if it contains k points which are pairwise separated at least by distance l .

Lemma 4.4. *Let $(X_t)_{t \geq 0}$ be a (λ, ξ) -process such that X_0 is finite almost surely. Then*

$$\mathbb{P}(|X_t| \geq 1, X_t \text{ is not } (k, l)\text{-spread out}) \rightarrow 0 \quad \text{as } t \rightarrow \infty \quad (16)$$

for all $k, l \in \mathbb{N}$.

Proof. We consider two cases and begin with $\xi((0, \infty) \times \{1\}) > 0$. In this case reproduction events of the form $(t, x, r, 1)$ have a positive rate and a quick eradication of a very large but spatially concentrated population is possible. Fix $k, l \in \mathbb{N}$ and let $(t_n)_{n \in \mathbb{N}}$ be a sequence converging to infinity such that, for each $n \in \mathbb{N}$, X_{t_n} is not $(k+1, l)$ -spread out. Then there exist points $x_1^{(n)}, \dots, x_k^{(n)} \in \mathbb{R}^d$ such that $X_{t_n} \subseteq \cup_{i=1}^k B_l(x_i^{(n)})$. With positive probability one time step further the population $X_{t_{n+1}}$ is extinct as a suitable sequence of events of the form $(t, x, r, 1)$ leaves $\cup_{i=1}^k B_l(x_i)$ empty with positive probability. Therefore the population goes eventually to extinction which proves (16) in the first case.

Now we turn to the case $\xi((0, \infty) \times \{1\}) = 0$. Consider the (λ, ξ) -process which starts with n particles in zero and let $n \rightarrow \infty$. The limiting process $(X_t^{(\infty)})_{t \geq 0}$ has infinitely many particles in zero at all times as the thinning probability is smaller than one at each reproduction event. This constant source for newborn particles guarantees that for $k, l \in \mathbb{N}$

$$\mathbb{P}(X_t^{(\infty)} \text{ is } (k, l)\text{-spread out}) \rightarrow 1 \quad \text{as } t \rightarrow \infty. \quad (17)$$

For fixed $\varepsilon > 0$ choose first $t_0 \geq 0$ and then $n_0 = n_0(t_0) \in \mathbb{N}$ such that

$$\mathbb{P}^{n\delta_0}(X_{t_0} \text{ is } (k, l)\text{-spread out}) \geq \mathbb{P}(X_{t_0}^{(\infty)} \text{ is } (k, l)\text{-spread out}) - \frac{\varepsilon}{2} \geq 1 - \varepsilon \quad (18)$$

for all $n \geq n_0$. Now on the event of survival, the total mass $|X_t|$ converges to infinity according to Lemma 4.2. Thus there will be more than n_0 particles and the process is (k, l) -spread out t_0 time units later with high probability. More formally, for $t > t_0$

$$\begin{aligned} & \mathbb{P}(|X_t| \geq 1, X_t \text{ is not } (k, l)\text{-spread out}) \\ & \leq \mathbb{P}(1 \leq |X_{t-t_0}| < n_0) + \mathbb{E}(\mathbf{1}_{|X_{t-t_0}| \geq n_0} \mathbb{P}^{X_{t-t_0}}(X_{t_0} \text{ is not } (k, l)\text{-spread out})) \\ & \leq \mathbb{P}(1 \leq |X_{t-t_0}| < n_0) + \mathbb{E}(\mathbf{1}_{|X_{t-t_0}| \geq n_0} \mathbb{P}^{|X_{t-t_0}| \delta_0}(X_{t_0} \text{ is not } (k, l)\text{-spread out})) \\ & \leq \mathbb{P}(1 \leq |X_{t-t_0}| < n_0) + \varepsilon. \end{aligned} \quad (19)$$

The last but one inequality reflects the fact that it is more difficult to be spread out in space if all initial particles are concentrated in one point. The right-hand side of (19) converges to ε as $t \rightarrow \infty$ according to Lemma 4.2. As $\varepsilon > 0$ was arbitrary, the lemma follows. \square

Our proof now proceeds in two steps. First, in Lemma 4.5, we consider the case where the initial value X_0 is dominated by a Poisson process and has the property that if we take a sufficiently ‘spread out sample’ from that Poisson process we are (asymptotically) sure to find a point of X_0 . (If the initial condition were itself Poisson then this would follow easily, but if we are to prove Theorem 1.3 for a general translation invariant initial condition then we must work a little harder.) Using the decomposition

into ancestral and ‘new’ individuals (blue and red particles) of Lemma 2.3, we then complete the proof of Theorem 1.2.

Lemma 4.5. *Suppose that $\mathcal{L}(X_0) \leq \mathcal{L}(Z)$ where Z is a Poisson point process with intensity λdx . Suppose further that there is a coupling of X_0 and Z such that $X_0 \leq Z$ and writing $I(x) = Z(\{x\}) - X_0(\{x\})$, there is a sequence of real numbers $(l_k)_{k \in \mathbb{N}}$ with $0 \leq l_k \rightarrow \infty$ as $k \rightarrow \infty$ and a sequence of positive random variables $(\Theta_k)_{k \in \mathbb{N}}$ with $\mathbb{E}[\Theta_k] \rightarrow 0$ as $k \rightarrow \infty$ for which*

$$\mathbb{P}[I(x) = 1 \text{ for all } x \in A^{(k)} | Z] \leq \Theta_k \quad a.s. \quad (20)$$

for all $A^{(k)} = \{x_1, x_2, \dots, x_k\} \subseteq Z$ satisfying $\|x_i - x_j\| \geq l_k$ for $1 \leq i \neq j \leq k$ and for all $k \in \mathbb{N}$. Then equation (6) holds.

Proof. We combine the couplings of Lemma 2.4 and Lemma 2.1 so as to construct simultaneously the process $(X_t)_{t \geq 0}$, the process $(Y_t)_{t \geq 0}$ of Lemma 4.1 and the stationary (λ, ξ) -revival process $(Z_t)_{t \geq 0}$ of Lemma 2.1. To see how to do this, assuming that we have coupled X_0 and Z in such a way that $X_0 \leq Z$, colour the individuals of X_0 red and the individuals of $Z \setminus X_0$ green. At a reproduction event corresponding to $(t, x, r, u) \in \Pi$, if at time $t-$ the ball $B_r(x)$ is empty, we introduce a Poisson number of ghost individuals with intensity $u\lambda \mathbf{1}_{B_r(x)}$ and colour them white. If $B_r(x)$ contains at least one red particle, then all new individuals produced are red, whereas if it contains green but no red, then all new offspring are green. Then the red individuals evolve according to the law of $(X_t)_{t \geq 0}$, the combined red and green give a copy of the process $(Y_t)_{t \geq 0}$ and the union of all red, green and white individuals gives us the process $(Z_t)_{t \geq 0}$ of Lemma 2.1, whose marginal distribution at each time is a Poisson point process of intensity λdx .

Writing $(D_t)_{t \geq 0}$ for the process of green individuals (which is the difference between the X and Y processes), our goal is to show that $\mathcal{L}(D_t)$ converges weakly to $\delta_{\underline{0}}$ as $t \rightarrow \infty$. Then the processes $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ successfully couple and the result follows from Lemma 4.1. To do this, for a fixed bounded set $A \in \mathcal{B}(\mathbb{R}^d)$, we consider the individuals of Z_t in A . We must show that as $t \rightarrow \infty$, the probability that all such individuals are either white or red tends to one. Evidently it is enough to consider a single individual. We show that given $\theta > 0$, there exists $T > 0$ such that for $t \geq T$, the probability that an individual picked from a given set A at time t is white or red is at least $1 - \theta$.

Choose then an individual from the set A at time t . To establish its colour, we trace backwards in time to establish all its potential ancestors at time zero, just as in the proof of Lemma 2.1. Recall our colouring rule. If an individual is a new ‘offspring’ produced during a given reproduction event affecting a ball $B_r(x)$ at time s , it will be red if *any* of the potential parents for that event, that is the individuals in $B_r(x)$ at time $s-$, is red. Similarly, for each such potential parent, they will be red if any of their

potential parents is red and so on. If any one of the potential ancestors at time 0 is red, then our chosen individual is red. Our task then is to show that as $t \rightarrow \infty$ either an individual is white, meaning that it has no potential ancestors, or one of its potential ancestors is red.

The assumption (20) reads as follows in terms of coloured individuals: That all individuals in a (k, l_k) -spread out sample of Z are green has probability at most Θ_k . Given $\theta > 0$, choose k big enough that $\mathbb{E}[\Theta_k] < \theta/2$. For $0 \leq s \leq t$, let A_s denote the potential ancestors of our chosen individual at time $t - s$ and write $|A_s|$ for the number of potential ancestors at time $t - s$. Considering the transitions of $(A_s)_{0 \leq s \leq t}$ at a reproduction event, we see from the reversibility of $(Z_t)_{t \geq 0}$ that $(A_s)_{0 \leq s \leq t}$ is a (λ, ξ) -process run until time t and started in A_0 . Consequently according to Lemma 4.4 there exists a T such that for all $t \geq T$ either A_t is extinct or A_t is (k, l_k) -spread out with probability at least $1 - \frac{\theta}{2}$. However if A_t is (k, l_k) -spread out, then all individuals in A_t are green with probability at most $\frac{\theta}{2}$. Thus the probability that either the chosen individual is white ($|A_t| = 0$) or that at least one of its potential ancestors is red is at least $1 - \theta$. This completes the proof. \square

Proof of Theorem 1.3

To complete the proof we must show that equation (6) still holds for *arbitrary* initial conditions. Note that by Lemma 2.3, the process is asymptotically bounded by a Poisson point process with intensity λdx , we have immediately that for all $s > 0$,

$$\limsup_{t \rightarrow \infty} \mathcal{L}(X_t) = \limsup_{t \rightarrow \infty} \mathcal{L}^{\mathcal{L}(X_t)}(X_s) \leq \mathcal{L}^Z(X_s) \xrightarrow{s \rightarrow \infty} \bar{\nu} \quad (21)$$

where the convergence is given by Lemma 4.1 and Z , as usual, is a Poisson point process with intensity λdx .

For the reverse inequality, the idea is to reduce to the previous case by running our process until time one. Using the coupling from Lemma 2.3, particles have then partially coupled with a Poisson process Z_1 . If we leave out those that have not had time to couple yet (the blue particles in that coupling), then the monotonicity in the initial data provided by Lemma 2.4 (together with the Markov property) show that ignoring these blue particles can only make the limit bigger. The configuration of particles that have coupled with the Poisson process (the red particles in the coupling of Lemma 2.3) provides us with a configuration of particles to which the Lemma 4.5 may be applied. It thus suffices to check that condition (20) is satisfied by the configuration of red particles at time 1.

In the notation of Lemma 2.3, let $R_t = X_t \setminus \check{X}_t$ be the configuration of red particles at time t (recall that \check{X}_t is just the set of blue particles, i.e., ancestral particles from time 0 which have never been affected by a reproduction event). We adapt the notation

of Lemma 4.5 and write $I_1(x) := Z_1(x) - R_1(x)$. We will show that for every $\varepsilon > 0$, there exists an $l = l(\varepsilon)$ such that if $\phi \in \mathcal{N}$ is the initial configuration of X_0 , then

$$\mathbb{E}^\phi \left[\prod_{j=1}^k I_1(x_j) \middle| Z_1 \right] \leq \varepsilon k + \prod_{j=1}^k \mathbb{E}^\phi [I_1(x_j) \middle| Z_1] \quad (22)$$

for all $k \in \mathbb{N}$, all $\phi \in \mathcal{N}$ and all $\{x_1, \dots, x_k\} \subset Z_1$ such that $\|x_i - x_j\| \geq l$ for all $i \neq j$. The notation \mathbb{E}^ϕ indicates $\mathcal{L}(X_0) = \delta_\phi$.

We shall say that there is a chain of birth events linking $B_1(x_i)$ to $B_l^c(x_i)$ in the time interval $[0, 1]$ if the union of the regions $B_r(x)$ for $(t, x, r, u) \in \Pi$ with $0 \leq t \leq 1$ corresponding to events in which at least one new individual is born (which we recall happens with probability $1 - \exp(-\lambda u C_d r^d)$) contains a path from $B_1(x_i)$ to $B_l^c(x_i)$. For fixed $\varepsilon > 0$, choose $l = l(\varepsilon)$ large enough that the probability of such a chain is at most ε . That this is possible follows from Assumption (2). With this, we obtain for every measurable function $F: \mathcal{N} \rightarrow [0, 1]$ and $x \in Z_1$

$$\begin{aligned} \mathbb{E}^\phi \left[I_1(x) \cdot F((Z_1 - R_1)|_{B_l^c(x)}) \middle| Z_1 \right] \\ \leq \varepsilon + \mathbb{E}^\phi [I_1(x) \middle| Z_1] \mathbb{E}^\phi \left[F((Z_1 - R_1)|_{B_l^c(x)}) \middle| Z_1 \right]. \end{aligned} \quad (23)$$

Induction on k proves (22).

Finally, we derive (20) from (22). Let $A_k = \{x_1, \dots, x_k\} \subset Z_1$ satisfy $\|x_i - x_j\| \geq l_k := l(\frac{1}{k^2})$ for all $1 \leq i \neq j \leq k$. In the following calculation, $\phi(\cdot - x)$ is the point measure ϕ shifted in space by $x \in \mathbb{R}^d$. Applying (22) with $\varepsilon := \frac{1}{k^2}$ and writing $\rho := \mathcal{L}(X_0)$, we obtain

$$\begin{aligned} \mathbb{P}[I_1(x) = 1 \forall x \in A_k \middle| Z_1] &= \int \mathbb{E}^\phi \left[\prod_{x \in A_k} I_1(x) \middle| Z_1 \right] \rho(d\phi) \\ &\leq \frac{1}{k^2} k + \int \prod_{i=1}^k \mathbb{E}^\phi [I_1(x_i) \middle| Z_1] \rho(d\phi) \\ &= \frac{1}{k} + \int \prod_{i=1}^k \mathbb{E}^{\phi(\cdot - x_1 + x_i)} [I_1(x_1) \middle| Z_1] \rho(d\phi). \end{aligned}$$

Indeed, note that $\mathbb{E}^\phi [I_1(x_i) \middle| Z_1]$ does not depend on the configuration of Z_1 as long as $x_i \in Z_1$. Thus,

$$\mathbb{E}^\phi [I_1(x_i) \middle| Z_1] = \mathbb{E}^\phi [I_1(x_i) \middle| Z_1(\cdot + x_1 - x_i)] \quad \text{a.s.} \quad (24)$$

Therefore, applying Hölder's inequality and using translation invariance of ρ ,

$$\begin{aligned} \mathbb{P}[I_1(x) = 1 \forall x \in A_k \middle| Z_1] &\leq \frac{1}{k} + \prod_{i=1}^k \left(\int \Psi(\phi(\cdot - x_1 + x_i))^k \rho(d\phi) \right)^{\frac{1}{k}} \\ &= \frac{1}{k} + \int (\Psi(\phi))^k \rho(d\phi) \end{aligned} \quad (25)$$

almost surely where $\Psi(\phi) := \mathbb{E}^\phi [I_1(x_1)|Z_1]$. Denote the right-hand side of (25) by Θ_k and note that $\Psi(\phi) < 1$ whenever $\phi \neq \underline{0}$ which holds ρ -almost surely. Thus, $\mathbb{E}[\Theta_k] \rightarrow 0$ as $k \rightarrow \infty$ which proves (20) and finishes the proof of Theorem 1.3. \square

5 Extinction

Theorem 1.2 shows that the process $\{X_t\}_{t \geq 0}$ survives if λ is sufficiently large. In this section we prove the complementary result that for sufficiently small λ the process dies out. Evidently this will be the case for $\lambda = 0$ for then no births take place. Using the monotonicity with respect to the parameter λ of Lemma 2.5, we see that the upper invariant measure of Theorem 1.3 exhibits a phase transition: there exists a critical $\lambda_0 \in [0, \infty]$ such that $\bar{\nu} = \delta_{\underline{0}}$ for all $0 \leq \lambda < \lambda_0$ and $\bar{\nu} \neq \delta_{\underline{0}}$ for all $\lambda_0 < \lambda < \infty$. Theorem 1.2 shows that $\lambda_0 < \infty$. The substance of Theorem 1.5 is that $\lambda_0 > 0$.

Proof of Theorem 1.5

In fact we have already done the bulk of the work in the proof of Lemma 2.1. It suffices, in our previous notation, to check that for each bounded set $A \in \mathcal{B}(\mathbb{R}^d)$ the process of potential ancestors of all the individuals in that set dies out in finite time (with probability one). In the proof of Lemma 2.1 we showed that the number of such potential ancestors is dominated by a continuous time branching process with branching rate given by equation (7) and mean number of offspring given by equation (8). Noting that the branching rate is independent of λ , it suffices to choose λ small enough that the mean number of offspring is less than one and Theorem 1.5 follows. \square

References

- [1] O Angel, N Berestycki and V Limic. Global divergence of spatial coalescents. In preparation.
- [2] K B Athreya and P E Ney. Branching Processes. Springer, 1972.
- [3] N H Barton, F Depaulis, and A M Etheridge. Neutral evolution in spatially continuous populations. *Theor Pop Biol*, 61:31–48, 2002.
- [4] M Birkner and A Depperschmidt. Survival and complete convergence for a spatial branching system with local regulation. *Ann. Appl. Probab.*, 17: 1777–1807, 2007.
- [5] B M Bolker and S W Pacala. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theor Pop Biol*, 52(3):179–197, 1997.

- [6] Richard Durrett. *Lecture notes on particle systems and percolation*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1988.
- [7] E Eller, J Hawks and J H Relethford. Local extinction and recolonization, species effective population size and modern human origins. *Human Biology*, 76(5): 689–709, 2004.
- [8] A M Etheridge. Survival and extinction in a locally regulated population. *Ann Appl Probab*, 14(1):188–214, 2004.
- [9] Alison M. Etheridge. Drift, draft and structure: some mathematical models of evolution. *Banach Center Publ*, 80:121–144, 2008.
- [10] Felsenstein, J. (1975). A pain in the torus: some difficulties with the model of isolation by distance. *Amer Nat*, 109:359–368.
- [11] G R Grimmett. *Percolation*, 2nd Edition. Springer, 1999.
- [12] M Hutzenthaler and A Wakolbinger. Ergodic behaviour of locally regulated populations. *Ann Appl Probab*, 17(2):474–501, 2007.
- [13] Peter Jagers. Stabilities and instabilities in population dynamics. *J. Appl. Probab.*, 29(4):770–780, 1992.
- [14] T. Kamae, U. Krengel, and G. L. O’Brien. Stochastic inequalities on partially ordered spaces. *Ann. Probab.*, 5(6):899–912, 1977.
- [15] R Law, D J Murrell, and U Dieckmann. On population growth in space and time: spatial logistic equations. *Ecology*, 84: 252–262, 2003.
- [16] T M Liggett *Interacting Particle Systems*. Springer, 1985.
- [17] T M Liggett, R H Schonmann, and A Stacey. Domination by product measures. *Ann. Probab.*, 25:71–95 (1997).
- [18] M Möhle and S Sagitov. A classification of coalescent processes for haploid exchangeable models. *Ann. Probab.*, 29:1547–1562, 2001.
- [19] V Limic and A Sturm. The spatial Lambda-coalescent. *Electron. J. Probab.*, 11(15):363–393, 2006.
- [20] J Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27:1870–1902, 1999.
- [21] S Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 26:1116–1125, 1999.

- [22] Sawyer, S. and Fleischmann, J. (1978). The maximal geographical range of a mutant allele considered as a subtype of a Brownian branching random field. *Proc Nat Acad Sci U.S.A.*, 76(2):872–875.
- [23] J Schweinsberg. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, 5:1–50, 2000.

Nathanaël Berestycki
Statistical Laboratory, University of Cambridge
Wilberforce Rd.
Cambridge CB3 0WB, U.K.

Alison M. Etheridge
Department of Statistics, University of Oxford,
Mathematical Institute
1 South Parks Road, Oxford, OX1 3TG, UK

Martin Hutzenthaler
Institut für Stochastik und Mathematische Informatik
Universität Frankfurt
Robert-Mayer-Str. 6-10
60325 Frankfurt/Main