A heterogeneous spatial model in which savanna and forest coexist in a stable equilibrium

Rick Durrett and Ruibo Ma
Dept. of Math, Duke, U.
Box 90320, Durham NC 27708-0320

January 29, 2019

Abstract

In work with a variety of co-authors, Staver and Levin have argued that savanna and forest coexist as alternative stable states with discontinuous changes in density of trees at the boundary. Here we formulate a nonhomogeneous spatial model of the competition between forest and savanna. We prove that coexistence occurs for a time that is exponential in the size of the system, and that after an initial transient, boundaries between the alternative equilibria remain stable.

1 Introduction

In a 2011 paper published in Science [24], Carla Staver, Sally Archibald and Simon Levin argued that tree cover does not increase continuously with rainfall but rather is constrained to low ($< 50\%$, "savanna") or high ($> 75\%$, "forest") levels. In follow-up work published in Ecology [25], the American Naturalist [26] and Journal of Mathematical Biology [23], they studied the following ODE for the evolution of the fraction of land covered by grass G, saplings S, and trees T:

$$
\begin{aligned}
\frac{dG}{dt} &= \mu S + \nu T - \beta G T \\
\frac{dS}{dt} &= \beta G T - \omega(G) S - \mu S \\
\frac{dT}{dt} &= \omega(G) S - \nu T
\end{aligned}
\tag{1}
$$

Here $\mu \geq \nu$ are the death rates for saplings and trees, and $\omega(G)$ is the rate at which saplings grow into trees. Fires decrease this rate of progression, and the incidence of fires is an increasing function of the fraction of grass, so $\omega(G)$ is decreasing. Studies suggest (see [26] for references) that regions with tree cover below about 40% burn frequently but fire is rare above this threshold, so they used an ω that is close to a step function.

Inspired by this work, Durrett and Zhang [16] considered two stochastic spatial models in which each site can be in state 0, 1, or 2: Krone's [18] model in which 0 = vacant, 1 =

1

juvenile, and 2 = a mature individual capable of giving birth, and the Staver-Levin foresst model in which 0 = grass, 1 = sapling, and 2 = tree. Theorem 1 in [16] shows that if (0,0,1) is an unstable fixed point of (1) then when the range of interaction is large, there is positive probability of grass and trees surviving starting from a finite set and there is a stationary distribution in which all three types are present. The result they obtain is asymptotically sharp for Krone's model. However, in the Staver-Levin forest model, if $(0,0,1)$ is attracting then there may also be another stable fixed point for the ODE, and as their Theorem 3 shows, in some of these cases there is a nontrivial stationary distribution.

1.1 Models

Recently Touboul, Staver, and Levin [28] have investigated a number of modifications of the three species system (1). Variants of the ODE that add a fourth type called forest trees display a wide variety of behaviors including limit cycles, homoclinic, and heteroclinic orbits. Simulations of the spatial version of ODE systems with periodic orbits, see [9], suggest these systems will have stationary distributions that are patchy and with local densities oscillating when the scale of observation is smaller than what physicists call the correlation length, see Figure 4 in [13]. Proving that this occurs is a very difficult problem. Here we will instead focus our attention on a two species system analyzed in [28]. We have changed notation used in [28] to make it more easily comparable to the system studied here

$$\frac{dF}{dt} = \phi_1(F)G - \phi_0(G)F \quad \text{where} \quad G = 1 - F \tag{2}$$

In our stochastic spatial model, the state of site x at time t can be $\xi_t(x) = 0$ (grass) or $\xi_t(x) = 1$ (tree). In formulating our model we have two space scales: the dispersal scale L, which might be hundreds or thousands of feet, and the continental scale ML, which might be thousands of miles. From a biological point of view, it would be natural to have the dynamics taking place on a two-dimensional square with Dirichlet boundary conditions, i.e., the sites outside the square which correspond to oceans are vacant. However, points near the boundary do not have the same number of neighbors and translation invariance is lost.

To facilitate taking the limit as $L \to \infty$ we will define our model on a torus with small lattice spacing: $(\mathbb{Z}^2/L \bmod M)^2$ As in (2) each site x changes state at a rate dictated by the density of its competitors in a neighborhood of x. To have rotational symmetry in the limit $L \to \infty$, we will define the neighborhood to be $x + \mathcal{N}$ where $\mathcal{N} = \mathbb{Z}^2 \cap D(0,1)$ with $D(x, r) = \{y : \|y - x\| \le r\}$ and $\|x\| = (x_1^2 + x_2^2)^{1/2}$.

1.1.1 Concrete model

Motivated by the large variability of rainfall amounts in South America, we will consider a spatially heterogeneous model. Let $f_i(x)$ be the fraction of type i sites in $x + \mathcal{N}$ and let $\alpha, \beta > 0$. The transition rates in our first model are :

$$\begin{aligned} 0 \to 1 \quad &\text{at rate } A(x)f_1^\alpha(x) \\ 1 \to 0 \quad &\text{at rate } B(x)(1 - f_1(x))^\beta \end{aligned}$$

2

Suppose for the moment that the system is homogeneous: $A(x) \equiv A$ and $B(x) \equiv B$. In this case, the first step in analyzing the process is to write the mean-field ODE which comes from assuming that state of the system at time t is a product measure in which 1's have density $u(t)$. The mean field equation is

$$\frac{du}{dt} = A(1-u)u^\alpha - Bu(1-u)^\beta$$
$$= u(1-u)[Au^{\alpha-1} - B(1-u)^{\beta-1}] \tag{3}$$

From the second from of the equation, it is easy to see that

- If $0 < \alpha < 1$ then as $u \to 0$, $u^{\alpha-1} \to \infty$ and $-B(1-u)^{\beta-1} \to B$, so 0 is unstable.

- If $\alpha > 1$ then as $u \to 0$, $u^{\alpha-1} \to 0$ and $-B(1-u)^{\beta-1} \to -B$, so 0 is stable.

Likewise if $0 < \beta < 1$, 1 is unstable; if $\beta > 1$, 1 is stable.

Consider the four combinations, we show in Section 2

Case 1. If $\alpha, \beta \in (0,1)$, 0, and 1 are unstable fixed points. There is a unique interior fixed point and it is attracting.

Case 2. If $\alpha, \beta > 1$, 0 and 1 are stable fixed points. There is a unique interior fixed point and it is unstable.
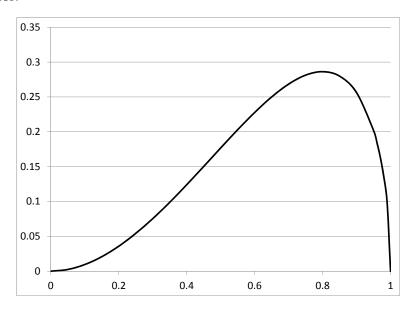

Figure 1: Graph of $x^{\alpha-1}(1-x)^{1-\beta}$ when $\alpha = 3$ and $\beta = 1/2$. The maximum occurs at $(3-1)/(3-0.5) = 0.8$. The value there is 0.286217, so solutions exist if $A/B > 3.4865$.

Case 3. If $\alpha > 1 > \beta > 0$, 0 is stable and 1 is unstable. let $\theta = A/B$, and let

$$w = (\alpha - 1)/(\alpha - \beta) \in (0,1),$$

be location of the local maximum of $x^{\alpha-1}(1-x)^{1-\beta}$. If

$$\theta w^{\alpha-1}(1-w)^{1-\beta} > 1$$

3

then there will be fixed points v_1 and v_2 with $0 < v_1 < w < v_2 < 1$. Fixed points 0 and v_2 are stable, while v_1 and 1 are unstable.

Case 4. If $\beta > 1 > \alpha > 0$ then the situation is the same as Case 3 but the roles of 0 and 1 are interchanged.

Heuristic analyses of Durrett and Levin [12] suggest that in Case 1, that 0's and 1's will coexist in a equilibrium with very little spatial structure, so this case is not relevant. In the other three cases we have bistability (two stable fixed points separated by an unstable one). so we will focus on Case 3. Case 4 is the same as Case 3 with the roles of 1 and 0 interchanged. The same methods will work in Case 2, but that situation is less interesting since the stable fixed points ar at 0 and 1.

1.1.2 General model

To obtain our conclusions it is not necessary to assume the dynamics have such a simple form, so we will consider a general formulation

$$
\begin{aligned}
0 \to 1 \quad &\text{at rate } A(x)G(f_1(x)) \\
1 \to 0 \quad &\text{at rate } B(x)H(1 - f_1(x))
\end{aligned}
\tag{4}
$$

Here G, $H \geq 0$ and $G(0) = H(0) = 0$. It is natural to impose some smoothness assumptions on G and H. We will not discuss what they should be because we will soon assume G and H are polynomials. For $y > 0$ let $g(y) = G(y)/y$ and $h(y) = H(y)/y$. Using this notation the mean-field equation is, when $u \in (0, 1)$,

$$
\begin{aligned}
\frac{du}{dt} &= A(1 - u)G(u) - BH(1 - u) \\
&= u(1 - u)[Ag(u) - Bh(1 - u)]
\end{aligned}
\tag{5}
$$

Our analysis will not work if there are more than two interior fixed points. Letting $\theta = A/B$ and writing the condition for a fixed point as

$$
\theta \frac{g(u)}{h(1 - u)} = 1
$$

we see that in order to have at most two interior fixed points for all C there must be a w so that

$$(M1) \qquad\qquad g(u)/h(1 - u) \text{ is } \begin{cases} \text{increasing} & u < w \\ \text{decreasing} & u > w \end{cases}$$

This assumption is not intuitive but it is necessary to have a transition from 0 interior fixed points to 2.

To carry out our proofs we need to reduce our generality a little.

(M2) We assume that there are $J, K < \infty$ so that $G(u)$ and $H(u)$ have the form

$$G(u) = \lambda_G \sum_{j=1}^{J} p_j^G \binom{J}{j} u^j (1-u)^{J-j}$$

$$H(u) = \lambda_H \sum_{k=1}^{K} p_k^H \binom{K}{k} u^k (1-u)^{K-k} \tag{6}$$

To explain the reason for this assumption, note that if $G(u)$ has this form, the birth dynamics can be constructed as follows: at rate λ we pick J neighbors. A change occurs with probability p_k (with $p_0 = 0$) if exactly k of the chosen neighbors are in state 1. A similar recipe works for the death dynamics.

As we will explain in Section 3, this will enable us to define a dual process an done in [11] and prove that when space and time are rescaled the system converges to the solution of an integro-differential equation. In a sense the assumption (M2) entails no loss of generality. If a is any continuous function and we set

$$a_m(u) = \sum_{k=0}^{m} a(k/m) \binom{m}{k} u^k (1-u)^{m-k}$$

then a result in Example 2.2.1 of [10] implies $a_m(u) \to a(u)$ uniformly on $[0,1]$ as $m \to \infty$. In words, we can approximate any continuous function by one of the form (6) with only a small error, so in a sense we can handle any function.

1.2   Results

To analyze the general model we will use a three step procedure that we have employed many times. See e.g., [2, 5, 7, 11, 16, 22]

Step 1. Show that the particle system converges to a deterministic limit that is a partial differential equation (PDE) or IDE.

Step 2. Obtain results that describe the asymptotic behavior of the limiting equation.

Step 3. Use a block construction (see [4] or [8]) to prove the existence of a stationary distribution.

The proof here will be based on the approach in [22]. The new feature here is the spatial heterogeneity.

We assume that the climatic conditions vary on a continental scale. To control the variability on intermediate scales, we make the first of several assumptions. Let $\mathcal{C}^2$ be the collection of functions f that have continuous derivatives $\partial f / \partial x$, $\partial f / \partial y$, $\partial^2 f / \partial x^2$, $\partial^2 f / \partial x \partial y$, and $\partial f / \partial y^2$,

(H1) $A(x) = a(x/M)$ and $B(x) = b(x/M)$ where a and b are positive and $\mathcal{C}^2$.

Since a and b are defined on $(\mathbb{R} \mod 1)^2$ it follows that their derivatives are bounded.

The goal of this study is to show:

- In our heterogeneous system it is possible to have stable coexistence of grassland and forest, which, as we will explain later, will not occur in a homogeneous system.

- Coexistence occurs because we will have regions that are grassland, i.e., mostly 0's, and regions of forest that are mostly 1's.

- The boundaries between the two regions can be predicted from the coefficients $A(x)$ and $B(x)$. Specifically, there is a constant θ_1 so that $\{x : A(x)/B(x) < \theta_1\}$ is grassland, and $\{x : A_1(x)/B_1(x) > \theta_1\}$ is forest.

- Viewed on the continental scale the boundaries will be stable once equilibrium is reached. On the dispersal scale, there will be transition zones between the two regions. However, our methods do not tell anything about the nature of these transition zones.

Our analysis is done in four steps, the first three of which are the same as above.

Step 1. We will let $L \to \infty$ and scale space by L so that the particle system can be replaced by an integro-differential equation. Before we can state our long-range limit theorem we have to explain what it means for a sequence of particle systems $\xi_t^L : \mathbb{Z}^2/L \to \{0,1\}$ to converge to a function $u(t,x)$. To do this we pick a $\gamma \in (0, 1/4]$ and tile the plane with squares $[y, y + L^{-\gamma}) \times [z, z + L^{-\gamma})$ in such a way that the origin $(0,0)$ is at the lower-left corner of one of the squares. Given an $x \in \mathbb{R}^2$ let $R_L(x)$ be the square to which x belongs and let $u_\xi^L(t,x)$ be the fraction of sites y in $R_L(x)$ with $\xi_t(y) = 1$. We say that ξ_t^L converges to $u(t,x)$ and write $\xi_t^L \to u(x,t)$ if for any $K < \infty$

$$\sup_{\|x\| \leq K} |u_\xi^L(t,x) - u(t,x)| \to 0 \quad \text{in probability as } L \to \infty. \tag{7}$$

Theorem 1. *Let $\bar{u}(t,x)$ be the average value of $u(t,y)$ over $D(x,1)$, the ball of radius 1 around x. If $\xi_0^L \to u_0(x)$ a continuous function then we have $\xi_t^L \to u(x,t)$ the solution of*

$$\frac{du}{dt} = A(x)(1 - u(t,x))G(\bar{u}(t,x))$$
$$- B(x)u(t,x)H(1 - \bar{u}(t,x)) \qquad u(0,x) = u_0(x) \tag{8}$$

Although it takes some work to prove this, the limit equation is what one should expect from the definition of the process. The site x changes $0 \to 1$ at rate $A(x)G(f_1(x))$, while x changes $1 \to 0$ at rate $B(x)H(1 - f_1(x))$. The main step in establishing the limit result, is to show that if L is large then $f_1(x) \approx \bar{u}(t,x)$.

Remark. While the statement claims convergence for each fixed t, and this all we need for the block construction, it is not hard to show that

$$\sup_{0 \leq t \leq T} \sup_{\|x\| \leq K} |u_\xi^L(t,x) - u(t,x)| \to 0 \quad \text{in probability as } L \to \infty. \tag{9}$$

To do this one uses convergence for $t_i = i\epsilon$, and then boundedness of the jump rates to conclude that in between t_i only a small fraction of the sites flip,

Step 2. Heuristic arguments in Durrett and Levin [12] predict that when there is bistability in the limiting ODE (i.e., two stable fixed points) then the density of the equilibrium of the spatial model will be the stronger equilibrium, which is defined by looking at the sign of the speed of the traveling wave solution for the one dimensional equation

$$u(t, x) = v(x - \rho t)$$

connecting the two fixed points. If $v(-\infty) = v_2$ and $v(\infty) = 0$ then v_2 is the stronger equilibrium when the speed $\rho > 0$ and 0 is the stronger equilbrium when the speed $\rho < 0$. Since in Step 1 we have taken a limit to get an IDE, we do not need to rely on this heuristic principle. We can use results of Weinberger [29] to prove that this is true for the IDE. That is, there is a constant θ_1 so that 0 is the stronger equilibrium and the limit of the IDE if $A/B < \theta_1$, and v_2 is the stronger equilibrium if $A/B > \theta_1$.

Step 3. To prove results about the homogeneous system with $A(x) \equiv A$ and $B(x) \equiv B$, we use a "block construction." This application is slightly different than usual since we are on a torus and we are proving prolonged survival rather than the existence of a stationary distribution. However, the general outline of the proof is the same.

Theorem 2. *Let $\epsilon > 0$. Suppose $A/B > \theta_1$ and the initial condition satisfies the hypothesis of Theorem 6. There is a $t_0 < \infty$ and an $\eta > 0$ so that if L and M are large enough then with high probability at times $t_0 \le t \le \exp(\eta M^2)$ the fraction of the torus occupied by 1's is $\in [v_2 - \epsilon, v_2 + \epsilon]$.*

Theorem 3. *Suppose $A/B < \theta_1$ and the initial condition satisfies the hypothesis of Theorem 7.There is a $t_0 < \infty$ and an $\eta > 0$ so that if L and M are large enough then with high probability at times $t_0 \le t \le \exp(\eta M^2)$ the fraction of the torus occupied by 1's is $\le \epsilon$.*

Together these result implies in particular that coexistence does not occur in the homogeneous system when $A/B \ne \theta_1$. Presumably there is no coexistence when $A/B = \theta_1$, but seems to be a difficult problem.

Step 4. To describe the behavior of the heterogeneous system now, we divide the plane into regions where $A(x)/B(x) < \theta_1$ and $A(x)/B(x) > \theta_1$. To avoid pathologies we assume

(H2) the boundaries $A(x)/B(x) = \theta_1$ between these regions are given by a finite collection of C^1 curves. The gradient $\nabla(A/B) \ne 0$ on each curve, so they do not touch each other.

To prove our result we need another assumption

(H3) There is a $\delta > 0$ so that: In each region $A(x)/B(x) > \theta_1 + \delta$ the initial condition satisfies the hypothesis of Theorem 6. In each region $A(x)/B(x) < \theta_1 - \delta$ the initial condition satisfies the hypothesis of Theorem 7.

Theorem 4. *Let $\delta > 0$ and $\epsilon > 0$. There is a time $t_0 < \infty$ and $\eta > 0$ so that if L and M are large then with high probability for $t_0 \le t \le \exp(\eta M^2)$*

(i) In each of the regions $A(x)/B(x) > \theta_1 + \delta$, the fraction of 1's is $\in [v_2 - \epsilon, v_2 + \epsilon]$

(ii) In each of the regions $A(x)/B(x) < \theta_1 - \delta$, the fraction of 1's is $\le \epsilon$

The remainder of the paper is devoted to the proofs of Theorems 2, 3, and 4. We will only give the details for Theorem 2, and part (i) of Theorem 4. The other results follow by interchanging the roles of 0 and 1.

2   Analysis of the mean-field ODE

In this section we study the ODE:

$$\frac{du}{dt} = A(1-u)u^\alpha - Bu(1-u)^\beta$$
$$= u(1-u)[Au^{\alpha-1} - B(1-u)^{\beta-1}]$$

If $\alpha < 1$ then as $u \to 0$, $u^{\alpha-1} \to \infty$ and $-B(1-u)^{\beta-1} \to B$, so 0 is unstable. If $\alpha > 1$ then as $u \to 0$, $u^{\alpha-1} \to 0$ and $-B(1-u)^{\beta-1} \to -B$, so 0 is stable. Likewise if $\beta < 1$, 1 is unstable; if $\beta > 1$, 1 is stable.

Caae 1. $\alpha < 1$, $\beta < 1$. By calculations above both boundary fixed points are unstable. Interior fixed point satisfies

$$D(u) \equiv Au^{\alpha-1} - B(1-u)^{\beta-1} = 0 \tag{10}$$

We have $D(0) = \infty$ and $D(1) = -\infty$. Differentiating gives

$$D'(u) = (\alpha-1)Au^{\alpha-2} + (\beta-1)B(1-u)^{\beta-2} < 0$$

so $D(u)$ is decreasing. From this we see that there is a unique interior fixed point. Since the two boundary fixed points are unstable, it is attracting.

Case 2. $\alpha > 1$, $\beta > 1$. By calculations above both boundary fixed points are stable. Again the interior fixed point satisfies

$$D(u) \equiv Au^{\alpha-1} - B(1-u)^{\beta-1} = 0$$

We have $D(0) = -B$ and $D(1) = A$ so there is a solution.

$$D'(u) = (\alpha-1)Au^{\alpha-2} + (\beta-1)B(1-u)^{\beta-2} > 0$$

so $D(u)$ is increasing and the solution is unique. Since the two boundary fixed points are unstable, it is attracting.

Case 3. $\alpha > 1$, $\beta < 1$. By calculations above 0 is stable and 1 is unstable. If we let $\theta = A/B$ the equation for an interior fixed point (10) can be rewritten as

$$Bu(1-u)^\beta f(u) = 0 \quad \text{where} \quad f(u) = [Cu^{\alpha-1}(1-u)^{1-\beta} - 1].$$

To find solutions of $f(u) = 0$, we again differentiate

$$\frac{d}{du}u^{\alpha-1}(1-u)^{1-\beta} = (\alpha-1)u^{\alpha-2}(1-u)^{1-\beta} - (1-\beta)u^{\alpha-1}(1-u)^{-\beta}$$
$$= u^{\alpha-2}(1-u)^{-\beta}[(\alpha-1)(1-u) - (1-\beta)u]$$
$$= u^{\alpha-2}(1-u)^{-\beta}[(\alpha-1) - (\alpha-\beta)u]$$

so the maximum occurs at $w = (\alpha-1)/(\alpha-\beta) \in (0,1)$.

$$f(w) = Bw(1-w)^\beta[Cw^{\alpha-1}(1-w)^{1-\beta} - 1]$$

If C is small $f(w) < 0$ and hence $f(u) < 0$ on $(0,1)$. In this case 0 is a stable fixed point, and 1 is an unstable fixed point If on the other hand

$$Cw^{\alpha-1}(1-w)^{1-\beta} > 1$$

then $f(w) > 0$. Since $f(0) < 0$ and $f(1) < 0$ there will be fixed points v_1 and v_2 with $0 < v_1 < w < v_2 < 1$. Since $f'(u) > 0$ for $u < w$ and $f'(u) < 0$ for $u > w$ there are no other fixed points and we have

$$
\begin{array}{cccc}
f(u) & < 0 & > 0 & < 0 \\
\text{on} & (0, v_1) & (v_1, v_2) & (v_2, 1)
\end{array}
$$

so 0 and v_2 are stable, v_1 and 1 are unstable. Plugging in the value of u_0 we see that the condition for bistablity (i.e., the existence of two attracting fixed points) is

$$\theta > \theta_0 \equiv \frac{(\alpha - \beta)^{\alpha - \beta}}{(\alpha - 1)^{\alpha - 1}(1 - \beta)^{1 - \beta}} \tag{11}$$

3 Hydrodynamic limit

We begin by considering the homogeneous case $A(x) \equiv A$ and $B(x) \equiv B$. Recall that we have assumed in (6) that

$$G(u) = \lambda_G \sum_{j=0}^{J} p_j^G \binom{J}{j} u^j (1-u)^{J-j}$$

$$H(u) = \lambda_H \sum_{k=0}^{K} p_k^H \binom{K}{k} u^k (1-u)^{K-k}$$

where $p_0^G = p_0^H = 0$. For each site x we have a Poisson process B_n^x with rate λ_G and a Poisson process D_n^x with rate λ_H. Let $\mathcal{N} = (\mathbb{Z}^2 / L) \cap (D(0,1)) - \{0\})$

- At each time B_n^x we have J random variables $U_n^{x,1}, \ldots U_n^{x,J}$ that are drawn at random from from $x + \mathcal{N}$ without replacement and a random variable $U_n^{x,0}$ that is uniform on $(0,1)$. If x is vacant and j of these J sites are occupied at time B_n^x then x becomes occupied if $U_n^{x,0} < p_j^G$.

- At each time D_n^x we have K random variables $\bar{U}_n^{x,1}, \ldots \bar{U}_n^{x,K}$ that are drawn at random from $x + \mathcal{N}$ without replacement and a random variable $\bar{U}_n^{x,0}$ that is uniform on $(0,1)$. If x is occupied and k of these K sites are vacant at time D_n^x then x becomes vacant if $\bar{U}_n^{x,0} < p_k^H$.

Given this structure, which is called a **graphical representation**, we can define a dual process $\mathcal{I}_s^{t,x}$ starting at x at time t and working backwards in order to determine the state of x at time t. Nothing happens until S_1 the first time s so that there is a point $y \in I_s^{t,x}$ and an n so that $B_n^y = t - s$ or $D_n^y = t - s$.

- In the first case we add $U_n^{y,1}, \ldots U_n^{y,J}$ to $\mathcal{I}_{S_1}^{t,x}$.

- In the second we add $\bar{U}_n^{y,1}, \ldots \bar{U}_n^{y,K}$ to $\mathcal{I}_{S_1}^{t,x}$.

We continue in this way: adding points when there is an arrival in a Poisson process at some $y \in I_s^{t,x}$. Durrett and Neuhauser [11] call this the **influence set** because we only need to know the values at $\mathcal{I}_t^{t,x}$ to determine the value of x at time t. To actually compute the state of x at time t we need the values of the random variables uniform on $(0,1)$ associated with the jumps in $\mathcal{I}_s^{t,x}$, $s \leq t$. Even though these were taken from the graphical representation, the result is the same in distribution if we use a new i.i.d. sequence. In a slight abuse of notation we will use $\mathcal{I}^{t,x}$, $s \leq t$ to denote the influence set plus the i.i.d. sequence. Following [5] we call this the **computation process**.

The analysis of the influence set in this example is particularly simple because points do not move. If at some time, a point that we want to add is already in $\mathcal{I}_s^{t,x}$ we say that a **collision occurs**. Let $\mathcal{C}_t^x$ be the event that this occurs by time t. Given two dual process $\mathcal{I}_s^{t,x}$ and $\mathcal{I}_s^{y,t}$ we say that a **collision occurs at time** s if some site $z \in \mathcal{I}_s^{t,x} \cup \mathcal{I}_s^{y,t}$ gives birth onto a site in $\mathcal{I}_s^{t,x} \cup \mathcal{I}_s^{y,t}$. Let $\mathcal{C}_t^{x,y}$ be the event that this occurs by time t.

Lemma 1. *As $L \to \infty$, $P(\mathcal{C}_t^0) \to 0$, and $\sup_x P(\mathcal{C}_t^{0,x}) \to 0$.*

Proof. It is clear from the definitions of $\mathcal{C}_t^0$ and $\mathcal{C}_t^{0,x}$ that it suffices to prove the second statement. To bound the growth of $|\mathcal{I}_s^{0,t} \cup \mathcal{I}_s^{t,x}|$ we compare with a branching process Z_t with $Z_0 = 2$ in which particles give birth at rate λ_G to J offspring, and at rate λ_H to K offspring. A standard calculation for branching processes shows that

$$EZ_t \leq 2e^{(\lambda_G J + \lambda_H K)t} \tag{12}$$

so $P(Z_t^x > L^{1/2} - 2) \to 0$ as $L \to \infty$. When there are $\leq L^{1/2} - 2$ jumps in Z_s we have $|Z_t| \leq (J+K)L^{1/2}$. Since $D(0,1)$ has area π, it follows that if L is large then the probability of a collision

$$\leq (J+K)L^{1/2} \cdot \frac{(J+K)L^{1/2}}{(\pi/2)L^2} \to 0. \tag{13}$$

The first factor bounds the number of births, the second assumes the worst case scenario being the situation in which all points are within distance 1 of each other. The upper bound is independent of x, so we have proved the desired result. $\qquad\square$

The developments in this section closely follow the proofs in [22]. The first step in proving Theorem 1 is to prove a simpler hydrodynamic limit. This version is not useful for the block construction because the assumption does not match what we can prove about the process at time t.

Theorem 5. *Suppose that the sites in ξ_0^L are independent with $P(\xi_0^L(x) = 1) = u^L(0,x)$, a continuous function, and let $u^L(t,x) = P(\xi_t^L(x) = 1)$. Then we have $u^L(t,x) \to u(t,x)$ the unique continuous solution of*

$$\frac{du}{dt} = A(x)(1 - u(t,x))G(\bar{u}(t,x))$$
$$- B(x)u(t,x)H(1 - \bar{u}(t,x)) \qquad u(0,x) = u_0(x)$$

taking values in $[0,1]$.

Proof. To take the limit as $L \to \infty$, it is useful to make the dependence on L explicit. With Lemma 1 established, it follows easily that if $x(L) \in \mathbb{Z}^d/L$ with $x(L) \to x$ then $\mathcal{I}_L^{t,x(L)}(s)$, $0 \le s \le t$ converges in distribution to a branching random walk $\mathcal{I}^{t,x}(s)$, $0 \le s \le t$ with values in $\mathbb{R}^2$ in which particles give birth at rate λ_G to J particles and at rate λ_H to K particles and new born particles are displaced from their parents by independent amounts that are uniform on $D(0,1)$.

From the last conclusion it follows easily that $u^L(t, x(L)) \to u(t, x)$, a limit that is jointly continuous in t and x. A simple generator calculation, see Section 2.3 in [22], shows that the limit satisfies the ODE. Uniqueness follows from the observation that if $u(t, x)$ satisfies the ODE then $u(t, x)$ can be computed from the initial condition $u(0, y)$ by running the dual process. $\qquad\square$

Proof of Theorem 1. Recall that we partition the space in small square with side $L^{-\gamma}$, with $\gamma \in (0, 1/4]$. We let $R_L(x)$ be the small square containing x and let $u_\xi^L(t, x)$ be the fraction of sites in $R_L(x)$ with $\xi_t(y) = 1$. There are two parts to the proof.

Part 1. Show that the mean $E u_\xi^L(t, x) \to u(t, x)$.

To begin to do this, we need with some definitions.

Let $\mathcal{Q}_L(\mathcal{I}_L^{t,x}, \xi_0^L)$ be the result $\in \{0, 1\}$ of using the computation process $\mathcal{I}_L^{t,x}$ on the initial configuration ξ_0^L. Let $q_L(t, x, \xi_0^L) = P(\mathcal{Q}_L(\mathcal{I}_L^{t,x}, \xi_0^L) = 1 | \xi_0^L)$

Let $\mathcal{Q}_L(\mathcal{I}_L^{t,x}, u_0)$ be the result $\in \{0, 1\}$ of using the computation process $\mathcal{I}^{t,x}$ on inputs that are product measure with density $u_0(x)$. Let $q_L(t, x, u_0) = P(\mathcal{Q}_L(\mathcal{I}_L^{t,x}, u_0) = 1)$.

In the absence of collisions the computation process gives the correct value for $\xi_t^L(x)$ so

$$q_L(t, x, u_0) - u^L(t, x) \to 0, \tag{14}$$

and it follows from Theorem 5 that $q_L(t, x, u_0) \to u(t, x)$, the solution of the ODE.

The next step is to compare the results of the computation process when it is applied to the initial configuration ξ_0^L and to product measure with density u_0. When the dual process converges to a branching Brownian motion this is not hard. If $\eta_L = L^{-\alpha}$ with $\alpha/2 < \gamma$ then with probability tending to 1 there is no branching in the dual on $[t-\eta_L, t]$. Let $y_L = i(t-\eta_L)$ be the location of a lineage in the dual at time $t-\eta_L$. In time η_L lineages will move a distance $L^{-\alpha/2} \gg L^{-\gamma}$. This implies that conditional on the lineage ending in a given small square $R_L(z)$ its location will almost be uniform on $R_L(z)$. Due to our assumption about the initial condition, the probability of the lineage landing on a 1 is close to $u_0(y_L)$.

When the lineages don't move after birth, we need another approach. Let $0 < \beta < \gamma$. Let $\mathcal{M} = (\mathbb{Z}^2/L^\beta) \cap (D(0,1) - \{0\})$. Let V_n^x be uniform on $x + \mathcal{M}$. Let

$$W_n^x \text{ be uniform on } (\mathbb{Z}^2/L) \cap [-L^{-\beta}/2, L^{-\beta}/2]^2.$$

The distribution of $V_n^x + W_n^x$ is very close to that of U_n^x. The only differences occur within distance $L^{-\beta}$ of the boundary of $D(0, 1)$.

Let $\mathcal{J}_L^{t,x}$ be the dual process in which the U_n^x have been replaced by V_n^x. Since the distances between corresponding points in $\mathcal{I}_L^{t,x}(t)$ and $\mathcal{J}_L^{t,x}$ are of order $L^{-\beta}$, if we let $\bar{q}_L(t, x, u_0) = P(\mathcal{Q}_L(\mathcal{J}_L^{t,x}, u_0) = 1)$ then

$$q_L(t, x, u_0) - \bar{q}_L(t, x, u_0) \to 0 \tag{15}$$

11

To complete the proof we will create a coupling of $\mathcal{Q}_L(\mathcal{I}_L^{t,x}, \xi_0^L)$ and $\mathcal{Q}_L(\mathcal{J}_L^{t,x}, u_0)$. To couple the influence sets the jumps in $\mathcal{I}_L^{t,x}$ are according to $V_n^x + W_n^x$. Number the particles in $\mathcal{I}_L^{t,x}(t)$ and $\mathcal{J}_L^{t,x}(t)$ in the order that they were created. Each particle $\mathcal{I}_L^{t,x}(i,t)$ has a unique path of jumps that led its location.

To simplify notation let $I_i = \mathcal{I}_L^{t,x}(i,t)$ and $J_i = \mathcal{J}_L^{t,x}(i,t)$. By construction $|I_i - J_i| = O(L^{-\beta})$ with the difference w_i being a sum of independent W_n^x that are independent of J_i). Intuitively, the small squares have side $L^{-\gamma} \ll L^{-\beta}$ so if we let $R_L(z)$ be one of the small squares near J_i (i.e., within $O(L^{-\beta})$ and condition on $w_i + \mathcal{J}_L^{t,x}(i,t) \in R_L(z)$ then the distribution on $R_L(z)$ will almost be uniform. To argue a little more carefully, suppose for concretenesss that w_i is the sum of two independent copies of W_n^x. In this case the density of w_i on $(\mathbb{Z}^2/L) \cap [-L^{-\beta}, L^{-\beta}]^2$ that has the form $\phi(z_1)\phi(z_2)$ where the $\phi(z)$ have a triangular shape. The claimed uniformity clearly will not hold if $R_L(z)$ overlaps the boundary of the support, but it will if the distance to the boundary is $\gg L^{-\gamma}$.

From the last paragraph we see that conditional on J_i the value of ξ_0^L, call it χ_i, will be 1 with probability $\approx u_0(J_i)$. To see that the χ_i are almost independent coin flips, work backwards in time starting with the last jump in the dual seen in $\mathcal{I}_L^{t,x}(s)$ before time t. The particles created at that time will have values of W_n^x that have not previously been seen in the dual. This means that the last set of coin flips are independent of each other and of the ones that came before them. Induction completes the proof that the coin flip variables we have constructed are asymptotically independent. This allows us to conclude that

$$\bar{q}_L(t, x, u_0) - q_L(t, x, \xi_0^L) \to 0 \tag{16}$$

and we have completed part 1 of the proof.

Part 2. Estimate the variance of $u_\xi^L(t, x)$.

Let $S_L(x) = \sum_{y \in R_L(x)} i_y$ where $i_y = 1_{\{\xi_t^L(y)=1\}}$. When the dual processes do not collide, the results of the dual computations are independent. Using (13) and noting that the sum defining $S_L(x)$ has $L^{2(1-\gamma)}$ terms, we conclude (recall that $\gamma \le 1/4$)

$$\text{var}\,(S_L) \le L^{2(1-\gamma)}(1 + L^{2(1-\gamma)}CL^{-1}) \le CL^{4(1-\gamma)-1}$$

Using Chebyshev's inequality, we have

$$P(|S_L - ES_L| > L^\alpha) \le \frac{CL^{4(1-\gamma)-1}}{L^{2\alpha}}$$

The number of small square in a region that is $O(1)$ is $O(L^{2\gamma})$ so we want

$$L^{2\gamma}L^{4(1-\gamma)-1-2\alpha} \to 0$$

for some $\alpha < 2(1-\gamma)$. If we take $\alpha = 1.8$ then we need $2\gamma + 0.4(1-\gamma) < 1$ which holds if $\gamma \le 1/4$. Combining our calculations we have established (7) and the proof of Theorem 1 is complete. $\qquad\square$

The proof is almost the same in the nonhomogeneous case. In that case the rates λ_G and λ_H depend on x, but they are bounded so (12) and (13) hold. The computations in Part 1, depend on the sizes of the jumps which don't change. Finally, part 2, only uses (13) so it remains valid.

4   Analysis of the homogeneous IDE

Weinberger [29] studied the asymptotic behavior as $n \to \infty$ of discrete time iterations $u_{n+1}(x) = (Qu_n)(x)$ where Q acts on functions from $u(x) : \mathbb{R}^d \to [0,1]$. Our evolution occurs in continuous time, but our system can be put into his setting by letting $(Qu)(x)$ be the solution at time 1, starting from initial condition $u(x)$. To develop a theory of these iterations Weinberger introduced a number of assumptions. We will use his Theorem 6.2 which requires (3.1) and (3.2).

(3.1) Let B be the collection of continuous functions taking values in $[0,1]$
   (i) $Q[u] \in B$ for all $u \in B$
   (ii) If T_y is translation by y, $Q[T_y[u]] = T_y[Q[u]]$
   (iii) There are constants $0 \le \pi_0 < \pi_1 \le \pi_+$ so that if α is the conttant function $u \equiv \alpha$ then $Q[\alpha] > \alpha$ for $\alpha \in (\pi_0, \pi_1)$, $Q[\pi_0] = \pi_0$. $Q[\pi_1] = \pi_1$.
   (iv) $u \le v$ implies $Q[u] \le Q[v]$
   (v) $u_n \to u$ on each bounded subset implies that $Q[u_n](x) \to Q[u](x)$.

(3.2) $Q[\alpha] < \alpha$ when $\alpha \in [\pi_1, \pi_+]$

In our example (ii) of (3.1) and (3.2) hold if we take $\pi_0 = v_1$, $\pi_1 = v_2$, $\pi_+ = 1$. The other conditions are easily checked.

In the theory of reaction diffusion equations, traveling wave solutions play an important role. These are solutions to the PDE of the form

$$u(t,x) = w(x \cdot \xi - \rho t) \tag{17}$$

where ξ is a fixed direction, i.e., $\xi \in \mathbb{S}^{d-1}$, the $d-1$ dimensional sphere of unit vectors in $\mathbb{R}^d$. This is a plane wave, i.e., at any time t the value $u(t,x)$ is constant on hyperplanes perpendicular to the direction of movement ξ. Note that the shape of the wave in (17) does not change and it moves with a constant speed ρ.

In the case of a PDE w satisfies an ordinary differential equation, but in the case of a discrete iteration or an IDE it is generally not possible to prove the existence of traveling solutions. (For an exception to this rule see [22].) Weinberger shows in Section 5 of his paper that under assumption (3.1) it is possible to define a wave speed $\rho(\xi)$ for each direction. In our case the equation is rotationally invariant so all the speeds are the same.

Define θ_1 so that $\rho > 0$ when $A/B > \theta_1$ and $\rho < 0$ when $A/B < \theta_1$. Given $\theta > \theta_1$ let v_1 be the unstable interior fixed point. The next result follows from Theorem 6.2 in [29].

Theorem 6. *Suppose $A/B > \theta_1$ so that $\rho > 0$, and let $r < \rho$. If $\sigma_1 > v_1$ then there is an N_1 so that if $u_0 \ge \sigma_1$ on $x_0 + [-N_1, N_1]^2$ then*

$$u_n(nx) \to v_2 \qquad uniformly \ on \ |x| \le rn \tag{18}$$

Interchanging the roles of 0 and 1

Theorem 7. *Suppose $\theta_0 < A/B < \theta_1$ so that $\rho < 0$, and let $r < -\rho$. If $\sigma_0 < v_1$ then there is an N_0 so that if $u_0 \le \sigma_0$ on $x_0 + [-N_0, N_0]^2$ then*

$$u_n(nx) \to 0 \qquad uniformly \ on \ |x| \le rn \tag{19}$$

13

When $\theta < \theta_0$ there is no interior fixed point so when the initial condition is not $\equiv 1$ it converges to 0. Apply 6.5 in [29] to $1 - u$. These results are laws of large numbers for the solution, i.e., it describes what we see if we scale space by n and let $n \to \infty$. It does not give information about what is happening near the front. For reaction diffusion equations such as

$$\frac{du}{dt} = u''(x) - x + \beta u^2 (1 - u)$$

then shape of the solution near the front converges to the traveling wave solution, see e.g., [17]. However the results for the IDE are not strong enough to prove this.

5 Block construction

5.1 Homogeneous case

Suppose that $A/B > \theta_1$. Let v_1 and v_2 be the interior fixed points of the ODE (5), Pick $\delta > 0$ so that $v_2 - v_1 > 3\delta$. Let $\sigma = v_1 + \delta$ and pick $N \geq N_\sigma$, the constant in Theorem 6. Using that result and noting that the distance from $(0,0)$ to $(3N, 3N)$ is $(3\sqrt{2})N = 4.242N$ we see that

Lemma 2. *Suppose the initial configuration of the IDE $u(0, x) \geq v_1 + \delta$ on $[-N, N]^2$ and let $T = 5N/\rho$. Theorem 6 implies that if N is sufficiently large then $u(T, x) \geq v_2 - \delta$ on $[-3N, 3N]^2$.*

Note that this implies that for all $k \geq 1$,

$$u(kT, x) \geq v_2 - \delta \quad \text{for } x \in [-N(1 + 2k), N(1 + 2k)]^2.$$

As in the formulation of Theorem 1, the statement that particle system ξ_t has density $\geq u_0$ on a square $[-aN, aN]^2$ means that if $\gamma \in (0, 1/4)$ is fixed and we tile the plane with squares $[y, y + L^{-\gamma}) \times [z, z + L^{-\gamma})$ in such a way that the origin is the lower-left corner of one of the squares, then the density of 1's in ξ_t in each small square contained in $[-aN, aN]^2$ is $\geq u_0$. For each $(i, j, n) \in \mathbb{Z}^3$ with $i + j + n$ even, define squares

$$I_{i,j} = (2Ni, 2Nj) + [-N, N]^2 \quad \text{and} \quad I_{i,j}^3 = (2Ni, 2Nj) + [-3N, 3N]^2.$$

We say that (i, j, n) is **wet** if ξ_{nT} has density $\geq v_2 - 2\delta$ on $I_{i,j}$ at time nT.

To prove Theorem 2 we will show that there is an $m < \infty$ so that if $\epsilon > 0$ and L and M are large then the wet sites dominate m-dependent oriented percolation in which sites are open with probability $1 - \epsilon$. To establish m-dependence, we compare the dual with a branching random walk (BRW). To build the BRW let $U^1, U^2, \dots$ be independent and uniform on $D(0, 1)$ and declare that (i) at rate A a particle at x gives birth to particles at $x + U^1, \dots x + U^J$, (ii) at rate B a particle at x gives birth to particles at $x + U^1, \dots x + U^K$. To control the movement of the dual we use the following well-known result, see e.g., [3].

Lemma 3. *Let η_t^0 be the particles in the branching random walk at time t starting from a single particle at the origin. If κ is chosen large enough, there is a $\eta > 0$ so that*

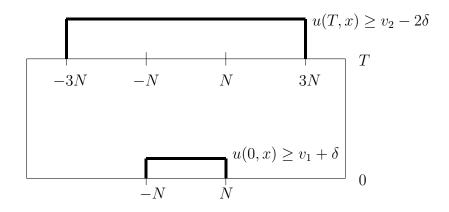$$P(\eta_s^0 \subset [-\kappa t, \kappa t]^2 \text{ for all } s \leq t) \geq 1 - \exp(-\eta t). \tag{20}$$

Figure 2: Picture of the block construction in the homogeneous case.

Recall $T = 5N/\rho$ and define the good event

$$G_0(i, j, n) = \{\zeta_s^{x,(n+1)T} \subset x + [-\kappa T, \kappa T]^2 \quad \text{for all } s \leq T \text{ and } x \in I_{i,j}^3\}.$$

Given $\epsilon > 0$ if N is large then G_0 holds with probability $> 1 - \epsilon/2$. If we pick m large enough so that

$$2m - 3 - \kappa \cdot \frac{5}{\rho} > 3 + \kappa \cdot \frac{5}{\rho} \tag{21}$$

then when $\|(i - i', j - j')\|_\infty \geq m$ the events $G_0(i, j, n)$ and $G_0(i', j', n)$ are independent. By construction good events on different levels are independent, so we have verified m-dependence of the G_0.

Part 2 of the proof of he hydrodynamic limit, Theorem 1, implies

Lemma 4. *Let $\epsilon > 0$ and $T = 5N/\rho$. If N is large then for any the initial configuration of the particle system with density $\geq v_1 + \delta$ in $[-N, N]^2$, ξ_T will have density $\geq v_2 - 2\delta$ on $[-3N, 3N]^2$ with with probability $\geq 1 - \epsilon/2$.*

To compare our process with oriented percolation on $\mathcal{L} = \{(i, n) : n \geq 0, i + n \text{ is even}\}$, we restrict our attention to the renormalized sites with $j = 0$. To define the $\eta(i, n)$ which are 1 of the sites is open, and 0 if they are closed, we use the following rules:

(i) If $(i, 0, n)$ is not wet we let $\eta(i, n)$ be a random variable independent of the process that is 1 with probability $p = 1 - \epsilon/2$, and 0 otherwise.

(ii) On $G_0(i, 0, n)^c$ we set $\eta(i, 0, n) = 0$.

(iii) If $G_0(i, 0, n)$ occurs and $(i, 0, n)$ is wet, we set $\eta(i, 0, n)$ if the event in Lemma 2 occurs. Note that on $G_0(i, 0, n)$ the events in Lemma 4, which can be determined from the dual CBRW are m-dependent.

1n 1989 Durrett and Schonmann [14] showed

Lemma 5. *If sites in supercritical one dimensional oriented percolation on $\mathcal{L} \cap (\{1, 2, \ldots \ell\} \times \mathbb{Z})$ are open with probability $1 - \epsilon$ and we start from a single site occupied, then the system will survive for time $\geq \exp(\eta(\epsilon)\ell)$ with a probability that tends to 1 as $\epsilon \to 0$.*

There are now results in higher dimensions and with sharp constants (see e.g., [20], and [21]). However, here we will compare with one dimensional percolation by embedding a copy of a long interval in our two-dimensional system. To do this, we will go back and forth across the rectangle, increasing the height by m each time so that sites on different horizontal segments are independent. To be precise, suppose we have a rectangle of renormalized sites $(x_0, y_0) + (2Ni, 2Nj)$ with $0 \leq i \leq i_0$ and $0 \leq j \leq j_0$ then the embedded curve is as drawn in Figure 3. An advantage of the construction is that it proves that the survival time of the contact process on the rectangle is $\geq \exp(ci_0 j_0)$, i.e., the exponent is proportional to the area of the rectangle.
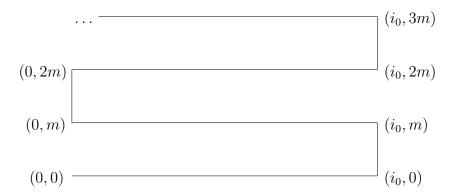

Figure 3: Comparison of the two-dimensional spatial process with one-dimensional oriented percolation.

5.2 Heterogeneous case

The outline of the proof of part (i) Theorem 4 is the same as the one just given for Theorem 2. However, in the absence of the translation invariance, we will need estimates that hold simultaneously for all blocks. Consider one component $\mathcal{R}$ of the open set $\{x : A(x)/B(x) > \theta_1 + \eta\}$. Eventually we will approximate $\mathcal{R}$ from inside by a finite union of rectangles. We begin by considering the case of one rectangle of renormalized sites $(x_0, y_0) + (2Ni, 2Nj)$ with $0 \leq i \leq i_0$ and $0 \leq j \leq j_0$ Assumption (H1) implies that in $\mathcal{R}$

$$0 < a_0 \leq A(x) \leq a_1 < \infty \quad \text{and} \quad 0 < b_0 \leq B(x) \leq b_1 < \infty. \tag{22}$$

Our first step is to show that the conclusion of Lemma 2 holds uniformly for a set of IDE with constant coefficients. By a change of time scale we can change the parameters from (A, B) to $(A/B, 1)$. Monotonicity implies that the worst case in $\mathcal{R}$ occurs for parameters $(\theta_1 + \eta, 1)$. Let v_1^* and v_2^* be the interior fixed points for the equation with parameters

$(\theta_1 + \eta/2, 1)$ and let ρ^* be the wave speed. Pick δ^* so that $v_2^* - v_1^* > 3\delta^*$. Noting that when $C' \geq C$ the solutions of the equation with parameters $(C', 1)$ lie above those with parameters $(C, 1)$ gives the following:

Lemma 6. *Suppose the initial configuration of the IDE has $u(0, x) \geq v_1^* + \delta$ on $[-N, N]^2$ and let $T = 5N/b_0\rho^*$. If N is sufficiently large then for all parameters (A, B) with $A/B \geq \theta_1 + \eta/2$, $u(T, x) \geq v_2^* - \delta^*$ on $[-3N, 3N]^2$.*

Note that v_1^* and v_2^* are the interior fixed points for the parameters $(\theta_1 + \eta/2, 1)$. The result holds for all larger A/B since v_1 is a decreasing function of A/B and v_2 is an increasing function of A/B. The choice of T is dictated by the fact that the speed is slowest for $(\theta_1 + \eta/2, 1)$, and we have to change time to return to (A, B).

Using (22), the proof of Lemma 3 can be repeated to show

Lemma 7. *Let η_t^x be the particles in the branching random walk at time t starting from a single particle at the x. If κ^* is chosen large enough, there is a $\eta^* > 0$ so that for all starting points $x \in \mathcal{R}$*

$$P(\eta_s^x \subset [-\kappa^* t, \kappa^* t]^2 \text{ for all } s \leq t) \geq 1 - \exp(-\eta^* t). \tag{23}$$

Define $G_0(i, j, n)$ as before but will κ^* instead of κ. As in the previous proof if we take $T = 5b_1 N/\rho^*$ then we will with high probability have this for all starting points for the dual $x \in I_{ij}^3$. This means that the good event $G_0(i, j, n)$ has high probability and is m-dependent (for a possibly larger value of m).

Using the calculation in part 2 of the hydrodynamic limit we have

Lemma 8. *Let $\epsilon > 0$ and let $T = 5N/b_0\rho^*$. If N is large then for any the initial configuration of the particle system with density $\geq v_1^* + \delta$ in $I_{i,j}$, ξ_T will have density $\geq v_2^* - 2\delta$ on $I_{i,j}^3$ with with probability $\geq 1 - \epsilon/2$.*

Combining Lemmas 7 and 8 with Lemma 5 we have prolonged survival on the rectangle. To get the conclusion of part (i) of Theorem 4 we note that the closure of the component $\mathcal{R} \{x : A(x)/B(x) > \theta + \delta\}$ can be covered by a finite number of overlapping rectangles of renormalizes sites that lie inside $\{x : A(x)/B(x) > \theta + \delta/2\}$. Our assumption implies that at least one renormalized site is wet at time 0. Using Lemma 5 the wet sites will spread to each of the rectangles and persist for an amount of time that is $\geq \exp(\eta M^2)$.

References

[1] Aronson, D. G. and Weinberger, H. F. (1978) Multidimensional nonlinear diffusion arising in population genetics. *Adv. in Math.* 30 (1978), 33-76

[2] Bessonov, M. and Durrett, R. (2017) Phase transitions for a planar quadratic contact process. *Adv. in Appl. Math.* 87, 82-107 [MBRD]

[3] Biggins, J.D.(1977) Chernoff's theorem in the branching random walk. *J. Appl. Prob.* 14, 630–636

[4] Bramson, M. and Durrett, R. (1988) A simple proof of the stability criterion of Gray and Griffeath. *Probab. Theory Related Fields.* 80, 293-298

[5] Cox, J.T., Durrett, R., and Perkins, E. (2013) *Voter model perturbations and reaction diffusion equations.* Asterique. Volume 349, (113 pages). Also available at arXiv:1103.1676

[6] Durrett, R. (1984) Oriented percolation in two dimensions. *Ann. Probab.* 12, 999–1040

[7] Durrett, R. (1992) Predator-prey systems. Pages 37-58 in *Asymptotic problems in probability theory: stochastic models and diffusions on fractals.* Edited by K.D. Elworthy and N. Ikeda. Pitman Research Notes 283. Longman, Essex, England

[8] Durrett, R. (1995) Ten Lectures on Particle Systems. Pages 97-201 in *St. Flour Lecture Notes.* Lecture Notes in Math 1608. Springer-Verlag, New York

[9] Durrett, R. (2008) Coexistence in Stochastic Spatial Models. (Wald Lecture Paper). *Ann. Appl. Prob.* 19, 477–496

[10] Durrett, R. (2010) *Probability: Theory and Examples.* 4th edition, Cambridge U. Press

[11] Durrett, R., and Neuhauser, C. (1994) Particle systems and reaction diffusion equations. *Ann. Prob.* 22, 289–333 [RDCN]

[12] Durrett, R., and Levin, S. (1994) The importance of being discrete (and spatial). *Theoret. Pop. Biol.* 46, 363–394

[13] Durrett, R., and Levin, L. (1998) Spatial aspects of interspecific competition. *Theoret. Pop. Biol.* 53, 30–43

[14] Durrett, R., and Schonmann, R.H. (1988) The contact process on a finite set, II. *Ann. Probab.*, 1570-1583

[15] Durrett, R., and Zähle, I. (2007) On the width of hybrid zones. *Stoch. Proc. Appl.* 117 (2007), 1751-1763

[16] Durrett, R., and Zhang, Y. (2015) Coexistence of Grass, Saplings and Trees in the Staver-Levin Forest Model. *Annals of Applied Probability.* 25, 104–115 [RDYZ]

[17] Fife, P.C. and McLeod, J.B. (1977) The approach of solutions of nonlinear diffusion equations to traveling frontsolutions. *Arch. Rat. mech. Anal.* 65, 335–361

[18] Krone, S. M. (1999). The two-stage contact process. *Ann. Appl. Probab.* 9, 331–351.

[19] Liggett, T. M. (1985) *Interacting particle systems.* Springer-Verlag, New York

[20] Mountford, T.S. (1993) A metastable result for the finite multidimensional contact process. *Canadian Math Bulletin.* 36, 216–222

[21] Mountford, T.S. (1999) Existence of a constant for finite system extinction. *J. Statistical Physics.* 96, 1331–1341

[22] Neuhauser, C. (1994) A long range sexual reproduction process. *Stochastic Process. Appl.* 53, 193220 [CNsex]

[23] Schertzer, E., Staver, A. C. and Levin S. (2015). Implications of the spatial dynamics of fire spread for the bistability of savanna and forest. *J Math Biol.* 70, 329–341

[24] Staver, A. C., Archibald, S., and Levin S. (2011). The global extent and determinants of savanna and forest as alternative biome states. *Science* 334, 230–232

[25] Staver, A. C., Archibald, S. and Levin S. (2011). Tree cover in sub-Saharan Africa: Rainfall and fire constrain forest and savanna as alternative stable states. *Ecology* 92, 1063–1072

[26] Staver, A. C. and Levin S. (2012). Integrating theoretical climate and fire effects on savanna and forest systems. *Amer. Nat.* 180, 211–224

[27] Swindle, G. (1990). A mean field limit of the contact process with large range. *Probab. Th. Rel. Fields* 85, 261–282 [Swindle]

[28] Touboul, J.D., Staver, A.C. and Levin, S.A. (2018) On the complex dynamics of savanna landscapes. *Proc. Natl. Acad. Sci.* 115 (7), E1336–E1345

[29] Weinberger, H.F. (1982) Long time behavior of a class of biological models. *SIAM J. Math. Anal.* 13, 353–396