COEXISTENCE OF GRASS, SAPLINGS AND TREES IN THE STAVER-LEVIN FOREST MODEL

By Rick Durrett* and Yuan Zhang*

Duke Univeristy

In this paper we prove the existence of a nontrivial stationary distribution for a forest model with Grass, Saplings and Trees. By comparing with the two type contact process model of Krone and considering the long range limit, we first shows that if a particle systems has states $\{0, 1, 2\}$ and is attractive, then coexistence occurs in the long-range model whenever the absorbing state (0,0) is an unstable fixed point of the mean-field ODE for (u_1, u_2). The result we obtain in this way is asymptotically sharp for Krone's model. In the Staver-Levin forest model, if $(0, 0)$ is attracting there may also be another stable fixed point for the ODE. We show that, as in the case of the quadratic contact process, in some of these cases there may still be a nontrivial stationary distribution.

1. Introduction. In a recent paper published in Science [17], 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 [18], the American Naturalist [19] and Journal of Mathematical Biology [16], they studied the following ODE for the evolution of the fraction of land covered by grass G, saplings S, and trees T:

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

Here $\omega(G)$ is a decreasing function of G which reflects death of saplings due to the spread of fires and $\mu \geq \nu$ are the death rates for saplings and trees. Studies suggest (see [19] for references) that regions with tree cover below

1

about 40% burn frequently but fire is rare above this threshold, so they used an ω that is close to a step function.

The ODE in (1) has very interesting behavior: it may have two stable fixed points, changing the values of parameters may lead to Hopf bifurcations, and if the system has an extra type of savanna trees, there can be period orbits. In this paper, we will begin the study of the corresponding spatial model. The state at time t is $\chi_t : \mathbb{Z}^d \to \{0, 1, 2\}$, where $0 =$ Grass, $1 =$ Sapling, and $2 =$ Tree. Given the application, it would be natural to restrict our attention to $d = 2$, but since the techniques we develop will be applicable to other systems we consider the general case.

In the forest model, it is natural to assume that dispersal of seeds is long range. To simplify our calculations, we will not use a continuous dispersal distribution for tree seeds, but instead let $f_i(x, L)$ denote the fraction of sites of type i in the box $x + [-L, L]^d$ and declare that site x changes

- $0 \to 1$ at rate $\beta f_2(x, L)$
- $1 \to 2$ at rate $\omega(f_0(x, \kappa L))$
- $1 \to 0$ at rate μ
- $2 \to 0$ at rate ν

The configuration with all sites 0 is an absorbing state. This naturally raises the question of finding conditions that guarantee the existence of a nontrivial stationary distribution. Our model has three states but it is "attractive," i.e., if $\chi_0(x) \le \chi'_0(x)$ for all x then we can construct the process so that this inequality holds for all time. From this, it follows from the usual argument that if we start from $\chi_0^2(x) \equiv 2$ then χ_t^2 converges to a limit χ_∞^2 that is a translation invariant stationary distribution, and there will be a nontrivial stationary distribution if and only if $P(\chi_\infty^2(0) = 0) < 1$. Since 2's give birth to 1's and 1's grow into 2's, if χ_∞^2 is nontrivial then both species will be present with positive density in χ_∞^2.

If $\omega \equiv \gamma$ is constant, $\nu = 1 + \delta$, and $\mu = 1$ then our system reduces to one studied by Krone [12]. In his model, 1's are juveniles who are not yet able to reproduce. Krone proved the existence of nontrivial stationary distributions in his model by using a simple comparison between the sites in state 2 and a discrete time finite-dependent oriented percolation. In the percolation process we have an edge from $(x, n) \to (x + 1, n + 1)$ if a 2 at x at time $n\epsilon$ will give birth to a 1 at $x + 1$, which then grows to a 2 before time $(n + 1)\epsilon$, and there are no deaths at x or $x + 1$ in $[n\epsilon, (n + 1)\epsilon]$. As the reader can imagine, this argument produces a very crude upper bound on the critical value of the system.

A simple comparison shows that if we replace the decreasing function

$\omega(G)$ in the Staver-Levin model, χ_t, by the constant $\omega = \omega(1)$, to obtain a special case η_t of Krone's model, then χ_t dominates η_t in the sense that given $\chi_0 \geq \eta_0$ the two processes can be coupled so that $\chi_t \geq \eta_t$ for all t. Because of this, we can prove existence of nontrivial stationary distribution in the Staver-Levin model by studying Krone's model. To do this under the assumption of long range interactions, we begin with the mean field ODE:

$$(2) \qquad \begin{aligned} \frac{dG}{dt} &= \mu S + \nu T - \beta GT \\ \frac{dS}{dt} &= \beta GT - (\omega + \mu)S \\ \frac{dT}{dt} &= \omega S - \nu T \end{aligned}$$

Since $G + S + T = 1$, we can set $G = 1 - S - T$ and reduce the system to two equations for S and T.

To guess a sufficient condition for coexistence in the long range limit we note that:

LEMMA 1.1. *For the mean-field ODE (2), $S = 0, T = 0$ is not a attracting fixed point if and only if*

$$(3) \qquad \mu\nu < \omega(\beta - \nu)$$

PROOF. When $(S, T) \approx (0, 0)$ and hence $G \approx 1$, the mean-field ODE is approximately:

$$\begin{pmatrix} dS/dt \\ dT/dt \end{pmatrix} \approx A \begin{pmatrix} S \\ T \end{pmatrix} \quad \text{where} \quad A = \begin{pmatrix} -[\omega + \mu] & \beta \\ \omega & -\nu \end{pmatrix}$$

The trace of A, which is the sum of its eigenvalues is negative, so $(0, 0)$ is not attracting if and only if the determinant of A, which is the product of the eigenvalues is negative. Since $(\omega + \mu)\nu - \beta\omega < 0$ if and only if $\mu\nu < (\beta - \nu)\omega$, we have proved the desired result. □

THEOREM 1. *Let η_t be Krone's model with parameters that satisfy (3) Then η_t has a nontrivial stationary distribution when L is large enough.*

On the other hand, in Krone's model we have the following converse, which does not require the assumption of long range.

THEOREM 2. *Suppose (3) does not hold. Then for any $L > 0$, Krone's model η_t dies out, i.e. for any initial configuration η_0 with finitely many non-zero sites:*

$$\lim_{t \to \infty} P(\eta_t(x) \equiv 0) = 1$$

This implies that the Staver-Levin model dies out when (3) fails in for $\omega = \omega(0)$.

1.1. *Survival when zero is stable.* Theorem 1 implies survival of the Staver-Levin model when the trivial fixed point $(0,0)$ is unstable, i.e.:

$$(4) \qquad \mu\nu < \omega(1)(\beta - \nu).$$

When (4) does not hold, the trivial fixed point of the mean-field ODE will be attracting. However, since the mean-field ODE (2) may still have another stable fixed point in the positive density region (and also a unstable fixed point in between), the Staver-Levin model, like the quadratic contact process studied by [7], [15] and [1] may have a nontrivial stationary distribution when $(0,0)$ is attracting.

Based on the observation in [18] mentioned above, it is natural to assume that $\omega(\cdot)$ is a step function. In our proof we let :

$$(5) \qquad \omega(G) = \begin{cases} \omega_0, \ G \in [0, 1 - \delta_0) \\ \omega_1, \ G \in [1 - \delta_0, 1] \end{cases}$$

where $\omega_0 > \omega_1$ and $\delta_0 \in (0,1)$. However, according to the monotonicity of χ_t, our result about the existence of a nontrivial stationary distribution can also hold if one replaces the "=" in (5) by "$\geq$" since the new process dominates the old one.

To prove the existence of a nontrivial stationary distribution under the assumption of long range, a natural approach would be showing that when $L \to \infty$ and space is rescaled by diving by L, the Staver-Levin model converges weakly to the solution of following integro-differential equation:

$$\frac{dS(x,t)}{dt} = \beta D_1^T(x,t)G - \mu S - \omega[D_\kappa^G(x,t)]S$$

$$(6) \qquad \frac{dT(x,t)}{dt} = \omega[D_\kappa^G(x,t)]S - \nu T$$

where $G = 1 - S - T$ and

$$D_1^T(x,t) = \frac{\int_{x+[-1,1]^d} T(y,t)\,dy_1 \cdots dy_d}{2^d}$$

$$D_\kappa^G(x,t) = \frac{\int_{x+[-\kappa,\kappa]^d} G(y,t)\,dy_1 \cdots dy_d}{(2\kappa)^d},$$

are the local densities of trees and grass on the rescaled lattice. The first problem with this approach is that since the density is computed by examining all sites in a square, there is not a good dual process, which was the

key to proofs in [20], [7], [15] and later in [1]. The second problem is that one does not know much about the limiting IDE. Results of Weinberger [21] show the existence of wave speeds and provide a convergence theorem in the case of a single equation, but we do not know of results for a pair of equations.

To avoid these difficulties, we will construct test functions S_{test} and T_{test}, so that under (6), the derivatives will always be positive for all x in $\overline{\{T_{test} > 0\}}$ and $\overline{\{S_{test} > 0\}}$, where $\bar{A}$ stands for the closure of set A. Figure 1 shows the tests functions when $d = 1$. The positive derivative implies that after a positive time the solution will dominate translates of the initial condition by positive and negative amounts. Monotonicity then implies that the solution will expand linearly and the result follows from a block construction. Details can be found in Section 7.

THEOREM 3. *Under condition*

$$\beta\omega_0 > 2^d\nu(\mu + \omega_0), \tag{7}$$

there is a constant $S_0(\beta, \omega_0, \nu, \mu)$ so that if

$$\delta_0 \in (0, 2^{-d}S_0)$$

the Staver-Levin forest model χ_t survives when L is large for any $\omega_1 \geq 0$ and $\kappa > 0$.

Combining Theorems 1-3 we have the following results for the Staver-Levin model

(a) When $\mu\nu < \omega(1)(\beta - \nu)$, χ_t has non-trivial stationary distribution when L is large.
(b) When $\mu\nu \geq \omega(0)(\beta - \nu)$, χ_t dies out for all $L \geq 1$.
(c) When $\mu\nu \geq \omega(1)(\beta - \nu)$, under the hypotheses of Theorem 3, χ_t can still have non-trivial stationary distribution when L is large, no matter how small is ω_1.
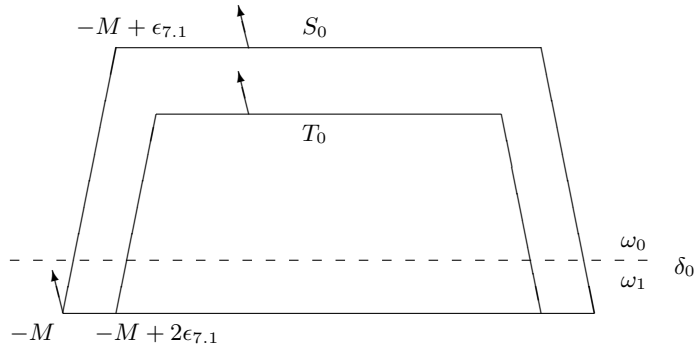
Fig 1: Test functions for $d = 1$

1.2. *Sketch of the Proof of Theorem 1.* Most of the remainder of the paper is devoted to the proof of Theorem 1. We will now describe the main ideas and then explain where the details can be found.

(i) The key idea is due to Grannan and Swindle [10]. They consider a model of a catalytic surface in which atoms of type $i = 1, 2$ land at vacant sites (0's) at rate p_i, while adjacent 1, 2 pairs turn into 0,0 at rate ∞. If after a landing event, several 1,2 pairs are created, one is chosen at random to be removed. The first type of event is the absorption of an atom onto the surface of the catalyst, while the second is a chemical reaction, e.g., carbon monoxide CO and oxygen O reacting to produce CO_2. The last reaction occurs in the catalytic converted in your car, but the appropriate model for that system is more complicated. An oxygen molecule O_2 lands and dissociates to two O bound to the surface when a pair of adjacent sites is vacant. See Durrett and Swindle [9] for more details about the phase transition in the system.

Suppose without loss of generality that $p_1 + p_2 = 1$. In this case Grannan and Swindle [10] showed that if $p_1 \neq p_2$ the only possible stationary distributions concentrate configurations that are $\equiv 1$ or $\equiv 2$. Mountford and Sudbury [14] later improved this result by showing that if $p_1 > 1/2$ and the initial configuration has infinitely many 1's then the system converges to the all 1's state.

The key to the Grannan-Swindle argument was to consider

$$Q(\eta_t) = \sum_x e^{-\lambda \|x\|} q[\eta_t(x)].$$

where $\|x\| = \sup_i |x_i|$ is the L^∞ norm, $q(0) = 0$, $q(1) = 1$, and $q(2) = -1$. If λ is small enough then $dEQ/dt \geq 0$ so Q is a bounded submartingale and hence converges almost surely to a limit. Since an absorption or chemical

reaction in $[-K, K]^d$ changes Q by an amount $\geq \delta_K$, it follows that such events eventually do not occur.

(ii) Recovery from small density is the next step. We will pick $\epsilon_0 > 0$ small, let $\ell = [\epsilon_0 L]$ be the integer part of $\epsilon_0 L$ and divide space into *small boxes* $\hat{B}_x = 2\ell x + (-\ell, \ell]^d$. To make the number of 1's and 2's in the various small boxes sufficient to describe the state of the process, we declare two small boxes to be neighbors if all of their points are within an L^∞ distance L. For the "truncated process", which is stochastically bounded by η_t, and in which births of trees can only occur between sites in neighboring small boxes, we will show that if $\kappa \in (d/2, d)$ and we start with a configuration that has L^κ non-zero sites in $\hat{B}_0$ and 0 elsewhere, then the system will recover and produce a small box $\hat{B}_x$ at time τ in which the density of nonzero sites is $a_0 > 0$ and $P(\tau > t_0 \log L) < L^{d/2-\kappa}$. See Lemma 3.1. To prove this, we use an analogue of Grannan and Swindle's Q. The fact that $(0,0)$ is an unstable fixed point implies $dEQ/dt > 0$ as long as the density in all small boxes is $\leq a_0$.

(iii) Bounding the location of the positive density box is the next step. To do this we use a comparison with branching random walk to show that the small box $\hat{B}_x$ with density a_0 constructed in Step 2 is not too far from 0, Random walk estimates will later be used to control how far it will wander as we iterate the construction. For this step it is important that the truncated process is invariant under reflection, so the mean displacement is 0. If we try to work directly with the original interacting particle system η_t then it is hard to show that the increments between box locations are independent and have mean 0.

(iv) Moving particles. The final ingredient in the block construction is to show that given a small block $\hat{B}_x$ with positive density and any y with $\|y - x\|_1 \leq [c \log L]$ then if c is small enough it is very likely that there will be $\geq L^\kappa$ particles in $\hat{B}_y$ at time $[c \log L]$. Choosing y appropriately and then using the recovery lemma we can get lower bounds on the spread of the process.

(v) Block construction. Once we have completed steps (ii), (iii), and (iv), it is straightforward to show that our system dominates a one-dependent oriented percolation and prove the existence of a nontrivial stationary distribution.

The truncated process is defined in Section 2 and a graphical representation is used to couple it, Krone's model, and the Staver-Levin model. In

Section 3 we use the Grannan-Swindle argument to prove the recovery from small density and the dying out result in Thoerem 2. In Section 4, we bound the movement of the positive density boxes produced by our construction. In Section 5, we show how to move particles. Finally in Section 6 we put all of the pieces together and complete the proof.

2. Box Process and Graphical Representation. For some fixed $\epsilon_0 > 0$ which will be specified in (15), let $l = [\epsilon_0 L]$ and divide space Z^d into small boxes:

$$\hat{B}_x = 2lx + (-l, l]^d \qquad x \in Z^d$$

For any $x \in Z^d$ there is a unique x' such that $x \in \hat{B}_{x'}$. Define the new neighborhood of interaction as follows: For any $y \in \hat{B}_{y'}$, $y \in \mathcal{N}(x)$ if and only if

$$\sup_{z_1 \in \hat{B}_{x'}, z_2 \in \hat{B}_{y'}} \|z_1 - z_2\| \le L.$$

It is easy to see that $\mathcal{N}(x) \subset B_x(L)$. To show that

(8) $$\mathcal{N}(x) \supset B_x((1 - 4\epsilon_0)L)$$

we note that if $\|x - y\| \le (1 - 4\epsilon_0)L$, $z_1 \in \hat{B}_{x'}$, and $z_2 \in \hat{B}_{y'}$, where $\hat{B}_{x'}$ and $\hat{B}_{y'}$ are the small boxes containing x and y:

$$\|z_1 - z_2\| \le \|z_1 - x\| + \|x - y\| + \|y - z_2\| \le 4([\epsilon_0 L]) + (1 - 4\epsilon_0)L \le L.$$

Given the new neighborhood $\mathcal{N}(x)$, we define the truncated version of Krone's model ξ_t by its transition rates:

transtion	at rate
$1 \to 0$	μ
$2 \to 0$	ν
$1 \to 2$	ω
$0 \to 1$	$\beta N_2[\mathcal{N}(x)]/(2L+1)^d$

where $N_i[S]$ stands for the number of i's in the set S.

For any $x \in Z^d$ and $\xi \in \{0, 1, 2\}^{Z^d}$, define $n_i(x, \xi)$ to be the number of type i's in the small box $\hat{B}_x$ in the configuration ξ. The box process

$$\zeta_t(x) = [n_1(x, \xi_t), n_2(x, \xi_t)], \forall x \in Z^d.$$

Then ζ_t is a Markov process on $\{(n_1, n_2) : n_1, n_2 \ge 0, n_1 + n_2 \le |\hat{B}_0|\}^{Z^d}$ in which:

transition	at rate
$\zeta_t(x) \to \zeta_t(x) - (1,0)$	$\mu\zeta_t^1(x)$
$\zeta_t(x) \to \zeta_t(x) - (0,1)$	$\nu\zeta_t^2(x)$
$\zeta_t(x) \to \zeta_t(x) + (-1,1)$	$\omega\zeta_t^1(x)$
$\zeta_t(x) \to \zeta_t(x) + (1,0)$	$\sum_{y:\hat{B}_y \subset \mathcal{N}(x)} \beta\zeta_t^2(y)$

Because ζ_t only records the number of particles in any small box, and the neighborhood is defined so that all sites in the same small box have the same neighbors, the distribution of ζ_t is symmetric under reflection in any axis. The main use for this observation is that the displacement of the location of the positive density box produced by the recovery lemma in Section 3 has mean 0.

2.1. *Graphical Representation.* We will use the graphical representation similar as in [12] to construct Krone's model η_t and the truncated version ξ_t on the same probability space, so that

(∗) If $\eta_0 \geq \xi_0$ then we will have $\eta_t \geq \xi_t$ for all t.

We use independent families of Poisson processes for each $x \in Z^d$, as follows:

- $\{V_n^x : n \geq 1\}$ with rate ν. We put an $\times$ at space-time point (x, V_n^x) and write a δ_{12} next to it to indicate a death will occur if x is occupied by a 1 or a 2.
- $\{U_n^x : n \geq 1\}$ with rate $\mu - \nu$. We put an $\times$ at space-time point (x, V_n^x) and write a δ_1 next to it to indicate a death will occur if x is occupied by 1.
- $\{W_n^x : n \geq 1\}$ with rate ω. We put an $\bullet$ at space-time point (x, V_n^x) and write a β next to it to indicate that if x is in state 1, it will become a 2.
- $\{T_n^{x,y} : n \geq 1\}$ with rate $\beta/|B_0|$ all $y \in \mathcal{N}(x)$. We draw a solid arrow from $(x, T_n^{x,y})$ to $(y, T_n^{x,y})$ to indicate that if x is occupied by a 2 and y is vacant, then a birth will occur at x in either process.
- $\{T_n^{x,y} : n \geq 1\}$ with rate $\beta/|B_0|$ all $y \in B_x(L) - \mathcal{N}(x)$. We draw a dashed arrow from $(x, T_n^{x,y})$ to $(y, T_n^{x,y})$ to indicate that if x is occupied by a 2 and y is vacant then a birth will occur at x in the process ξ_t.

Standard arguments that go back to Harris [11] over forty years ago, guarantee that we have constructed the desired processes and (∗) holds. We can construct the Staver-Levin model on the same space so that it has $\chi_t \geq \xi_t$ by adding another family of Poisson process $\{\hat{W}_n^x : n \geq 1\}$ with rate $1 - \omega$, and independent random variables $w_{x,n}$ uniform on $(0,1)$. At any time

W_n^x if x is in state 1, it will increase to state 2 if

$$w_{n,x} > \frac{\omega(f_0(x, \kappa L)) - \omega}{1 - \omega}$$

3. Recovery Lemma. Given (3), one can pick a θ, which must be > 1, such that

(9)
$$\frac{\mu + \omega}{\omega} < \theta < \frac{\beta}{\nu}$$

so we have

$$\theta\omega - (\omega + \mu) > 0, \beta - \theta\nu > 0$$

and since the inequalities above are strict, we can pick some $a_0 > 0$ and $\rho > 0$ such that

(10)
$$\theta\omega - (\omega + \mu) \geq \rho, \beta(1 - 4a_0) - \theta\nu \geq \theta\rho.$$

Now we can let the undetermined ϵ_0 in the definition of ξ_t in Section 2 be a positive constant such that $(1 - 4\epsilon_0)^d > 1 - 2a_0$.

We start with an initial configuration in Ξ_0, the ξ_0 that have $\xi_0(x) = 0$ for all $x \notin \hat{B}_0$ and the number of nonzero sites in $\hat{B}_0$ is at least L^α for some $\alpha \in (d/2, d)$. We define a stopping time τ:

(11)
$$\tau = \inf\{t : \exists x \in Z^d \ s.t. n_1(x, \xi_t) + n_2(x, \xi_t) \geq a_0|\hat{B}_0|\}$$

LEMMA 3.1. **Recovery Lemma.** *Suppose we start the truncated version of Krone's model from a $\xi_0 \in \Xi_0$. Let $t_0 = 2d/\rho$. When L is large,*

(12)
$$P(\tau > t_0 \log L) < L^{d/2 - \alpha}$$

PROOF. As mentioned in the introduction, we consider

$$Q(\xi_t) = \lambda^d \sum_{x \in Z^d} e^{-\lambda\|x\|} w[\xi_t(x)]$$

where $\lambda = L^{-1}a_0/2$ and

$$w[\xi(x)] = \begin{cases} 0 & \text{if } \xi(x) = 0 \\ 1 & \text{if } \xi(x) = 1 \\ \theta & \text{if } \xi(x) = 2 \end{cases}$$

If we imagine $\mathbb{R}^d$ divided into cubes with centers at $\lambda \mathbb{Z}^d$ and think about sums approximating an integral then we see that

$$(13) \qquad \lambda^d \sum_{x \in Z^d} e^{-\lambda \|x\|} \leq e^{\lambda/2} \int_{\mathbb{R}^d} e^{-\|z\|} \, dz \leq U_0$$

for all $\lambda \in (0, 1]$. From this it follows that $Q(\xi_t) \leq \theta U_0$. Here and in what follows U's are upper bounds that are independent of $\lambda \in (0, 1]$.

Our next step towards Lemma 3.1 is to study the infinitesimal mean

$$\mu(\xi) = \lim_{\delta t \downarrow 0} \frac{E[Q(\xi_{t+\delta t}) - Q(\xi_t)|\xi_t = \xi]}{\delta t}.$$

LEMMA 3.2. *For all ξ such that $n_1(x,\xi)+n_2(x,\xi) \leq a_0|\hat{B}_0|$ for all $x \in Z^d$ and any $t \geq 0$, $\mu(\xi) \geq \rho Q(\xi)$ where ρ is defined in (10).*

PROOF. Straightforward calculation gives

$$(14)$$
$$\frac{\mu(\xi)}{\lambda^d} = \sum_{\xi(x)=1} [(\theta-1)\omega - \mu]e^{-\lambda\|x\|} + \sum_{\xi(x)=0} \beta \frac{N_2[\mathcal{N}(x)]}{(2L+1)^d} e^{-\lambda\|x\|} - \sum_{\xi(x)=2} \theta\nu e^{-\lambda\|x\|}.$$

For the second term in the equation above, we interchange the roles of x and y then rearrange the sum:

$$\sum_{\xi(x)=0} \beta \frac{N_2[\mathcal{N}(x)]}{(2L+1)^d} e^{-\lambda\|x\|} = \sum_{\xi(x)=2} (2L+1)^{-d} \sum_{y \in \mathcal{N}(x), \xi(y)=0} \beta e^{-\lambda\|y\|}.$$

Noting that $\lambda = L^{-1}a_0/2$, and that for any x and $y \in \mathcal{N}(x) \subset B_x$, $-L \leq \|y\| - \|x\| \leq L$, we have

$$e^{-\lambda\|y\|} \geq e^{-a_0/2}e^{-\lambda\|x\|} \geq (1-a_0)e^{-\lambda\|x\|}.$$

Using this with $n_1(x,\xi) + n_2(x,\xi) < a_0|\hat{B}_0|$, and $B_x[(1-4\epsilon_0)L] \subset \mathcal{N}(x)$ from (8),

$$\sum_{\xi(x)=0} \beta \frac{N_2[\mathcal{N}(x)]}{(2L+1)^d} e^{-\lambda\|x\|} \geq (1-a_0) \sum_{\xi(x)=2} \beta \frac{N_0(B_x[(1-4\epsilon_0)L])}{(2L+1)^d} e^{-\lambda\|x\|}$$

$$\geq (1-a_0)[(1-4\epsilon_0)^d - a_0] \sum_{\xi(x)=2} \beta e^{-\lambda\|x\|}.$$

So we can now specify ϵ_0 be small enough so that

$$(15) \qquad\qquad\qquad (1-4\epsilon_0)^d > 1 - 2a_0.$$

This choice implies

$$\sum_{\xi(x)=0} \beta \frac{N_2[\mathcal{N}(x)]}{(2L+1)^d} e^{-\lambda\|x\|} > (1-a_0)(1-3a_0) \sum_{\xi(x)=2} \beta e^{-\lambda\|x\|} > (1-4a_0) \sum_{\xi(x)=2} \beta e^{-\lambda\|x\|}.$$

Combining inequality above with (14) and (10) gives

$$\mu(\eta) \geq \lambda^d \sum_{\eta(x)=1} [(\theta-1)\omega-\mu]e^{-\lambda\|x\|} + \lambda^d \sum_{\eta(x)=2} [(1-4a_0)\beta-\theta\nu]e^{-\lambda\|x\|} \geq \rho Q(\eta).$$

which proves the desired result.   □

Let $\sigma^2(\xi)$ be the infinitesimal variance of $Q(\xi_t)$:

$$\sigma^2(\xi) = \lim_{\delta t \downarrow 0} \frac{E([Q(\xi_{t+\delta t}) - Q(\xi_t)]^2 | Q(\xi_t) = \xi)}{\delta t}.$$

LEMMA 3.3.   *There is a constant U_1 independent of $\lambda \in (0,1]$, so that* $\sigma^2(\xi) \leq U_1 L^{-d}$.

PROOF. By considering the rates at which jumps occur and the changes they cause we see that
(16)
$$\frac{\sigma^2(\xi)}{\lambda^{2d}} = \sum_{\xi(x)=1} [(\theta-1)^2\omega+\mu]e^{-\lambda\|x\|} + \sum_{\xi(x)=0} \beta \frac{N_2[\mathcal{N}(x)]}{(2L+1)^d} e^{-\lambda\|x\|} + 4 \sum_{\xi(x)=2} \theta\nu e^{-\lambda\|x\|}.$$

Using (13) and recalling $\lambda = L^{-1} a_0/2$ now gives the desired result.   □

For any initial configuration ξ_0, define

$$(17) \qquad\qquad M_t = Q(\xi_t) - Q(\xi_0) - \int_0^t \mu(\xi_s)ds$$

According to Dynkin's formula, M_t is a martingale with $EM_t = 0$.

LEMMA 3.4. $EM_t^2 = E \int_0^t \sigma^2(\xi_s)\,ds \leq U_1 L^{-d}$ *and hence*

$$(18) \qquad\qquad E\left(\sup_{s \leq t} M_s^2\right) \leq 8U_1 L^{-d}t.$$

PROOF. Using (14) and (13) we see that

$$(19) \qquad\qquad |\mu(\xi_t)| \leq U_2 = (\omega + \beta + \nu + \mu)U_0.$$

To calculate EM_t^2, let $t_i^n = it/n$.

(20)
$$EM_t^2 = \sum_{i=0}^{n-1} E(M_{t_{i+1}^n} - M_{t_i^n})^2 = \sum_{i=0}^{n-1} E\left[Q(\xi_{t_{i+1}^n}) - Q(\xi_{t_i^n}) - \int_{t_i^n}^{t_{i+1}^n} \mu(\xi_s)ds\right]^2.$$

Using (19) we have

$$\sum_{i=0}^{n} \left[\int_{t_i^n}^{t_{i+1}^n} \mu(\xi_s)ds\right]^2 \leq U_2 \sum_{i=1}^{n}(t_{i+1}^n - t_i^n)^2 = U_2 \frac{t^2}{n} \to 0.$$

From the definition of the infinitesimal variance, and the fact that the semi-group associated with a Markov process is a contraction operator, we get

$$0 \leftarrow \left|(n/t)E[Q(\xi_{t/n}) - Q(\xi_0)]^2 - \sigma^2(\xi_0)\right|$$
$$\geq E\left|(n/t)E[Q(\xi(t_{i+1}^n)) - Q(\xi(t_i^n)]^2 - \sigma^2(\xi(t_i^n))\right|$$

which implies that

$$\sum_{i=0}^{n-1} E[Q(\xi_{t_{i+1}^n}) - Q(\xi_{t_i^n})]^2 \to E\int_0^t \sigma^2(\xi_s)\,ds.$$

The last two results and the Cauchy-Schwartz inequality imply that the other term generated by the square in (20) tends to 0 and we have established the first equality. Using lemma 3.3 and the L^2 maximal inequality now, completes the proof. □

At this point, we have all the tools needed in the proof of Lemma 3.1. If $\xi_0 \in \Xi_0$ there is a $u_4 > 0$ such that for all ξ_0 in Lemma 3.1:

$$u_4 L^{-d+\alpha} \leq Q(\xi_0).$$

Using (18) now

$$E\left(\sup_{s \leq t_0 \log L} M_s^2\right) \leq 4U_1 L^{-d} t_0 \log L$$

so by Chebyshev's inequality and the fact that $\alpha > d/2$:

$$P\left(\left|\sup_{s \leq t_0 \log L} M_s\right| \geq u_4 L^{-d+\alpha}/2\right) \leq \frac{16U_1 L^{-d} t_0 \log L}{u_4^2 L^{-2(d-\alpha)}} = O(L^{-2\alpha+d} \log L) = o(L^{d/2-\alpha}) \to 0.$$

Consider the event $\{\tau > t_0 \log L\}$. For any $s \leq t_0 \log L$, $n_1(x, \xi_s) + n_2(x, \xi_s) < a_0|\hat{B}_0|$, for all $x \in Z^d$, so by Lemma 3.2, $\mu(\xi_s) \geq \rho Q(\xi_s)$. Consider the set:

$$A = \left\{ \left| \sup_{s \leq t_0 \log L} M_s \right| < u_4 L^{-d+\alpha}/2 \right\} \cap \{\tau > t_0 \log L\}.$$

On A we will have that for all $t \in [0, t_0 \log L]$,

$$Q(\xi_t^0) \geq u_4 L^{-d+\alpha} + \rho \int_0^t Q(\xi_s^0)ds.$$

If we let $f(t) = e^{\rho t} u_4 L^{-d+\alpha}/2$, then

$$f(t) = u_4 L^{-d+\alpha}/2 + \rho \int_0^t f(s)ds.$$

Reasoning as in the proof of Gronwall's inequality:

LEMMA 3.5. $Q(\xi_t^0) \geq f(t)$ for all $t \in [0, t_0 \log L]$

PROOF. Suppose the lemma does not hold. Let $t_1 = \inf\{t \in [0, t_0 \log L] : Q(\xi_t^0) < f(t)\}$. By right-continuity of $Q(\xi_t^0)$, $Q(\xi_{t_1}^0) \leq f(t_1)$ and $t > 0$. However, by definition of t_1, we have $Q(\xi_t^0) \geq f(t)$ on $[0, t_1)$, and by right-continuity of $Q(\xi_t^0)$ near $t = 0$, the inequality can be strict in a neighborhood of 0. Thus

$$Q(\xi_{t_1}^0) \geq u_4 L^{-d+\alpha}/2 + \rho \int_0^{t_1} Q(\xi_s^0)ds > u_4 L^{-d+\alpha}/2 + \rho \int_0^{t_1} f(s)ds = f(t_1)$$

contradiction.   □

Recalling that $t_0 = 2d/\rho$

$$f(t_0 \log L) = e^{\rho t_0 \log L} u_4 L^{-d+\alpha}/2 = u_4 L^{d+a}/2.$$

When L is large this will be $\geq \theta U_0$ the largest possible value of $Q(\xi_t)$, a contradiction, and we have finished the proof of Lemma 3.1.   □

3.1. *Proof of Theorem 2* .   Using the same technique as in Lemma 3.2, we are able to prove the extinction result in Theorem 2, which does not require the assumption of long range.

PROOF. When (3) does not hold, if $\beta \leq \nu$, the system dies out since η_t can be bounded by a subcritical contact process with birth rate β and death rate ν (the special case of η_t when $\omega = \infty$). Otherwise, we can find a θ' such that

$$\frac{\mu + \omega}{\omega} > \theta' > \frac{\beta}{\nu} > 1.$$

For η_t starting from η_0 with a finite number of non-zero sites, consider

$$S(\eta_t) = \sum_{x \in Z^d} 1_{\eta_t(x)=1} + \theta' 1_{\eta_t(x)=2}.$$

Similarly, let $\mu(\eta_t)$ be the infinitesimal mean of $S(\eta_t)$, according to same calculation as in the proof of Lemma 3.2, we have

$$\mu(\eta_t) = \sum_{x \in Z^d} [\omega(\theta' - 1) - \mu] 1_{\eta_t(x)=1} + [-\theta'\nu + f_0(x, \eta_t)\beta] 1_{\eta_t(x)=2}.$$

Note that

$$\omega(\theta' - 1) - \mu < 0$$

and that

$$-\theta'\nu + f_0(x, \eta_t)\beta \leq -\theta'\nu + \beta < 0$$

we showed that $\mu(\eta_t) \leq 0 \forall t \geq 0$, $S(\eta_t)$ is a nonnegative supermartingale. By martingale convergence theorem, $S(\eta_t)$ converge to some limit as $t \to \infty$. Note that each jump in η_t will change $S(\eta_t)$ by $1, \theta$ or $\theta - 1 > 0$. Thus to have η_t converges, with probability one there must be only finite jumps in each path of η_t, which implies that with probability one η_t will end up at configuration of all 0's, which is the absorbing state.   □

4. Spatial Location of the Positive Density Box. The argument in the previous section proves the existence of a small box $\hat{B}_x$ with positive density, but this is not useful if we do not have control over its location. To do this, we note that the graphical representation in Section 2, shows that box process ξ_t can be stochastically bounded by Krone's Model η_t starting from a same initial configuration. Krone's model can in turn be bounded by a branching random walk γ_t in which there are no deaths, 2's give birth to 2 at rate β, and births are not suppressed if the site is occupied.

LEMMA 4.1. *Suppose we start from γ_0 such that $\gamma_0(x) = 2$ for all $x \in \hat{B}_0$, $\gamma_0(x) = 0$ otherwise. Let $M_k(t)$ be the largest kth coordinate among the occupied sites at time t. If L is large enough then for any $m > 0$ we have*

$$(21) \qquad P(M_k(t) \geq (2\beta + m)Lt) \leq e^{-mt} |\hat{B}_0|.$$

From this it follows that there is a $C_{4.1} < \infty$ so,

$$(22) \qquad\qquad E(M_k^2(t_0 \log L)) \leq C_{4.1}(L \log L)^2$$

Remark: Here and later in Section 5-7, we may subscript variables according to the lemmas/ formulas where they are first introduced/defined, so it will be easier for the reader to find where they are defined.

PROOF. First we will start from the case where γ_0 has only one particle at 0. Rescale space by dividing by L. In the limit as $L \to \infty$ we have a branching random walk $\bar{\gamma}_t$ with births displaced by an amount uniform on $[-1, 1]^d$. It suffices to show that the corresponding maximum has $E\bar{M}_k^2(t_0 \log L) \leq C(\log L)^2$. To this we note that mean number of particles in A at time t

$$E(\bar{M}(t, A)) = e^{\beta t} P(S_t \in A)$$

where S_t is a random walk that makes jumps at rate β.

$$E \exp(\theta S_t) = \exp(\beta[\phi(\theta) - 1]) \quad \text{with} \quad \phi(\theta) = (e^\theta + e^{-\theta})/2\theta.$$

Large deviations implies that for any $\theta > 0$

$$P(S_t^i \geq x) \leq e^{-\theta x} \exp(\beta t[\phi(\theta) - 1])$$

and hence that

$$P(\bar{M}_k(t) \geq x) \leq e^{-\theta x} e^{\beta t} \exp(\beta t[\phi(\theta) - 1]).$$

When $\theta = 1$, $\phi(1) - 1 = 0.543$ so if L is large $\phi(1) - 1 \leq 1$ and it follows that

$$P(\bar{M}_k(t) \geq (2\beta + m)t) \leq e^{-mt}.$$

Thus if we start with $|\hat{B}_0|$ particles in $\hat{B}_0/L \subset [-1, 1]^d$ in $\bar{\gamma}_0$ then

$$P(\bar{M}_k(t) \geq 1 + (2\beta + m)t) \leq |\hat{B}_0|e^{-mt}.$$

Taking $t = t_0 \log L$ now gives the desired result. $\qquad\qquad\square$

LEMMA 4.2. *For any $a > 0$, let $M^j(t_0 \log L)$ $1 \leq j \leq L^a$ be independent and identically distributed. There is a $C_{4.2} < \infty$ so that for large L*

$$P\left(\max_{1 \leq j \leq L^a} \|M^j(t_0 \log L)\| \geq C_{4.2} L \log L \right) \leq 1/L$$

PROOF. Taking $t = t_0 \log L$ in (21) the quantity in the theorem is bounded by $L^{d+a} \exp(-mt_0 \log L)$. Taking the constant m to be large enough gives thee desired result. $\qquad\qquad\square$

5. Moving Particles in η_t. Let $H_{t,x}$ be the number of nonzero sites of η_t in $\hat{B}_x$ at time t. In this section, we will use the graphical representation in Section 2 and the technique in Durrett and Lanchier [5] to show that

LEMMA 5.1. *There are constants $\delta_{5.1} > 0$ and an $L_{5.1} < \infty$ such that for all $L > L_{5.1}$ and any initial configuration η_0 with $H_{0,0} \geq L^{d/2}$*

$$P(H_{1,\delta x} \geq \delta_{5.1} H_{0,0}) > 1 - e^{-L^{d/4}}$$

for any $v \in \{0, \pm e_1, \cdots, \pm e_d\}$

PROOF. We begin with the case $v = 0$ which is easy. Define G_0^0 to be the set of of points $x \in \hat{B}_0$, with (a) $\eta_0(x) \geq 1$, and (b) no death marks $\times$'s occur in $\{x\} \times [0,1]$. We have $\xi_t(x) \geq 1$ on $S_0 = H_{0,0} \cap G_0$, and $|S_0| \sim Binomial(H_{0,0}, e^{-\mu})$, so the desired result follows from large deviations for the Binomial.

For $v \neq 0$, define G_0 to be the set of of points in G_0^0 for which (c) there exists a (•) in $\{x\} \times [0, 1/2]$. We define G_v to be the set of points y in $\hat{B}_v$ so that there are no $\times$'s in $\{y\} \times [0, 1]$. For any $x \in \hat{B}_0$ and $y \in \hat{B}_v$ we say that x and y are connected (and write $x \to y$) if there is an arrow from x to y in $(1/2, 1)$. By definition of our process $\eta_1(y) \geq 1$ for all y in

$$S = \{y : y \in G_v, \text{ there exists an } x \in G_0 \text{ so that } x \to y\}.$$

It is easy to see that

(23) $$|G_0| \sim Binomial[H_{0,0}, e^{-\mu}(1 - e^{-\omega/2})]$$

and conditional on $|G_0|$, we can determine the conditional distribution of S:

(24) $$|S| \sim Binomial(|\hat{B}_v|, e^{-\mu}[1 - e^{-\beta|G_0|/2|B_0|}]).$$

Since being the recipient of a birth from $\hat{B}_0$ are independent for different sites in G_v by Poisson thinning.

Since that binomial distribution decays exponentially fast away from the mean, there is some constant $c > 0$ such that

(25) $$P(|G_0| > H_{0,0}e^{-\mu}(1 - e^{-\omega/2})/2) \geq 1 - e^{-cL^{d/2}}.$$

To simplify the next computation, we note that $1 - e^{-\beta r} \sim \beta r$ as $r \to 0$ so if the ϵ_0 in the definition of the small box is small enough $1 - e^{-\beta|G_v|/2|B_0|} \geq$

$\beta|G_v|/4|B_0|$. Let $p = e^{-\mu}\beta|G_0|/4|B_0|$. A standard large deviations result, see e.g., Lemma 2.8.5 in [4] shows that if $X = \text{Binomial}(N, p)$ then

$$P(X \leq Np/2) \leq \exp(-Np/8)$$

from which the desired result follows.   □

Iterating this result $O(\log L)$ we can move a positive density box to one with at least L^α particles. Let $\|\cdot\|_1$ be the L^1-norm on Z^d.

LEMMA 5.2.   *For any* $\alpha \in (d/2, d)$, *let* $C_{5.2}$ *be a constant such that* $C_{5.2} \log \delta_{5.1} > \alpha - d$. *There is a finite* $L_{5.2} > L_{5.1}$ *such that for all* $L > L_{5.2}$, *any initial configuration* η_0 *with* $H_{0,0} \geq a_0|\hat{B}_0|$, *and any* $x \in Z^d$ *such that* $\|x\|_1 \leq C_{5.2} \log L$, *where* ,

$$P(H_{x,[C_{5.2}\log L]} \geq L^\alpha) \geq 1 - e^{-L^{d/4}/2}$$

PROOF.   Let $n = [C_{5.2} \log L]$. We can find a sequence $x_0 = 0, x_1, \cdots, x_n = x$ such that for all $i = 0, \cdots, n-1$, $x_{i-1} - x_i \in \{0, \pm e_1, \cdots, \pm e_d\}$. For any $i = 1, \cdots, n$ define the event

$$A_i = \{H_{x_i,i} \geq \delta_{5.1}^i H_{0,0}\}.$$

By the definition of $C_{5.2}$, $H_{x,[C_{5.2}\log L]} \geq L^\alpha$ on A_n. To estimate $P(A_n)$ note that by Lemma 5.1

$$P(A_n) \geq 1 - \sum_{i=1}^n P(A_i^c) \geq 1 - \sum_{i=1}^n P(A_i^c|A_{i-1}) \geq 1 - C_{5.2}(\log L)e^{-L^{d/4}} \geq 1 - e^{-L^{d/4}/2}$$

when L is large.   □

6. Block Construction.   At this point,we have all the tools to construct the block event and complete the proof of Theorem 1. Let $0 < a < \alpha/2 - d/4$, $K = L^{1+2a/3}$, $\Gamma_m = 2mKe_1 + [-K, K]^d$, and $\Gamma'_m = 2mKe_1 + [-K/2, K/2]^d$. If $m + n$ is even, we say that (m, n) is wet if there is a positive density small box in Γ'_m at some time in $[nL^a, nL^a + C_6 \log L]$, where $C_6 = C_{5.2} + t_0$. Our goal is to show

LEMMA 6.1.   *If* (m, n) *is wet then with high probability so is* $(m+1, n+1)$, *and the events which produce this are measurable with respect to the graphical representation in* $(\Gamma_m \cup \Gamma_{m+1}) \times [nL^a, (n+1)L^a + C_6 \log L]$.

Once this is done, Theorem 1 follows. See [2] for more details.

PROOF. To prove Lemma 6.1, we will alternate two steps, starting from the location of the initial positive density box $\hat{B}_{y_0}$ at time T_0. Let $A_0 = \{T_0 < \infty\}$ which is the whole space. Assume given a deterministic sequence δ_i. with $\|\delta_i\| \leq C_{5.2} \log L$. If we do not ever meet a failure, the construction will terminate at the first time that $T_i > (n+1)L^a$. The actual number steps will be random but the number is $\leq N = \lceil L^a/[C_{5.2} \log L] \rceil$. We will do our estimate of the probability of success supposing that N steps are required, to lower bound the probability of success when we stop at the first time $T_i \geq (n+1)L^a$. Suppose $i \geq 1$.

Deterministic Moving. If at the stopping time $T_{i-1} < \infty$ we have a positive density small box $\hat{B}_{y_{i-1}}$, then we use results in Section 5 to produce a small box $\hat{B}_{y_{i-1}+\delta_i}$ with at least L^α nonzero sites at time $S_i = T_{i-1} + [C_{5.2} \log L]$. If we fail we let $S_i = \infty$ and the construction terminates. Let $A_i^+ = \{S_i < \infty\}$.

Random Recovery. If at the stopping time $S_i < \infty$ we have a small box $\hat{B}_{y_{i-1}+\delta_i}$ with at least L^α nonzero sites then we set all of the sites outside the box to 0, and we use the recovery lemma to produce a positive density small box $\hat{B}_{y_i}$ at time $S_i \leq T_i \leq S_i + t_0 \log L$. Again if we fail we let $T_i = \infty$ and the construction terminates. Let $A_i = \{T_i < \infty\}$. Let $\Delta_i(\omega) = y_i - (y_{i-1} - \delta_i)$ on A_i, and $= 0$ on A_i^c.

If we define the partial sums $\bar{y}_i = y_0 + \sum_{j=1}^i \delta_i$ and $\Sigma_i = \sum_{j=1}^i \Delta_j$ then we have $y_i = \bar{y}_i + \Sigma_i$. We think of $\bar{y}_i$ as the mean of the location of the positive density box and Σ_i as the random fluctuations in its location. We make no attempt to adjust the deterministic movements δ_i to compensate for the fluctuations. Let $y_{end} = (y_{end}^1, 0, \cdots, 0) \in Z^d$ be such that $2K(m+1)e_1 \in \hat{B}_{y_{end}}$. We define the δ_i to reduce the coordinates y_0^k, $k = 2, \ldots d$ to 0 and then increase y_0^1 to y_{end}^1, in all cases using steps of size $\leq C_{5.2} \log L$. Note that

$$\|y_0 - y_{end}\|_1 = O(K)/\ell = O(L^{2a/3}) = o(N)$$

so we can finish the movements well before N steps. And once this is done we set the remaining δ_i to 0. Moreover, note that each successful step in our iteration takes a time at most $C_6 \log L$. Thus we will get to y_{end} by $t = nL^a + O(L^{2a/3})C_6 \log L < (n+1)L^a$. At the first time $T_i \geq (n+1)L^a$, we already have

(26) $$\delta_i = 0, \ \bar{y}_i = y_{end}$$

At this point, we are ready to state the main lemma of this section that controls the spatial movement in our iteration:

LEMMA 6.2. *For any initial configuration $\eta(T_0)$ so that there is a small box $\hat{B}_{y_0} \subset \Gamma'_m$, and any sequence δ_i, $i \leq N$ with $\|\delta_i\| \leq C_{5.2} \log L$ and any $\epsilon > 0$, there is a good event G_N with $G_N \to 1$ as $L \to \infty$ so that (a) $G_N \subset A_N$, (b) on G_N, $\|y_i - \bar{y}_i\| < \epsilon L^{2a/3}$ for $1 \leq i \leq N$. (c) G_N depends only on the gadgets of graphical representation in $\Gamma_m \cup \Gamma_{m+1}$*

PROOF. The first step is to show that $P(A_N) \to 1$ as $L \to \infty$. For the ith deterministic moving step, using the strong Markov property and Lemma 5.2, we have
$$P(A_i^+|A_{i-1}) > 1 - e^{-L^{d/4}/2}.$$
Then for the random recovery phase, according to Lemma 3.1, we have the conditional probability of success:
$$P(A_i|A_i^+) = P(\tau_i < t_0 \log L) > 1 - L^{d/2 - \alpha}.$$
Combining the two observations, we have

(27) $P(A_i|A_{i-1}) > (1 - e^{-L^{d/4}/2})(1 - L^{d/2-\alpha}) > 1 - e^{-L^{d/4}/2} - L^{d/2-\alpha}$

which implies

$$P(A_N) \geq 1 - \sum_{i=1}^{N} P(A_i^c) \geq 1 - \sum_{i=1}^{N} P(A_i^c|A_{i-1})$$
$$\geq 1 - L^a(e^{-L^{d/4}/2} + L^{d/2-\alpha}) \geq 1 - 2L^{d/4-\alpha/2} \to 1.$$

The next step is to control the fluctuations in the movement of our box.

LEMMA 6.3. *Let $\mathcal{F}(T_i)$ be the filtration generated by events in the graphical representation up to stopping time T_i. For any $1 \leq k \leq d$, $\{\Sigma_i^k\}_{i=1}^{N}$ is a martingale with respect to $\mathcal{F}(T_i)$. $E(\Sigma_i^k) = 0$, and $var(\Sigma_N^k) \leq C_{6.3} L^a \log L$ so for any ϵ we have*

$$P\left(\max_{i \leq N} \|\Sigma_i\| > \epsilon L^{2a/3}\right) \to 0$$

PROOF. Consider the conditional expectation of Δ_i^k under $\mathcal{F}(S_i)$. By remarks in Section 2, we have $E(\Delta_i^k|\mathcal{F}(S_i)) = 0$. Noting that $2\ell|\Delta_i^k|$ can be bounded by the largest kth coordinate among the occupies sites of the corresponding branching random walk at time $t_0 \log L$, Lemma 4.1 implies that

$E((\Delta_i^k)^2|\mathcal{F}(S_i)) \leq C(\log L)^2$. By orthogonality of martingale increments $\text{var}(\Sigma_N^k) \leq NC(\log L)^2$. Since $N \leq L^a/[C_{5.2}\log L] + 1$, we have the desired bound on the variances and the desired result follows from L^2 maximal inequality for martingales. $\qquad\square$

To check (c) now note that under A_i the success of A_{i+1}^+ depends only on gadgets in:

$$(2\ell y_i + [-\ell C_{5.2}\log L, \ell C_{5.2}\log L]^d) \times [T_i, S_{i+1}]$$

and that when the ith copy of the truncated process never wander outside of

$$D_i = 2\ell(y_{i-1} + \delta_i) + [-C_{4.2}(L\log L), C_{4.2}(L\log L)]^d$$

the success of A_{i+1} under A_{i+1}^+ depends only on gadgets in D_i. According to Lemma 4.2 and the fact that $N < L^a$, with probability $\geq 1 - L^{-1} = 1 - o(1)$, our construction only depends on gadgets in the box:
(28)
$$\cup_{i=0}^{N-1} \big[(2\ell y_i + [-\ell C_{5.2}\log L, \ell C_{5.2}\log L]^d) \times [T_i, S_{i+1}] \cup$$
$$(2\ell(y_i + \delta_{i+1}) + [-C_{4.2}(L\log L), C_{4.2}(L\log L)]^d) \times [S_{i+1}, T_{i+1}] \big].$$

The locations of the y_i are controlled by Lemma 6.3 so that one can easily found the box defined in (28) a subset of $\Gamma_n \cup \Gamma_{n+1}$, and proof of Lemma 6.2 is complete. $\qquad\square$

Back to the proof of Lemma 6.1. On G_N it follows from (26) and Lemma 6.2, when we stop at the first time $T_i \geq (n+1)L^a$:

$$\|y_i - 2K(m+1)e_1\| \leq \ell(1 + 2\|y_i - \bar{y}_i\|) \leq 4\epsilon K$$

which implies

$$\hat{B}_{y_i} = 2\ell y_i + (-\ell, \ell]^d \subset \Gamma'_{n+1}.$$

Noting that the success of G_N only depends on gadgets in $\Gamma_n \cup \Gamma_{n+1}$, we have proved that G_N is a subset of the block event in Lemma 6.1, which completes the proof of Lemma 6.1 and Theorem 1. $\qquad\square$

7. Proof of Theorem 3. Our **First Step** is to construct the test functions for S and T and show that, under (6), they have positive derivatives for all sites in the region of interest. For any $\kappa > 0$, according to (7), we can choose $\Sigma_0 \in (0, 1)$ such that

$$(29) \qquad\qquad \frac{\nu}{\omega_0} < \frac{\beta(1 - \Sigma_0)}{2^d(\mu + \omega_0)}.$$

Let

$$(30) \qquad \gamma_0 \in \left(\frac{\nu}{\omega_0}, \frac{\beta(1 - \Sigma_0)}{2^d(\mu + \omega_0)} \right)$$

and let:

$$(31) \qquad T_0 = \frac{\Sigma_0}{1 + \gamma_0}, \qquad S_0 = \frac{\gamma_0 \Sigma_0}{1 + \gamma_0}.$$

Note that

$$(32) \qquad S_0 + T_0 = \Gamma_0, \qquad S_0/T_0 = \gamma_0.$$

Recall that for any $x \in Z^d$ and $r \geq 0$, $B(x, r)$ is defined in Section 2 to be the L^∞ neighborhood of x with range r. With S_0, T_0 defined as above and $\epsilon_{7.1}$ to be specified later, define the test functions $S_{test}(x, 0)$ and $T_{test}(x, 0)$ as follows: let $S_{test}(x, 0) = S_0$ on $B(0, M - \epsilon_{7.1})$, $S_{test}(x, 0) = 0$ on $B(0, M)^c$, and

$$(33) \qquad S_{test}(x, 0) = \frac{d[x, B(0, M)^c]S_0}{d[x, B(0, M)^c] + d[x, B(0, M - \epsilon_{7.1})]}$$

for $x \in B(0, M - \epsilon_{7.1})^c \cap B(0, M)$. Similarly, let $T_{test}(x, 0) = T_0$ on $B(0, M - 3\epsilon_{7.1})$, $T_{test}(x, 0) = 0$ on $B(0, M - 2\epsilon_{7.1})^c$, and

$$(34) \qquad T_{test}(x, 0) = \frac{d[x, B(0, M - 2\epsilon_{7.1})^c]T_0}{d[x, B(0, M - 2\epsilon_{7.1})^c] + d[x, B(0, M - 3\epsilon_{7.1})]}$$

for $x \in B(0, M - 3\epsilon_{7.1})^c \cap B(0, M - 2\epsilon_{7.1})$. In the definitions above, $M = \max\{4, 4\kappa\}$, $d(x, A)$ be the distance between $x \in R^d$ and $A \subset R^d$ under L^∞-norm, $\epsilon_{7.1} = \epsilon_{7.1}(\beta, \omega_0, \mu, \nu, \kappa)$ is some positive constant that will be specified later in the proof of Lemma 7.1. The following lemma shows that the test functions have positive derivatives under IDE (6).

Remark: Throughout the discussion in this section, all the ϵ's, δ's, t's and c's introduced are constants independent to the choice of L.

LEMMA 7.1.   *Under the conditions in Theorem 3, there are $\epsilon_{7.1}$ and $\epsilon_1 >$ 0 so that under IDE (6):*

$$(35) \qquad \begin{aligned} \frac{dS_{test}(x, 0)}{dt} &\geq 4\epsilon_1 \text{ for all } x \in B(0, M) = \overline{\{S_{test} > 0\}} \\ \frac{dT_{test}(x, 0)}{dt} &\geq 4\epsilon_1 \text{ for all } x \in B(0, M - 2\epsilon_{7.1}) = \overline{\{T_{test} > 0\}} \end{aligned}$$

PROOF. With T_{test} and S_{test} defined as above, for any $x \in B(0, M)$, the local grass density can be upper bounded as follows (see Figure 2 for the case when $d = 2$):

$$D_\kappa^G(x, 0) \leq 1 - \left(\frac{\kappa - \epsilon_{7.1}}{2\kappa}\right)^d S_0 \leq 1 - 2^{-d}\left(1 - d\frac{\epsilon_{7.1}}{\kappa}\right)S_0.$$
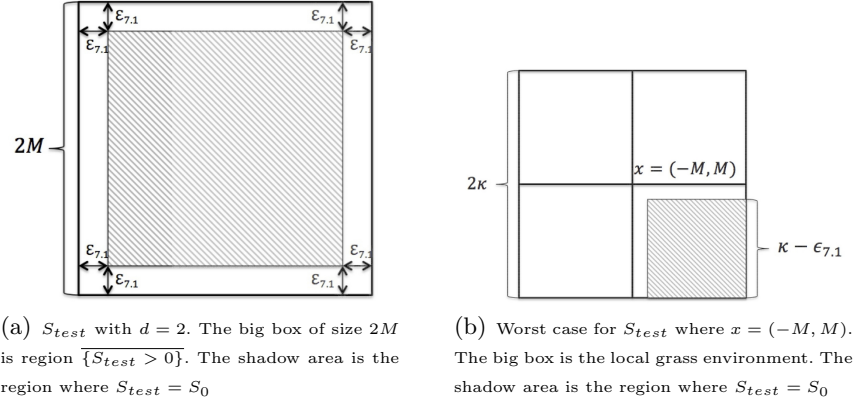

(a) S_{test} with $d = 2$. The big box of size $2M$ is region $\overline{\{S_{test} > 0\}}$. The shadow area is the region where $S_{test} = S_0$

(b) Worst case for S_{test} where $x = (-M, M)$. The big box is the local grass environment. The shadow area is the region where $S_{test} = S_0$

Fig 2: S_{test} with $d = 2$

Noting that $\delta_0 < 2^{-d}S_0$, let

(36)
$$\epsilon_{(36)} = \frac{\kappa(1 - 2^d\delta_0/S_0)}{4d}.$$

Then it is easy to check that when $\epsilon_{7.1} \leq \epsilon_{(36)}$

$$D_\kappa^G(x, 0) < 1 - \delta_0$$

which implies that

(37)
$$\omega[D_\kappa^G(x, 0)] \equiv \omega_0$$

for all $x \in B(0, M)$. I.e. all sites in the region of test function live in a environment with a higher growth rate ω_0. It is easy to see that the derivative of T_{test} can be lower bounded by its derivative at the top corner, i.e. for any $x \in B(x, M - 2\epsilon_{7.1})$

(38)
$$\frac{dT_{test}(x, 0)}{dt} \geq \frac{dT_{test}(M - 3\epsilon_{7.1}, 0)}{dt} \geq \omega_0 S_0 - \nu T_0$$

and this holds for all $\kappa > 0$. Note that $\gamma_0 = S_0/T_0$ according to (31). Combining this observation with the definition of γ_0 in (30)

$$(39) \qquad \omega_0 S_0 - \nu T_0 = \omega_0 T_0 \left(\frac{S_0}{T_0} - \frac{\nu}{\omega_0} \right) = \omega_0 T_0 \left(\gamma_0 - \frac{\nu}{\omega_0} \right) > 0.$$

Thus we have the derivative of T_{test} is always positive for all $x \in B(0, M - 2\epsilon_{1,1})$.

Similarly, we can control the lower bound of derivative for test function S_{test} as follows: for any $x \in B(x, M)$

$$(40) \qquad \begin{aligned} \frac{dS_{test}(x,0)}{dt} &\geq \frac{\beta T_0 (1 - 3\epsilon_{7.1})^d}{2^d} (1 - S_0 - T_0) - (\mu + \omega_0) S_0 \\ &\geq \frac{\beta T_0 (1 - 3d\epsilon_{7.1})}{2^d} (1 - S_0 - T_0) - (\mu + \omega_0) S_0 \end{aligned}$$

For the right side of (40) , according to (32)

$$\begin{aligned} \frac{dS_{test}(x,0)}{dt} &\geq \frac{\beta T_0 (1 - 3d\epsilon_{7.1})}{2^d} (1 - S_0 - T_0) - (\mu + \omega_0) S_0 \\ &= T_0 (\mu + \omega_0) \left[(1 - 3d\epsilon_{7.1}) \frac{\beta(1 - S_0 - T_0)}{2^d(\mu + \omega_0)} - \frac{S_0}{T_0} \right] \\ &= T_0 (\mu + \omega_0) \left[(1 - 3d\epsilon_{7.1}) \frac{\beta(1 - \Sigma_0)}{2^d(\mu + \omega_0)} - \gamma_0 \right] \end{aligned}$$

Again recalling the definition in (30) that

$$\frac{\beta(1 - \Sigma_0)}{2^d(\mu + \omega_0)} > \gamma_0$$

let

$$(41) \qquad \epsilon_{(41)} = \left[1 - \frac{2^d(\mu + \omega_0)\gamma_0}{\beta(1 - \Sigma_0)} \right] / (6d).$$

So for any $\epsilon_{7.1} \leq \epsilon_{(41)}$ and $x \in B(0, M)$

$$(42) \qquad \frac{dS_{test}(x,0)}{dt} \geq T_0(\mu + \omega_0) \left[(1 - 3d\epsilon_{(41)}) \frac{\beta(1 - \Sigma_0)}{2^d(\mu + \omega_0)} - \gamma_0 \right] > 0.$$

Thus let

$$(43) \qquad \epsilon_{7.1} = \min\{\epsilon_{(36)}, \epsilon_{(41)}\} > 0$$

and

$$\epsilon_1 = \frac{1}{4} \min \left\{ \omega_0 T_0 \left(\gamma_0 - \frac{\nu}{\omega_0} \right), T_0(\mu + \omega_0) \left[(1 - 3d\epsilon_{(41)}) \frac{\beta(1 - \Sigma_0)}{2^d(\mu + \omega_0)} - \gamma_0 \right] \right\} > 0.$$

Combining (39) and (42) and the proof is complete. $\qquad \square$

With the test functions constructed, our **Second Step** is similar to the proof of Theorem 1. We introduce the truncated (box) version of the Staver-Levin model, and as before, denote the process by $\bar{\xi}_t$. For $\ell = \epsilon L$, where the exact value of ϵ is specified later in Lemma 7.4, $\bar{\xi}_t$ has birth rate $\beta|\mathcal{N}(x,L)|/(2L+1)^d$, where $\mathcal{N}(x,L)$ in the truncated neighborhood defined in Section 2. Type 1's and 2's in $\bar{\xi}_t$ die at the same rates as in the original χ_t, while a growth of a sapling into a tree occurs at rate:

$$\bar{\omega}[\bar{G}(x,\xi)] = \begin{cases} \omega_0 & \bar{G}(x,\xi) \in [0, 1-\delta) \\ \omega_1 & \bar{G}(x,\xi) \in [1-\delta, 1] \end{cases}$$

where

$$\bar{G}(x,\xi) = \frac{\text{\# of 0's in } \mathcal{N}(x,K)}{|\mathcal{N}(x,K)|},$$

$\delta = \delta_0 + 4d\epsilon$. First of all, with the same argument as in Section 2, one can immediately see that the number of different types in each small box forms a Markov process $\bar{\zeta}_t$. According to (8):

$$\bar{G}(x,\xi) \leq 1 - \frac{\text{\# of 1+2's in } \mathcal{N}(x,K)}{|B(x,K)|} \leq G(x,\xi) + 1 - (1-4\epsilon)^d \leq G(x,\xi) + 4d\epsilon,$$

combining this with the definition of δ, we have for any $\xi' \geq \xi$, $\omega[G(x,\xi')] \geq \bar{\omega}[\bar{G}(x,\xi)]$, which implies that the truncated $\bar{\xi}_t$ once again is dominated by the original χ_t. Thus in order to prove Theorem 3, it suffices to show that $\bar{\xi}_t$ survives.

The **Third Step** is to construct a initial configuration of $\bar{\xi}_0$ according to the test functions defined in (33) and (34). For any $x \in Z^d$:

- If $2\ell x \notin B(0, ML)$ there is no saplings or trees in $\hat{B}_x$ under $\bar{\xi}_0$.
- If $2\ell x \in B(0, ML)$, $n_1(x, \bar{\xi}_0) = |\hat{B}_0| S_{test}(2\ell x/L)$, $n_2(x, \bar{\xi}_0) = |\hat{B}_0| T_{test}(2\ell x/L)$.

As noted earlier in the box process $\bar{\xi}_t$, the locations of the 1's and 2's inside each small box makes no difference.

We then look at $f_i(x, \bar{\xi}_t)$, the densities of type i in each small box. For any x, the infinitesimal means of f_1 and f_2 can be written as follows:
(44)

$$\mu_1(x, \bar{\xi}) = (2L+1)^{-d}|\hat{B}_0| \left(\sum_{y:\hat{B}_y \subset \mathcal{N}(x,L)} f_2(y, \bar{\xi}) \right) f_0(x, \bar{\xi})\beta - [\bar{\omega}(\bar{G}(x,\bar{\xi})) + \mu]f_1(x, \bar{\xi})$$

$$\mu_2(x, \bar{\xi}) = \bar{\omega}(\bar{G}(x, \bar{\xi}))f_1(x, \bar{\xi}) - \nu f_2(x, \bar{\xi}).$$

We prove that

LEMMA 7.2.   *There is a $\epsilon_{7.2} > 0$ such that for any $\ell = \epsilon L \leq \epsilon_{7.2} L$, under $\bar{\xi}_0$:*

- *$\mu_1(x, \bar{\xi}_0) \geq 2\epsilon_1$ for all $2\ell x \in B(0, ML + 4\ell)$.*
- *$\mu_2(x, \bar{\xi}_0) \geq 2\epsilon_1$ for all $2\ell x \in B(0, ML - 2\epsilon_{7.1} L + 4\ell)$.*

PROOF. First noting that $\delta = \delta_0 + 4d\epsilon$, $\delta_0 < 2^{-d} S_0$, for

$$(45) \qquad\qquad \epsilon_{(45)} = (2^{-d} S_0 - \delta_0)/(8d)$$

$\delta = \delta_0 + 4d\epsilon < 2^{-d} S_0$ for $\epsilon \leq \epsilon_{(45)}$. Moreover, for all x such that $2\ell x \in B(0, ML + 4\ell)$,

$$(46) \qquad \begin{aligned} \bar{G}(x, \bar{\xi}_0) &= \frac{\text{\# of 0's in } \mathcal{N}(x, K)}{|\mathcal{N}(x, K)|} \\ &\leq 1 - \left(\frac{\kappa - \epsilon_{7.1} - 10\epsilon}{2\kappa} \right)^d S_0 \leq 1 - 2^{-d} S_0 (1 - d\epsilon_{7.1} - 10d\epsilon). \end{aligned}$$

Let

$$(47) \qquad\qquad \epsilon_{(47)} = \frac{S_0(1 - d\epsilon_{7.1}) - 2^d \delta_0}{20dS_0 + 2^{d+2}d} > 0.$$

It is easy to see that $G(x, \bar{\xi}_0) < 1 - \delta_0 - 4d\epsilon = 1 - \delta$ for all $\epsilon \leq \epsilon_{(47)}$, which implies $\bar{\omega}(\bar{G}(x, \bar{\xi}_0)) \equiv \omega_0$ for all $2\ell x \in B(0, ML + 4\ell)$. Thus

$$\mu_2(x, \bar{\xi}_0) = \bar{\omega}(\bar{G}(x, \bar{\xi}_0)) f_1(x, \bar{\xi}_0) - \nu f_2(x, \bar{\xi}_0) \geq \omega_0 S_0 - \nu T_0 > 4\epsilon_1$$

for all $2\ell x \in B(0, ML - 2\epsilon_{7.1} L + 4\ell)$. Then for the infinitesimal mean of type 1:

$$\mu_1(x, \bar{\xi}_0) = (2L+1)^{-d} |\hat{B}_0| \left(\sum_{y : \hat{B}_y \subset \mathcal{N}(x, L)} f_2(y, \bar{\xi}_0) \right) f_0(x, \bar{\xi}_0) \beta - [\omega_0 + \mu] f_1(x, \bar{\xi}_0)$$

Note that

$$\begin{aligned} \inf_{2\ell x \in B(0, ML + 4\ell)} &(2L+1)^{-d} |\hat{B}_0| \left(\sum_{y : \hat{B}_y \subset \mathcal{N}(x, L)} f_2(y, \bar{\xi}_0) \right) f_0(x, \bar{\xi}_0) \beta \\ &\geq 2^{-d} (1 - 3\epsilon_{7.1} - 10\epsilon)^d T_0 (1 - S_0 - T_0) \beta \\ &\geq 2^{-d} (1 - 3d\epsilon_{7.1} - 10d\epsilon) T_0 (1 - S_0 - T_0) \beta. \end{aligned}$$

Recalling (42), (43) and the definition of ϵ_1, let

$$(48) \qquad \epsilon_{(48)} = \frac{2^{d-1}\epsilon_1}{5dT_0(1 - S_0 - T_0)\beta}.$$

For all $\epsilon \leq \epsilon_{(48)}$ and any $2\ell x \in B(0, ML + 4\ell)$,

$$\mu_2(x, \bar{\xi}_0) \geq 2^{-d}(1 - 3d\epsilon_{7.1})T_0(1 - S_0 - T_0) - (\omega_0 + \mu)S_0 - 2\epsilon_1 \geq 2\epsilon_1.$$

Overall, let

$$\epsilon_{7.2} = \min\{\epsilon_{(45)}, \epsilon_{(47)}, \epsilon_{(48)}\}.$$

It satisfies the condition of this lemma by definition. $\qquad\square$

Moreover, since that the inequalities for $\bar{G}$'s in the proof above are strict and that all other terms in the infinitesimal mean are continuous, we have:

LEMMA 7.3. *There is some $\delta_{7.3} > 0$ so that for any configuration $\bar{\xi}_0'$ with*

$$|f_i(x, \bar{\xi}_0) - f_i(x, \bar{\xi}_0')| \leq \delta_{7.3}, \ i = 1, 2; \ 2\ell x \in B(0, 2ML)$$

we have

- $\mu_1(x, \bar{\xi}_0') \geq \epsilon_1$ *for all $2\ell x \in B(0, ML + 4\ell)$.*
- $\mu_2(x, \bar{\xi}_0') \geq \epsilon_1$ *for all $2\ell x \in B(0, ML - 2\epsilon_{7.1}L + 4\ell)$.*

for all $\ell = \epsilon L \leq \epsilon_{7.2}L$.

PROOF. First note that for any x such that $2\ell x \in B(0, ML + 4\ell)$,

$$|\bar{G}(x, \bar{\xi}_0) - \bar{G}(x, \bar{\xi}_0')| \leq 2\delta_{7.3}.$$

Let

$$(49) \qquad \delta_{(49)} = \left[2^{-d}S_0(1 - d\epsilon_{7.1} - 10d\epsilon_{7.2}) - \delta_0 - 4d\epsilon_{7.2}\right]/4 > 0.$$

For any $\delta_{7.3} \leq \delta_{(49)}$, recalling (46) and (47), we have

$$\begin{aligned}
\bar{G}(x, \bar{\xi}_0') &\leq \bar{G}(x, \bar{\xi}_0) + 2\delta_{7.3} \\
&\leq 1 - 2^{-d}S_0(1 - d\epsilon_{7.1} - 10d\epsilon_{7.2}) + 2\delta_{(49)} \\
&< 1 - \delta_0 - 4d\epsilon_{7.2} \leq 1 - \delta_0 - 4d\epsilon = 1 - \delta
\end{aligned}$$

which implies that

$$(50) \qquad \bar{\omega}(G(x, \bar{\xi}_0')) \equiv \omega_0$$

for all x such that $2\ell x \in B(0, ML + 4\ell)$. Furthermore, under (50), (44) implies that:

(51) $$|\mu_2(x, \bar{\xi}_0) - \mu_2(x, \bar{\xi}'_0)| \leq \omega_0 \delta_{7.3} + \nu \delta_{7.3} = (\omega_0 + \nu)\delta_{7.3}$$

(52)

$$|\mu_1(x, \bar{\xi}_0) - \mu_1(x, \bar{\xi}'_0)| \leq (2L+1)^{-d}|\hat{B}_0| \left(\sum_{y:\hat{B}_y \subset \mathcal{N}(x,L)} f_2(y, \bar{\xi}_0) \right) 2\beta \delta_{7.3}$$

$$+ (2L+1)^{-d}|\hat{B}_0| \left(\sum_{y:\hat{B}_y \subset \mathcal{N}(x,L)} \delta_{7.3} \right) f_0(x, \bar{\xi}_0)\beta$$

$$+ 2\delta_{7.3}^2 + (\omega_0 + \mu)\delta_{7.3}$$

for all x such that $2\ell x \in B(0, ML + 4\ell)$. Noting that $f_i \leq 1$, $\delta_{7.3} \leq 1$, (52) can be simplified as:

$$|\mu_1(x, \bar{\xi}_0) - \mu_1(x, \bar{\xi}'_0)| \leq (2 + 3\beta + \omega_0 + \mu)\delta_{7.3}.$$

Thus let

$$\delta_{7.3} = \min \left\{ \delta_{(49)}, \frac{\epsilon_1}{2(\omega_0 + \nu)}, \frac{\epsilon_1}{2(2 + 3\beta + \omega_0 + \mu)} \right\}.$$

(50)-(52) show that $\delta_{7.3}$ satisfies the conditions in our lemma. $\qquad\square$

Both S_{test} and T_{test} are Lipchitz with constants $S_0 \epsilon_{7.1}^{-1}$ and $T_0 \epsilon_{7.1}^{-1}$. Let C_{lip} be the max of these two constants. At this point, we are ready to specify the size of our small box and have the lemma as follows:

LEMMA 7.4.   *For $\epsilon = \epsilon_{7.4}$, where*

(53) $$\epsilon_{7.4} = \min \left\{ \frac{\delta_{7.3}\epsilon_1}{16(\beta + \omega_0 + \mu)C_{lip}}, \frac{\epsilon_{7.2}}{2} \right\},$$

$\bar{\xi}_t$ be the truncated process starting from $\bar{\xi}_0$. At time

$$t_{7.4} = \frac{\delta_{7.3}}{2(\beta + \omega_0 + \mu)}$$

there is some $C_{7.4} < \infty$ such that the probability that

- *$f_1(x, \bar{\xi}_{t_{7.4}}) \geq f_1(x, \bar{\xi}_0) + c_{7.4}$, any x such that $2\ell x \in B(0, ML + 4\ell)$*

- $f_2(x, \bar{\xi}_{t_{7.4}}) \geq f_2(x, \bar{\xi}_0) + c_{7.4}$, any x such that $2\ell x \in B(0, ML - 2\epsilon_{7.1}L + 4\ell)$

is greater than $1 - C_{7.4}L^{-d}$ when L is large, where

$$c_{7.4} = \frac{\delta_{7.3}\epsilon_1}{8(\beta + \omega_0 + \mu)}.$$

PROOF. Consider the stopping time
(54)
$$\bar{\tau} = \min\{t : \exists x : 2\ell x \in B(0, 2ML), i = 1 \text{ or } 2, |f_i(x, \bar{\xi}_0) - f_i(x, \bar{\xi}_0')| > \delta_{7.3}\}.$$

Note that each site in our system flip at a rate no larger than $\beta + \omega_0 + \mu$. According to standard large deviations result as we used in Lemma 5.1, there is some $c_{5.1}, C_{5.1} \in (0, \infty)$ independent to L such that

$$P(\bar{\tau} \leq t_{7.4}) \leq C_{5.1} \exp(-c_{5.1}L^d) < C_{5.1}L^{-d}$$

when L is large. Now consider, $\sigma_i^2(x, \bar{\xi}_t)$, the infinitesimal variances of the local densities. According to exactly the same calculation as we did in Lemma 3.3, there is a $C_{3.3} < \infty$ such that

(55) $$\sigma_i^2(x, \bar{\xi}) \leq C_{3.3}L^{-d}$$

for all $x \in Z^d$, $i = 1, 2$ and all configurations $\bar{\xi}$. Thus we can again define Dynkin's martingale:

(56) $$\bar{M}_i(x, t) = f_i(x, \bar{\xi}_t) - f_i(x, \bar{\xi}_0) - \int_0^t \mu_i(x, \bar{\xi}_t)dt$$

and Lemma 3.4 implies that there is a $C_{3.4}$

$$P(|M_i(x, t)| > c_{7.4}) < C_{3.4}L^{-d}.$$

Consider the event

(57) $$A_i(x) = \{\tau > t_{7.4}\} \cap \{|M_i(x, t)| < c_{7.4}\}.$$

By definition, there is some $U_{7.4} < \infty$, independent to L, such that

$$P(A_i(x)) > 1 - U_{7.4}L^{-d}, \forall x \ s.t. 2\ell x \in B(0, ML + 4\ell).$$

For any x such that $2\ell x \in B(0, ML + 4\ell)$, when $A_1(x)$ holds, Lemma 7.3 implies that for any x

(58) $$f_1(x, \bar{\xi}_{t_{7.4}}) \geq \int_0^{t_{7.4}} \epsilon_1 dt - c_{7.4} > c_{7.4}.$$

Similarly, for any x such that $2\ell x \in B(0, ML - 2\epsilon_{7.1}L + 4\ell)$, when $A_2(x)$ holds,

$$(59) \qquad f_2(x, \bar{\xi}_{t_{7.4}}) \geq \int_0^{t_{7.4}} \epsilon_1 dt - c_{7.4} > c_{7.4}.$$

So let

$$(60) \quad A = \left(\bigcap_{x: 2\ell x \in B(0, ML+4\ell)} A_1(x) \right) \cap \left(\bigcap_{x: 2\ell x \in B(0, ML-2\epsilon_{7.1}L+4\ell)} A_2(x) \right).$$

The conditions in our lemma are satisfied on the event A. Noting that

$$P(A) \geq 1 - \sum_{x: 2\ell x \in B(0, ML+4\ell)} [P(A_1^c(x)) + P(A_2^c(x))] \geq 1 - \frac{4M^d}{\epsilon_{7.4}^d} U_{7.4} L^{-d}$$

let $C_{7.4} = 4M^d U_{7.4}/\epsilon_{7.4}^d$ and the proof is complete. $\qquad \square$

For any $x \in Z^d$ and any $\xi \in \{0, 1, 2\}^{Z^d}$, define $\text{shift}(\xi, x)$ to be the configuration that for any $y \in Z^d$:

$$\text{shift}(\xi, x)(y) = \xi(y - x).$$

Recalling the definition of C_{lip}, on the event A, for any $i = 1, \cdots, d$,

$$\bar{\xi}_{t_{7.4}} \geq \text{shift}(\bar{\xi}_0, \pm 2\ell e_i),$$

Monotonicity enables us to restart the construction above from anyone among the shifts. Note that the success probability of such a construction is of $1 - O(L^{-d})$. So when L is large, with high probability we can do it for $2d \log L$ times without a failure. This will give us a "copy" of $\bar{\xi}_0$ at $\pm 2\ell(\log L)e_i$ for each i and will take time $T = (\log L)t_{7.4}$.

Thus we can have out block construction as follows: let

$$\Gamma_x = 2\ell(\log L)x + [-\ell \log L, \ell \log L]^d, \ \forall x \in Z^d$$

and

$$T_n = nT, n \geq 0.$$

We say (x, n) is wet if

$$\bar{\xi}_{T_n} \geq \text{shift}(\bar{\xi}_0, 2\ell(\log L)x).$$

From the construction above, one can immediately see that (x, n) is wet then with high probability $(x \pm e_i, n + 1)$ are all wet for $i = 1, \cdots, d$.

To check that with high-probability, the block events are finite-dependent, note that $\bar{\xi}_t$ is dominated by a branching random walk with birth rate β and initial configuration $\bar{\xi}_0$. Lemma 4.1 shows that for any $m > 0$

$$P(M_k(T_1) \geq (2\beta + m)LT_1) \leq e^{-mT_1}|B(0, (M+1)L)|$$

when L is large enough, where $M_k(t)$ is the largest kth coordinate among the accepted sites at time t. Noting that $T_1 = (\log L)t_{7.4}$ and that the choice of $t_{7.4}$ is independent to the choice of L, let

(61) $$m = (d+1)/t_{7.4}.$$

We can control the probability that $\bar{\xi}_t$ wanders too far as follows:

$$P_{\bar{\xi}_0}\left(\max_{t \in [0,T_0]}\{\|x\| : \bar{\xi}_t(x) \neq 0\} \geq (2\beta + m)LT_1\right) \leq 2de^{-mT_1}|B(0, (M+1)L)|.$$

Noting that

$$e^{-mT_1} = e^{-(\log L)t_{7.4}(d+1)/t_{7.4}} = e^{-(d+1)\log L} = L^{-d-1}$$

we have
(62)
$$P_{\bar{\xi}_0}\left(\max_{t \in [0,T_0]}\{\|x\| : \bar{\xi}_t(x) \neq 0\} \geq (2\beta + m)LT_1\right) = L^{-d-1}O(L^d) = O(L^{-1}) \to 0$$

as $L \to \infty$. Thus noting that $\ell = \epsilon_{7.4}L$, let

$$R = \frac{(2\beta + m)LT_1}{2\ell \log L} = \frac{(2\beta + m)t_{7.4}}{2\epsilon_{7.4}}$$

which is a finite constant independent to the choice of L. As L goes large, we have that with high probability $\bar{\xi}_t$ cannot exit the following finite union of blocks by time T_1:

$$\bigcup_{x:\|x\|\leq R} \Gamma_x$$

which implies that the block events we constructed has finite range of dependence. Then again according to standard block argument in [2] and [11], we complete the proof of survival for $\bar{\xi}_t$ and this implies Theorem 3.

REFERENCES

[1] BESSONOV, M. and DURRETT, R. (2013). Phase transitions for a planar quadratic contact process. arXiv:1312.353.

[2] 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.

[3] DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation.* Wadsworth, Inc., California.

[4] DURRETT, R. (2007). *Random Graph Dynamics.* Cambridge U. Press, New York.

[5] DURRETT, R. and LANCHIER, N. (2008). Coexistence in host-pathogen systems. *Stoch. Proc. Appl.* **118** 1004–1021.

[6] DURRETT, R. and LEVIN, S. (1998). *Theor Popul Biol.* **53** 30–43.

[7] DURRETT, R. and NEUHAUSER, C. (1991). Epidemics with recovery in d = 2. *Ann. Appl. Probab.* **1** 189–206.

[8] DURRETT, R. and SWINDLE, G. (1991). Are there bushes in a forest? *Stoch. Proc. Appl.* **37** 19–31.

[9] DURRETT, R. and SWINDLE, G. (1994). Coexistence results for catalysts. *Probab. Th. Rel. Fields* **98** 489–515.

[10] GRANNAN, E. R. and SWINDLE, G. (1990). Rigorous result on mathematical models of catalytic surfaces. *J. Stat. Phys.* **61** 1085–1103.

[11] HARRIS, T. (1972). Nearest neighbor Markov interaction processes on multidimensional lattices. *Adv. Math.* **9** 66–89.

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

[13] LIGGETT, T. M. (1985). *Interacting Particle Systems.* Springer-Verlag, New York.

[14] MOUNTFORD, T. S. and SUDBURY, A. (1992). An extension of a result of Grannan and Swindle on the poisoning of catalytic surface. *J. Stat. Phys.* **67** 1219–1222.

[15] NEUHAUSER, C. (1994). A long range sexual reproduction process. *Stoch. Proc. Appl.* **53** 193–220.

[16] SCHERTZER, E., STAVER, A. C. and LEVIN S. (2014). Implications of the spatial dynamics of fire spread for the bistability of savanna and forest. Accepted *J Math Biol.*

[17] 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.

[18] 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.

[19] STAVER, A. C. and LEVIN S. (2012). Integrating Theoretical Climate and Fire Effects on Savanna and Forest Systems. *Amer. Nat.* **180** 211–224.

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

[21] WEINBERGER, H. F. (1982). Long-Time Behavior of a Class of Biological Models. *SIAM J. Math. Anal.* **13** 353–396.

DEPARTMENT OF MATHEMATICS, BOX 90320
DUKE U., DURHAM, NC 27708-0320
E-MAIL: rtd@math.duke.edu
         yzhang@math.duke.edu