

## COEXISTENCE RESULTS FOR SOME COMPETITION MODELS

BY RICK DURRETT<sup>1</sup> AND CLAUDIA NEUHAUSER<sup>2,3</sup>

*Cornell University and University of Minnesota*

Barley yellow dwarf is a widespread disease that affects small grains and many grass species, as well as wheat, barley and oat. The disease is caused by an aphid transmitted virus. Rochow conducted a study near Ithaca, New York, which showed that a shift in the dominant strain occurred between 1957 and 1976. Motivated by this phenomenon, we develop a model for the competition between different strains of the barley yellow dwarf virus. Our main goal is to understand the phase diagram of the model, that is, to identify parameter values where one strain competitively excludes the other strain and where both strains coexist. Our analysis applies to a number of other systems as well, for example to a model of competition of water flea species studied by Hanski and Ranta and Bengtsson.

**1. Introduction.** This paper began as an attempt to understand the competition of different strains of the barley yellow dwarf virus. Barley yellow dwarf (BYD) is a serious widespread disease of small grains and grasses caused by a group of aphid transmitted viruses. Its symptoms are chlorosis (i.e., yellowing of plant tissue) and stunting of the affected plant. BYD is an important agricultural disease since it affects large numbers of different grains throughout the world. The total yield loss in the United States is around 1 to 3 percent each year, but under favorable conditions, losses of 40% are not uncommon. The disease was first reported in the United States by Galloway and Southwood (1890) but only much later recognized by Oswald and Houston (1951) as being caused by a virus.

The barley yellow dwarf virus (BYDV) is a member of the luteoviruses. This group includes bean leaf roll, beet western yellows, carrot red leaf and potato leaf roll. These viruses are transmitted by aphids (the “vector” for the disease) and they are typically very host-specific [see, e.g., Duffus, Falk, and Johnstone (1987)].

---

Received April 1996; revised October 1996.

<sup>1</sup> Research supported in part by NSF Grants DMS-93-01070 and BIR-94-23339.

<sup>2</sup> Alfred P. Sloan Research Fellow. Research supported in part by NSF Grant DMS-94-03644. This author also acknowledges support from the Department of Ecology and Evolutionary Biology, Princeton University, and the Institute for Advanced Study, Princeton, during 1995/1996.

<sup>3</sup> Also at University of Wisconsin, Madison.

AMS 1991 *subject classifications*. Primary 60K35; secondary 92B05.

*Key words and phrases*. Interacting particle systems, contact process, successional model, rescaling, coexistence, competitive exclusion, barley yellow dwarf, *Daphnia*.

Five distinct BYDV strains have been identified. They can be classified using enzyme linked immunoabsorbent assay (ELISA). The virus strains are highly vector specific; that is, different strains of the virus can only be transmitted by specific aphids (although one aphid may be able to transmit more than one virus and one virus may be transmitted by more than one aphid). Table 1, which appeared in Gildow (1987), indicates by +'s the strains that each aphid can transmit. To make the table easier to read, we have indicated the genus names *Macrosiphum*, *Rhopalosiphum*, and *Sitobion* by their first letters.

To motivate the definition of our model, we will now briefly describe some aspects of the transmission of BYD. See, for instance, Gildow (1987). Starting with infected plant tissue, aphids may ingest virus particles suspended in phloem sap during feeding. (Phloem is the part of the vascular tissue which is responsible for transporting substances produced in the plant's metabolism.) Before the aphid can transmit the virus to other plants, the virus particles need to be transported through the body of the aphid to the salivary glands. This part of the acquisition process is responsible for the fact that each aphid species can transmit only a few virus strains. Once they have acquired the virus, aphids can infect the plant during feeding.

Mixed infections with luteoviruses are common in the field [see, e.g., Waterhouse, Gildow and Johnstone (1988), Rochow (1965), Falk and Duffus (1981)] but there seems to be some evidence that cross-protection among different strains of BYDV is possible [see Wen, Lister and Fatou (1991)]. In this paper we will, for simplicity, concentrate on the interaction of MAV and PAV, the two most prevalent strains in New York State. Figure 1 is a graph of data collected by Rochow (1979) near Ithaca. It shows that between 1957 and 1976 the dominant strain shifted from MAV to PAV. This graph motivates our main question: can the two strains coexist in equilibrium or will one always competitively exclude the other?

As for the shift in the dominant strain, Rochow (1979) concluded that it was not caused by a change of the predominant aphid species. He cited various factors, such as changes in cultivars and acreages of small grains, but no definite conclusion was reached. The results below will show that in our model a region in parameter space exists where the two strains coexist but

TABLE 1  
*Vector transmission patterns*

Aphid species	MAV	PAV	RMV	RPV	SGV
<i>M. dirhodum</i>	+	+	—	—	—
<i>R. maidis</i>	—	—	+	—	—
<i>R. padi</i>	—	+	—	+	—
<i>R. rufiabdominalis</i>	—	+	+	+	—
<i>S. graminum</i>	—	+	—	+	+
<i>S. avenae</i>	+	+	—	—	—

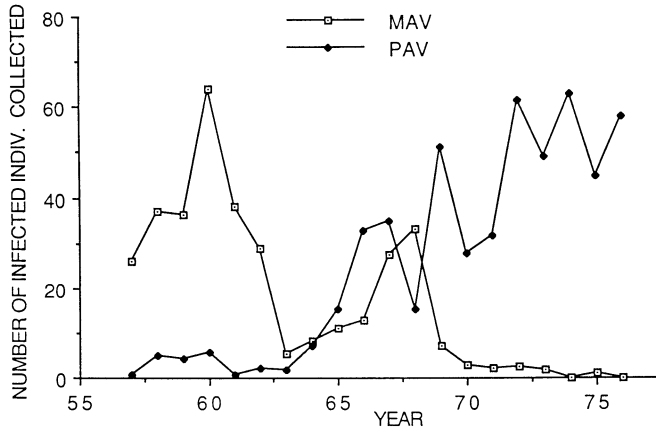


FIG. 1. Shift of dominance from MAV to PAV according to data collected by Rochow (1979).

give no quantitative information about the size of the region. If the region is narrow, then a shift in the cultivated species could shift the system from one exclusion region to the other.

To formulate the dynamics we let  $\mathcal{N} \subset \mathbf{Z}^2$  define the neighbors of 0, let  $x + \mathcal{N}$  be the neighbors of  $x$  and let  $f_i$  the the fraction of neighbors in state  $i$ . Since most of the infection is spread by apterous (i.e., wingless) aphids, it is natural to choose  $\mathcal{N}$  to be the four nearest neighbors or some other small neighborhood. However, for our last two results we will have to consider long range interactions. To accommodate all the choices that appear in our paper and to rule out trivialities (such as absence of irreducibility) we will assume throughout that  $\mathcal{N} = \{x: \|x\| \leq r\}$  where  $\|x\|$  is some norm on  $\mathbf{R}^2$ , for example,  $\|x\|_1 = |x_1| + |x_2|$ ,  $\|x\|_2 = (x_1^2 + x_2^2)^{1/2}$  or  $\|x\|_\infty = \max\{|x_1|, |x_2|\}$ . The model is a continuous time Markov process whose state at time  $t$  is a function  $\xi_t: \mathbf{Z}^2 \rightarrow \{0, 1, 2, 3\}$ . We interpret this as follows: if  $\xi_t(x) = 0$ , we say  $x$  is vacant; if  $\xi_t(x) = 1$ ,  $x$  is infected with strain 1; if  $\xi_t(x) = 2$ ,  $x$  is infected with strain 2; if  $\xi_t(x) = 3$ ,  $x$  is infected with both strains.

The notation we have just introduced allows us to write down all the transition rates of the model:

$$\begin{array}{ll}
 0 \rightarrow 1 & \beta_1(f_1 + r_{11}f_3) & 1 \rightarrow 0 & \delta_1 \\
 0 \rightarrow 2 & \beta_2(f_2 + r_{21}f_3) & 2 \rightarrow 0 & \delta_2 \\
 2 \rightarrow 3 & r_{12}\beta_1(f_1 + r_{11}f_3) & 3 \rightarrow 2 & \delta_1/r_{13} \\
 1 \rightarrow 3 & r_{22}\beta_2(f_2 + r_{21}f_3) & 3 \rightarrow 1 & \delta_2/r_{23}.
 \end{array}$$

The overall structure of this model is like the contact process: infection rates are proportional to the number of infected neighbors, while recovery rates are constant. The parameter  $r_{ij}$  describes the degree of interspecific competition. Here, we will choose  $r_{ij} \in [0, 1]$  and interpret them as reduction coefficients.

In the first rate, that is,  $\beta_1(f_1 + r_{11}f_3)$ ,  $r_{11}$  reflects the fact that the amount of virus 1 in a doubly infected plant is less than in a singly infected plant, so the transmission rate is less. The second pair of rates is the first pair multiplied by  $r_{i2}$  to account for the fact that a singly infected plant's immune system is working and hence has a reduced rate of acquiring a second infection. The third pair of rates says that plants recover at a constant rate, while in the fourth we divided the third rates by  $r_{i3} \leq 1$  to account for the fact that the two infections compete within the host.

Our model can also be applied to a number of other competitive situations. One situation that has been extensively studied is the competition of water flea species (*Daphnia*) in rock pools in Scandinavia. See Hanski and Ranta (1983), Bengtsson (1991) and references therein. When considering the competition of two species, these authors assign each rock pool a state: 0 = vacant, 1, 2 = occupied by one species, 3 = occupied by both species, much as we have above. However, they ignore the spatial arrangements of the pools and postulate that migration between any two patches occurs at a constant rate. Taking a limit in which the number of pools tends to infinity, they arrive at a system of ordinary differential equations for the fraction of pools in various states, which was first studied by Slatkin (1974).

Figure 2 shows the study site of Hanski and Ranta (1983). Vacant pools are white. The species inhabiting the black occupied pools are indicated by M = *D. magma*, P = *D. pulex*, and L = *D. longispina*. The arrangement of pools does not seem to be consistent with the assumption that all migration rates are equal but instead seems to suggest the use of a one-dimensional

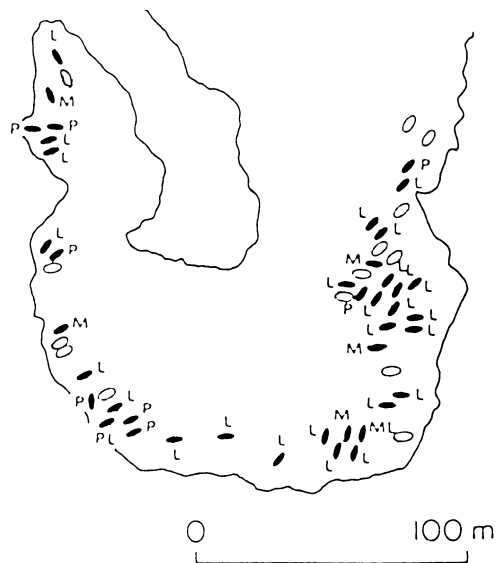


FIG. 2. Spatial arrangements of rock pools in the study of competition of *Daphnia*. [From Hanski and Ranta (1983).]

chain of sites. In their situation, however, this oversight does not have serious consequences since the use of a spatially explicit model does not change the qualitative predictions. Our analysis will show that when the range of interaction in the particle system is large, then the densities in the particle system behave like those in the corresponding ordinary differential equation; in particular, coexistence is possible. When the range is small there are correlations between the states of adjacent sites that shift the equilibrium densities in the particle system away from those predicted by the ODE. This changes the parameter values at which coexistence occurs but should not eliminate the possibility of coexistence. In making the last conclusion, we are relying on the physicists' notion that the qualitative properties of the model do not depend on the neighborhood chosen.

Turning to the analysis of the model, our first step is to get rid of some of its 10 parameters. Our first simplification is to equate all the reduction coefficients for a given species; that is, we set  $r_{ij} = c_i$  for  $j = 1, 2, 3$ . This eliminates the *Daphnia* model as a special case but, as we will explain later, the coexistence result (Theorem 4) extends in a straightforward way to the general model. Our second simplification is to set  $\delta_1 = \delta_2$  and, without further loss of generality, to set the common value equal to 1. The qualitative behavior of the system should not be different when  $\delta_1 \neq \delta_2$ , but the equality of  $\delta$ 's is needed in the proof of Theorem 2 (and only there).

Our final modification is simply a change of perspective. In what follows, we will consider  $\beta_1$ ,  $\beta_2$  and  $\delta_1 = \delta_2 = 1$  fixed and vary the competition coefficients  $c_1$  and  $c_2$ . When  $c_1 = c_2 = 1$ , neither disease feels the presence of the other. So if each disease has a sufficiently high transmission rate to survive on its own, then there will be coexistence. To explain the qualifying phrase in the previous sentence, note that if only one disease is present, the system reduces to the contact process with neighborhood set  $\mathcal{N}$ . In this case there is a critical value  $\beta_c > 1$ , which depends on  $\mathcal{N}$ , so that if  $\beta \leq \beta_c$ , the infection dies out, but if  $\beta > \beta_c$  there is a unique nontrivial stationary distribution that is translation invariant. For more on the contact process, see Liggett (1985), Durrett (1988) or Durrett and Levin (1994a) and references therein.

To state our first result we need one more definition: we say there is coexistence if there is a stationary distribution which concentrates on configurations with infinitely many sites in each state. In all cases below when we prove coexistence, we construct a translation invariant stationary distribution.

**THEOREM 1.** *If  $\beta_1 c_1^3, \beta_2 c_2^3 > \beta_c$ , then there is coexistence.*

To explain the  $c_i^3$ , note that if we pretend that all sites are always occupied by 2's, then we have a system with only two states 2 and 3, which makes transitions  $2 \rightarrow 3$  at rate  $\beta_1 c_1^2 f_3$  and  $3 \rightarrow 2$  at rate  $1/c_1$ . This is a time change of the basic contact process which survives if  $\beta_1 c_1^3 > \beta_c$ . To get from

the last observation to the conclusion of Theorem 1 we need a comparison result.

**PROPOSITION 1.1.** *Let  $\xi_t$  denote the process with parameters  $(\beta_1, \beta_2, \delta_1, \delta_2, c_1, c_2)$  and  $\xi'_t$  denote the process with  $(\beta'_1, \beta'_2, \delta'_1, \delta'_2, c'_1, c'_2)$ . Suppose  $\beta_1 \leq \beta'_1$ ,  $\delta_1 \geq \delta'_1$ ,  $c_1 \leq c'_1$ ,  $\beta_2 \geq \beta'_2$ ,  $\delta_2 \leq \delta'_2$ , and  $c_2 \geq c'_2$ . Let  $\eta_t = \{x: \xi_t(x) = 1 \text{ or } 3\}$ ,  $\zeta_t = \{x: \xi_t(x) = 2 \text{ or } 3\}$  and define  $\eta'_t$  and  $\zeta'_t$  accordingly in terms of  $\xi'_t$ . If  $\eta_0 \subset \eta'_0$  and  $\zeta_0 \supset \zeta'_0$ , then we can construct processes  $(\eta_t, \zeta_t)$  and  $(\eta'_t, \zeta'_t)$  on the same probability space with the two parameter sets so that  $\eta_t \subseteq \eta'_t$  and  $\zeta_t \supset \zeta'_t$  for all  $t \geq 0$ .*

If we let  $\zeta_0 = \mathbf{Z}^2$  and  $\delta_2 = 0$  while leaving the other aspects of the process unchanged, we get the comparison for Theorem 1.

Proposition 1.1 also allows us to deduce properties of the phase diagram of the model, but first we need a definition. Given an initial distribution  $\xi_0$  we say that type  $i$  dies out if  $P(\xi_t(x) = i) \rightarrow 0$  as  $t \rightarrow \infty$  for all  $x$ . If we take  $\beta_i = \beta'_i$  and  $\delta_i = \delta'_i$ , then we see that if the 1's die out for some  $(c_1, c_2)$ , then they die out when  $c'_1 \leq c_1$  and  $c'_2 \geq c_2$ .

From the corner  $c_1 = c_2 = 1$ , we jump now to the corner  $c_1 = c_2 = 0$ . In this case transitions into state 3 have rate 0 and transitions out of state 3 have rate  $\infty$ , so the set of possible states reduces to  $\{0, 1, 2\}$  and the model reduces to the multitype contact process of Neuhauser (1992) which makes transitions as follows:

$$\begin{array}{ll} 0 \rightarrow 1 & \beta_1 f_1 \quad 1 \rightarrow 0 \quad 1 \\ 0 \rightarrow 2 & \beta_2 f_2 \quad 2 \rightarrow 0 \quad 1. \end{array}$$

In this case her results show that if  $\beta_1 > \beta_2$ , then the 2's die out for any translation invariant initial distribution that concentrates on configurations with infinitely many 1's. Our next result extends this conclusion to our new system and strengthens the conclusion.

**THEOREM 2.** *Suppose  $\beta_1 > \beta_2$ . There is an  $\varepsilon > 0$  so that if  $0 \leq c_2 \leq \varepsilon$  and  $0 \leq c_1 \leq 1$ , then the 2's die out starting from any initial distribution that concentrates on configurations with infinitely many 1's.*

Theorem 2 also covers the corner  $c_1 = 1, c_2 = 0$  so we turn now to the remaining one:  $c_2 = 1, c_1 = 0$ . When  $c_1 = 0$ ,  $2 \rightarrow 3$  transitions are impossible and  $3 \rightarrow 2$  transitions occur at rate  $\infty$  so again state 3 disappears but this time the rates are

$$\begin{array}{ll} 0 \rightarrow 1 & \beta_1 f_1 \quad 1 \rightarrow 0 \quad 1 \\ 0 \rightarrow 2 & \beta_2 f_2 \quad 2 \rightarrow 0 \quad 1 \\ 1 \rightarrow 2 & c_2 \beta_2 f_2. \end{array}$$

When  $c_2 = 1$ , this further reduces to the grass–bushes–trees system studied by Durrett and Swindle (1991). In this case, the 2's do not feel the presence of the 1's and are a contact process, so they will survive if  $\beta_2 > \beta_c$ . Durrett and Swindle (1991) showed that if  $\beta_1 > \beta_2^2 > 1$  and the range of interaction is large, then there is coexistence. Our next result generalizes the more difficult converse first proved by Durrett and Schinazi (1993). For concreteness we suppose that the norm  $\|x\|_\infty = \max\{|x_1|, |x_2|\}$  is used to define the neighborhood. However, the choice of norm is not important and the conclusion can be generalized to dispersal distributions of the form  $c_r \phi(|y - x|/r)$  with  $r$  large provided  $\phi(\cdot)$  has exponential tails.

**THEOREM 3.** *Suppose  $\beta_1 < \beta_2^2$ . If the range of the interaction  $r \geq r_0$  and  $\varepsilon \leq \varepsilon_0(r)$ , then for  $c_1 \leq \varepsilon$  and  $c_2 \geq 1 - \varepsilon$  the 1's die out.*

**PROOF.** This result is a corollary of the block argument in Durrett and Schinazi (1993). That paper constructs a block event for the case  $c_1 = 0$  and  $c_2 = 1$ . Once we have an event in a specified box which guarantees survival when its probability is large, it is immediate that the same conclusion holds when  $c_1 \leq \varepsilon$  and  $c_2 \geq 1 - \varepsilon$ .  $\square$

The result is almost certainly true with an  $\varepsilon_0$  independent of  $r$ , but this version of the result is enough to show that the third possibility, 1's die out, occurs in the phase diagram. See Figure 3 for a sketch of what we have proved. The monotonicity of the boundary curves follows from the remark after Proposition 1.1.

When one relies on large range limits to prove the existence of a phenomenon in a particle system, it is natural to try to find the exact limiting behavior of the phase diagram. The first step in doing this is to look at the system through the eyes of mean field theory, that is, to pretend that all sites are independent and see how the densities evolve. Letting  $u_i$  be the fraction of sites in state  $i$ , this leads to the following system of ordinary differential equations:

$$\begin{aligned}
 \frac{du_1}{dt} &= \beta_1 u_0 u_1 + c_1 \beta_1 u_0 u_3 - c_2 \beta_2 u_1 u_2 \\
 &\quad - c_2^2 \beta_2 u_1 u_3 - u_1 \delta_1 + u_3 \delta_2 / c_2, \\
 \frac{du_2}{dt} &= \beta_2 u_0 u_2 + c_2 \beta_2 u_0 u_3 \\
 (1.1) \quad &\quad - c_1 \beta_1 u_1 u_2 - c_1^2 \beta_1 u_2 u_3 - u_2 \delta_2 + u_3 \delta_1 / c_1, \\
 \frac{du_3}{dt} &= (c_1 \beta_1 + c_2 \beta_2) u_1 u_2 + c_1^2 \beta_1 u_2 u_3 \\
 &\quad + c_2^2 \beta_2 u_1 u_3 - u_3 (\delta_1 / c_1 + \delta_2 / c_2).
 \end{aligned}$$

Let  $\Gamma = \{(u_1, u_2, u_3): u_i \geq 0, u_1 + u_2 + u_3 \leq 1\}$  be the collection of values we are interested in. Intuitively, if the ODE has a fixed point in  $\Gamma^\circ$ , the interior of  $\Gamma$ , and that fixed point is the limit from any initial condition in  $\Gamma^\circ$ ,

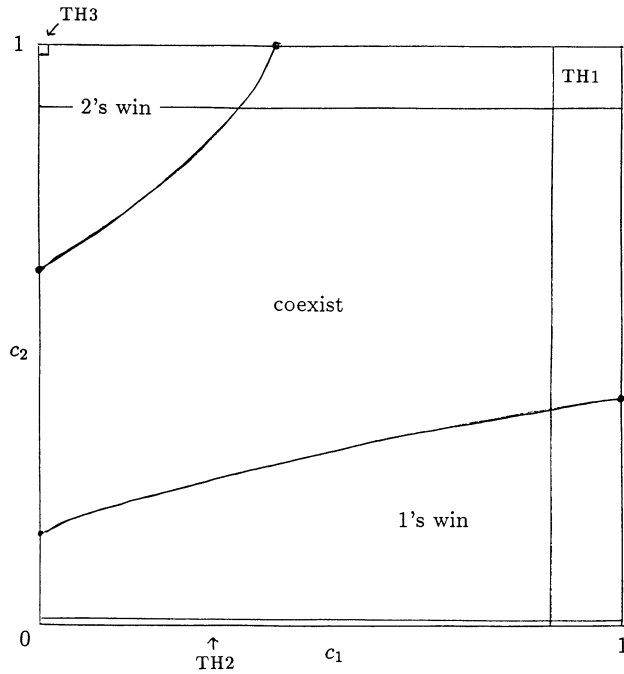


FIG. 3. Summary of theorems and phase diagram.

we expect coexistence to occur in the particle system when the range is large; see Durrett and Levin (1994b). Durrett (1992b) used this idea in the study of a predator-prey system with rapid stirring. In that case it was possible to calculate explicitly the location of the equilibrium for the mean field ODE and prove that it was globally attracting, but neither problem seems tractable for (1.1).

To circumvent this difficulty, we will use an approach that is commonly used in the biology literature to determine if two species can coexist; namely, two species coexist if either species can invade the other in its equilibrium state. First, consider the case in which strain 1 is absent. In this case  $u_0 = 1 - u_2$  and

$$\frac{du_2}{dt} = \beta_2 u_2 (1 - u_2) - \delta_2 u_2.$$

So if  $\beta_2 > \delta_2$ , the density of 2's equilibrates at  $(\beta_2 - \delta_2)/\beta_2$ . Suppose now that the first strain is present at a small density [i.e.,  $u_1(t) + u_3(t)$  is small]. Using the previous conclusion now and continuity of the solution as a function of the starting point, which is true whenever the right-hand side is a Lipschitz continuous function of  $(u_1, u_2, u_3)$ , we see that after a fixed amount of time we will come to  $u_2 \approx (\beta_2 - \delta_2)/\beta_2$ ,  $u_0 \approx \delta_2/\beta_2$ .

This motivates the following definition: we say that “1's can invade 2's when they are in equilibrium” if the boundary equilibrium  $u_1 = 0$ ,  $u_2 =$



$(\beta_2 - \delta_2)/\beta_2$ ,  $u_3 = 0$  is unstable. To be precise, the linearization of the ordinary differential equation at this point has an unstable direction that points into  $\Gamma^\circ$ . The next result says this condition prevents the 1's from dying out in the particle system.

**THEOREM 4.** *Suppose that  $c_1 > 0$  and that 1's can invade 2's when they are in equilibrium. Consider the particle system starting from a product measure in which state 1 is present with positive probability. If the range  $r \geq r_0$ , then any subsequential limit concentrates on configurations in which 1's have density  $\varepsilon(r) > 0$ .*

When  $c_1 = 0$ , the rate for  $3 \rightarrow 2$  transitions becomes infinite and the state space of the model changes. A number of modifications are needed in the argument but the general approach does not change and the conclusion remains the same.

One can interchange the roles of 1's and 2's in Theorem 4 to get a result for the persistence of 2's and then combine the two results to get conditions for coexistence. We believe that the resulting conditions are asymptotically sharp as  $r \rightarrow \infty$ . That is we

**CONJECTURE.** *In (1.1) if 1's cannot invade 2's when they are in equilibrium, then 1's die out when the range is large.*

Proving the conjecture in general seems to be a difficult problem. The first difficulty here is to show that if 1's cannot invade 2's when they are in equilibrium (which means that all eigenvalues of the linearization around the fixed point are nonpositive), then all trajectories that start in  $\Gamma^\circ$  converge to the fixed point  $u_1 = 0$ ,  $u_2 = (\beta_2 - \delta_2)/\beta_2$ ,  $u_3 = 0$ . Even if this is established, there is the second problem of showing that this behavior of the ordinary differential equation implies that the 1's die out for large range.

For the rest of the section we will concentrate on what happens on the sides of the square. This will give insight into the structure of the phase diagram.

*The left side*,  $c_1 = 0$ . As noted after the statement of Theorem 2, state 3 disappears in this case. Eliminating the terms that include  $u_3$  or  $c_1$ , and adding a new term to the  $u_2$  equation to account for the transitions from  $1 \rightarrow 3$  and then instantaneously from  $3 \rightarrow 2$ , the mean field equation simplifies to

$$(1.2) \quad \begin{aligned} \frac{du_1}{dt} &= \beta_1 u_0 u_1 - \delta_1 u_1 - c_2 \beta_2 u_1 u_2, \\ \frac{du_2}{dt} &= \beta_2 u_0 u_2 - \delta_2 u_2 + c_2 \beta_2 u_1 u_2. \end{aligned}$$

In this case we can factor  $u_i$  out of the  $i$ th equation and substitute  $u_0 = 1 - u_1 - u_2$  to get

$$(1.3) \quad \begin{aligned} \frac{du_1}{dt} &= u_1\{(\beta_1 - \delta_1) - \beta_1 u_1 - (c_2 \beta_2 + \beta_1) u_2\}, \\ \frac{du_2}{dt} &= u_2\{(\beta_2 - \delta_2) - \beta_2(1 - c_2) u_1 - \beta_2 u_2\}. \end{aligned}$$

In this case the equilibria satisfy two linear equations in two unknowns, so one can solve. Taking  $\delta_1 = \delta_2 = 1$  to simplify the answer we have

$$(1.4) \quad \begin{aligned} \bar{u}_1 &= \frac{(\beta_1 - \beta_2) - c_2 \beta_2 (\beta_2 - 1)}{\beta_2 c_2 (\beta_1 - \beta_2 + \beta_2 c_2)}, \\ \bar{u}_2 &= \frac{c_2 \beta_2 (\beta_1 - 1) - (\beta_1 - \beta_2)}{\beta_2 c_2 (\beta_1 - \beta_2 + \beta_2 c_2)}. \end{aligned}$$

To see what these equations say, we will suppose  $\beta_2 < \beta_1 < \beta_2^2$  so that there are two phase transitions on the left edge. In this case by examining the signs of the numerators we see that if  $r = \beta_1/\beta_2$ ,  $a = (r - 1)/(\beta_1 - 1)$  and  $b = (r - 1)/(\beta_2 - 1)$  then

for $c_2$ in	equilibrium density of 1	equilibrium density of 2
$[0, a]$	$(\beta_1 - 1)/\beta_1$	0
$[a, b]$	$\bar{u}_1$	$\bar{u}_2$
$[b, 1]$	0	$(\beta_2 - 1)/\beta_2$ .

We do not know how to prove that the densities in the table are the limiting equilibrium densities as the range  $r \rightarrow \infty$ . However simulation results suggest that this is true and that  $r$  does not have to be very large for mean field theory to provide a good approximation. Figure 4 gives the results of simulations for the  $L^\infty$  neighborhoods with radius 2 ( $\times$ 's) and radius 4 ( $\diamond$ 's) when  $\beta_1 = 3$  and  $\beta_2 = 2$  and compares them with the predictions of mean field theory given above (solid line). When  $r = 4$ , the equilibrium densities for the particle system are quite close to those predicted. Of course, range 4 does mean that there are 80 sites in the neighborhood.

*The top edge*,  $c_2 = 1$ , should be easy since  $c_2 = 1$  means that the 2's do not notice that the 1's are present, and thus the ODE equilibrium satisfies

$$\bar{u}_2 + \bar{u}_3 = (\beta_2 - \delta_2)/\beta_2 \quad \text{and} \quad \bar{u}_0 + \bar{u}_1 = \delta_2/\beta_2.$$

However, this does not seem to be enough information to solve for the equilibrium densities, so we simplify further by linearizing around the fixed point  $u_1 = 0$ ,  $u_2 = (\beta_2 - \delta_2)/\beta_2$ ,  $u_3 = 0$ . If we suppose  $u_1$  and  $u_3$  are small, then  $u_0 \approx \delta_2/\beta_2$ ,  $u_2 \approx (\beta_2 - \delta_2)/\beta_2$ , and terms involving  $u_1 u_3$  can be ig-

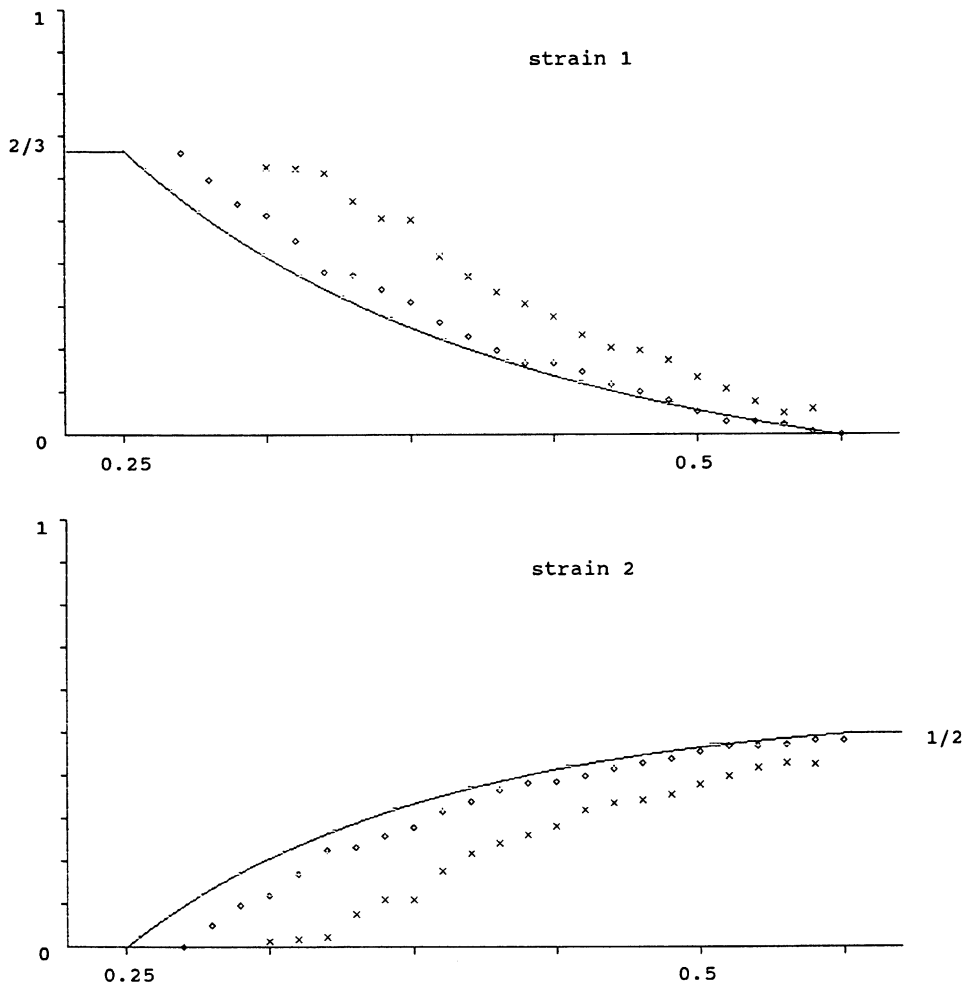


FIG. 4. Simulation results for the  $L^\infty$  neighborhood with radius 2 ( $\times$ 's) and radius 4 ( $\diamond$ 's) when  $\beta = 3$  and  $\beta_2 = 2$  and the prediction from mean field theory (solid line).

nored. When these simplifications are made, we get the following ODE, which we have written for a general value of  $c_2$ :

$$\begin{aligned}
 \frac{du_1}{dt} &= u_1 \left( \beta_1 \frac{\delta_2}{\beta_2} - c_2 \beta_2 \frac{\beta_2 - \delta_2}{\beta_2} - \delta_1 \right) \\
 &\quad + u_3 \left( c_1 \beta_1 \frac{\delta_2}{\beta_2} + \frac{\delta_2}{c_2} \right), \\
 \frac{du_3}{dt} &= u_1 \left( (c_1 \beta_1 + c_2 \beta_2) \frac{\beta_2 - \delta_2}{\beta_2} \right) \\
 &\quad + u_3 \left( c_1^2 \beta_1 \frac{\beta_2 - \delta_2}{\beta_2} - \frac{\delta_1}{c_1} - \frac{\delta_2}{c_2} \right).
 \end{aligned}
 \tag{1.5}$$

By recalling that for a  $2 \times 2$  matrix  $A$  with eigenvalues  $\lambda_1$  and  $\lambda_2$ , we have  $\det(A) = \lambda_1 \lambda_2$  and  $\text{trace}(A) = \lambda_1 + \lambda_2$ , it is easy to compute numerically, for a given example when the maximum eigenvalue is positive, the condition for the 1's to invade 2's. In the case  $c_2 = \delta_1 = \delta_2 = 1$ , we are able to obtain a condition for the 1's to invade 2's in terms of the coefficients, namely,

$$(1.6) \quad \beta_1 > \frac{\beta_2(c_1 + \beta_2)}{1 + c_1 + c_1^2(\beta_2 - 1)(2 + c_1\beta_2)}.$$

We leave it up to the reader to decide if (1.6) is intuitive or not. Plotting the right-hand side of (1.6) in the  $(c_1, \beta_1)$  plane for fixed  $\beta_2$  yields a monotonically decreasing graph (as a function of  $c_1$ ) through  $(0, \beta_2^2)$  and  $(1, 1)$ . See Figure 5. For points  $(c_1, \beta_1)$  above this curve, 1's and 2's coexist; below this curve, 2's competitively exclude 1's. In particular, if we fix  $\beta_1$  and vary  $c_1$ , there is a critical value of  $c_1^*$  so that for  $c_1 > c_1^*$ , 1's and 2's coexist, whereas for  $c_1 < c_1^*$ , 2's competitively exclude 1's.

Though it is difficult to compute the values of  $\beta_1$  and  $c_1$  that allow coexistence, one thing is very simple on the top edge. Thus when  $c_2 = 1$  and the range is large, the 2's will be present with density close to  $(\beta_2 - 1)/\beta_2$ . Since the 2's do not feel the presence of the 1's, it should not be too hard to prove, using techniques of Durrett and Schinazi (1993), that the condition given in Theorem 4 is asymptotically the correct one for coexistence.

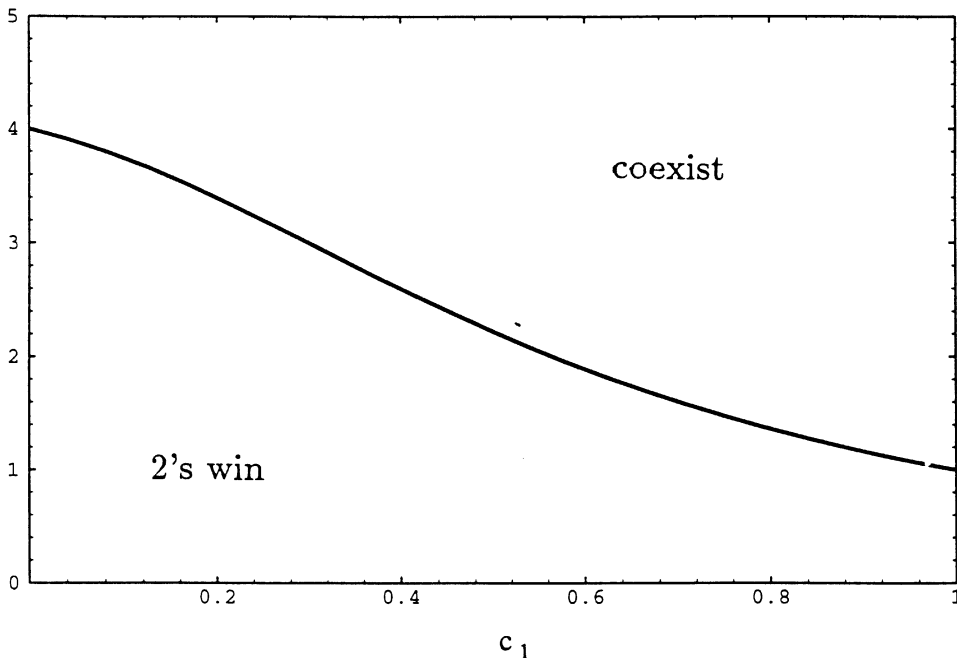


FIG. 5. The right-hand side of (1.6) in the  $c_1 - \beta_1$  plane when  $\beta_2 = 2$ .

The rest of the paper is devoted to proofs. Proofs of Proposition 1.1 and Theorem 1 are given in Section 2, Theorem 2 in Section 3 and Theorem 4 in Section 4. The authors would like to express their appreciation to Susi Remold, a graduate student in Ecology and Systematics at Cornell, for introducing us to this system. The observation that the range of the interaction does not have to be very large for mean field values to be accurate was suggested by simulations of Peter Calabrese, a participant in a Research Experience for Undergraduates held at Cornell in the summer of 1994.

**2. Proofs of Proposition 1.1 and Theorem 1.** As noted in the introduction, it suffices to prove Proposition 1.1. To do this, we need to construct the two processes on the same space. For  $x \in \mathbf{Z}^2$  and  $y \in x + \mathcal{N}$ , we introduce independent Poisson processes with the indicated rates:

process	$T_n^{1,x,y}$	$T_n^{2,x,y}$	$S_n^{1,x}$	$S_n^{2,x}$
rate	$\beta_1/ \mathcal{N} $	$\beta_2/ \mathcal{N} $	$\delta_1/c_1$	$\delta_2/c_2$

and independent Uniform  $(0, 1)$  random variables  $U_n^{1,x,y}, U_n^{2,x,y}, V_n^{1,x}, V_n^{2,x}$ . The  $T_n^{i,x,y}$  are times of a potential  $i$  infection from  $y$  to  $x$ , while the  $S_n^{j,x}$  are times when we potentially lose infection by  $j$  at  $x$ .

The uniform random variables are used to determine if the event should actually occur and hence thin the Poisson processes so that events are happening at the right rates. For example, if at time  $t = T_n^{1,x,y}$  we have  $\xi'_t(x) = 2$  and  $\xi'_t(y) = 1$ , then the infection will occur if  $U_n^{1,x,y} < c'_1$ , while if  $\xi_t(x) = 2$  and  $\xi_t(y) = 3$ , then the infection will occur if  $U_n^{1,x,y} < c_1^2 \beta_1 / \beta'_1$ . There are too many possible combinations for us to list formulas for every situation, but the rules should be clear from the example.

To prove Proposition 1.1 we have to argue that every flip preserves  $\eta_t \subset \eta'_t$  and  $\zeta_t \supset \zeta'_t$ . To do this we have to consider the four possible transitions. However, it is easy to see that when the inclusions hold, (i) the birth rate of 1's is higher in  $\xi'_t$ , (ii) the birth rate of 2's is higher in  $\xi_t$ , (iii) the death rate of 1's is higher in  $\xi_t$ , (iv) the death rate of 2's is higher in  $\xi'_t$ . This implies that every flip preserves the inclusions and the proof is complete.

**3. Proof of Theorem 2.** This section is devoted to proving Theorem 2, which describes the behavior of the process for  $c_2$  close to 0 and  $c_1 \in [0, 1]$ . As already explained in the introduction, the process with  $c_1 = c_2 = 0$  is known as the multitype contact process, which was investigated in Neuhauser (1992). Her results show that if  $c_1 = c_2 = 0$  and  $\beta_1 > \beta_2$ , then  $P(\xi_t(0) = 2) \rightarrow 0$  and  $t \rightarrow \infty$  for any translation invariant initial distribution that concentrates on configurations with infinitely many 1's. To prove Theorem 2, we need to strengthen and extend this result to a region near  $c_1 = c_2 = 0$ ; the comparison result (Proposition 1.1) then implies that the result holds for all  $c_1 \in [0, 1]$ . Our goal is to show that if  $\beta_1 > \beta_2$ , then there exists an in time

expanding “cone” which is void of 2’s and in which 1’s are in the corresponding contact process equilibrium for  $c_1$  and  $c_2$  close to 0.

The proof is based on a rescaling argument supplemented by a result in Durrett (1992a). The rescaling argument is by now a standard technique and has been applied frequently. [See, e.g., Bramson (1989), Bramson and Durrett (1988), Bramson and Neuhauser (1993), Durrett (1992a, b) and Durrett and Neuhauser (1994); the argument is reviewed in Durrett (1991, 1995).] The basic idea is to show that for given  $\delta > 0$ , members of the family of processes under consideration, when viewed on suitable length and time scales, dominate an oriented site percolation process in which sites are open with probability  $1 - \delta$ . (The sites may be  $j$ -dependent.) It is a well-known fact that oriented site percolation percolates for  $\delta$  close enough to 0. This almost produces our result. Our problem remains, namely, oriented site percolation has a positive density of unoccupied sites. To show that there exists an in-all-directions expanding region on which the processes have the desired properties, we therefore need to show that unoccupied sites within this region do not behave badly (they could, for instance, contain 2’s). To deal with this problem, we apply a result from Durrett (1992a) which shows that unoccupied sites do not percolate for  $\delta$  close enough to 0. Since particles of either type cannot appear spontaneously, once a region is void of one type, this type can only reappear in the region through invasion from the outside. This then implies that our processes have the desired properties. We refer the reader to Durrett (1992a, 1995) for details on the procedure.

To apply the rescaling argument, we need to strengthen Neuhauser’s results in two ways. First, we need to remove the condition of translation invariance of the initial configuration. Second, we need estimates on the rate of convergence of  $P(\xi_t(0) = 2)$  as  $t \rightarrow \infty$ . These estimates need to be good enough so that a perturbation argument can be applied to extend the results away from the corner.

To compare our process with oriented site percolation, to begin by introducing a grid and boxes in space:

$$\begin{aligned} \phi(z) &= (Lz_1, Lz_2), & A &= (-2L, 2L]^2, & A(z) &= \phi(z) + A, \\ & & B &= [-L, L]^2, & B(z) &= \phi(z) + B \end{aligned}$$

and similar definitions in space–time:

$$\begin{aligned} \psi(z, k) &= (Lz_1, Lz_2, kT), & C &= [-L - r, L + r]^2 \times [T - r, 2T), \\ & & C(z, k) &= \psi(z, k) + C \end{aligned}$$

for  $z = (z_1, z_2) \in \mathbf{Z}^2$  and  $L \in \mathbf{Z}^+$ . The squares  $A(z)$  are centered at  $\phi(z)$  and are the bases for boxes of height  $2T$  with  $T = L^2$ ; that is, the boxes are of the form  $A(z) \times [kT, (k + 2)T)$  for integer  $k$ , which contain  $C(z, k)$ ; here, we choose  $z_1$  and  $z_2$  both even for even  $k$ , and we choose  $z_1$  and  $z_2$  both odd for odd  $k$ . See Figure 6 for a slice through the rescaling boxes.

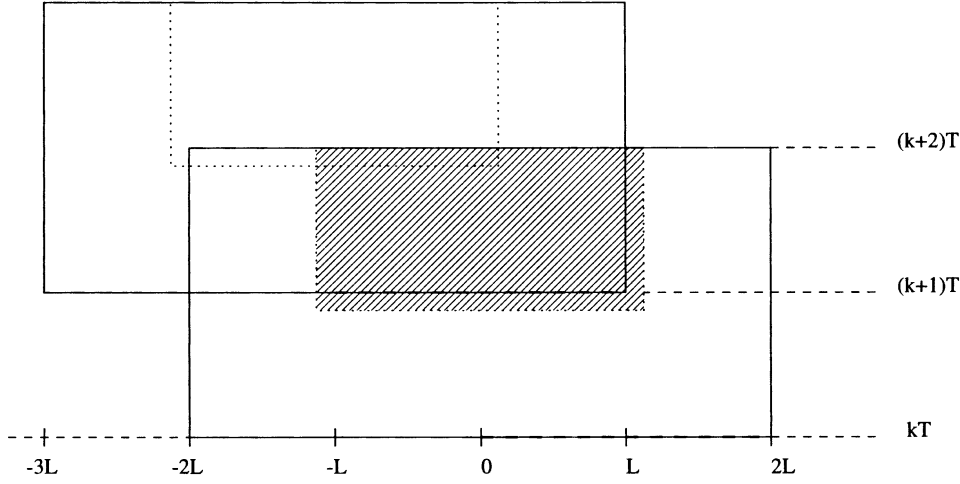


FIG. 6. A slice through the rescaling boxes for Theorem 2. The two big boxes are  $A(0, 0) \times [kT, (k+2)T)$  and  $A(-1, 0) \times [(k+1)T, (k+3)T)$ . The shaded box is  $C(0, 0, k)$ .

We partition  $B$  [and likewise,  $B(z)$ ] further. We tile  $B$  with  $L^{0.1} \times L^{0.1}$  squares, that is, we set

$$\begin{aligned} \pi(w) &= (L^{0.1}w_1, L^{0.1}w_2) \quad \text{for } w = (w_1, w_2) \in \mathbf{Z}^2 \\ D &= (-L^{0.1}/2, L^{0.1}/2]^2 \\ D(w) &= \pi(w) + D \\ I_z &= \{w : D(w) \subseteq B(z)\}. \end{aligned}$$

We are now ready to define certain “good events” in  $A(z) \times [kT, (k+2)T)$ . We say  $B(z)$  is  $G_1$  at time  $t$  if  $B(z)$  is void of 2's at time  $t$  and has at least one particle of type 1 in each of the squares  $D(w)$  with  $w \in I_z$ . We say that  $C(z, k)$  is  $G_2$  if the space-time box  $C(z, k)$  does not contain any 2's. For  $z = (z_1, z_2) \in \mathbf{Z}^2$  with  $z_1$  and  $z_2$  both even for even  $k$ , and  $z_1$  and  $z_2$  both odd for odd  $k$ , we say that  $(z, k)$  is *occupied* if  $B(z)$  is  $G_1$  at time  $(k+1)T$  and  $C(z, k)$  is  $G_2$ . The main goal of this section is to show the following two propositions. For each we suppose that:

(A)  $c_1 = c_2 = 0$ ,  $\beta_1 > \beta_2$ ,  $\beta_1 > \beta_c$  and  $B(z)$  is  $G_1$  at time 0 for some  $z \in \mathbf{Z}^2$  with  $\|z\|_\infty = 1$ .

PROPOSITION 3.1. Assume (A). If  $\delta > 0$ , then when  $L$  is large

$$(3.1) \quad P(B(0) \text{ is not } G_1 \text{ at time } T) \leq \delta/3.$$

PROPOSITION 3.2. Assume (A). If  $\delta > 0$ , then when  $L$  is large

$$(3.2) \quad P(C(0, 0) \text{ is not } G_2) \leq \delta/3.$$

The main ingredient in the proof of Proposition 3.1 is the following lemma which provides an estimate on the rate of convergence of  $P(\xi_t(x) = 2)$  to 0 as  $t \rightarrow \infty$ .

LEMMA 3.1. *Assume (A) and set  $T = L^2$ . There is a constant  $C > 0$  so that*

$$(3.3) \quad \max_{x \in B(0)} P(\xi_T(x) = 2) \leq CT^{-4}.$$

To compute  $P(\xi_t(x) = 2)$  when  $c_1 = c_2 = 0$ , we need to understand the ancestry of  $(x, t)$ . This is provided by the dual process of the multitype contact process, which was described in Neuhauser (1992). To define the dual we begin by constructing the process from a graphical representation. For  $x, y \in \mathbf{Z}^2$  with  $y - x \in \mathcal{N}$ , let  $\{T_n^{x,y}: n \geq 1\}$  and  $\{V_n^x: n \geq 1\}$  be the arrival times of Poisson processes with rates  $\beta_1/|\mathcal{N}|$  and 1, respectively. At times  $V_n^x$ , we put a  $\delta$  at  $x$ . The effect of a  $\delta$  is to kill the particle at  $x$  (if there is one). At times  $T_n^{x,y}$ , we draw an arrow from  $y$  to  $x$  to indicate that if  $x$  is vacant and  $y$  is occupied, a birth may occur. When  $y$  is occupied by a 1, a birth will always occur. However, since 2's are supposed to give birth at rate  $\beta_2 < \beta_1$ , we toss a coin with success probability  $(\beta_1 - \beta_2)/\beta_1$  at each arrow. If a success occurs, we label the arrow with a "1" and only allow 1's to give birth through these arrows. Since our birth rates have finite range, an idea of Harris (1972) then allows one to construct the process for all times starting from any initial configuration.

We say that there is a *path* from  $(x, s)$  to  $(y, t)$ ,  $0 \leq s \leq t$ , if there is a sequence of times  $s_0 = s < s_1 < s_2 \cdots < s_n < s_{n+1} = t$  and spatial locations  $x_0 = x, x_1, \dots, x_n = y$  so that the following hold:

1. For  $i = 1, 2, \dots, n$ , there is an arrow from  $x_{i-1}$  to  $x_i$  at time  $s_i$ .
2. The vertical segments  $\{x_i\} \times (s_i, s_{i+1})$ ,  $i = 0, 1, \dots, n$  do not contain any  $\delta$ 's.

Note that a path may contain both labelled and unlabelled arrows. Below we will need the notion of a *dual path*. We say that there is a dual path from  $(x, t)$  to  $(y, t - s)$ ,  $0 \leq s \leq t$ , if there is a path from  $(y, t - s)$  to  $(x, t)$ . That is, dual paths move against the direction of time.

We are now ready to define the *dual process*. For  $0 \leq s \leq t$ , we set

$$(3.4) \quad \hat{\xi}_s^{(x,t)} = \{y: \text{there is a dual path from } (x, t) \text{ to } (y, t - s)\}.$$

We call the elements of  $\hat{\xi}_s^{(x,t)}$  *ancestors*. If the initial configuration contains only 1's, then  $(x, t)$  is occupied by a 1 if and only if

$$\hat{\xi}_0^{(x,t)} \cap \{y: \xi_0(y) = 1\} \neq \emptyset.$$

If the initial configuration contains only 2's, the situation is similar, except that now the dual process may not use any 1-arrows in its paths. It is more complicated to determine the state of a site  $x$  at time  $t$  when the initial configuration contains both types.



It was shown in Neuhauser (1992) that the ancestors can be arranged according to the order they determine the state at  $(x, t)$ . To describe this hierarchy, assume for the moment that initially all sites are occupied by 1's. If  $(x, t)$  is occupied, we can identify exactly one site at time 0 which is the ancestor of the particle at  $(x, t)$ . We call the site of this particle the *first ancestor* in the hierarchy and denote it by  $\hat{\xi}_t^{(x,t)}(1)$ . We refer to the path it takes as the *path of the first ancestor*. Since the hierarchy of ancestors does not depend on the initial configuration, we will provide an algorithm for locating the first ancestor (and of all subsequent ancestors) without referring to a particular initial configuration. We refer to Figure 7 to explain the algorithm. Start at  $(x, t)$ . Go down the graphical representation until the first time a death mark is encountered. Go back up until the first time the tip of an arrow is encountered. [In Figure 7, this is the first 1-arrow in the graph when starting at  $(x, t)$ .] Follow this arrow against its direction to the branch the arrow is attached to. Take this branch until the first time a death mark is encountered. Repeat the above procedure until you reach time 0. The site you reached is the location of the first ancestor. (In Figure 7, this is  $x - 3$ .) To find the second ancestor, start at  $(x - 3, 0)$  and trace back your steps (only going upwards and across arrows) until the first time you encounter the tip of an arrow. Follow this arrow against its direction to the branch this arrow is attached to and repeat the algorithm for the first ancestor starting at this location until you reach time 0. The site you land on is the second ancestor. (In Figure 7, this is  $x - 1$ .) Continuing in this way, one obtains the dual process as an ordered set of ancestors

$$(3.5) \quad \hat{\xi}_s^{(x,t)} = \left( \hat{\xi}_s^{(x,t)}(1), \hat{\xi}_s^{(x,t)}(2), \dots \right)$$

for  $0 \leq s \leq t$ . This set is, of course, finite for  $t < \infty$ . We wish to point out, however, that the same site will typically appear repeatedly in the set of ancestors since different paths may lead to the same site. In Figure 7, the ancestor set is  $(x - 3, x - 1, x, x + 1, x - 1, x, x + 3, x + 4, x + 2, x + 3)$ . Note that the ancestor set  $\hat{\xi}_t^{(x,t)}$  depends solely on the graphical representation for the one-type contact process with parameter  $\beta_1$  and is independent of the initial configuration and of the labelling of the 1-arrows.

Using Figure 7, we will now explain how to determine the state of  $(x, t)$  when starting with an arbitrary initial configuration that contains both types. In Figure 7, the first ancestor is  $x - 3$ . If  $x - 3$  is occupied by a 1,  $(x, t)$  will be occupied by a 1 with the particle at  $x - 3$  being its ancestor. If  $x - 3$  is vacant,  $x - 1$  is the next candidate to determine the state of  $(x, t)$ , and so on. If  $x - 3$  is occupied by a 2, then the 2 can move up until the first time it encounters a 1-arrow. In this case, not only  $x - 3$  fails to determine the state of  $(x, t)$ , but also  $x - 1$  and  $x$  are blocked. That is, all ancestors of the site where the 1-arrow is attached are excluded. The next candidate now is  $x + 1$ . If  $x + 1$  fails (either because it is empty or occupied by a 2 which would be blocked by the following 1-arrow),  $x - 1$  would be the next candidate. For instance, if  $x - 1$  was occupied by a 2, this 2 would now be able to

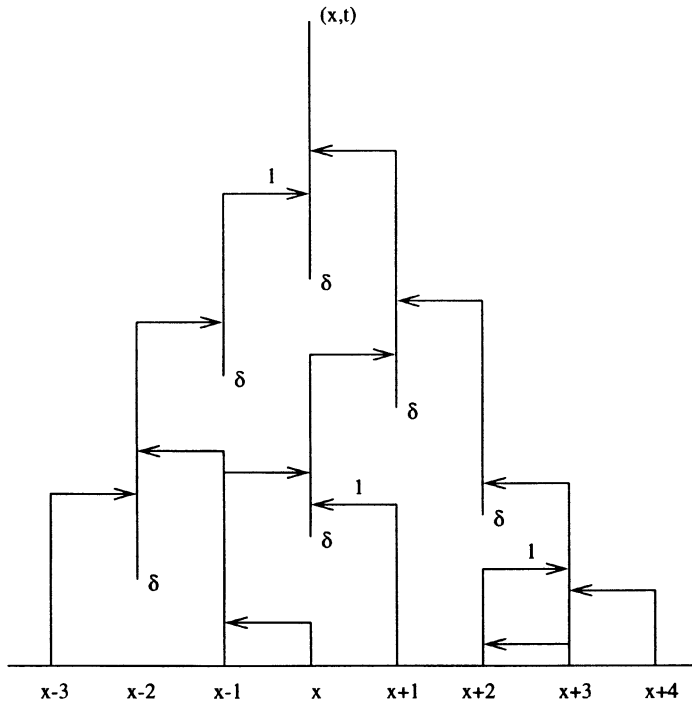


FIG. 7. A realization of the dual process for the multitype contact process.

paint  $(x, t)$  its color if no other ancestor has done so earlier, even though the first time it appeared in the ancestor set it would have been blocked by the upmost 1-arrow. Proceeding in this manner, one can determine the state of  $(x, t)$  and its ancestor.

The dual process is a complicated object. Fortunately, the path of the first ancestor, which is crucial to our analysis, is manageable. Its path can be broken up at certain points, which we call *renewal points*. These points are defined as follows. Whenever the first ancestor jumps to a site where the process starting at this site does not die out, the site is called a renewal point. We will also say that this site *lives forever*. We call a renewal point *associated with a 1-arrow* if the first arrow a particle crosses on its way up the designated path of the first ancestor starting at the renewal point is a 1-arrow.

We denote the spatial displacement between consecutive renewal points by  $X_i$ , and the corresponding temporal displacement by  $\tau_i$ . It was shown in Neuhauser (1992), using an idea of Kuczek (1989), that conditioned on survival,  $\{(X_i, \tau_i)\}_{i \geq 1}$  for an i.i.d. family of random vectors on  $\mathbf{Z}^d \times \mathbf{R}^+$  (here,  $d = 2$ ). Furthermore,

$$(3.6) \quad P(\|X_i\| > t) \leq C_1 e^{-\gamma_1 t} \quad \text{and} \quad P(\tau_i > t) \leq C_1 e^{-\gamma_1 t}$$

for appropriate constants  $C_1 < \infty$  and  $\gamma_1 > 0$ . ( $\|\cdot\|$  may stand for any  $L_p$  norm.) That is, the renewal points perform a random walk. This property enables us to control the location of the first ancestor. For example it gives the following large deviations estimate.

LEMMA 3.2. *Let  $\mathcal{X}_0 = 0$  and  $\mathcal{X}_n = \sum_{i=1}^n \mathcal{X}_i$ ,  $n \geq 1$ , be a renewal process whose interarrival times  $\mathcal{X}_i$  are i.i.d. with mean  $\Lambda$  and  $P(\mathcal{X}_i > z) \leq C_2 e^{-\gamma_2 z}$ . Then for any  $a > 0$ , there are appropriate  $C_3 < \infty$  and  $\gamma_3 > 0$  so that*

$$(3.7) \quad \begin{aligned} P(\mathcal{X}_n \geq (1+a)n\Lambda) &\leq C_3 e^{-\gamma_3 n\Lambda}, \\ P(\mathcal{X}_n \leq (1-a)n\Lambda) &\leq C_3 e^{-\gamma_3 n\Lambda}. \end{aligned}$$

We are now ready to embark on the proof of Lemma 3.1. (To simplify notation, we translate in space to assume  $B(0)$  is  $G_1$ . The key to the proof is a procedure we call the *repositioning algorithm*. As in Bramson and Griffeath's (1980), (1981) studies of the biased voter model, the idea is to select a dual path, denoted by  $A_t$ , that will, with probability close to one, land in the *target region*  $[-2L/5, 2L/5]^2$  after  $T$  units of time when starting at some  $(x, T)$  with  $x \in B(z)$ ,  $z \in \mathbf{Z}^2$ ,  $\|z\|_\infty = 1$ . That is,  $A_0 = (x, T)$ . This path will block 2's from determining the type of  $(x, T)$ . In particular, a 1 can move up this path to  $(x, T)$  if no other 1 has done so earlier. Into the path  $A_t$  we embed a jump process  $S_t$  which stays put except at times when  $A_t$  jumps to a renewal point that is associated with a 1-arrow. At such times, denoted by  $\{\sigma_i\}_{i \geq 1}$ , we apply the "repositioning algorithm." Between times  $\sigma_i$ ,  $A_t$  follows the path determined by the algorithm of a first ancestor. We set  $\sigma_0 = 0$ .

We define the repositioning algorithm inductively. Assume that at time  $\sigma_i$ , some  $i \geq 1$ ,  $A_{\sigma_i}$  is at the location of a renewal point associated with a 1-arrow. If we pretend that a 2 was able to come up all the way to this 1-arrow, then this 2 would now be blocked by this 1-arrow and consequently some other site at time  $\sigma_i$ , if it exists, would be the next candidate for a path determining the type of the particle at  $(x, T)$ . This site does not necessarily live forever, but there is a positive probability (independent of  $L$ ) that within one unit of time, this new path jumps to a site for which the dual will live forever. We call this site  $B_{\sigma_i}$ . If we denote by  $J = [-L/10, L/10]^2$  and by  $\text{dist}(z, J)$  the Euclidean distance between a point  $z$  and the set  $J$ , then we use the following rule:

$$(3.8) \quad \begin{aligned} &\text{If } \text{dist}(B_{\sigma_i}, J) < \text{dist}(A_{\sigma_i}, J), \text{ proceed with the new path,} \\ &\text{otherwise continue with the old path.} \end{aligned}$$

If there is no new path or if the path does not jump to a site which will live forever within one unit of time, continue with the old path. If we select the new path, we set  $S_{\sigma_i} = B_{\sigma_i}$ , otherwise  $S_{\sigma_i} = A_{\sigma_i}$ . In either case,  $A_t$  continues starting at  $S_{\sigma_i}$  and uses the path determined by the algorithm of the first ancestor until time  $\sigma_{i+1}$  where the repositioning algorithm is applied again. If both  $B_{\sigma_i}$  and  $A_{\sigma_i}$  are contained in  $J$ , toss a fair coin to determine which path to continue with.

In Neuhauser (1992) it was shown that the spatial displacement of the embedded random walk in the path of the first ancestor has mean 0. From (3.8) and the fact that there is positive probability that the repositioning algorithm selects  $B_{\sigma_i}$  instead of  $A_{\sigma_i}$ , it follows that  $\|S_{\sigma_{i+1}}\| - \|S_{\sigma_i}\|$  has negative mean as long as  $S_t$  is outside of the set  $J$ . Intuitively, this should cause  $S_t$  to drift towards the set  $J$ . We make this precise in the following.

We begin by introducing the following quantity. Assume that the dual process starting at  $(x, t)$  survives and set

$$(3.9) \quad \tau = \inf\{s: \hat{\xi}_s^{(x,t)}(1) \text{ jumps to a renewal point associated with a 1-arrow}\}.$$

The following result shows that the distribution of  $\tau$  has exponential tails.

LEMMA 3.3. *Assume  $c_1 = c_2 = 0$  and  $\beta_1 > \beta_2$ . There exist  $C_4 < \infty$  and  $\gamma_4 > 0$  so that*

$$(3.10) \quad P(\tau > t) \leq C_4 e^{-\gamma_4 t}.$$

PROOF. This follows immediately from (3.6) and  $\beta_1 > \beta_2$  since every time the first ancestor goes through a renewal point, there is probability  $(\beta_1 - \beta_2)/\beta_1$  that the associated arrow is a 1-arrow. Since the 1-arrows are independent of each other, a geometric number of trials suffices which, together with (3.6), implies (3.10).  $\square$

We say that the repositioning algorithm was *applied successfully* if the new path was chosen. We denote by  $N(t)$  the number of times the repositioning algorithm has been applied successfully by time  $t$ .

LEMMA 3.4. *Assume  $c_1 = c_2 = 0$  and  $\beta_1 > \beta_2$ . There exists  $b > 0$  so that*

$$P(N(t) \leq bt) \leq C_5 \exp(-\gamma_5 t)$$

for appropriate  $C_5 < \infty$  and  $\gamma_5 > 0$ .

PROOF. This follows from Lemma 3.3 and the second large deviations estimate in (3.7) since each time the repositioning algorithm is applied, there is a positive probability that it is applied successfully.  $\square$

We still need an estimate on the maximum size of the spatial displacement of the embedded jump process  $S_t$ . We denote the spatial displacement  $S_{\sigma_i} - S_{\sigma_{i-1}}$  by  $Y_i$ , for  $i \geq 1$ . Let  $M(t)$  denote the number of times the repositioning algorithm has been applied by time  $t$  (regardless of whether it was successful).

LEMMA 3.5. *Set  $T = L^2$ . Assume  $c_1 = c_2 = 0$  and  $\beta_1 > \beta_2$ . For any  $\gamma_6 < \infty$  there exist  $C_6 < \infty$  and  $\gamma_7 > 0$  so that for  $L$  sufficiently large*

$$(3.11) \quad P\left(\max_{1 \leq i \leq M(T)} \|Y_i\| > \frac{\gamma_7}{\gamma_4} \log L\right) \leq C_6 L^{-\gamma_6}.$$

PROOF. This is, with some minor modifications, the proof of Lemma 9 in Neuhauser (1992). Set  $m = EM(t)/t$ . We decompose the event on the left-hand side of (3.11) according to  $\{M(T) > 2mT\}$  and  $\{M(T) \leq 2mT\}$ . Then for any  $\gamma_7 > 0$ ,

$$(3.12) \quad \begin{aligned} & P\left(\max_{1 \leq i \leq M(T)} \|Y_i\| > \frac{\gamma_7}{\gamma_4} \log L\right) \\ & \leq P(M(T) > 2mT) + 2mTP\left(\|Y_i\| > \frac{\gamma_7}{\gamma_4} \log L\right) \end{aligned}$$

The first large deviations estimate in (3.7) takes care of the first term on the right-hand side of (3.12). For the second term on the right-hand side of (3.12), we use Lemma 3.3 combined with an argument in Neuhauser (1992) which shows that an estimate similar to (3.10) holds for the spatial displacement. (This argument essentially uses the fact that the contact process spreads out at most linearly in time and hence the spatial displacement is of the same order as the temporal displacement.) Hence we can bound (3.12) by

$$\leq C_3 \exp(-\gamma_3 mT) + 2mTC_4 \exp(-\gamma_7 \log L).$$

For given  $\gamma_6 > 0$ , we can choose  $\gamma_7 > 0$  so that (3.11) holds for appropriate  $C_6 < \infty$  and  $L$  sufficiently large.  $\square$

The next lemma shows that, with probability close to 1, the selected path will land in the target region by time  $\sigma_K$  where

$$K = \min\{k : \sigma_k \geq T - 2L^{0.5}\}$$

It follows from Lemma 3.3 that except for an exponentially small probability ( $\leq C_4 \exp(-\gamma_4 L^{0.5})$ ),  $\sigma_K \leq T - L^{0.5}$ . Here the subscripts on  $P$  indicate the starting site for the process  $S_t$ .

LEMMA 3.6. *Assume  $x \in B(z)$  for some  $z \in \mathbf{Z}_2$  with  $\|z\|_\infty = 1$ ,  $c_1 = c_2 = 0$  and  $\beta_1 > \beta_2$ . For any  $\gamma_8 > 0$ ,*

$$(3.13) \quad P_x\left(S_{\sigma_K} \notin [-2L/5, 2L/5]^2\right) \leq C_8 L^{-\gamma_8}$$

for appropriate  $C_8 < \infty$  and  $L$  sufficiently large.

PROOF. We observe that

$$(3.14) \quad \begin{aligned} & P_x\left(S_{\sigma_K} \notin [-2L/5, 2L/5]^2\right) \\ & \leq P_x\left(S_{\sigma_k} \notin J \text{ for all } k \leq K\right) \\ & \quad + \sup_{y \in J} P_y\left(S_{\sigma_k} \notin [-2L/5, 2L/5]^2 \text{ for some } k \leq M(T)\right). \end{aligned}$$

To estimate the first term on the right-hand side of (3.14), note that

$$\begin{aligned}
 & P_x(S_{\sigma_k} \notin J \text{ for all } k \leq K) \\
 & \leq P_x(S_{\sigma_k} \notin J \text{ for all } k \leq mT/2) + P(K \leq mT/2) \\
 (3.15) \quad & \leq P_x\left(S_{\sigma_k} \notin J \text{ for all } k \leq mT/2; \max_{1 \leq k \leq mT/2} \|Y_k\| \leq \frac{\gamma_7}{\gamma_4} \log L\right) \\
 & \quad + P\left(\max_{1 \leq k \leq mT/2} \|Y_k\| > \frac{\gamma_7}{\gamma_4} \log L\right) + P(K \leq mT/2).
 \end{aligned}$$

We restrict the jump size  $\|Y_k\|$  to ensure that  $S_{\sigma_k}$  will not miss the set  $J$ . Since there exists  $\mu_1 > 0$  so that  $E(\|S_{\sigma_k}\| - \|S_{\sigma_{k-1}}\| | \mathcal{F}_{\sigma_{k-1}}) \leq -\mu_1$  on  $\{S_{\sigma_{k-1}} \notin J\}$ , it follows that there exists  $\mu_2$  with  $0 < \mu_2 \leq \mu_1$ , so that

$$E\left(\left(\|S_{\sigma_k}\| - \|S_{\sigma_{k-1}}\|\right) \mathbf{1}_{(\|Y_k\| \leq (\gamma_7/\gamma_4)\log L)} | \mathcal{F}_{\sigma_{k-1}}\right) \leq -\mu_2 \quad \text{on } \{S_{\sigma_{k-1}} \notin J\}$$

for  $\gamma_7$  and  $L$  sufficiently large, where  $\mathcal{F}_{\sigma_{k-1}} = \sigma(S_{\sigma_0}, \dots, S_{\sigma_{k-1}})$ . Furthermore, we have  $\|\|S_{\sigma_k}\| - \|S_{\sigma_{k-1}}\|\| \leq \|Y_k\|$  so it follows that

$$\begin{aligned}
 (3.16) \quad & P_x\left(S_{\sigma_k} \notin J \text{ for all } k \leq mT/2; \max_{1 \leq k \leq mT/2} \|Y_k\| \leq \frac{\gamma_7}{\gamma_4} \log L\right) \\
 & \leq C_9 \exp(-\lambda_9 L)
 \end{aligned}$$

for appropriate  $C_9 < \infty$  and  $\lambda_9 > 0$ . The second term on the right-hand side of (3.15) can be bounded using (3.11) and the third term on the right-hand side of (3.15) can be bounded using the second inequality in (3.7). Combining the estimates, it follows that for any  $\gamma_{10} > 0$  there exists  $C_{10} < \infty$  so that

$$P_x(S_{\sigma_k} \notin J \text{ for all } k \leq K) \leq C_{10} L^{-\gamma_{10}}$$

for  $\gamma_7$  and  $L$  sufficiently large.

Once the path is in  $J$ , we need to keep it inside of  $[-2L/5, 2L/5]^2$  for the remainder of the time. Assume now that at time 0, the path starts at some site in  $J$ . We define the following stopping times. Let  $s_0 = 0$  and define for  $l \geq 1$ ,

$$\begin{aligned}
 t_l &= \inf\{t > s_{l-1}; S_t \notin [-L/5, L/5]^2\} \\
 s_l &= \inf\{t > t_l; S_t \in J\}
 \end{aligned}$$

and set

$$R = \inf\{t; S_t \notin [-2L/5, 2L/5]^2\} \quad \text{and} \quad l_0 = \inf\{l; t_l > T\}.$$

Then for any  $y \in J$

$$\begin{aligned}
 & P_y(S_{\sigma_k} \notin [-2L/5, 2L/5]^2 \text{ for some } k \leq M(T)) \\
 & \leq P\left(\bigcup_{l=1}^{2mT} \{s_l > R\}\right) + P\left(\max_{1 \leq k \leq 2mT} \|Y_k\| > \frac{\gamma_7}{\gamma_4} \log L\right) + P(l_0 > 2mT) \\
 & \leq 2mTP(s_1 > R) + P\left(\max_{1 \leq k \leq 2mT} \|Y_k\| > \frac{\gamma_7}{\gamma_4} \log L\right) + P(M(T) > 2mT).
 \end{aligned}$$

To estimate  $P(s_1 > R)$ , we only need to estimate the probability that  $S_t$  will leave  $[-2L/5, 2L/5]^2$  before returning to  $J$ . As in the estimate for (3.16), we use that  $\|S_{\sigma_k}\| - \|S_{\sigma_{k-1}}\|$  has negative drift and bounded increments on the set where  $\max_{1 \leq k \leq 2mT} \|Y_k\| \leq (\gamma_7/\gamma_4)\log L$ . Hence,

$$P(s_1 > R) \leq C_{11} \exp(-\gamma_{11}L)$$

for appropriate  $C_{11} < \infty$  and  $\gamma_{11} > 0$ . The other two terms are estimated using (3.11) and (3.7) as above and the lemma follows.  $\square$

So far we have demonstrated that there exists, with probability close to 1, a path that lands in the target region  $[-2L/5, 2L/5]^2$  at time  $\sigma_K$  and blocks 2's from determining the type of  $(x, T)$ . Since this path does not necessarily land at an occupied site, we still need to show that if  $(x, T)$  is occupied, one of the sites in  $B(0)$  (which is only occupied by 1's) determines the type of  $(x, T)$  provided no other 1 succeeded earlier. The idea is the following. We run the dual process starting at  $(S_K, \sigma_K)$  for another  $L^{0.5} - L^{0.2}$  units of time. (Recall that except for an exponentially small probability,  $\sigma_K \leq T - L^{0.5}$ .) It follows from the properties of the dual of the contact process that except for exponentially small probability ( $\leq C_{12} \exp(-\gamma_{12}L^{0.1})$ ), we can select  $L^{0.1}$  sites at about time  $L^{0.2}$  which are contained in the dual process starting at  $(S_K, \sigma_K)$  such that (i) all these sites are contained in  $[-3L/5, 3L/5]^2$ , (ii) they are at least  $L^{0.3}$  units apart from each other, and (iii) none of the duals starting at these sites interferes with any of the other duals for the remaining  $L^{0.2}$  units of time. Recall that the configuration in  $B(0)$  is  $G_1$  at time 0. In particular, this means that every  $L^{0.1} \times L^{0.1}$  square  $D(w)$ ,  $w \in I_0$ , contains at least one site occupied by a 1. This, together with the fact that each of the duals has positive probability of surviving, implies that each of these  $L^{0.1}$  sites has probability  $\eta > 0$  of being occupied. On the set where the duals starting at these  $L^{0.1}$  sites do not interfere, these trials are independent. Therefore,

$$\begin{aligned} & P((S_K, \sigma_K) \text{ is not occupied by a 1}) \\ (3.17) \quad & \leq C_4 \exp(-\gamma_4 L^{0.5}) + C_{12} \exp(-\gamma_{12} L^{0.1}) + (1 - \eta)^{L^{0.1}} \\ & \leq C_{13} \exp(-\gamma_{13} L^{0.1}) \end{aligned}$$

for appropriate  $C_{13} < \infty$  and  $\gamma_{13} > 0$ . We can now finish the proof of Lemma 3.1.

PROOF OF LEMMA 3.1. Combining (3.13) and (3.17) it follows that for any  $x \in B(0)$ ,

$$(3.18) \quad P(\xi_T(x) = 2) \leq C_8 L^{-\gamma_8} + C_{13} \exp(-\gamma_{13} L^{0.1}).$$

Since  $\gamma_8$  was arbitrary, (3.3) follows.  $\square$

It is now straightforward to prove Propositions 3.1 and 3.2.

PROOF OF PROPOSITION 3.1. There are at most  $(2L + 1)^2$  sites in  $[-L, L]^2$ . This together with Lemma 3.1 implies

$$(3.19) \quad P(\xi_T(x) = 2 \text{ for some } x \in B(0)) \leq (2L + 1)^2 CT^{-4} \leq C_{14} L^{-6}.$$

for appropriate  $C_{14} < \infty$ . Furthermore, since the process dominates a one-color contact process with parameter  $\beta_2 > \beta_c$ , the probability that there exists an  $L^{0.1} \times L^{0.1}$  square  $D(w)$ ,  $w \in I_0$ , that is empty, can be bounded by

$$(3.20) \quad \leq C_{15} L^{1.8} \exp(-\gamma_{15} L^{0.2})$$

Combining (3.19) and (3.20) yields Proposition 3.1.  $\square$

PROOF OF PROPOSITION 3.2. To show that  $[-L - r, L + r]^2 \times [T - r, 2T]$  remains void of 2's provided  $B(z)$  is  $G_1$  at time 0 for some  $z \in \mathbf{Z}^2$  with  $\|z\|_\infty = 1$ , we apply Lemma 3.1 to all sites in the region

$$E = \left( ([-L - r, L + r]^2 \times [T, 2T]) - ([-L, L]^2 \times [T, 2T]) \right) \\ \cup ([-L - r, L + r]^2 \times [T - r, T]).$$

There are at most  $(2L + r + 1)^2(T + r + 1) \leq 8L^3$  sites in  $E$ . Hence,

$$P(\xi_t(x) = 2 \text{ for some } (x, t) \in E) \leq 8L^3 CT^{-4} \leq C_{16} L^{-5}.$$

Since the range is  $r$ , this implies that on the set  $\{\xi_t(x) \neq 2 \text{ for all } (x, t) \in E\}$ , the whole region  $C(0, 0)$  is void of 2's. This completes the proof of Proposition 3.2.  $\square$

To make the comparison with  $M$ -dependent oriented site percolation, we still need to show that boxes that are sufficiently far apart are independent of each other with high probability. This is contained in the following lemma.

LEMMA 3.7. *Assume  $c_1 = c_2 = 0$  and  $\beta_1 > \beta_2$ . Set  $T = L^2$ . For any  $\varepsilon_1 > 0$  we can find  $M > 0$  so that for  $L$  sufficiently large,*

$$(3.21) \quad P(\text{any of the selected paths is not contained in} \\ [-ML/3, ML/3]^2 \text{ for some } t \leq 2T) \leq \varepsilon_1.$$

PROOF. This follows from the fact that the selected paths have a drift toward the target region. A large deviations estimate then shows (3.21).  $\square$

To make the comparison we assume that all the sites outside of  $[-ML, ML]^2$  are occupied by 2's throughout  $[0, 2T]$ . By stacking the boxes as in Durrett and Schinazi (1993) and using Durrett (1992a), it follows that there is a cone void of 2's linearly growing in time. Inside the cone, the distribution of 1's is close to the equilibrium distribution for the corresponding one-color contact process. This follows from a coupling argument.

To remove the condition of translation invariance of the initial configuration, we note that if initially there are infinitely many 1's, then, with probability 1, there will be infinitely many occupied squares  $B(z)$ ,  $z =$



$(z_1, z_2) \in \mathbf{Z}^2$  with  $z_1$  and  $z_2$  both odd, at time  $T$ . This is a simple consequence of the Borel–Cantelli lemma since each square that contains 1’s at time 0 has a positive probability of being occupied at time  $T$  and sufficiently far apart squares  $B(z)$  do not influence each other with high probability for all  $t \leq T$ . The random set

$$A = \{z = (z_1, z_2) \in \mathbf{Z}^2: z_1 \text{ and } z_2 \text{ are both odd and } B(z) \text{ is occupied at time } T\}$$

can thus serve as the “source” for the rescaling argument and we can start the iteration at time  $T$ .

The proof of Theorem 2 is now a straightforward perturbation argument whose main ingredient is contained in the following lemma. We say a site is *doubly occupied* if it is occupied by both strains (i.e., it is in state 3). We set

$$F = \{\text{some site in } B(0) \text{ becomes doubly occupied during } [0, T]\}.$$

LEMMA 3.8. *Given  $\beta_1$  and  $L$ . for any  $\delta > 0$  we can choose  $c_1$  and  $c_2$  sufficiently small so that  $P(F) \leq \delta/3$ .*

PROOF. There are at most  $(2L + 1)^2$  sites in  $B(0)$ . The total rate at which sites can become doubly occupied is thus bounded by  $(c_1 \vee c_2)\beta_1 5L^2$ . Hence,

$$P(F) \leq 1 - \exp\{-(c_1 \vee c_2)\beta_1 5L^2 T\} \leq 5(c_1 \vee c_2)\beta_1 L^4 \leq \delta/3$$

for  $c_1$  and  $c_2$  sufficiently small.  $\square$

Combining Propositions 3.1 and 3.2 with Lemma 3.8, Theorem 2 follows for  $c_1$  and  $c_2$  close to 0. Using the monotonicity in Proposition 1.1, then implies that we can extend our result to all  $c_1 \leq 1$ . This concludes the proof of Theorem 2.  $\square$

**4. Proof of Theorem 4.** Our general approach in this section is to consider the process on the scaled lattice  $\mathbf{Z}^2/r$  and let  $r \rightarrow \infty$ . To get a lower bound on the growth of the set of 1 infections  $\eta_t^r$  we will start with all sites occupied by 2’s. To set the stage, we begin by considering the behavior of the set of sites occupied by 2’s,  $\zeta_t^r$ , when there are no 1 infections. In this case,  $\zeta_t^r$  is a long-range contact process. If we let  $\bar{\zeta}_t^r$  and  $\zeta_t^{r,x}$  denote the long-range contact process with  $\bar{\zeta}_0^r = \mathbf{Z}^2/r$  and  $\zeta_0^{r,x} = \{x\}$  where  $x \in \mathbf{Z}^2/r$ , then the duality equation for the contact process implies

$$(4.1) \quad P(x \in \bar{\zeta}_t^r) = P(\zeta_t^{r,x} \neq \emptyset),$$

a quantity which is independent of  $x$  and that we will call  $\rho_r(t)$ .

Let  $Z_t^x$  be a branching random walk that starts with one particle at  $x$  and in which a particle at  $y$  dies at rate  $\delta_2$  and at rate  $\beta_2$  gives birth to an offspring that is sent to  $y + U$  where  $U$  is uniform on  $[-1, 1]^2$ . In the proof of

Theorem 3 of Durrett (1991) it is shown that as  $r \rightarrow \infty$   $\{\zeta_t^{r,x}, t \geq 0\}$  converges weakly to  $\{Z_t^x, t \geq 0\}$  and hence

$$(4.2) \quad \rho_r(t) \rightarrow \rho(t) = P(Z_t^x \neq \emptyset).$$

The reasoning just applied to the density of occupied sites also applies to a pair of sites. Duality implies

$$(4.3) \quad P(x, y \in \bar{\zeta}_t^r) = P(\zeta_t^{r,x} \neq \emptyset, \zeta_t^{r,y} \neq \emptyset).$$

The limiting processes  $Z_t^x$  and  $Z_t^y$  are independent so it should not be surprising that the events on the right-hand side are asymptotically independent. Using the methods of the proof of Theorem 3 in Durrett (1991), one can easily show that as  $r \rightarrow \infty$

$$(4.4) \quad v_r(t) = \sup_{x \neq 0} |P(0 \in \zeta_t^r, x \in \zeta_t^r) - \rho_r^2(t)| \rightarrow 0.$$

If we imagine sites occupied by 2 as white and the others black, then (4.2) and (4.4) imply (with a little help from Chebyshev's inequality) that in the limit as  $r \rightarrow \infty$ ,  $\zeta_t^r$  will be a featureless plane with the shade of gray determined by  $\rho(t)$ .

Let  $\eta_1^r(t)$  and  $\eta_3^r(t)$  denote the set of sites occupied by 1's and 3's at time  $t$ . In the following lemma we will establish the fact that  $(\eta_1^r(t), \eta_3^r(t))$  converges to a (temporarily inhomogeneous) two type branching random walk  $(\eta_1(t), \eta_3(t))$ . The number of 1's and 3's at time  $t$  in the limit process,  $N_t^i = |\eta_i(t)|$ , are a two-type branching process with the following transactions:

transition	rate at time $t$
$0 \rightarrow 1$	$(N_t^1 + c_1 N_t^3) \beta_1 (1 - \rho(t))$
$2 \rightarrow 3$	$(N_t^1 + c_1 N_t^3) c_1 \beta_1 \rho(t)$
$1 \rightarrow 3$	$N_t^1 c_2 \beta_2 \rho(t)$
$3 \rightarrow 1$	$N_t^3 \delta_2 / c_2$
$1 \rightarrow 0$	$N_t^1 \delta_1$
$3 \rightarrow 2$	$N_t^3 \delta_1 / c_1$ .

Here we have written the transitions as they occur in the particle system. The first two ( $0 \rightarrow 1$  and  $2 \rightarrow 3$  transitions) result in new individuals displaced from their parents by an amount uniformly distributed over  $[-1, 1]^2$ . Type changes ( $1 \rightarrow 3$ ,  $3 \rightarrow 1$ ) and deaths ( $1 \rightarrow 0$ ,  $3 \rightarrow 2$ ) occur at the particle's location.

**LEMMA 4.1.** *Start with 2 infections at all sites. Add a fixed number of 1 infections to sites  $x_1^r, \dots, x_k^r$  and suppose  $x_i = \lim x_i^r$  exists for each  $i$ . Let  $\eta_1^r(t)$  and  $\eta_3^r(t)$  denote the set of sites occupied by 1's and 3's at time  $t$ . Then  $(\eta_1^r(t), \eta_3^r(t))$  converges to the (temporally inhomogeneous) two-type branch-*

ing random walk  $(\eta_1(t), \eta_3(t))$  defined above with  $\eta_1(0) = \emptyset$  and  $\eta_3(0) = \{x_1, x_2, \dots, x_K\}$ .

PROOF. The transition rates defined above are obtained by assuming that if  $r$  is large, then in the neighborhood of any point (i) the fractions of 1's and 3's are small and (ii) the fractions of 0's and 2's are equal to their equilibrium values. To prove Lemma 4.1 we simply have to justify these assumptions. For (i) we note that even if we ignore the deaths of 1's and 3's and assume that each individual gives birth at rate  $\beta_1$  to a new 1 to get an upper bound on the total number of births in the real process, we have for any  $T < \infty$

$$E \sup_{t \leq T} (N_t^1 + N_t^3) \leq \exp(\beta_1 T)(N_0^1 + N_0^3).$$

From this and Chebyshev's inequality, it follows that, for any fixed  $T$ , if  $M_r \rightarrow \infty$  then

$$P\left(\sup_{t \leq T} (N_t^1 + N_t^3) > M_r\right) \leq M_r^{-1} \exp(\beta_1 T)(N_0^1 + N_0^3) \rightarrow 0$$

as  $r \rightarrow \infty$ .

The process of 1's only interferes with the 2's at the sites that it occupies by reducing the birth rate of 2's and increasing the death rate of 2's at those sites. Thus the last estimate implies that the 2 process will with high probability lose at most  $M_r$  particles from this interference. In view of this, if we let  $M_r \rightarrow \infty$  slowly, we can ignore the 1 infections in checking (ii) and suppose that the set of 2's is a long-range contact process.

Suppose that all the initial 1 infections lie in  $[-A, A]^2$  (on the scaled lattice  $\mathbf{Z}^2/r$ ). To prove (ii) now we will let  $h_r \rightarrow 0$  slowly and look at the number of 2's in the boxes  $B_{ij}^r = (ih_r, jh_r) + [0, h_r)^2$  where  $-l(r) \leq i, j \leq l(r)$  and  $l(r) = (A + 2 + M_r)/h_r$ . To explain the range of indices, note that when the number of births is less than or equal to  $M_r$ , all sites occupied by 1's or 3's lie in  $[-A - M_r, A + M_r]^2$  and all of their neighbors lie in  $[-A - 1 - M_r, A + 1 + M_r]^2$ . Replacing 1 by 2 takes care of the boxes that intersect the edge of the square.

Let  $S_{ij}^r(t) = |\zeta_t^r \cap B_{ij}^r|$ . To prove (ii) we will show that for  $\varepsilon > 0$ :

$$(4.5) \quad P\left(\sup_{\substack{0 \leq t \leq T \\ -l(r) \leq i, j \leq l(r)}} \left| \frac{S_{ij}^r(t)}{|B_{ij}^r|} - \rho_r(t) \right| > 3\varepsilon\right) \rightarrow 0.$$

We begin by estimating the probability for one square at a fixed time. By Chebyshev's inequality

$$P\left(\left| \frac{S_{ij}^r(t)}{|B_{ij}^r|} - \rho_r(t) \right| > \varepsilon\right) \leq \varepsilon^{-2} \frac{\text{var}(S_{ij}^r(t))}{|B_{ij}^r|^2}.$$

Equation (4.4) gives a bound on the covariance of the terms in  $S_{ij}^r(t)$  which implies that

$$\text{var}(S_{ij}^r(t)) \leq |B_{ij}^r| + v_r |B_{ij}^r|^2.$$

Combining this with the Chebyshev bound we have

$$P\left(\left|\frac{S_{ij}^r(t)}{|B_{ij}^r|} - \rho_r(t)\right| > \varepsilon\right) \leq \varepsilon^{-2}(|B_{ij}^r|^{-1} + v_r).$$

Our next goal is to estimate the probability of

$$G_t = \left\{ \sup_{-l(r) \leq i, j \leq l(r)} \left| \frac{S_{ij}^r(t)}{|B_{ij}^r|} - \rho_r(t) \right| \leq \varepsilon \right\}.$$

Adding up the error probabilities, we have for fixed  $t$  that

$$P(G_t^c) \leq (2l(r) + 1)^2 \varepsilon^{-2}(|B_{ij}^r|^{-1} + v_r).$$

Now  $|B_{ij}^r| \geq [rh_r]^2$ , where  $[x]$  denotes the integer part of  $x$  and  $l(r) \sim M_r/h_r$  as  $r \rightarrow \infty$ , so

$$(4.6) \quad P(G_t^c) \leq C \left( \frac{M_r}{h_r} \right)^2 ((rh_r)^{-2} + v_r).$$

If we let  $h_r \rightarrow 0$  and  $M_r \rightarrow \infty$  slowly enough, the right-hand side tends to 0 as  $r \rightarrow \infty$ .

The estimate in (4.6) is for one fixed time but trivially implies that if  $\tau$  and  $L$  are fixed

$$(4.7) \quad P(G_t^c \text{ for some } t \in \{0, \tau, \dots, L\tau\}) \rightarrow 0.$$

To handle intermediate times let  $D_{ij,k}^r$  be the number of sites in  $B_{ij}^r$  hit by “2-deaths” in the time interval  $[k\tau, (k+1)\tau]$ . Pick  $\tau$  so that  $1 - \exp(-\tau\delta_2) \leq \varepsilon/2$  and if  $|s - t| \leq \tau$  then  $|\rho_r(s) - \rho_r(t)| \leq \varepsilon$ . The latter is possible since  $t \rightarrow \rho_r(t)$  is decreasing and continuous. Since the events that a 2-death hits a site are independent for different sites, Chebyshev’s inequality implies

$$P(D_{ij,k}^r > \varepsilon | B_{ij}^r) \leq \frac{\text{var}(D_{ij,k}^r)}{(\varepsilon/2)^2 |B_{ij}^r|^2} \leq C(\varepsilon/2)^{-2} |B_{ij}^r|^{-1}.$$

So if we let  $H = \{D_{ij,k}^r \leq \varepsilon | B_{ij}^r \text{ for all } -l(r) \leq i, j \leq l(r) \text{ and } 0 \leq k < L\}$ , then using the previous estimate and the asymptotics for  $|B_{ij}^r|$  and  $l(r)$  we have

$$(4.8) \quad \begin{aligned} P(H^c) &\leq L(2l(r) + 1)^2 \cdot 4\varepsilon^{-2} |B_{ij}^r|^{-1} \\ &\leq CL \left( \frac{M_r}{h_r} \right)^2 (rh_r)^{-2} \rightarrow 0 \end{aligned}$$

by the choices of (4.6).

To combine (4.7) and (4.8) to get (4.5) and complete the proof of Lemma 4.1 we will show:

On  $H \cap (\cap_{k=0}^L G_{k\tau})$ , we have

$$(4.9) \quad \sup_{\substack{0 \leq t \leq T \\ -l(r) \leq i, j, \leq l(r)}} \left| \frac{S_{ij}^r(t)}{|B_{ij}^r|} - \rho_r(t) \right| \leq 3\varepsilon.$$

For one direction we observe that if  $S_{ij}^r(k\tau) \geq (\rho_r(k\tau) - \varepsilon)|B_{ij}^r|$  and  $D_{ij,k}^r \leq \varepsilon|B_{ij}^r|$  then  $S_{ij}^r(t) \geq (\rho_r(t) - 2\varepsilon)|B_{ij}^r|$  for all  $t \in [k\tau, (k+1)\tau]$ . [Here we use the fact that  $t \rightarrow \rho_r(t)$  is decreasing.] On the other hand, if  $S_{ij}^r(t) > (\rho_r(t) + 3\varepsilon)|B_{ij}^r|$  for some  $t \in [k\tau, (k+1)\tau]$  and  $D_{ij,k}^r \leq \varepsilon|B_{ij}^r|$ , then  $S_{ij}^r((k+1)\tau) > (\rho_r((k+1)\tau) + \varepsilon)|B_{ij}^r|$ , a contradiction. [Recall our choice of  $\tau$  implies  $\rho_r((k+1)\tau) \geq \rho_r(t) - \varepsilon$ .]  $\square$

The next result should explain the reason for the assumption in Theorem 4.

LEMMA 4.2. *The instability assumption in Theorem 4 implies that the branching process is supercritical for large  $t$ .*

PROOF. As  $t \rightarrow \infty$ ,  $\rho(t)$  converges to the probability the branching process  $Z_t^x$  survives for all time  $t$ . It is well known [see, e.g., Bramson, Durrett, and Swindle (1989) or Durrett (1991)] and easy to compute that this probability is  $(\beta_2 - \delta_2)/\beta_2$ . If we replace  $\rho(t)$  in the transition rates of the two-type branching process by its limiting value, then we get a two-type branching process in which the mean number of particles of type  $i$ ,  $v_i(t)$ , satisfies

$$(4.10) \quad \begin{aligned} \frac{dv_1}{dt} &= v_1 \left( \beta_1 \cdot \frac{\delta_2}{\beta_2} - c_2 \beta_2 \cdot \frac{\beta_2 - \delta_2}{\beta_2} - \delta_1 \right) \\ &\quad + v_3 \left( c_1 \beta_1 \cdot \frac{\delta_2}{\beta_2} + \frac{\delta_2}{c_2} \right), \\ \frac{dv_3}{dt} &= v_1 \left( (c_1 \beta_1 + c_2 \beta_2) \cdot \frac{\beta_2 - \delta_2}{\beta_2} \right) \\ &\quad + v_3 \left( c_1^2 \beta_1 \cdot \frac{\beta_2 - \delta_2}{\beta_2} - \frac{\delta_1}{c_1} - \frac{\delta_2}{c_2} \right). \end{aligned}$$

It is easy to see that this coincides with the linearization of (1.1) about the fixed point given in (1.5). Indeed, given that the linearization is based on assuming  $u_2 \approx (\beta_2 - \delta_2)/\beta_2$ ,  $u_0 \approx \delta_2/\beta_2$  and  $u_1 = \varepsilon v_1$ ,  $u_3 = \varepsilon v_3$  with  $\varepsilon$  small, the two must coincide.

If we write (4.10) in vector form:  $dv/dt = Qv$ , then we can write the solution as

$$v(t) = e^{Qt}v(0) \quad \text{where } e^{Qt} = \sum_{n=0}^{\infty} \frac{Q^n t^n}{n!}.$$

The assumption of Theorem 4 is that there is an eigenvector

$$Q \begin{pmatrix} v_1 \\ v_3 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_3 \end{pmatrix}$$

with  $v_1, v_3$ , and  $\lambda$  positive. Plugging this into the formula for the solution we have

$$e^{Qt} \begin{pmatrix} v_1 \\ v_3 \end{pmatrix} = e^{\lambda t} \begin{pmatrix} v_1 \\ v_3 \end{pmatrix}$$

If  $\Gamma = e^Q$  then  $\Gamma_{ij}$  is the mean number of individuals of type  $j$  at time 1 when we start with one individual of type  $i$ . The interpretation implies  $\Gamma_{ij} > 0$  so the Perron–Frobenius theorem for positive matrices [see, e.g., Section 1.1 of Seneta (1973)] implies that there is a positive eigenvalue  $\sigma$  larger than the modulus of the other eigenvalue (which of course must be real in this  $2 \times 2$  case) and an associated strictly positive eigenvector  $(w_1, w_3)$ . To see that  $(w_1, w_3) = c(v_1, v_3)$  for some  $c > 0$  and hence  $\sigma = e^\lambda > 1$ , we note that the other eigenvector of  $\Gamma$  is perpendicular to  $(w_1, w_3)$  and hence has a negative component. The fact that the largest eigenvalue of  $\Gamma$  is bigger than 1 implies that the multitype branching process which occurs in the limit as  $t \rightarrow \infty$  is supercritical. The desired result now follows from the fact that the maximum eigenvalue of a positive matrix is a continuous function of its entries.  $\square$

To estimate the behavior of the two-type branching process, we will bound it below by a two-stage process which follows one law on  $[0, T_0]$  and another on  $[T_0, \infty)$ . Pick  $T_0$  and  $\delta > 0$  (a fudge factor we will need later) so that the multitype branching process which arises from assuming 2's have density  $\bar{\rho} = \rho(T_0) + \delta$  is supercritical. This is the process we will use on  $[T_0, \infty)$ . On  $[0, T_0]$  we will not allow new 1 infections to occur and subject the existing ones to death at rate  $\delta_1/c_1$ . We expect this to reduce the fraction of 1 infected sites to  $\exp(-T_0 \delta_1/c_1)$  times the original number at time  $T_0$ . However, we will be able to recover our losses in the second stage when the process is supercritical. This tactic was used in Durrett and Swindle (1991), but here the branching process has two types of particles.

To be able to recover our losses, we need estimates on the mean number of particles in the second stage branching random walk. For the next definition and for the proof of Lemma 4.3, we will suppose that the second stage law is used for all time. Let  $\mu_{ij}^t(x, D)$  be the mean number of individuals of type  $j$  in the set  $D$  at time  $t$  when we start with one individual of type  $i$  at  $x$ . Let  $\mu_i^t = \mu_{i1}^t + \mu_{i3}^t$  and let  $e_1 = (1, 0)$ .

LEMMA 4.3. Let  $J_t = [-\sqrt{t}, \sqrt{t}]^2$ :

$$\lim_{t \rightarrow \infty} \inf_i \inf_{x \in J_t} \mu_i^t(x, 2\sqrt{t}e_1 + J_t) = \infty.$$

PROOF. Intuitively this is true since  $\mu_{ij}^t(x, \mathbf{R}^2)$  grows exponentially and the central limit theorem implies that the fraction of the particles that land in the target set is bounded below uniformly in  $x \in J_t$  for large  $t$ . It takes some work to write down a formula that allows us to conclude this. However, we will give the details of the proof of this “standard result” here since we will also need a generalization of the situation in which births are not allowed outside  $(-4\sqrt{t}, 4\sqrt{t})$ .

Let  $b_{ij}$  be the rate at which type  $i$  particles give birth to type  $j$  particles without dying themselves. Consulting the rates given in Lemma 4.1 and recalling the definition of the second stage, we have

$$\begin{aligned} b_{11} &= \beta_1(1 - \bar{\rho}), & b_{31} &= c_1 \beta_1(1 - \bar{\rho}), \\ b_{13} &= c_1 \beta_1 \bar{\rho}, & b_{33} &= c_1^2 \beta_1 \bar{\rho}. \end{aligned}$$

In all the cases in the table the new offspring is displaced from the parent by an amount that is uniform on  $[-1, 1]^2$ . Particles can also change type without moving or die. Let  $c_{ij}$  be the rate at which particles of type  $i$  change into type  $j$  and let  $d_i$  be the death rate for particles of type  $i$ :

$$c_{13} = c_2 \beta_2 \bar{\rho}, \quad c_{31} = \delta_2/c_2, \quad d_1 = \delta_1, \quad d_3 = \delta_1/c_1.$$

To reduce things to discrete time, let

$$r_1 = b_{11} + b_{13} + c_{13} + d_1, \quad r_3 = b_{31} + b_{33} + c_{31} + d_3, \quad r = \max\{r_1, r_3\}.$$

If we let  $U$  be uniform on  $[-1, 1]^2$ , then a particle of type 1 at  $x$  waits for an exponential amount of time with mean  $1/r$  and then:

becomes	with probability
a 1 at $x$ , a 1 at $x + U$	$b_{11}/r$
a 1 at $x$ , a 3 at $x + U$	$b_{13}/r$
a 3 at $x$	$c_{13}/r$
a 1 at $x$	$1 - r_1/r$
dead, i.e., $\emptyset$	$d_1/r$ .

There is a completely analogous table for the actions of a type 3 particle. Let  $m_{ij}^1(x, D)$  be the mean number of  $j$  offspring in  $D$  produced in one transition by a particle of type  $i$  at  $x$ . Iterating we can define the mean result of  $n$  transitions inductively by

$$m_{ij}^n(x, D) = \sum_k \int m_{ik}^{n-1}(x, dy) m_{kj}^1(y, D)$$

and recover the mean for the continuous time process by

$$(4.11) \quad \mu_{ij}^t(x, D) = \sum_{n=0}^{\infty} e^{-rt} \frac{(rt)^n}{n!} m_{ij}^n(x, D).$$

Let  $\bar{m}_{ij}^n = m_{ij}^n(x, \mathbf{R}^2)$ , which does not depend on  $x$ , and  $p_{ij}^n(x, D) = m_{ij}^n(x, D)/\bar{m}_{ij}^n$ . Then  $\bar{m}^n$  is the  $n$ th power of the matrix  $\bar{m}$  and  $p^n$  is the  $n$ th power of the transition probability  $p_{ij}^1(x, D) = g_{ij}(D - x)$ .

Let  $\bar{Q}$  be the matrix which results if we take the definition of  $Q$  in Lemma 4.2 and replace  $(\beta_2 - \delta_2)/\beta_2$  by  $\bar{\rho}$  and  $\delta_2/\beta_2$  by  $1 - \bar{\rho}$ . Let  $\bar{\Gamma} = e^{\bar{Q}t}$ .  $\bar{\Gamma}_{ij}$  is the mean number of individuals of type  $j$  at time 1 when we start the second stage process with one individual to type  $i$ . Repeating the argument for Lemma 4.2, we see that by the Perron–Frobenius theorem  $\bar{m}_{ij}$  has a largest eigenvalue  $\kappa$  with a strictly positive eigenvector  $(w_1, w_3)$  which must be a positive multiple of the one for  $\bar{\Gamma}$ . Using (4.11) with  $D = \mathbf{R}^2$ , multiplying by  $w_j$  and summing over  $j$ , we see that if  $e^{\bar{\lambda}}$  is the largest eigenvalue of  $\bar{\Gamma}$  then

$$\exp(\bar{\lambda}t)(w_1, w_3) = (w_1, w_3) \sum_{n=0}^{\infty} \exp(-rt) \frac{(\kappa rt)^n}{n!} = \exp((\kappa - 1)rt)(w_1, w_3)$$

and hence  $\bar{\lambda} = (\kappa - 1)r$ . Noting that

$$\sum_{n=0}^{\infty} \exp(-rt) \frac{(\kappa rt)^n}{n!} = \exp((\kappa - 1)rt) \sum_{n=0}^{\infty} \exp(-\kappa rt) \frac{(\kappa rt)^n}{n!}$$

and applying the local central limit theorem to the Poisson distribution, we see that the  $n_t = \lceil \kappa rt \rceil$  term in the sum on the left is asymptotically  $Ct^{-1/2}$ . Applying the Perron–Frobenius theorem to  $\bar{m}$ , it follows that for all  $i, j$

$$(4.12) \quad \exp(-rt) \frac{(rt)^{n_t}}{n_t!} \bar{m}_{ij}^{n_t} \sim C_{ij} t^{-1/2} \exp(\bar{\lambda}t)$$

The transition probability  $p_{ij}^{n_t}$  is the result of  $n_t$  steps chosen from one of the four distributions  $g_{ij}$ , all of which have positive finite variance, so the central limit theorem implies

$$(4.13) \quad \liminf_{t \rightarrow \infty} \inf_{i, j} \inf_{x \in J_t} p_{ij}^{n_t}(x, 2\sqrt{t}e_1 + J_t) > 0.$$

Combining this with (4.12) gives the desired result.  $\square$

We are now almost ready for the block argument. Let  $\mathcal{L} = \{(m, n) \in \mathbf{Z}^2: m + n \text{ is even}\}$ , and make the following definitions in  $\mathbf{R}^2 \times [0, \infty)$ . The quantity  $T_0$  was chosen above and  $T_1$  will be chosen in a moment. Define

$$\phi(m, n) = (2mLe_1, nT) \quad \text{where } L = \sqrt{T_1} \text{ and } T = T_0 + T_1,$$

$$B = (-4L, 4L)^2 \times [0, T], \quad B_{m, n} = \phi(m, n) + B,$$

$$I = [-L, L]^2, \quad I_m = 2mLe_1 + I.$$



Consider a truncation of our two-stage branching random walk in which births are not allowed outside of  $(-4L, 4L)^2$ . The truncation does not affect the first stage since no births occur then. Our next step is to compute what the truncation does to the mean of the second stage process.

Let  $\hat{p}_{ij}(x, D) = p_{ij}(x, D \cap (-4L, 4L)^2)$  and let  $\hat{p}_{ij}^n(x, D)$  be the  $n$ th iterate. Let  $n_t = \lceil \kappa r t \rceil$  as in the proof of Lemma 4.3. The original transition probability  $p_{ij}^{n_t}$  is the result of  $n_t$  steps chosen from one of the four distributions  $g_{ij}$ , all of which have positive finite variance, so using Donsker's theorem instead of the central limit theorem in the proof of (4.13), we have

$$(4.14) \quad \liminf_{t \rightarrow \infty} \inf_{i,j} \inf_{x \in J_t} \hat{p}_{ij}^{n_t}(x, 2\sqrt{t}e_1 + J_t) > 0.$$

Let  $\hat{\mu}_{ij}^t(x, D)$  be the mean number of individuals of type  $j$  in the set  $D$  at time  $t$  when we start the modified branching random walk with one individual of type  $i$  at  $x$  and let  $\bar{\mu}_i^t = \hat{\mu}_{i1}^t + \hat{\mu}_{i3}^t$ . From (4.14) it follows easily that if  $T_1$  is large

$$(4.15) \quad \inf_i \inf_{x \in J_{T_1}} \mu_i^t(x, 2\sqrt{T_1}e_1 + J_{T_1}) \geq 4 \exp(T_0 \delta_1 / c_1).$$

To make the connection with the definitions above, notice that  $L = \sqrt{T_1}$ ,  $I = J_{T_1}$ , and  $I_1 = 2\sqrt{T_1}e_1 + J_{T_1}$ .

Let  $A \subset I$ , and let  $(\hat{\eta}_1^A(t), \hat{\eta}_3^A(t))$  be the sets of sites occupied by 1's and 3's in the truncated two stage branching process started from  $\hat{\eta}_1^A(0) = \emptyset$ ,  $\hat{\eta}_3^A(0) = A$ . Recall that we start with all sites occupied by 2's so at time 0, 1's are impossible. Let  $\hat{\eta}^A(t) = \hat{\eta}_1^A(t) \cup \hat{\eta}_3^A(t)$ .

LEMMA 4.4. *For any  $\varepsilon > 0$  we can pick  $K$  large enough so that for any  $A \subset I_0$  with  $|A| = K$ ,*

$$P(|\hat{\eta}^A(T) \cap I_1| \geq K) \geq 1 - \varepsilon.$$

PROOF. Let  $V$  be the number of 1 infections that survive the first stage.

$$EV = K \exp(-T_0 \delta_1 / c_1)$$

so (4.15) implies that  $E|\hat{\eta}^A(T) \cap I_1| \geq 2K$ . Since the progeny sets of different individuals are independent,

$$\text{var}(|\hat{\eta}^A(T) \cap I_1|) \leq KC_T,$$

where  $C_T < \infty$ . Using Chebyshev's inequality now we have

$$P(|\hat{\eta}^A(T) \cap I_1| < K) \leq \frac{\text{var}(|\hat{\eta}^A(T) \cap I_1|)}{K^2} \leq \frac{C_T}{K},$$

from which the desired result follows.  $\square$

Our next step is to extend Lemma 4.4 to the particle system with large range. Before doing this, however, we have to impose one last truncation to get a finite range of dependence in the block events to be defined below. For

$t \leq T_0$ , we pretend that all sites are always occupied by 2's. For  $t \geq T_0$ , we put a 2 at  $x$  at time  $t$ , if "the dual of the long-range contact process of 2's," starting at  $x$  at time  $t$  survives for  $T_0$  units of time or escapes from the box of radius  $L = \sqrt{T_1}$  centered at  $x$  by time  $T_0$ . To explain what we mean by the process in quotes, note that if we use Poisson processes to construct the process in the usual way [see Section 2 of Durrett (1995)], then this allows us to construct on the same space the process starting from any initial configuration. In particular we can construct the process starting from no 1's when it reduces to the long-range contact process and construct in the usual way [see Section 3 of Durrett (1995)] the dual of that process.

If  $T_1$  is chosen large enough, then the probability that the dual escapes from the box of radius  $L = \sqrt{T_1}$  is smaller than the fudge factor  $\delta$  introduced in the definition of  $\bar{\rho}$  given before Lemma 4.3. Let  $A \subset I$ , and let  $(\hat{\eta}_1^{r,A}(t), \hat{\eta}_3^{r,A}(t))$  denote the set of sites occupied by 1's and 3's in the particle system on  $\mathbf{Z}^2/r$  started from  $\hat{\eta}_1^{r,A}(0) = \emptyset$ ,  $\hat{\eta}_3^{r,A}(0) = A$  and truncated so that 1 infections are not allowed outside  $(-4L, 4L)^2$ . Let  $\hat{\eta}^{r,A}(t) = \hat{\eta}_1^{r,A}(t) \cup \hat{\eta}_3^{r,A}(t)$ .

LEMMA 4.5. *For any  $\varepsilon > 0$  we can pick  $K$  and then  $r_0$  large enough so that if  $r \geq r_0$  then for any  $A \subset I_0$  with  $|A| = K$ ,*

$$P(|\hat{\eta}^{r,A}(t) \cap I_1| \geq K) \geq 1 - 2\varepsilon.$$

PROOF. Given  $\varepsilon > 0$ , pick  $K$  as in Lemma 4.4. Suppose now that there is a sequence  $r_n \rightarrow \infty$  and sets  $A_n = \{x_1^n, \dots, x_K^n\}$  so that the conclusion is false. By considering a subsequence  $n_i$  we can have  $x_j^{n_i}$  converging to a limit  $x_j$  for  $1 \leq j \leq K$ . A straightforward generalization of the proof of Lemma 4.1 implies that if we let  $(\tilde{\eta}_1^i, \tilde{\eta}_3^i)$  denote the processes with  $r = r_{n_i}$  and  $A = A_{n_i}$  then  $(\tilde{\eta}_1^i, \tilde{\eta}_3^i)$  converges weakly to the truncated two stage process.

The probability of a particle landing exactly on the boundary of  $(-4L, 4L)^2$  or on the boundary of one of the  $I_i$ ,  $i = -1, 1$  is 0 so  $P(|\tilde{\eta}_1^i(T) \cup \tilde{\eta}_3^i(T)| \geq K)$  converges to the probability for the limit process starting from  $\{x_1, \dots, x_K\}$ . Lemma 4.4 implies that this probability is greater than or equal to  $1 - \varepsilon$ , so we have a contradiction which proves the result.  $\square$

Lemma 4.5 shows that for large range, the particle system satisfies the comparison assumptions on page 140 of Durrett (1995). Invoking Theorem 4.3 on page 141 of that paper now we see that if  $\varepsilon$  is chosen small enough in Lemma 4.5 and we let  $X_n = \{m: (m, n) \in \mathcal{L}_0 \text{ and the number of 1 infections in } I_m \text{ at time } nT \text{ is } \geq K\}$ , then  $X_n$  dominates a supercritical 2-dependent oriented percolation process with initial set  $W_0 = X_0$ .

For the initial configuration given in Theorem 4,  $W_0$  is product measure with density  $p$  so Theorem A.3 on page 195 of Durrett (1995) implies

$$\liminf_{n \rightarrow \infty} P(\{-2K, \dots, 2K\} \cap W_{2n} \neq \emptyset) \geq 1 - \varepsilon_K$$

with  $\varepsilon_K \rightarrow 0$  as  $K \rightarrow \infty$ . The proof of Theorem 4 can now be completed by combining this observation with the standard techniques of taking the Cesaro

average of the distribution of  $\xi_s$ ,  $0 \leq s \leq t$  and extracting a convergent subsequence to construct a stationary distribution for the process [see (2.13) on page 124 of Durrett (1995)].  $\square$

## REFERENCES

- BENGTSSON, J. (1991). Interspecific competition in metapopulations. *Biol. J. of the Linnean Soc.* **42** 219–237.
- BRAMSON, M. (1989). Survival of nearest particle systems with low birth rate. *Ann. Probab.* **17** 433–443.
- BRAMSON, M. and DURRETT, R. (1988). A simple proof of the stability criterion of Gray and Griffeath. *Probab. Theory Related Fields* **80** 293–298.
- BRAMSON, M., DURRETT, R. and SWINDLE, G. (1989). Statistical mechanics of crabgrass. *Ann. Probab.* **17** 444–481.
- BRAMSON, M. and GRIFFEATH, D. (1980). On the Williams–Bjerknes tumor growth model, II. *Math. Proc. Cambridge Philos. Soc.* **88** 339–357.
- BRAMSON, M. and GRIFFEATH, D. (1981). On the Williams–Bjerknes tumor growth model. I. *Ann. Probab.* **9** 173–185.
- BRAMSON, M. and NEUHAUSER, C. (1993). Survival of one-dimensional cellular automata under random perturbations. *Ann. Probab.* **21** 244–263.
- DUFFUS, J. E., FALK, B. W. and JOHNSTONE, G. R. (1987). Luteoviruses—one system, many variations. In *World Perspectives on Barley Yellow Dwarf* (P. A. Burnett, ed.) 86–104. Proceedings of the International Workshop 1987, Udine, Italy.
- DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth and Brooks/Cole, Belmont, CA.
- DURRETT, R. (1991). A new method for proving the existence of phase transitions. In *Spatial Stochastic Processes* 141–169. Birkhäuser, Boston.
- DURRETT, R. (1992a). Multicolor particle systems with large threshold and range. *J. Theoret. Probab.* **5** 127–152.
- DURRETT, R. (1992b). Predator-prey systems. In *Asymptotic Problems in Probability Theory: Stochastic Models and Diffusions on Fractals* (K. D. Elworthy and N. Ikeda, eds.) 37–58. Longman, Essex, England.
- DURRETT, R. (1995). Ten lectures on particle systems. *Lectures on Probability Theory (Saint-Flour, 1993). Lecture Notes in Math.* **1608** 97–201. Springer, Berlin.
- DURRETT, R. and LEVIN, S. A. (1994a). Stochastic spatial models: a user’s guide to ecological applications. *Philos. Trans. Roy. Soc. Ser. B* **343** 329–350.
- DURRETT, R. and LEVIN, S. A. (1994b). The importance of being discrete (and spatial). *Theoret. Population Biol.* **46** 363–394.
- DURRETT, R. and NEUHAUSER, C. (1994). Particle systems and reaction–diffusion equations. *Ann. Probab.* **22** 289–333.
- DURRETT, R. and SCHINAZI, R. (1993). Asymptotic critical value for a competition model. *Ann. Appl. Probab.* **3** 1047–1066.
- DURRETT, R. and SWINDLE, G. (1991). Are there bushes in a forest? *Stochastic Process. Appl.* **37** 19–31.
- FALK, B. W. and DUFFUS, J. E. (1981). Epidemiology of helper-dependent persistent aphid transmitted virus complexes. In *Plant Diseases and Vectors* (K. Harris and K. Maramorosch, eds.) 161–179. Academic Press, New York.
- GALLOWAY, B. T. and SOUTHWOOD, E. A. (1890). Preliminary notes on a new and destructive disease. *Journal of Mycology* **6** 72–73.
- GILDOW, F. E. (1987). Barley yellow dwarf. Virus-aphid vector interactions associated with virus transmission and vector specificity. In *World Perspectives on Barley Yellow Dwarf* (P. A. Burnett, ed.) 111–122. Proceedings of the International Workshop 1987, Udine, Italy.

- HANSKI, I. and RANTA, E. (1983). Coexistence in a patchy environment: three species of *Daphnia* in rock pools. *J. Animal Ecology* **52** 263–279.
- HARRIS, T. E. (1972). Nearest neighbor Markov interaction processes. *Adv. Math.* **9** 66–89.
- KUCZEK, T. (1989). The central limit theorem for the right edge of supercritical percolation. *Ann. Probab.* **17** 1322–1332.
- LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.
- NEUHAUSER, C. (1992). Ergodic theorems for the multitype contact process. *Probab. Theory Related Fields* **91** 467–506.
- OSWALD, J. W. and HOUSTON, B. R. (1951). A new virus disease of cereals, transmissible by aphids. *Plant Disease Reporter* **35** 471–475.
- ROCHOW, W. F. (1965). Apparent loss of vector specificity following double infection by two strains of barely yellow dwarf virus. *Phytopathology* **55** 62–68.
- ROCHOW, W. F. (1979). Field variants of barley yellow dwarf virus: detection and fluctuation during twenty years. *Phytopathology* **69** 655–660.
- SENETA, E. (1973). *Nonnegative Matrices*. Wiley, New York.
- SLATKIN, M. (1974). Competition and regional coexistence. *Ecology* **55** 128–134.
- WATERHOUSE, P. M., GILDOW, F. E. and JOHNSTONE, G. R. (1988). Luteovirus group, description of plant viruses. No. 339, Commonwealth Mycological Institute/Association of Applied Biologists.
- WEN, F., LISTER, R. M. and FATTOUH, F. A. (1991). Cross-protection among strains of barley yellow dwarf virus. *Journal of General Virology* **72** 791–799.

DEPARTMENT OF MATHEMATICS  
CORNELL UNIVERSITY  
WHITE HALL  
ITHACA, NEW YORK 14853  
E-MAIL: rtd1@cornell.edu

SCHOOL OF MATHEMATICS  
UNIVERSITY OF MINNESOTA  
127 VINCENT HALL  
MINNEAPOLIS, MINNESOTA 55455  
E-MAIL: nhauser@math.umn.edu