The Annals of Probability
1994, Vol. 22, No. 1, 289–333

PARTICLE SYSTEMS AND REACTION–DIFFUSION EQUATIONS[1]

BY R. DURRETT AND C. NEUHAUSER

Cornell University and University of Southern California

In this paper we will consider translation invariant finite range particle systems with state space $\{0, 1, \dots, \kappa - 1\}^S$ with $S = \varepsilon\mathbb{Z}^d$. De Masi, Ferrari and Lebowitz have shown that if we introduce stirring at rate ε^{-2}, then the system converges to the solution of an associated reaction diffusion equation. We exploit this connection to prove results about the existence of phase transitions when the stirring rate is large that apply to a wide variety of examples with state space $\{0, 1\}^S$.

1. Introduction. The point of this paper is to describe a general method for proving the existence of phase transitions in interacting particle systems in which the particles are stirred at a fast rate. To be precise, we consider processes $\xi_t^\varepsilon\colon \varepsilon\mathbb{Z}^d \to \{0, 1, \dots, \kappa - 1\}$ that evolve as follows:

(i) There are *translation invariant finite range flip rates*

$$c_i(x, \xi) = h_i(\xi(x), \xi(x + \varepsilon y_1), \dots, \xi(x + \varepsilon y_N))$$

so that if $\xi_t(x) \neq i$, then

$$P(\xi_{t+s}(x) = i|\xi_t) \sim c_i(x, \xi_t)s$$

as $s \to 0$, where $f(s) \sim g(s)$ means $f(s)/g(s) \to 1$. In words, $c_i(x, \xi)$ is the rate that site x changes to state i when the configuration is ξ. The values of $c_i(x, \xi)$ on $\{\xi\colon \xi(x) = i\}$ are not relevant to the definition so we set them equal to 0.

(ii) For each $x, y \in \varepsilon\mathbb{Z}^d$ with $\|x - y\|_1 = \varepsilon$ we exchange the values at x and y at rate $\varepsilon^{-2}/2$.

Here $\|z\|_1 = |z_1| + \cdots + |z_d|$. We will also use $\|z\|_\infty = \sup_i |z_i|$ and $|z| = (z_1^2 + \cdots + z_d^2)^{1/2}$.

The reader should note that in (i), changing ε scales the lattice but does not change the interaction between the sites. In (ii), we superimpose "stirring" in such a way that the individual values will be moving according to Brownian motions in the limit. The motivation for modifying the system in this way comes from the following *mean field limit theorem* of De Masi, Ferrari and Lebowitz (1986).

Received January 1992; revised November 1992.

[1]The research of both authors was partially supported by the National Science Foundation and the Army Research Office through the Mathematical Sciences Institute at Cornell University.

AMS 1991 *subject classifications*. Primary 60K35; secondary 35K35.

Key words and phrases. Contact process, sexual reproduction model, mean field limit theorem, hydrodynamic limit, reaction diffusion equation, metastability.

THEOREM 1. *Suppose $\xi_0^\varepsilon(x)$ are independent and let $u_i^\varepsilon(t,x) = P(\xi_t^\varepsilon(x) = i)$. If $u_i^\varepsilon(0,x) = g_i(x)$ is continuous, then as $\varepsilon \to 0$, $u_i^\varepsilon(t,x) \to u_i(t,x)$ the bounded solution of*

$$(1.1) \qquad \partial u_i/\partial t = \tfrac{1}{2}\Delta u_i + f_i(u), \qquad u_i(0,x) = g_i(x),$$

where

$$(1.2) \qquad f_i(u) = \langle c_i(0,\xi)\rangle_u - \sum_{j\neq i}\langle c_j(0,\xi)1_{(\xi(0)=i)}\rangle_u$$

and $\langle\phi(\xi)\rangle_u$ denotes the expected value of $\phi(\xi)$ under the product measure in which state j has density u_j, that is, when $\xi(x)$ are i.i.d. with $P(\xi(x) = j) = u_j$.

Theorem 1 is easy to understand. The stirring mechanism [i.e., (ii)] has product measures as its stationary distributions. When ε is small, stirring operates at a fast rate and keeps the system close to a product measure. The rate of change of the densities can then be computed assuming adjacent sites are independent. In many investigations, the proof of results like Theorem 1 is the end of the story. Here, we view it as a starting point for investigating the stationary distribution of the particle system with fast stirring. The techniques we will use can be applied to a wide variety of examples. In this paper we will confine our attention to the special case in which $\kappa = 2$ and we have the following:

(a) when $\xi(x) = 1$, $c_0(x,\xi) = a(\xi(x+\varepsilon y_1),\dots,\xi(x+\varepsilon y_N))$;
(b) when $\xi(x) = 0$, $c_1(x,\xi) = b(\xi(x+\varepsilon y_1),\dots,\xi(x+\varepsilon y_N))$;
(c) if $u, v \in \{0,1\}^N$ with $u_i \le v_i$, then $a(u) \ge a(v)$ and $b(u) \le b(v)$;
(d) $b(0,\dots,0) = 0$, $a(1,\dots,1) > 0$.

Assumption (d) implies that $\xi(x) \equiv 0$ is an absorbing state so δ_0, the point mass on the $\equiv 0$ configuration, is a trivial stationary distribution. The basic question we would like to answer about our systems is: "When do nontrivial stationary distributions exist?" There are some simple general results that can help us answer this question. [See Liggett (1985) or Durrett (1988).] Assumption (c) implies that the system is *attractive*, that is, if $\xi_0(x) \le \xi_0'(x)$ for all x, then we can define the two processes on the same space so that $\xi_t(x) \le \xi_t'(x)$ for all t, x. An immediate consequence is that if we start from $\bar{\xi}_0^\varepsilon(x) \equiv 1$, then as $t \to \infty$, $\bar{\xi}_t^\varepsilon$ converges weakly to $\bar{\xi}_\infty^\varepsilon$, a stationary distribution that is translation invariant. Since we have started from the largest possible initial state, the limit $\bar{\xi}_\infty^\varepsilon$ is the largest stationary distribution. Of course, we may have $\bar{\xi}_\infty^\varepsilon = \delta_0$ and in this case there is no nontrivial stationary distribution.

Before trying to answer our basic question, it is useful to look at some examples. Since we have taken $\kappa = 2$, $u_0 = 1 - u_1$. So if we let $u = u_1$, we have a single equation:

$$(1.3) \qquad \partial u/\partial t = \tfrac{1}{2}\Delta u + f(u).$$

In describing the examples it is useful to think of 0's as vacant sites and 1's as occupied sites. In all cases we will have $a \equiv 1$. That is, particles die at rate 1 independent of their surroundings.

EXAMPLE 1 (The basic contact process). Here $b(u) = \lambda \sum_i u_i$, that is, the birth rate is proportional to the number of occupied neighbors. In this case if we let $\beta = N\lambda$, then

$$f(u) = -u + (1-u)\beta u.$$

EXAMPLE 2 (Threshold contact process). Here $b(u) = \beta$ if $\sum_i u_i > 0$, that is, the birth rate is β if at least one neighbor is occupied. In this case

$$f(u) = -u + (1-u)\beta\{1 - (1-u)^N\}.$$

EXAMPLE 3 (Sexual reproduction). See Durrett and Gray (1985) and Chen (1992). In these processes one is given a list of pairs of neighbors (y_i, z_i), $1 \le i \le l$, and one has

$$c_1(x, \xi) = \lambda \sum_{i=1}^{l} 1_{(\xi(x+y_i)=1,\ \xi(x+z_i)=1)} \tag{1.4}$$

when $\xi(x) = 0$. That is, each occupied pair contributes a birth rate λ. One possibility in $d = 2$ is to take $l = 4$ and use the pairs $\{(\pm 1, 0), (0, \pm 1)\}$. In any case, if we let $\beta = l\lambda$, then

$$f(u) = -u + (1-u)\beta u^2.$$

EXAMPLE 4 (Autocatalytic reactions). Dickman and Tomé (1991) have simulated the following systems in $d = 1$ when $k = 2, 3$:

$$c_1(x, \xi) = \lambda\{1_{(\xi(x-i)=1 \text{ for } 1 \le i \le k)} + 1_{(\xi(x+i)=1 \text{ for } 1 \le i \le k)}\}$$

when $\xi(x) = 0$. In words, a string of particles of length at least k produces a new particle at rate 2λ and the new particle appears at a randomly chosen end. When $k = 2$ this is a special case of sexual reproduction. One can generalize this system to include Example 3 by writing down a list of k-tuples $(y_i^1, \dots, y_i^k)$, $1 \le i \le l$, and defining birth rates by imitating (1.4). In this case if we let $\beta = l\lambda$, then

$$f(u) = -u + (1-u)\beta u^k.$$

If we set $k = 1$, we get the equation in Example 1, but when we refer to Example 4 we will always suppose that $k \ge 2$.

EXAMPLE 5 (Quadratic birth rate). Combining Examples 1 and 3 we can define a two-parameter family of systems in which

$$b(u) = \lambda_1 n + \lambda_2 \binom{n}{2}, \quad \text{where } n = \sum_i u_i.$$

Letting $\alpha = N\lambda_1$ and $\beta = \binom{N}{2}\lambda_2$, we have

$$f(u) = -u + \alpha u(1-u) + \beta u^2(1-u).$$

In each of the first four examples there is one parameter β and increasing β makes it easier to have a nontrivial stationary distribution. The last observation motivates the definition of a critical value

$$\beta_c(\varepsilon) = \inf\{\beta: \xi_\infty^\varepsilon \neq \delta_0\}.$$

To guess the limiting behavior of $\beta_c(\varepsilon)$ as $\varepsilon \to 0$, we note that Theorem 1 implies that if we let $\varepsilon \to 0$, then $P(\bar{\xi}_t^\varepsilon(x) = 1) \to v(t)$, the solution of $v'(t) = f(v(t))$ with $v(0) = 1$. As $t \to \infty$, $v(t)$ decreases to ρ_0, the largest root of $f(u) = 0$ in $[0,1]$, so by interchanging our two limits ($\varepsilon \to 0$, $t \to \infty$), we conclude that if ε is small, $P(\bar{\xi}_\infty^\varepsilon(x) = 1)$ is close to the largest root of $f(\rho) = 0$ in $[0,1]$. As we will see below this reasoning (which we will refer to as *mean field theory*) is correct in Examples 1 and 2 but not in Examples 3 and 4.

EXAMPLE 1 (The basic contact process). Here $f(u) = -u + \beta u(1-u)$. The roots are 0 and $\rho_0 = (\beta - 1)/\beta$, the latter being positive only if $\beta > 1$.

THEOREM 2. *As $\varepsilon \to 0$, $\beta_c(\varepsilon) \to 1$. If $\beta > 1$, then*

$$P\big(\bar{\xi}_\infty^\varepsilon(x) = 1\big) \to \rho_0 \quad \textit{as } \varepsilon \to 0.$$

EXAMPLE 2 (The threshold contact process). In this case

$$f(u) = -u + (1-u)\beta\{1 - (1-u)^N\},$$
$$f'(u) = -1 - \beta\{1 - (1-u)^N\} + N\beta(1-u)^N,$$
$$f''(u) = -\beta(N + N^2)(1-u)^{N-1},$$

so f is strictly concave on $[0,1)$. If $\beta \le 1/N$, then $f'(0) \le 0$ and there are no positive roots, while if $\beta > 1/N$, there is exactly one root $\rho_0 \in (0,1)$. [Notice $f(1) = -1$.]

THEOREM 3. *As $\varepsilon \to 0$, $\beta_c(\varepsilon) \to 1/N$. If $\beta > 1/N$, then*

$$P\big(\bar{\xi}_\infty^\varepsilon(x) = 1\big) \to \rho_0 \quad \textit{as } \varepsilon \to 0.$$

EXAMPLE 3 (Sexual reproduction). Here $f(u) = -u + (1-u)\beta u^2$. We always have $u = 0$ as a root and if $\beta \ge 4$, we have two other roots $\rho_1 \le \rho_0$ given by $(1 \pm \sqrt{1 - 4/\beta})/2$.

THEOREM 4. *As $\varepsilon \to 0$, $\beta_c(\varepsilon) \to 4.5$. If $\beta > 4.5$, then*

$$P\big(\bar{\xi}_\infty^\varepsilon(x) = 1\big) \to \rho_0 \quad \textit{as } \varepsilon \to 0.$$

The first question we should answer is "Why 4.5 instead of 4?" This can be seen from two different viewpoints. The first, due to Schlögl (1972), is to rewrite $f(x) = -V'(x)$, that is, the force is minus the gradient of the potential function V, and observe that for $\beta < 4.5$, 0 is the global minimum of V, while for $\beta > 4.5$, ρ_0 is the global minimum. The last analogy makes it possible to guess the answer but to prove it we have to use a more complicated approach, which has the additional advantage that it generalizes to systems in which the nonlinearity $(f_1(x), \dots, f_n(x))$ does not happen to be the gradient of a function.

Suppose for the moment that $d = 1$. The equation

$$\partial u/\partial t = \tfrac{1}{2}\Delta u + f(u) \tag{1.5}$$

has travelling wave solutions $u(t, x) = w(x - rt)$, where w is a decreasing solution of the ordinary differential equation

$$-rw' = \tfrac{1}{2}w'' + f(w)$$

with $w(\infty) = 0$ and $w(-\infty) = \rho_0$. [See Section 2 of Fife and McLeod (1977) for this and other facts we need about travelling waves.] Multiplying both sides of the last equation by w' and integrating from $-\infty$ to ∞ gives

$$\begin{aligned} -r\int w'(x)^2\,dx &= \tfrac{1}{2}\int w'(x)w''(x)\,dx + \int f(w(x))w'(x)\,dx \\ &= 0 - \int_0^{\rho_0} f(y)\,dy \end{aligned} \tag{1.6}$$

after changing variables $y = w(x)$ in the second integral, since $((w')^2/2)' = w'w''$ and $w'(x) \to 0$ as $|x| \to \infty$. The last equation implies that the sign of the speed r is the same as that of the integral of f from 0 to ρ_0. When $\beta = 4.5$, $\rho_1 = \rho_0/2$ and symmetry implies that the integral is 0. A little calculus shows that $r < 0$ when $\beta < 4.5$ and $r > 0$ when $\beta > 4.5$.

From the last calculation and the statement of Theorem 4, the reader can probably leap to the conclusion that

$$\beta_c(\varepsilon) \to \beta_1 \equiv \inf\{\beta : r(\beta) > 0\}. \tag{1.7}$$

To see that this is consistent with Theorems 2 and 3, we note that in Examples 1 and 2 there are travelling waves with $w(-\infty) = \rho_0$ and $w(\infty) = 0$ when $\beta > 1$ [see, e.g., Aronson and Weinberger (1975)]. In this case the calculation above implies $r > 0$. In a moment we will explain the significance of the sign of r for the proofs of our theorems. First, we want to use (1.7) to compute the asymptotic behavior of the critical value in the two remaining examples.

EXAMPLE 4 (Autocatalytic reactions). $f(u) = -u + (1 - u)\beta u^k$. As always 0 is a root. Factoring out the trivial root, we have to solve $\phi(u) \equiv u^{k-1} - u^k = 1/\beta$. Now $\phi'(u) = (k-1)u^{k-2} - ku^{k-1}$, so if we let $u_0 = (k-1)/k$, then ϕ is increasing on $[0, u_0]$ and decreasing on $[u_0, 1]$. When $\beta = \beta_0 \equiv 1/\phi(u_0)$ the equation $\phi(u) = 1/\beta$ has a double root at u_0, while for $\beta > \beta_0$ we have exactly two roots $\rho_1 < \rho_0$ in $[0, 1]$. It is comforting to note that when $k = 2$, $u_0 = 1/2$ and $\beta_0 = 4$. When $k = 3$, $u_0 = 2/3$ and $\beta_0 = 27/4$.

To calculate the asymptotic behavior of $\beta_c(\varepsilon)$, we have to find the value of β where

$$(1.8)\qquad 0 = \int_0^{\rho_0} \left(-x + \beta x^k - \beta x^{k+1}\right) dx = -\frac{\rho_0^2}{2} + \frac{\beta\rho_0^{k+1}}{k+1} - \frac{\beta\rho_0^{k+2}}{k+2}.$$

Since ρ_0 is a root we also have

$$(1.9)\qquad -\rho_0 + \beta\rho_0^k - \beta\rho_0^{k+1} = 0.$$

Dividing (1.8) by $-\rho_0/2$ and then adding to (1.9) gives

$$\beta\left\{\rho_0^k\left(1 - \frac{2}{k+1}\right) - \rho_0^{k+1}\left(1 - \frac{2}{k+2}\right)\right\} = 0,$$

so $\rho_0 = (k-1)(k+2)/(k(k+1))$ and the corresponding value of β is $\beta_1 \equiv 1/(\rho_0^{k-1}(1-\rho_0))$. Again, it is comforting to note that when $k = 2$, $\rho_0 = 4/6$ and $\beta_1 = 9/2$. When $k = 3$, $\rho_0 = 5/6$ and $\beta_1 = 6^3/5^2 = 8.64$.

THEOREM 5. *As $\varepsilon \to 0$, $\beta_c(\varepsilon) \to \beta_1$. If $\beta > \beta_1$, then*

$$P\left(\bar{\xi}_\infty^\varepsilon(x) = 1\right) \to \rho_0 \quad as\ \varepsilon \to 0.$$

REMARK. It is interesting to compare the results of Theorem 5 with the simulations in Dickman and Tomé. When $k = 2$ and 95% of the steps are devoted to stirrings they find $\beta_c \approx 5.15$, in contrast to our limit of 4.5. When $k = 3$ and 85% of the steps are stirrings they find $\beta_c \approx 10.415$ in contrast to our limit of 8.64. We mention these results not to cast doubt on the simulations but rather to indicate how slowly $\beta_c(\varepsilon)$ converges to its limit. In contrast, Dab, Lawniczak, Boon and Kapral (1990) have a nonrigorous scheme of approximating reaction diffusion equations by cellular automata that reproduces the limiting critical values almost exactly.

EXAMPLE 5 (Quadratic birth rate). The first step is to determine when $f(u) = 0$ has a nontrivial solution. Factoring out the trivial root gives $0 = -1 + \alpha(1-u) + \beta u(1-u)$ or $\beta u^2 + (\alpha - \beta)u + (1-\alpha) = 0$, which has roots

$$u = \frac{(\beta - \alpha) \pm \sqrt{(\alpha - \beta)^2 - 4\beta(1-\alpha)}}{2\beta}.$$

When $\alpha > 1$, the contact process part of the birth rate is supercritical, the positive square root is larger than $|\beta - \alpha|$ and there is exactly one positive root. When $\alpha \le 1$, we have two roots $\rho_1 < \rho_0$ in $(0,1)$ if and only if $\beta > \alpha$ and the expression under the square root is positive, that is,

$$(1.10)\qquad \beta > 2 - \alpha + \sqrt{4 - 4\alpha}.$$

However, from our analysis of Example 3, we see that $r > 0$ if and only if

$\rho_0 > 2\rho_1$, that is,

$$(\alpha - \beta)^2 - 4\beta(1 - \alpha) > (\beta - \alpha)^2/9$$

or, solving the quadratic equation,

$$\beta > \left(9 - 5\alpha + \sqrt{81 - 90\alpha + 9\alpha^2}\right)/4. \tag{1.11}$$

It is comforting to note that when $\alpha = 0$, the right-hand sides of (1.10) and (1.11) are 4 and 4.5, respectively. At the other end it is interesting that when $\alpha = 1$ both of these quantities are 1. Let $\mathscr{A}$ (for above) be the set of (α, β) so that $\alpha > 1$ or $\alpha \in [0, 1]$ and (1.11) holds. Let $\mathscr{B} = [0, \infty)^2 - \mathscr{A}$. See Figure 1 for a picture. The constraint in (1.11) is indicated by a solid line and in (1.10) by a dotted one.

THEOREM 6. *If* $(\alpha, \beta) \in \mathscr{A}$, *then* $P(\bar{\xi}_\infty^\varepsilon(x) = 1) \to \rho_0$ *as* $\varepsilon \to 0$. *If* $(\alpha, \beta) \in \mathscr{B}$, *then* $P(\bar{\xi}_\infty^\varepsilon(x) = 1) = 0$ *for small* ε.

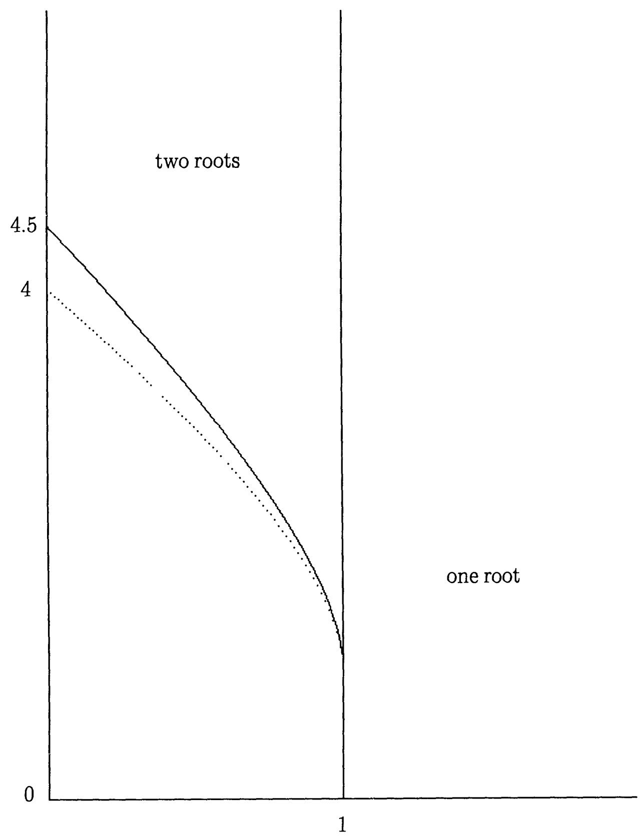

FIG. 1.

Theorems 2–6 follow from the fact that we can determine the asymptotic behavior of $P(\bar{\xi}^{\varepsilon}_{\infty}(x) = 1)$ if we impose the following assumptions in addition to (a)–(d) above. Let $0 = z_1 < \cdots < z_M$ be the roots of f in [0, 1]. [Assumptions (a) and (b) imply f is a polynomial of degree at most $N + 1$ so $M \le N + 1$.]

(e) If $0 \le i \le M$, then $f'(z_i) \neq 0$.
(f) If $0 \le i < j \le M$, then $\int_{z_i}^{z_j} f(u)\,du \neq 0$.
(g) $M \le 3$.

By proving suitable generalizations of Theorems 2.7 and 3.3 in Fife and McLeod (1977) one can drop assumption (g). However, this paper is already quite lengthy so we will give the details in a subsequent publication.

Turning to the proofs, we will now sketch the arguments in the special case in which $M = 3$. As the reader will see from the proofs, the details are similar but simpler when $M \le 2$. The keys to our analysis are two facts about the solution of the limiting p.d.e. that can be proved using the methods of Fife and McLeod (1977). (They proved their results in $d = 1$. The extension to $d > 1$ is carried out in the appendix.) Here r is the wave speed defined above, and $0 < \rho_1 < \rho_0 < 1$ are the two nonzero roots.

Proposition 1. *Suppose $r > 0$, $u(0, x) \ge 0$ for all x, and $u(0, x) \ge \rho_1 + \delta$ when $|x| \le L$. If $L \ge L_\delta$ and $\eta > 0$, then there are constants $0 < c, C < \infty$ so that*

$$u(t, x) \ge \rho_0 - Ce^{-ct} \quad \text{for } |x| \le (r - \eta)t.$$

In words, if the initial configuration is large enough, then inside a ball that grows at rate less than r the density is not much smaller than ρ_0.

Proposition 2. *Suppose $r < 0$, $u(0, x) \le 1$ for all x, and $u(0, x) \le \rho_1 - \delta$ when $|x| \le L$. If $L \ge L_\delta$ and $\eta > 0$, then there are constants $0 < c, C < \infty$ so that*

$$u(t, x) \le Ce^{-ct} \quad \text{for } |x| \le (-r - \eta)t.$$

In words, a large enough hole grows linearly even if all sites outside are occupied. The exponential rate of convergence to 0 should not be surprising since $f'(0) < 0$ and hence for small u, $f(u) \le (f'(0)/2)u$.

Let $\rho_\varepsilon = P(\bar{\xi}^{\varepsilon}_{\infty}(x) = 1)$. There are three things to show.

(I) $\limsup_{\varepsilon \to 0} \rho_\varepsilon \le \rho_0$.

This is a simple consequence of Theorem 1. See Section 3.

(II) If $r > 0$, then $\liminf_{\varepsilon \to 0} \rho_\varepsilon \ge \rho_0$.

To get this from Proposition 1 we use a method that is simple, at least in concept. Pick $\delta < (\rho_0 - \rho_1)/10$ and $L \ge L_\delta$. Proposition 1 shows that if we start the p.d.e. with density at least $\rho_1 + \delta$ in $[-L, L]^d$, then at some time T we will have density at least $\rho_0 - \delta > \rho_1 + \delta$ in $2Le_1 + [-L, L]^d$ and in $-2Le_1 + [-L, L]^d$, where $e_1 = (1, 0, \dots, 0)$. Theorem 1, when supplemented by a second moment computation, shows that if ε is small, then the particle system will do this with high probability. Then the existence of a stationary distribution with density close to ρ_0 follows by comparing with a mildly dependent oriented percolation process in which sites are open with probability close to 1.

REMARK. Noble (1989) proved (II) for Example 3 in his thesis. The proof we give in Section 3 is similar but incorporates one substantial improvement: by looking at "reasonable" sites we avoid studying the p.d.e. in a finite interval with Dirichlet boundary conditions.

(III) If $r < 0$, then $\rho_\varepsilon = 0$ for small ε.

The proof of (III) follows the outline of the proof of (II) but requires more work. The first and most difficult step is to show that the second moment computation referred to above works at times $t \le \theta \log(1/\varepsilon)$ for some $\theta > 0$. Proposition 2 then implies that if L is large and we start with density at most $\rho_1/2$ in $[-L, L]^d$, then at time $\theta \log(1/\varepsilon)$ we will have density at most ε^λ in a very large cube. Once the density gets this low, we can ignore the times at which two particles are in the neighborhood of some point and the system behaves like a branching process in which the birth rate minus the death rate is $f'(0) < 0$. Using the last observation and some simple estimates on random walk and subcritical branching processes, we can show that at time $C \log(1/\varepsilon)$ we have a large vacant cube. Putting the pieces together, one concludes for suitable L and T that if initially there are no particles in $[-L, L]^d$, then at time T there are with high probability no particles in $[-3L, 3L]^d$, and the desired result follows from another comparison with a mildly dependent oriented percolation.

Before embarking on the proofs, we would like to mention two open problems. In Example 3, or more generally in Example 4, $\rho_\varepsilon(\beta) \equiv P(\bar{\xi}^\varepsilon_\infty(x) = 1)$ converges to a discontinuous function of β, so it is natural to ask if ρ_ε is discontinuous when ε is small. Dickman and Tomé claim that for the one-dimensional systems they investigated the answer to this question is "No" when $k = 2$ and "Yes" when $k = 3$. We disagree with their conclusion.

CONJECTURE 1. *For the systems in Example* 4, ρ_ε *is always continuous in* $d = 1$.

HEURISTIC ARGUMENT. Let $\hat{\xi}^\varepsilon_t$ be the system starting with $\hat{\xi}^\varepsilon_0(x) = 1$ if and only if $x \le 0$. Let $r^\varepsilon_t = \sup\{x\colon \hat{\xi}^\varepsilon_t(x) = 1\}$. It follows from results in Durrett (1980) that $\lim_{t\to\infty} r^\varepsilon_t/t = \alpha^\varepsilon(\beta)$ exists almost surely. Based on results in that paper and the discussion after Theorem 4, it is reasonable to guess that

$$(1.12) \qquad \beta_c(\varepsilon) = \inf\{\beta: \alpha^\varepsilon(\beta) > 0\},$$

but all we know how to show is that if $\alpha^\varepsilon(\beta) < 0$, then $\beta < \beta_c(\varepsilon)$. Assuming (1.12) and taking poetic license, we have the following picture of the evolution for β slightly larger than β_c. "Holes" form at a small rate and their size evolves like a random walk with a small negative drift. When we are close enough to β_c, the rate at which holes are created is larger than the drift in the random walk, so most of the line is covered by holes and the equilibrium density is low.

To explain the discrepancy between the simulation results and our heuristic argument, we turn to a seemingly unrelated question: "What meaning can be attached to $\rho_0(\beta)$ when $\beta \in (\beta_0, \beta_1)$ in Example 4?" We believe that this curve gives the densities of "metastable states."

CONJECTURE 2. *Suppose $\beta > \beta_0$. Let $N_t^\varepsilon = |\{x \in [-1/2, 1/2]^d: \bar{\xi}_t^\varepsilon(x) = 1\}|$ and let $\tau^\varepsilon = \inf\{t: N_t^\varepsilon \le (\rho_0 - \delta)\varepsilon^{-d}\}$. If $\delta > 0$, there is a constant $\gamma > 0$ so that*

$$P\big(\tau^\varepsilon \le \exp(\gamma\varepsilon^{-d})\big) \to 0 \quad \textit{as } \varepsilon \to 0.$$

To explain this we note that the ordinary differential equation $v_t' = f(v_t)$ converges to ρ_0 as long as $v_0 > \rho_1$. Based on this one can leap to the conclusion that if ε is small, then product measure with density ρ_0 is stable under small perturbations. That is, the density will stay close to ρ_0 until some "large deviation" near the origin drives the density below $\rho_0 - \delta$ or some hole created by a large deviation event in another part of the space grows and reaches the origin. Perhaps the most convincing argument in support of this picture is the fact that the second author has succeeded in making this argument rigorous for a long range version of sexual reproduction. Using ideas from the proof of Lemma 4.3 one can show that

$$P\big(\tau_\varepsilon \le \varepsilon^{-0.29/(d+1)}\big) \to 0,$$

but this falls far short of Conjecture 2.

Given Conjecture 2, it is easy to explain the discrepancy between the simulation results and Conjecture 1. Holes form at a very small rate [$\approx \exp(-\gamma\varepsilon^{-d})$] so the heuristic argument applies to a very narrow range of values. Dickman and Tomé were well aware of the metastability. Indeed they use the existence of "multiple locally stable phases" to argue for the existence of a discontinuous transition. They define two transition points by "for $\beta < \beta_+$ the vacuum (all 0's state) is locally stable while the active state remains stable for $\beta > \beta_-$." They report that when stirring accounts for 95% of the steps $\beta_- \approx 9.67$ and $\beta_+ \approx 17$. We believe that in the limit as $\varepsilon \to 0$, $\beta_- \to \beta_0 = 27/4$ and $\beta_+ \to \infty$.

In Section 2 we prove Theorem 1. We give the details for three reasons: (i) De Masi, Ferrari and Lebowitz (1986) only give the details in one special case,

(ii) we prove convergence directly rather than checking tightness and using uniqueness results for the limiting equation, and (iii) we need estimates on the rate of convergence for later proofs. In Section 3, we prove (I) and (II). The proof of (III) is carried out in Sections 4 and 5. In what follows, c and C are positive finite constants whose values are unimportant and will in general change from line to line, and we use the phrase "Markov's inequality" to refer to the following obvious estimate: if $X \geq 0$, then $P(X \geq t) \leq EX^r/t^r$.

REMARK. Since this paper was prepared, a very nice monograph by De Masi and Presutti (1992) has appeared. Chapters 6 to 10 treat the mean field limit theorem and various refinements concerning phase separation, interface dynamics and escape from an unstable equilibrium.

2. Mean field limit theorem. In this section we will prove Theorem 1. The ideas behind the proof are simple: We will give an explicit construction that allows us to define dual processes by asking the question "What is the state of x at time t?" and working backwards in time. The answer to this question can be determined by looking at the states of the sites in the "dual process" $I_\varepsilon^{x,t}(s)$ at time $t - s$. The particles in $I_\varepsilon^{x,t}(s)$ move according to stirring at a fast rate and give birth to new particles at rate

$$c^* = \sup_{\xi} \sum_i c_i(x, \xi).$$

So for small ε the dual process is almost a branching random walk and converges to a branching Brownian motion as $\varepsilon \to 0$. The last observation leads in a straightforward way to the convergence of the $u_i^\varepsilon(t, x)$ to limits $u_i(t, x)$ and to the asymptotic independence of adjacent sites, which implies that the $u_i(t, x)$ satisfy the limiting equations.

a. The dual process. The first step in the proof is to construct the process from a number of Poisson processes, all of which are assumed to be independent. For each $x \in \varepsilon\mathbb{Z}^d$, let $\{T_n^x, n \geq 1\}$ be a Poisson process with rate c^* and let $\{U_n^x, n \geq 1\}$ be a sequence of independent random variables that are uniform on $(0, 1)$. At time T_n^x we compute the flip rates $r_i = c_i(x, \xi(T_n^x))$ and use U_n^x to determine what (if any) flip should occur at x at time T_n^x. To be precise we let $p_i = \sum_{j \leq i} r_j / c^*$ for $i = 0, \dots, \kappa - 1$ with $p_{-1} = 0$ and flip to i if $U_n^x \in (p_{i-1}, p_i)$. To move the particles around, we let $\{S_n^{x,y}, n \geq 1\}$ be Poisson processes with rate $\varepsilon^{-2}/2$ when $x, y \in \varepsilon\mathbb{Z}^d$ with $\|x - y\|_1 = \varepsilon$, and we declare that at time $S_n^{x,y}$ the contents of x and y are exchanged. Even though there are infinitely many Poisson processes in the *graphical representation* $(\{S_n^{x,y}, n \geq 1\}, \{T_n^x, n \geq 1\}, \{U_n^x, n \geq 1\})$, and hence no first arrival, an idea of Harris (1972) allows us to construct the process starting from any initial configuration. We omit the proof of this assertion since we will now explain how to compute the state of x at time t by working backwards in time.

The dual process $I_\varepsilon^{x,t}(s)$ is naturally defined only for $0 \leq s \leq t$ but for a number of reasons it is convenient to assume that the Poisson processes and uniform random variables in the construction are defined for negative times and define $I_\varepsilon^{x,t}(s)$ for all $s \geq 0$. Let $\mathscr{N} = \{\varepsilon y_1, \dots, \varepsilon y_N\}$ be the set of neighbors of 0. The dual process makes transitions as follows:

> If $y \in I_\varepsilon^{x,t}(s)$ and $T_n^y = t - s$, then we add all the points of $y + \mathscr{N}$ to $I_\varepsilon^{x,t}(s)$.
>
> If $y \in I_\varepsilon^{x,t}(s)$ and $S_n^{y,z} = t - s$, then we move the particle at y to z.

It is easy to see that we can compute the state of x at time t by knowing the states of the y in $I_\varepsilon^{x,t}(s)$ at time $t - s$. We start with the values in $I_\varepsilon^{x,t}(s)$ at time $t - s$ and work up to time t. At S arrivals we perform the indicated stirrings. When an arrival T_n^y occurs at a point of the dual, we look at the value of the process on $y + \mathscr{N}$, compute the flip rates r_i and use U_n^x to determine what (if any) flip should occur.

To prepare for the proof of the convergence of $u_i^\varepsilon(t, x)$, we will now give a more detailed description of $I_\varepsilon^{x,t}(s)$. Let $X_\varepsilon^0(0) = x$, let R_ε^1 be the smallest value of s so that we have a T arrival at $X_\varepsilon^0(s)$ at time $t - s$, and set $X_\varepsilon^i(s) = \varepsilon y_i + X_\varepsilon^0(s)$. Finally, we set $\mu_\varepsilon^1 = 0$ to indicate that 0 is the mother of the N particles created at time R_ε^1. Passing now to the inductive step of the definition, suppose that we have defined the process up to time R_ε^m with $m \geq 1$. The $mN + 1$ existing particles move as dictated by stirring until R_ε^{m+1}, the first time $s > R_\varepsilon^m$ that a T arrival occurs at the location of one of our moving particles $X_\varepsilon^k(s)$, and then we set $X_\varepsilon^{mN+i}(s) = \varepsilon y_i + X_\varepsilon^k(s)$ for $1 \leq i \leq N$, and $\mu_\varepsilon^{m+1} = k$. The new particles may be created at the locations of existing particles. If so we say that a *collision* occurs and call the new particle *fictitious*. We will prove later that collisions can be ignored, but for proving the convergence of $u_i^\varepsilon(t, x)$ it is convenient to allow the fictitious particles to move and give birth like other particles, so for each $m \geq 1$ we define an independent copy of the graphical representation which we use for the births and movement of the mth particle if it is fictitious. By definition all the offspring of fictitious particles are also fictitious. One final bookkeeping detail: we set $X_\varepsilon^k(s) = \Upsilon$ before the particle is born.

b. The dual process is almost a branching random walk. The point of introducing fictitious particles is that $\mathscr{K}_t = mN + 1$ for $t \in [R_m^\varepsilon, R_{m+1}^\varepsilon)$ defines a branching process in which each particle gives birth to N additional particles at rate c^*. Our next goal is to show that if ε is small, then $I_\varepsilon^{x,t}(s)$ is almost a branching random walk in which particles jump to a randomly chosen neighbor at rate $d\varepsilon^{-2}$ and give birth as above. To do this we will couple X_ε^k to independent random walks Y_ε^k that start at the same location at time β_k = the birth time of X_ε^k, and jump to a randomly chosen neighbor at rate $d\varepsilon^{-2}$.

We say X_ε^k is *crowded* at time s if for some $j \neq k$, $\|X_\varepsilon^j(s) - X_\varepsilon^k(s)\|_1 \leq \varepsilon$. When X_ε^k is not crowded, we define the displacements of Y_ε^k to be equal to those of X_ε^k. When X_ε^k is crowded we use independent Poisson processes to determine the jumps of Y_ε^k. To estimate the difference between X_ε^k and Y_ε^k, we need to estimate the amount of time X_ε^k is crowded. Let $j \neq k$, $V_s^\varepsilon = X_\varepsilon^k(s) - X_\varepsilon^j(s)$ and W_s^ε be a random walk that jumps to a randomly chosen neighbor at rate $2d\varepsilon^{-2}$. The transition probabilities of V_s^ε differ slightly from those of W_s^ε when $\|x\|_1 = \varepsilon$ (here y denotes any point $\neq -x$ with $\|y\|_1 = \varepsilon$):

jumps from x to	rate in V	rate in W
$-x$	$\varepsilon^{-2}/2$	0
0	0	ε^{-2}
$x+y$	ε^{-2}	ε^{-2}

From this it should be clear that if we cut the visits to 0 out of the sample paths of W_s^ε and call the result $\hat{W}_s^\varepsilon$ then $\{\|\hat{W}_s^\varepsilon\|_1: s \geq 0\}$ and $\{\|V_s^\varepsilon\|_1: s \geq 0\}$ have the same distribution. It follows that for any integer $M \geq 1$, $v_t^{M\varepsilon} = |\{s \leq t: \|V_s^\varepsilon\|_1 \leq M\varepsilon\}|$ is stochastically smaller than $w_t^{M\varepsilon} = |\{s \leq t: \|W_s^\varepsilon\|_1 \leq M\varepsilon\}|$. That is, the two random variables can be constructed on the same space so that $v_t^{M\varepsilon} \leq w_t^{M\varepsilon}$. Well known asymptotic results for random walks imply that, when $t\varepsilon^{-2} \geq 2$,

$$Ew_t^{M\varepsilon} \leq \begin{cases} CM^d\varepsilon^2, & d \geq 3, \\ CM^2\varepsilon^2 \log(t\varepsilon^{-2}), & d = 2, \\ CM\varepsilon t^{1/2}, & d = 1. \end{cases} \tag{2.1}$$

Let $\chi_\varepsilon^k(t)$ be the amount of time X_ε^k is crowded in $[0, t]$. It is easy to see that

$$E\big(\chi_\varepsilon^k(t) | \mathscr{K}_t = K\big) \leq KEw_t^\varepsilon, \tag{2.2}$$

$$E\mathscr{K}_t = \exp(\nu t), \quad \text{where } \nu = c^*N, \tag{2.3}$$

$$E\big(\chi_\varepsilon^k(t)\big) \leq \exp(\nu t) Ew_t^\varepsilon. \tag{2.4}$$

To estimate the difference between $X_\varepsilon^k(s)$ and $Y_\varepsilon^k(s)$, we observe that if $\chi_\varepsilon^k(t) = \tau$, then the number of "independent jumps" in the ith coordinate of Y_ε^k that occur in $[0, t]$ has a Poisson distribution with mean $\varepsilon^{-2}\tau$. Let $\Delta_Y^i(s)$ be the net effect of the independent jumps on coordinate i up to time s. Recalling that changes in the ith coordinate of Y_ε^k have mean 0 and variance ε^2, it follows that $E\Delta_Y^i(s) = 0$ and

$$E\big(\Delta_Y^i(s)^2\big) = E\chi_\varepsilon^k(s). \tag{2.5}$$

Since $\Delta_Y^i(s)$ is a martingale, Kolmogorov's inequality implies

$$E\Big(\max_{0 \leq s \leq t} \Delta_Y^i(s)^2\Big) \leq 4E\big(\Delta_Y^i(t)^2\big). \tag{2.6}$$

Using Markov's inequality followed by (2.6), (2.5), (2.4) and (2.1) gives

$$\begin{aligned} &P\Big(\max_{0 \leq s \leq t} |\Delta_Y^i(s)| \geq \varepsilon^{0.3}\Big) \\ &\quad \leq \varepsilon^{-0.6} E\Big(\max_{0 \leq s \leq t} \Delta_Y^i(s)^2\Big) \leq C\varepsilon^{0.4} t^{1/2} \exp(\nu t). \end{aligned} \tag{2.7}$$

The arguments leading to the last inequality also apply to $\Delta_X^i(t)$, the net effect of jumps in $[0, t]$ while X_ε^k is crowded, so

$$P\Big(\max_{0 \le s \le t} \|X_\varepsilon^k(s) - Y_\varepsilon^k(s)\|_\infty \ge 2\varepsilon^{0.3}\Big) \le C\varepsilon^{0.4} t^{1/2} \exp(\nu t) \tag{2.8}$$

The last estimate shows that the X_ε^k are close to independent random walks. To see that with high probability no collisions occur, we note that by repeating the derivation of (2.4) it follows that the expected number of births from X_ε^k while there is some other X_ε^j in $X_\varepsilon^k + \mathscr{N}$ is smaller than

$$C\varepsilon t^{1/2} \exp(\nu t). \tag{2.9}$$

Equation (2.3) and Markov's inequality imply that

$$P(\mathscr{K}_t > K) \le K^{-1} \exp(\nu t). \tag{2.10}$$

When $\mathscr{K}_t \le K$, (2.9) implies that the expected number of collisions is smaller than

$$KC\varepsilon t^{1/2} \exp(\nu t). \tag{2.11}$$

Combining the last two results and setting $K = \varepsilon^{-0.2}$ shows that the probability of a collision is smaller than

$$C\varepsilon^{0.2} t^{1/2} \exp(\nu t). \tag{2.12}$$

Having shown that collisions are unlikely, we no longer have to worry about the labels μ_m^ε that tell us the mother of the N particles created at time R_m^ε since this will be clear from the evolution of the dual. A more significant consequence of the results in this subsection is that dual processes for different sites are asymptotically independent. To argue this, we say the two duals *collide* if a particle in one dual gives birth when crowded by a particle in the other one. The arguments leading to (2.12) show that with high probability two duals do not collide, and (2.8) implies that the movements of all the particles can be coupled to independent random walks.

c. Weak convergence of branching random walk to branching Brownian motion. This is a well known fact. However, for the proof of (III) we will need estimates of the rate of convergence, so we will give a complete proof. As in the previous subsection we will bound the difference by coupling the ith components of the random walks $Y_\varepsilon^{k,i}$ to one dimensional Brownian motions $Z^{k,i}$ that start at the same locations at the birth time β_k. Since the increments $Y_\varepsilon^{k,i}(\beta_k + t) - Y_\varepsilon^{k,i}(\beta_k)$ are independent random walks that start at 0, it suffices to show how to couple one such random walk S_t to a one-dimensional Brownian motion B_t that starts at 0. To do this we use the Skorokhod embedding, which in our case is particularly simple. Let $\tau_0^\varepsilon = 0$ and for $n \ge 1$ let

$$\tau_{n+1}^\varepsilon = \inf\{t > \tau_n^\varepsilon : |B(t) - B(\tau_n^\varepsilon)| = \varepsilon\}.$$

It is easy to see that $W_n = B(\tau_n^\varepsilon)$ is a discrete time random walk on $\varepsilon\mathbb{Z}$ that jumps $\pm\varepsilon$ with probability $1/2$ each. To get a continuous time random walk we let $N(t)$ be a Poisson process with rate ε^{-2} and let $S_t = W_{N(t)}$.

Our next step is to estimate $|S_t - B_t|$ by first showing that $\tau^\varepsilon_{N(t)}$ is close to t and then estimating the oscillations of Brownian motion. Let $t^\varepsilon_n = \tau^\varepsilon_n - \tau^\varepsilon_{n-1}$ for $n \geq 1$. The strong Markov property of Brownian motion implies that the t^ε_n are i.i.d. Now, $B_t^2 - t$ and $B_t^4 - 6B_t^2 t + 3t^2$ are martingales, so using the optional stopping theorem at time τ^ε_1 gives

$$E(\tau^\varepsilon_1) = \varepsilon^2, \qquad E(\tau^\varepsilon_1)^2 = 5\varepsilon^4/3.$$

Since $N(t)$ is independent of τ^ε_n, well known formulas for random sums give

$$E\tau^\varepsilon_{N(t)} = EN(t)E\tau^\varepsilon_1 = t,$$

$$\operatorname{Var}(\tau^\varepsilon_{N(t)}) = EN(t)\operatorname{Var}(\tau^\varepsilon_1) + (E\tau^\varepsilon_1)^2 \operatorname{Var}(N(t)) = C\varepsilon^2 t.$$

Now

$$\tau^\varepsilon_{N(t)} - t = \{\tau^\varepsilon_{N(t)} - \varepsilon^2 N(t)\} + \{\varepsilon^2 N(t) - t\}$$

is a martingale, so Kolmogorov's inequality implies

$$E\Big(\max_{0 \leq s \leq t} |\tau^\varepsilon_{N(s)} - s|^2\Big) \leq 4E(\tau^\varepsilon_{N(t)} - t)^2 \leq C\varepsilon^2 t, \tag{2.13}$$

and Markov's inequality gives

$$P\Big(\max_{0 \leq s \leq t} |\tau^\varepsilon_{N(s)} - s| > \varepsilon^{0.8}\Big) \leq \varepsilon^{-1.6} E\Big(\max_{0 \leq s \leq t} |\tau^\varepsilon_{N(s)} - s|^2\Big) \leq C\varepsilon^{0.4} t. \tag{2.14}$$

To estimate the oscillations of Brownian motion we observe that using the reflection principle, Brownian scaling and then a standard estimate on the tail of the normal distribution [see, e.g., Durrett (1990), page 9] gives

$$\begin{aligned} &P\Big(\max_{s \leq r \leq s+h} |B_s - B_r| > h^{0.4}\Big) \\ &\quad \leq 2P(|B_h| > h^{0.4}) \\ &\quad = 2P(|B_1| > h^{-0.1}) \leq 2h^{0.1} \exp(-h^{-0.2}/2). \end{aligned} \tag{2.15}$$

Since $B_{s+h} - B_{s+h-r}$, $0 \leq r \leq h$, is also a Brownian motion, the last estimate implies

$$P\Big(\max_{s \leq r \leq s+h} |B_{s+h} - B_r| > h^{0.4}\Big) \leq 2h^{0.1} \exp(-h^{-0.2}/2). \tag{2.16}$$

Using the last two estimates for $h = \varepsilon^{0.8}$ and $s = k\varepsilon^{0.8}$, where k is an integer with $0 \leq k \leq t\varepsilon^{-0.8}$, shows

$$\begin{aligned} &P(|B_u - B_v| \geq 2\varepsilon^{0.32} \text{ for some } 0 \leq u < v \leq t \text{ with } |v - u| \leq \varepsilon^{0.8}) \\ &\quad \leq (t\varepsilon^{-0.8} + 1) 2\varepsilon^{0.08} \exp(-\varepsilon^{-0.16}/2) \end{aligned} \tag{2.17}$$

since $u, v \in [(j-1)\varepsilon^{0.8}, (j+1)\varepsilon^{0.8}]$ for some j with $1 \leq j \leq [t\varepsilon^{-0.8}]$ and

$$|B_u - B_v| \leq |B_u - B_{j\varepsilon^{0.8}}| + |B_{j\varepsilon^{0.8}} - B_v|.$$

Combining (2.14) and (2.17) shows that when ε is small we have

$$(2.18)\qquad P\big(|S_s - B_s| \geq 2\varepsilon^{0.32} \text{ for some } s \leq t\big) \leq C\varepsilon^{0.4}t \quad \text{for } t \geq \varepsilon^{0.8}.$$

d. Convergence of $u_i^\varepsilon(t,x)$. The last two subsections have shown that if ε is small, the dual process is close to a branching Brownian motion. Indeed, (2.10) gives a bound on $P(\mathscr{K}_t > K)$, and (2.8) and (2.18) imply that, when $t \geq \varepsilon^{0.8}$,

$$(2.19)\qquad \begin{aligned} &P\Big(\max_{0\leq s\leq t} \big\| X_\varepsilon^k(s) - Z^k(s)\big\|_\infty > 4\varepsilon^{0.3} \text{ for some } k \leq K\Big) \\ &\quad \leq KC\varepsilon^{0.4}t^{1/2}\exp(\nu t). \end{aligned}$$

To compute the state of x at time t, we need not only the dual process $I_\varepsilon^{x,t}(s)$, $s \leq t$, but also the labels μ_n^ε and the uniform random variables U_n^x. However, the uniform random variables are independent of the dual process and, as we pointed out in a remark after (2.12), the μ_n^ε are only needed when a collision occurs.

The results in the last paragraph make it easy to show that $u_a^\varepsilon(t,x) \to u_a(t,x)$ as $\varepsilon \to 0$. Here and in what follows we will use a and b to denote possible states of the sites to ease the burden on the middle of the alphabet. The first step is to describe $u_a(t,x)$. Let Z_s be a branching Brownian motion starting with a single particle at x and let $\mathscr{K}_t$ be the number of particles at time t. For $0 \leq k < \mathscr{K}_t$, we let $\zeta_0(k)$ be independent and equal a with probability $\phi_a(Z_t^k)$. Once the ζ_0 are defined, we work up the space time set $\{Z_{t-s}^k\} \times \{s\}$. The values of $\zeta_s(k)$ stay constant as long as only stirring occurs. When $N+1$ branches $Z_{t-s}^i, Z_{t-s}^{kN+1}, \dots Z_{t-s}^{(k+1)N}$ come together (corresponding to a birth in the dual), we compute the flip rate at Z_{t-s}^i assuming it is in state $\zeta_s(i)$ and its neighbors are in states $\zeta_s(kN+j)$, $1 \leq j \leq N$. We generate an independent random variable uniform on $(0,1)$ to determine what (if any) flip should occur at Z_{t-s}^i. After we decide if we should change $\zeta_s(i)$, we can ignore $\zeta_s(kN+j)$ for $1 \leq j \leq N$. When we reach time t we will only be looking at the value of $\zeta_t(0)$. We call this value the *result of the computation* and let $u_a(t,x) = P(\zeta_t(0) = a)$.

The description in the last paragraph is much like the one given earlier for the dual, with one exception: The uniform random variables come from an auxiliary i.i.d. sequence instead of being read off the graphical representation. When there are no collisions in the dual, then the family structure of the influence set and the branching Brownian motion are the same. In this case if the inputs $\zeta_0(k)$ and the uniform random variables used are the same, the two computations have the same result. We have supposed that the initial functions $\phi_b(x)$ are continuous so (2.19) implies that as $\varepsilon \to 0$,

$$\max_k \big|\phi_b\big(X_\varepsilon^k(t)\big) - \phi_b\big(Z^k(t)\big)\big| \to 0,$$

where the maximum is taken over particles alive at time t. The last observation implies that we can with high probability arrange for all the inputs to be the same and it follows that $u_a^\varepsilon(t,x) \to u_a(t,x)$. The last proof extends trivially to show that if $x_\varepsilon \to x$, then $u_a^\varepsilon(t, x_\varepsilon) \to u_a(t, x)$. At the end of subsection (b), we observed that the influence sets from different points are asymptotically independent. Combining that observation with the proofs in this subsection implies that if $x_\varepsilon \to x$, then

$$P\big(\xi_t^\varepsilon(x_\varepsilon + \varepsilon y_j) = c_j, 0 \le j \le N\big) \to \prod_{j=0}^{N} u_{c_j}(t, x). \tag{2.20}$$

We are interested in statements that allow $x_\varepsilon \to x$ since this form of the conclusion implies that convergence occurs uniformly on compact sets.

e. The limit satisfies the p.d.e. The first step is to write the limiting equation in integral form.

(2.21) LEMMA. *Suppose f_a, $0 \le a < \kappa$, are continuous and g_a, $0 \le a < \kappa$, are bounded and continuous. The following statements are equivalent*:

(i) *The functions $u_a(t, x)$ are a classical solution of*

$$\frac{\partial u_a}{\partial t} = \frac{1}{2}\Delta u_a - f_a(u), \qquad u_a(0, x) = g_a(x),$$

that is, the indicated derivatives exist and are continuous.

(ii) *The functions $u_a(t, x)$ are bounded and satisfy*

$$u_a(t, x) = \int p_t(x, y) g_a(y)\, dy + \int_0^t ds \int p_s(x, y)\, f_a(u(t - s, y))\, dy,$$

where $p_t(x, y)$ is the transition probability density for Brownian motion.

PROOF. Statement (i) implies that $Z_s^a \equiv u_a(t - s, B_s) - \int_0^s f_a(u(t - r, B_r))\, dr$ is a bounded martingale, so $Z_0^a = EZ_t^a$ and (ii) follows from Fubini's theorem. To prove the converse, we begin by observing that if (ii) holds, then $u_a(t, x)$ has the necessary derivatives and Z_s^a is a martingale, so (i) follows from Itô's formula.

To get (ii) we will use the integration by parts formula. Let S_t^ε be the semigroup for the stirred particle system and T_t^ε be the semigroup for pure stirring. The integration by parts formula implies that for nice functions ψ we have

$$S_t^\varepsilon \psi(\xi) = T_t^\varepsilon \psi(\xi) + \int_0^t ds\, S_{t-s}^\varepsilon L T_s^\varepsilon \psi(\xi), \tag{2.22}$$

where L is the generator for the particle system with no stirring. We use (2.22) with $\psi_{x,a}(\xi) = 1$ if $\xi(x) = a$ and 0 otherwise. Now, for this choice of ψ,

$$T_s^\varepsilon \psi_{x,a}(\xi) = \sum_y p_s^\varepsilon(x, y)\psi_{y,a}(\xi) \tag{2.23}$$

where $p_s^\varepsilon(x, y)$ is the transition probability of a random walk that jumps from

y to z at rate $\varepsilon^{-2}/2$ if $\|y - z\|_1 = \varepsilon$. Now if $c_b(y, \xi) = h_b(\xi(y + \varepsilon y_0), \dots, \xi(y + \varepsilon y_N))$, then

$$L\psi_{y,a} = -\sum_b h_{b_0}(a, b_1, \dots, b_N)\psi_{y,a} \prod_{j=1}^N \psi_{y+\varepsilon y_j, b_j} + \sum_b h_a(b_0, b_1, \dots, b_N)\psi_{y,b_0} \prod_{j=1}^N \psi_{y+\varepsilon y_j, b_j}, \tag{2.24}$$

where the sums are over $b_0, \dots, b_N \in \{0, 1, \dots, \kappa - 1\}$. Substituting (2.23) and (2.24) into (2.22) gives

$$\begin{aligned} P(\xi_t^\varepsilon(x) = a) = {} & \sum_y p_t^\varepsilon(x,y) g_a(y) + \int_0^t ds \sum_y p_s^\varepsilon(x,y) \\ & \times E\Bigg\{ -\sum_b h_{b_0}(a, b_1, \dots, b_N)\psi_{y,a}(\xi_{t-s}^\varepsilon) \prod_{j=1}^N \psi_{y+\varepsilon y_j, b_j}(\xi_{t-s}^\varepsilon) \\ & + \sum_b h_a(b_0, b_1, \dots, b_N)\psi_{y,b_0}(\xi_{t-s}^\varepsilon) \prod_{j=1}^N \psi_{y+\varepsilon y_j, b_j}(\xi_{t-s}^\varepsilon) \Bigg\}. \end{aligned} \tag{2.25}$$

The local central limit theorem implies

$$\sum_y \left| \varepsilon^d p_s(x,y) - p_s^\varepsilon(x,y) \right| \to 0 \tag{2.26}$$

as $\varepsilon \to 0$. As we observed at the end of subsection (d),

$$E\psi_{y,c_0}(\xi_{t-s}^\varepsilon) \prod_{j=1}^N \psi_{y+\varepsilon y_j, c_j}(\xi_{t-s}^\varepsilon) \to \prod_{j=0}^N u_{c_j}(t-s, y),$$

and this convergence occurs uniformly on compact sets. Using (2.25), (2.26) and the dominated convergence theorem gives

$$\begin{aligned} u_a(t,x) = {} & \int p_t(x,y) g_a(y)\, dy + \int_0^t ds \int dy\, p_s(x,y) \\ & \times \Bigg\{ -\sum_b h_{b_0}(a, b_1, \dots, b_N) u_a(t-s, y) \prod_{j=1}^N u_{b_j}(t-s, y) \\ & + \sum_b h_a(b_0, b_1, \dots, b_N) u_{b_0}(t-s, y) \prod_{j=1}^N u_{b_j}(t-s, y) \Bigg\}. \end{aligned} \tag{2.27}$$

The term in braces is

$$\begin{aligned} & -\sum_{b \neq a} \left\langle c_b(0, \xi) 1_{\{\xi(0)=a\}} \right\rangle_{u(t-s,y)} + \left\langle c_a(0, \xi) \right\rangle_{u(t-s,y)} \\ & = f_a(u(t-s, y)). \end{aligned} \tag{2.28}$$

Combining this with (2.21) gives the conclusion of Theorem 1. □

3. Existence of nontrivial stationary distributions. In this section we will prove (I) and (II) for our five examples. The proof of (I) is simple and very general.

LEMMA 3.1. $\limsup_{\varepsilon\to 0}\rho_\varepsilon(\beta)\le\rho_0(\beta)$.

PROOF. Since our systems are attractive, $t\to E\bar{\xi}_t^\varepsilon(x)$ is decreasing. If we fix t, Theorem 1 implies that as $\varepsilon\to 0$, $E\bar{\xi}_t^\varepsilon(x)\to\bar{u}(t,x)$, where $\bar{u}(t,x)$ solves (1.3) with $\bar{u}(0,x)\equiv 1$ $\bar{u}(t,x)=v(t)$ and $v(t)\downarrow\rho_0$, the largest root of $f(u)=0$ in $[0,1]$. If $\delta>0$ and we pick t large and then ε small,

$$\rho_0+2\delta\ge v(t)+\delta\ge E\bar{\xi}_t^\varepsilon(x)\ge E\bar{\xi}_\infty^\varepsilon(x)$$

and the proof is complete. □

The proof of (II) requires much more work. The first step is the following input from p.d.e.'s.

LEMMA 3.2. *There are constants $p_1<p_2$, L and T, so that if $u(0,x)\ge p_1$ on $[-L+1,L-1]^d$, then $u(T,x)\ge p_2$ on $[-3L,3L]^d$.*

For Examples 1 and 2 we use the following result of Aronson and Weinberger [(1978), page 41].

THEOREM 3.1. *Suppose $f(0)=0$, $f(u)>0$ for $u\in(0,\alpha)$ with $\alpha>0$, and*

$$\lim_{u\to 0}f(u)/u^{1+2/d}>0.$$

If $u\ge 0$ is a solution of (1.3) with $u\not\equiv 0$, then for any compact set K

$$\liminf_{t\to\infty}\inf_{x\in K}u(t,x)\ge\alpha.$$

For Examples 3 and 4, Lemma 3.2 follows from Proposition 1. Example 5 can be treated by using Theorem 1 in the region with one root and Proposition 1 in the region with two.

Once we have the conclusion of Lemma 3.2., the rest of the proof of (II) only relies on results in the proof of the mean field limit theorem and hence is independent of the process under consideration. The plan is simple: Let $\delta<(p_2-p_1)/10$. Lemma 3.2 shows that if we start the p.d.e. with density at least p_1 in $[-L+1,L-1]^d$, then at time T we will have density at least p_2 in $2Le_1+[-L,L]^d$ and in $-2Le_1+[-L,L]^d$. Theorem 1, when supplemented by a second moment computation, will show that if ε is small, then the particle system will do this with high probability. The existence of a nontrivial stationary distribution will then follow by comparison with a mildly dependent oriented percolation process.

To begin carrying out this outline, we need to say what it means for the particle system to have density at least σ. Let

$$D_t^\varepsilon(x) = (\varepsilon/a_\varepsilon)^d \sum_{y:\, 0 \le y_i - x_i < a_\varepsilon} \xi_t^\varepsilon(y)$$

be the empirical density of particles in the cube $x + [0, a_\varepsilon)^d$. We say that ξ_t^ε has density at least σ in A if $D_t^\varepsilon(x) \ge \sigma$ for all points $x \in a_\varepsilon \mathbb{Z}^d \cap A$. We will pick $a_\varepsilon \to 0$ in the proof of Lemma 3.3. [See (3.5).] To make our little cubes fit neatly into the larger cubes I_k that we will define in a moment, we will suppose that $1/a_\varepsilon$ and L (to be defined later) are integers.

To achieve a finite range of dependence in the construction that we will use to prove (II), we have to restrict our attention to sites whose dual processes do not move too far. We say that y is *reasonable* at time t if $I_\varepsilon^{y,t}(s) \subset y + (-(M-3)L, (M-3)L)^d$ for all $s \le T + 1$. Since T is fixed it is clear that if we pick M large, then the probability that a site is reasonable is at least $1 - \delta$. Let $\tilde{\xi}_t^\varepsilon(y) = 1$ if $\xi_t^\varepsilon(y) = 1$ and y is reasonable at time t, and let

$$\tilde{D}_t^\varepsilon(x) = (\varepsilon/a_\varepsilon)^d \sum_{y:\, 0 \le y_i - x_i < a_\varepsilon} \tilde{\xi}_t^\varepsilon(y).$$

We say that ξ_t^ε has a nice density of at least σ in A if $\tilde{D}_t^\varepsilon(x) \ge \sigma$ for all points $x \in a_\varepsilon \mathbb{Z}^d \cap A$. Let $I_k = 2Lke_1 + [-L, L)^d$.

LEMMA 3.3. *Let $b_\varepsilon = a_\varepsilon^{1.9}$. If ξ_0^ε has density at least $p_1 + 2\delta$ in I_0 and ε is small, then with high probability $\xi_{T+b_\varepsilon}^\varepsilon$ will have a nice density of at least $p_2 - 4\delta$ in I_1 and in I_{-1}.*

To prove Lemma 3.3, we will first show the following.

LEMMA 3.4. *Suppose $\xi_0^\varepsilon(x)$ are independent with $P(\xi_0^\varepsilon(x) = 1) = \phi(x)$ and $\phi(x) \ge p_1 + \delta$ for $x \in [-L+1, L-1]^d$. If ε is small, then with high probability ξ_T^ε will have a nice density of at least $p_2 - 3\delta$ in I_1 and in I_{-1}.*

PROOF. We will prove this result by computing the mean and variance and then using Chebyshev's inequality. Lemma 3.2 and the definition of reasonable imply that

$$E\tilde{D}_T^\varepsilon(x) \ge p_2 - 2\delta \quad \text{for } x \in \mathscr{X}_\varepsilon \equiv a_\varepsilon \mathbb{Z}^d \cap I_1. \tag{3.1}$$

It follows from the remark after (2.12) that, as $\varepsilon \to 0$,

$$\mathrm{cov}_\varepsilon \equiv \sup_{x \ne y} \mathrm{Cov}\big(\tilde{\xi}_T^\varepsilon(x), \tilde{\xi}_T^\varepsilon(y)\big) \to 0. \tag{3.2}$$

Now $\tilde{D}_T^\varepsilon(x)$ is the average of $(a_\varepsilon/\varepsilon)^d$ random variables that take values in $\{0, 1\}$ and hence have variance $\le 1/4$, so

$$\mathrm{Var}\big(\tilde{D}_T^\varepsilon(x)\big) \le \frac{1}{4(a_\varepsilon/\varepsilon)^d} + \mathrm{cov}_\varepsilon. \tag{3.3}$$

Using (3.1) and Chebyshev's inequality now gives

$$P\big(\tilde{D}_T^\varepsilon(x) \le p_2 - 3\delta \text{ for some } x \in \mathscr{X}_\varepsilon\big) \le (2L/a_\varepsilon)^d \{\varepsilon^d/4a_\varepsilon^d + \mathrm{cov}_\varepsilon\}/\delta^2. \tag{3.4}$$

If we let $a_\varepsilon \to 0$ so that

$$\varepsilon/a_\varepsilon^2 \to 0 \quad \text{and} \quad \mathrm{cov}_\varepsilon/a_\varepsilon^d \to 0, \tag{3.5}$$

then the right-hand side of (3.4) goes to 0 and the desired conclusion follows. □

PROOF OF LEMMA 3.3. There are three steps in the proof. The first step is to show the following:

(a) Starting from a fixed configuration at time 0 is almost the same as starting from a product measure at time b_ε.

We begin by describing the product measure. For each $x \in \varepsilon\mathbb{Z}^d$, let $S_\varepsilon^x(t)$ be an independent continuous time random walk that starts at $S_\varepsilon^x(0) = x$ and jumps from z to each point $z + y$ with $\|y\|_1 = \varepsilon$ at rate $\varepsilon^{-2}/2$. Let $\zeta_\varepsilon(x) = \xi_0^\varepsilon(S_\varepsilon^x(b_\varepsilon))$. It is easy to see that the random variables $\zeta_\varepsilon(x)$ are independent. (Note: For this it is important that the initial configuration is nonrandom.) Consider a modified dual process that evolves according to the usual rules up to time T and then uses S_ε^y to move any particle at y at time T. The answer to the question "Is x occupied at time $T + b_\varepsilon$?" obtained from the modified dual process is obviously the same as the answer when we start with the configuration $\zeta_\varepsilon(y)$ at time b_ε.

To prove (a), we will now show that the modified dual process is almost the same as the dual process. To do this we need to show that during $[T, T + b_\varepsilon]$ the particles in the dual behave like independent random walks. This happens when five good events called $G_1, \dots, G_5$ occur. We will use $F_1, \dots, F_5$ to denote their complements. Let $\beta_\varepsilon = a_\varepsilon^{0.6}$ and

$$G_1 = \{\mathscr{K}_{T-\beta_\varepsilon} \le K\},$$

where $\mathscr{K}_t$ is the number of particles in the dual at time t. Inequality (2.10) implies that

$$P(F_1) \le K^{-1}\exp(\nu T). \tag{3.6}$$

Let

$$G_2 = \{\text{no birth occurs in } [T - \beta_\varepsilon, T + b_\varepsilon]\}.$$

It is easy to see that

$$P(G_1 \cap F_2) \le Kc^*(\beta_\varepsilon + b_\varepsilon) \le KCa_\varepsilon^{0.6}. \tag{3.7}$$

Now $\beta_\varepsilon = a_\varepsilon^{0.6} \ge \varepsilon^{0.3}$ for small ε, by (3.5), so well known results for random walk imply that for $t \ge \beta_\varepsilon$,

$$\sup_y P(S_\varepsilon^x(t) = y) \le C\big(\beta_\varepsilon \varepsilon^{-2}\big)^{-d/2}. \tag{3.8}$$

(The local central limit theorem gives the result for $t = \beta_\varepsilon$ and the Markov property extends it to $t > \beta_\varepsilon$.) Let

$$G_3 = \left\{ \| Y_\varepsilon^j(T) - Y_\varepsilon^k(T) \|_\infty \ge a_\varepsilon^{0.9} \text{ for all } j < k \le K \right\},$$

where the Y_ε^k are the independent random walks defined in Section 2. Inequality (3.8) implies

$$P(F_3) \le K^2 (2 a_\varepsilon^{0.9} \varepsilon^{-1})^d C \beta_\varepsilon^{-d/2} \varepsilon^d = K^2 C a_\varepsilon^{0.6d} \tag{3.9}$$

since $\beta_\varepsilon = a_\varepsilon^{0.6}$. Let

$$G_4 = \left\{ \| X_\varepsilon^k(T) - Y_\varepsilon^k(T) \|_\infty \le 2\varepsilon^{0.3} \text{ for all } k \le K \right\}.$$

Inequality (2.8) implies that

$$P(F_4) \le K C \varepsilon^{0.4} T^{1/2} \exp(\nu T). \tag{3.10}$$

After time T, the particles in the dual will behave like independent random walks until the first time two particles collide (i.e., occupy adjacent sites) but this is unlikely to occur by time $T + b_\varepsilon$. Let $S_\varepsilon^{0,i}(t)$ denote the ith component of $S_\varepsilon^0(t)$. Kolmogorov's inequality implies

$$P\left(\max_{0 \le t \le b_\varepsilon} |S_\varepsilon^{0,i}(t)| > a_\varepsilon^{0.9}/2 \right) \le C E |S_\varepsilon^{0,i}(b_\varepsilon)|^{2p} \Big/ (a_\varepsilon^{0.9}/2)^{2p}.$$

It is well known that in the case of simple random walk, the central limit theorem can be strengthened to conclude that if $t\varepsilon^{-2} \to \infty$, then

$$E |S_\varepsilon^{0,i}(t)/t^{1/2}|^p \to E\chi^p, \tag{3.11}$$

where χ has the standard normal distribution, so recalling $b_\varepsilon = a_\varepsilon^{1.9}$ we have

$$E |S_\varepsilon^{0,i}(b_\varepsilon)|^{2p} \Big/ (a_\varepsilon^{0.9}/2)^{2p} \le C a_\varepsilon^{0.1p}.$$

Now if we let

$$\begin{aligned} G_5 = \{&\text{all of the coordinates of the first } K \text{ particles change} \\ &\text{by less than } a_\varepsilon^{0.9}/3 \text{ in } [T, T + b_\varepsilon]\} \end{aligned}$$

and take $p = 6$, it follows that

$$P(F_5) \le C K a_\varepsilon^{0.1p} \le C K a_\varepsilon^{0.6}. \tag{3.12}$$

On $\bigcap_{i=1}^5 G_i$ the dual particles behave like independent random walks during $[T, T + b_\varepsilon]$, so if $\zeta_{T+b_\varepsilon}^\varepsilon$ gives the state of the process at time $T + b_\varepsilon$ starting from $\zeta_\varepsilon(x)$ at time b_ε, then adding the bounds on the $P(F_i)$ and noticing that the sum of the bounds in (3.12), (3.9) and (3.7) is smaller than $K^2 C a_\varepsilon^{0.6}$ gives

$$\begin{aligned} &\left| P(\xi_{T+b_\varepsilon}^\varepsilon(x) = 1) - P(\zeta_{T+b_\varepsilon}^\varepsilon(x) = 1) \right| \\ &\qquad \le K^{-1} \exp(\nu T) + K^2 C a_\varepsilon^{0.6} + K C \varepsilon^{0.4} T^{1/2} \exp(\nu T). \end{aligned} \tag{3.13}$$

The last inequality generalizes easily to show that if we are interested in the joint distribution of the values at x and y, then the difference of the probabilities in ξ and ζ is bounded by 2 times the right-hand side of (3.13).

Let $\phi(x) = P(\xi_0^\varepsilon(S_\varepsilon^x(b_\varepsilon)) = 1)$. The second step is to show the following:

(b) If ξ_0^ε has density $p_1 + 2\delta$ in $[-L, L)^d$ and ε is small, then $\phi(x) \geq p_1 + \delta$ when $x \in [-L+1, L-1]^d$.

To simplify computations we will remove the extra particles from the initial configuration: Let $\bar{\xi}_0^\varepsilon \leq \xi_0^\varepsilon$ have no particles outside $[-L, L)^d$ and have exactly $[(p_1 + 2\delta)(a_\varepsilon/\varepsilon)^d] + 1$ particles in $y + [0, a_\varepsilon)^d$ for all $y \in a_\varepsilon \mathbb{Z}^d \cap [-L, L)^d$. To bound $\phi(x)$, we write

$$P(S_\varepsilon^x(b_\varepsilon) = y) = \pi_1(x, y) + \pi_2(x, y),$$

where $\pi_1(x, y)$ is constant for $y \in Q(z) \equiv z + [0, a_\varepsilon)^d$ for each $z \in a_\varepsilon \mathbb{Z}^d$, $\pi_2(x, y) \geq 0$ and π_1 is chosen as large as possible. Let

$$q_2(x) = \sum_y \pi_2(x, y).$$

We have chosen $b_\varepsilon \geq a_\varepsilon^{1.9} \geq \varepsilon^{0.95}$, so $\varepsilon^{-2} b_\varepsilon \to \infty$. The local central limit theorem tells us that $P(S_\varepsilon^x(b_\varepsilon) = y)$ will be almost constant on cubes of side $o(b_\varepsilon^{1/2}) = o(a_\varepsilon^{0.95})$ so

$$\sup_x q_2(x) \to 0 \quad \text{as } \varepsilon \to 0. \tag{3.14}$$

Let

$$q_1(x, z) = \sum_{y \in Q(z)} \pi_1(x, y).$$

To calculate $P(\zeta_\varepsilon(x) = 1)$, we will generate random variables $\zeta_\varepsilon'(x) =_d \zeta_\varepsilon(x)$ in a special way. We first flip a coin with probability $q_2(x)$ of tails and $1 - q_2(x)$ of heads. When tails comes up we pick a site y_x according to the distribution $\pi_2(x, y)/q_2(x)$ and set $\zeta_\varepsilon'(x) = 1$ if that site is occupied, $= 0$ otherwise. When heads comes up, we pick z_x according to the distribution $q_1(x, z)/(1 - q_2(x))$, and let $w_x = 1$ if $z_x \in [-L, L)^d$ and $= 0$ otherwise. The random variable z_x tells us the cube $Q(z)$ that S_ε^x lands in. Now the definition of $\pi_1(x, y)$ implies that the conditional distribution given $y \in Q(z)$ is uniform, and we have supposed that there are exactly $[(p_1 + 2\delta)(a_\varepsilon/\varepsilon)^d] + 1$ particles in $Q(z)$ when $z \in [-L, L)^d$, so we let v_x be an independent random variable that is 1 with probability $p_1 + 2\delta$, 0 otherwise, and let $\zeta_\varepsilon'(x) = w_x v_x$. From the definitions above it should be clear that

$$P(\zeta_\varepsilon'(x) = 1) \geq \left\{P\left(S_\varepsilon^x(b_\varepsilon) \in [-L, L)^d\right) - q_2(x)\right\}(p_1 + 2\delta). \tag{3.15}$$

Relation (3.12) implies that, as $\varepsilon \to 0$,

$$\inf_{x \in [-L+1, L-1]^d} P\left(S_\varepsilon^x(b_\varepsilon) \in [-L, L)^d\right) \to 1, \tag{3.16}$$

so using (3.14) we see that, for small ε,

$$\phi(x) = P(\zeta_\varepsilon'(x) = 1) \geq p_1 + \delta \quad \text{for all } x \in [-L+1, L-1]^d, \tag{3.17}$$

completing the proof of (b).

To complete the proof of Lemma 3.3, we will now compute the mean and variance and use Chebyshev's inequality. Let $S = T + b_\varepsilon$. Statements (3.13), (b) and (3.1) imply that if ε is small, then

$$(3.18) \qquad E\tilde{D}_S^\varepsilon(x) \ge p_2 - 3\delta \quad \text{for } x \in \mathscr{X}_\varepsilon \equiv a_\varepsilon \mathbb{Z}^d \cap I_1.$$

To bound the covariance we observe that the remark after (3.13) and (3.2) imply that

$$(3.19) \qquad \text{cov}_\varepsilon \equiv \sup_{x \ne y} \text{Cov}\big(\tilde{\xi}_S^\varepsilon(x), \tilde{\xi}_S^\varepsilon(y)\big) \to 0.$$

With (3.18) and (3.19) established, the rest of the proof is almost identical to the proof of Lemma 3.4, so the details are omitted. □

PROOF OF (II). With Lemma 3.3 established, (II) will follow from a comparison with a mildly dependent oriented percolation. To prepare for the proof of (III) and future applications, we will state a general result.

COMPARISON ASSUMPTIONS. We suppose the following ingredients to be given: a process ξ_t^ε: $\varepsilon\mathbb{Z}^d \to \{0, 1, \dots \kappa - 1\}$ that is constructed from a graphical representation (i.e., a family of independent Poisson processes), a real number L that is an integer multiple of ε, and a collection H of good configurations ξ: $\varepsilon\mathbb{Z}^d \cap [-L, L]^d \to \{0, 1, \dots \kappa - 1\}$ with the following property: If $\xi \in H$, then there is an event G_ξ measurable with respect to the graphical representation in $[-ML, ML]^d \times [0, MT]$ and with $P(G_\xi) \ge (1 - \theta)$ so that if the restriction of ξ_0^ε to $[-L, L]^d$ is ξ, then on G_ξ the restrictions of ξ_T^ε to $2Le_1 + [-L, L]^d$ and to $-2Le_1 + [-L, L]^d$ lie in (the appropriate translates of) H.

The parenthetical phrase refers to the fact that the restriction of ξ_T^ε to $2Le_1 + [-L, L]^d$ is a function from $\varepsilon\mathbb{Z} \cap 2Le_1 + [-L, L]^d$ to $\{0, 1, \dots, \kappa - 1\}$ but H is by definition a collection of functions from $\varepsilon\mathbb{Z} \cap [-L, L]^d$ to $\{0, 1, \dots, \kappa - 1\}$. We will ignore this small point in other definitions below.

Our goal is to show that when the comparison assumptions hold our process dominates an M dependent oriented percolation process on $\mathscr{L} = \{(x, m) \in \mathbb{Z}^2$: $x + m$ is even$\}$ with density at least $1 - \theta$. Our first step is to describe the percolation process. Given random variables $\omega(x, n)$, $(x, n) \in \mathscr{L}$, that indicate whether the sites are open (1) or closed (0), we say that (y, n) can be reached from (x, m) and write $(x, m) \to (y, n)$ if there is a sequence of points $x = x_m, \dots, x_n = y$ so that $|x_k - x_{k-1}| = 1$ for $m < k \le n$ and $\omega(x_k, k) = 1$ for $m \le k \le n$. Up to this point the $\omega(x, n)$ could be arbitrary random variables. The phrase "M dependent with density at least $1 - \theta$" means that if (x_i, n_i), $1 \le i \le I$, is a sequence with $|x_i - x_j| \ge M$ or $|n_i - n_j| \ge M$ whenever $i \ne j$, then

$$(3.20) \qquad P\big(\omega(x_i, n_i) = 0 \text{ for } 1 \le i \le I\big) \le \theta^I.$$

To compare with oriented percolation, we will say that $(x, n) \in \mathscr{L}$ is *occupied* if ζ = the restriction of ξ_{nT}^ε to $2xLe_1 + [-L, L]^d$ is in H and let

$V_n = \{x\colon (x, n) \text{ is occupied}\}$. To define the random variables $\omega(x, n)$ for $(x, n) \in \mathscr{L}$ with $n \geq 0$ we consider two cases. If (x, n) is occupied, we let $\omega(x, n) = 1$ if the graphical representation in $(2xLe_1 + [-ML, ML]^d) \times [nT, (n + M)T]$ lies in the corresponding G_ζ, and 0 otherwise. If (x, n) is vacant, we define $\omega(x, n)$ by flipping an independent coin with probability $1 - \theta$ of heads and probability θ of tails. Let $W_0 = V_0$ and $W_n = \{y\colon (x, 0) \to (y, n) \text{ for some } x \in W_0\}$. From the comparison assumptions and induction it follows easily that $V_n \supset W_n$. We prove (3.20) by induction on I. The conclusion for $I = 1$ is an immediate consequence of the definition. To do the induction step, suppose without loss of generality that n_I is the largest value of n_i and let $\mathscr{F}$ be the information contained in the graphical representation up to time $n_I T$ or in one of the boxes $(2x_i Le_1 + [-ML, ML]^d) \times [n_i T, (n_i + M)T]$ with $1 \leq i < I$. The comparison assumptions imply that

$$P(\omega(x_I, n_I) = 0 | \mathscr{F}) \leq \theta.$$

Integrating the last inequality over $\{\omega(x_i, n_i) = 0,\ 1 \leq i < I\}$ and using the induction hypothesis gives (3.20).

The next result is a straightforward extension of the arguments in Section 10 of Durrett (1984).

LEMMA 3.5. *Suppose $W_0 = 2\mathbb{Z}$. If $\theta \leq 6^{-4(2M-1)^2}$, then*

$$P(0 \notin W_{2n}) \leq \theta + 162\theta^{1/(4(2M-1)^2)}.$$

PROOF. We say that (x, m) can be reached from (y, n) by a dual path [and write $(y, n) \to_* (x, m)$ if there is a sequence of points $x = x_m, \dots, x_n = y$ so that $|x_k - x_{k-1}| = 1$ for $m < k \leq n$ and $\omega(x_k, k) = 1$ for $m \leq k \leq n$. It should be clear from the definition that $(x, m) \to (y, n)$ if and only if $(y, n) \to_* (x, m)$, so $\{0 \in W_{2n}\} = \{(0, 2n) \to_* (\cdot, 0)\}$, where $(\cdot, 0)$ is short for $\{(x, 0)\colon x \in 2\mathbb{Z}\}$. To estimate the probability of $(0, 2n) \not\to_* (\cdot, 0)$ we observe that this happens if $(0, 2n)$ is closed, which has probability at most θ, and this gives the first term on the right-hand side. When $(0, 2n)$ is open and $(0, 2n) \not\to_* (\cdot, 0)$ occurs, we can let $C_n = \{z\colon (0, 2n) \to_* z\}$ and define a contour associated with C_n exactly as on page 1026 of Durrett (1984). As in the source cited, (i) for a contour of length m to exist there must be $m/4$ closed sites, (ii) we can find a subset of size at least $m/4(2M - 1)^2$ that satisfies the hypotheses of (3.20), (iii) there are at most 3^m contours of length m and (iv) the shortest contour has length 4, so

$$P((0, 2n) \text{ is open}, (0, 2n) \not\to_* (\cdot, 0)) \leq \sum_{m=4}^{\infty} 3^m \theta^{m/4(2M-1)^2} \leq 2 \cdot 3^4 \theta^{1/(2M-1)^2}$$

since $3\theta^{1/4(2M-1)^4} \leq 1/2$. □

With Lemma 3.5 in hand the rest is a standard argument. We start our system with all sites occupied so $\bar{\xi}_0^\varepsilon$ has density $> p_1 + \delta$ in all the cubes $2xLe_1 + [-L, L]^d$. As noted in the introduction, $P(\bar{\xi}_t^\varepsilon(x) = 1)$ is a decreasing function of t, and $\bar{\xi}_t^\varepsilon$ converges weakly to a limit that is a stationary distribution $\bar{\xi}_\infty^\varepsilon$. To check that the limit is nontrivial we note that when $0 \in V_{2n}$, $\bar{\xi}_{2nS}^\varepsilon$ has density at least $p_2 - 4\delta$ in $[-L, L]^d$, so

$$\frac{1}{(2L/\varepsilon)^d} E\Bigg(\sum_{x \in [-L, L)^d} \bar{\xi}_{2nS}^\varepsilon(x) \Bigg) \geq (p_2 - 4\delta) P(0 \in V_{2n}).$$

Using translation invariance and our comparison it follows that

$$P\big(\bar{\xi}_{2nS}^\varepsilon(x) = 1\big) \geq (p_2 - 4\delta) P(0 \in W_{2n}).$$

Using Lemma 3.5 now and recalling that δ and θ are arbitrary completes the proof of (II). □

4. Nonexistence of nontrivial stationary distributions. In this section we will do the first half of the proof of (III). Let $a_\varepsilon = \varepsilon^\alpha$ with $\alpha = 0.1/d$, and as in Section 3, let $b_\varepsilon = a_\varepsilon^{1.9}$ and $\beta_\varepsilon = a_\varepsilon^{0.6}$. Let $\gamma = \alpha/80 = 1/(800d)$, and let $T = (\gamma/\nu)\log(1/\varepsilon)$, where $\nu = c^* N$ is the constant defined in (2.3), so that $\exp(\nu T) = \varepsilon^{-\gamma}$. Let $\omega = -r/2 > 0$, where $r < 0$ is the wave speed described in the introduction. Let $u_L(t, x)$ be the solution of (1.3) with initial condition $u(0,x) = \phi_L(x)$, where $\phi_L(x) = \rho_1/2$ for $\|x\|_\infty \leq L - 1$ and 1 otherwise. Let $S = T + b_\varepsilon$. Taking $\eta = -r/4$ in Proposition 2 it follows that if $L \geq L_0$ and $T \geq T_0$, then there are constants $0 < \lambda, C < \infty$ so that, for small ε,

$$u_L(T, x) \leq C\varepsilon^\lambda \quad \text{for } x \in [-L - wS, L + wS]^d. \tag{4.1}$$

Note that we have used L instead of $L - 1$ and S instead of T but $b_\varepsilon \leq 1$ so this can be done for $T \geq T_0$. To simplify computations below we can and will suppose that $\lambda < \gamma$. As in Section 3, we let

$$D_t^\varepsilon(x) = (\varepsilon/a_\varepsilon)^d \sum_{y:\, 0 \leq y_i - x_i < a_\varepsilon} \xi_t^\varepsilon(y)$$

and say that ξ_t^ε has density at most σ in A if $D_t^\varepsilon(x) \leq \sigma$ for all $x \in a_\varepsilon \mathbb{Z}^d \cap A$. We will eventually set $L = C\log(1/\varepsilon)$. Our first goal is to establish the following.

LEMMA 4.1. *If ξ_0^ε has density at most $\rho_1/3$ in $[-L, L)^d$, where $L_0 \leq L \leq \varepsilon^{-0.005/d}$, then*

$$P\big(\xi_S^\varepsilon(x) = 1\big) \leq C\varepsilon^\lambda \quad \textit{for } x \in [-L - wS, L + wS]^d.$$

PROOF. Let $S_\varepsilon^x(t)$ be the independent random walks defined in the proof of (a) in Section 3, let $\zeta_\varepsilon(x) = \xi_0^\varepsilon(S_\varepsilon^x(b_\varepsilon))$ and let $\zeta_{T+b_\varepsilon}^\varepsilon$ be the state at time $T + b_\varepsilon$ starting from the product measure $\zeta_\varepsilon(y)$ at time b_ε. Taking $K = \varepsilon^{-2\gamma}$ in (3.13) and recalling $\exp(\nu T) = \varepsilon^{-\gamma}$ gives

$$\begin{aligned} &\big|P\big(\xi_{T+b_\varepsilon}^\varepsilon(x) = 1\big) - P\big(\zeta_{T+b_\varepsilon}^\varepsilon(x) = 1\big)\big| \\ &\qquad \leq \varepsilon^\gamma + C\varepsilon^{-4\gamma + 0.6\alpha} + C\varepsilon^{0.4 - 3\gamma}\log^{1/2}(1/\varepsilon) \leq C\varepsilon^\gamma \end{aligned} \tag{4.2}$$

since $\alpha = 0.1/d$ and $\gamma = 1/(800d)$. From the proof of (3.17) one concludes easily that

$$(4.3)\quad P(\zeta_\varepsilon(y) = 1) \le \rho_1/3 + q_2(y) + P\big(S_\varepsilon^y(b_\varepsilon) \notin [-L, L]^d\big) \le \rho_1/2$$

for all $y \in [-L+1, L-1]^d$ when ε is small.

Let $\hat{\zeta}^\varepsilon_{T+b_\varepsilon}$ be the state at time $T + b_\varepsilon$ when we start at time b_ε from a product measure with $P(\hat{\zeta}^\varepsilon_{b_\varepsilon}(x) = 1) = \phi_L(x)$. Inequality (4.3) and attractiveness imply

$$(4.4)\qquad P\big(\zeta^\varepsilon_{T+b_\varepsilon}(x) = 1\big) \le P\big(\hat{\zeta}^\varepsilon_{T+b_\varepsilon}(x) = 1\big).$$

To bound $P(\hat{\zeta}^\varepsilon_{T+b_\varepsilon}(x) = 1)$ we will show that the dual process $I_\varepsilon^{x, T+b_\varepsilon}(t)$, $t \le T$, is close to a branching Brownian motion. Taking $t = T$ and $K = \varepsilon^{-2\gamma}$ in (2.10) gives

$$(4.5)\qquad P\big(\mathscr{K}_T > \varepsilon^{-2\gamma}\big) \le \varepsilon^\gamma.$$

When $\mathscr{K}_T \le \varepsilon^{-2\gamma}$ it follows from (2.11) that the probability of a collision is smaller than [recall $\exp(\nu T) = \varepsilon^{-\gamma}$]

$$(4.6)\qquad C\varepsilon^{1-3\gamma}\log^{1/2}(1/\varepsilon),$$

and (2.19) implies that, for small ε,

$$(4.7)\qquad \begin{aligned} &P\Big(\max_{0 \le s \le T+b_\varepsilon} \|X_\varepsilon^k(s) - Z^k(s)\|_\infty > 4\varepsilon^{0.3} \text{ for some } k \le K\Big) \\ &\quad \le C\varepsilon^{0.4-3\gamma}\log^{1/2}(1/\varepsilon). \end{aligned}$$

When $\|X_\varepsilon^k(s) - Z^k(s)\|_\infty \le 4\varepsilon^{0.3}$ for all k, we can have a "ϕ-error," that is, $\phi_L(X_\varepsilon^k(T)) \ne \phi_L(Z^k(T))$, only if $Z^k(T)$ lands in J, the set of points within a distance $4\varepsilon^{0.3}$ of $\partial[-L+1, L-1]^d$. To estimate the probability that $Z^k(T) \in J$, we note that if our branching random walk starts at x, then the mean number of particles in J at time T

$$(4.8)\qquad \begin{aligned} m_T(x, J) &= e^{\nu T}P_x(B_T \in J) \\ &\le e^{\nu T}(2\pi T)^{-d/2}2d(8\varepsilon^{0.3})(2L)^{d-1} \le C\varepsilon^{0.295-\gamma} \end{aligned}$$

since $T \ge 1$ and we have supposed $L \le \varepsilon^{-0.005/d}$.

When no collisions occur, the family structure of the dual process is the same as that of the branching Brownian motion, and when no ϕ-error occurs, we can suppose that the coin flips that determine the occupancy or vacancy of the site where the kth particle lands are the same as in the two processes. Now if we run a branching Brownian motion starting with one particle at x until time T, attach independent uniforms to the birth events as in subsection (d) of Section 2, flip a coin with probability $\phi_L(Z^k(T))$ of heads to determine if the kth site is occupied and compute as we do for the dual process, then the probability that x is occupied is $u_L(T, x)$. Combining the last observation with (4.5)–(4.8) and observing that (4.6) is smaller than the right-hand side of (4.7) shows

$$(4.9)\qquad \begin{aligned}&\left|P\left(\hat{\zeta}^{\varepsilon}_{T+b_\varepsilon}(x)=1\right)-u_L(T,x)\right|\\ &\qquad\le \varepsilon^{\gamma}+C\varepsilon^{0.4-3\gamma}\log^{1/2}(1/\varepsilon)+C\varepsilon^{0.295-\gamma}\end{aligned}$$

Combining (4.2), (4.4), (4.9) and (4.1) now and recalling $\lambda<\gamma=1/(800d)$ gives

$$(4.10)\qquad P\left(\xi^{\varepsilon}_{T+b_\varepsilon}(x)=1\right)\le C\varepsilon^{\lambda}\quad \text{for } x\in[-L-wS,L+wS]^d,$$

which is the desired conclusion. □

To prepare for the next step, we need a preliminary result about $\mathscr{K}_t$, the number of births by time t. The proof is a straightforward generalization of the well known special case $r=2$ and can be skipped without loss.

LEMMA 4.2. *If $r\ge 1$ is an integer, then*

$$E\mathscr{K}_t^r\le C_r(E\mathscr{K}_t)^r.$$

PROOF. The result is trivial for $r=1$. To prove it for $r\ge 2$ we will use induction. We will first prove the result for integer t. To do this, we observe that $Z_n=\mathscr{K}_n$ is a discrete time branching process in which the offspring distribution has mean $\theta=\exp(\nu)>1$. Corollary 1 on page 111 of Athreya and Ney (1972) implies that $E\mathscr{K}_t^r<\infty$ for any t, so the rth moment of the offspring distribution is finite. For concreteness we suppose Z_n is constructed from i.i.d. ξ_i^n by

$$Z_{n+1}=\sum_{i=1}^{Z_n}\xi_i^{n+1},$$

with $Z_{n+1}=0$ if $Z_n=0$. It is well known that if we let $\mathscr{F}_n=\sigma(\xi_i^m\colon m\le n)$, then $X_n=Z_n/\theta^n$ is a martingale. Written in terms of X_n, what we want to show is $EX_n^r\le C_r$. To compute EX_n^r we will look at

$$E\left(X_{n+1}^r|\mathscr{F}_n\right)=\theta^{-(n+1)r}E\left(\left(\sum_{i=1}^{Z_n}\left\{\theta+\left(\xi_i^{n+1}-\theta\right)\right\}\right)^r\middle|\mathscr{F}_n\right)$$

If we let $S_k=\xi_1^{n+1}+\cdots+\xi_k^{n+1}-k\theta$, then on $\{Z_n=k\}$ the right-hand side is

$$\begin{aligned}&\theta^{-(n+1)r}E\left(\sum_{q=0}^{r}\binom{r}{q}(\theta k)^{r-q}S_k^q\middle|\mathscr{F}_n\right)\\ &\qquad=\theta^{-nr}E\sum_{q=0}^{r}\binom{r}{q}\theta^{-q}k^{r-q}ES_k^q.\end{aligned}$$

Now $ES_k=0$ and the reasoning that leads to (3.11) shows $|ES_k^q|\le B_qk^{q/2}$, so substituting Z_n for k we have

$$E(X_{n+1}^r|\mathscr{F}_n) = X_n^r + \sum_{q=2}^{r} \theta^{-q} B_q \binom{r}{q} Z_n^{r-q/2}/\theta^{nr}$$

$$\leq X_n^r + A_r X_n^{r-1} \theta^{-n},$$

where $A_r = \Sigma_{q=2}^r \theta^{-q} B_q \binom{r}{q}$. Taking expected values and using the result for $r-1$ gives

$$EX_{n+1}^r \leq EX_n^r + A_r C_{r-1} \theta^{-n}.$$

Since $EX_0^r = 1$ it follows that

$$EX_{n+1}^r \leq 1 + A_r C_{r-1}/(1 - 1/\theta).$$

This proves the result for integer t. To extend the result to general t we observe that since there are no deaths in $\mathscr{K}_t$,

$$E\mathscr{K}_t^r \leq E\mathscr{K}_{[t]+1}^r,$$

where $[t]$ is the largest integer less than or equal to t. □

Lemma 4.1 gives an estimate for the probability that a site is occupied. To get a large region in which the density is low we will apply the next result several times.

LEMMA 4.3. *Let $S = T + b_\varepsilon$. If ξ_0^ε has density at most $\rho_1/3$ in $[-L, L)^d$, where $L_0 \leq L \leq \varepsilon^{-0.005/d}$, then with probability at least $1 - \varepsilon^{0.005}$, ξ_S^ε has density at most $\rho_1/3$ in $[-L - wS, L + wS)^d$.*

PROOF. Without loss of generality we can suppose that there are exactly $[(\rho_1/3)(a_\varepsilon/\varepsilon)] + 1$ particles in each cube $y + [0, a_\varepsilon)^d$ with $y \in a_\varepsilon \mathbb{Z}^d$ that lies in $[-L, L)^d$ and that all sites outside $[-L, L)^d$ are occupied. Let $D_t^\varepsilon(x)$ be the density defined at the beginning of the section, let $n_\varepsilon = (a_\varepsilon/\varepsilon)^d$ and let q be an even integer.

$$E\left(D_S^\varepsilon(x)^q\right) \leq n_\varepsilon^{-q}\left(C_q n_\varepsilon^{q-1} + \sum E \prod_i \xi_S^\varepsilon(y_i)\right), \tag{4.11}$$

where the sum is over distinct $y_1, \dots, y_q \in x + [0, a_\varepsilon)^d$, since there are at most $C_q n_\varepsilon^{q-1}$ terms in which two y_i's are equal and $\xi_S^\varepsilon(y_i) \in \{0, 1\}$. To bound the terms in the second sum, we begin by observing that if we replaced the duals by independent branching Brownian motions and the initial configuration by coin flips with probability $\phi_L(y)$ of heads that occur at time b_ε, then we would get independent random variables ζ_y with means $u_L(T, y)$. Let $\xi_y = \xi_S^\varepsilon(y)$. To bound the expectation of the product we will construct the ξ_y and ζ_y on the same space so that

$$\xi_y = \zeta_y + \Delta_y^1 + \Delta_y^2,$$

where $\zeta_y + \Delta_y^1 \in \{0, 1\}$, and the (ζ_y, Δ_y^1) are independent. Hence

$$E\prod_i \xi_{y_i} \le E\Big\{\prod_i \left(\zeta_{y_i} + \Delta^1_{y_i}\right)\Big\} + (2^q - 1)\max_i P\left(\Delta^2_{y_i} \neq 0\right).$$

Now the (ζ_y, Δ_y^1) are independent and $E\zeta_y = u_L(S, y)$, so

$$E\prod_i \left(\zeta_{y_i} + \Delta^1_{y_i}\right) = \prod_i E\left(\zeta_{y_i} + \Delta^1_{y_i}\right) \le \left(\max_i\left\{u_L(T, y_i) + P\left(\Delta^1_{y_i} \neq 0\right)\right\}\right)^q$$

and it follows that

$$E\prod_i \left(\xi_{y_i} - u_{y_i}\right) \le \left(\max_y\left\{u_L(T, y) + P\left(\Delta^1_y \neq 0\right)\right\}\right)^q \tag{4.12}$$
$$+ (2^q - 1)\max_y P\left(\Delta^2_y \neq 0\right),$$

where the maxima are taken over $y \in x + [0, a_\varepsilon)^d$.

To estimate the difference between ξ_x and ζ_x, we observe that the dual process and the branching Brownian motion compute the same result when the following good events happen:

(i) $\mathscr{K}_T \le K$;
(ii) $\max_{0 \le s \le T}\|X_\varepsilon^k(s) - Z^k(s)\|_\infty \le 2\varepsilon^{0.3}$ for all $k \le K$;
(iii) no collision occurs during $[0, T]$;
(iv) no birth occurs in $[T - \beta_\varepsilon, T + b_\varepsilon]$;
(v) $\|Z_\varepsilon^j(T) - Z_\varepsilon^k(T)\|_\infty \ge a_\varepsilon^{0.9}$ for all $j < k \le K$;
(vi) none of the X_ε^k moves by more than $a_\varepsilon^{0.9}/3$ in $[T, T + b_\varepsilon]$;
(vii) $\phi_L(X_\varepsilon^k(T)) = \phi_L(Z^k(T))$ for $1 \le k \le K$.

We begin with the three that concern the branching Brownian motion Z_ε^k and hence are independent for different sites. Markov's inequality says that for any $X \ge 0$ and $r > 0$ one has

$$P(X \ge x) \le EX^r/x^r.$$

Using this estimate and Lemma 4.2 it follows that

$$P(\mathscr{K}_T > K) \le K^{-r}E(\mathscr{K}_T^r) \le K^{-r}C_r\exp(\nu Tr).$$

Setting $K = \exp(2\nu T) = \varepsilon^{-2\gamma}$gives

$$P(\mathscr{K}_T > K) \le CK^{-r/2} = C\varepsilon^{r\gamma}. \tag{4.13}$$

Turning to the two troublesome estimates we observe that (3.7) implies

$$P\left(\mathscr{K}_{T-\beta_\varepsilon} \le K, \text{a birth occurs in } [T - \beta_\varepsilon, T + b_\varepsilon]\right) \le KCa_\varepsilon^{0.6}, \tag{4.14}$$

while repeating the proof of (3.9) gives

$$P\left(\left\|Z_\varepsilon^j(T) - Z_\varepsilon^k(T)\right\|_\infty \le a_\varepsilon^{0.9} \text{ for some } j < k \le K\right) \le K^2Ca_\varepsilon^{0.6d}. \tag{4.15}$$

Combining (4.13)–(4.15) and noticing that if r is large all of the upper bounds are smaller than $K^2 C a_\varepsilon^{0.6}$ gives

$$(4.16)\qquad P\big(\Delta_y^1 \neq 0\big) \leq K^2 C a_\varepsilon^{0.6} = C\varepsilon^{0.6\alpha - 4\gamma} = \varepsilon^{0.55\alpha}$$

since $\gamma = \alpha/80$.

Turning to the other error events, we note that when there are at most K births in each dual, it follows from (2.12) that the probability of a collision in one dual or between the two duals before time T is smaller than

$$(4.17)\qquad KC\varepsilon T^{1/2}\exp(\nu T) \leq C\varepsilon^{1-3\gamma}\log^{1/2}(1/\varepsilon)$$

since $K = \exp(2\nu T) = \varepsilon^{-2\gamma}$. Inequality (2.19) implies that in each dual

$$(4.18)\qquad \begin{aligned} &P\Big(\max_{0\leq s\leq T}\big\|X_\varepsilon^k(s) - B_\varepsilon^k(s)\big\|_\infty > 4\varepsilon^{0.3} \text{ for some } k \leq K\Big) \\ &\quad \leq KC\varepsilon^{0.4}T^{1/2}\exp(\nu T) \leq C\varepsilon^{0.4-3\gamma}\log^{1/2}(1/\varepsilon). \end{aligned}$$

Relation (4.8) implies that when the event in (4.18) does not occur, the probability of a ϕ-error is at most

$$(4.19)\qquad C\varepsilon^{0.295-\gamma}.$$

Finally, (3.12) implies

$$(4.20)\qquad \begin{aligned} &P\big(\text{one of the } K \text{ particles moves by more than } a_\varepsilon^{0.9}/3 \text{ in } [T, T+b_\varepsilon]\big) \\ &\quad \leq KCa_\varepsilon^{0.1p}. \end{aligned}$$

Combining (4.17)–(4.20), taking p large and recalling $\gamma = 1/(800d)$ gives

$$(4.21)\qquad P\big(\Delta_x^2 \neq 0\big) \leq C\varepsilon^{0.293}.$$

Plugging (4.1), (4.17) and (4.21) into (4.12) gives

$$E\prod_i \xi_{y_i} \leq C\Big\{\big(\varepsilon^\lambda + \varepsilon^{1.1d}\big)^q + \varepsilon^{0.293}\Big\}.$$

Now $\alpha = 0.1/d$ so if q is large, it follows from (4.11) that

$$E\big(D_S^\varepsilon(x)\big)^q \leq C\Big\{\big(\varepsilon/a_\varepsilon\big)^d + \varepsilon^{0.293}\Big\} \leq C\varepsilon^{0.293},$$

and using Markov's inequality gives

$$(4.22)\qquad P\big(D_S^\varepsilon(x) > \rho_1/3\big) \leq C\varepsilon^{0.293}.$$

For the conclusion of Lemma 4.3, we are concerned with the density in $(CL/a_\varepsilon)^d$ cubes and $L \leq \varepsilon^{-0.005/d} = \varepsilon^{-0.05\alpha}$ so $(CL/a_\varepsilon)^d \leq C\varepsilon^{-1.05d\alpha} = C\varepsilon^{-0.105}$ and the conclusion follows from (4.22). □

5. Killing off the particles and the process. The next step in the proof of (III) is to show that once the density of occupied sites is below $C\varepsilon^\lambda$, it will drop to zero. To do this we need an estimate (Lemma 5.2) on the number of particles that come into a cube from outside. To prove that estimate we begin with a result about branching random walks.

LEMMA 5.1. *Let $\mathscr{Y}_\varepsilon^0(s,x)$ denote the number of particles at x at time s in the branching random walk Y_ε starting with a single particle at* 0. *If $\varepsilon \le 1$, then*

$$P\big(\mathscr{Y}_\varepsilon^0(t,x) > 0 \text{ for some } x \notin (-\rho t, \rho t)^d\big) \le 2d \exp(-(\rho - \nu - 1)t).$$

REMARK. For a fixed value of ε this result and the proof we give below are well known. We have to give the proof because we need to show that the bound holds uniformly for $\varepsilon \in (0,1]$.

PROOF OF LEMMA 5.1. By considering the coordinates separately, it suffices to prove the result in $d = 1$. Let $m(t,x) = E\mathscr{Y}_\varepsilon^0(t,x)$. By considering the effects of the possible transitions, we see that m solves

$$\frac{\partial}{\partial t} m(t,x) = (\nu - \varepsilon^{-2}) m(t,x) + \frac{\varepsilon^{-2}}{2}(m(t,x+\varepsilon) + m(t,x-\varepsilon)), \tag{5.1}$$

with initial condition $m(0,0) = 1$, $m(0,x) = 0$ otherwise. It is easy to check that the solution of the last equation is given by

$$m(t,x) = \exp(\nu t) P(S_t^\varepsilon = x), \tag{5.2}$$

where S_t^ε is a random walk on $\varepsilon\mathbb{Z}$ that starts at 0 and jumps to a randomly chosen neighbor at rate ε^{-2}. To estimate the probability that the random walk moves by more than ρt we note that if $\theta > 0$, Chebyshev's inequality implies

$$\exp(\theta\rho t) P_0(S_t^\varepsilon > \rho t) \le E \exp(\theta S_t^\varepsilon), \tag{5.3}$$

and direct computation gives

$$\begin{aligned} E \exp(\theta S_t^\varepsilon) &= \sum_{n=0}^\infty \exp(-t\varepsilon^{-2}) \frac{(t\varepsilon^{-2})^n}{n!} \left(\frac{e^{\theta\varepsilon} + e^{-\theta\varepsilon}}{2}\right)^n \\ &= \exp\left(-t\varepsilon^{-2}\left(1 - \frac{e^{\theta\varepsilon} + e^{-\theta\varepsilon}}{2}\right)\right) \\ &= \exp\left(t \sum_{k=1}^\infty \frac{\theta^{2k}\varepsilon^{2k-2}}{(2k)!}\right) \\ &\le \exp(t(e^\theta - 1 - \theta)) \end{aligned} \tag{5.4}$$

when $\theta > 0$ and $0 < \varepsilon \le 1$. Combining (5.2)–(5.4) gives

$$\sum_{x \ge \rho t} m(t,x) \le \exp(t\{-\theta\rho + \nu + (e^\theta - 1 - \theta)\}). \tag{5.5}$$

Now $e < 3$ so taking $\theta = 1$ and multiplying by 2 to take care of $y \le -\rho t$ gives the desired result. □

Recall that $T = (\gamma/\nu)\log(1/\varepsilon)$, where $\gamma = 1/(800)d$, and $S = T + b_\varepsilon$, where $b_\varepsilon = a_\varepsilon^{1.9}$ and $a_\varepsilon = \varepsilon^\alpha$ with $\alpha = 0.1/d$.

LEMMA 5.2. *If we pick A_1 so that $(A_1 - \nu - 2)(\gamma/\nu) \ge 0.1$, then, for small ε,*

$$P\Big(I_\varepsilon^{0,S}(s) \not\subset (-A_1 S, A_1 S)^d \text{ for some } s \le S\Big) \le 4d\varepsilon^{0.1}.$$

PROOF. Using this choice of A_1 in Lemma 5.1 and observing $S \ge T = (\gamma/\nu)\log(1/\varepsilon)$ gives

$$\begin{aligned} &P\Big(\mathscr{Y}_\varepsilon^0(S,x) > 0 \text{ for some } x \notin \big(-(A_1-1)S, (A_1-1)S\big)^d\Big) \\ &\qquad \le 2d\exp\big(-(A_1-\nu-2)S\big) \le 2d\varepsilon^{0.1}. \end{aligned} \tag{5.6}$$

To extend the last conclusion to the dual processes, note that $\exp(\nu T) = \varepsilon^{-\gamma}$ and $b_\varepsilon \le 1$, so using (2.10) and (2.8) with $K = \varepsilon^{-0.2}$ gives

$$P(\mathscr{K}_S > K) \le C\varepsilon^{0.2-\gamma}, \tag{5.7}$$

$$\begin{aligned} &P\Big(\max_{0\le s\le S} \|X_\varepsilon^k(s) - Y_\varepsilon^k(s)\|_\infty \ge 2\varepsilon^{0.3} \text{ for some } 1 \le k \le K\Big) \\ &\qquad \le C\varepsilon^{0.2-\gamma}\log^{1/2}(1/\varepsilon). \end{aligned} \tag{5.8}$$

Combining (5.6)–(5.8) and observing $A_1 S \ge 2\varepsilon^{0.3}$ for small ε gives the desired result. □

Lemmas 5.1 and 5.2 are valid in general. For the next result we will use the assumption that $f'(0) < 0$. The first step is to observe that, by considering (1.2) with small u, it is easy to see that

$$f'(0) = -a(0,\dots,0) + \sum_{j=1}^N b(e_j),$$

where e_j is the jth unit vector in $\mathbb{R}^N$. Letting $\alpha = a(0,\dots,0)$, $\beta = \sum_{j=1}^N b(e_j)$ and $\mu = \alpha - \beta = -f'(0)$, we have the following.

LEMMA 5.3. *Suppose $L \le \varepsilon^{-0.005/d}$ and $0 < \sigma < d$. Let $\eta = \min\{\sigma/12d, \mu\gamma/2\nu\}$. If there are fewer than $\varepsilon^{\sigma-d}$ particles in $[-L, L]^d$ at time 0, then with high probability there are fewer than $\varepsilon^{\sigma+\eta-d}$ particles in $[-L + A_1 S, L - A_1 S]^d$ at time S.*

PROOF. The restriction on L and Lemma 5.2 imply that with high probability

$$I_\varepsilon^{x,S}(s) \subset (-L, L)^d \quad \text{for all } s \le S,\ x \in [-L + A_1 S, L - A_1 S]^d.$$

When no dual escapes from $(-L, L)^d$, the state at time S in $[-L + A_1 S, L - A_1 S]^d$ agrees with $\hat\xi_S$, the state at time S when $\hat\xi_0(x) = \xi_0(x)$ for $x \in (-L, L)^d$ and $\hat\xi_0(x) = 0$ for $x \notin (-L, L)^d$.

To prove the lemma we will show that the number of particles in $\hat\xi_t$ can, with high probability, almost be dominated by a branching process in which particles die at rate α and give birth at rate β. Call an arrival in a Poisson process in the graphical representation an *unusual event* if it occurs at a site with two or more occupied neighbors. Here we consider the site itself to be one of its neighbors so this definition embraces two possibilities: an arrival at an occupied site with one other occupied neighbor or at a vacant site with two occupied neighbors. At unusual events we may have a birth in the particle system or a death in the branching process that does not occur in the other, so to get an upper bound we suppose that each unusual event adds a particle to the branching process which we call an *immigrant*. To estimate the number of immigrants, we begin by observing that (2.1) implies that the expected amount of time two stirred particles spend within distance $M\varepsilon$ of each other up to time t is bounded by

$$\begin{cases} CM\varepsilon t^{1/2}, & d = 1, \\ CM^2\varepsilon^2 \log(t\varepsilon^{-2}), & d = 2, \\ CM^d\varepsilon^2, & d \geq 3, \end{cases} \tag{5.9}$$

when $t\varepsilon^{-2} \geq 2$. Let

$$\tau = \inf\left\{t: t \geq S \text{ or the number of births in } \hat{\xi}_t \geq \varepsilon^{\sigma-\eta-d}\right\}.$$

We will estimate the behavior of the process up to time τ and then show that $\tau = S$ with high probability. To see the reason for restricting our attention to what happens up to time τ, we observe that (5.9) implies that in $d \leq 2$ the expected number of unusual events up to time τ is smaller than

$$C\left(\varepsilon^{\sigma-\eta-d}\right)^2 \varepsilon^{d-\eta} = C\varepsilon^{2\sigma-3\eta-d}. \tag{5.10}$$

The last estimate in (5.9) is too crude in $d \geq 3$, but this is due to overcounting close encounters at very small times, so we can fix it by refining our bookkeeping. We will imagine that the particles initially present are green for the first $3\varepsilon^{\sigma/d}$ units of time, those born at positive time are green for the first $3\varepsilon^{\sigma/d}$ time units of their life and then the particles turn white. Relaxing the earlier definition, we say that an unusual event occurs if there is a Poisson arrival at a site with a green particle in its neighborhood or at a site with two white occupied neighbors. We call the two types of unusual events green and white, respectively. It is easy to see that the expected number of green unusual events by time τ is at most

$$C\varepsilon^{\sigma-\eta-d}\varepsilon^{\sigma/d}. \tag{5.11}$$

To estimate the expected number of white unusual events, we first need to estimate the amount of time two stirred particles spend near each other after time $3\varepsilon^{\sigma/d}$ using the comparison introduced in part (b) of Section 2. Let V_s^ε be the difference of the locations of two stirred particles and let W_s^ε be a random walk that jumps to a randomly chosen neighbor at rate $2d\varepsilon^{-2}$. As noted in subsection (b) of Section 2, a process with the same distribution as $\|V_s^\varepsilon\|_1$ can be made by taking $\|W_s^\varepsilon\|_1$ and removing the visits to 0. Writing $v^{M\varepsilon}(s,t) = |\{r \in [s,t]: \|V_r^\varepsilon\|_1 \le M\varepsilon\}|$ and $w^{M\varepsilon}(s,t) = |\{r \in [s,t]: \|W_r^\varepsilon\|_1 \le M\varepsilon\}|$, we have

$$\begin{aligned}(5.12)\qquad & P\big(v^{M\varepsilon}(3\varepsilon^{\sigma/d}, S) \ge v\big) \\ & \quad \le P\big(w^{M\varepsilon}(\varepsilon^{\sigma/d}, S) \ge v\big) + P\big(w^0(0, 3\varepsilon^{\sigma/d}) \ge 2\varepsilon^{\sigma/d}\big).\end{aligned}$$

To bound $Ev^{M\varepsilon}(3\varepsilon^{\sigma/d}, S)$ we will integrate the last inequality from 0 to S. To bound the second term we observe that $\|W_r^\varepsilon\|_1$ jumps out of 0 at rate $2d\varepsilon^{-2}$ and jumps back in at rate at most ε^{-2}, so a simple large deviations estimate for exponential random variables shows that there are constants $0 < c$, $C < \infty$ independent of ε so that

$$(5.13)\qquad SP\big(w^0(0, 3\varepsilon^{\sigma/d}) \ge 2\varepsilon^{\sigma/d}\big) \le SC\exp(-c\varepsilon^{-2+\sigma/d}),$$

which is incredibly small since $\sigma/d \le 1$ and $S \le 1 + (\gamma/\nu)\log(1/\varepsilon)$. To bound the first term in (5.12) we observe that

$$\begin{aligned}(5.14)\qquad Ew^{M\varepsilon}(\varepsilon^{\sigma/d}, S) &\le C\varepsilon^2 \int_{\varepsilon^{-2+\sigma/d}}^{\varepsilon^{-2}S} t^{-d/2}\,dt \\ &\le C\varepsilon^2\varepsilon^{d-2}\varepsilon^{-\sigma(d-2)/(2d)} \le C\varepsilon^d\varepsilon^{-\sigma/2}.\end{aligned}$$

Combining (5.12)–(5.14), we see that the expected number of white unusual events before time τ is smaller than

$$(5.15)\qquad C\big(\varepsilon^{\sigma-\eta-d}\big)^2\varepsilon^{d-(\sigma/2)} = C\varepsilon^{3\sigma/2-2\eta-d}.$$

Adding (5.15) and (5.11) and comparing with (5.10) gives a bound on the number of unusual events that is valid in any dimension. To combine these estimates we observe that $\eta \le \sigma/(12d) \le \sigma/12$, so

$$\sigma - 3\eta \ge 3\sigma/4 \ge 9\eta, \qquad \sigma/d - \eta \ge 11\eta, \qquad \sigma/2 - 2\eta \ge \sigma/3 \ge 4\eta.$$

Hence the expected number of unusual events is always smaller than

$$(5.16)\qquad C\varepsilon^{\sigma+4\eta-d},$$

and Markov's inequality implies

$$(5.17)\qquad P\big(\text{more than } \varepsilon^{\sigma+\eta-d} \text{ unusual events in } [0,\tau]\big) \le C\varepsilon^{3\eta}.$$

Our next goal is to show that with high probability the number of births in the branching process plus the number of children of the immigrants does not exceed $\varepsilon^{\sigma-\eta-d}$ and hence $\tau \ge S$. This is a routine computation for subcritical branching processes. Let Z_0 be the number of particles at time 0 plus the number of immigrants, and suppose that $Z_0 \le 2\varepsilon^{\sigma-d}$. [(5.17) implies that this is valid with high probability.] Since each particle gives birth at rate β and dies at rate α, it has a geometrically distributed number of children with mean $\beta/\alpha < 1$. If we let Z_n be the number of particles in generation n, then $EZ_n = (\beta/\alpha)^n EZ_0$ and hence

$$(5.18)\qquad E\sum_{n=0}^{\infty} Z_n \le C\varepsilon^{\sigma-d}.$$

Using the last result and Markov's inequality it follows that $\tau \geq S$ with high probability.

Having established that $\tau \geq S$ with high probability, the rest is routine. The mean number of particles at time S in the branching process is at most

$$\varepsilon^{\sigma-d} e^{-\mu S} \leq \varepsilon^{\sigma+2\eta-d}$$

since $\eta \leq \mu\gamma/2\nu$. So Markov's inequality implies that for the branching process

$$P\left(\text{more than } \tfrac{1}{2}\varepsilon^{\sigma+\eta-d} \text{ particles at time } S\right) \leq C\varepsilon^{\eta}. \tag{5.19}$$

As for the immigrants and their children, (5.16) bounds the expected number of immigrants and, since the branching process is subcritical, it also bounds the expected number of children they have at time S. Another use of Markov's inequality tells us that for the immigrants and their children

$$P\left(\text{more than } \tfrac{1}{2}\varepsilon^{\sigma+\eta-d} \text{ particles at time } S\right) \leq C\varepsilon^{3\eta}. \tag{5.20}$$

Adding (5.19) and (5.20) gives the desired result. □

We are now ready to put the pieces together and prove (III). Suppose that ξ_0^ε has density at most $\rho_1/3$ in $[-S, S]^d$. If ε is small, it follows from Lemma 4.1 and Lemma 4.3 that, for any j,

$$P\left(\xi_{jS}^\varepsilon(x) = 1\right) \leq C(\varepsilon^\lambda + j\varepsilon^{0.005}) \quad \text{for } x \in \left[-S(1+jw), S(1+jw)\right]^d. \tag{5.21}$$

From the last result and Chebyshev's inequality it follows [recall $\lambda < \gamma = 1/(800d)$] that if we pick J and then ε small, then with high probability

$$\xi_{jS}^\varepsilon \text{ has } \leq C\varepsilon^{\lambda/2-d} \text{ particles in } \left[-S(1+jw), S(1+jw)\right]^d \tag{5.22}$$

for $1 \leq j \leq J$. Now if $\sigma \geq \lambda/2$ in Lemma 5.3, then $\eta \geq \delta > 0$. Pick j_0 so that $(j_0 - 1)\delta > d$ and then pick J_0 so that $(J_0 w - j_0 A_1) \geq 2$. Using Lemma 5.3 repeatedly now shows that with high probability

$$\xi_{jS}^\varepsilon \text{ has no particles in } \left[-S\{1+jw-(j_0-1)A_1\}, S\{1+jw-(j_0-1)A_1\}\right]^d \tag{5.23}$$

for $J_0 \leq j \leq 2J_0 - 1$. Letting

$$\mathscr{D} = \bigcup_{j=J_0}^{2J_0-1} \left[-S\{1+jw-j_0A_1\}, S\{1+jw-j_0A_1\}\right]^d \times [jS, (j+1)S] \tag{5.24}$$

and using Lemma 5.2 now gives that with high probability

$$\xi_t^\varepsilon \text{ has no particles in } \mathscr{D}. \tag{5.25}$$

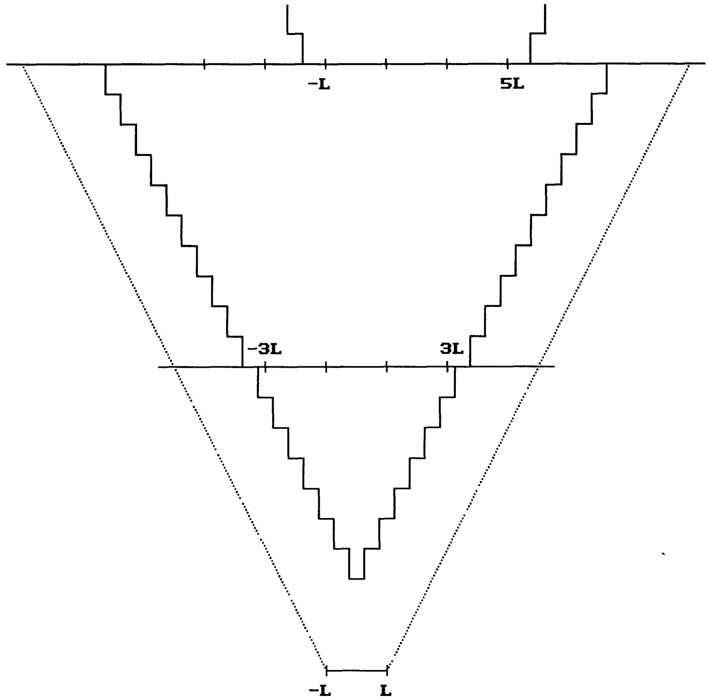

FIG. 2.

The choice of J_0 implies that the bottom of $\mathscr{D}$ contains $[-3S, 3S]^d \times \{J_0 S\}$ and the top contains $[-5S, 5S]^d \times \{2J_0 S\}$ [since $(2J_0 w - j_0 A_1) \geq 2(J_0 w - j_0 A_1) \geq 4$]. See Figure 2 for a picture of the case $d = 1$.

To prove (III) we will compare with an M dependent oriented percolation process to show that if there is a large enough vacant region, then with high probability it will grow linearly and so the process will die out. We will give the proof first in the simple case $d = 1$. As in Section 3, to achieve a finite range of dependence we have to restrict our attention to duals that don't move too much. To see that this can be done, we note that all the events of interest in (5.22)–(5.25) can be determined from dual processes $I_\varepsilon^{x,t}(s)$, $s \leq S$, where

$$(x, t) \in \bigcup_{j=1}^{2J_0} [-S(1 + jw), S(1 + jw)] \times \{jS\}$$

and that Lemma 5.2 implies that with high probability all these duals will stay in

$$\mathscr{E} = \bigcup_{j=1}^{2J_0} [-(1 + jw + A_1)S, (1 + jw + A_1)S] \times [(j - 1)S, jS].$$

To make the comparison with oriented percolation on the lattice $\mathscr{L} = \{(m,n) \in \mathbb{Z}^2\colon m+n \text{ is even}\}$ defined in Section 4, we will use the general setup from the end of Section 3. This time there is only one good configuration $\xi \equiv 0$ and the good event G_0 has two parts: All the duals of interest stay in $\mathscr{E}$ and there are no particles in $\mathscr{D}$. We apply the comparison to the system starting from an initial configuration with the same distribution as $\bar{\xi}_S^{\varepsilon}$, the state at time S when we start with all sites occupied at time 0. Let $L = S$, $T = J_0 S$ and $V_n = \{(x,n) \in \mathscr{L}\colon$ there are no particles in $xL + [-L, L]$ at time $nT\}$. If $x \in V_0$, let $W_n^x = \{y\colon (x,0) \to (y,n)\}$, where $\to$ is the connectivity relation defined in Section 3. Let $l_n^x = \inf W_n^x$ and $r_n^x = \sup W_n^x$. It follows from the remark after (5.22) that the next lemma holds.

Lemma 5.4. *If $W_n^x \neq \varnothing$, all the sites in $[Ll_n^x - 3L, Lr_n^x + 3L]$ are vacant at time $(n+1)J_0 S$.*

Proof. Let $(x_0, 0), (x_1, 1), \dots, (x_n, n)$ be an open path from $(x,0)$ to (r_n^x, n). For $1 \le k \le n-1$, $|x_k - x_{k-1}| = 1$ so the top of $(Lx_{k-1}, (k-1)J_0 S) + \mathscr{D}$ intersects the bottom of $(Lx_k, kJ_0 S) + \mathscr{D}$ in an interval of length at least $6L$. The union of the $(Lx_k, kJ_0 S) + \mathscr{D}$ for $0 \le k \le n$ gives a region Δ_r of width greater than or equal to $6L$ (r is for right) in which there are no particles. Repeating the construction for l_n^x gives another region Δ_l that we know to be vacant. There are no "spontaneous births" so it follows that there are no particles inside $\Delta_l \cup \Delta_r$, that is, in the bounded components of $(0, (n+1)J_0 S) \cap (\Delta_l \cup \Delta_r)^c$. To see that this contains the indicated interval note that $\Delta_l \cup \Delta_r$ contains $[(r_n^x - 3)L, (r_n^x + 3)L] \times \{(n+1)J_0 S\}$ and $[(l_n^x - 3)L, (l_n^x + 3)L] \times \{(n+1)J_0 S\}$.

It follows from Lemma 3.5 that if $\theta < 6^{-4(2M-1)^2}$, then

$$(5.26) \qquad P(W_{2n}^x \neq \varnothing) \ge 1 - \theta - 162\theta^{1/4(2M-1)^2}.$$

To complete the proof of (III) we will show that

$$(5.27) \qquad l_n^x \to -\infty, \qquad r_n^x \to \infty \quad \text{a.s. on } \Omega^x \equiv \{W_n^x \neq \varnothing \text{ for all } n\}$$

and then show that the last event occurs for some $x \in V_0$. To prove (5.27) we introduce $W_n^- = \{y\colon (x,0) \to (y,n) \text{ for some } x \le 0\}$ and $r_n = \sup W_n^-$. It is easy to verify that

$$(5.28) \qquad r_n^0 = r_n \quad \text{on } \{W_n^0 \neq \varnothing\}.$$

[See (2) on page 1003 in Durrett (1984).] A straightforward modification of the proof of (2) on page 1030 in Durrett (1984) gives the following.

Lemma 5.5. *If $\theta < 6^{-8(2M-1)^2}$, then $P(r_{2n} \le n) \le 6 \cdot 2^{-2n}$.*

Proof. As indicated in the source cited, (i) there is a contour that starts at $(0, -1)$ and ends at $(r_{2n} + 1, 2n)$, (ii) any possible contour has length $2n + 1 + k$ for some integer $k \ge 0$, (iii) for such a contour to exist there must be at least $(n + k)/4$ closed sites, (iv) we can find a subset of size $(n + k)/4(2M - 1)^2$ that satisfies the hypotheses of (3.20) and (v) there are at most 3^m contours of length m. So

$$P(r_{2n} \le n) \le \sum_{k=0}^{\infty} 3^{2n+1+k} \theta^{(n+k)/4(2M-1)^2} \le 3 \cdot 2^{-2n} \sum_{k=0}^{\infty} 12^{-k}. \qquad \square$$

Lemma 5.5 and the Borel–Cantelli lemma imply that

$$\liminf_{n \to \infty} r_{2n}/n \ge 1/2 \quad \text{a.s.} \tag{5.29}$$

Combining the last result with (5.28) and using translation invariance and symmetry gives (5.27). To complete the proof of Theorem 4, it suffices to show the following.

LEMMA 5.6. *With probability* 1 *there is an* $x \in V_0$ *with* $\{W_n^x \ne \varnothing$ *for all* $n\}$.

PROOF. To prove this it is convenient to suppose that the Poisson processes used in the construction given in Section 2 are defined on $[-S, \infty)$, so that we can use the arrivals in $[-S, 0)$ to compute the initial state of the particle system, and then use the other arrivals to compute the evolution. Let

$$\mathscr{P}_x = (\{S_n^{x, x+\varepsilon}, n \ge 1\}, \{T_n^x, n \ge 1\}, \{U_n^x, n \ge 1\})$$

be the ingredients for the graphical representation associated with the site x. The indicator function of the event $\{0 \in V_0\} \cap \Omega^0$ is a function ϕ of the graphical representation $\mathscr{P}_x$, $x \in \varepsilon\mathbb{Z}$. If we translate the graphical representation by $2xL$ and apply ϕ, then we get the indicator function of the event $\{x \in V_0\} \cap \Omega^x$. The family $\mathscr{P}_x$, $x \in \varepsilon\mathbb{Z}$, is an i.i.d. sequence (taking values in a rather large but nice space) and hence ergodic: Any event invariant under all the shifts θ_{2xL}, $x \in 2\mathbb{Z}$, is trivial. It is well known that functions of ergodic sequences are ergodic. [See Breiman (1968), Proposition 6.31 on page 119, or Durrett (1992), Theorem 1.3 on page 295.] From the last observation and the ergodic theorem it follows that, as $L \to \infty$,

$$\frac{1}{2L+1} \sum_{k=-L}^{L} 1_{\{2k \in V_0\} \cap \Omega^{2k}} \to P(\{0 \in V_0\} \cap \Omega^0) > 0 \quad \text{a.s.}$$

This proves the lemma and completes the proof of (III) in $d = 1$. $\square$

The proof of Lemma 5.4 relies heavily on the fact that we are in $d = 1$. To get our vacant region to grow in $d > 1$ we have to compare with a different percolation process. Let $D = d + 1$ and let A be a $D \times D$ matrix so that (i) if $x_1 + \cdots + x_D = 1$, then $(Ax)_d = 1$, and (ii) if x and y are orthogonal, then Ax and Ay are. Let $e_1, \dots, e_D$ be the D unit vectors and let $v_i = Ae_i$. Let

$\mathscr{L}_D = \{Az: z \in \mathbb{Z}^D\}$ and make $\mathscr{L}_D$ into a graph by drawing an oriented arc from x to $x + v_i$ for $1 \le i \le D$.

Let u_i, $1 \le i \le d$, be the vectors in $\mathbb{R}^d$ that consist of the first d components of the v_i. It is easy to see that there is a constant U so that $\|u_i\|_2 = U$. Let $B_2(x, r) = \{y \in \mathbb{R}^d: \|y - x\|_2 \le r\}$ and $L = S$. If $(z, n) \in \mathscr{L}_D$ we say z is good at time n and put $z \in V_n$ if all the sites in $B_2(zL/U, L)$ are vacant at time nS. It is straightforward to define oriented percolation on $\mathscr{L}_D$ and follow the approach of the proof for $d = 1$ to show that the wet sites on level W_n satisfy $W_n \subset V_n$ if $W_0 = V_0$. However, this result is not good enough for the desired conclusion. Let w_i, $1 \le i \le D(D-1)$, be the vectors in $\mathbb{R}^D$ that have one $+1$ component, one -1 component and $D - 2$ zeros. These are the vectors in the hyperplane $x_1 + \cdots + x_D = 0$ that are the closest to 0. Let $\mathscr{W} = \{Aw_i: 1 \le i \le D(D-1)\}$ and $\mathscr{V} = \{v_i: 1 \le i \le D\}$. A little thought reveals that if $(z, n) \in \mathscr{L}_D$ and there is an occupied site in $B_2(zL/U, L)$ at time nS, then we can find a sequence (z_m, k_m), $0 \le m \le m_0$, of points not in $W = \bigcup_n W_n$ so that $z_0 = z$, $k_0 = n$, $k_{m_0} = 0$, k_m is nonincreasing, $z_m - z_{m-1} \in \mathscr{V}$ when $k_m < k_{m-1}$, and $z_m - z_{m-1} \in W$ when $k_m = k_{m-1}$. If such a path exists we say that z is *exposed at time* n. The next result follows easily from the proof of Lemma 4.3 in Durrett (1992). Here θ is the probability a site in W is closed.

LEMMA 5.7. *If θ is sufficiently small, there is a $c > 0$ so that on Ω_x if n is large, then there are no sites in $B_2(x, cn)$ that are exposed at time n.*

With this in hand the proof can be completed by generalizing Lemma 5.6 to $d > 1$.

APPENDIX

In this section we will prove Propositions 1 and 2 for a suitable general class of equations

$$(*) \qquad \frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + f(u).$$

We begin by stating our assumptions on f. The function f has a continuous derivative. There are $p_0 < p_1 < p_2$ so that $f(p_i) = 0$ for $i = 0, 1, 2$, $f(x) > 0$ for $x \in (p_1, p_2)$, $f(x) < 0$ for $x \in (p_0, p_1)$, $f'(p_0) < 0$ and $f'(p_2) < 0$. Under these assumptions there is, in $d = 1$, a travelling wave solution $u(t, x) = w(x - rt)$ that is a strictly decreasing function from $(-\infty, \infty)$ onto (p_0, p_2) and has

$$\text{(A.1)} \qquad |w'(x)| \le Ke^{-\lambda|x|}.$$

[See Fife and McLeod (1977), Section 2 and page 349.] Recall from the introduction that the sign of r agrees with the sign of $\int_{p_0}^{p_2} f(x)\,dx$. Our first result assumes more about the initial condition than Proposition 1 does but easily implies that result and Proposition 2 as well.

LEMMA A.1. *Suppose $r > 0$ and let $\delta < r$. Let $\eta > 0$ so that $p_2 - 2\eta > p_1$ and $f(u) > 0$ for $u \in [-2\eta, 0)$. If $L \geq L_\eta$, then there are constants $0 < c$, $C < \infty$ so that, for any initial condition with $u(0,x) \geq p_0 - \eta$ for all x and $u(0,x) \geq p_2 - \eta$ when $|x| \leq L$, we have*

$$\inf_{x:\,|x| \leq (r-\delta)t} u(t,x) \geq p_2 - Ce^{-ct}.$$

PROOF. Our approach is to generalize the proof of Lemma 4.1 in Fife and McLeod (1977) to dimensions $d > 1$. Based on that proof, one's first guess is to look at $w(|x| - rt + \tau(t)) - q(t)$, but $|x|$ is not smooth at 0 so we need to modify the solution there. Let

$$h(x) = \begin{cases} x^2 - (x^3/3), & \text{if } 0 \leq x \leq 1, \\ x - 1/3, & \text{if } x \geq 1. \end{cases}$$

To see the reason for this choice note that

$$h'(x) = \begin{cases} 2x - x^2, & \text{if } 0 \leq x \leq 1, \\ 1, & \text{if } x \geq 1; \end{cases}$$

$$h''(x) = \begin{cases} 2 - 2x, & \text{if } 0 \leq x \leq 1, \\ 0, & \text{if } x \geq 1. \end{cases}$$

so h is C^2. For the future note that

$$h'(x) \leq 1, \qquad h''(x) \leq 2, \qquad h'(x) \leq 2x.$$

Let $v(t,x) = w(h(|x|) - rt + \tau(t)) - q(t)$. To prove our result, we will pick $q_0 = 2\eta$; let

$$\tau(t) = -A + B\log(1+t), \qquad q(t) = q_0 e^{-\mu t}$$

and show that if we choose A, B and μ appropriately, then $v(t,x)$ is a subsolution of $(*)$, that is,

$$Nv \equiv \frac{\partial v}{\partial t} - \frac{1}{2}\Delta v - f(v) \leq 0.$$

The definition of v guarantees that if $L \geq L_\eta$, then

$$u(0,x) \geq v(0,x) = w(h(|x|) - A) - 2\eta;$$

so application of a standard comparison theorem shows $u(t,x) \geq v(t,x)$ and the desired conclusion follows since (A.1) implies that $w(x) \to p_2$ exponentially rapidly as $x \to -\infty$.

The first step in computing Nv is to observe that

$$\frac{d}{dx_i}g(h(|x|)) = g'(h(|x|))h'(|x|)\left(\sum x_i^2\right)^{-1/2}x_i,$$

$$\frac{d^2}{dx_i^2}g(h(|x|)) = g''(h(|x|))h'(|x|)^2\left(\sum x_i^2\right)^{-1}x_i^2 + g'(h(|x|))h''(|x|)\left(\sum x_i^2\right)^{-1}x_i^2$$
$$+ g'(h(|x|))h'(|x|)\left\{-\left(\sum x_i^2\right)^{-3/2}x_i^2 + \left(\sum x_i^2\right)^{-1/2}\right\},$$

$$\Delta g(h(|x|)) = g''(h(|x|))h'(|x|)^2 + g'(h(|x|))\left\{h''(|x|) + h'(|x|)\frac{d-1}{|x|}\right\}.$$

Taking $g(y) = w(y - rt + \tau(t))$, writing ζ for $h(|x|) - rt + \tau(t)$, recalling the definition of N and then using $-w''/2 = rw' + f(w)$, we have

$$Nv = (-r + \tau')w'(\zeta) - q'(t) - \frac{1}{2}w''(\zeta)h'(|x|)^2$$
$$- \frac{1}{2}w'(\zeta)\left\{h''(|x|) + h'(|x|)\frac{d-1}{|x|}\right\} - f(w(\zeta) - q)$$

$$(\text{A.2}) \quad = \frac{1}{2}w'(\zeta)\left\{-2r + 2\tau' + 2rh'(|x|)^2 - h''(|x|) - h'(|x|)\frac{d-1}{|x|}\right\}$$
$$+ f(w(\zeta))h'(|x|)^2 - f(w(\zeta) - q) - q'(t)$$
$$= \frac{1}{2}w'(\zeta)\left\{2\tau' - h'(|x|)\frac{d-1}{|x|} - h''(|x|) + \left(h'(|x|)^2 - 1\right)2r\right\}$$
$$+ \left(h'(|x|)^2 - 1\right)f(w(\zeta)) + \{f(w(\zeta)) - f(w(\zeta) - q)\} - q'(t).$$

Three of the terms are easy to deal with since $h'(|x|) = 1$ and $h''(|x|) = 0$ for $|x| \geq 1$. To begin to deal with the others let

$$F(u,q) = \begin{cases} \{f(u-q) - f(u)\}/q, & \text{if } q > 0, \\ -f'(u), & \text{if } q = 0. \end{cases}$$

Now $F(p_2, q) > 0$ for $0 \leq q \leq q_0$ (since $p_2 - q_0 > p_1$) and F is continuous, so there are constants μ_0, $\alpha_0 > 0$ (with $\alpha_0 < q_0$) so that $F(u,q) \geq \mu_0$ when $0 \leq q \leq q_0$ and $p_2 - \alpha_0 \leq u \leq p_2$. That is,

$$(\text{A.3}) \quad f(u) - f(u-q) \leq -\mu_0 q \quad \text{for } 0 \leq q \leq q_0,\ p_2 - \alpha_0 \leq u \leq p_2.$$

The last argument also works near 0. Since $q_0 = 2\eta$, our assumptions imply $F(0,q) > 0$ for $0 \leq q \leq q_0$ and F is continuous, so there are constants μ_1, $\alpha_1 > 0$ so that

$$(\text{A.4}) \qquad f(u) - f(u-q) \leq -\mu_1 q \quad \text{for } 0 \leq q \leq q_0,\ 0 \leq u \leq \alpha_1.$$

We will now choose our constants. The reasons for these choices will become clear as we proceed. Let

$$\mu \le \min\{\mu_0/2, \mu_1, \lambda r/3\}. \tag{A.5}$$

To pick B, we begin by observing that w is strictly decreasing and w' is continuous (in fact w'' exists), so there is a $\beta > 0$ so that

$$w'(z) \le -\beta < 0 \quad \text{when } \alpha_1 \le w(z) \le p_2 - \alpha_0. \tag{A.6}$$

Let $\kappa = \sup\{f'(z): z \in [p_0 - q_0, p_2]\}$ and note that

$$f(w(\zeta)) - f(w(\zeta) - q) \le \kappa q \quad \text{for } q \le q_0. \tag{A.7}$$

Pick B large enough so that, for all t,

$$\frac{B}{t+1} - \frac{d-1}{1 + rt/3} \ge 0, \tag{A.8}$$

$$\frac{B\beta}{2(t+1)} \ge \kappa q_0 e^{-\mu t}. \tag{A.9}$$

Next pick D large enough so that

$$w(-D) \ge p_2 - \alpha_0, \tag{A.10}$$

$$(d + r)K \exp(-\lambda D) \le \mu_0 q_0/2, \tag{A.11}$$

and pick A large enough so that

$$\tau(t) = -A + B\log(1 + t) \le -1 - D + (rt/3) \quad \text{for all } t. \tag{A.12}$$

To verify that $Nv \le 0$ we suppose first that $|x| \le 1 + (rt/3)$ and note that $h(|x|) \le |x|$ and (A.12) imply

$$\zeta = h(|x|) - rt + \tau(t) \le -D - (rt/3), \tag{A.13}$$

and hence $w(\zeta) \ge p_2 - \alpha_0$ by the choice of D in (A.10). Since τ is increasing, $h''(x) \le 2$, and $h'(x) \le 2x$, we have

$$\begin{aligned} &\left\{2\tau' - h'(|x|)\frac{d-1}{|x|} - h''(|x|) + \left(h'(|x|)^2 - 1\right)2r\right\} \\ &\qquad \ge -2(d-1) - 2 - 2r = -2d - 2r. \end{aligned} \tag{A.14}$$

Now $h'(x) \le 1$ and $f(w(\zeta)) > 0$ when $w(\zeta) \ge p_2 - \alpha_0 \ge p_1$ (by the choices of α_0 and q_0), so

$$\left(h'(|x|)^2 - 1\right) f(w(\zeta)) \le 0. \tag{A.15}$$

Combining the last two inequalities with (A.3) and recalling $q(t) = q_0 e^{-\mu t}$ (so $q' = -\mu q$) converts (A.2) into [recall $w'(x) < 0$]

$$\begin{aligned} Nv &\le -(d + r)w'(\zeta) + 0 - \mu_0 q + \mu q \\ &\le (d + r)K \exp(\lambda\{-D - (rt/3)\}) - (\mu_0 - \mu)q_0 e^{-\mu t} \le 0, \end{aligned} \tag{A.16}$$

by (A.13), (A.1), the choice of μ in (A.5) and the choice of D in (A.11).

The next case to consider is $w(\zeta) \geq p_2 - \alpha_0$ and $|x| \geq 1 + (rt/3)$. In this case $h''(|x|) = 0$ and $h'(|x|) = 1$, so the upper bound in (A.2) simplifies considerably. To bound what is left recall $\tau(t) = -A + B\log(1+t)$; note that our choice of B in (A.8) implies

$$(A.17)\qquad 2\tau' - h'(|x|)\frac{d-1}{|x|} \geq \frac{2B}{t+1} - \frac{d-1}{1+(rt/3)} \geq \frac{B}{t+1} > 0,$$

so repeating the proof of (A.16) but using (A.17) and $w'(\zeta) < 0$ instead of (A.14) gives

$$(A.18)\qquad Nv \leq -(\mu_0 - \mu)q_0 e^{-\mu t} \leq 0.$$

Turning to $w(\zeta) \leq p_2 - \alpha_0$, we recall that $|x| \leq 1 + (rt/3)$ implies $w(\zeta) \geq p_2 - \alpha_0$, so in this case we must have $|x| \geq 1$. When $w(\zeta) \leq \alpha_1$, repeating the proof of (A.18) but using (A.4) instead of (A.3) shows

$$(A.19)\qquad Nv \leq -(\mu_1 - \mu)q_0 e^{-\mu t} \leq 0,$$

by the choice of μ in (A.5). To deal with the intermediate range $\alpha_1 < w(\zeta) < p_2 - \alpha_0$, we note that (A.6), (A.17), (A.7) and $q'(t) \geq 0$ imply that

$$(A.20)\qquad Nv \leq -\frac{\beta B}{2(1+t)} + \kappa q \leq 0,$$

by the choice of B in (A.9). Combining (A.16) and (A.18)–(A.20), we have shown that $Nv \leq 0$ in all cases and this gives the desired result. □

To get Proposition 1 now it suffices to prove the following.

LEMMA A.2. *Let η_0 be a possible choice of η in Lemma* A.1. *Let $\beta \in (p_1, p_2)$ and $\alpha < p_0$ be such that $f(x) > 0$ for $x \in [\alpha, p_0)$. There are constants T and $L_{\alpha,\beta}$ so that if $u(0,x) \geq \alpha$ for all x and $u(0,x) \geq \beta$ when $|x| \leq L_{\alpha,\beta} + L_{\eta_0}$, then $u(T, X)$ satisfies the hypotheses of Lemma A.1 with $\eta = \eta_0$.*

PROOF. Our assumption on α implies that the solution of $v_\alpha'(t) = f(v_\alpha)$ with $v_\alpha(0) = \alpha$ has $v_\alpha(t) \geq p_0 - \eta_0$ for $t \geq T_1$. Since $u(0,x) \geq v_\alpha(0)$, a standard comparison argument implies that for $t \geq T_1$ we have $u(t,x) \geq v_\alpha(t) \geq p_0 - \eta_0$.

Our assumption on β implies that the solution of $v_\beta'(t) = f(v_\beta)$ with $v_\beta(0) = \beta$ has $v_\beta(t) \geq p_2 - \eta_0/2$ for $t \geq T_2$. Let $T = \max\{T_1, T_2\}$. Having fixed T we can now pick $L_{\alpha,\beta}$ large enough so that, for the branching Brownian motion considered in Section 2, the probability that all the particles stay within $L_{\alpha,\beta}$ of the starting point up to time T is at least $\eta_0/2$. It follows from results at the end of subsection (d) in Section 2 that when $|x| \leq L_{\eta_0}$, we have $u(T,x) \geq v_\beta(T) - \eta_0/2 \geq p_2 - \eta_0$, by the choice of T, and the proof of the lemma is complete.

□

To prove Proposition 2 we let $\bar{u}(t,x) = -u(t,x)$ and $\bar{f}(x) = -f(-x)$, so

$$(**)\qquad \frac{\partial \bar{u}}{\partial t} = \frac{1}{2}\Delta\bar{u} + \bar{f}(\bar{u}).$$

Letting $\bar{p}_i = -p_{2-i}$ it is easy to check that $\bar{f}$ satisfies our assumptions. Having multiplied f by -1, there is a travelling wave solution $\bar{w}(x + rt)$ of $(**)$ that moves at rate $-r$, and the desired result follows from our two lemmas.

REFERENCES

ARONSON, D. G. and WEINBERGER, H. F. (1975). Nonlinear diffusion in population genetics, combustion and nerve propagation. *Partial Differential Equations and Related Topics. Lecture Notes in Math.* **446** 5–49. Springer, New York.

ARONSON, D. G. and WEINBERGER, H. F. (1978). Multidimensional nonlinear diffusion arising in population genetics. *Adv. in Math.* **30** 33–76.

ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, New York.

BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, MA.

CHEN, H. N. (1992). On the stability of a population growth model with sexual reproduction on $\mathbb{Z}^d$. *Ann. Probab.* **20** 232–285.

DAB, D., LAWNICZAK, A., BOON, J. P. and KAPRAL, R. (1990). Cellular-automaton model for reactive systems. *Phys. Rev. Lett.* **64** 2462–2465.

DE MASI, A., FERRARI, P. A. and LEBOWITZ, J. L. (1985). Rigorous derivation of reaction–diffusion equations with fluctuations. *Phys. Rev. Lett.* **55** 1947–1949.

DE MASI, A., FERRARI, P. A. and LEBOWITZ, J. L. (1986). Reaction–diffusion equations for interacting particle systems. *J. Statist. Phys.* **44** 589–644.

DE MASI, A. and PRESUTTI, E. (1992). *Mathematical Methods and Hydrodynamic Limits. Lecture Notes in Math.* **1501**. Springer, New York.

DICKMAN, R. and TOMÉ, T. (1991). First order phase transition in a one dimensional nonequilibrium model. *Phys. Rev. A* **44** 4833.

DURRETT, R. (1980). On the growth of one dimensional contact processes. *Ann. Probab.* **8** 890–907.

DURRETT, R. (1984). Oriented percolation in two dimensions. *Ann. Probab.* **12** 999–1040.

DURRETT, R. (1985). Some peculiar properties of a particle system with sexual reproduction. *Stochastic Spatial Processes. Lecture Notes in Math.* **1212**. Springer, New York.

DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth, Pacific Grove, CA.

DURRETT, R. (1990). *Probability: Theory and Examples*. Wadsworth, Pacific Grove, CA.

DURRETT, R. (1992). Multicolor particle systems with large threshold and range. *J. Theoret. Probab.* **5** 127–152.

DURRETT, R. and GRAY, L. (1985). Some peculiar properties of a particle system with sexual reproduction. Unpublished manuscript. [For a statement of the results see Durrett (1985).]

FIFE, P. C. and MCLEOD, J. B. (1977). The approach of solutions of nonlinear diffusion equations to travelling front solutions. *Arch. Rat. Mech. Anal.* **65** 335–361.

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

LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.

NEUHAUSER, C. (1992). The long range sexual reproduction process. Preprint.

NOBLE, C. (1989). Ph.D. dissertation, Cornell Univ.

SCHLÖGL, F. (1972). Chemical reaction models for non-equilibrium phase transitions. *Z. Phys.* **253** 147–161.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CALIFORNIA 90089-1113