

# THE 1996 WALD MEMORIAL LECTURES

## STOCHASTIC MODELS OF INTERACTING SYSTEMS<sup>1</sup>

BY THOMAS M. LIGGETT

*University of California, Los Angeles*

Interacting particle systems is by now a mature area of probability theory, but one that is still very active. We begin this paper by explaining how models from this area arise in fields such as physics and biology. We turn then to a discussion of both older and more recent results about them, concentrating on contact processes, voter models, and exclusion processes. These processes are among the most studied in the field, and have the virtue of relative simplicity in their description, which permits us to address the fundamental issues about their behavior without dealing with the extra complications that models from specific areas of application would require.

**1. Introduction.** It is customary to begin the Wald Lectures by drawing attention to connections between the subject matter of the lectures and the work of Abraham Wald. While Wald's major impact was on the field of statistics, he also made significant contributions to probability theory, particularly in the area of limit theorems. Almost fifty years have now passed since his death, however, and this passage of time and the phenomenal progress that has occurred during the past half century make it increasingly difficult to draw close parallels to current work. So, I would like to begin with some personal comments about connections with the Wald Lectures themselves, and about some of the giants of my field.

The first Wald Memorial Lectures were given by Samuel Karlin in 1957. Ten years later, I became his Ph.D. student. While I ended by working on invariance principles, which are in fact rather close in spirit to Wald's own work on limit theorems, Karlin initially encouraged me to work on total positivity, an area closely related to the subject of his Wald Lectures: "Pólya type theory." Nearly forty years after those lectures, Sam is still very active, now working in mathematical biology.

Many people have thought that my thesis advisor was Frank Spitzer. While I did not meet Frank until a year or two after receiving my degree, there is good reason for that impression. At least half of my work has been fairly closely related to models he created and questions he asked about them. Frank was the twenty-first Wald Lecturer, and, in fact, his 1979

---

Received June 1996; revised August 1996.

<sup>1</sup> Preparation of this paper was supported in part by NSF Grant DMS-94-00644.

AMS 1991 *subject classification*. 60K35.

*Key words and phrases*. Stochastic Ising model, contact processes, voter models, exclusion processes.

lectures were partly based on work he and I had done together. Many of us lost a good friend and mentor when Frank died in 1992.

In my book on interacting particle systems, I said that “much of the impetus” for this field came from the work of Frank Spitzer in the United States and of Roland Dobrushin in the Soviet Union in the late 1960’s. I did not meet Roland until much later, but had several opportunities to be with him in recent years, including a wonderful dinner in his apartment in Moscow in 1992. I was greatly saddened by his death there last fall. My lectures this week will give us an opportunity to remember not only Abraham Wald, one of the founders of modern statistics, but also Frank Spitzer and Roland Dobrushin, who like Wald, were leaders in the development of a mathematical subject with strong connections to, and motivations from, the natural and social sciences.

When I started working on interacting particle systems twenty-five years ago, there were only a handful of papers on the subject. By 1980, papers were appearing at the rate of one per month, and, by now, the rate is even higher. It has become impossible to survey the entire subject in a series of lectures, or even in a book for that matter, so I will only be able to touch on what I regard as some of the high points of recent developments. I propose to begin by describing some of the areas of application in which models of this type arise naturally. I will then specialize to the contact process, a model for the spread of infection that has played a central role in the theory since its introduction by Ted Harris twenty years ago and which still maintains its preeminent position in the field. Following this, we will discuss voter models, which have seen a resurgence of interest since new forms of them were introduced by Ted Cox and Rick Durrett five years ago. In the final lecture, I will describe some recent developments in work on the exclusion process, including connections to partial differential equations, Ulam’s problem on the length of the longest increasing sequence in a random permutation and queueing systems.

**2. The general setup and some examples.** Suppose  $S$  is a countable set of sites (usually  $Z^d$ ), and  $K$  is a finite set of possible states at a site. Then  $X = K^S$  is the set of configurations of a system and will be the state space for the models that we will consider. The models will be continuous time Markov processes  $\eta_t$ , which are described by specifying the rates at which the system changes from one configuration to another. Changes are generally local, in that only one or two sites change state at any given time, and the rates for such transitions depend on the configuration near those sites. If  $|K| = 2$  and changes occur at one site at a time, for example, then the rate at which a change at  $x \in S$  occurs when the configuration is  $\eta \in X$  will be denoted by  $c(x, \eta)$ :

$$P^n(\eta_t(x) \neq \eta) = c(x, \eta)t + o(t), \quad t \downarrow 0.$$

As we will see, many of the models we will discuss and the questions we will ask about them (and even the corresponding answers) can be described in rather simple terms. The proofs that these are in fact the correct answers

are generally not so easy, so we will not dwell on them here. Nevertheless, this contrast between simple statements and hard proofs remains one of the charms of the subject. A general treatment of this subject as of a decade ago can be found in my book [Liggett (1985)]. More recent books on special models or particular aspects of the subject include Durrett (1988), Chen (1992), DeMasi and Presutti (1991), Konno (1994) and Spohn (1991).

Dobrushin's (1971a, b) interest in models of this type came from statistical physics. One can model a magnet, for example, in the following way: take  $S = Z^d$ , the  $d$ -dimensional integer lattice (the set of iron atoms) and  $K = \{-1, +1\}$  (the possible spins of an atom). Let  $\beta$  be a positive parameter (the reciprocal of the temperature), and suppose that in configuration  $\eta \in \{-1, +1\}^{Z^d}$ , the spin at site  $x \in Z^d$  flips at rate

$$(2.1) \quad c(x, \eta) = \exp\left(-\beta\eta(x) \sum_{|y-x|=1} \eta(y)\right).$$

This is the simplest example of what is known as a "stochastic Ising model." Note that the flip rate in (2.1) is relatively large if the spin at  $x$  is different from the spins at most of its neighbors, and smaller if it agrees with most of its neighbors. The effect of this is to make spins want to line up with neighboring spins.

Now suppose that we place the iron in a strong positive magnetic field, so that most of the spins are  $+1$ , and then turn off the field, letting the evolution given in (2.1) take over. To idealize, take  $\eta_0 \equiv +1$ , and consider the distribution of spins at large times. The first question that arises is whether

$$\lim_{t \rightarrow \infty} P(\eta_t(0) = +1)$$

is greater than  $\frac{1}{2}$  or equal to  $\frac{1}{2}$ . We would say that the iron is magnetized in the first case, but is not magnetized in the second.

It is an observed fact in nature and a proven property of this model (for  $d \geq 2$ ), that magnetization occurs at low temperatures (large  $\beta$ ), but not at high temperatures (small  $\beta$ ). While we are stating this in the context of the stochastic Ising model, it is (a simple consequence of) one of the fundamental results for the Ising model, whose extensive study in the mathematics and physics literature preceded the "stochastic" version by several decades. "Ising" refers to the equilibrium model, while "stochastic Ising" refers to the time evolution.

The value of  $\beta$  that separates the regimes of presence and absence of magnetization is denoted by  $\beta_c$  and called the critical value for the model. In the language of Markov processes, this situation can be described by saying that  $\eta_t$  has a unique stationary distribution (which has the same density of  $+$  and  $-$  spins) if  $\beta < \beta_c$  and has at least two stationary distributions (in one of which  $+$  spins predominate, and in the other  $-$  spins do) if  $\beta > \beta_c$ . We will denote by  $\mu_+$  the stationary distribution that is the limiting distribution starting from the configuration  $\eta_0 \equiv +1$  and by  $\mu_-$  the one that is the limiting distribution starting from the configuration  $\eta_0 \equiv -1$ . These station-

ary distributions are all examples of what are known as Gibbs random fields. Gibbs random fields are probability distributions  $\nu$  which are given formally in terms a potential  $\{U_A(\cdot), A \subset S\}$  by

$$\nu(\{\eta\}) = \text{constant} \exp\left[-\sum_{A \subset S} U_A(\eta)\right].$$

Of course, real systems are finite, and finite (irreducible) Markov chains have unique stationary distributions, so one can reasonably ask why this issue of multiplicity of stationary distributions is relevant to real systems. The answer is that over the time periods during which we observe a large system, it may or not be possible for it to “realize” that it is finite, and it therefore may behave like an infinite system for those time periods. The real issue is the relative sizes of the relevant temporal and spatial scales, and it does turn out that infinite models usually capture the essential features of large finite systems better than do finite models. We will see examples of the relevance of temporal scales in Theorem 2.3 and then again in Section 3.2.

The result described above for the stochastic Ising model is but the first step in an investigation that is now twenty-five years old and that has led to extensive rigorous results about this and related models. One of the main contributors to the theory during this period has been Holley; see Holley (1987, 1991) and Holley and Stroock (1987, 1989), for example. More recently, Schonmann (1994a, b) has obtained very detailed information about the stochastic Ising model in the parameter region in which there are multiple stationary distributions. To state one of his results, modify the rates in (2.1) to include the effect on the evolution of an external magnetic field of strength  $h$ :

$$c(x, \eta) = \exp\left(-\beta\eta(x) \left[\sum_{|y-x|=1} \eta(y) + h\right]\right).$$

The effect of the extra term is to favor spins with the same sign as  $h$ . A fact that is (initially, at least) rather surprising is that if  $h \neq 0$ , the corresponding process  $\eta_t^h$  has a unique stationary distribution  $\mu_h$  (for any  $\beta$ ). As has long been known, this distribution satisfies

$$(2.2) \quad \lim_{h \uparrow 0} \mu_h = \mu_-, \quad \lim_{h \downarrow 0} \mu_h = \mu_+.$$

The fact that there is uniqueness when  $h \neq 0$  becomes less surprising if one thinks of a particle moving under the influence of a double well potential. In the symmetric case (corresponding to  $h = 0$  in our context), the two wells have equal depth, and this leads to two equally “desirable” configurations for the particle. The least bit of asymmetry makes one well deeper than the other, and hence makes one of these configurations “better” than the other.

Consider now this process with initial configuration  $\eta_0^h \equiv -1$ . If  $h$  is small and positive, then for some period of time, the process presumably does not notice that  $h \neq 0$ , and will therefore try to be close to  $\mu_-$ . At larger times, the effect of  $h$  comes into play, and the process tries to be near  $\mu_+$  because of

(2.2). The mechanism by which this occurs is roughly the following: initially, small  $+$  clusters form because of the random flipping. It takes a certain amount of time for a  $+$  cluster that is “sufficiently” large (how large depends on the value of  $h$ ) to form. After that time, the positive  $h$  leads this cluster to grow, so that the distribution of the process will be essentially in the  $+$  phase. Schonmann’s theorem answers the natural question of the relative magnitudes of  $h$  and  $t$  that correspond to the two possible types of behavior described above.

**THEOREM 2.3.** *Suppose  $d \geq 2$  and  $\beta$  is sufficiently large. There are two constants  $0 < C_1 \leq C_2 < \infty$  so that if  $h \downarrow 0$  and  $t \rightarrow \infty$  jointly, then*

$$\eta_t^h \Rightarrow \begin{cases} \mu_-, & \text{if } \limsup h^{d-1} \log t < C_1, \\ \mu_+, & \text{if } \liminf h^{d-1} \log t > C_2, \end{cases}$$

where  $\Rightarrow$  denotes convergence in distribution.

The phenomenon of appearing to be in equilibrium for a long time, and only much later converging to the true stationary distribution, is known as metastability. We will encounter another instance of it in Section 3.2.

This brief introduction to the stochastic Ising model suggests a framework for the study of more general interacting particle systems. The first natural problem is to determine the structure of the set  $\mathcal{I}$  of stationary distributions for  $\eta_t$ . Then one wants to find the domain of attraction of each  $\nu \in \mathcal{I}$ , that is, the set of probability distributions  $\mu$  on  $X$  so that

$$(2.4) \quad \lim_{t \rightarrow \infty} \mu_t = \nu,$$

where  $\mu_t$  is the distribution of the process at time  $t$  when the initial distribution is  $\mu$ . Finally, it is important to understand the nature of the convergence in (2.4). How fast is it? What is the mechanism that leads to it?

Most readers are probably familiar with the theory of finite and countable state Markov chains. Superficially, the questions raised above for interacting particle systems appear very similar to those that are central to Markov chain theory. On closer examination, however, one finds significant differences. First, the issue of existence of a stationary distribution does not usually arise in interacting particle systems, since the state space of the system  $K^S$  is compact—there is essentially always a stationary distribution. Much of countable state Markov chain theory deals with the issue of positive recurrence vs. null recurrence and transience, which is essentially the issue of existence of a stationary distribution. When there is a stationary distribution for a Markov chain, it is automatically unique once a mild irreducibility condition is shown to hold. Most interesting interacting particle systems have more than one stationary distribution, and it is this fact that leads to the richness and broad applicability of this subject. For example, the problem of determining domains of attraction of stationary distributions is clearly more complex when these distributions are not unique.

An area in which close relatives of the stochastic Ising model have played an important role in recent years is image analysis. In a common set-up [see Geman (1990) or Winkler (1995), for example], both the original and the observed (corrupted) images are modeled by Gibbs random fields on  $\{0, 1\}^S$  for a large  $S$ , and a Bayesian approach requires the computation of means of posterior distributions. The size of the space  $\{0, 1\}^S$  is so large that direct computation is not feasible. The Monte Carlo techniques that are used instead involve letting a process similar to the stochastic Ising model (e.g., the Gibbs and Metropolis samplers) evolve until it is close to its stationary distribution, which is the desired Gibbs random field.

If instead of the mean of the posterior distribution, one uses the mode (which is known as the MAP estimate), then one needs to perform a minimization over a large space. A commonly used technique to accomplish this is known as simulated annealing, which again is closely related to stochastic Ising models. [See the previous references, or Holley and Stroock (1988).] Suppose that one wishes to find the location  $\zeta$  of the minimum value of some complicated function  $U$  on a large space  $X$ . The idea behind simulated annealing algorithms is to consider a family of Markov processes  $\eta_t^\beta$  on  $X$  indexed by a parameter  $\beta$  (which is again interpreted as the reciprocal of the temperature) with the property that the stationary distribution  $\mu_\beta$  of  $\eta_t^\beta$  becomes increasingly concentrated at  $\zeta$  as  $\beta \rightarrow \infty$  (i.e., as the temperature tends to zero). Then one takes a function  $\beta(t)$  that satisfies

$$(2.5) \quad \lim_{t \rightarrow \infty} \beta(t) = \infty$$

and defines a temporally inhomogeneous Markov process  $\eta_t$  using the transition mechanism for the process  $\eta_s^{\beta(t)}$  at time  $t$ . If the speed at which the convergence in (2.5) occurs is chosen correctly, then the distribution of  $\eta_t$  should tend to the pointmass at  $\zeta$  at a reasonable rate. The speed should be fast enough so that  $\mu_{\beta(t)}$  converges rapidly to the pointmass at  $\zeta$ . On the other hand, if it is too fast, then there is no reason to expect the distribution of  $\eta_t$  to be close to its stationary distribution  $\mu_{\beta(t)}$ . The analysis of this competition between conflicting goals is similar in spirit to that carried out in Theorem 2.3.

While physics provided much of the initial motivation for the development of the field of interacting particle systems, in more recent years, biological considerations have led to a number of new models and issues. To illustrate, we describe a model studied recently by Durrett and Neuhauser (1997). Their motivating context is barley yellow dwarf, a disease of certain grains which is caused by viruses transmitted by aphids. A study by Rochow (1979) found that the dominant strain of the virus in an area near Cornell University shifted from one strain to another rather quickly during the 1960's. The model seeks to contribute to the understanding of mechanisms that might explain this shift. Here is its description: take  $S = \mathbb{Z}^2$  (the set of plants) and  $K = \{0, 1, 2, 3\}$  (the set of possible states of a plant), where  $\eta(x) = 0$  means the site  $x$  is not infected,  $\eta(x) = 1$  means that  $x$  is infected by strain 1,

$\eta(x) = 2$  means that  $x$  is infected by strain 2 and  $\eta(x) = 3$  means that  $x$  is infected by both strains. To describe the transition rates, let  $f_i$  denote the fraction of the neighbors of  $x$  that are in state  $i$ ,  $0 \leq i \leq 3$ . Then the rates at  $x$  are given by

$$\begin{array}{ll} 0 \rightarrow 1 & \beta_1(f_1 + c_1 f_3) & 1 \rightarrow 0 & 1 \\ 0 \rightarrow 2 & \beta_2(f_2 + c_2 f_3) & 2 \rightarrow 0 & 1 \\ 2 \rightarrow 3 & c_1 \beta_1(f_1 + c_1 f_3) & 3 \rightarrow 2 & c_1^{-1} \\ 1 \rightarrow 3 & c_2 \beta_2(f_2 + c_2 f_3) & 3 \rightarrow 1 & c_2^{-1}. \end{array}$$

The  $\beta_i$  parameters are the underlying infection rates for the two strains, ignoring the effects of the presence of the other strain. The effect of the other strain is built into the parameters  $c_i \in [0, 1]$ . In the  $0 \rightarrow 1$  rate, for example, the presence of  $c_1$  reflects the fact that the effect of strain 1 in a doubly infected plant is less than it is in a singly infected plant.

Consider the rates  $\beta_1 > \beta_2$  to be fixed, and ask for what choices of the parameters  $0 \leq c_1, c_2 \leq 1$  there exist stationary distributions in which both strains coexist, and for what choices one or the other strain takes over as  $t \rightarrow \infty$ . The answer is given in Figure 1, which is essentially taken from Durrett and Neuhauser (1997). The crosshatched regions represent theorems

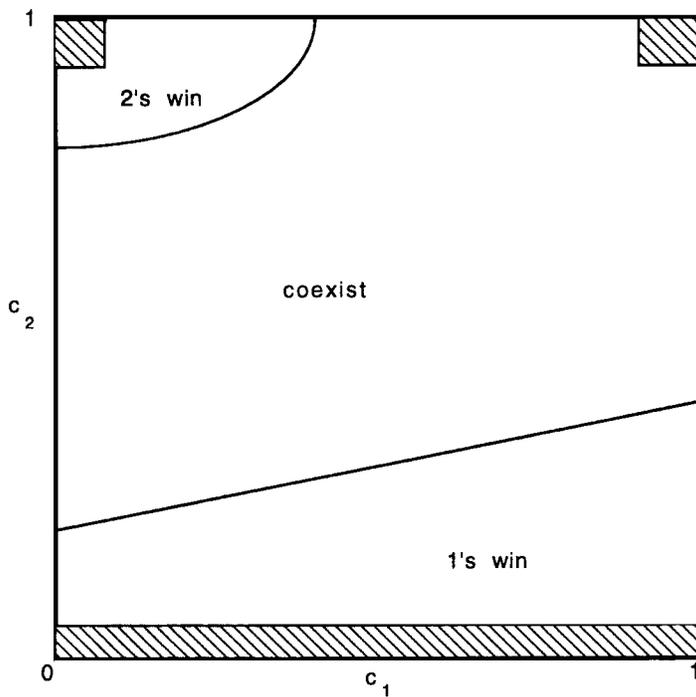


FIG. 1.

(in some cases the proofs require the neighborhood of a site to be sufficiently large), while the rest of the figure represents the expected qualitative behavior of the system. So there are definitely regions of coexistence, as well as regions in which each strain takes over. There are no significant rigorous results about the size of the coexistence region. However, Durrett and Neuhauser argue that if it is relatively narrow, this mechanism could explain the change in dominant strain observed by Rochow: small changes in the values of  $c_1$ ,  $c_2$  would lead to a dramatic shift from one strain to the other.

Note that when only one strain is present, say strain 1, so that  $f_2 = f_3 = 0$ , the transition rates become

$$0 \rightarrow 1 \quad \beta_1 f_1 \qquad 1 \rightarrow 0 \quad 1.$$

This reduces to a model known as the contact process, which we focus on next.

**3. The contact process.** The contact process has played a central role in the theory of interacting particle systems since it was introduced by Harris (1974) over twenty years ago. It is simple to describe, has a rich structure and lends itself to results that are easy to state, but often quite difficult to prove. The contact process has served as a fruitful testing ground for results and techniques that were later extended and applied to other classes of interacting particle systems. It has important properties, which are known as additivity, attractiveness and self-duality. On the other hand, it fails to be reversible; if it were, it would be easier to analyze and hence probably of less mathematical interest. One indication of the importance of the contact process in the field of interacting particle systems is that a number of survey papers and book chapters have been devoted to it: Griffeath (1979, 1981), Liggett (1985), Durrett (1988, 1991) and Konno (1994), for example. Interest in the contact process continues to grow. There have been more papers published or written about it since 1990 (at least 40) than in the entire period 1974–1989.

To describe the  $d$ -dimensional contact process, take  $S = Z^d$  and  $K = \{0, 1\}$ . Zeros represent healthy individuals, while ones represent infected individuals. Infected sites wait an exponential time with parameter 1 and then become healthy, while healthy sites become infected at a rate proportional to the number of infected neighbors. It is common to identify  $\eta \in \{0, 1\}^{Z^d}$  with the set  $A \subset Z^d$  of infected sites:  $A = \{x \in Z^d: \eta(x) = 1\}$ . This leads to the Markov process  $A_t$  on the collection of subsets of  $Z^d$  which has the following transitions:

$$\begin{aligned} A &\rightarrow A \setminus \{x\} \quad \text{for } x \in A \text{ at rate } 1, \text{ and} \\ A &\rightarrow A \cup \{x\} \quad \text{for } x \notin A \text{ at rate } \lambda \#\{y \in A: |y - x| = 1\}. \end{aligned}$$

Note that  $\emptyset$  is a trap for the process. The process (or infection) is said to survive if

$$P^{(0)}(A_t \neq \emptyset \forall t) > 0.$$

(The superscript refers to the initial configuration.) A special case of the self-duality statement [Liggett (1985), Theorem 1.7, Chapter VI] asserts that

$$P^{Z^d}(\mathbf{0} \in A_t) = P^{(0)}(A_t \neq \emptyset).$$

Since the distributions  $P^{Z^d}(A_t \in \cdot)$  are stochastically decreasing in  $t$ , it follows that there is a nontrivial (i.e.,  $\neq$  the pointmass on  $\emptyset$ ) stationary distribution for the process if and only if the process survives. This limiting distribution is called the upper invariant measure, and denoted by  $\nu$ . An easy coupling argument shows that there is a critical value  $\lambda_c \in [0, \infty]$ , so that survival occurs for  $\lambda > \lambda_c$  and extinction occurs for  $\lambda < \lambda_c$ .

Many of the problems of interest revolve around this issue of survival and extinction: What can be said about the critical value? Does the process survive at the critical value? How can one identify the subcritical and supercritical phases from the behavior of the process on a finite set? When survival occurs, what is the limiting distribution of the process for arbitrary initial configurations? How rapidly does convergence occur? How do the answers to some of these questions change if the infection parameter  $\lambda$  is allowed to change from site to site? What if we replace  $Z^d$  by other graphs, such as trees? The remainder of this section is devoted to tracing the development of answers to some of these questions. Among the topics that we will mention only briefly or not at all are metastability, edge fluctuations, rates of convergence, pointwise ergodic theorems and multitype systems.

**3.1. The critical value.** The proof that  $\lambda_c > 0$  is easy: if  $2d\lambda < 1$ , the cardinality of  $A_t$  is dominated by a subcritical branching process (points in  $A_t$  disappear at rate 1 each, while new points are added at rate at most  $2d\lambda|A_t|$ ), and hence dies out. The resulting lower bound for  $\lambda_c$  (which is  $1/2d$ ) can easily be improved to

$$\lambda_c \geq \frac{1}{2d-1}.$$

In one dimension, the best known lower bound is 1.539; see Grillenberger and Ziezold (1988).

Proving survival (and hence upper bounds for  $\lambda_c$ ) is substantially harder than proving extinction. Harris (1974) was the first to prove that the critical value is finite. While he did not give an explicit upper bound, Durrett (1988) used a version of Harris's technique to show  $\lambda_c < 1328$  in one dimension. Other upper bounds that have been given in one dimension are:  $\lambda_c \leq 2$  [Holley and Liggett (1978)],  $\lambda_c \leq 7$  [Gray and Griffeath (1982)],  $\lambda_c \leq 3.95$  [Durrett (1992a)],  $\lambda_c \leq 2.27$  [Stacey (1994)] and  $\lambda_c \leq 1.942$  [Liggett (1995a)]. Note that these bounds are not chronologically monotone. The primary reason for this is that there has been an interest in finding the best bound that a given technique will produce, since the various techniques that have been used apply with different degrees of success in different (i.e., other than the basic contact process) contexts. In particular, the Holley–Liggett technique, which is the best in the sense of giving a good bound when it applies, appears to have the most limited range of applicability.

This technique is based on the following idea: choose a renewal measure  $\nu$  on  $\{0, 1\}^Z$  (i.e., one for which the spacings between successive 1's are i.i.d.) as initial distribution, and then show that if  $\nu$  is chosen appropriately, the distribution at time  $t$  is increasing in  $t$  in a certain sense. If this is the case, the limiting distribution cannot be the pointmass on  $A = \emptyset$ . If  $\lambda = 2$ , the choice that works has spacing distribution with tails

$$P(Z > n) = \frac{(2n)!}{n!(n+1)!4^n}.$$

If  $\lambda < 2$ , there is no renewal measure that has the desired properties.

In higher dimensions, there are the following upper bounds:  $\lambda_c \leq 2/d$  [Holley and Liggett (1978)],  $\lambda_c \leq (1 - \rho)/\rho$  [Holley and Liggett (1981)] and  $\lambda_c \leq (1/2)d(2\rho - 1)$  [Griffeath (1983)]. In the last two bounds,  $\rho$  is the probability that a simple random walk on  $Z^d$  never returns to its starting point. The main advantage of the latter two bounds (which apply only for  $d \geq 3$ ) is that they imply that  $\lambda_c \sim (2d)^{-1}$  for large  $d$ . The best upper bound in one dimension is 1.942 [Liggett (1995a)] and in two dimensions is 0.79 [Stacey (1994)]. One might reasonably ask why it is important to have good bounds on critical values. An example that illustrates this importance will be discussed in Section 4.

Techniques developed to bound the critical value of the contact process have been used in other areas as well. For example, Liggett (1995b) used the technique used to prove  $\lambda_c \leq 2$  to prove that the critical value of two-dimensional oriented percolation is at most  $\frac{2}{3}$  in the bond case and at most  $\frac{3}{4}$  in the site case.

**3.2. The process on a finite set.** The real systems that are modeled by the contact and similar processes are finite (though large) in space and time. Experience shows, however, that the important properties of these finite real systems are better modeled by infinite than by finite mathematical objects. Finite systems do not survive for any value of  $\lambda$ , while infinite ones do for large values of  $\lambda$ . As we see next, however, the finite systems corresponding to supercritical infinite systems survive for much longer than do those which correspond to subcritical infinite systems. It is in this sense that the infinite system (in both space and time) is a good model for a large finite system. The spatial size of the real system is so large that it is effectively infinite for time periods that arise naturally. We would like to make these ideas precise.

The contact process on the finite set  $\{1, \dots, N\}$  is a finite state Markov chain which is eventually absorbed into  $\emptyset$ . (Infections at 0 and at  $N + 1$  are suppressed.) A natural question is how the subcritical, critical and supercritical phases of the infinite system can be observed in the finite system. This problem was first treated by Griffeath (1981); more complete results were then obtained by Durrett and Liu (1988), Durrett and Schonmann (1988) and Durrett, Schonmann and Tanaka (1989). To state them, let  $\sigma_N$  be the

extinction time for the process on  $\{1, \dots, N\}$ , when the initial configuration is  $A_0 = \{1, \dots, N\}$ . Then each of the following statements holds in probability as  $N \rightarrow \infty$ :

$$(3.1) \quad \begin{aligned} \frac{\sigma_N}{\log N} &\rightarrow C_1 \quad \text{if } \lambda < \lambda_c, \\ \frac{\sigma_N}{N} &\rightarrow \infty \quad \text{and} \quad \frac{\sigma_N}{N^4} \rightarrow 0 \quad \text{if } \lambda = \lambda_c, \\ \frac{\log \sigma_N}{N} &\rightarrow C_2 \quad \text{if } \lambda > \lambda_c. \end{aligned}$$

Here  $C_1, C_2$  are positive constants, and  $\lambda_c$  is the critical value for the contact process on  $Z^1$ . Thus extinction occurs at times that are logarithmic in system size in the subcritical regime, while they are algebraic in the critical case, and exponential in the supercritical regime. Note that in the simple case  $\lambda = 0$ ,  $\sigma_N$  is the maximum of  $N$  independent unit exponentials, so that

$$\frac{\sigma_N}{\log N} \rightarrow 1.$$

Thus the first part of (3.1) says that subcritical systems behave much the same as pure death systems.

A closely related type of result involves metastability—a phenomenon that we first encountered in the discussion of Theorem 2.3. It states that in the supercritical case,

$$\frac{\sigma_N}{E\sigma_N} \Rightarrow \text{unit exponential distribution,}$$

and that for “most” times before  $\sigma_N$ , the process is close to  $\nu$ . Results of this type were proved in one dimension for large  $\lambda$  by Cassandro, Galves, Olivieri and Vares (1984) and for all  $\lambda > \lambda_c$  by Schonmann (1985). Analogous results in higher dimensions have been proved more recently by Mountford (1993) and Simonis (1996).

**3.3. Convergence.** The complete convergence theorem for the contact process asserts that

$$(3.2) \quad P^A(A_t \in \cdot) \Rightarrow P^A(\sigma < \infty)\delta_{\emptyset}(\cdot) + P^A(\sigma = \infty)\nu(\cdot)$$

for all initial configurations  $A$ . Here  $\sigma$  is the hitting time of  $\emptyset$  and  $\nu$  is the upper invariant measure for the process. One consequence of such a result is that the only extremal stationary distributions for the process are  $\delta_{\emptyset}$  and  $\nu$ . As we will see next, the statement in (3.2) is deceptively simple—its proof presents a significant challenge.

The complete convergence theorem was proved in a number of stages. Harris (1976) considered translation invariant initial distributions instead of deterministic initial configurations. Griffeath (1978) extended Harris’s result to somewhat more general initial distributions, and proved (3.2) for general

$A$  in one dimension for sufficiently large  $\lambda$ . Durrett (1980) showed that in the one-dimensional supercritical case, the rightmost point in  $A_t$  moves to the right with positive speed, and used this result to prove (3.2) in this case. Durrett and Griffeath (1982) proved (3.2) for all  $A$  in higher dimensions, but only for large values of  $\lambda$ . Schonmann (1987) gave a simplified proof of this result. Andjel (1988) enlarged the set of  $\lambda$ 's for which (3.2) could be proved. The final version of (3.2) on  $Z^d$ , valid in all dimensions and for all  $\lambda$  and initial configurations is a consequence of techniques developed by Bezuidenhout and Grimmett (1990) to prove that the critical contact process dies out in all dimensions. This latter result had been an open problem for a number of years, even in one dimension.

**3.4. Inhomogeneous and random environments.** In many contexts, it is important to consider systems with spatially varying dynamics. As an example of this, we consider the inhomogeneous contact process  $A_t$  with the following transition rates:

$$\begin{aligned} A &\rightarrow A \setminus \{x\} && \text{for } x \in A \text{ at rate } \delta(x), \text{ and} \\ A &\rightarrow A \cup \{x\} && \text{for } x \notin A \text{ at rate } \sum_{y \in A: |y-x|=1} \lambda(x, y), \end{aligned}$$

so that the infection and recovery rates are allowed to depend on the sites in question. This dependence can alter the behavior of the process in several respects. For example, Bramson, Durrett and Schonmann (1991) showed that with  $\lambda(x, y) \equiv 1$  and  $\delta(x)$  i.i.d. in one dimension, it is possible to have survival without having linear growth of  $A_t$ . We saw in Section 3.3 that this cannot occur in the homogeneous case. More recently, Madras, Schinazi and Schonmann (1994) proved that there are (deterministic) choices of  $\delta(x)$  so that the process with  $\lambda(x) \equiv \lambda$  survives at the critical point.

To discuss briefly the class of problems that are most directly analogous to the determination of critical values in the homogeneous case, take  $\delta \equiv 1$  and  $\lambda(x, y)$  i.i.d. An ideal result would be of the following form: there are reasonably simple increasing functions  $f$  and  $g$  that are “almost” the same so that

$$\begin{aligned} A_t &\text{ dies out if } Ef(\lambda) \text{ is sufficiently small, and} \\ A_t &\text{ survives if } Eg(\lambda) \text{ is sufficiently large,} \end{aligned}$$

with quantified versions of the word “sufficiently.” Results obtained so far fall significantly short of this ideal. In one dimension, Liggett (1991) observed that  $A_t$  dies out if  $E \log \lambda < 0$ , and Liggett (1992) proved that  $A_t$  survives if  $E(2\lambda^{-1} + \lambda^{-2}) < 1$  using a variant of the Holley–Liggett technique. In higher dimensions  $d$ , Klein (1994) proved that  $A_t$  dies out if  $E(\log(1 + \lambda))^{\beta(d)}$  is sufficiently small, where  $\beta(d)$  is of order  $2d^2$  for large  $d$ . Andjel (1992), on the other hand, showed that for any  $\beta < d$ ,  $A_t$  can survive even if  $E(\log(1 + \lambda))^\beta$  is arbitrarily small. It would be interesting to know what the correct cutoff for  $\beta$  is in this problem.

In a recent paper, Newman and Volchan (1996) take  $\lambda(x, x+1)$  and  $\lambda(x, x-1)$  deterministic and independent of  $x$  (and not both zero) in one dimension, and  $\delta(x)$  i.i.d. Under a condition that is slightly stronger than

$$(3.3) \quad E([- \log \delta(x)]^+) = \infty,$$

they prove that the process survives. Since the process dies out if

$$E \log \delta(x) > \log \max(\lambda(x, x+1), \lambda(x, x-1)),$$

it would be of interest to know whether (3.3) itself is sufficient for survival.

**3.5. Processes on trees.** Let  $T_d$  be the homogeneous (connected) tree in which each vertex has  $d+1$  neighbors. The contact process is defined as before, with  $|x-y|=1$  meaning that  $x$  and  $y$  are neighbors on the tree. We will say that  $A_t$  survives strongly if

$$P^{(x)}(x \in A_t \text{ for arbitrarily large } t) > 0,$$

and that it survives weakly if it survives but does not survive strongly. Weak survival would mean that even though the infection may not disappear, it moves off to “infinity.” As a consequence of (3.2), weak survival cannot occur for (symmetric) contact processes on  $Z^d$ .

We mention parenthetically that the word “symmetric” in the last sentence is important. For the totally asymmetric contact process on  $Z^1$ , Griffeath (1979) observed that weak survival can occur. Asymmetric contact processes were further studied by Schonmann (1986). Schinazi (1994) proved an analogue of (3.1) in that context, leaving one case open which was recently settled by Sweet (1997).

Following related work on percolation by Grimmett and Newman (1990), Pemantle (1992) discovered that weak survival does occur for (symmetric) contact processes on  $T_d$  for an appropriate range of  $\lambda$ 's, at least if  $d \geq 3$ . This difference in behavior between  $Z^d$  and  $T_d$  is an important reason for the interest in the contact process on  $T_d$ .

Since weak survival can occur on  $T_d$ , it is natural to define critical values  $\lambda_1 \leq \lambda_2$  by the statement that  $A_t$  survives strongly for  $\lambda > \lambda_2$ , survives weakly for  $\lambda_1 < \lambda < \lambda_2$  and dies out for  $\lambda < \lambda_1$ . Thus Pemantle's result is that  $\lambda_1 < \lambda_2$  if  $d \geq 3$ . He first proved

$$\lambda_1 \leq \frac{1}{d-1} \quad \text{and} \quad \lambda_2 \geq \frac{1}{2\sqrt{d}},$$

which implies that  $\lambda_1 < \lambda_2$  if  $d \geq 6$ . He then improved these bounds to extend the conclusion to  $d \geq 3$ . Liggett (1996) further improved the bounds to cover the case  $d=2$ . Stacey (1996) then gave a different proof of  $\lambda_1 < \lambda_2$  which works for  $d \geq 2$  and even for some inhomogeneous trees. Here are

bounds on the two critical values, which show that  $\lambda_1 < \lambda_2$  if  $d \geq 2$  (recall that  $\lambda_1 = \lambda_2 = \lambda_c$  if  $d = 1$ ):

$$\begin{aligned} d = 1 & \quad \lambda_1 \leq 1.942 & \lambda_2 \geq 1.539 \\ d = 2 & \quad \lambda_1 \leq 0.605 & \lambda_2 \geq 0.609 \\ d = 3 & \quad \lambda_1 \leq 0.391 & \lambda_2 \geq 0.425 \\ d = 4 & \quad \lambda_1 \leq 0.279 & \lambda_2 \geq 0.354 \\ d = 5 & \quad \lambda_1 \leq 0.218 & \lambda_2 \geq 0.309 \end{aligned}$$

Case  $d = 1$  was mentioned earlier (note that  $T_1 = Z^1$ ), while the other cases are due to Liggett (1996) ( $d = 2$ ) and Pemantle (1992) ( $d \geq 3$ ). The bounds mentioned above also give some information about the asymptotics of the critical values for large  $d$ :

$$\lim_{d \rightarrow \infty} d\lambda_1^{(d)} = 1$$

and

$$2 - \sqrt{2} \leq \liminf_{d \rightarrow \infty} \sqrt{d} \lambda_2^{(d)} \leq \limsup_{d \rightarrow \infty} \sqrt{d} \lambda_2^{(d)} \leq e.$$

It would be interesting to evaluate the limit in the second statement.

Here are some other results that have been proved for the contact process on  $T_d$ :

(a) The critical contact process ( $\lambda = \lambda_1$ ) dies out. This was proved by Pemantle (1992) for  $d \geq 3$  and by Morrow, Schinazi and Zhang (1994) for  $d \geq 2$ . In the latter paper, this result is derived from the statement that  $\sup_{t > 0} E^{(x)} |A_t| < \infty$  in the critical case. (Whether or not this statement holds if  $d = 1$  is not known, but the general feeling seems to be that it does not.)

(b) Wu (1995) proved that the survival probability satisfies

$$\liminf_{\lambda \downarrow \lambda_1} \frac{P^{(x)}(A_t \neq \emptyset \ \forall t)}{\lambda - \lambda_1} > 0, \quad \limsup_{\lambda \downarrow \lambda_1} \frac{P^{(x)}(A_t \neq \emptyset \ \forall t)}{\lambda - \lambda_1} < \infty,$$

provided that  $d \geq 5$ . This is the statement that the “critical exponent” for survival is 1.

(c) Durrett and Schinazi (1995) proved that there are infinitely many extremal stationary distributions for the process if  $\lambda_1 < \lambda < \lambda_2$ . Each stationary distribution  $\mu$  that they construct satisfies

$$\lim_{x \rightarrow \theta} \mu\{\eta: \eta(x) = 1\} > 0$$

for  $\theta$  in a significant fraction of the boundary of  $T_d$ . (The boundary of  $T^d$  can be identified with the set of semi-infinite non-self-intersecting paths emanating from a fixed vertex, that is, the set of ways of “getting to  $\infty$ .”) For example, take  $A$  to be the set of all sites that are closer to  $x$  than to  $y$ , where  $x, y$  are two neighboring sites. Consider the process with initial configuration  $A_0 = A$ . Then the limiting distribution as  $t \rightarrow \infty$  is a stationary distribution that is asymptotic to the upper invariant measure as one approaches the boundary of  $T^d$  through  $A$  and to the pointmass on  $\emptyset$  as one approaches it through  $A^c$ .

Liggett (1997) constructed another class of stationary distributions (for some values of  $\lambda$ ) that are spherically symmetric and satisfy

$$\lim_{x \rightarrow \theta} \mu\{\eta: \eta(x) = 1\} = 0$$

for all  $\theta \in \partial T_d$ . (A spherically symmetric measure is one that is invariant under automorphisms of  $T^d$  that fix a particular vertex.) It is not known what all the invariant measures are for the contact process on  $T_d$ ,  $d \geq 2$ .

(d) Zhang (1997) has proved that the complete convergence theorem (3.2) holds for  $\lambda > \lambda_2$  and that there is no strong survival at  $\lambda = \lambda_2$ .

(e) Let

$$u(n) = P^{(x)}(y \in A_t \text{ for some } t)$$

for  $|x - y| = n$ . It is not hard to see that

$$u(n + m) \geq u(n)u(m),$$

so that

$$\rho(\lambda) = \lim_{n \rightarrow \infty} u(n)^{1/n}$$

exists and  $u(n) \leq [\rho(\lambda)]^n$ . Liggett (1997) has proved that

$$\rho(\lambda) \leq \frac{1}{d}$$

if  $\lambda \leq \lambda_1$ , and conjectured that

$$\rho(\lambda) \leq \frac{1}{\sqrt{d}}$$

if  $\lambda \leq \lambda_2$ . The primary interest in this conjecture is the following: it is not hard to check that  $u(n)$  is bounded below if  $\lambda > \lambda_2$ , and hence that  $\rho(\lambda) = 1$  in that case. Thus the conjecture would imply that  $\rho(\lambda)$  cannot take values in  $(1/\sqrt{d}, 1)$ . If true, this would give another proof of Zhang's result in (d) above that there is no strong survival at  $\lambda = \lambda_2$ .

Contact processes on inhomogeneous trees and other more general graphs have a rich structure, which has recently been investigated by Salzano and Schonmann (1998).

One of the roles of the contact process is to serve as a comparison system in the analysis of other models. The next section contains an example of this type of application to threshold voter models.

**4. Voter models.** By a voter model, we mean a process on  $X = \{0, 1\}^{Z^d}$  in which 0's and 1's flip (individually) at rates that depend on the states of the neighboring sites and that have the following two properties.

1.  $\eta \equiv 0$  and  $\eta \equiv 1$  are fixed points for the evolution.
2. The evolution is unchanged by interchanging the roles of 0's and 1's.

One can imagine that there is a "voter" at each point in  $Z^d$ , and that his position on some issue (0 or 1) changes at random times under the influence

of the opinions of his neighbors. By property (1) above, a voter model has two trivial extremal stationary distributions, the pointmasses on  $\eta \equiv 0$  and  $\eta \equiv 1$ , which represent consensus. The main question we will discuss is whether or not there are others, which would then represent coexistence of different opinions in equilibrium. If this occurs, we will say that the process coexists. On the other hand, if

$$\lim_{t \rightarrow \infty} P[\eta_t(x) \neq \eta_t(y)] = 0$$

for all  $x, y \in Z^d$  and all initial configurations, we will say that the process clusters, since the configuration at large times is made up of large clusters of zeros and large clusters of ones. This dichotomy between coexistence vs. clustering is similar to the one discussed in Section 2 in the context of the barley yellow dwarf model.

Until five years ago, the only voter models that had been considered were the linear ones, in which the rate for  $0 \rightarrow 1$  at  $x \in Z^d$  is given by a linear function of  $\eta$ :

$$(4.1) \quad c(x, \eta) = \sum_y p(x, y) \eta(y).$$

They were introduced independently by Clifford and Sudbury (1973) and Holley and Liggett (1975). The theory was developed quite completely in that case; most of it can be found in Chapter V of Liggett (1985) and Chapters 2 and 10 of Durrett (1988). In particular, if the  $p(x, y)$  in (4.1) is given by

$$p(x, y) = \begin{cases} 1, & \text{if } |x - y| \leq N, \\ 0, & \text{if } |x - y| > N, \end{cases}$$

then the process clusters if  $d \leq 2$  and coexists if  $d \geq 3$ . It is important to note that this aspect of its behavior depends on the dimension  $d$ , but not on the size of the neighborhood, which is determined by  $N$ . We will contrast this observation with the answer to the corresponding question for the nonlinear voter models that were introduced by Cox and Durrett (1991).

Before turning to the nonlinear case, we will describe briefly the argument that leads to the above answer in the linear case. A special case of the basic duality relation for the linear voter model says that

$$P^n[\eta_t(x) \neq \eta_t(y)] = P[\eta(X_t) \neq \eta(Y_t)]$$

for all initial configurations  $\eta \in X$ , where  $X_t$  and  $Y_t$  are (continuous time) random walks on  $Z^d$  with  $X_0 = x$ ,  $Y_0 = y$ , which make uniformly chosen jumps of size at most  $N$  at exponential rates, which are independent until the first time they hit each other, and are equal after that time. By recurrence of the random walk  $X_t - Y_t$ , if  $d \leq 2$ ,  $X_t$  and  $Y_t$  will hit eventually with probability 1, and hence

$$P^n[\eta_t(x) \neq \eta_t(y)] \leq P[X_t \neq Y_t] \rightarrow 0$$

as  $t \rightarrow \infty$ . On the other hand, if  $d \geq 3$ , there is positive probability that the random walks never hit, and hence if the initial distribution for  $\eta_t$  is the product measure  $\nu_\rho$  with density  $\rho \in (0, 1)$ , then

$$\lim_{t \rightarrow \infty} P[\eta_t(x) \neq \eta_t(y)] = 2\rho(1 - \rho) \lim_{t \rightarrow \infty} P[X_t \neq Y_t] > 0$$

for  $x \neq y$ . In fact, all the extremal stationary distributions are obtained by taking limits of the distribution at time  $t$  of the process whose initial distribution is  $\nu_\rho$  for  $\rho, \rho \in [0, 1]$ .

The dimensional dependence of the behavior of the linear voter model is actually richer than indicated above. A result that illustrates this dependence is due to Cox and Griffeath (1983): consider again the linear voter model with initial distribution  $\nu_\rho$ . Then as  $t \rightarrow \infty$ , the ‘‘occupation time’’ at site 0 satisfies

$$\text{var}\left(\int_0^t \eta_s(0) ds\right) \sim C(d, N)\rho(1 - \rho) \begin{cases} t^2, & \text{if } d = 1, \\ t^2/\log t, & \text{if } d = 2, \\ t^{3/2}, & \text{if } d = 3, \\ t \log t, & \text{if } d = 4, \\ t, & \text{if } d \geq 5. \end{cases}$$

The corresponding central limit theorem holds, with a Gaussian limit if (and only if)  $d \geq 2$ . The size of the variance is of course a measure of the strength of the correlations between  $\eta_t(0)$  for different  $t$ 's.

Now consider a special class of nonlinear voter models, known as threshold voter models, in which a site  $x$  changes value at rate 1 if there is at least one site  $y$  in its neighborhood (i.e.,  $|x - y| \leq N$ ) that has the opposite opinion, and rate 0 otherwise. The duality that was so useful in the linear case does not occur in nonlinear models, but we will see that coupling arguments are useful substitutes for it.

A simple computation in the case  $d = N = 1$  shows that if the initial distribution of the process is translation invariant, then

$$\frac{d}{dt} P[\eta_t(k) \neq \eta_t(k+1)] = -2P[\eta_t(k-1) \neq \eta_t(k) \neq \eta_t(k+1)].$$

It follows that any stationary distribution  $\mu$  that is translation invariant must satisfy

$$\mu(101) = \mu(010) = 0$$

(the cylinder probabilities above refer to three consecutive sites), and from this it follows that such a measure must be a mixture of the pointmasses on  $\eta \equiv 0$  and  $\eta \equiv 1$ . This is the core of the proof that if  $d = N = 1$ , the threshold voter model clusters.

In contrast to this, we have the following result, which was proved for large  $N$  by Cox and Durrett (1991) and for all  $N$  by Liggett (1994).

**THEOREM 4.2.** *If  $(N, d) \neq (1, 1)$ , then the threshold voter model coexists.*

This statement is unusual in that an answer can be given for all values of the parameters  $(N, d)$ . Most results in this field can only be proved for certain regions of the parameter space. (Recall the Durrett–Neuhauser results discussed in Section 2, for example.) Another feature of this result that is somewhat surprising is that it shows that the threshold voter model is quite different from the linear voter model, in that coexistence occurs even in one dimension, provided that the neighborhood is not too small. It becomes less surprising when one notes that the threshold model has a drift toward the “local minority” which is not present in the linear case.

The proof of Theorem 4.2 illustrates a number of ideas that are used frequently in interacting particle systems, so we will say a bit about it at this point. A key step is to compare the threshold voter model  $\eta_t$  to a threshold contact process  $\zeta_t$  in such a way that survival of the contact process (with  $\lambda = 1$ ) implies coexistence of the voter model. The threshold contact process has the following transitions at site  $x$ :

$$\begin{aligned} 1 &\rightarrow 0 && \text{at rate } 1, \\ 0 &\rightarrow 1 && \text{at rate } \lambda \text{ if } \eta(y) = 1 \text{ for some } |x - y| \leq N, \end{aligned}$$

and rate 0 otherwise. Note that the  $0 \rightarrow 1$  transitions have the same rates for the two processes, while the  $1 \rightarrow 0$  rates are smaller for the voter model than for the contact process (with  $\lambda = 1$ ). This means that one can couple the two processes together so that

$$(4.3) \quad \zeta_t \leq \eta_t$$

for all  $t$ , provided that the initial distributions can be so coupled. If  $\zeta_t$  survives, then  $\eta_t$  (with initial distribution the product measure with density  $\frac{1}{2}$ , say) will in the limit have infinitely many 1's with probability 1. By  $0 \leftrightarrow 1$  symmetry, this limit will also have infinitely many 0's with probability 1, and hence the threshold voter model coexists.

Therefore we see that it is sufficient to prove survival of the threshold contact process for  $(N, d) \neq (1, 1)$ ,  $\lambda = 1$ . This survival statement is “monotone” in  $(N, d)$ —survival for one choice of  $(N, d)$  implies survival for any larger choice. This statement is easy to check for the threshold contact process, since the  $1 \rightarrow 0$  rates are constant and the  $0 \rightarrow 1$  rates are increasing in  $(N, d)$ . The corresponding argument cannot be made for the threshold voter model, and it is for that reason that it is important to make the comparison in (4.3).

By this monotonicity in  $(N, d)$ , it is enough to prove survival for  $(N, d) = (2, 1)$  and for  $(N, d) = (1, 2)$ . Another comparison argument shows that the former implies the latter. The key to this comparison is that the “projection”  $\pi: Z^2 \rightarrow Z^1$  defined by

$$\pi(m, n) = m + 2n$$

maps the  $N = 1$  neighborhoods in  $Z^2$  to the  $N = 2$  neighborhoods in  $Z^1$ . (Note that sites have four neighbors in both cases.) The final and hardest step

is then to prove survival of the threshold contact process with  $\lambda = 1$ ,  $d = 1$ ,  $N = 2$ . This is done by using a version of the Holley–Liggett (1978) argument, which works at most down to  $\lambda = 0.985$ , thus showing the importance of finding very tight bounds on critical values. Note that having an upper bound greater than 1 for the critical value in this context would be of no use at all.

The proof of survival in this case is significantly more difficult than in the context of the basic (nearest neighbor) contact process. Recall that in the earlier proof, we were able to write down explicitly the distribution of the relevant renewal measure. In the current situation, the corresponding distribution is determined by a set of equations that (apparently) cannot be solved explicitly. In fact, the analysis of these equations is computer aided. It would be interesting to have an analytic treatment of these equations. The proof of monotonicity in time of the distribution of the process is also made harder by the fact that the interaction is second nearest neighbor. The details are in Liggett (1994).

There are clearly other classes of nonlinear voter models that can be considered, but results about them are much less complete. The most studied process is the  $T$ -threshold voter model. Here one allows flips at  $x$  at rate 1 only if there are at least  $T$   $y$ 's satisfying  $|x - y| \leq N$ ,  $\eta(y) \neq \eta(x)$ . Here are some of the results that have been proved for this class ( $M$  denotes the cardinality of the neighborhood  $\{x: |x| \leq N\}$ ).

1. If  $d = 1$  and  $T = N(= (M - 1)/2)$ , then the process clusters [Andjel, Liggett and Mountford (1992)].
2. If  $T = \theta M$  with  $\theta < \frac{1}{4}$  and  $N$  is sufficiently large, then the process coexists [Durrett (1992b)].
3. If  $T > \frac{1}{2}(M - 1)$ , then the process fixates, in the sense that each site flips only finitely often [Durrett and Steif (1993)].

Durrett (1995) conjectures that if  $T = \theta M$  with  $\frac{1}{4} < \theta < \frac{1}{2}$  and  $N$  is sufficiently large, then the process clusters.

**5. The exclusion process.** All of the systems we have considered so far have the property that only one site changes value at a time. Exclusion processes describe particle motion (there are other interpretations—traffic flow, for example), and when a particle moves from  $x$  to  $y$ , both  $\eta(x)$  and  $\eta(y)$  change values. This, together with the related fact that the “number” of particles does not change with time, means that the analysis of these processes is quite different from that of contact and voter models.

We will consider here only the nearest neighbor exclusion processes in one dimension, which already have a very rich theory, with connections to partial differential equations and queueing theory. The theory of more general exclusion processes as of a decade ago is presented in Chapter VIII of Liggett (1985).

Here is a description of the process: the state space is  $\{0, 1\}^Z$ . Ones represent occupied sites, while zeros represent vacant sites. There is always at most one particle per site. A particle at  $x \in Z$  waits a unit exponential

time, and then tries to move to  $x + 1$  with probability  $p$  and to  $x - 1$  with probability  $q = 1 - p$ , where  $0 \leq p \leq 1$ . Moves to occupied sites are suppressed (which is the “exclusion” interaction). Spitzer (1970) introduced the exclusion process, and discovered a duality for it in case  $p = 1/2$  (and more generally, for symmetric systems) that is analogous to the contact process duality used in Section 3, and the voter model duality used in Section 4.

There is a parallel between the tools used in the analysis of exclusion processes and those used for voter models. For voter models, duality is heavily used in the linear case, while coupling comes to the fore for nonlinear models. In the exclusion world, the corresponding dichotomy is between symmetric systems and asymmetric ones—duality is used in the symmetric case, and coupling is used in the asymmetric case.

It is not hard to check that the product measure  $\nu_\rho$  with density  $\rho$  is invariant for the process for any  $0 \leq \rho \leq 1$ . Spitzer (1974) proved that these are all the extremal stationary distributions in the symmetric case  $p = 1/2$ :  $\mathcal{I}_e = \{\nu_\rho, 0 \leq \rho \leq 1\}$ . If  $p \neq 1/2$ , there are other extremal stationary distributions, known as “blocking” measures. However, all of them concentrate on the uninteresting part of the state space

$$\left\{ \eta: \sum_x \eta(x)[1 - \eta(x+1)] < \infty \right\}$$

in which  $\eta$  is constant outside of a finite set, as was proved by Liggett (1976) by using coupling techniques.

In order to illustrate how one uses coupling to prove such results, we will outline the proof that every extremal shift invariant measure  $\mu \in \mathcal{I}$  is some  $\nu_\rho$ . The coupled process  $(\eta_t, \zeta_t)$  is defined by the following.

1. The exponential waiting times are coupled. In other words, the *same* exponential random variables are used by a particle at  $x$  at time  $t$  in the two processes  $\eta_t$  and  $\zeta_t$  to decide when to try to move.
2. The right/left choices are coupled, so that the *same* Bernoulli random variables are used by a particle at  $x$  at time  $t$  in the two processes  $\eta_t$  and  $\zeta_t$  to decide where to try to move.
3. The actual transitions are then determined by each process’s occupancy. For example, if the particle at  $x$  at time  $t$  tries to move to  $x + 1$ , it may be that  $\eta_{t-}(x+1) = 1$  and  $\zeta_{t-}(x+1) = 0$ . In this case, the particle at  $x$  will move in the  $\zeta$  process, but not in the  $\eta$  process.

With these rules, here are some of the possible transitions for the coupled process:

$$\begin{array}{l} \eta: \quad 1 \quad 0 \quad \rightarrow \quad 0 \quad 1 \quad 1 \quad 0 \quad \rightarrow \quad 0 \quad 1 \quad 1 \quad 1 \quad \rightarrow \quad 1 \quad 1 \\ \zeta: \quad 1 \quad 0 \quad \rightarrow \quad 0 \quad 1 \quad 0 \quad 1 \quad \rightarrow \quad 0 \quad 1 \quad 1 \quad 0 \quad \rightarrow \quad 0 \quad 1 \end{array}$$

The rate at which each of these transitions occurs is  $p$ . By a “discrepancy,” we will mean a site at which  $\eta(x) \neq \zeta(x)$ . A key observation is that, with our definition of the coupling, discrepancies can move and disappear, but they cannot be created. (This would not be true if  $\eta_t$  and  $\zeta_t$  were allowed to evolve

independently, for example.) Therefore, the density of discrepancies does not increase. It follows that in equilibrium, the coupled process cannot have discrepancies of both types  $(\eta, \zeta) = (1, 0), (0, 1)$ , since if it did, there would be a positive rate of decrease of discrepancies, which cannot occur in equilibrium. (Actually, this argument only works if the distributions are shift invariant, so that we can talk about “densities” of discrepancies. The basic principle is still useful without assuming shift invariance, however, but its implementation becomes more difficult.) Here is how the argument goes.

1. Find an extremal shift invariant stationary distribution for the coupled process  $\gamma$  with marginals  $\mu$  and  $\nu_\rho$ .
2. Check that  $\gamma(\{(\eta, \zeta): \eta \leq \zeta\}) = 1$  or  $\gamma(\{(\eta, \zeta): \eta \geq \zeta\}) = 1$ .
3. Show that for every  $\rho \in [0, 1]$ , either  $\mu \leq \nu_\rho$  or  $\mu \geq \nu_\rho$ .
4. Conclude that  $\mu = \nu_\rho$  for some  $\rho \in [0, 1]$ .

The next natural problem is to find the domain of attraction of each  $\nu_\rho$ . Here we begin to see more clearly the differences between the symmetric ( $p = 1/2$ ) and asymmetric processes. In the symmetric case, a sufficient (and nearly necessary) condition for  $\mu_t \Rightarrow \nu_\rho$  is

$$\sum_y p_t(x, y) \eta(y) \rightarrow \rho$$

in probability relative to  $\mu = \mu_0$  for every  $x \in Z$ , where  $p_t(x, y)$  are the transition probabilities for the simple symmetric random walk on  $Z$  [Spitzer (1974)].

In the asymmetric case, the situation is much more delicate. To illustrate, take the initial distribution to be the product measure  $\nu_{\lambda, \rho}$  with density

$$(5.1) \quad \nu_{\lambda, \rho}\{\eta: \eta(x) = 1\} = \begin{cases} \lambda, & \text{if } x < 0, \\ \rho, & \text{if } x \geq 0. \end{cases}$$

This is about the simplest spatially inhomogeneous initial distribution that could be chosen. A consequence of Spitzer’s theorem above is that

$$\lim_{t \rightarrow \infty} \mu_t = \nu_{(\lambda + \rho)/2}$$

if  $p = 1/2$  and  $\mu_0 = \nu_{\lambda, \rho}$ . By contrast, we have the following result for asymmetric systems, in which the dependence of the limit on  $\lambda$  and  $\rho$  is more elaborate:

**THEOREM 5.2.** *Suppose  $p > 1/2$  and  $\mu_0 = \nu_{\lambda, \rho}$ . Then*

$$\lim_{t \rightarrow \infty} \mu_t = \begin{cases} \nu_{1/2}, & \text{if } \lambda \geq \frac{1}{2} \text{ and } p \leq \frac{1}{2}, \\ \nu_\rho, & \text{if } \rho \geq \frac{1}{2} \text{ and } \lambda + \rho > 1, \\ \nu_\lambda, & \text{if } \lambda \leq \frac{1}{2} \text{ and } \lambda + \rho < 1, \\ \frac{1}{2}\nu_\lambda + \frac{1}{2}\nu_\rho, & \text{if } \lambda < \rho \text{ and } \lambda + \rho = 1. \end{cases}$$

The first three cases were proved by Liggett (1975, 1977), and the more difficult final statement was proved by Andjel, Bramson and Liggett (1988).

The most interesting aspect of this last statement is that the limit is a nontrivial mixture of extremal stationary distributions in this case.

In order to understand the statement of Theorem 5.2, we will begin by making a formal computation that relates the exclusion process to a well-studied partial differential equation. Let  $u(x, t) = \mu_t\{\eta: \eta(x) = 1\}$ , and pretend that  $\mu_t$  is a product measure for all  $t$ . While this is clearly not true (for  $\lambda \neq \rho$ ), the fact that the initial distribution is a product measure and the (relevant) extremal stationary distributions are product measures indicates that this may not be far from the truth. With this assumption,

$$\begin{aligned} \frac{d}{dt}u(x, t) &= p\mu_t\{\eta: \eta(x) = 0, \eta(x-1) = 1\} + q\mu_t\{\eta: \eta(x) = 0, \eta(x+1) = 1\} \\ &\quad - p\mu_t\{\eta: \eta(x) = 1, \eta(x+1) = 0\} - q\mu_t\{\eta: \eta(x) = 1, \eta(x-1) = 0\} \\ &\approx pu(x-1, t)[1 - u(x, t)] + qu(x+1, t)[1 - u(x, t)] \\ &\quad - pu(x, t)[1 - u(x+1, t)] - qu(x, t)[1 - u(x-1, t)], \end{aligned}$$

which is a discrete approximation to Burgers' partial differential equation:

$$(5.3) \quad \frac{\partial u}{\partial t} + (p - q) \frac{\partial}{\partial x} [u(1 - u)] = 0.$$

When  $p \neq 1/2$ , (5.3) is nonlinear, and this nonlinearity can give rise to discontinuities ("shocks") in the solutions. (When  $p = 1/2$ , the scaling we have used is not the correct one—the correct one leads to the heat equation  $u_t = \frac{1}{2}u_{xx}$ .) Take  $p < 1/2$  and the initial profile

$$u(x, 0) = \begin{cases} \lambda, & \text{if } x < 0, \\ \rho, & \text{if } x \geq 0, \end{cases}$$

which corresponds to the density of the initial distribution (5.1) for the exclusion process. Then if  $\lambda \geq \rho$ , the (weak "entropic") solution  $u(x, t)$  to (5.3) does not have shocks, and

$$\lim_{t \rightarrow \infty} u(x, t) = \begin{cases} \rho, & \text{if } \rho \geq \frac{1}{2}, \\ \lambda, & \text{if } \lambda \leq \frac{1}{2}, \\ \frac{1}{2}, & \text{if } \rho \leq \frac{1}{2} \leq \lambda. \end{cases}$$

On the other hand, if  $\lambda < \rho$ , the discontinuity in the initial condition persists at later times, and moves at rate  $v = (p - q)(1 - \lambda - \rho)$ , so that

$$u(x, t) = u(x - vt, 0).$$

These limiting results for the solution to (5.3) suggest the answers given in Theorem 5.2, except for the fourth case. In that case,  $v = 0$  and the shock in  $u(x, t)$  remains at the origin. The statement in Theorem 5.2 suggests that what is happening in the exclusion process is that there is a shock whose

location behaves something like a simple random walk, in that with probability about  $1/2$  it is to the left of the origin, with probability  $1/2$  it is to the right of the origin, and its distance from the origin is greater than the size of the “disturbance” near the shock (so that the distribution near 0 will look like  $\nu_\lambda$  in one case and like  $\nu_\rho$  in the other).

Recently, Ferrari and Fontes (1994), building on earlier work by Ferrari, Kipnis and Saada (1991) and Ferrari (1992) have developed a very precise picture of this sort. To describe it, take  $\rho > \lambda$  and use the initial distribution  $\nu_{\lambda, \rho}$ , except that a “second-class” particle is placed at the origin initially. This particle moves like the others, except that if one of the other particles tries to move to its position, then the two particles interchange their positions. “Second class” reflects the fact that the other particles have precedence over it. This definition has the important property that both the process of first-class particles and the process of all particles are Markovian versions of the exclusion process with the corresponding initial distribution.

Let  $X_t$  be the position of the second-class particle at time  $t$ . The  $X_t$  identifies the position of the shock in the following sense: the process viewed from  $X_t$  is Markov, and its distribution converges as  $t \rightarrow \infty$  to a stationary distribution  $\gamma$  that behaves asymptotically (as  $|x| \rightarrow \infty$ ) like  $\nu_{\lambda, \rho}$  in the sense that

$$\lim_{n \rightarrow +\infty} \gamma \circ \tau_n = \nu_\rho, \quad \lim_{n \rightarrow -\infty} \gamma \circ \tau_n = \nu_\lambda,$$

where  $\tau_n$  denotes the spatial shift by  $n$  units. Furthermore

$$EX_t = vt = (p - q)(1 - \lambda - \rho)t,$$

$$D = \lim_{t \rightarrow \infty} \frac{\text{var } X_t}{t} = (p - q) \frac{\rho(1 - \rho) + \lambda(1 - \lambda)}{\rho - \lambda},$$

and the corresponding central limit theorem for  $X_t$  holds. As a consequence,

$$\lim_{t \rightarrow \infty} \mu_t \circ \tau_{vt + a\sqrt{t}} = (1 - \alpha)\nu_\lambda + \alpha\nu_\rho,$$

where  $\alpha = P(Z \leq a)$  and  $Z$  has the normal distribution with mean 0 and variance  $D$ . Taking  $v = a = 0$ , we get the fourth case of Theorem 5.2.

The version of the exclusion process that contains both first- and second-class particles has been studied for its own sake, as well as for the insight it has given to the behavior of the exclusion process itself. Explicit computations for this “two-species” system have been carried out by Derrida, Janowski, Lebowitz and Speer (1993). These then led to an analysis of the stationary distributions of the two-species system by Ferrari, Fontes and Kohayakawa (1994).

The connections between the exclusion process and the partial differential equation (5.3) extend well beyond the understanding of the behavior described in Theorem 5.2. The area of interacting particle systems that deals with these connections is known as hydrodynamics. A very rough description of a hydrodynamical result for the asymmetric exclusion process is the

following: suppose the initial distribution  $\mu_0^N$  is sufficiently close to being a product measure with

$$\mu_0^N\{\eta: \eta(x) = 1\} = u\left(\frac{x}{N}\right)$$

for some reasonable function  $u$  on  $\mathbb{R}^1$ , and let  $\mu_t^N$  be the distribution at time  $t$  of the exclusion process that is speeded up by a factor of  $N$ . Then for large  $N$ ,  $\mu_t^N$  will be approximately a product measure with

$$\mu_t^N\{\eta: \eta(x) = 1\} = u\left(\frac{x}{N}, t\right),$$

where  $u(x, t)$  is the (entropic) solution of (5.3) with  $u(x, 0) = u(x)$ .

Hydrodynamics has flourished during the past fifteen years, and many results for various interacting systems have been obtained. Of the many papers and books on this subject, we mention here only Rost (1981), Kipnis, Olla and Varadhan (1989), DeMasi and Presutti (1991), Rezakhanlou (1991), Varadhan (1993) and Kipnis and Landim (1997).

The one-dimensional exclusion process and similar systems have been used to investigate some seemingly unrelated models. One very nice application of hydrodynamics was given recently by Aldous and Diaconis (1995). They consider Ulam's problem on the length  $L_n$  of the longest increasing subsequence in a random permutation of  $\{1, \dots, n\}$ . It has been known for twenty years that

$$(5.4) \quad \lim_{n \rightarrow \infty} \frac{L_n}{\sqrt{n}} = 2$$

in probability. In their paper, Aldous and Diaconis showed that one can deduce this result from a hydrodynamical theorem for a process they call "Hammersley's interacting particle process," which is similar to the asymmetric exclusion process.

Here is a brief description of Hammersley's process  $\zeta_t$ , and of the heuristic that leads to the number 2 in (5.4). The states  $\zeta$  of the process are locally finite collections of points in  $(0, \infty)$ . At points  $(x, t)$  of a space-time Poisson process on  $(0, \infty)^2$ , the point in  $\zeta_t$  that is closest to  $x$  on the right is moved to  $x$ . If there is no point in  $\zeta_t$  to the right of  $x$ , a new point is put there. [Formally, the process is defined on the class of finite collections of points in  $(0, A)$  for fixed  $A$ ; the definitions for different  $A$ 's are clearly compatible. This allows us to start the process off from the empty set, which we will do.]

The hydrodynamic heuristic asserts that  $\zeta_t$  is approximately distributed as an inhomogeneous Poisson process with rate  $\lambda(x, t)$  at time  $t$  and location  $x$ , where  $\lambda(x, t)$  is determined by some evolution rule. This is analogous to the assumption that led to our "derivation" of (5.3). Let  $D_{x,t}$  be the distance from  $x$  to the nearest point in  $\zeta_t$  to the left of  $x$ . Then

$$\frac{d}{dt} E|\zeta_t(0, x)| = ED_{x,t}.$$

For a Poisson process, we would expect

$$ED_{x,t} \sim \frac{1}{\lambda(x,t)} \quad \text{and} \quad \frac{d}{dx} E|\zeta_i(0,x)| = \lambda(x,t).$$

Putting these together, and letting  $w(x,t) = E|\zeta_i(0,x)|$ , we conclude that the following should be approximately true:

$$\frac{\partial w}{\partial t} \frac{\partial w}{\partial x} = 1, \quad w(x,0) = w(0,t) = 0.$$

The solution to this PDE is  $w(x,t) = 2\sqrt{xt}$ . It is the 2 appearing in this solution that leads to the limit in (5.4).

Another fruitful connection has been with queueing theory. Kipnis (1986) exploited the following relation between the exclusion process and a series of queues: label the initial positions of the particles as  $\dots < x_{-1} < x_0 < x_1 < \dots$ , and look at their positions at time  $t$ :  $\dots < x_{-1}(t) < x_0(t) < x_1(t) < \dots$ . Let  $\zeta_i(i) = x_{i+1}(t) - x_i(t) - 1 =$  the number of empty sites between the  $i$ th and  $(i+1)$ st particles. Then  $\zeta_i(i)$  can be thought of as the length of the  $i$ th queue in a series of queues. The dynamics of the exclusion process translates into the following dynamics for the queues: at rate 1, a customer in the  $i$ th queue (provided there is one) moves, and chooses to move to the  $(i-1)$ st queue with probability  $p$  and to the  $(i+1)$ st queue with probability  $q = 1 - p$ . With this identification, the fact that the  $\nu_p$ 's are invariant for the exclusion process translates into the well-known statement that products of geometric distributions are stationary for the system of queues. In his paper, Kipnis (1986) used this connection to deduce a central limit theorem for a tagged particle in the exclusion process from results for the system of queues.

More recently, Mountford and Prabhakar (1995) have used techniques developed for the analysis of exclusion-type particle systems to solve an old problem in queueing theory. To describe their result, let  $M_\alpha$  be the class of all stationary ergodic point processes on the line of intensity  $\alpha < 1$ . Use  $\Lambda \in M_\alpha$  as the arrival process for a single server queue with exponential service times of parameter one. The resulting exit process is again a point process in  $M_\alpha$ . This operation defines a mapping  $T$  on  $M_\alpha$ . The Poisson process  $\Pi_\alpha$  of intensity  $\alpha$  is in  $M_\alpha$  and is a fixed point for  $T$ :  $T\Pi_\alpha = \Pi_\alpha$  (this is known as Burke's theorem). Anantharam (1993) proved that  $\Pi_\alpha$  is the only fixed point for  $T$  in  $M_\alpha$ . The Mountford–Prabhakar result is the corresponding convergence statement:

$$(5.5) \quad \lim_{n \rightarrow \infty} T^n \Lambda = \Pi_\alpha$$

for every  $\Lambda \in M_\alpha$ .

The similarity between this setup and exclusion processes can be seen by laying out sample paths of the arrival process and corresponding departure process on the line, identifying a departure time with the earlier arrival time of that same customer. Thus the sample path version of  $T$  moves particles (times) to the right by various random amounts, which is something like a discrete time, continuous state space version of the one-sided exclusion process.

The proof that  $T^n\Lambda$  is close to  $\Pi_a$  for large  $n$  is based on an elaborate coupling argument, which is reminiscent of the couplings used to characterize the stationary distributions of the asymmetric exclusion process. Here is a simplified version of the basic idea: a pathwise version of  $T$  can be constructed from a Poisson process  $\Pi$  of intensity 1, which represents the times at which service would be completed if the queue were always nonempty. If this same Poisson process  $\Pi$  is applied to both  $\Lambda$  and  $\Pi_\alpha$ , the result is a new pair  $\Lambda'$  and  $\Pi'_\alpha$ . (This is the coupling.) The intensity of points in  $\Lambda' \cap \Pi'_\alpha$  is greater than the intensity of points in  $\Lambda \cap \Pi_\alpha$ , except in trivial cases. (This part of the argument is the key to the proof of Anantharam's theorem.) Iterating this construction, one obtains a sequence of pairs  $(T^n\Lambda, T^n\Pi_\alpha)$ , and Mountford and Prabhakar prove that the intensity of  $T^n\Lambda \cap T^n\Pi_\alpha$  tends to  $\alpha$  as  $n \rightarrow \infty$ , which proves (5.5). They do this by exploiting the following coloring scheme: "customers" are colored yellow, blue or red according to whether they are in

$$T^n\Lambda \cap T^n\Pi_\alpha, \quad T^n\Lambda \setminus T^n\Pi_\alpha \quad \text{or} \quad T^n\Pi_\alpha \setminus T^n\Lambda.$$

A (joint) queue discipline is imposed that has the properties that yellow customers remain yellow forever, and order is preserved among customers which never become yellow. Finally, they argue that if there were a positive density of customers that never become yellow, then there must have been a situation initially which contradicts ergodicity.

#### REFERENCES

- ALDOUS, D. and DIACONIS, P. (1995). Hammersley's interacting particle process and longest increasing subsequences. *Probab. Theory Related Fields* **103** 199–213.
- ANANTHARAM, V. (1993). Uniqueness of stationary ergodic fixed point for a  $\cdot/M/k$  node. *Ann. Appl. Probab.* **3** 154–173.
- ANDJEL, E. D. (1988). The contact process in high dimensions. *Ann. Probab.* **16** 1174–1183.
- ANDJEL, E. D. (1992). Survival of multidimensional contact process in random environments. *Bol. Soc. Brasil. Mat.* **23** 109–119.
- ANDJEL, E. D., BRAMSON, M. and LIGGETT, T. M. (1988). Shocks in the asymmetric exclusion process. *Probab. Theory Related Fields* **78** 231–247.
- ANDJEL, E. D., LIGGETT, T. M. and MOUNTFORD, T. (1992). Clustering in one-dimensional threshold voter models. *Stochastic Process. Appl.* **42** 73–90.
- BEZUIDENHOUT, C. and GRIMMETT, G. (1990). The critical contact process dies out. *Ann. Probab.* **18** 1462–1482.
- BRAMSON, M., DURRETT, R. and SCHONMANN, R. H. (1991). The contact process in a random environment. *Ann. Probab.* **19** 960–983.
- CASSANDRO, M., GALVES, A., OLIVIERI, E. and VARES, M. E. (1984). Metastable behavior of stochastic dynamics: a pathwise approach. *J. Statist. Phys.* **35** 603–628.
- CHEN, M. F. (1992). *From Markov Chains to Non-Equilibrium Particle Systems*. World Scientific, Singapore.
- CLIFFORD, P. and SUDBURY, A. (1973). A model for spatial conflict. *Biometrika* **60** 581–588.
- COX, J. T. and DURRETT, R. (1991). Nonlinear voter models. In *Random Walks, Brownian Motion and Interacting Particle Systems. A Festschrift in Honor of Frank Spitzer* 189–201. Birkhäuser, Boston.
- COX, J. T. and GRIFFEATH, D. (1983). Occupation time limit theorems for the voter model. *Ann. Probab.* **11** 876–893.
- DEMASI, A. and PRESUTTI, E. (1991). *Mathematical Methods for Hydrodynamic Limits*. Springer, New York.

- DERRIDA, B., JANOWSKY, S. A., LEBOWITZ, J. L. and SPEER, E. R. (1993). Exact solution of the totally asymmetric simple exclusion process: shock profiles. *J. Statist. Phys.* **73** 813–842.
- DOBRUSHIN, R. L. (1971a). Markov processes with a large number of locally interacting components: existence of a limit process and its ergodicity. *Problems Inform. Transmission* **7** 149–164.
- DOBRUSHIN, R. L. (1971b). Markov processes with many locally interacting components—the reversible case and some generalizations. *Problems Inform. Transmission* **7** 235–241.
- DURRETT, R. (1980). On the growth of one-dimensional contact processes. *Ann. Probab.* **8** 890–907.
- DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth, Belmont, CA.
- DURRETT, R. (1991). The contact process, 1974–1989. In *Proceedings of the 1989 AMS Seminar on Random Media* 1–18. Amer. Math. Soc., Alexandria, VA.
- DURRETT, R. (1992a). Stochastic growth models—bounds on critical values. *J. Appl. Probab.* **29** 11–20.
- DURRETT, R. (1992b). Multicolor particle systems with large threshold and range. *J. Theoret. Probab.* **5** 127–152.
- DURRETT, R. (1995). Ten lectures on particle systems. *Ecole d'Été de Probabilité de Saint-Flour. Lecture Notes in Math.* **1608** 97–201. Springer, Berlin.
- DURRETT, R. and GRIFFEATH, D. (1982). Contact processes in several dimensions. *Z. Wahrsch. Verw. Gebiete* **59** 535–552.
- DURRETT, R. and LIU, X. (1988). The contact process on a finite set. *Ann. Probab.* **16** 1158–1173.
- DURRETT, R. and NEUHAUSER, C. (1997). Coexistence results for some competition models. *Ann. Appl. Probab.* **7**.
- DURRETT, R. and SCHINAZI, R. (1995). Intermediate phase for the contact process on a tree. *Ann. Probab.* **23** 668–673.
- DURRETT, R. and SCHONMANN, R. (1988). The contact process on a finite set. II. *Ann. Probab.* **16** 1570–1583.
- DURRETT, R., SCHONMANN, R. and TANAKA, N. (1989). The contact process on a finite set. III. The critical case. *Ann. Probab.* **17** 1303–1321.
- DURRETT, R. and STEIF, J. E. (1993). Fixation results for threshold voter systems. *Ann. Probab.* **21** 232–247.
- FERRARI, P. (1992). Shock fluctuations in asymmetric simple exclusion. *Probab. Theory Related Fields* **91** 81–101.
- FERRARI, P. and FONTES, L. R. G. (1994). Shock fluctuations in the asymmetric simple exclusion process. *Probab. Theory Related Fields* **99** 305–319.
- FERRARI, P., FONTES, L. R. G. and KOHAYAKAWA, Y. (1994). Invariant measures for a two-species asymmetric process. *J. Statist. Phys.* **76** 1153–1177.
- FERRARI, P., KIPNIS, C. and SAADA, E. (1991). Microscopic structure of travelling waves in the asymmetric simple exclusion process. *Ann. Probab.* **19** 226–244.
- GEMAN, D. (1990). Random fields and inverse problems in imaging. *Ecole d'Été de Probabilité de Saint-Flour XVIII. Lecture Notes in Math.* **1427** 117–193. Springer, Berlin.
- GRAY, L. and GRIFFEATH, D. (1982). A stability criterion for attractive nearest-neighbor spin systems on  $Z$ . *Ann. Probab.* **10** 67–85.
- GRIFFEATH, D. (1978). Limit theorems for nonergodic set-valued Markov processes. *Ann. Probab.* **6** 379–387.
- GRIFFEATH, D. (1979). *Additive and Cancellative Interacting Particle Systems. Lecture Notes in Math.* **724**. Springer, Berlin.
- GRIFFEATH, D. (1981). The basic contact process. *Stochastic Process. Appl.* **11** 151–186.
- GRIFFEATH, D. (1983). The binary contact path process. *Ann. Probab.* **11** 692–705.
- GRILLENBERGER, C. and ZIEZOLD, H. (1988). On the critical infection rate of the one-dimensional basic contact process: numerical results. *J. Appl. Probab.* **25** 1–8.
- GRIMMETT, G. and NEWMAN, C. (1990). Percolation in  $\infty + 1$  dimensions. In *Disorder in Physical Systems* 167–190. Oxford Univ. Press.

- HARRIS, T. E. (1974). Contact interactions on a lattice. *Ann. Probab.* **2** 969–988.
- HARRIS, T. E. (1976). On a class of set-valued Markov processes. *Ann. Probab.* **4** 175–194.
- HOLLEY, R. (1987). One-dimensional stochastic Ising models. In *Percolation Theory and Ergodic Theory of Infinite Particle Systems, IMA Vol. Math. Appl.* **8** 187–202. Springer, New York.
- HOLLEY, R. (1991). On the asymptotics of the spin–spin autocorrelation function in stochastic Ising models near the critical temperature. In *Spatial Stochastic Processes: a Festschrift in Honor of Ted Harris on His Seventieth Birthday* 89–104. Birkhäuser, Boston.
- HOLLEY, R. and LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting systems and the voter model. *Ann. Probab.* **3** 643–663.
- HOLLEY, R. and LIGGETT, T. M. (1978). The survival of contact processes. *Ann. Probab.* **6** 198–206.
- HOLLEY, R. and LIGGETT, T. M. (1981). Generalized potlatch and smoothing processes. *Z. Wahrsch. Verw. Gebiete* **55** 165–195.
- HOLLEY, R. and STROOCK, D. W. (1987). Logarithmic Sobolev inequalities and stochastic Ising models. *J. Statist. Phys.* **46** 1159–1194.
- HOLLEY, R. and STROOCK, D. W. (1988). Simulated annealing via Sobolev inequalities. *Comm. Math. Phys.* **115** 553–569.
- HOLLEY, R. and STROOCK, D. W. (1989). Uniform and  $L^2$  convergence in one dimensional stochastic Ising models. *Comm. Math. Phys.* **123** 85–93.
- KIPNIS, C. (1986). Central limit theorems for infinite series of queues and applications to simple exclusion. *Ann. Probab.* **14** 397–408.
- KIPNIS, C. and LANDIM, C. (1997). *Hydrodynamic Limits of Interacting Particle Systems*. Springer, New York.
- KIPNIS, C., OLLA, S. and VARADHAN, S. R. S. (1989). Hydrodynamics and large deviation for simple exclusion processes. *Comm. Pure Appl. Math.* **42** 115–137.
- KLEIN, A. (1994). Extinction of contact and percolation processes in a random environment. *Ann. Probab.* **22** 1227–1251.
- KONNO, N. (1994). *Phase Transitions of Interacting Particle Systems*. World Scientific, Singapore.
- LIGGETT, T. M. (1975). Ergodic theorems for the asymmetric exclusion process. *Trans. Amer. Math. Soc.* **213** 237–261.
- LIGGETT, T. M. (1976). Coupling the simple exclusion process. *Ann. Probab.* **4** 339–356.
- LIGGETT, T. M. (1977). Ergodic theorems for the asymmetric exclusion process. II. *Ann. Probab.* **5** 795–801.
- LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.
- LIGGETT, T. M. (1991). Spatially inhomogeneous contact processes. In *Spatial Stochastic Processes: a Festschrift in Honor of Ted Harris on His Seventieth Birthday* 105–140. Birkhäuser, Boston.
- LIGGETT, T. M. (1992). The survival of one-dimensional contact processes in random environments. *Ann. Probab.* **20** 696–723.
- LIGGETT, T. M. (1994). Coexistence in threshold voter models. *Ann. Probab.* **22** 764–802.
- LIGGETT, T. M. (1995a). Improved upper bounds for the contact process critical value. *Ann. Probab.* **23** 697–723.
- LIGGETT, T. M. (1995b). Survival of discrete time growth models, with applications to oriented percolation. *Ann. Appl. Probab.* **5** 613–636.
- LIGGETT, T. M. (1996). Multiple transition points for the contact process on the binary tree. *Ann. Probab.* **24** 1675–1710.
- LIGGETT, T. M. (1997). Branching random walks and contact processes on homogeneous trees. *Probab. Theory Related Fields*. To appear.
- MADRAS, N., SCHINAZI, R. and SCHONMANN, R. (1994). On the critical behavior of the contact process in deterministic inhomogeneous environment. *Ann. Probab.* **22** 1140–1159.
- MORROW, G., SCHINAZI, R. and ZHANG, Y. (1993). The critical contact process on a homogeneous tree. *J. Appl. Probab.* **31** 250–255.

- MOUNTFORD, T. S. (1993). A metastable result for the finite multidimensional contact process. *Canad. Math. Bull.* **36** 216–226.
- MOUNTFORD, T. S. and PRABHAKAR, B. (1995). On the weak convergence of departures from an infinite series of  $M/M/1$  queues. *Ann. Appl. Probab.* **5** 121–127.
- NEWMAN, C. M. and VOLCHAN, S. B. (1996). Persistent survival of one-dimensional contact processes in random environments. *Ann. Probab.* **24** 411–421.
- PEMANTLE, R. (1992). The contact process on trees. *Ann. Probab.* **20** 2089–2116.
- REZAKHANLOU, F. (1991). Hydrodynamic limit for attractive particle systems on  $Z^d$ . *Comm. Math. Phys.* **140** 417–448.
- ROCHOW, W. F. (1979). Field variants of barley yellow dwarf virus: detection and fluctuation during twenty years. *Phytopathology* **69** 655–660.
- ROST, H. (1981). Non-equilibrium behaviour of a many particle process: density profile and local equilibria. *Z. Wahrsch. Verw. Gebiete* **58** 41–53.
- SALZANO, M. and SCHONMANN, R. H. (1998). The second lowest extremal invariant measure of the contact process. Unpublished manuscript.
- SCHINAZI, R. (1994). The asymmetric contact process on a finite set. *J. Statist. Phys.* **74** 1005–1016.
- SCHONMANN, R. H. (1985). Metastability for the contact process. *J. Statist. Phys.* **41** 445–464.
- SCHONMANN, R. H. (1986). The asymmetric contact process. *J. Statist. Phys.* **44** 505–534.
- SCHONMANN, R. H. (1987). A new proof of the complete convergence theorem for contact processes in several dimensions with large infection parameter. *Ann. Probab.* **15** 382–387.
- SCHONMANN, R. H. (1994a). Slow droplet-driven relaxation of stochastic Ising models in the vicinity of the phase coexistence region. *Comm. Math. Phys.* **161** 1–49.
- SCHONMANN, R. H. (1994b). Theorems and conjectures on the droplet-driven relaxation of Stochastic Ising models. In *Probability and Phase Transition* (G. Grimmett, ed.) 265–301. Kluwer, Dordrecht.
- SIMONIS, A. (1991). Metastability for the d-dimensional contact process. *J. Statist. Phys.* **83** 1225–1239.
- SPITZER, F. (1970). Interaction of Markov processes. *Adv. in Math.* **5** 246–290.
- SPITZER, F. (1974). Recurrent random walk of an infinite particle system. *Trans. Amer. Math. Soc.* **198** 191–199.
- SPOHN, H. (1991). *Large Scale Dynamics of Interacting Particles*. Springer, New York.
- STACEY, A. M. (1994). Bounds on the critical probability in oriented percolation models. Ph.D. thesis, Cambridge Univ.
- STACEY, A. M. (1996). The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24**.
- SWEET, T. (1997). The asymmetric contact process at its second critical value. *J. Statist. Phys.* **86**.
- VARADHAN, S. R. S. (1993). Entropy methods in hydrodynamical scaling. *Nonequilibrium Problems in Many-Particle Systems. Lecture Notes in Math.* **1551** 112–145. Springer, Berlin.
- WINKLER, G. (1995). *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Springer, Berlin.
- WU, C. C. (1995). The contact process on a tree: behavior near the first phase transition. *Stochastic Process. Appl.* **57** 99–112.
- ZHANG, Y. (1997). The complete convergence theorem of the contact process on trees. *Ann. Probab.* **25**. To appear.

DEPARTMENT OF MATHEMATICS  
 UNIVERSITY OF CALIFORNIA  
 LOS ANGELES, CALIFORNIA 90095  
 E-MAIL: tml@math.ucla.edu