Stochastic models for large interacting systems and related correlation inequalities

Thomas M. Liggett[1]

Department of Mathematics, University of California, Los Angeles, CA 90095

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2008.

Contributed by Thomas M. Liggett, August 3, 2010 (sent for review July 1, 2010)

A very large and active part of probability theory is concerned with the formulation and analysis of models for the evolution of large systems arising in the sciences, including physics and biology. These models have in their description randomness in the evolution rules, and interactions among various parts of the system. This article describes some of the main models in this area, as well as some of the major results about their behavior that have been obtained during the past 40 years. An important technique in this area, as well as in related parts of physics, is the use of correlation inequalities. These express positive or negative dependence between random quantities related to the model. In some types of models, the underlying dependence is positive, whereas in others it is negative. We give particular attention to these issues, and to applications of these inequalities. Among the applications are central limit theorems that give convergence to a Gaussian distribution.

contact process | exclusion process | Glauber dynamics | voter models

Models for Interacting Systems

During the past half century, mathematical models for the evolution of large interacting systems arising in a number of scientific areas have been proposed and analyzed. Here are some of these areas, together with a sampling of the many papers and books in which such models have been discussed: magnetic systems (1), high energy scattering (2), dynamics of mutation in a structured population (3), tumor growth (4, 5), competition between different strains of viruses (6), mutations of pathogens (7), biopolymers (8), epidemics (9, 10), ecology (11, 12), hydrology (13), cooperative behavior (14–16), spatial distribution of unemployment (17), and the analysis of traffic flow (18–20).

The main objective in the study of these models is to describe their long-time behavior. Usually, the models contain one or more parameters. An important issue is to determine how the long-time behavior depends on these parameters. Often there is a sharp transition in the nature of the behavior at some particular parameter value. This situation is described by saying that a phase transition occurs there.

Some of the analysis of these systems has been mathematical, whereas other approaches have been based on simulations and heuristics. Dobrushin (21) and Spitzer (22) are usually credited with initiating the mathematical developments about 40 years ago. The modern theory of models of this type is treated in my two monographs (23, 24).

Typically, the model is a random process η_t with state space $\{0,1\}^{Z^d}$ of binary configurations on the d-dimensional integer lattice Z^d. The interpretation of the values 0 and 1 at a site $x \in Z^d$ depends on the model, and on the area that motivated it. The process satisfies the Markov property, which means that once one knows the state of the system at a given time t, the evolution of the system after that time does not depend on its behavior before time t. It follows that the evolution rules can be described by specifying how the process will behave in an infinitesimal time period $(t, t+dt)$ as a (random) function of the state η_t at time t. This is analogous to describing a deterministic function $y(t)$ by a differential equation that it satisfies.

In the present context, the evolution rules are given by certain transition rates. To say that the transition $\eta \to \zeta$ from one configuration to another occurs at rate $\lambda > 0$ means that in a short time period of length ϵ, the transition occurs with probability approximately $\lambda\epsilon$. Usually, the transition rate will depend on η, and this dependence leads to interactions among various parts of the system.

Our focus here is on models on the graph Z^d. However, in many contexts, such as communications and social systems, it is more natural to consider more general graphs. For related evolutions on random graphs, see ref. 25, for example. Some of the results mentioned below have extensions to cases in which Z^d is replaced by a general countable set.

A very useful tool in the mathematical analysis of interacting systems is that of correlation inequalities—inequalities that assert that the state of one random quantity has a positive (or negative) influence on the state of another. These inequalities often make it possible to treat dependent random quantities as if they were independent. This is, of course, a great simplification. We will see a number of specific instances of this simplification in the present paper.

Here is the plan for this paper: I begin by describing some of the most important models in this area—voter, contact, magnetic, and exclusion—and give a sampling of the most important results about them. Then I discuss the associated correlation inequalities (positive for the first three models and negative for exclusion) and present some consequences that follow from them.

Before getting started, I need to introduce a bit of notation and terminology from probability theory. The probability of an event A is denoted by $P(A)$. If it appears with a superscript, as in $P^\eta(A)$, the superscript η is the initial state of the process. Similarly, $E^\eta X$ is the expected value, or mean value, of the random quantity X, when the initial state of the system is η.

Bernoulli random variables are random variables that take only two values, typically 0 and 1. Thus a probability distribution on $\{0,1\}^{Z^d}$ gives the joint distribution of a collection of (generally not independent) Bernoulli random variables indexed by Z^d.

Voter Models. The simplest models in this area are known as voter models. They were introduced in refs. 26 and 27. Later it was realized that they are very similar to the earlier "stepping stone" model of population genetics introduced in ref. 3. A biased version was proposed as a model for tumor growth in ref. 4.

In ref. 27, the idea was to model conflict between populations. Sites x for which $\eta(x) = 1$ represent areas controlled by one population; those for which $\eta(x) = 0$ are controlled by the other. A site controlled by one group is taken over by the other at a rate that is proportional to the number of neighbors controlled by the opposing group.

The voter interpretation of ref. 26 was not our motivation in that paper—the actual motivation was of a more mathematical

Author contributions: T.M.L. wrote the paper.

The author declares no conflict of interest.

[1]E-mail: tml@math.ucla.edu.

nature. The idea was to identify a class of models that had properties such as Eq. **2** below that permitted an essentially complete mathematical analysis of the process.

Even though I do not claim that this is a good model for electoral behavior, I will describe the process in electoral terms. Each site in Z^d represents a person, who at any given time, has one of two possible opinions on an issue, labeled 0 and 1. Each person waits a unit exponentially distributed time T, i.e., one for which $P(T > t) = e^{-t}$. At that time, he chooses one of his $2d$ neighbors at random and adopts that neighbor's opinion.

Here is the main question: Is it the case that the system reaches a consensus (in the voter interpretation), or that one population takes over the entire space (in the spatial conflict interpretation), in the sense that

$$\lim_{t\to\infty} P^{\eta}[\eta_t(x) = \eta_t(y)] = 1 \quad \textbf{[1]}$$

for all $x, y \in Z^d$ and all initial configurations η?

The key to the answer lies in a connection between the voter model and a classical random walk $X(t)$, which moves on Z^d in the following way: It waits where it is for a unit exponential time, then moves to a randomly chosen neighbor, and continues in this way. Here is a special case of the connection. Suppose that initially each voter independently chooses opinion 1 with probability α and opinion 0 with probability $1 - \alpha$. Then the probability that the individuals at x and y share the same opinion at time t can be expressed in terms of a probability related to the random walk:

$$P[\eta_t(x) = \eta_t(y)] = 1 - 2\alpha(1 - \alpha)P^{y-x}[X(s) \neq 0 \text{ for all } s < t]. \quad \textbf{[2]}$$

A classical result in probability theory states that $X(t)$ is recurrent (i.e., hits 0 eventually with probability 1) if $d = 1$ or 2, but not if $d \geq 3$. It follows that the limiting statement Eq. **1** holds if and only if $d \leq 2$.

A key issue for all the models we consider is understanding the nature of their stationary distributions. A probability distribution μ on $\{0,1\}^{Z^d}$ is said to be stationary for η_t if the process with that initial distribution continues to have distribution μ at all later times. The importance of stationary distributions comes from the fact that any limiting distribution of the process as $t \to \infty$ is stationary. Thus the identification of stationary distributions is the first step in the analysis of the limiting behavior of η_t.

When Eq. **1** holds, the voter model has only trivial stationary distributions: If μ is stationary, then it concentrates on configurations representing consensus;

$$\mu\{\eta : \eta \equiv 0 \text{ or } \eta \equiv 1\} = 1.$$

When $d \geq 3$, the situation is quite different (26). Both opinions can coexist in equilibrium:

Theorem 1. *Suppose* $d \geq 3$.

(a) For every $0 \leq \alpha \leq 1$, *there is a stationary distribution* μ_α *in which the proportion of 1's is exactly* α. *It is obtained by starting the system with all voters having opinion 1 independently with probability* α, *and then passing to the limit as* $t \to \infty$.

(b) Every stationary distribution can be expressed as an average of the distributions μ_α.

For extensions of this result to voter models on a general countable set, see Chapter V of ref. 23.

Rather than starting the system with a distribution in which all voters are equivalent, one could start the voter model with a single voter with opinion 1 and ask for the behavior of the number N_t of voters with that opinion at time t. It turns out that $EN_t = 1$ for all t, and it is then easy to check that $P(N_t \geq 1) \to 0$ as $t \to \infty$. In ref. 28, it is proved that this probability has the following order of magnitude as $t \to \infty$:

$$P(N_t \geq 1) \sim \begin{cases} 1/\sqrt{t} & \text{if } d = 1; \\ \log t/t & \text{if } d = 2; \\ 1/t & \text{if } d \geq 3. \end{cases}$$

Furthermore, conditioned on the event $\{N_t \geq 1\}$ of nonextinction of opinion 1, $N_t P(N_t \geq 1)$ converges to a unit exponential distribution if $d \geq 2$.

In the biased version of the voter model that was proposed to model tumor growth, rates for the transitions $0 \to 1$ are larger than the corresponding rates for the transitions $1 \to 0$. The interpretation now is that 1's correspond to cancerous cells and 0's to normal cells. The process starts with a single cancerous cell. There is positive probability that the tumor disappears, but as a result of the bias, there is also a positive probability that it continues to grow forever. One of the important results for this model (29) is that the growth of the tumor is linear in time and that it takes on a deterministic asymptotic shape as $t \to \infty$.

Contact Models. The contact process was introduced in ref. 9. Here the interpretation is one of spread of infection. Later it was realized that the model is closely related to a field theory in high energy physics (2). This is surprising, because nothing in the description of the model suggests that there might be such a connection.

With the infection interpretation, sites with the value 1 are infected, whereas those with the value 0 are healthy. Infected sites remain infected for a unit exponential time, independently of the states of their neighbors, and then become healthy. Healthy sites become infected at rate

$$\lambda \times (\text{the number of infected neighbors}),$$

where λ is a positive parameter. This transition mechanism is deceptively similar to that of the voter model, but the analysis is much harder because connections such as Eq. **2** no longer hold.

Now a type of phase transition occurs. For small values of λ, the infection dies out, in the sense that

$$\lim_{t\to\infty} P^{\eta}[\eta_t(x) = 1] = 0$$

for all initial configurations η and all sites x. For larger λ, this is not the case, and there is a probability distribution ν on $\{0,1\}^{Z^d}$ with a positive density of infected sites that is stationary for the evolution. The threshold value λ_d that separates the regimes of survival and extinction of the infection cannot be computed exactly, even in one dimension, but it can be approximated numerically. It does satisfy the rigorous bounds (9, 30)

$$\frac{1}{2d - 1} \leq \lambda_d \leq \frac{2}{d}.$$

Thus $1 \leq \lambda_1 \leq 2$, for example. Somewhat better bounds are available in low dimensions: $1.539 \leq \lambda_1 \leq 1.942$. For large d, the lower bound above is asymptotically correct: $2d\lambda_d \to 1$ as $d \to \infty$.

At this point, it is reasonable to ask the following question: In view of the fact that real systems are finite, why is it reasonable to study infinite models at all? The contact process provides a clear answer to this question. If the set of sites for the mathematical model is taken to be the finite box $\{1, \ldots, N\}^d$ instead of Z^d, then the infection dies out for all values of λ. This is a consequence of an elementary result about finite state Markov chains. However, how long it takes to die out depends strongly on the value of λ. If $\lambda < \lambda_d$ (where λ_d is the critical value for the process on Z^d), then the extinction time is logarithmic in the system size N^d, whereas if $\lambda > \lambda_d$, it is exponential in the system size (31). Therefore, if one observes a large finite system for a large finite time, one will see the infection die out in the subcritical case, but survive in the supercritical case. So, the process on Z^d is a better model for a large finite system than a process on a finite set would be!

Magnetic Models. In this case, it is more natural to let the possible values of $\eta(x)$ be ± 1 rather than 0 and 1, because they represent magnetic spins. The central objects of study in statistical mechanics are the Gibbs distributions for the Ising model, which are probability distributions μ on $\{-1,+1\}^{Z^d}$. They are described by specifying the conditional probabilities for the state at $x \in Z^d$, given the states at other sites:

$$\mu[\eta(x)=+1|\eta(y)=\zeta(y) \text{ for all } y \neq x] = \frac{e^{\beta\sum_{y\sim x}\zeta(y)}}{e^{\beta\sum_{y\sim x}\zeta(y)}+e^{-\beta\sum_{y\sim x}\zeta(y)}},$$

where the sums are over the neighbors y of x. Here β is a positive parameter that represents the reciprocal of the temperature of the system. Classical results include the fact that these conditional probabilities determine μ uniquely for all β in one dimension, whereas in higher dimensions the Gibbs distribution is unique for small β, but not for large β.

The transition rates for the random evolution, which is known as the Glauber dynamics (1), are chosen so that the Gibbs distributions are stationary (and in fact reversible) for the evolution. There are many choices with this property; in a simple one, the rate of flipping the state at x from $\eta(x)$ to $-\eta(x)$ is taken to be

$$e^{-\beta\eta(x)\sum_{y\sim x}\eta(y)}$$

when the configuration is η. Note that these rates are large if $\eta(x)$ differs from the states at most of its neighbors, and small if it largely agrees with them. This means that spins prefer to align themselves with their neighbors, which is certainly reasonable to expect in this context.

A natural question is whether all stationary distributions for the time evolution are Gibbs distributions. This is known to be the case if $d=1$ (which is easy because the Gibbs distribution is unique) or $d=2$ (which is much harder—see ref. 32). This remains an open problem in higher dimensions.

Although the original motivation for these models comes from physics, they have also led to important techniques known as Markov Chain Monte Carlo or Gibbs sampling. Here the objective is to simulate a Gibbs distribution on a large but finite set of sites. Rather than doing it directly, which is difficult given the large size of the system, the evolution is run for a long time t, and the distribution at that time is used as an approximation to the limiting Gibbs distribution. This is a huge field with many applications. Two references are refs. 33 and 34. The latter is an example of an application in computational biology.

Exclusion Processes. These are of a different nature than the models described so far. Transitions change the values at two sites rather than only one. Now the states 0 and 1 represent occupancy by particles (or cars in the traffic flow context). Particles move on Z^d in such a way that there is at most one particle per site. A particle at x moves to y, if it is vacant (hence the name exclusion), at rate $p(y-x)$, where $p(x) \geq 0$ for each x and $\sum_x p(x)=1$. An alternative description is that a particle at x waits a unit exponential time and then chooses a y to try to move to with probability $p(y-x)$. If y is vacant, it moves there, whereas if y is occupied, it remains at x.

Although exclusion processes seem natural in the contexts of particle motion and traffic flow, it is interesting to note that perhaps the earliest appearance of them was in a biological situation —see ref. 8. In this case, the "particles" are ribosomes that move along a messenger RNA template reading genetic information.

Again we are interested in stationary distributions. A probability distribution on $\{0,1\}^{Z^d}$ is called exchangeable if it does not change when finitely many coordinates of η are permuted. It is not hard to check that all exchangeable distributions are stationary for the exclusion process. It is harder to determine when these are all the stationary distributions. Here is one of the early results about this problem (35, 36):

Theorem 2. *Suppose $p(\cdot)$ is symmetric, i.e., $p(-x)=p(x)$ for all x. Then all stationary distributions are exchangeable.*

For extensions of this result to exclusion processes on a general countable set, see Chapter VIII of ref. 23.

The above conclusion is often false for asymmetric systems. For example, take the case in which $d=1$, $p(1)=p$, $p(-1)=1-p$, and $p(x)=0$ otherwise. If $p>\frac{1}{2}$, so particles experience a drift to the right, there are stationary distributions with respect to which there are only finitely many particles to the left of the origin, and only finitely many empty sites to the right of the origin. In one example, the coordinates $\{\eta(x), x \in Z^1\}$ are independent, with

$$P[\eta(x)=1]=\frac{p^x}{p^x+(1-p)^x}. \quad [3]$$

In fact, all stationary distributions can be constructed from these and the exchangeable ones in this case. Generalizations of this statement to one-dimensional systems with long-range jumps can be found in ref. 37. In this more general context, explicit formulas such as Eq. **3** are usually not available. This is a source of much of the difficulty that arises in the analysis.

To describe a rather surprising consequence of the asymmetry, we continue with the one-dimensional nearest-neighbor case. Suppose the initial distribution is of the following type: negative sites are independently occupied with probability λ, and nonnegative sites with probability ρ. If $\lambda=\rho$, this distribution is exchangeable, and hence stationary. What happens in the limit as $t\to\infty$ if $\lambda\neq\rho$? Here is the answer (38, 39), which is substantially more complex in the asymmetric case:

Theorem 3.

(a) If $p=\frac{1}{2}$, then

$$\lim_{t\to\infty} P[\eta_t(x)=1]=\frac{\lambda+\rho}{2}.$$

(b) If $p>\frac{1}{2}$, then

$$\lim_{t\to\infty} P[\eta_t(x)=1]=\begin{cases} \frac{1}{2} & \text{if } \lambda\geq\frac{1}{2} \text{ and } \rho\leq\frac{1}{2}; \\ \lambda & \text{if } \lambda\leq\frac{1}{2} \text{ and } \lambda+\rho<1; \\ \rho & \text{if } \rho\geq\frac{1}{2} \text{ and } \lambda+\rho>1; \\ \frac{1}{2} & \text{if } \lambda\leq\frac{1}{2} \text{ and } \lambda+\rho=1. \end{cases}$$

These results can be predicted by the behavior of associated partial differential equations (PDEs)—the heat equation

$$\frac{\partial u}{\partial t}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2}$$

if $p=\frac{1}{2}$, and Burgers' equation

$$\frac{\partial u}{\partial t}+(2p-1)\frac{\partial}{\partial x}[u(1-u)]=0 \quad [4]$$

if $p>\frac{1}{2}$. The more elaborate and interesting limiting behavior in the asymmetric case is a consequence of the nonlinearity in Eq. **4**.

The connection between the exclusion process and the PDE arises in the following way: If the evolution equations for the exclusion process are simplified by assuming that the coordinate variables $\eta_t(x)$ are independent for different x's, the result is a discrete form of the corresponding PDE. If one solves the PDE with the initial condition

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

then $\lim_{t\to\infty} u(t,x)$ takes the form given in Theorem 3.

The limiting (in distribution) occupation variables $\eta_\infty(x)$ in Theorem 3 are independent for different x's in all of these cases except $p > \frac{1}{2}$, $\lambda \leq \frac{1}{2}$, and $\lambda + \rho = 1$, when the covariances are given by

$$\mathrm{Cov}[\eta_\infty(x), \eta_\infty(y)] = \frac{1}{4}(\rho - \lambda)^2, \qquad x \neq y.$$

Exclusion processes on finite sets have been of substantial interest as well—see ref. 40, for example. To describe one recent result, suppose S is a set with n points, and place n distinguishable particles on it, one at each point. For each pair $x, y \in S$, interchange the particles at x and y at a rate that depends on the locations of the two particles. There are various Markov chains that are embedded in this structure. By following the motion of only one of the particles, one obtains a chain with n states. More generally, following the positions of $k \leq n$ particles gives rise to a chain with many more states: $n(n-1)\cdots(n-k+1)$. In this case, if one makes the particles indistinguishable, the k particles move according to a symmetric exclusion process on S.

For a concrete example, consider shuffling a standard 52 card deck. Then $n = 52$, and S is the set of possible positions of a card in the deck. The shuffling is done by interchanging the kth and lth cards at a rate that depends on k and l. For example, the rate might be higher if the two cards are closer together in the deck than if they are farther apart. If one follows the position of the ace of spades, say, the chain has 52 possible states. If one follows the positions of all 52 cards, the corresponding chain has $52! \approx 10^{68}$ states.

The rate of convergence to the stationary distribution (which is a perfectly shuffled deck in the shuffling context) is determined by the smallest nontrivial eigenvalue of a matrix made up of the transition rates. This eigenvalue can be computed easily when the chain has 52 states, say, but cannot be computed for a chain of anything like 10^{68} states. Recently, Caputo et al. (41) were able to prove the 1992 conjecture of D. Aldous that the principal eigenvalues for the large ($n!$ states) and small (n states) chains are the same for any n and any choice of rates. It follows that computing the eigenvalue for the smaller chain is enough to determine the rate of convergence to equilibrium for the larger chain.

Here is the barest outline of our approach. The proof is by induction on n. To carry out the induction step, it is necessary to take the set of size n with transition rates associated to pairs of points in that set and construct from it a set of size $n-1$, together with a new collection of rates on pairs of those points. This is done by generalizing the series, parallel, and star-triangle reductions used in electrical network theory. Using the induction hypothesis on the smaller set, the problem becomes one of showing that a particular $n! \times n!$ matrix is positive semidefinite. This is done by a careful analysis of the structure of a related large matrix.

Correlation Inequalities

There is a natural (partial) order on $\{0,1\}^{Z^d}$:

$$\eta \leq \zeta \quad \text{if } \eta(x) \leq \zeta(x) \quad \text{for all } x.$$

A real valued continuous function f on $\{0,1\}^{Z^d}$ is said to be increasing if $\eta \leq \zeta$ implies $f(\eta) \leq f(\zeta)$. An important problem is to determine the evolutions and initial distributions for which

$$Ef(\eta_t)g(\eta_t) \geq Ef(\eta_t)Eg(\eta_t)$$

for all increasing f and g and all $t > 0$. This means that the random variables $f(\eta_t)$ and $g(\eta_t)$ are positively correlated in the usual sense. This section is devoted to a discussion of this question, together with the analogous question for negative correlations.

Positive Association. A probability distribution μ on $\{0,1\}^{Z^d}$ is said to be positively associated if

$$Ef(\eta)g(\eta) \geq Ef(\eta)Eg(\eta) \quad \text{for all increasing } f \text{ and } g, \qquad \textbf{[5]}$$

when η has distribution μ. The best known result related to this concept is the Fortuin–Kasteleyn–Ginibre (FKG) theorem (42), which gives a sufficient condition (known as the FKG lattice condition) for positive association. It is easy to check this condition when the coordinates $\eta(x)$ are independent [in which case positive association was known earlier—(43)] and can often be verified for Gibbs distributions. However, the FKG lattice condition can essentially never be checked for the distribution at time t of any of the evolutions we are considering. In fact, it is often false, even if it turns out that the distribution is positively associated.

To check that the distribution at time t of an evolution is positively associated, one uses the following result (44), which applies to a very general class of processes on $\{0,1\}^S$, including the voter, contact, and magnetic (but not exclusion) models described above:

Theorem 4. *Suppose the process satisfies the following two properties:*

(a) Individual transitions affect the state at only one site.
(b) For every continuous increasing function f and every $t > 0$*, the function* $\eta \to E^\eta f(\eta_t)$ *is increasing.*

Then, if the initial distribution is positively associated, so is the distribution at all later times.

It follows from this that the limiting distribution as $t \to \infty$, if it exists, is also positively associated.

Negative Association. In the analogous definition for negative association, Eq. **5** is replaced by

$$Ef(\eta)g(\eta) \leq Ef(\eta)Eg(\eta)$$
$$\text{for all increasing } f \text{ and } g \text{ that depend on}$$
$$\text{disjoint sets of coordinates.} \qquad \textbf{[6]}$$

This last constraint is necessary, because if $f = g$, the opposite inequality Eq. **5** automatically holds for any μ.

One might hope that negative association is related to the exclusion process in much the same way that positive association is related to voter, contact, and magnetic models. Here is the reason: In the exclusion process, particles are neither created nor destroyed. Therefore, if one knows that a certain subset of Z^d has many particles, it is likely that disjoint subsets have relatively fewer particles. It turns out that in order for something like this to actually be true, $p(\cdot)$ must be symmetric: $p(-x) = p(x)$ for all x.

Whereas the intuition is fairly clear, it took 35 years to find the correct version of the connection between the symmetric exclusion process and negative association (45–47). Here is one consequence of the general statement for the symmetric exclusion process that is proved in ref. 47:

Theorem 5. *Suppose that initially, the random variables* $\{\eta(x) x \in Z^d\}$ *are independent. Then*

(a) the distribution of the process at time $t > 0$ *is negatively associated, and*
(b) if S is a subset of Z^d*, the number* $\sum_{x\in S} \eta_t(x)$ *of particles in S at time t has the same distribution as a sum* $\sum_{x\in S} \zeta_t(x)$ *of appropriately chosen independent Bernoulli random variables.*

Part (b) is a very useful property for proving limit theorems, as we will see in the next section.

Given the form of Theorem 4, one might suspect that negative association itself is preserved by the symmetric exclusion evolution. This is not the case (48). The key to Theorem 5 is finding another property that is preserved, and that implies properties (a) and (b) in this result.

The property that works is a rather unintuitive one known as stability. To describe it, suppose the exclusion process is evolving on a finite set $S = \{1,\ldots,n\}$. The random variables $\{\eta(x), x \in S\}$ are said to be stable if the (generating) function of n complex variables

$$f(z_1,\ldots,z_n) = Ez_1^{\eta(1)}\cdots z_n^{\eta(n)}$$

is not zero whenever all the z_i's have strictly positive imaginary parts. It turns out that the property of stability is preserved by the symmetric exclusion process. The fact that independent Bernoulli random variables are stable is easy to check. The fact that stable random variables are negatively associated is much more difficult to establish. On the other hand, the fact that stable random variables have property (b) of Theorem 5 is easy to see: Take $z_1,\ldots,z_n$ to be equal. Then

$$f(z,\ldots,z) = Ez^{\eta(1)+\cdots+\eta(n)}$$

is the generating function of the sum $\eta(1)+\cdots+\eta(n)$. This is a polynomial in one variable of degree n, whose zeros cannot have positive imaginary parts by the stability property and therefore cannot have negative imaginary part, because the zeros occur in conjugate pairs. They are therefore real, and in fact ≤ 0, because the polynomial is strictly positive on the positive real axis. Therefore, it can be factored in the form

$$f(z,\ldots,z) = (p_1z+1-p_1)\cdots(p_nz+1-p_n), \qquad [7]$$

where $0 \leq p_i \leq 1$ for each i. Now take $\zeta(i)$ to be independent with $P[\zeta(i)=1] = p_i$. Then $\zeta(1)+\cdots+\zeta(n)$ has generating function Eq. **7** as well, so $\eta(1)+\cdots+\eta(n)$ and $\zeta(1)+\cdots+\zeta(n)$ have the same distribution.

Consequences of Correlation Inequalities

In this section, we describe a few of the many results concerning interacting systems that are related to correlation inequalities.

Voter Models. It follows from Theorem 4 that when $d \geq 3$, the nontrivial stationary distributions μ_α for the voter model are positively associated. In fact, using Eq. **2**, one can show that the covariances for the coordinate random variables relative to μ_α are given by

$$\mathrm{Cov}[\eta(x),\eta(y)] = \alpha(1-\alpha)\frac{G(y-x)}{G(0)},$$

where

$$G(x) = \int_0^\infty P^0[X(t)=x]dt,$$

which is the expected total amount of time the random walk spends at x.

Looking ahead to comments about central limit theorems for contact and magnetic models below, note that

$$\sum_x \mathrm{Cov}[\eta(x),\eta(0)] = \infty.$$

This is an indication that the (positive) correlations among voter opinions are quite strong.

Contact Models. It took 15 years to prove that the critical contact process (the one with $\lambda = \lambda_d$) dies out. The proof (49, 50) uses several times the fact that collections of independent Bernoulli random variables are positively associated.

The nontrivial stationary distribution ν for the supercritical ($\lambda > \lambda_d$) contact process does not satisfy the FKG lattice condition (51). However, it is positively associated by Theorem 4. Theorem 4.20 of Chapter I of ref. 23 and Theorem 2.30 of Part I of ref. 24 combine to show that the covariances of $\eta(x)$ and $\eta(y)$ relative to ν decay exponentially rapidly as a function of the distance $|y-x|$. It then follows from results in refs. 52 or 53 that ν satisfies the following central limit theorem:

Theorem 6. *Let* $S_n = \sum_{|x|\leq n}\eta(x)$. *Then*

$$\frac{S_n - ES_n}{\sqrt{\mathrm{Var}(S_n)}} \Rightarrow N(0,1).$$

In this statement, $\Rightarrow$ denotes convergence in distribution, Var stands for variance, and $N(0,\sigma^2)$ represents the Gaussian distribution with mean 0 and variance σ^2.

The FKG lattice condition is equivalent to the statement that the distribution is positively associated, even after conditioning on the values of $\{\eta(x) x \in S\}$ for any S. This raises the question of whether ν is associated after some special type of conditioning. It is not when the conditioning is on the event $\eta(0) = 1$. In fact, if $d = 1$, the conditional distribution satisfies Eq. **6** rather than Eq. **5** if f depends on $\{\eta(x), x < 0\}$ and g depends on $\{\eta(x), x > 0\}$ (54). The intuition behind this is that if the origin is known to be infected, the infection must have come from somewhere. If it did not come from the left, it must have come from the right.

Nevertheless, ν is positively associated after conditioning on the event $\{\eta(x) = 0 x \in S\}$ (55, 56). A consequence of this (together with other known properties of the contact process) is that if $\{\eta(x) x \in Z^d\}$ have distribution ν and $\lambda \geq 2$, then there exist independent Bernoulli random variables $\{\zeta(x), x \in Z^d\}$ with density

$$P(\zeta(x)=1) = \frac{\lambda-2}{\lambda}$$

so that $\zeta(x) \leq \eta(x)$ for all x (57). As in Theorem 5(b), this is a connection between nonindependent Bernoulli random variables and independent ones that is very useful is analyzing the former collection.

For example, consider the site percolation model, in which one asks whether there is positive probability that infinitely many sites are connected to the origin by paths that travel only through sites for which $\eta(x) = 1$ (respectively, $\zeta(x) = 1$). Classical results for independent percolation imply that if $d \geq 2$ and λ is sufficiently large, percolation occurs for the above ζ's. The comparison result implies that it also occurs for the nonindependent η's. Motivated by ref. 57 and the biological application in ref. 58, properties that have long been known for independent percolation have recently been extended to percolation in ν for $d = 2$ in ref. 59.

Magnetic Models. Suppose that initially all spins are $+1$. Then for every $t > 0$, the covariances $\mathrm{Cov}[\eta_t(x),\eta_t(y)]$ decay exponentially rapidly as a function of $|y-x|$ by Proposition 4.18 of Chapter I of ref. 23. The random variables $\eta_t(x)$ are positively associated by Theorem 4. It then again follows that the spin variables satisfy the central limit theorem. If the (distributional) limiting random variables $\eta_\infty(x)$ satisfy

$$\sum_x \mathrm{Cov}[\eta_\infty(x),\eta_\infty(0)] < \infty, \qquad [8]$$

the same argument applies. Condition **8** holds often, but not always.

Exclusion Processes. Assume throughout that the model is symmetric, $p(-x) = p(x)$ for all x, because it is only then that useful correlation inequalities are available.

The proof of part of Theorem 2 begins with an extension of the symmetry property, which is known as duality. Consider two copies of the exclusion process, η_t and ζ_t, with initial configurations η and ζ, respectively. Then

$$P^{\eta}(\eta_t \geq \zeta) = P^{\zeta}(\eta \geq \zeta_t) \quad \textbf{[9]}$$

for all $t > 0$. When η has infinitely many particles and ζ has finitely many particles, this duality property reduces many problems for the infinite system to corresponding problems for the finite system.

By Theorem 5(a),

$$P^{\zeta}[\zeta_t(x_1) = 1, \dots, \zeta_t(x_n) = 1] \leq P^{\zeta}[\zeta_t(x_1) = 1] \cdots P^{\zeta}[\zeta_t(x_n) = 1] \quad \textbf{[10]}$$

for distinct points $x_1, \dots, x_n \in Z^d$. The right side can be interpreted as the probability that n independent (by Eq. **9** and the fact that it is a product of probabilities) particles starting at $x_1, \dots, x_n$ will be in the set $\{x : \zeta(x) = 1\}$ at time t. Thus problems relating to n particles moving with the exclusion interaction can often be reduced to problems relating to n independent particles, which is a great simplification.

Consider now the problem of the motion of a tagged particle. The tagged particle is initially placed at the origin; other sites are initially occupied with probability $\frac{1}{2}$ each. The problem concerns the asymptotic behavior of the position $X(t)$ of the tagged particle at time t. The presence of the other particles has the effect of slowing down the tagged particle. The question is, by how much is it slowed down? The following situation is special, but particularly interesting in view of the unusual scaling (60):

Theorem 7. *Suppose $d = 1$ and $p(1) = p(-1) = \frac{1}{2}$. Then $X(t)$ obeys the central limit theorem*

$$\frac{X(t)}{t^{1/4}} \Rightarrow N(0, \sqrt{2/\pi}). \quad \textbf{[11]}$$

In essentially all other cases, $X(t)$ is asymptotically Gaussian, but with a variance that is of order t rather than $\sqrt{t}$ (61–63). The proof of **11** is based on Eq. **10** as well. A key point is that the variance of the sum of negatively correlated Bernoulli random variables is at most equal to its mean.

The two applications above use only the weak form Eq. **10** of negative association that has been known since 1974. Here is an application of the more elaborate version proved in ref. 47 only recently. Suppose $d = 1$, and that initially all negative sites are occupied and all positive sites are vacant. Let $W(t)$ be the number of particles that are to the right of the origin at time t:

$$W(t) = \sum_{x>0} \eta_t(x).$$

By Theorem 5, for each $t > 0$, the summands above are negatively correlated, and there are independent Bernoulli random variables $\zeta_t(x)$ so that $W(t)$ has the same distribution as

$$\sum_{x>0} \zeta_t(x).$$

This makes it possible to apply classical central limit theorems to the sum directly, once one proves that $\mathrm{Var}[W(t)] \to \infty$ as $t \to \infty$. This fact is intuitively obvious, but is not particularly easy to prove. The difficulty comes from the fact that in the expression

$$\mathrm{Var}[W(t)] = \sum_{x,y>0} \mathrm{Cov}[\eta_t(x), \eta_t(y)],$$

the summands corresponding to $x = y$ are positive, whereas those corresponding to $x \neq y$ are negative, and may cancel the positive contributions and lead to a bounded variance.

The proof that $\mathrm{Var}[W(t)] \to \infty$ is again based on comparisons between finite interacting systems and the corresponding independent systems. Here is the result proved in ref. 64:

Theroem 8. *If $\sum_x x^2 p(x) < \infty$, then*

$$\frac{W(t) - EW(t)}{\sqrt{\mathrm{Var}[W(t)]}} \Rightarrow N(0,1), \quad \textbf{[12]}$$

with both the mean and the variance of $W(t)$ being of order $\sqrt{t}$.

The central limit theorem **12** has been extended to some choices of $p(\cdot)$ with infinite variance in ref. 65.

Discussion

In this paper, I have described some of the important results from the area of probability theory that is known as interacting particle systems—an area with motivations from, and connections to, a number of the sciences. Among the various techniques that have been important in the analysis of models in this area are the following:

(a) Coupling, in which two or more copies of the process are defined on the same probability space. This leads to conclusions about one of the processes based on known properties of the others.
(b) Duality, in which algebraic relations between two processes are exploited. Examples are given by Eqs. **2** and **9**. Duality is used in the proofs of Theorems 1 and 2.
(c) Renormalization, in which finite boxes in Z^d are regarded as individual sites. This is a key tool in the proof of the extinction of the critical contact process, for example.
(d) The correlation inequalities discussed here.

I have focussed on the latter technique in this paper for a number of reasons, including (*i*) the fact that major progress has been made in the past 2 years in the case of negative association and (*ii*) my own involvement in the proof and use of both positive and negative correlations in interacting particle systems in recent years.

In various ways, correlation inequalities often allow one to treat dependent random variables as if they were independent and therefore to apply classical results on independent random variables to obtain results in situations in which dependence occurs. Applications of this technique that we have discussed here include the following:

(a) the proof of Theorem 2 and its extensions on stationary distributions for the symmetric exclusion process,
(b) extinction of the critical contact process,
(c) the existence of percolation for the nontrivial stationary distribution of the contact process, and
(d) central limit theorems for a tagged particle in the exclusion process, and for the number of particles in large boxes for several models.

Even in the proofs of the central limit theorems, the way in which the correlation inequalities are used varies from case to case. This is a very versatile tool.

ACKNOWLEDGMENTS. I appreciate helpful comments by R. Durrett and C. Newman on an early version of this paper.

1. Glauber RJ (1963) Time-dependent statistics of the Ising model. *J Math Phys* 4:294–307.
2. Grassberger P, de la Torre A (1979) Reggeon field theory (Schlögl's first model) on a lattice: Monte Carlo calculations of critical behaviour. *Ann Phys* 122:373–396.
3. Kimura M (1953) "Stepping stone" model of population. *Annu Rep Natl Inst Genet (Japan)* 3:62–63.
4. Williams T, Bjerknes R (1972) Stochastic model for abnormal clone spread through epithelial basal layer. *Nature* 236:19–21.
5. Komarova NL (2006) Spatial stochastic models for cancer initiation and progression. *Bull Math Biol* 68:1573–1599.
6. Durrett R, Neuhauser C (1997) Coexistence results for some competition models. *Ann Appl Probab* 7:10–45.
7. Liggett TM, Schinazi RB, Schweinsberg J (2008) A contact process with mutations on a tree. *Stoch Proc Appl* 118:319–332.
8. MacDonald CT, Gibbs JH, Pipkin AC (1968) Kinetics of biopolymerization on nucleic acid templates. *Biopolymers* 6:1–25.
9. Harris TE (1974) Contact interactions on a lattice. *Ann Probab* 2:969–988.
10. Mollison D, ed. (1995) *Epidemic Models: Their Structure and Relation to Data* (Cambridge Univ Press, Cambridge, UK).
11. Krone SM, Neuhauser C (2000) A spatial model of range-dependent succession. *J Appl Probab* 37:1044–1060.
12. Satake A, Leslie HM, Iwasa Y, Levin SA (2007) Coupled ecological-social dynamics in a forested landscape: Spatial interactions and information flow. *J Theor Biol* 246:695–707.
13. Rodríguez-Iturbe I, Rinaldo A (1997) *Fractal River Basins: Chance and Self-Organization* (Cambridge Univ Press, Cambridge, UK).
14. Nowak MA, Bonhoeffer S, May RM (1994) More spatial games. *Int J Bifurcat Chaos* 4:33–56.
15. Liggett TM, Rolles SWW (2004) An infinite stochastic model of social network formation. *Stoch Proc Appl* 113:65–80.
16. Durrett R, Levin SA (2005) Can stable social groups be maintained by homophilous imitation alone? *J Econ Behav Organ* 57:267–278.
17. Topa G (2001) Social interactions, local spillovers and unemployment. *Rev Econ Studies* 68:261–295.
18. Lee H-W, Popkov V, Kim D (1997) Two-way traffic flow: Exactly solvable model of traffic jam. *J Phys A* 30:8497–8513.
19. Karimipour V (1999) Multispecies asymmetric simple exclusion process and its relation to traffic flow. *Phys Rev E* 59:205–212.
20. Helbing D (2001) Traffic and related self-driven many-particle systems. *Rev Mod Phys* 73:1067–1141.
21. Dobrushin RL (1971) Markov processes with a large number of locally interacting components: Existence of a limit process and its ergodicity. *Probl Inf Transm* 7:149–164.
22. Spitzer F (1970) Interaction of Markov processes. *Adv Math* 5:246–290.
23. Liggett TM (1985) *Interacting Particle Systems* (Springer, New York).
24. Liggett TM (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* (Springer, Berlin).
25. Durrett R (2010) Some features of the spread of epidemics and information on a random graph. *Proc Natl Acad Sci USA* 107:4491–4498.
26. Holley R, Liggett TM (1975) Ergodic theorems for weakly interacting systems and the voter model. *Ann Probab* 3:643–663.
27. Clifford P, Sudbury A (1973) A model for spatial conflict. *Biometrika* 60:581–588.
28. Bramson M, Griffeath D (1980) Asymptotics for interacting particle systems on Z^d. *Z Wahrsch Verw Gebiete* 53:183–196.
29. Bramson M, Griffeath D (1980) On the Williams-Bjerknes tumour growth model. II. *Math Proc Cambridge Philos Soc* 88:339–357.
30. Holley R, Liggett TM (1978) The survival of contact processes. *Ann Probab* 6:198–206.
31. Durrett R, Liu X (1988) The contact process on a finite set. *Ann Probab* 16:1158–1173.
32. Holley R, Stroock D (1977) In one and two dimensions, every stationary measure for a stochastic Ising model is a Gibbs state. *Commun Math Phys* 55:37–45.
33. Winkler G (2003) *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods a Mathematical Introduction* (Springer, Berlin), 2nd Ed.
34. Keith JM (2006) Segmenting eukaryotic genomes with the generalized Gibbs sampler. *J Comput Biol* 13:1369–1383.
35. Liggett TM (1973) A characterization of the stationary measures for an infinite particle system with interactions. *Trans Amer Math Soc* 179:433–453.
36. Spitzer F (1974) Recurrent random walk of an infinite particle system. *Trans Amer Math Soc* 198:191–199.
37. Bramson M, Liggett TM, Mountford T (2002) Characterization of stationary measures for one dimensional exclusion processes. *Ann Probab* 30:1529–1575.
38. Liggett TM (1975) Ergodic theorems for the asymmetric simple exclusion process. *Trans Amer Math Soc* 213:237–261.
39. Andjel ED, Bramson MD, Liggett TM (1988) Shocks in the asymmetric exclusion process. *Probab Theory Relat Fields* 78:231–247.
40. Derrida B, Lebowitz JL (1998) Exact large deviation function in the asymmetric exclusion process. *Phys Rev Lett* 80:209–213.
41. Caputo P, Liggett TM, Richthammer T (2010) Proof of Aldous' spectral gap conjecture. *J Am Math Soc* 23:831–851.
42. Fortuin CM, Kasteleyn PW, Ginibre J (1971) Correlation inequalities on some partially ordered sets. *Commun Math Phys* 22:89–103.
43. Harris TE (1960) A lower bound for the critical probability in a certain percolation process. *Proc Cambridge Philos Soc* 56:13–20.
44. Harris TE (1977) A correlation inequality for Markov processes in partially ordered state spaces. *Ann Probab* 5:451–454.
45. Liggett TM (1974) A characterization of the stationary measures for an infinite particle system with interactions II. *Trans Amer Math Soc* 198:201–213.
46. Andjel ED (1988) A correlation inequality for the symmetric exclusion process. *Ann Probab* 16:717–721.
47. Borcea J, Brändén P, Liggett TM (2009) Negative dependence and the geometry of polynomials. *J Am Math Soc* 22:521–567.
48. Liggett TM (2002) Negative correlations and particle systems. *Markov Process Related Fields* 8:547–564.
49. Barsky DJ, Grimmett GR, Newman CM (1991) Percolation in half-spaces: Equality of critical densities and continuity of the percolation probability. *Probab Theory Relat Fields* 90:111–148.
50. Bezuidenhout C, Grimmett G (1990) The critical contact process dies out. *Ann Probab* 18:1462–1482.
51. Liggett TM (1994) Survival and coexistence in interacting particle systems. *Probability and Phase Transition*, ed G Grimmett (Kluwer, Dordrecht), pp 209–226.
52. Cox JT, Grimmett G (1984) Central limit theorems for associated random variables and the percolation model. *Ann Probab* 12:514–528.
53. Newman CM (1980) Normal fluctuations and the FKG inequalities. *Comm Math Phys* 74:119–128.
54. van den Berg J, Häggström O, Kahn J (2006) Proof of a conjecture of N. Konno for the 1D contact process. *Dynamics and Stochastics. Festschrift in Honor of M. S. Keane*, eds D Denteneer, F den Hollander, and E Verbitskiy (Inst Math Stat, Beachwood, OH), pp 16–23.
55. van den Berg J, Häggström O, Kahn J (2006) Some conditional correlation inequalities for percolation and related processes. *Random Struct Algor* 29:417–435.
56. Liggett TM (2006) Conditional association and spin systems. *ALEA* 1:1–19.
57. Liggett TM, Steif JE (2006) Stochastic domination: The contact process, Ising models and FKG measures. *Ann Inst H Poincaré Probab Statist* 42:223–243.
58. Kéfi S, et al. (2007) Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature* 449:213–217.
59. van den Berg J Sharpness of the percolation transition in the two-dimensional contact process. *Ann Appl Probab*, in press.
60. Arratia R (1983) The motion of a tagged particle in the simple symmetric exclusion system on Z. *Ann Probab* 11:362–373.
61. Kipnis C, Varadhan SRS (1986) Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun Math Phys* 104:1–19.
62. Varadhan SRS (1995) Self-diffusion of a tagged particle in equilibrium for asymmetric mean zero random walk in simple exclusion. *Ann Inst Henri Poincaré* 31:273–285.
63. Sethuraman S, Varadhan SRS, Yau H-T (2000) Diffusive limit of a tagged particle in asymmetric simple exclusion processes. *Comm Pure Appl Math* 53:972–1006.
64. Liggett TM (2009) Distributional limits for the symmetric exclusion process. *Stoch Proc Appl* 119:1–15.
65. Vandenberg-Rodes A (2010) A limit theorem for particle current in the symmetric exclusion process. *Elect C Probab* 15:240–252.