NEGATIVE CORRELATIONS AND PARTICLE SYSTEMS

Thomas M. Liggett

University of California, Los Angeles

September 4, 2002

Abstract. We consider the symmetric exclusion process and systems of independent Markov chains. For each of these, we prove that certain classes of distributions with negative dependence are preserved by the evolution. We also show by example that the class of negatively associated distributions is not preserved by the symmetric exclusion process.

1. Introduction

Harris (1977) proved that the class of probability measures with positive correlations is preserved (i.e., mapped into itself) by fairly general types of continuous time Markov processes. The purpose of this paper is to investigate analogues of his theorem for negative correlations.

We begin by stating an important special case of Harris' theorem. Let S be a finite or countable set, and put $X = \{0,1\}^S$ or $X = \{0,1,...\}^S$. Then X is a partially ordered set with partial order given by

$$\eta \leq \zeta \quad \text{iff} \quad \eta(x) \leq \zeta(x) \text{ for all } x \in S.$$

A function f on X is increasing if $\eta \leq \zeta$ implies $f(\eta) \leq f(\zeta)$. A continuous time Markov process η_t on X is said to be monotone if $E^\eta f(\eta_t)$ is an increasing function of η for every $t \geq 0$ and every increasing function f. A probability measure μ on X is said to have positive correlations if

$$\int fg\,d\mu \geq \int f\,d\mu \int g\,d\mu \tag{1.1}$$

1991 *Mathematics Subject Classification.* 60K35, 60E15.

Key words and phrases. Correlation inequalities, symmetric exclusion process, negative association, independent particle systems.

Research supported in part by NSF Grant DMS 0070465.

for all bounded increasing functions f, g. Harris proved that if η_t is a monotone Markov process on X and all transitions for η_t are between comparable states (i.e., if a single transition takes η to ζ, then either $\eta \leq \zeta$ or $\zeta \leq \eta$), then the distribution at time t has positive correlations whenever the initial distribution has positive correlations. (See Cox (1984) and Liggett (1985) for two other versions of the proof.) In particular, this result applies to attractive spin systems. An immediate consequence is that if η_0 is deterministic and η_t converges weakly to μ as $t \to \infty$, then μ has positive correlations. Such results have been useful in the study of the contact process, for example.

In trying to find situations in which negative correlations are preserved by a Markov process, we run into two problems:

(a) deciding what definition of negative correlations to use,

and

(b) finding classes of Markov chains that preserve them.

There are certainly many possible definitions of negative dependence. In fact, recently Pemantle (2000) has made a systematic study of measures with various negative dependence properties, and focussed on several definitions, giving implications among them. We will consider other definitions here that appear to be better suited to our problem of preservation of this property by particle systems. One fact that Pemantle's paper makes quite clear is that issues related to negative correlations are much more subtle than the corresponding issues for positive correlations. This is also clear in the present paper. We have been unable to find a simply stated, general analogue to Harris' theorem for negative dependence.

The processes we will consider here are the symmetric exclusion process and systems of independent Markov chains. To define the exclusion process, let $p(x, y)$ be the transition probabilities for a discrete time Markov chain on S. A configuration $\eta \in \{0, 1\}^S$ is interpreted in terms of particle occupancy — x is occupied or vacant according to whether $\eta(x) = 1$ or $\eta(x) = 0$. Particles wait independent unit exponential times. A particle at x then attempts to move to y with probability $p(x, y)$. The transition occurs if y is vacant, but not otherwise. The process is said to be symmetric if $p(x, y) = p(y, x)$ for all $x, y \in S$. For more on this process, see Chapter VIII of Liggett (1985) or Part III of Liggett (1999). One should expect that the symmetric exclusion process has some form of negative correlation propery for the following reason: particles are neither created nor destroyed, so if there are many particles in one part of S, there will tend to be fewer particles on other parts of S. Other known results support this idea. For example, every extremal stationary distribution μ has the property that

$$\mu\{\eta : \eta(x) = \eta(y) = 1\} \leq \mu\{\eta : \eta(x) = 1\}\mu\{\eta : \eta(y) = 1\}$$

for $x \neq y$. (See Lemma 1.36 of Chapter VIII of Liggett(1985).)

The process of independent Markov chains is simply what the name suggests: $\eta(x)$ is the number of particles at $x \in S$. Particles move on S independently according to continuous time Markov chains with transition probabilities $p_t(x, y)$. The resulting process η_t is well defined on the set of configurations $\eta \in \{0, 1, ...\}^S$ that satisfy

$$\sum_{x \in S} \eta(x)p_t(x, y) < \infty$$

for each $y \in S$ and $t > 0$.

The symmetry assumption for the exclusion process is needed because asymmetric systems essentially never preserve any type of negative dependence. A simple example that illustrates the problem is given by $S = \{1, 2\}$ with $p(1,1) = p(2,1) = 1, p(1,2) = p(2,2) = 0$. If the initial distribution is the symmetric product measure given by $\mu(\{\eta\}) = \frac{1}{4}$ for each η, it should have negative correlations according to any reasonable definition. However, the limiting distribution ν as $t \to \infty$ is given by $\nu(\{11\}) = \frac{1}{4}, \nu(\{10\}) = \frac{1}{2}, \nu(\{00\}) = \frac{1}{4}$, so that

$$\int \eta(1)\eta(2)d\nu = \frac{1}{4} > \frac{3}{16} = \int \eta(1)d\nu \int \eta(2)d\nu.$$

Here is a more interesting example: Take $S = Z^1$ and $p(x, x+1) = p, p(x, x-1) = 1-p$ for all x. Take the initial distribution μ to be the product measure with marginals

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

Andjel, Bramson and Liggett (1988) proved that if $p > \frac{1}{2}$ and $0 < \lambda < \rho, \lambda + \rho = 1$, then the distribution μ_t at time t satisfies

$$\lim_{t\to\infty} \mu_t = \frac{1}{2}\nu_\lambda + \frac{1}{2}\nu_\rho,$$

where ν_ρ is the product measure with density ρ. This result was later extended in a series of papers by Ferrari and Fontes. See Part III of Liggett (1999) for details. In this case, even though μ is a product measure, the limiting distribution ν satifies

$$\int \eta(1)\eta(2)d\nu = \frac{1}{2}\lambda^2 + \frac{1}{2}\rho^2 > \left(\frac{\lambda+\rho}{2}\right)^2 = \int \eta(1)d\nu \int \eta(2)d\nu.$$

If $p = \frac{1}{2}$, the limiting distribution is the product measure $\nu_{\frac{1}{2}}$, however, so this problem does not arise in the symmetric case. The strict positive correlations in the limiting distribution that appear in the case $p > \frac{1}{2}$ can be understood as follows: For large t there is a random location X_t (the *shock*) so that the distribution is approximately ν_λ to the left of X_t and ν_ρ to the right of X_t. Therefore, knowing that there are many particles in one region suggests that X_t is to the left of that region, and therefore that there will be many particles in nearby regions.

Some negative correlation results for the symmetric exclusion process have been proved and found to be quite useful. In the 1970's, Liggett proved that for distinct $x_1, ..., x_n \in S$,

$$P^\eta(\eta_t(x_i) = 1 \text{ for } 1 \leq i \leq n) \leq \prod_{i=1}^{n} P^\eta(\eta_t(x_i) = 1).$$

(This is a consequence of Proposition 1.7 in Chapter VIII of Liggett (1985) and duality.) This result was extended by Andjel (1988) in the following form: For disjoint $A, B \subset S$,

$$P^\eta(\eta_t \equiv 1 \text{ on } A \cup B) \leq P^\eta(\eta_t \equiv 1 \text{ on } A)P^\eta(\eta_t \equiv 1 \text{ on } B). \tag{1.2}$$

This and other extensions have played a role in proofs of many results; see for example Arratia (1985), Andjel (1988), Ferrari, Galves and Landim (2000), and Asselah and Ferrari (2002). Note though that (1.2) is not in the form of a statement about the preservation of a negative dependence property by the symmetric exclusion process. In particular, it says nothing about the distribution of the process at time t if the initial distribution is not deterministic.

In Section 2, we consider the case of independent Markov chains, and prove that two concepts of negative dependence are preserved by this process. A consequence of these results is that if the initial distribution is deterministic, then the distribution μ_t at time t is negatively associated, in the sense that

$$\int fg d\mu_t \leq \int f d\mu_t \int g d\mu_t \tag{1.3}$$

for all bounded increasing functions f, g on $\{0, 1, ...\}^S$ that depend on disjoint sets of coordinates. Of course this last restriction is essential – if $f = g$ and (1.3) holds, then $f =$ constant a.s. (μ_t). For more on negative association, see Joag-Dev and Proschan (1983) and Pemantle (2000), for example.

A consequence of our results for random initial configurations is that

$$\begin{aligned} E\eta_t(x)\eta_t(y) &\leq E\eta_t(x)E\eta_t(y) \quad \text{for } x \neq y \\ E\eta_t(x)[\eta_t(x) - 1] &\leq [E\eta_t(x)]^2 \quad \text{for all } x, \end{aligned} \tag{1.4}$$

is satisfied for all $t > 0$ if it is satisfied at $t = 0$.

Section 3 considers symmetric exclusion processes. Our first result is a strengthening of Andjel's theorem (1.2). Rather than considering only deterministic initial configurations, we prove that if the initial distribution μ satisfies

$$\mu\{\eta : \eta \equiv 1 \text{ on } A \cup B\} \leq \mu\{\eta : \eta \equiv 1 \text{ on } A\}\mu\{\eta : \eta \equiv 1 \text{ on } B\}$$

for all disjoint A, B, then the distribution at time t also has this property. This statement is more in the spirit of Harris' result on the preservation of positive correlations. In addition, the proof is a bit easier than Andjel's – the form of the statement facilitates the proof.

In Section 3, we also give an example to show that the property of negative association is *not* preserved by the symmetric exclusion process. That leads to a statement and proof of a result that another (less natural) property is preserved. This property implies that

$$\mu\{\eta : \eta \equiv 1 \text{ on } A, \eta \equiv 0 \text{ on } B\} \geq \mu\{\eta : \eta \equiv 1 \text{ on } A\}\mu\{\eta : \eta \equiv 0 \text{ on } B\}$$

for disjoint A, B. Unfortunately, it is not satisfied by deterministic measures.

This discussion suggests an open problem: Is it true that the distribution of a symmetric exclusion process at time t with deterministic initial configuration is negatively associated? In particular, is it the case that

$$P^\eta(\eta_t \equiv 1 \text{ on } A, \eta_t \equiv 0 \text{ on } B) \geq P^\eta(\eta_t \equiv 1 \text{ on } A)P^\eta(\eta_t \equiv 0 \text{ on } B) \tag{1.5}$$

for disjoint A, B? This is a consequence of (1.2) if A or B is a singleton, but not otherwise. The proof of (1.2) does not work to prove (1.5), and in fact, (1.5) appears to be much more subtle than (1.2).

There is a dual form of (1.5) which one can try to prove. Suppose (A_t, B_t) are coupled copies of the finite exclusion process satisfying $A_t \cap B_t = \emptyset$. The coupling is the one induced by the stirring construction: at rate $p(x, y)$, if A_t contains exactly one of $\{x, y\}$, then A_t is replaced by $A_t \triangle \{x, y\}$, while otherwise A_t remains unchanged; B_t evolves in the same way. Then (1.5) is equivalent to

$$P^{A,B}(A_t \subset D, B_t \subset D^c) \geq P^A(A_t \subset D)P^B(B_t \subset D^c),$$

where $D = \{x : \eta(x) = 1\}$.

At the end of Section 3, we prove a result which provides some evidence for (1.5): If the initial distribution of the symmetric exclusion process is a product measure, then the limiting distribution is negatively associated. Additional evidence is given by Corollary 2.6, which gives answers our question affirmatively for independent Markov chains.

2. Independent Markov Chains

In this section, we treat the problem of preservation of negative correlations for the process of independent Markov chains. For $x \in S$, let e_x be the configuration defined by

$$e_x(y) = \begin{cases} 1 & \text{if } y = x, \\ 0 & \text{if } y \neq x. \end{cases}$$

For our first result, let $\mathcal{M}$ be the set of probability measures on $\{0, 1, ...\}^S$ that are the distributions of random configurations η that can be represented in the following way:

$$\eta = \sum_{i=1}^{\infty} e_{X_i}\Big|_S \tag{2.1}$$

where $X_1, X_2, ...$ are independent random variables with values in $S \cup \{\infty\}$. Of course, every deterministic measure is in $\mathcal{M}$.

Theorem 2.2. *If $\mu \in \mathcal{M}$, then $\mu_t \in \mathcal{M}$, where μ_t is the distribution at time t of the independent Markov chain process with initial distribution μ.*

Proof. Write η with distribution μ as in (2.1). Let Y_i be the location at time t of a Markov chain starting at X_i. These Markov chains are taken to be independent of each other and of the X_i's. Then the Y_i's are independent, and μ_t is the distribution of

$$\sum_{i=1}^{\infty} e_{Y_i}\Big|_S$$

Therefore $\mu_t \in \mathcal{M}$.

Next we show that every measure in $\mathcal{M}$ is negatively associated, so that $\mathcal{M}$ is a reasonable class of measures to consider in showing that negative dependence is preserved by systems of independent Markov chains. In the case that the X_i's are also identically distributed, this result was proved by Dubhashi and Ranjan (1998) – see also pages 1380–81 of Pemantle (2000). An interesting feature of our proof is that this *negative* association statement is a consequence of the *positive* correlations of product measures.

Theorem 2.3. *Suppose $\mu \in \mathcal{M}$, and let f and g be increasing functions on $\{0,1,...\}^S$. If for some $T \subset S$, f is just a function of $\{\eta(x), x \in T\}$ and g is just a function of $\{\eta(x), x \notin T\}$, then*

$$\int fg\,d\mu \le \int f\,d\mu \int g\,d\mu. \tag{2.4}$$

Proof. Let $X_1, X_2, ...$ be the random variables that realize representation (2.1) for μ, and let $Y_1, Y_2, ...$ be a sequence independent of the X_i's with the same joint distributions. Then (2.4) is equivalent to

$$E\Big[f\Big(\sum_i e_{X_i}\Big) - f\Big(\sum_i e_{Y_i}\Big)\Big]\Big[g\Big(\sum_i e_{X_i}\Big) - g\Big(\sum_i e_{Y_i}\Big)\Big] \le 0. \tag{2.5}$$

(Here and below, we omit the restriction $|_S$ statement to simplify notation.) Now, let A and B be random subsets of $\{1,2,...\}$ defined by

$$A = \{i : X_i \in T\} \quad \text{and} \quad B = \{i : Y_i \in T\}.$$

They are of course independent and identically distributed, and their distribution is a product measure, in the sense that the events $\{i \in A\} = \{X_i \in T\}$ are independent. Define functions F and G on the subsets of $\{1,2,...\}$ by the following conditional expectations:

$$F(C) = E\Big[f\Big(\sum_{i\in C} e_{X_i}\Big)\Big| X_i \in T \text{ for all } i \in C\Big]$$

and

$$G(C) = E\Big[g\Big(\sum_{i\notin C} e_{X_i}\Big)\Big| X_i \notin T \text{ for all } i \notin C\Big].$$

Write the expression on the left of (2.5) as an iterated expectation, conditioning with respect to A and B. Since f depends only on coordinates in T and g depends only on coordinates not in T,

$$\begin{aligned}
&E\Big\{\Big[f\Big(\sum_i e_{X_i}\Big) - f\Big(\sum_i e_{Y_i}\Big)\Big]\Big[g\Big(\sum_i e_{X_i}\Big) - g\Big(\sum_i e_{Y_i}\Big)\Big]\Big| A, B\Big\}\\
&\quad = E\Big\{\Big[f\Big(\sum_{i\in A} e_{X_i}\Big) - f\Big(\sum_{i\in B} e_{Y_i}\Big)\Big]\Big[g\Big(\sum_{i\notin A} e_{X_i}\Big) - g\Big(\sum_{i\notin B} e_{Y_i}\Big)\Big]\Big| A, B\Big\}\\
&\quad = \big[F(A) - F(B)\big]\big[G(A) - G(B)\big],
\end{aligned}$$

so the left side of (2.5) is just

$$E\big[F(A) - F(B)\big]\big[G(A) - G(B)\big] = 2\big[EF(A)G(A) - EF(A)EG(A)\big].$$

This is ≤ 0, since the distribution of A has positive correlations by Corollary B18 of Liggett (1999), while F is increasing, and G is decreasing. To check the monotonicity of F, for example, let Z_i be independent random variables, with Z_i having the conditional distribution of X_i, given the event $\{X_i \in T\}$. Then

$$F(C) = Ef\left(\sum_{i\in C} e_{Z_i}\right).$$

But

$$C \subset D \quad \text{implies} \quad \sum_{i\in C} e_{Z_i} \leq \sum_{i\in D} e_{Z_i},$$

so $F(C) \leq F(D)$, since f is increasing.

Remark. One might reasonably object to our choice of $\mathcal{M}$ in the present context, since it does not include all product measures. However, we must exclude many product measures if our class is to be preserved by independent Markov chains and be negatively associated at later times. To see this, suppose $S = \{x, y\}$, and consider a particle system in which particles move from x to y and from y to x independently at rate 1. Suppose the initial distribution assigns a random number M of particles to x and none to y. The limiting distribution of $(\eta(x), \eta(y))$ as $t \to \infty$ is that of

$$\left(\sum_{i=1}^{M} X_i, M - \sum_{i=1}^{M} X_i\right),$$

where $\{X_i, i \geq 1\}$ are independent Bernoulli random variables with parameter $\frac{1}{2}$. If this distribution were negatively associated, then for each increasing function f on $[0, \infty)$ we would have

$$Ef\left(\sum_{i=1}^{M} X_i\right) f\left(M - \sum_{i=1}^{M} X_i\right) \leq Ef\left(\sum_{i=1}^{M} X_i\right) Ef\left(M - \sum_{i=1}^{M} X_i\right).$$

Replacing M by $2Mn$, rescaling f by a factor of n and passing to the limit leads to

$$Ef^2(M) \leq \left[Ef(M)\right]^2.$$

This would imply that M is constant. So, some constraints must be placed on the distribution of M.

Theorems 2.2 and 2.3 combine to give the following statement:

Corollary 2.6. *If η_t is a system of independent Markov chains and the initial configuration is deterministic, then the distribution of η_t is negatively associated.*

We turn now to the preservation of another type of negative dependence for independent Markov chains – one based on factorial moments. For $\eta, \zeta \in \{0, 1, ...\}^S$, let

$$H(\eta, \zeta) = \prod_{x\in S} \binom{\eta(x)}{\zeta(x)} \zeta(x)!,$$

where as usual, the binomial coefficient appearing above is defined to be zero if $\zeta(x) > \eta(x)$. Note that if η is random and ζ is deterministic, then $EH(\eta, \zeta)$ is the joint ζ^{th} factorial moment of η. For $i, j \geq 0$, let $\mathcal{M}_{i,j}$ be the set of all probability measures on $\{0, 1, ...\}^S$ so that

$$\int H(\eta, \zeta + \gamma)\mu(d\eta) \leq \int H(\eta, \zeta)\mu(d\eta) \int H(\eta, \gamma)\mu(d\eta) \tag{2.7}$$

for all ζ and γ such that $\sum_x \zeta(x) = i$ and $\sum_x \gamma(x) = j$. For example, if μ is a product measure, then $\mu \in \mathcal{M}_{1,1}$ if and only if the marginal distributions have the property that the variance is at most the mean. A product measure is in $\cap_{i,j}\mathcal{M}_{i,j}$ if and only if the factorial moments of the marginal distributions are sub-multiplicative.

For the next result, let $q(x, y)$ be the rate at which a particle in the system moves from x to y (with $q(x, x) = 0$ for all x), and assume

$$\sup_x \sum_y q(x, y) < \infty \quad \text{and} \quad \sup_y \sum_x q(x, y) < \infty.$$

Theorem 2.8. *Fix nonnegative integers i, j. Suppose $\mu \in \mathcal{M}_{i,j}$, and let μ_t be the distribution at time t of an independent Markov chain process with initial distribution μ. Then $\mu_t \in \mathcal{M}_{i,j}$ for every $t > 0$.*

Proof. It suffices to prove this in case μ concentrates on configurations with a bounded number particles. Once this is done, one can write a general $\mu \in \mathcal{M}_{i,j}$ as a monotone limit of measures on finite configurations, and pass to the limit in (2.7) using monotone convergence. Of course, the resulting quantities may be infinite, but that is not excluded in the definition of $\mathcal{M}_{i,j}$.

Let η_t be the process of independent Markov chains with initial distribution μ. Then, for determinstic ζ,

$$\begin{aligned}
\frac{d}{dt}EH(\eta_t, \zeta) &= \sum_{x,y} q(x, y)E\eta_t(x)\big[H(\eta_t - \epsilon_x + \epsilon_y, \zeta) - H(\eta_t, \zeta)\big] \\
&= \sum_{x,y} q(x, y)EH(\eta_t, \zeta)\frac{\eta_t(x)\zeta(y) - \zeta(x)\eta_t(y) - \zeta(x)}{\eta_t(y) - \zeta(y) + 1} \\
&= \sum_{x,y} q(x, y)\big[\zeta(y)EH(\eta_t, \zeta + \epsilon_x - \epsilon_y) - \zeta(x)EH(\eta_t, \zeta)\big].
\end{aligned} \tag{2.9}$$

Using (2.9), take deterministic ζ and γ, and compute

$$\begin{aligned}
&\frac{d}{dt}\big[EH(\eta_t, \zeta)EH(\eta_t, \gamma) - EH(\eta_t, \zeta + \gamma)\big] = \\
&\sum_{x,y} q(x, y)\zeta(y)\big[EH(\eta_t, \zeta + \epsilon_x - \epsilon_y)EH(\eta_t, \gamma) - EH(\eta_t, \zeta + \gamma + \epsilon_x - \epsilon_y)\big] \\
&+ \sum_{x,y} q(x, y)\gamma(y)\big[EH(\eta_t, \zeta)EH(\eta_t, \gamma + \epsilon_x - \epsilon_y) - EH(\eta_t, \zeta + \gamma + \epsilon_x - \epsilon_y)\big] \\
&- \big[EH(\eta_t, \zeta)EH(\eta_t, \gamma) - EH(\eta_t, \zeta + \gamma)\big]\sum_{x,y} q(x, y)[\zeta(x) + \gamma(x)].
\end{aligned} \tag{2.10}$$

Now write the index (ζ,γ) with $\sum_x \zeta(x)=i$ and $\sum_x \gamma(x)=j$ that appears in this system of differential equations as k, and let

$$f_k(t)=EH(\eta_t,\zeta)EH(\eta_t,\gamma)-EH(\eta_t,\zeta+\gamma).$$

Note that these are uniformly bounded by the simplifying assumption we made at the beginning of the proof. Then (2.10) can be written as

$$f_k'(t)=\sum_l a_{k,l}f_l(t)-b_kf_k(t), \tag{2.11}$$

where $a_{k,l}\ge 0$ and $b_k\ge 0$ for each k,l. Note that

$$b_k=\sum_{x,y}q(x,y)[\zeta(x)+\gamma(x)]\le (i+j)\sup_x\sum_y q(x,y) \tag{2.12}$$

and

$$\sum_l a_{k,l}=\sum_{x,y}q(x,y)[\zeta(y)+\gamma(y)]\le (i+j)\sup_y\sum_x q(x,y). \tag{2.13}$$

Since $\mu\in\mathcal{M}_{i,j}$, $f_k(0)\ge 0$ for every k. Since b_k and $\sum_l a_{k,l}$ are uniformly bounded in k, the unique bounded solution to the system (2.11) with nonnegative initial conditions is nonnegative for all $t>0$. Therefore $\mu_t\in\mathcal{M}_{i,j}$ for all $t>0$.

Remark. Equation (2.9) can be rewritten as

$$\frac{d}{dt}EH(\eta_t,\zeta)=\sum_{x,y}q(x,y)\zeta(y)\left[EH(\eta_t,\zeta+\epsilon_x-\epsilon_y)-EH(\eta_t,\zeta)\right]$$
$$+\sum_{x,y}q(x,y)[\zeta(y)-\zeta(x)]EH(\eta_t,\zeta).$$

If q is symmetric, the last term above is zero. In this case, we would conclude that η_t is self-dual with respect to the duality function H. (See Section II.3 of Liggett (1985).) We are not aware of this type of duality having been noticed earlier. We do not need this duality to prove Theorem 2.8, however, so there is no need to assume the symmetry of q.

In view of Theorems 2.2 and 2.8, it is of interest to compare the classes $\mathcal{M}$ appearing in the first of these and $\cap_{i,j}\mathcal{M}_{i,j}$ in the second. We conclude this section with the result.

Proposition 2.14. *(a)* $\mathcal{M}\subset\cap_{i,j}\mathcal{M}_{i,j}$.
(b) $\mathcal{M}\ne\cap_{i,j}\mathcal{M}_{i,j}$.

Proof. For part (a), take η with distribution $\mu\in\mathcal{M}$, and write

$$\eta(x)=\sum_k 1_{\{X_k=x\}}.$$

Then for $\eta(x) \geq m$,

$$\begin{aligned}\binom{\eta(x)}{m} m! =& \eta(x)[\eta(x)-1]\cdots[\eta(x)-m+1] \\ =& \sum_{k_1} 1_{\{X_{k_1}=x\}}\left[\sum_{k_2} 1_{\{X_{k_2}=x\}}-1\right]\cdots\left[\sum_{k_m} 1_{\{X_{k_m}=x\}}-m+1\right] \\ =& \sum_{k_1,\dots,k_m \text{ distinct}} 1_{\{X_{k_1}=x\}}\cdots 1_{\{X_{k_m}=x\}}.\end{aligned}$$

Therefore, $EH(\eta,\zeta)$ is a sum of positive terms, indexed by $(k_1,\dots,k_i)$, where $i=\sum_x \zeta(x)$, with the restriction that all the k_l are distinct. So, we see that

$$EH(\eta,\zeta+\gamma) \quad \text{and} \quad EH(\eta,\zeta)EH(\eta,\gamma)$$

are sums of the same positive terms. The only difference is that all the indexes in the first sum are distinct, while only the indexes in the second sum that correspond to ζ or to γ separately are required to be distinct. Therefore,

$$EH(\eta,\zeta+\gamma) \leq EH(\eta,\zeta)EH(\eta,\gamma).$$

For part (b), consider a random η with $\eta(u) \leq 2$ for some fixed u and $\eta(x)=0$ for all $x \neq u$. Let $P(\eta(u)=2)=\alpha$ and $P(\eta(u)=1)=\beta$. Then the distribution of η is in $\cap_{i,j}\mathcal{M}_{i,j}$ if and only if $2\alpha = E\eta(u)[\eta(u)-1] \leq [E\eta(u)]^2 = (2\alpha+\beta)^2$. On the other hand, the distribution is in $\mathcal{M}$ if and only if $\eta(u)$ can be written as the sum of two independent Bernoulli random variables with parameters p_1 and p_2 respectively. This would mean that $\alpha = p_1p_2$ and $\beta = p_1+p_2-2p_1p_2$. These equations can be solved for p_1,p_2 if and only if $(2\alpha+\beta)^2 \geq 4\alpha$. So, for example, if $\alpha=\frac{1}{2},\beta=\frac{1}{4}$, then the distribution of η is in $\cap_{i,j}\mathcal{M}_{i,j}$ but not in $\mathcal{M}$.

3. Symmetric Exclusion Processes

If μ is a probability measure on $\{0,1\}^S$ and $A \subset S$ is finite, define

$$\hat{\mu}(A) = \mu\{\eta : \eta \equiv 1 \text{ on } A\}.$$

In this section we will let $\mathcal{M}$ be the set of all probability measures on $\{0,1\}^S$ that satisfy the following property:

$$A \cap B = \emptyset \quad \text{implies} \quad \hat{\mu}(A\cup B) \leq \hat{\mu}(A)\hat{\mu}(B).$$

Since pointmasses are in $\mathcal{M}$, the following generalizes Andjel's result (1.2).

Theorem 3.1. *Suppose $\mu \in \mathcal{M}$ and μ_t is the distribution at time t for a symmetric exclusion process with initial distribution μ. Then $\mu_t \in \mathcal{M}$ for all $t \geq 0$.*

Proof. If $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$ are bounded generators of the Markov semigroups $S_1(t), S_2(t)$ and $S(t)$ respectively, the Trotter product formula gives

$$S(t)f = \lim_{t\to\infty} \left[S_1\left(\frac{t}{n}\right)S_2\left(\frac{t}{n}\right)\right]^n f.$$

Therefore, if $\mu \in \mathcal{M}$ implies $\mu S_i(t) \in \mathcal{M}$ for $i = 1, 2$ and $t \geq 0$, it follows that $\mu \in \mathcal{M}$ implies $\mu S(t) \in \mathcal{M}$. Combining this observation with the Trotter-Kurtz semigroup convergence theorem (Theorem 2.12 of Chapter I of Liggett (1985)), it follows that it suffices to prove the theorem in case the exclusion process has transitions only between u and v for some fixed $u, v \in S$. So, suppose the coordinates $\eta(u)$ and $\eta(v)$ are interchanged at rate 1. Then by duality,

$$\hat{\mu}_t(A) = E^A\hat{\mu}(A_t) = \alpha(t)\hat{\mu}(A) + [1-\alpha(t)]\hat{\mu}(A_{u,v}),$$

where $A_{u,v}$ is the set obtained from A by interchanging sites u and v, A_t is the exclusion process at time t, and $\alpha(t)$ is the probability that a Markov chain on $\{u, v\}$ with transition rate 1 in each direction will be at its initial site at time t. Thus we need to show that if $\mu \in \mathcal{M}$, it follows that

$$\begin{aligned}\alpha\hat{\mu}(A\cup B)+&(1-\alpha)\hat{\mu}((A\cup B)_{u,v}) \leq \\ &\left[\alpha\hat{\mu}(A) + (1-\alpha)\hat{\mu}(A_{u,v})\right]\left[\alpha\hat{\mu}(B) + (1-\alpha)\hat{\mu}(B_{u,v})\right]\end{aligned} \tag{3.2}$$

for $0 \leq \alpha \leq 1$ and $A \cap B = \emptyset$. This is immediate if $u, v \in A$ or $u, v \in B$ or $u, v \in (A\cup B)^c$, since then $A_{u,v} = A$, $B_{u,v} = B$, and $(A\cup B)_{u,v} = A\cup B$, so (3.2) reduces to $\hat{\mu}(A\cup B) \leq \hat{\mu}(A)\hat{\mu}(B)$. By symmetry, this leaves two cases: $u \in A, v \in B$, and $u \in A, v \in (A \cup B)^c$.

Case 1: Assume $u \in A, v \in B$. Then $A_{u,v} \cup B_{u,v} = (A \cup B)_{u,v} = A \cup B$. Therefore, since $\mu \in \mathcal{M}$, we have

$$\hat{\mu}(A\cup B) \leq \hat{\mu}(A)\hat{\mu}(B) \quad \text{and} \quad \hat{\mu}(A\cup B) \leq \hat{\mu}(A_{u,v})\hat{\mu}(B_{u,v}). \tag{3.3}$$

If $\hat{\mu}(A) = 0$ or $\hat{\mu}(A_{u,v}) = 0$ then $\hat{\mu}(A\cup B) = 0$, so (3.2) is automatically satisfied. If neither is zero, the right side of (3.2) is

$$\geq \left[\alpha\hat{\mu}(A) + (1-\alpha)\hat{\mu}(A_{u,v})\right]\left[\alpha\frac{\hat{\mu}(A\cup B)}{\hat{\mu}(A)} + (1-\alpha)\frac{\hat{\mu}(A\cup B)}{\hat{\mu}(A_{u,v})}\right]$$

by (3.3). Multiplying this out leads to

$$\hat{\mu}(A\cup B)\left[\alpha^2 + (1-\alpha)^2 + \alpha(1-\alpha)\left(\frac{\hat{\mu}(A)}{\hat{\mu}(A_{u,v})} + \frac{\hat{\mu}(A_{u,v})}{\hat{\mu}(A)}\right)\right] \geq \hat{\mu}(A\cup B),$$

since the sum of fractions above is ≥ 2. Since $(A\cup B)_{u,v} = A\cup B$, the proof is complete in this case.

Case 2: Assume $u \in A, v \in (A \cup B)^c$. Then $B_{u,v} = B$, so the right side of (3.2) becomes

$$\big[\alpha\hat{\mu}(A) + (1-\alpha)\hat{\mu}(A_{u,v})\big]\hat{\mu}(B),$$

which is $\geq$

$$\alpha\hat{\mu}(A \cup B) + (1-\alpha)\hat{\mu}((A \cup B)_{u,v})$$

as required, since $\mu \in \mathcal{M}$, $B \cap A_{u,v} = \emptyset$, and $(A \cup B)_{u,v} = A_{u,v} \cup B$.

Remark. It follows from the proof of Theorem 3.1 that for each k, l, the evolution of the symmetric exclusion process does preserve the property

$$|A| = k, |B| = l, A \cap B = \emptyset \quad \text{imply} \quad \hat{\mu}(A \cup B) \leq \hat{\mu}(A)\hat{\mu}(B).$$

We now turn to a negative result concerning the preservation of negative association.

Theorem 3.5. *There exists a negatively associated measure μ and a symmetric exclusion process so that μ_t is not negatively associated for some $t > 0$.*

Proof. Take $S = \{1, 2, 3, 4\}$, and consider the exclusion process in which particles move between states 2 and 3 at rate 1. Let μ be the measure that assigns mass

$$\begin{array}{rll}
0 & \text{to configurations} & 1111, 0000, \\
v + w & \text{to configurations} & 1110, 0001, 1101, 0010, 1001, 0110, \\
u + v & \text{to configurations} & 1011, 0100, 0111, 1000, \\
\dfrac{1}{8} - 2v & \text{to configurations} & 1100, 0011, \\
\dfrac{1}{8} - u - w & \text{to configurations} & 1010, 0101,
\end{array}$$

where $u = .0374..., v = .0410..., w = .0446...$, solve the three equations

$$u + 3v + 2w = \frac{1}{4}, \tag{3.6}$$

$$\left(\frac{1}{8} + 2v\right)^2 = \frac{1}{8} - u - w, \tag{3.7}$$

$$\left(\frac{1}{8} + 2w\right)\left(\frac{1}{8} + 2u\right) = \frac{1}{8} - 2v. \tag{3.8}$$

Note that (3.6) implies that the total mass of μ is 1. Since u, v, w are all smaller than $\frac{1}{16}$, the masses are all nonnegative. Also, note that μ is invariant under the transformations

$$(\eta(1), \eta(2), \eta(3), \eta(4)) \to (1 - \eta(1), 1 - \eta(2), 1 - \eta(3), 1 - \eta(4)) \tag{3.9}$$

and

$$(\eta(1), \eta(2), \eta(3), \eta(4)) \to (\eta(2), \eta(1), \eta(4), \eta(3)). \tag{3.10}$$

We will now check that μ is negatively associated. Take f, g increasing and depending on disjoint sets of coordinates. We need to show that

$$\int fg d\mu \leq \int f d\mu \int g d\mu.$$

There are two cases to consider: (i) f depends on three coordinates and g depends on the fourth coordinate, and (ii) f and g each depends on two coordinates. In the first case, using the invariance of μ under (3.9) and (3.10), we see that it suffices to show that

$$\int [1 - 2\eta(1)] f(\eta(2), \eta(3), \eta(4)) d\mu \geq 0, \tag{3.11}$$

and

$$\int [1 - 2\eta(3)] f(\eta(1), \eta(2), \eta(4)) d\mu \geq 0. \tag{3.12}$$

The left side of (3.11) is

$$\begin{aligned} &\left(\frac{1}{8} - 2v\right)\left[f(111) + f(010) - f(101) - f(000)\right] \\ +&\left(\frac{1}{8} - 2w\right)\left[f(111) + f(100) - f(011) - f(000)\right]. \end{aligned}$$

(Recall (3.6)–(3.8) in checking this.) This is nonnegative since v and w are less than $\frac{1}{16}$, and

$$f(111) \geq f(101), \quad f(010) \geq f(000), \quad f(111) \geq f(011), \quad f(100) \geq f(000).$$

The left side of (3.12) is

$$\begin{aligned} &\left(2u + v + w - \frac{1}{8}\right)\left[f(111) + f(100) - f(011) - f(000)\right] \\ &+(w - u)\left[f(111) + f(101) - f(010) - f(000)\right] \\ &+\left(w + 3v - \frac{1}{8}\right)\left[f(111) + f(001) - f(110) - f(000)\right]. \end{aligned}$$

Again, this is nonnegative since f is increasing and $\frac{1}{32} < u < v < w$.

Turning to case (ii), it suffices to consider the cases in which f and g is each an indicator function of an event of the form $\{\eta(i) = \eta(j) = 1\}$ or $\{\eta(i) + \eta(j) \geq 1\}$ for $i \neq j$. (To see this, note that every increasing function of two variables can be written as a positive linear combination of these two indicators, together with increasing functions that only depend on one variable.) If they are both of the same type, f and g are negatively correlated since

$\mu(1111) = \mu(0000) = 0$. Keeping in mind the invariance of μ under (3.9) and (3.10), there are three cases to consider:

$$\begin{gathered}\mu(1100) \geq \mu(11??)\mu(??00), \quad \mu(1010) \geq \mu(1?1?)\mu(?0?0), \\ \mu(1001) \geq \mu(1??1)\mu(?00?).\end{gathered} \tag{3.13}$$

Here we have used ? to indicate a site at which no constraint is placed. The first two inequalities in (3.13) are in fact equalities – these are (3.8) and (3.7) respectively. The third is equivalent to

$$v + w \geq \frac{1}{16},$$

which is true. Thus μ is negatively associated.

To check that μ_t is not negatively associated for some t, let

$$\nu = \lim_{t\to\infty} \mu_t.$$

Then

$$\begin{aligned}\nu(1100) =& \frac{1}{2}\mu(1100) + \frac{1}{2}\mu(1010) = \frac{1}{8} - \frac{u+w}{2} - v, \\ \nu(11??) =& \frac{1}{2}\mu(11??) + \frac{1}{2}\mu(1?1?) = \frac{1}{8} + w + v, \\ \nu(??00) =& \frac{1}{2}\mu(??00) + \frac{1}{2}\mu(?0?0) = \frac{1}{8} + u + v.\end{aligned}$$

Therefore

$$\nu(1100) - \nu(11??)\nu(??00) = -(w-v)(v-u) < 0.$$

(We have used identities (3.7) and (3.8) to simplify this expression.) It follows that ν is not negatively associated.

The example in Theorem 3.5 shows that the evolution of the symmetric exclusion process does not in general preserve the property

$$\mu\{\eta : \eta \equiv 1 \text{ on } A, \eta \equiv 0 \text{ on } B\} \geq \mu\{\eta : \eta \equiv 1 \text{ on } A\}\mu\{\eta : \eta \equiv 0 \text{ on } B\} \tag{3.14}$$

for disjoint sets A, B, even if A and B are taken to be doubletons. (By Theorem 3.1 and the remark following its proof, this property is preserved if we restrict A or B to be a singleton.)

This suggests looking for a stronger condition than (3.14) that is preserved by the evolution. That would imply that if the initial measure satisfies this stronger condition, then the distribution at time t will satisfy (3.14). With this objective in mind, for nonnegatve integers i, j and real $\gamma > 0$, define $\mathcal{M}_{i,j}(\gamma)$ to be the collection of all probability measures μ on $\{0,1\}^S$ for which

$$4\mu\{\eta : \eta \equiv 1 \text{ on } A, \eta \equiv 0 \text{ on } B\} \geq \left[\gamma\mu\{\eta : \eta \equiv 1 \text{ on } A\} + \gamma^{-1}\mu\{\eta : \eta \equiv 0 \text{ on } B\}\right]^2 \tag{3.15}$$

for all disjoint A, B with cardinalities $|A| = i$, $|B| = j$. Note that (3.15) implies (3.14). This stronger condition is preserved by the evolution, as we will now check.

Theorem 3.16. *Fix i, j, γ as above. Suppose $\mu \in \mathcal{M}_{i,j}(\gamma)$ and μ_t is the distribution at time t for a symmetric exclusion process with initial distribution μ. Then $\mu_t \in \mathcal{M}_{i,j}(\gamma)$ for all $t \geq 0$.*

Proof. As in the proof of Theorem 3.1, it suffices to consider the exclusion process in which the coordinates $\eta(u)$ and $\eta(v)$ are interchanged at rate 1, and all other coordinates remain fixed. Let

$$\Gamma(A, B) = \mu\{\eta : \eta \equiv 1 \text{ on } A, \eta \equiv 0 \text{ on } B\},$$
$$\Gamma_1(A) = \mu\{\eta : \eta \equiv 1 \text{ on } A\} \quad \text{and} \quad \Gamma_0(B) = \mu\{\eta : \eta \equiv 0 \text{ on } B\}.$$

We need to check that if $\mu \in \mathcal{M}_{i,j}(\gamma)$, then

$$4\alpha\Gamma(A, B) + 4(1-\alpha)\Gamma(A_{u,v}, B_{u,v}) \geq$$
$$\left[\gamma\alpha\Gamma_1(A) + \gamma(1-\alpha)\Gamma_1(A_{u,v}) + \gamma^{-1}\alpha\Gamma_0(B) + \gamma^{-1}(1-\alpha)\Gamma_0(B_{u,v})\right]^2$$

for disjoint A, B of cardinalities i, j respectively. Since $\mu \in \mathcal{M}_{i,j}(\gamma)$, it suffices to check that

$$\begin{aligned} &\alpha\left[\gamma\Gamma_1(A) + \gamma^{-1}\Gamma_0(B)\right]^2 + (1-\alpha)\left[\gamma\Gamma_1(A_{u,v}) + \gamma^{-1}\Gamma_0(B_{u,v})\right]^2 \geq \\ &\left[\gamma\alpha\Gamma_1(A) + \gamma(1-\alpha)\Gamma_1(A_{u,v}) + \gamma^{-1}\alpha\Gamma_0(B) + \gamma^{-1}(1-\alpha)\Gamma_0(B_{u,v})\right]^2. \end{aligned} \tag{3.17}$$

But this is true, since (the left side of (3.17) – the right side of (3.17)) can be written as

$$\alpha(1-\alpha)\left[\gamma\Gamma_1(A) + \gamma^{-1}\Gamma_0(B) - \gamma\Gamma_1(A_{u,v}) - \gamma^{-1}\Gamma_0(B_{u,v})\right]^2.$$

Unfortunately, Theorem 3.16 does not allow us to solve the problem posed at the end of Section 1: Does (1.5) hold for all disjoint A, B? The reason is that the pointmass at η is typically not in $\mathcal{M}_{i,j}(\gamma)$ for any γ.

In order to understand Theorem 3.16, we need to better understand the classes $\mathcal{M}_{i,j}(\gamma)$. We do this within the context of exchangeable measures. Exchangeable measures are invariant for any symmetric exclusion process, so it is not of any interest to ask about the evolution with an exchangeable initial distribution. However, if an exchangeable measure on a finite S satisfies the inequalities in (3.15) with strict inequalities, then neighboring measures will be in $\mathcal{M}_{i,j}(\gamma)$.

Suppose then that μ is exchangeable on a set of cardinality m. Define

$$a_k = \mu\{\eta : \eta \equiv 1 \text{ on } A, \eta \equiv 0 \text{ on } A^c\},$$

where $|A| = k$. In his 2000 paper, Pemantle considers a number of possible definitions of negative correlations, and showed that many of them agree for exchangeable measures, and to the following: The sequence a_k is logconcave, i.e.,

$$a_k^2 \geq a_{k-1}a_{k+1}, \quad 1 \leq k < m.$$

Since μ is exchangeable, we may modify the notation used in the proof of Theorem 3.16, and write

$$\Gamma(i, j) = \Gamma(A, B), \quad \Gamma_1(i) = \Gamma_1(A), \quad \Gamma_0(j) = \Gamma_0(B),$$

where $|A| = i, |B| = j$, and the two sets are disjoint. Then

$$\Gamma(i, j) = \sum_{k=i}^{m-j} \binom{m-i-j}{k-i} a_k, \ \Gamma_1(i) = \sum_{k=i}^{m} \binom{m-i}{k-i} a_k, \ \Gamma_0(j) = \sum_{k=0}^{m-j} \binom{m-j}{k} a_k.$$

Proposition 3.18. *If μ is exchangeable with logconcave sequence $\{a_k\}$, then $\mu \in \mathcal{M}_{i,j}(\gamma)$ for*

$$\gamma = \sqrt{\Gamma_0(j)/\Gamma_1(i)}.$$

Proof. With this choice of γ, $\mu \in \mathcal{M}_{i,j}(\gamma)$ is equivalent to

$$\Gamma(i,j) \geq \Gamma_1(i)\Gamma_0(j).$$

This follows from Theorem 2.7 of Pemantle (2000), part of which says that the logconcavity of $\{a_k\}$ implies that μ is negatively associated.

We conclude with a result that supports the idea that for a reasonable class of initial distributions, the distribution at later times of a symmetric exclusion process is negatively associated.

Theorem 3.19. *Suppose the initial distribution for an irreducible symmetric exclusion process on the finite set S is a product measure on $\{0,1\}^S$, and let μ_t be the distribution of the process at time t. Then*

$$\nu = \lim_{t\to\infty} \mu_t$$

is negatively associated.

Proof. Of course, ν is simply the exchangeable measure for which the distribution of the number of particles

$$\sum_{x\in S} \eta(x)$$

is the same as it is under μ. We prove the stronger result that the sequence $\{a_k\}$ corresponding to ν is logconcave by induction on the cardinality of S. Suppose we know the result for S, and we want to prove it for $S' = S \cup \{u\}$. Then we need to know that if ν is exchangeable on $\{0,1\}^S$ with logconcave sequence $\{a_k\}$ and $\rho \in [0,1]$, then the exchangeable measure ν' on $\{0,1\}^{S'}$ that is obtained by starting with $\mu \times [\rho\delta_1 + (1-\rho)\delta_0]$ and permuting the sites at random, has logconcave sequence $\{a'_k\}$. But, letting $m = |S|$,

$$a'_k = \frac{\rho k a_{k-1} + (1-\rho)(m-k+1)a_k}{m+1},$$

so

$$\begin{aligned}(m+1)^2\big[(a'_k)^2 - a'_{k-1}a'_{k+1}\big] &= \\ &\rho^2(k^2-1)(a_{k-1}^2 - a_{k-2}a_k) + \rho(1-\rho)(k-1)(m-k)\big[a_{k-1}a_k - a_{k-2}a_{k+1}\big] \\ &+ (1-\rho)^2[(m-k+1)^2-1]\big[a_k^2 - a_{k-1}a_{k+1}\big] + \big[\rho a_{k-1} - (1-\rho)a_k\big]^2.\end{aligned}$$

This is nonnegative since $\{a_k\}$ is logconcave.

Remark. A more general form of the above argument can be found in Theorem 1 of Liggett (1997).

References

E. D. Andjel, *A correlation inequality for the symmetric exclusion process*, Ann. Probab. **16** (1988), 717–721.

E. D. Andjel, M. D. Bramson and T. M. Liggett, *Shocks in the asymmetric exclusion process*, Probab. Th. Rel. Fields **78** (1988), 231–247.

R. Arratia, *Symmetric exclusion processes: A comparison inequality and a large deviations result*, Ann. Probab. **13** (1985), 53–61.

A. Asselah and P. A. Ferrari, *Regularity of quasi-stationary measures for simple exclusion in dimension* $d \geq 5$, Ann. Probab. (2002).

J. T. Cox, *An alternate proof of a correlation inequality of Harris*, Ann. Probab. **12** (1984), 272–273.

D. Dubhashi and D. Ranjan, *Balls and bins: a study in negative dependence*, Ran. Struct. Algor. **13** (1998), 99–124.

P. A. Ferrari, A. Galves and C. Landim, *Rate of convergence to equilibrium of symmetric simple exclusion processes*, Markov Proc. Rel. Fields **6** (2000), 73–88.

T. E. Harris, *A correlation inequality for Markov processes in partially ordered state spaces*, Ann. Probab. **5** (1977), 451–454.

K. Joag-Dev and F. Proschan, *Negative association of random variables with applications*, Ann. Statist. **11** (1983), 286–295.

T. M. Liggett, *Interacting Particle Systems*, Springer, New York, 1985.

T. M. Liggett, *Ultra logconcave sequences and negative dependence*, J. Comb. Th. A **79** (1997), 315–325.

T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer, Berlin, 1999.

R. Pemantle, *Towards a theory of negative dependence*, J. Math. Phys. **41** (2000), 1371–1390.

Department of Mathematics
University of California
Los Angeles CA 90095
tml@math.ucla.edu