Some Conditional Correlation Inequalities for Percolation and Related Processes

J. van den Berg,* O. Häggström, J. Kahn†
CWI, Chalmers University of Technology, and Rutgers University;
e-mail: J.van.den.Berg@cwi.nl; olleh@math.chalmers.se; jkahn@math.rutgers.edu

Received 12 August 2004; accepted 3 January 2005; received in final form 13 June 2005
Published online 12 December 2005 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/rsa.20102

ABSTRACT: Consider ordinary bond percolation on a finite or countably infinite graph. Let s, t, a, and b be vertices. An earlier paper (J. Van den Berg and J. Kahn, Ann Probab 29 (2001), 123–126) proved the (nonintuitive) result that, conditioned on the event that there is no open path from s to t, the two events "there is an open path from s to a" and "there is an open path from s to b" are positively correlated. In the present paper we further investigate and generalize the theorem of which this result was a consequence. This leads to results saying, informally, that, with the above conditioning, the open cluster of s is conditionally positively (self-)associated and that it is conditionally negatively correlated with the open cluster of t.

We also present analogues of some of our results for (a) random-cluster measures and (b) directed percolation and contact processes and observe that the latter lead to improvements of some of the results in a paper of Belitsky et al. (Stoch Proc Appl 67 (1997), 213–225). Random Struct. Alg., 29, 417–435, 2006

Keywords: correlation inequalities; percolation; contact process; random-cluster model; Ahlswede–Daykin Theorem

---

Correspondence to: J. Kahn, Department of Mathematics, Rutgers University, Piscataway, NJ 08854.
*Part of vdB's research is financially supported by BRICKS Project AFM2.2.
†Supported by NSF Grant DMS0200856.

1. INTRODUCTION AND RESULTS FOR ORDINARY PERCOLATION

This paper is concerned with positive and negative correlation and the stronger notion of positive association. Recall that events A, B (in some probability space) are *positively correlated* if $\Pr(AB) \geq \Pr(A)\Pr(B)$ and negatively correlated if the reverse inequality holds. The stronger notion of positive association will be defined below (following Theorem 1.3).

We begin in this section with results for ordinary bond percolation. Our original motivation here was Theorem 1.4. All results in this section have analogues for site percolation; our reason for focusing on bond rather than site percolation is that this is the natural setting for the extensions we consider in the following sections. These extensions deal with (i) the random cluster model (Section 2) and (ii) percolation on directed graphs, together with applications to the contact process (Section 3).

A few words about proofs may be in order. The approach given for percolation in Section 1 is similar to that of [3] (see the proof of the present Theorem 1.1). This approach does not seem applicable to the random cluster model, and Section 2 takes a completely different route, based on Markov chains, to extend the results of Section 1 to this more general setting. We also describe, in Section 2.2, a different way of getting at some of the random cluster results. This is based on a connection with the so-called fuzzy Potts model and is included here despite handling only a subset of what's covered by the Markov chain approach, because we think the connection is interesting. Most of the results for "directed percolation" in Section 3 can again be obtained using either the approach of Section 1 or the Markov chain approach of Section 2, and we do not repeat the arguments. A partial exception is Theorem 3.4, one part of which we do prove, both because it requires the least routine extension of earlier ideas and because it is the one place in the present work where the Markov chain approach seems not to apply.

Consider bond percolation on a (finite or countably infinite, locally finite) graph $G = (V, E)$, where each edge e is, independently of all other edges, open with probability p_e and closed with probability $1 - p_e$. For $a, b \in V$ the event that there is an open path from a to b is denoted by $a \leftrightarrow b$ and the complement of this event by $a \not\leftrightarrow b$. For $X, Y \subset V$ we write $X \not\leftrightarrow Y$ for the event $\{x \not\leftrightarrow y \ \forall x \in X, y \in Y\}$.

In an earlier paper [3] we showed that, for any vertices s, t, a, b,

$$\Pr(s \leftrightarrow a,\ s \leftrightarrow b \mid s \not\leftrightarrow t) \geq \Pr(s \leftrightarrow a \mid s \not\leftrightarrow t) \Pr(s \leftrightarrow b \mid s \not\leftrightarrow t). \tag{1}$$

This was a consequence (really a special case) of Theorem 1.2 of [3], to which we will return below.

Here we show, among other results, a sort of complement of (1), *viz.*,

$$\Pr(s \leftrightarrow a,\ t \leftrightarrow b \mid s \not\leftrightarrow t) \leq \Pr(s \leftrightarrow a \mid s \not\leftrightarrow t) \Pr(t \leftrightarrow b \mid s \not\leftrightarrow t). \tag{2}$$

In this section we will prove the quite intuitive (2) by way of a generalization of the *not* very intuitive (1). Before giving this generalization, we need some further definitions and notation.

Let s be a fixed vertex. By the *open cluster*, C_s, of s we mean the set of all edges that are in open paths starting at s. As in [3] we define, for $X \subseteq V$, the event

$$R_X := \{s \not\leftrightarrow X\} = \{s \not\leftrightarrow x \ \forall x \in X\}.$$

Let $\Omega = \{0, 1\}^E$ be the set of realizations; elements of Ω will typically be denoted ω. We write $\omega' \geq \omega$ as short for $\omega'(e) \geq \omega(e) \forall e \in E$. For two events $A, B \subseteq \Omega$, we abbreviate

$A \cap B$ as AB. Recall that an event A is *increasing* (really, *nondecreasing*) if $\omega' \geq \omega \in A$ implies $\omega' \in A$. We also say that A is *increasing and determined by the open cluster of* s if $\omega \in A$ and $C_s(\omega') \supseteq C_s(\omega)$ imply $\omega' \in A$ (note that such an event is increasing in the sense above) and analogously that A is *decreasing and determined by the open cluster of* s if $\omega \in A$ and $C_s(\omega') \subseteq C_s(\omega)$ imply $\omega' \in A$. A simple example of an event that is increasing and determined by C_s, is $\{s \leftrightarrow a\}$.

The following statement is a natural generalization of Theorem 1.2 of [3]. To avoid technicalities concerning regularity conditions on events (and functions) we restrict here, as well as in Theorems 1.2–1.4 below, to finite graphs. Our most general result in this section, Theorem 1.5, does include infinite graphs; but its formulation, in terms of positive association, avoids the aforementioned technicalities.

Theorem 1.1. *Let the graph G be finite, and let A and B be increasing events determined by the open cluster of s. Then, for all $X, Y \subseteq V \setminus \{s\}$,*

$$\Pr(A\, R_X)\Pr(B\, R_Y) \leq \Pr(A\, B\, R_{X\cap Y})\,\Pr(R_{X\cup Y}). \tag{3}$$

Remark. Theorem 1.2 in [3] is the special case where each of A, B is of the form $\{s \leftrightarrow w\ \forall w \in U\}$ for some $U \subset V$. The proof of the present more general result is almost the same and we present it in a slightly abbreviated form, emphasizing the parts that need extra attention because of the generalization. (One may say that the key idea (in both cases) is generalizing from statements like (1) to the form (3), which supports an inductive proof.)

Proof. The proof is by induction on the number of vertices. When G has only one vertex, the result is obvious; so we suppose, for some $n \geq 1$, that the result holds for graphs with at most n vertices and consider G with $n+1$ vertices.

For $X \subseteq V$, write E_X for the set of edges with at least one endpoint in X. With notation as in the theorem, it is easy to see that there is an event $\tilde{A} \subseteq A$ with the following properties: it is increasing and determined by C_s; it does not depend on E_X (that is, if $\omega'_e = \omega_e$ for all $e \notin E_X$, then $\omega \in \tilde{A}$ iff $\omega' \in \tilde{A}$); and, finally, $\tilde{A}\, R_X = A\, R_X$. A similar remark holds for B and Y. So we may assume that A does not depend on E_X and B does not depend on E_Y.

If $X \cap Y = \emptyset$, the r.h.s. of (3) is $\Pr(A\, B)\ \Pr(R_{X\cup Y})$, and two applications of Harris' inequality give the result

$$\begin{aligned}\Pr(A\, R_X)\Pr(B\, R_Y) &\leq \Pr(A)\,\Pr(B)\,\Pr(R_X)\,\Pr(R_Y)\\ &\leq \Pr(A\, B)\,\Pr(R_{X\cup Y})\end{aligned} \tag{4}$$

(note $R_{X\cup Y} = R_X R_Y$).

Now suppose $Z := X \cap Y \neq \emptyset$. Let N be the set of all vertices outside Z with at least one neighbor in Z. Let $\mathbf{S}$ be the (random) set of those vertices of N connected to Z by at least one open edge. We have

$$\Pr(A\, R_X) = \sum_S \Pr(\mathbf{S} = S)\,\Pr(A\, R_X | S),$$

where the sum is over $S \subseteq N$ and we write $\Pr(\cdot|S)$ for $\Pr(\cdot|\mathbf{S} = S)$. Similar expressions hold for the other terms in (3). Moreover, clearly,

$$\Pr(\mathbf{S} = S)\,\Pr(\mathbf{S} = T) = \Pr(\mathbf{S} = S\cap T)\,\Pr(\mathbf{S} = S\cup T) \qquad \forall S, T \subseteq N.$$

So according to the Ahlswede–Daykin ("Four Functions") Theorem ([1] or, e.g., [5]), (3) will follow if we show that, for all $S, T \subseteq N$,

$$\Pr(A\,R_X|S)\ \Pr(B\,R_Y|T) \le \Pr(A\,B\,R_{X\cap Y}|S\cap T)\,\Pr(R_{X\cup Y}|S\cup T). \tag{5}$$

Now it is easy to see that, for any set of vertices $W \supseteq Z$, and any event D that does not depend on E_W,

$$\Pr(D\,R_W|S) = \Pr'(D\,R_{(W\setminus Z)\cup S}), \tag{6}$$

where $\Pr'$ refers to the induced model on the graph G' obtained from G by removing Z. (Strictly speaking, the D on the r.h.s. of (6) is not the same as that on the l.h.s., since it is a subset of $\{0,1\}^{E\setminus E_Z}$ rather than $\{0,1\}^E$; but since D does not depend on E_W (and hence not on E_Z), the two events are essentially the same, so we ignore the irrelevant distinction.)

Applying (6) to each of the four terms in (5), we have

$$\begin{aligned}\Pr(AR_X|S)\Pr(BR_Y|T) &= \Pr'(AR_{(X\setminus Z)\cup S})\,\Pr'(BR_{(Y\setminus Z)\cup T})\\ &\le \Pr'(ABR_{((X\setminus Z)\cup S)\cap((Y\setminus Z)\cup T)})\,\Pr'(R_{(X\setminus Z)\cup S\cup(Y\setminus Z)\cup T})\\ &\le \Pr'(ABR_{((X\cap Y)\setminus Z)\cup(S\cap T)})\,\Pr'(R_{((X\cup Y)\setminus Z)\cup(S\cup T)})\\ &= \Pr(ABR_{X\cap Y}|S\cap T)\,\Pr(R_{X\cup Y}|S\cup T),\end{aligned}$$

where the first inequality follows from our inductive hypothesis (applicable since G' has fewer vertices than G) and the second from

$$((X\setminus Z)\cup S)\cap((Y\setminus Z)\cup T) \supseteq ((X\cap Y)\setminus Z)\cup(S\cap T)$$

and

$$(X\setminus Z)\cup S\cup(Y\setminus Z)\cup T = ((X\cup Y)\setminus Z)\cup(S\cup T).$$

∎

In particular we have the promised generalization of (1):

Theorem 1.2. *For G, s, A, B, and X as in Theorem* 1.1,

$$\Pr(A\,B \mid s \not\leftrightarrow X) \ge \Pr(A \mid s \not\leftrightarrow X)\,\Pr(B \mid s \not\leftrightarrow X). \tag{7}$$

Proof. Take $Y = X$ in Theorem 1.1. ∎

Remarks.

1. It is easy to see that Theorem 1.2 is equivalent to the special case where $|X| = 1$. (To reduce to this, simply identify the vertices of X, retaining all edges connecting them to $V\setminus X$ (edges internal to X may be deleted, but are anyway irrelevant).) We have used the present form both because it will be convenient for the proof of Theorem 1.5 and because it is natural from the point of view of the contact process application in Section 3. Similarly, we could replace s in all results of this section and t in Theorems 1.4 and 1.5 by *sets* of vertices. The same easy equivalence holds for the directed graph results of Section 3; but in the case of the random cluster measures

of Section 2 the more general statements, while still true, do not seem to follow in the same way from their specializations.

2. The derivation of Theorem 1.2 may give the impression that it is less general than Theorem 1.1, but in fact there is an easy way to derive Theorem 1.1 from Theorem 1.2. To see this, first note that if an event is increasing and determined by C_x, then its complement is decreasing and determined by C_x. It follows that Theorem 1.2 also holds when A and B are both decreasing rather than increasing, while the inequality (7) reverses if one of A, B is increasing and the other decreasing. Thus, for A, B, X, Y as in Theorem 1.1, Theorem 1.2 implies that conditioned on $R_{X\cap Y}$, each of the pairs $(A, R_{X\setminus Y})$, $(B, R_{Y\setminus X})$ is negatively correlated, while each of (A, B), $(R_{X\setminus Y}, R_{Y\setminus X})$ is positively correlated. So, writing $\Pr'$ for our percolation measure conditioned on $R_{X\cap Y}$, we have (compare (4))

$$\begin{aligned}\Pr'(A\,R_{X\setminus Y})\Pr'(B\,R_{Y\setminus X}) &\le \Pr'(A)\Pr'(R_{X\setminus Y})\Pr'(B)\Pr'(R_{Y\setminus X})\\ &\le \Pr'(AB)\Pr'(R_{X\setminus Y}R_{Y\setminus X}),\end{aligned}$$

which is equivalent to (3).

It will be helpful to have the "functional extension" of Theorem 1.2:

Theorem 1.3. *Let G, s, and X be as in Theorem* 1.1*, and let f and g be increasing functions of C_s. Then*

$$\mathsf{E}[f\,g \mid s \nleftrightarrow X] \ge \mathsf{E}[f \mid s \nleftrightarrow X]\,\mathsf{E}[g \mid s \nleftrightarrow X]. \tag{8}$$

The inequality is reversed if one of f,g is increasing and the other decreasing.

Proof. The first assertion is a standard reduction: Theorem 1.2 immediately implies that (8) holds when f and g are indicators of increasing events. Since any increasing f and g can be written as sums of such indicators with positive weights, (8) follows. The second assertion of the theorem follows from (8), applied to the pair $(f, -g)$. ∎

Recall that a finite collection of $\{0,1\}$-valued random variables $\sigma_1, \cdots, \sigma_n$ is said to be *positively associated* if for any two functions f, g of the σ_i's that are either both increasing or both decreasing, one has $\mathsf{E}fg \ge \mathsf{E}f\,\mathsf{E}g$; the corresponding measure on $\{0,1\}^n$ is then also said to exhibit positive association. The simplest non-trivial example is when the σ_i's are independent (Harris' inequality). A countably infinite collection of $\{0,1\}$-valued random variables is said to be positively associated if each finite subcollection is positively associated.

If we define a random subset W of a set T to be positively associated if the collection $\{\eta(a) = \mathbf{1}_{\{a\in W\}} : a \in T\}$ is positively associated, then Theorem 1.3 says that the open cluster of s is conditionally positively associated given the event $\{s \nleftrightarrow X\}$. We will see further, similar examples later.

Positive association is often derived from the FKG inequality, which generalizes Harris' inequality and says that positive association holds for measures on $\{0,1\}^n$ satisfying the "positive lattice condition" (also called "FKG lattice condition"), *viz.*,

$$\mu(\sigma)\mu(\tau) \le \mu(\sigma\wedge\tau)\mu(\sigma\vee\tau) \tag{9}$$

(where $(\sigma \wedge \tau)_x$ and $(\sigma \vee \tau)_x$ are the minimum and maximum of σ_x and τ_x). The positive lattice condition is much stronger than positive association. It says that the conditional probability that $\sigma_x = 1$, given the values of $\sigma_y, y \neq x$, is increasing in those values.

Let us also recall here that for measures ν and ν' on $\{0, 1\}^n$ (or some other partially ordered set), ν *stochastically dominates* ν' ($\nu \succ \nu'$) if $\nu(f) \geq \nu'(f)$ for every increasing function f (where $\nu(f)$ is the expectation of f w.r.t. ν).

As suggested earlier, we find Theorem 1.3 somewhat counterintuitive. (For instance, as noted in [3], it is easy to see that the analogous statement with $s \leftrightarrow t$ in place of $s \not\leftrightarrow t$ is false.) Nonetheless, it turns out to imply (see the proof of Theorem 1.5 below) the following intuitively more natural statement, which says, informally, that conditioned on nonexistence of an open (s, t)-path, the clusters C_s and C_t are negatively correlated.

Theorem 1.4. *Let G and s be as in Theorem 1.1, t a vertex distinct from s, and f and g increasing functions of C_s and C_t, respectively. Then*

$$\mathsf{E}[f\,g \mid s \not\leftrightarrow t] \leq \mathsf{E}[f \mid s \not\leftrightarrow t]\,\mathsf{E}[g \mid s \not\leftrightarrow t].$$

Note that (2) is the special case where f is the indicator of the event $\{s \leftrightarrow a\}$ and g that of the event $\{t \leftrightarrow b\}$.

We have stated Theorem 1.4 above largely because, as mentioned earlier, it was the original motivation for this work; but the next statement, which contains Theorem 1.3 as well as Theorem 1.4, seems to be the correct level of generality here.

Theorem 1.5. *Let G be a finite or countably infinite, locally finite graph and s and t (distinct) vertices. Let, for each edge e, $X_e = \mathbf{1}_{\{e \in C_s\}}$ and $Y_e = \mathbf{1}_{\{e \in C_t\}}$. Then, conditional on $\{s \not\leftrightarrow t\}$, the collection*

$$\{X_e : e \in E\} \cup \{1 - Y_e : e \in E\}$$

is positively associated.

Proof of Theorem 1.5. We give the proof for finite G. The inifinite case then follows from standard limit arguments.

Let f and g be functions of (C_s, C_t), each increasing in C_s and decreasing in C_t. We have to show that

$$\mathsf{E}[f\,g \mid s \not\leftrightarrow t] \geq \mathsf{E}[f \mid s \not\leftrightarrow t]\,\mathsf{E}[g \mid s \not\leftrightarrow t]. \tag{10}$$

Note that the l.h.s. of (10) can be written as

$$\mathsf{E}[f\,g \mid s \not\leftrightarrow t] = \sum_W \Pr(C_s = W \mid s \not\leftrightarrow t)\,\mathsf{E}[f\,g \mid C_s = W], \tag{11}$$

where we may restrict to W containing no (s, t)-path. Write $\bar{W}$ for the union of W and its "boundary"; that is, $\bar{W}$ consists of all edges having at least one vertex in common with some edge of W.

When we condition on $\{C_s = W\}$, f and g become decreasing functions of C_t, and the (conditional) distribution of C_t is the same as that for the restriction of our percolation model to the graph obtained from G by deleting all edges in $\bar{W}$. Thus (on $\{C_s = W\}$), f, g are decreasing functions of the independent r.v.'s $(\omega_e : e \in E \setminus \bar{W})$, and by Harris' inequality we have $\mathsf{E}[fg|C_s = W] \geq \mathsf{E}[f|C_s = W]\,\mathsf{E}[g|C_s = W]$.

On the other hand, the conditional distribution of C_t given $\{C_s = W\}$ is stochastically decreasing in W (to couple these distributions, choose *all* ω_e's independently according to their p_e's and then for conditioning on $\{C_s = W\}$ simply ignore those ω_e's with $e \in \bar{W}$); so in particular $\mathsf{E}[f|C_s = W]$ and $\mathsf{E}[g|C_s = W]$ are increasing functions of W, and it then follows from Theorem 1.3 that the right-hand side of (11) is not less than

$$\left(\sum_W \Pr(C_s = W \mid s \not\leftrightarrow t)\,\mathsf{E}[f|C_s = W]\right)\left(\sum_W [\Pr(C_s = W \mid s \not\leftrightarrow t)\,\mathsf{E}[g|C_s = W]\right)$$
$$= \mathsf{E}[f \mid s \not\leftrightarrow t]\,\mathsf{E}[g \mid s \not\leftrightarrow t],$$

so we have (10). ∎

As just shown, Theorem 1.5, and hence its special case Theorem 1.4, follows easily from Theorem 1.3. While one might expect that in a similar (or *some*) way the reverse implication (that is, Theorem 1.3 from Theorem 1.4) can be shown, we do not see this. In Section 2 we will (as mentioned earlier) take a completely different approach that, even for the more general class of random-cluster measures, gives Theorems 1.3 and 1.4 "simultaneously."

2. RANDOM-CLUSTER MEASURES

2.1. Definitions and a Markov Chain Proof

A well-known generalization of the bond percolation model is the *random-cluster model* introduced by Fortuin and Kasteleyn circa 1970 (see, e.g., [7] and [9] for additional background and references).

Let $G = (V, E)$ be a finite graph. In addition to the parameters p_e, $e \in E$, of the ordinary bond percolation model, the random-cluster model is equipped with a positive parameter q. To avoid trivialities we assume that $0 < p_e < 1$ for all $e \in E$. The *random-cluster measure* (r.c.m.) with the above parameters on $\Omega = \{0, 1\}^E$ is then given by

$$\varphi_q(\omega)(= \varphi_{G,q}(\omega)) \propto q^{k(\omega)} \prod_{e \in E : \omega_e = 1} p_e \prod_{e \in E : \omega_e = 0} (1 - p_e), \quad \omega \in \Omega, \tag{12}$$

where $k(\omega)$ is the number of connected components in ω, and, as usual, $f(\omega) \propto g(\omega)$ means $f(\omega) = Cg(\omega)$ for some (positive) constant C. (For the present discussion we regard the p_e's as given once and for all and omit them from our notation.)

Thus, $q = 1$ gives the ordinary bond percolation model. We have, despite serious attempts, not been able to adapt the approach of Section 1 to $q > 1$. (We do not consider $q < 1$, for which the correlation properties of the model are quite different). Here we take a different, "dynamical" approach, based on the introduction of a Markov chain whose states are pairs of clusters (this is not the only possibility; see the remark following the proof of Theorem 2.1), which converges to a measure (on pairs of clusters) corresponding to (12), and for intermediate stages of which the correlation properties we are after can be derived from known properties of the random-cluster model. Similar uses of Markov chains have turned out to be quite useful in this field; see [13] for a pioneering example.

For the following extension of Theorem 1.5 to the random-cluster model we replace the vertices s and t by sets S and T, recalling that the remark following Theorem 1.2 regarding the easy reduction from sets to singletons is not valid here. Extending our earlier notation, we use C_S for the set of edges belonging to open paths starting at vertices of S.

Theorem 2.1. *Consider a distribution* (12) *with* $q \geq 1$. *Let* S *and* T *be disjoint sets of vertices and* f *and* g *functions of* (C_S, C_T), *each increasing in* C_S *and decreasing in* C_T. *Then, conditional on* $\{S \not\leftrightarrow T\}$,

$$\mathsf{E}fg \geq \mathsf{E}f\,\mathsf{E}g. \tag{13}$$

Following the Markov chain proof of this result, we also give, in Section 2.2, a different argument, which unfortunately seems only to work when $q \geq 2$ and $|S| = |T| = 1$. So, somewhat strangely, we have separate (and distinct) proofs for the cases $q = 1$ and $q \geq 2$, but for the intermediate case $1 < q < 2$, no alternative to the Markov chain approach.

For the proof of Theorem 2.1, we first give some additional notation and state some (well-known) lemmas we will need. If F is a subset of E (the set of edges of our graph G), and $\omega \in \Omega = \{0,1\}^E$, we write ω_F for the restriction of ω to F ($\omega_F = (\omega_e : e \in F)$), and $V(F)$ for the set of vertices incident with edges of F. We continue to use the notation $\bar{W}$ introduced following (11).

Lemma 2.2. *For* $q \geq 1$, *the random-cluster measure* (12) *satisfies the positive lattice condition* (9).

(See, e.g., [8] for a proof.)

When $\varphi_{G,q}$ is conditioned on the values of some of the variables ω_e, the remaining variables are distributed as they would be under the (natural) r.c.m. on the graph obtained from G by deleting e's with $\omega_e = 0$ and contracting those with $\omega_e = 1$. For our purposes the relevant cases of this are given by the following.

Lemma 2.3. *Fix* $A \subset V$ *and* $F \subset E$ *such that the event* $\{C_A = F\}$ *is nonempty. The restriction of* $\varphi_{G,q}$ *to* $\{0,1\}^{E \setminus \bar{F}}$ *under conditioning on either of the events* $\{C_A = F\}$, $\{\omega_{\bar{F}} \equiv 0\}$ *(i.e.,* $\{\omega_e = 0\, \forall e \in \bar{F}\}$*) is the r.c.m. with parameter* q *on* $G - \bar{F}$ *(the graph obtained from* G *by deleting all edges in* $\bar{F}$*); more formally,*

$$\varphi_{G,q}(\omega_{E\setminus\bar{F}} = \cdot \mid C_A = F) = \varphi_{G,q}(\omega_{E\setminus\bar{F}} = \cdot \mid \omega_{\bar{F}} \equiv 0) = \varphi_{G-\bar{F},q}(\cdot).$$

If A, F are as in Lemma 2.3, and $B \subseteq V \setminus V(F)$, then $\{C_A = F\} \subseteq \{A \not\leftrightarrow B\}$; so Lemma 2.3 implies

Lemma 2.4. *If* A, F *are as in Lemma 2.3,* $B \subseteq V \setminus V(F)$, *and* φ^* *is* $\varphi_{G,q}$ *conditioned on* $\{A \not\leftrightarrow B\}$, *then*

$$\varphi^*(\omega_{E\setminus\bar{F}} = \cdot \mid C_A = F) = \varphi_{G,q}(\omega_{E\setminus\bar{F}} = \cdot \mid \omega_{\bar{F}} \equiv 0) = \varphi_{G-\bar{F},q}(\cdot).$$

We now turn to the proof of Theorem 2.1. We consider a Markov chain with state space $\hat{\Omega}$ consisting of pairs (C_S, C_T) satisfying $Q := \{S \not\leftrightarrow T\}$. (So the states are pairs (C, C')

such that $C, C' \subseteq E$; C (resp., C') is a union of paths beginning at vertices of S (resp., vertices of T); and $V(C) \cap V(C') = \emptyset$.)

We write φ for the measure $\varphi_{G,q}$ conditioned on Q and $\hat{\varphi}$ for the measure that φ induces on $\hat{\Omega}$.

Initially our chain is in some fixed state $(C_S^0, C_T^0) \in \hat{\Omega}$. Given (C_S^{i-1}, C_T^{i-1}), the state of the chain at time $i-1$, we choose (C_S^i, C_T^i) in two steps, first choosing C_T^i according to φ conditioned on $\{C_S = C_S^{i-1}\}$—that is,

$$\Pr(C_T^i = \cdot) = \varphi(C_T = \cdot | C_S = C_S^{i-1})$$

—and then, similarly, C_S^i according to

$$\Pr(C_S^i = \cdot) = \varphi(C_S = \cdot | C_T = C_T^i).$$

It is clear (for instance by noting that the chain is a Gibbs sampler (see, e.g., [11]) for $\hat{\varphi}$) that $\hat{\varphi}$ is stationary for this chain and that the chain is irreducible and aperiodic, so to prove Theorem 2.1 it is enough to show:

Claim 2.5. *For f, g as in the statement of Theorem 2.1 and any n,* (13) *holds for expectation taken with respect to the law of (C_S^n, C_T^n).*

Let X_e^i and Y_e^i be the indicators of the events $\{e \notin C_T^i\}$ and $\{e \in C_S^i\}$ $(e \in E,\ i = 0, 1, \ldots)$. These are, of course, not independent, but we will show, using the following presumably well-known observation, that they are positively associated.

Lemma 2.6. *Suppose $W_1, \ldots, W_a$ and $Z_1, \ldots, Z_b$ are (say) $\{0,1\}$-valued r.v.'s with joint distribution ψ satisfying*

(i) *$W_1, \ldots, W_a$ are positively associated;*
(ii) *$Z_1, \ldots, Z_b$ are conditionally positively associated given $W_1, \ldots, W_a$; and*
(iii) *for $W, W' \in \{0,1\}^a$ with $W' \geq W$, $\psi(\cdot|W') \succ \psi(\cdot|W)$, where $\psi(\cdot|W)$ is the conditional distribution of $(Z_1, \ldots, Z_b)$ given $(W_1, \ldots, W_a) = W$.*

Then $W_1, \ldots, W_a, Z_1, \ldots, Z_b$ are positively associated.

Proof. Suppose f, g are increasing functions of $W_1, \ldots, Z_b$, and for $W \in \{0,1\}^a$, set $F(W) = \mathsf{E}[f|W]$ $(:= \mathsf{E}[f|(W_1, \ldots, W_a) = W])$ and $G(W) = \mathsf{E}[g|W]$. Then

$$\begin{aligned} \mathsf{E}fg &= \mathsf{E}\{\mathsf{E}[fg|W]\} \\ &\geq \mathsf{E}\{\mathsf{E}[f|W]\mathsf{E}[g|W]\} \\ &\geq \mathsf{E}F\mathsf{E}G \\ &= \mathsf{E}f\mathsf{E}g, \end{aligned}$$

where the first inequality follows from (ii) and the second from (iii) and (i). ∎

Lemma 2.7. *The collection*

$$\bigcup_{i\geq 1}\bigcup_{e\in E}\{X_e^i, Y_e^i\} \tag{14}$$

is positively associated.

Note that this is enough for Claim 2.5 due to the following trivial observation.

Remark 2.8. *For each n, C_S^n is increasing in the variables X_e^i, Y_e^i, and C_T^n is decreasing in these variables.*

Proof of Lemma 2.7. Of course it's enough to show positive association for finite subsets of the collection (14). We will show by induction on n that for each n, each of the collections

$$\{X_e^i : e\in E, i\leq n\}\cup\{Y_e^i : e\in E, i<n\} \tag{15}$$

and

$$\{X_e^i : e\in E, i\leq n\}\cup\{Y_e^i : e\in E, i\leq n\} \tag{16}$$

is positively associated. (The base cases—those with $n=0$—are, of course, trivial.) Actually we just give the argument for (15), that for (16) being essentially the same.

We want to apply Lemma 2.6 with

$$(W_1,\ldots,W_a)=\{X_e^i : e\in E, i<n\}\cup\{Y_e^i : e\in E, i<n\}$$

and $(Z_1,\ldots,Z_b)=(X_e^n : e\in E)$, so we need to verify conditions (i)–(iii) of the lemma. Of course (i) is just our inductive hypothesis, so our concern is really with (ii) and (iii).

Consider a possible value W of $(W_1,\ldots,W_a)$, with F the corresponding value of C_S^{n-1}. Under conditioning on $\{(W_1,\ldots,W_a)=W\}$, we have X_e^n fixed for $e\in\bar{F}$ (namely $X_e^n=1$ $\forall e\in F$ and $X_e^n=0$ $\forall e\in\bar{F}\setminus F$), while, by Lemma 2.4, the remaining X_e^n's are distributed as the variables $\mathbf{1}_{\{e\notin C_T(\omega)\}}$, where $(\omega_e : e\in E\setminus\bar{F})$ is chosen according to $\varphi_{G-\bar{F},q}$. Conditional positive association of these remaining X_e^n's follows from Lemma 2.2, so we have (ii).

Now let W' be a second possible value of $(W_1,\ldots,W_a)$, with $W'\geq W$ and F' the corresponding value of C_S^{n-1}. According to Remark 2.8 we have $F\subseteq F'$. So (iii) amounts to saying that for $F\subseteq F'\subseteq E$ and h any increasing function of C_T,

$$\varphi(h|C_S=F)\geq\varphi(h|C_S=F') \tag{17}$$

(note h is a *decreasing* function of the X_e^n's). But using Lemmas 2.3 and 2.4, we may rewrite the left- and right-hand sides of (17) as

$$\varphi_{G-\bar{F},q}(h)$$

and

$$\varphi_{G-\bar{F}',q}(h)\varphi_{G-\bar{F},q}(h|\omega_{\bar{F}'\setminus\bar{F}}\equiv 0);$$

and then (17) follows from Lemma 2.2 (which gives positive association for the measure $\varphi_{G-\bar{F},q}$). ∎

This completes the proof of Theorem 2.1. We end this section by briefly indicating an alternative proof of Theorem 2.1, again using a Markov chain and based on a similar idea. This, our original proof, is perhaps more natural than that given above, but does not seem as easily adapted to prove the directed version of Theorem 1.5 (Theorem 3.6).

We again use φ for $\varphi_{G,q}$ conditioned on $\{S \nleftrightarrow T\}$. Our chain in this case is $\omega^0, \omega^1, \omega^2, \ldots$ drawn from the state space $\hat{\Omega} := \{\omega \in \Omega : S \nleftrightarrow T\}$. Initially the chain is in some fixed state ω^0. Given ω^{i-1}, the state of the chain at time $i-1$, we choose ω^i in two steps, first choosing an intermediate configuration τ^i according to φ conditioned on $\{C_S = C_S(\omega^{i-1})\}$—that is, for $\zeta \in \hat{\Omega}$ with $C_S(\zeta) = C_S(\omega^{i-1})$,

$$\Pr(\tau^i = \zeta) = \varphi(\omega = \zeta | C_S(\omega) = C_S(\omega^{i-1}))$$

—and then, similarly, ω^i according to

$$\Pr(\omega^i = \zeta) = \varphi(\omega = \zeta | C_T(\omega) = C_T(\tau^i)).$$

It is clear that φ is stationary for this chain and that the chain is irreducible and aperiodic; so to prove Theorem 2.1 it's enough to show:

Claim. *For f, g as in the statement of Theorem 2.1 and any n,* (13) *holds for expectation taken with respect to the law of* (ω^n).

To prove this we introduce independent r.v.'s X_e^i, Y_e^i ($e \in E$, $i = 1, \ldots$), each uniform on $[0,1]$, and some fixed ordering, "$\prec$," of E. Then to decide the value of τ_e^i we compute the conditional probability, say α, that $\tau_e^i = 1$ given the values of the ω_e^{i-1}'s (or just the value of $C_S(\omega^{i-1})$) and those $\tau_{e'}^i$'s with $e' \prec e$ and set $\tau_e^i = 1$ iff $X_e^i < \alpha$. For ω^i we proceed analogously, with the requirement for $\omega_e^i = 1$ now being $Y_e^i > 1 - \alpha$.

It is then not hard to show, again using Lemmas 2.2–2.4, that (for each n) ω^n is increasing in the variables X_e^i, Y_e^i, so that the claim follows from Harris' inequality applied to these variables, and Theorem 2.1 follows.

2.2. A Separate Proof for $q \geq 2$

As mentioned earlier, it turns out, somewhat curiously, that for $q \geq 2$ and S and T consisting of single vertices s and t, we can prove Theorem 2.1 in a different way by exploiting a connection between the random cluster model and the fuzzy Potts model. (The corresponding connection involving the ordinary Potts model again goes back to Fortuin and Kasteleyn.) Before doing so, we need to review some classical and more recent facts concerning this connection.

Let $q = \alpha + \beta$ with $\alpha, \beta > 0$. Using the random-cluster measure φ we generate a random spin configuration $\sigma \in \{0,1\}^V$ as follows.

(i) Choose $\omega \in \{0,1\}^E$ according to φ_q.
(ii) For each component C of ω, let σ take the value 1 (resp., 0) on all vertices of C with probability α/q (resp., β/q), independently of the values of σ on other components. Let $\mu_{\alpha,\beta}$ denote the distribution of σ.

(In [10] this is called the *fractional fuzzy Potts model*.) This procedure produces a coupling measure $\mathcal{P}$ of ω and σ, or, rather, of φ_q and $\mu_{\alpha,\beta}$. So we may also think of first choosing σ and then drawing from the conditional distribution $\mathcal{P}(\cdot \mid \sigma)$ to obtain a typical (with distribution φ_q) edge configuration ω. It is known (and easy to check) that this "reversed" procedure can be described as follows.

(iii) Choose σ according to $\mu_{\alpha,\beta}$.
(iv) For $i = 1, 0$, let $G(i) = G[\sigma^{-1}(i)]$ (the (*induced*) subgraph consisting of vertices in $\sigma^{-1}(i)$ and edges of G contained in this set). Set $\omega_e = 0$ whenever σ assigns

different values to the ends of $e \in E$, and choose the restrictions of ω to $E(G(1))$ and $E(G(0))$ (independently) according to $\varphi_{G(1),\alpha}$ and $\varphi_{G(0),\beta}$.

Furthermore, if in (iii) we choose σ according to the *conditional* distribution

$$\hat{\mu}_{\alpha,\beta}(\cdot) = \mu_{\alpha,\beta}(\cdot \mid \sigma(s) = 1, \sigma(t) = 0),$$

then ω (in (iv)) has the distribution we want, namely $\varphi_{G,q}(\cdot \mid s \nleftrightarrow t)$; so for Theorem 2.1 we may take ω to be chosen in this way.

The salient points for our purposes are then as follows. (For clarity we now add subscripts to the expectation symbol E to indicate measures with respect to which expectation is taken.)

(a) For any graph H and $c \geq 1$, $\varphi_{H,c}$ satisfies the positive lattice condition (9) (see Lemma 2.2).
(b) It is shown in [10] that for any $\alpha, \beta \geq 1$, $\mu = \mu_{\alpha,\beta}$ satisfies (9), whence, by the FKG inequality, $\mu_{\alpha,\beta}$ and $\hat{\mu}_{\alpha,\beta}$ are positively associated.
(c) If $\alpha, \beta \geq 1$ and f is a function of (C_s, C_t) that is increasing in C_s and decreasing in C_t, then $\mathsf{E}_{\mathcal{P}}[f \mid \sigma]$ is increasing in σ on the event $\{\sigma(s) = 1, \sigma(t) = 0\}$. (This follows from (iv) and (a)).

Alternative proof of Theorem 2.1 for $q \geq 2$, $S = \{s\}$ and $T = \{t\}$. Let f, g be as in the statement of the theorem. Fix some $\alpha, \beta \geq 1$ with $\alpha + \beta = q$. For simplicity we write μ for $\mu_{\alpha,\beta}$ and φ for φ_q. The connections described above give

$$\begin{aligned}
\mathsf{E}_{\varphi}[fg \mid s \nleftrightarrow t] &= \sum_{\sigma} \hat{\mu}(\sigma)\, \mathsf{E}_{\mathcal{P}}[f(\omega)g(\omega)|\sigma] \\
&\geq \sum_{\sigma} \hat{\mu}(\sigma)\, \mathsf{E}_{\mathcal{P}}[f(\omega)|\sigma]\, \mathsf{E}_{\mathcal{P}}[g(\omega)|\sigma] \\
&\geq \sum_{\sigma} \hat{\mu}(\sigma)\, \mathsf{E}_{\mathcal{P}}[f(\omega)|\sigma] \sum_{\sigma} \hat{\mu}(\sigma)\, \mathsf{E}_{\mathcal{P}}[g(\omega)|\sigma] \\
&= \mathsf{E}_{\varphi}[f \mid s \nleftrightarrow t]\mathsf{E}_{\varphi}[g \mid s \nleftrightarrow t],
\end{aligned}$$

where the first inequality follows from (a) (and (iv)) and the second from (b) and (c). ∎

3. DIRECTED PERCOLATION AND CONTACT PROCESSES

In this section we consider another generalization of ordinary percolation: as in Section 1 we have a product distribution on $\{0, 1\}^E$, but now some (or all, or none) of the edges of our graph are oriented. Graphs of this type (allowing both directed and undirected edges) are sometimes called *mixed graphs.*

There are (at least) two natural ways to try to extend the results of Section 1 to this setting, corresponding to two possible extensions of the conditioning event $\{s \nleftrightarrow t\}$. As we will see, both extensions are reasonable for Theorems 1.1–1.3, but only one of them makes sense for Theorem 1.4 (and Theorem 1.5). The first set of extensions yield in particular improvements of some of the results of Belitsky et al. [2] regarding the contact process (defined below).

We will first indicate these extensions (proofs of which are essentially identical to the proofs of the corresponding statements in Section 1) and discuss their relevance to the contact process, before turning to the second set of extensions.

We will need the following additional notation. Unoriented edges will be denoted by $\{v,w\}$ and oriented edges by (v,w) (where the orientation is from v to w). When we speak of a *path*, we will now mean one that respects the orientations of its oriented edges. We write $\{s \to t\}$ for the event that there is an open path from s to t and $\{s \nrightarrow t\}$ for the complement of this event. The *open cluster*, C_s, of s is again the set of all edges contained in open paths starting at s. As in Section 1, we fix a vertex s and set $R_X = \{s \nrightarrow x\, \forall x \in X\}$ for each $X \subseteq V \setminus \{s\}$. Of course all these definitions collapse to those of Section 1 in case there are no oriented edges; so the next result contains Theorem 1.1.

Theorem 3.1. *With the preceding modified definitions, Theorem 1.1 holds for directed percolation.*

Proof. The proof is essentially the same as that of Theorem 1.1, the only difference being that we should now take N to be the set of those $i \notin Z$ for which there is at least one edge (i,j) or $\{i,j\}$ with $j \in Z$ and modify the definition of S similarly. ∎

Write C_X^- for the set of edges in paths *ending in X*. The most straightforward extension of Theorem 1.5 to directed percolation is:

Theorem 3.2. *Let G as above be finite, $s \in V(G)$ and $X \subseteq V(G) \setminus \{v\}$, and let f and g be functions of (C_S, C_X^-), each increasing in C_S and decreasing in C_X^-. Then*

$$\mathsf{E}[f\, g \mid s \nrightarrow X] \geq \mathsf{E}[f \mid s \nrightarrow X]\, \mathsf{E}[g \mid s \nrightarrow X]. \tag{18}$$

This can be derived beginning with Theorem 3.1 in the same way as Theorem 1.3 was derived beginning with Theorem 1.1. It can also be proved using Markov chains (following either the proof of Theorem 2.1 or the alternate sketched afterward).

Remarks and consequences for Contact Processes.

(i) Analogously to what we said in Section 1, Theorem 3.2 can be stated in terms of (conditional) positive association; namely for any $X \subseteq V$, the random variables $\eta(y) := \mathbf{1}_{\{s \to y\}}$, conditioned on the event $\{\eta \equiv 0 \text{ on } X\}$, are positively associated.

(ii) Taking $A = B = \Omega$ in Theorem 3.1 gives

$$\Pr(R_X)\Pr(R_Y) \leq \Pr(R_{X\cup Y})\Pr(R_{X\cap Y}). \tag{19}$$

(iii) Belitsky et al. ([2], Theorem 1.5) proved a special case of (19) involving a particular graph on the vertex set $\mathbb{Z}^2$. Their argument actually applies whenever V admits a partition $(V_0 = \{s\}) \cup V_1 \cup \cdots$ such that each edge is directed from V_{i-1} to V_i for some i and $X \cup Y$ is contained in some V_i, but does seem to depend essentially on these properties.

(iv) Much of [2] deals with the *contact process* on a countable set S. See [15] and [16] for background on this model; very briefly: Each site (individual) in S can be in either of the states 1 (ill and contagious) or 0 (healthy, noncontagious). Time is continuous, with $\eta_t(x)$ denoting the state of site x at time t. An infected site x becomes healthy at rate δ_x, and a healthy site becomes ill at rate $\sum_y \lambda(x,y)\eta_t(y)$. Here $\delta_x, x \in S$, and $\lambda(x,y), x,y \in S$, are the parameters of the model. They are assumed to be non-negative and, if S is infinite, to satisfy the following conditions (see [2]): $\sup_{x\in S}\delta(x) < \infty$ and $\sup_{x\in S}\sum_{y\in S}[\lambda(x,y)+\lambda(y,x)] < \infty$.

A nice aspect of the model is that it can be viewed in terms of percolation, via a graphical representation (see, e.g., [16], pages 32–34): being ill at some given time corresponds to the existence of an appropriate path in space-time. In fact, as is well-known, the process can, by time-discretization and standard limit arguments, be approximated by a directed percolation model on a finite graph. (See the subsection on correlation inequalities, in particular page 11, of [16] for the general idea of how correlation inequalities for collections of independent Bernoulli random variables can be extended to continuous-time interacting particle systems and page 65 of [16] for a concrete example for the contact process).

Combining this with the present results, one obtains, in a straightforward way, contact process analogues of the conditional association property stated in (i) above. In particular this gives the following theorem.

Theorem 3.3. *Suppose* $(\eta_t : t \geq 0)$ *is a contact process as above, with deterministic initial configuration* η_0. *Then, for each finite* $W \subset S$ *and* $t \geq 0$, *the collection* $(\eta_t(x) : x \in S \setminus W)$ *is conditionally positively associated given* $\{\eta_t \equiv 0 \text{ on } W\}$.

An example of Liggett [17] shows that if we instead condition on $\{\eta_t \equiv 1 \text{ on } W\}$, then the above positive association need not hold. In fact, building on the present ideas, we recently showed [4] that in dimension 1, conditioning on $\eta_t(0) = 1$ leads to *negative* dependence between the η_t-values of the sites to the left of 0 and those to the right of 0. (This contains a conjecture of Konno [14, Conjecture 4.5.2].)

Suppose now that at time 0 *each* site is ill. Let ν_t be the law of η_t $(= (\eta_t(x) : x \in S))$. It is well known (and follows easily from standard monotonicity arguments) that as $t \to \infty$, ν_t tends to a limit, called the *upper invariant measure* of the process and denoted ν. Clearly the preceding conditional association property for finite times extends to ν. So, for each finite $W \subset S$,

$$\nu(\cdot \mid \eta \equiv 0 \text{ on } W) \text{ is positively associated.} \tag{20}$$

This is a considerable strengthening of a conjecture of Konno ([14], Conjecture 3.4.13, itself a strengthening of a theorem of Harris [12]), which was proved in—and seems to have been the main motivation for—[2] (see inequality (1.3) in [2]), namely: for any $K, L \subseteq S$,

$$\nu(K \cap L)\nu(K \cup L) \geq \nu(K)\nu(L), \tag{21}$$

where, for $M \subseteq S$, $\nu(M) := \nu\{\eta : \eta \equiv 0 \text{ on } M\}$. (Of course (21) is easily obtained from (20): First note that it is sufficient to prove (21) for finite K, L. Next, for such K and L, take $W = K \cap L$ and consider the events $A = \{\eta \equiv 0 \text{ on } K \setminus L\}$, $B = \{\eta \equiv 0 \text{ on } L \setminus K\}$, and $Q = \{\eta \equiv 0 \text{ on } K \cap L\}$. Then (20) says that

$$\nu(AB \mid Q) \geq \nu(A \mid Q)\nu(B \mid Q),$$

which is equivalent to (21).)

Recently, Liggett and Steif (see [18], Section 2, in particular Theorem 2.1) have used our result (20) as a key ingredient in their proof that the upper invariant measure of the contact process dominates certain product measures. It is interesting to note that for that application (21) seems not to be sufficiently strong.

Random Initial Configurations

Recall that a measure h on 2^V (V any finite set, although here it will be $V(G)$) is called *log-supermodular* (lsm) if for any $X, Y \subseteq V$,

$$h(X \cap Y)\, h(X \cup Y) \geq h(X)\, h(Y).$$

For a probability measure μ on 2^V, define $h = h_\mu : 2^V \to \mathbb{R}$ by $h(A) = \mu(\mathbf{S} \cap A = \emptyset)$, where $\mathbf{S}$ is chosen according to μ. Say μ is *Rlsm* if h_μ is lsm. (The perhaps suboptimal name is for "Radon log-supermodular," the stated condition being equivalent to log-supermodulariy of the "Radon transform" of μ: $\overline{\mu}(X) = \sum\{\mu(Y) : Y \subseteq X\}$.) Following Liggett and Steif [18] we say that μ is *downward FKG* (dFKG) if the random variables $\mathbf{1}_{\{x \in \mathbf{S}\}}$ are conditionally positively associated given any event $\{\mathbf{S} \cap A = \emptyset\}$. (Note this implies Rlsm.)

Note that (21) says ν is Rlsm: the (standard) contact process notation of (21) recycles ν, using it for what we would call h_ν. Belitsky et al. actually prove something more general than (21): if the law of η_0 (no longer assumed deterministic) is Rlsm, then so is the law of η_t for each t. (They also prove a corresponding generalization of the special case of (19) mentioned in item (iii) following (19).) It is natural to ask—as did one of the referees of this paper—whether Theorem 3.3 can be similarly generalized. As we shall see, this and more is true.

Let $\mathbf{S}$ be a random subset of $V = V(G)$ with law μ, η a random edge configuration, chosen independently of $\mathbf{S}$, with law φ, and μ^* the law of $C_\eta(\mathbf{S}) := \{x \in V : \mathbf{S} \to_\eta x\}$ (where "$\to_\eta$" has the obvious meaning, and in particular, "$\to$" is the same as "$\leftrightarrow$" if G is undirected).

Theorem 3.4 (With notation as above).

(a) *if G is a mixed graph and φ an ordinary percolation measure, then each of the properties Rlsm, dFKG holds for μ^* if it holds for μ;*
(b) *if G is undirected and φ is a random cluster measure* (12) *with $q \geq 1$, then dFKG for μ implies dFKG for μ^*.*

"Two-sided" versions extending Theorems 1.5, 2.1, and 3.2 are also true, but we omit the slightly awkward statements. Part (a) gives the corresponding extension of Theorem 3.3:

Theorem 3.5. *Suppose $(\eta_t : t \geq 0)$ is a contact process, with μ_t the law of η_t. Then each of the properties Rlsm, dFKG holds for every μ_t provided it holds for μ_0.*

The Rlsm part of Theorem 3.5 is of course the aforementioned result of Belitsky et al. [2].

Partial proof of Theorem 3.4. We will prove only the "Rlsm" portion of (a). We include this partly because, as mentioned in the Introduction, it is the one result in the present paper that seems unlikely to be susceptible to the Markov chain approach. (It is also appealing as a natural extension to general graphs of the version of (19) proved in [2].) The dFKG portion (of (a)) can be proved either using the approach of Section 2 or by combining parts of the following argument with the ideas of Section 1; and (b) can again be proved using Markov chains. (These extensions are somewhat less straightforward than, for instance, what's needed for Theorem 3.1.)

Let $\Pr = \mu \times \varphi$, and for $X \subseteq V$, set $h^*(X) = \Pr(\mathbf{S} \nrightarrow_\eta X)$ (that is, $h^* = h_{\mu^*}$). Our task is to show that, for any $X, Y \subseteq V$,

$$h^*(X)h^*(Y) \le h^*(X \cup Y)h^*(X \cap Y). \tag{22}$$

Suppose first that $X \cap Y = \emptyset$. For $W \subseteq V$ we may write

$$h^*(W) = \sum_\eta \varphi(\eta)h(D_\eta(W)),$$

where $D_\eta(W) = \{v \in V : v \rightarrow_\eta W\}$ and $h = h_\mu$. Thus, we want

$$\sum_\eta \varphi(\eta)h(D_\eta(X)) \sum_\eta \varphi(\eta)h(D_\eta(Y)) \le \sum_\eta \varphi(\eta)h(D_\eta(X \cup Y)).$$

Set $f(\eta) = h(D_\eta(X))$, $g(\eta) = h(D_\eta(Y))$. Noting that $D_\eta(X \cup Y) = D_\eta(X) \cup D_\eta(Y)$, we have, since μ is Rlsm,

$$\begin{aligned} f(\eta)g(\eta) &\le h(D_\eta(X) \cup D_\eta(Y))h(D_\eta(X) \cap D_\eta(Y)) \\ &\le h(D_\eta(X \cup Y)); \end{aligned}$$

and then, since f, g are decreasing functions of η, Harris' inequality gives

$$\begin{aligned} \mathsf{E}f\,\mathsf{E}g \le \mathsf{E}fg &= \sum \varphi(\eta)f(\eta)g(\eta) \\ &\le \sum \varphi(\eta)h(D_\eta(X \cup Y)), \end{aligned}$$

which is what we want.

Now suppose $Z := X \cap Y \neq \emptyset$. In this case we proceed by induction on $|V|$. Let $\mathbf{A}$ be the set of vertices in $V \setminus Z$ from which Z can be reached by an open edge. We use p for the law of $\mathbf{A}$, noting that for any A, B,

$$p(A)p(B) = p(A \cap B)p(A \cup B). \tag{23}$$

Since

$$h(W) = \sum_A p(A)\Pr(\mathbf{S} \nrightarrow_\eta W|A)$$

(where $\Pr(\cdot|A)$ means $\Pr(\cdot|\mathbf{A} = A)$), it follows from (23) and the Ahlswede–Daykin Theorem that to establish (22) it's enough to show that (for any A, B)

$$\begin{aligned} &\Pr(\mathbf{S} \nrightarrow_\eta X|A)\Pr(\mathbf{S} \nrightarrow_\eta Y|B) \\ &\quad \le \Pr(\mathbf{S} \nrightarrow_\eta (X \cup Y)|A \cup B)\Pr(\mathbf{S} \nrightarrow_\eta (X \cap Y)|A \cap B). \end{aligned} \tag{24}$$

We would like to rewrite this in terms of the smaller graph $G' = G - Z$. Let φ' be the percolation measure that φ induces on G' and η' the corresponding configuration. Let Q be the event $\{\mathbf{S} \cap Z = \emptyset\}$. Then (e.g.),

$$\Pr(\mathbf{S} \nrightarrow_{\eta} X|A) = \mu(Q)(\mu \times \varphi')(\mathbf{S} \nrightarrow_{\eta'} X \cup A|Q),$$

and we can further rewrite

$$(\mu \times \varphi')(\mathbf{S} \nrightarrow_{\eta'} X \cup A|Q) = (\mu' \times \varphi')(\mathbf{S}' \nrightarrow_{\eta'} X \cup A),$$

where $\mathbf{S}'$ is the random subset of $V \setminus Z$ with law $\mu'(\cdot) = \mu(\cdot|Q)$.

So, setting $\Pr' = \mu' \times \varphi'$ and $h'(W) = \Pr'(\mathbf{S}' \nrightarrow_{\eta'} W)$, we will have (24) if we can show

$$h'(X \cup A)h'(Y \cup B) \leq h'((X \cup Y) \cup (A \cup B))h'((X \cap Y) \cup (A \cap B)). \tag{25}$$

This will follow from our inductive hypothesis.

Notice that μ' is again Rlsm: with $h = h_{\mu}$ as earlier,

$$\mu'(\mathbf{S}' \cap K = \emptyset) = \frac{\mu(\mathbf{S}' \cap K = \emptyset, Q)}{\mu(Q)} = \frac{h(K \cup Z)}{\mu(Q)};$$

so Rlsm for μ' is equivalent to the statement that for all $K, L \subseteq V \setminus Z$,

$$h(K \cup Z)h(L \cup Z) \leq h(K \cup L \cup Z)h((K \cap L) \cup Z),$$

which (since $(K \cap L) \cup Z = (K \cup Z) \cap (L \cup Z)$) is contained in Rlsm for μ.

So h' is lsm by induction, and this gives (25):

$$\begin{aligned} h'(X \cup A)h'(Y \cup B) &\leq h'((X \cup A) \cup (Y \cup B))h'((X \cup A) \cap (Y \cup B)) \\ &\leq h'((X \cup Y) \cup (A \cup B))h'((X \cap Y) \cup (A \cap B)) \end{aligned}$$

(where we used $(X \cup A) \cap (Y \cup B) \supseteq (X \cap Y) \cup (A \cap B)$). ∎

A Directed Version of Theorem 1.5

For a sensible generalization of Theorems 1.4 and 1.5 to the present setting we need a different substitute for $\{s \nleftrightarrow t\}$. It is easy to see that neither $\{s \nrightarrow t\}$ nor $\{s \nrightarrow t \nrightarrow s\}$ will do here (e.g., consider the graph on $\{s, t, v, a\}$ with (oriented) edges $(s, v), (t, v), (v, a)$ and events $A = \{s \rightarrow a\}$, $B = \{t \rightarrow a\}$); but there is another natural choice that does work, at least when we assume there are no undirected edges. Recall that $V(F)$ is the set of vertices incident with edges of F. In our formulation we restrict to finite graphs.

Theorem 3.6. *Assume G is a finite digraph in the usual sense (that is, all its edges are directed). Let s and t be (distinct) vertices and f and g functions of (C_s, C_t), each increasing in C_s and decreasing in C_t. Then, conditional on the event $Q := \{V(C_s) \cap V(C_t) = \emptyset\}$,*

$$\mathsf{E}fg \geq \mathsf{E}f\mathsf{E}g. \tag{26}$$

As mentioned in Section 1, this can be proved along the lines of either Theorem 1.5 or Theorem 2.1, with Theorem 3.2 a crucial ingredient in either case. Here we only give (sketchily) the second argument, leaving the reader to fill in the first (which, like the second, depends on Observation 3.7 below).

It's a little strange that we can so far prove Theorem 3.6 only in the absence of undirected edges, and we conjecture that it remains true without this restriction. (The difficulties in extending the proof below—those for the other version are essentially the same—are the (related) failures of Observation 3.7 and of the validity of the hypothesis (iii) when we come to apply Lemma 2.6.)

Proof. We will not repeat the proof of Theorem 2.1, but just indicate what changes are needed in the present situation.

The state space $\hat{\Omega}$ and transitions for our Markov chain are essentially as before. (Here we have chosen to say C_s, C_t rather than C_S, C_T, but as noted earlier (following Theorem 1.2) this really makes no difference.) Of course $\{V(C_s) \cap V(C_t) = \emptyset\}$ now replaces $\{S \not\leftrightarrow T\}$ as the conditioning event Q.

Let us write ψ for our (unconditioned) percolation measure, φ for our ψ conditioned on Q, and $\hat{\varphi}$ for the measure which φ induces on $\hat{\Omega}$. The argument here then follows that for Theorem 2.1 *verbatim* until, in proving positive association of the collection (15), we come to establishing conditions (ii) and (iii) of Lemma 2.6. For these we need the easily verified (but crucial):

Observation 3.7. *For $U \subseteq V$ the distribution of C_s is the same under φ conditioned on $\{V(C_t) = U\}$ as under ψ conditioned on $\{s \not\rightarrow U\}$ (and similarly with the roles of s and t reversed).*

Note now that (ii) is an immediate consequence of Observation 3.7 in conjunction with Theorem 3.2 (the relevant information from conditioning on W being just the value of C_s^{n-1} that results from W).

For (iii) we first observe that if W, W' are possible values of $(W_1, \ldots, W_a)$ with $W \leq W'$, and U, U' are the corresponding values of $V(C_s^{n-1})$, then according to Remark 2.8 we have $U \subseteq U'$. Moreover, by Observation 3.7, the distribution of $(Z_1, \ldots, Z_b)$ $(= (X_e^n : e \in E))$ given W is simply the distribution of the indicators of $E \setminus C_t$ under ψ conditioned on $\{t \not\rightarrow U\}$. So, writing $\hat{\varphi}_U$ for this distribution on $(X_e^n : e \in E)$, we need to show that $U' \supseteq U$ implies $\hat{\varphi}_{U'} \succ \hat{\varphi}_U$ (note that increasing C_t^n corresponds to *decreasing* the X_e^n's). This follows from Theorem 3.2: Note that $\hat{\varphi}_{U'}$ is the same as $\hat{\varphi}_U$ conditioned on $B := \{t \not\rightarrow U' \setminus U\}$. (More accurately, $\hat{\varphi}_{U'}$ is the distribution of the collection $(\mathbf{1}_{\{e \notin C_t\}} : e \in E)$ induced by $\psi(\cdot | t \not\rightarrow U)$ conditioned on B.) But then, since B is a decreasing event determined by C_t, Theorem 3.2 says that under $\hat{\varphi}_U$, B is positively correlated with any decreasing event determined by C_t; that is, $\hat{\varphi}_{U'} \succ \hat{\varphi}_U$. ∎

Remark. The choice of $\hat{\Omega}$ is a key to the preceding argument. For instance, taking the state space to be the analogue of that in the alternative proof of Theorem 2.1 sketched at the end of Section 2.1—namely $\{\omega \in \{0, 1\}^E : Q \text{ holds for } \omega\}$—gets in trouble because we lose some positive correlations, e.g., of events $\{\omega_e = 1\}$.

Note added in proof. In a recent preprint (Conditional Association and Spin Systems (2005)), Tom Liggett gives, among other things, a more direct proof of the dFKG part of Theorem 3.5.

ACKNOWLEDGMENTS

We thank Tom Liggett for drawing our attention, after publication of [3], to his paper [2] with Belitsky *et al.* and an anonymous referee for asking about the possibility of something like Theorem 3.4.

JvdB thanks the Isaac Newton Institute for its support and hospitality during a three-week visit in the fall of 2003.

REFERENCES

[1] R. Ahlswede and D. E. Daykin, An inequality for the weights of two families of sets, their unions and intersections, Z Wahrsch Verw Geb 43 (1978), 183–185.

[2] V. Belitsky, P. A. Ferrari, N. Konno and T. M. Liggett, A strong correlation inequality for contact processes and oriented percolation, Stoch Proc Appl 67 (1997), 213–225.

[3] J. van den Berg and J. Kahn, A correlation inequality for connection events in percolation, Ann Probab 29 (2001), 123–126.

[4] J. van den Berg, O. Häggström, and J. Kahn, Proof of a conjecture of N. Konno for the 1D contact process, preprint, 2005.

[5] B. Bollobás, Combinatorics, Cambridge University Press, Cambridge, 1986.

[6] L. Chayes, Percolation and ferromagnetism on Z^2: The q-state Potts cases, Stoch Proc Appl 65 (1996), 209–216.

[7] G. Grimmett, Percolation, 2nd ed., Springer-Verlag, 1999.

[8] G. Grimmett, Percolative problems, in Probability and Phase Transition, G. Grimmett (editor), Kluwer, Dordrecht, 1994, pp. 69–86.

[9] G. Grimmett, The random-cluster model, in Probability on Discrete Structures, H. Kesten (editor), Encyclopedia of Mathematical Sciences, Vol. 110, Springer-Verlag, 2003, pp. 73–123.

[10] O. Häggström, Positive correlations in the fuzzy Potts model, Ann Appl Probab 9 (1999), 1149–1159.

[11] O. Häggström, Finite Markov Chains and Algorithmic Applications, Cambridge University Press, Cambridge, 2002.

[12] T. E. Harris, Contact interactions on a lattice, Ann Probab 2 (1974), 969–988.

[13] R. Holley, Remarks on the FKG inequalities, Comm Math Phys 36 (1974), 227–231.

[14] N. Konno, Phase Transitions of Interacting Particle Systems, World Scientific, Singapore, 1994.

[15] T. M. Liggett, Interacting Particle Systems, Springer-Verlag, 1985.

[16] T. M. Liggett, Stochastic interacting systems: Contact, Voter and Exclusion Processes, Springer-Verlag, 1999.

[17] T. M. Liggett, Survival and coexistence in interacting particle systems, Probability and Phase Transition, Kluwer, Dordrecht, 1994, pp. 209–226.

[18] T. M. Liggett and J. E. Steif, Stochastic Domination: The Contact Process, Ising Models and FKG Measures, Annales Institut H. Poincaré, Probabilités et Statistiques, to appear.