# STRICT INEQUALITY FOR CRITICAL VALUES OF POTTS MODELS AND RANDOM-CLUSTER PROCESSES

C. E. Bezuidenhout, G. R. Grimmett, H. Kesten

Abstract. We prove that the critical value $\beta_c$ of a ferromagnetic Potts model is a *strictly* decreasing function of the strengths of interaction of the process. This is achieved in the (more) general context of the random-cluster representation of Fortuin and Kasteleyn, by deriving and utilizing a formula which generalizes the technique known in percolation theory as Russo's formula. As a byproduct of the method, we present a general argument for showing that, at any given point on the critical surface of a multiparameter process, the values of a certain critical exponent do not depend on the direction of approach of that point. Our results apply to all random-cluster processes satisfying the FKG inequality.

## 1. Introduction

In the general study of phase transitions, it is commonly the case that the value of the critical point is a *monotone* function of the strengths of local interactions (and hence of the structure of the underlying graph). For example, the critical probability of a percolation process cannot increase if new edges are added to the lattice. Similarly, the critical temperature of an Ising model cannot decrease if the intensity of any pair-interaction is increased. Such monotonicity, when true, is usually easy to prove. *Strict* monotonicity, on the other hand, presents new difficulties. A general approach to the problem of proving *strict* monotonicity of the critical point was described by Aizenman and Grimmett (1991). Using this approach, the latter authors obtained necessary and sufficient conditions for strict monotonicity in percolation, and in addition proved strict monotonicity for Ising models with pair-interactions. In this paper, we prove such a result for Potts models, thereby generalizing the latter conclusion of Aizenman and Grimmett (1991).

Rather than working directly with Potts models, we shall work with the corresponding random-cluster processes (sometimes known as Fortuin–Kasteleyn processes), and our main result may be stated loosely as follows. For any random-cluster process satisfying the FKG inequality (*i.e.*, with cluster-weighting factor $q$ satisfying $q \geq 1$) the critical value is a strictly monotone function of the bond-interactions. Since a Potts model with $q$ states (where $q$ is an integer satisfying $q \geq 2$) corresponds to a random-cluster process with parameter $q$, this result implies strict monotonicity for the critical temperatures of Potts models.

---

1991 *Mathematics Subject Classification.* 60K35, 82B20.

*Key words and phrases.* Potts model, Ising model, random-cluster process, critical temperature.

G.R.G. acknowledges support from Cornell University, and also partial support by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University. H.K. was supported in part by the N.S.F. through a grant to Cornell University.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Here is one example of the result applied to Potts models. Let $q$ be an integer satisfying $q \geq 2$, and write $\Sigma = \{1, 2, \ldots, q\}^{\mathbb{Z}^d}$; we think of $\sigma \in \Sigma$ as an allocation of a spin $\sigma(y)$ to each point $y \in \mathbb{Z}^d$, where $\sigma(y) \in \{1, 2, \ldots, q\}$. Let $\Sigma_\Lambda$ be the set of all elements $\sigma$ ($\in \Sigma$) that satisfy $\sigma(y) = 1$ if $y \notin \Lambda$, where $\Lambda$ is some finite box of $\mathbb{Z}^d$. Partition the set of nearest-neighbour pairs $\{u, v\}$ of $\mathbb{Z}^d$ into a union $\mathbb{E}_1 \cup \mathbb{E}_2$ of periodic classes (for example, $\mathbb{E}_1$ might contain all pairs of nearest neighbours which differ only in their first coordinates). Let $\mathbf{J} = (J_1, J_2)$ where $J_1, J_2 > 0$, and let

$$H_\Lambda(\sigma) = \sum_{\{u,v\}} J_{uv}\big(1 - \delta_{\sigma(u),\sigma(v)}\big), \quad \sigma \in \Sigma_\Lambda, \tag{1.1}$$

where $J_{uv} = J_1$ if $\{u, v\} \in \mathbb{E}_1$, and $J_{uv} = J_2$ otherwise; the sum runs over all nearest-neighbour pairs $\{u, v\}$, and $\delta_{i,j}$ is the Kronecker delta. We then consider weak limits as $\Lambda \uparrow \mathbb{Z}^d$ of the Potts probability measures

$$\pi_{\Lambda,\mathbf{J},q}(\sigma) = \frac{1}{Y_\Lambda} \exp\big(-\beta H_\Lambda(\sigma)\big), \quad \sigma \in \Sigma_\Lambda,$$

where $\beta > 0$ and $Y_\Lambda$ is the appropriate normalizing factor. There is a standard definition of the critical inverse-temperature $\beta_c(\mathbf{J}, q)$ for this model (see the discussion after Theorem 3 for more details). A special case of the forthcoming Theorem 2 is that $\beta_c(\mathbf{J}, q)$ is *strictly* decreasing in $\mathbf{J}$, *i.e.*, $\beta_c(\mathbf{J}, q) > \beta_c(\mathbf{J}', q)$ if $\mathbf{J} \leq \mathbf{J}'$ but $\mathbf{J} \neq \mathbf{J}'$. Similarly, the critical value $\beta_c$ is *strictly* diminished if we add edges to our graph in a periodic manner. As an example of this last operation, if the summation in (1.1) were to include all pairs $\{u, v\}$ which are either neighbours or next-nearest neighbours, each extra summand having a factor $J_3$ ($> 0$), then the corresponding value of $\beta_c$ would be strictly smaller than before.

For accounts of the relationship between Potts models and random-cluster processes, see Aizenman, Chayes, Chayes, and Newman (1988), Edwards and Sokal (1988), Grimmett (1993), and the references therein.

The general philosophy of Aizenman and Grimmett (1991) may be sketched in the following way. Let us consider some process of interaction on the hypercubic lattice $\mathbb{Z}^d$ with $d \geq 2$, parametrized by a family $\mathbf{J} = (J_i : 1 \leq i \leq m)$ of real variables; we think of this family as a vector which takes values in a parameter space $\mathcal{J}$ ($\subseteq \mathbb{R}^m$). Denote by $\theta = \theta(\mathbf{J})$ the 'order parameter' of the process (such as the percolation probability, or the magnetization). Suppose that $\theta$ is a non-decreasing function of $\mathbf{J}$, and that, for all $i, j \in \{1, 2, \ldots, m\}$, there exists a continuous function $\alpha_{ij}(\mathbf{J})$ taking values in $[0, \infty]$ such that

$$\frac{\partial \theta}{\partial J_i} \leq \alpha_{ij}(\mathbf{J}) \frac{\partial \theta}{\partial J_j}, \tag{1.2}$$

$$\alpha_{ij} \text{ is finite on } \operatorname{int}(\mathcal{J}), \tag{1.3}$$

where $\operatorname{int}(\mathcal{J})$ is the interior of $\mathcal{J}$. It follows that the critical surface of the process may be expressed in the form $\psi(\mathbf{J}) = 0$ for some function $\psi$ which is continuous and strictly increasing on $\operatorname{int}(\mathcal{J})$.

It is not practicable to prove (1.2) directly, since (1.2) begs the question of the differentiability of $\theta$ on $\mathrm{int}(\mathcal{J})$. However, it suffices to prove (1.2) for finite-box approximations to $\theta$. Let $(\Lambda(n) : n \geq 1)$ be an increasing sequence of boxes in $\mathbb{Z}^d$ such that $\Lambda(n) \uparrow \mathbb{Z}^d$ as $n \to \infty$, and let $\theta_\Lambda$ be the finite-box approximation to $\theta$ associated with a bounded box $\Lambda$. Instead of (1.2) it suffices to prove that

$$\frac{\partial \theta_{\Lambda(n)}}{\partial J_i} \leq \alpha_{ij}(\mathbf{J}) \frac{\partial \theta_{\Lambda(n)}}{\partial J_j}, \quad \text{for all } 1 \leq i, j \leq m \text{ and for all large } n. \tag{1.4}$$

See Aizenman and Grimmett (1991) for further details of the general argument.

In the case of percolation, (1.4) may be established by way of Russo's formula, which expresses the derivatives of $\theta_\Lambda$ in terms of the mean numbers of pivotal edges; this method was first used by Menshikov (1987). Whereas the inequality of (1.4) involves a *global* function (*i.e.*, the function $\theta_\Lambda$), this use of Russo's formula enables one to work with *local* estimates (*i.e.*, to perform calculations on regions which are small in size, uniformly in the value of $n$). In the case of the Ising model, Aizenman and Grimmett (1991) obtained (1.4) by using the 'random current' representation rather than a version of Russo's formula ('random current' formulae have been exploited very successfully by Aizenman (1982) and others). One of the two main steps in our proof of (1.4) for random-cluster processes is a generalization of Russo's formula which expresses the derivatives as sums of covariances (*cf.* Proposition 4 and equation (4.1)). These covariances are non-negative, and are 'global' quantities. A principal novelty of the present paper is a method for expressing these covariances in terms of quantities *which are still non-negative and which have a 'local' nature*; in this way, we shall compare the contributions appearing in the two sides of (1.4). The mechanism for doing this comparison is provided by Proposition 5.

In Section 2 we present notation for random-cluster and Potts processes, and we state our main results. This is followed in Section 3 by the new form of Russo's formula, and in Section 4 by the remaining proofs. We note that inequality (1.4) has some implications for critical exponents, given at the end of Section 2. Finally, many of the arguments described here may be applied also to processes with many-body interactions; this will be the subject of a forthcoming article, Grimmett (1992).

## 2. Notation and results

We begin by defining the random-cluster process. As usual, $\mathbb{Z}^d$ is the set of all $d$–vectors of integers, where $d \geq 2$. For $x \in \mathbb{Z}^d$, we generally write $x = (x_1, x_2, \ldots, x_d)$. Let $K$ be a *finite* subset of $\mathbb{Z}^d$, not containing the origin 0. We place an edge between two points $x, y \in \mathbb{Z}^d$ if and only if either $x - y \in K$ or $y - x \in K$; such an edge is denoted by $\langle x, y \rangle$, and the ensuing graph, called the *interaction graph*, is denoted by $G$. We make the following assumptions about $K$:

$$\text{if } x \in K \text{ then } -x \notin K, \tag{2.1}$$

$$G \text{ is connected.} \tag{2.2}$$

Clearly the definition of $G$ is unchanged by assumption (2.1), whilst assumption (2.2) is more important. We write $\mathbb{E}$ for the set of edges of $G$, so that $G = (\mathbb{Z}^d, \mathbb{E})$.

A configuration $\omega = (\omega(e) : e \in \mathbb{E}) \in \{0, 1\}^{\mathbb{E}}$ is an allocation of 0 or 1 to each edge of $G$, and $\Omega$ denotes the set of all such configurations. With $\Omega$ we associate the $\sigma$-field $\mathcal{F}$

generated by the finite-dimensional cylinders. For $\omega \in \Omega$, we write $\eta(\omega) = \{e \in \mathbb{E} : \omega(e) = 1\}$, the set of 'open' edges. We shall sometimes confuse the roles of $\omega$ and $\eta(\omega)$, and may speak of '$\omega$ containing the edge $f$' when we mean that $f \in \eta(\omega)$. For subsets $A$ and $B$ of $\mathbb{Z}^d$, and configurations $\omega$, we write $A \leftrightarrow B$ if there is a path of $(\mathbb{Z}^d, \eta(\omega))$ joining some point in $A$ to some point in $B$.

Let $\Lambda$ be a finite box of $\mathbb{Z}^d$, *i.e.*,

$$\Lambda = \prod_{i=1}^{d} [x_i, y_i], \quad \text{for some } x, y \in \mathbb{Z}^d.$$

The *boundary* $\partial\Lambda$ of $\Lambda$ is the set of points $x \in \Lambda$ for which there exists $y \notin \Lambda$ satisfying $\langle x, y\rangle \in \mathbb{E}$. We write $\mathbb{E}_\Lambda$ for the set of edges of $G$ both of whose endpoints lie in $\Lambda$, $\Omega_\Lambda$ for the subset of $\Omega$ containing all configurations $\omega$ satisfying $\omega(e) = 1$ if $e \notin \mathbb{E}_\Lambda$, and we write $\mathcal{F}_\Lambda$ for the $\sigma$-field of all subsets of $\Omega_\Lambda$.

Let $\mathbf{p} = (p_x : x \in K)$ be a real vector satisfying $0 \le p_x \le 1$ for all $x \in K$, and let $q$ be a positive real number. For $e \in \mathbb{E}$, where $e = \langle u, u + x\rangle$ and $x \in K$, we define $p_e = p_x$. The parameters $\mathbf{p}, q$ specify a random-cluster process in the following way. For any finite box $\Lambda$, we let $\phi_{\Lambda,\mathbf{p},q}$ be the probability measure on $(\Omega_\Lambda, \mathcal{F}_\Lambda)$ given by

$$\phi_{\Lambda,\mathbf{p},q}(\omega) = \frac{1}{Z_\Lambda}\left\{\prod_{e\in\mathbb{E}_\Lambda} p_e^{\omega(e)}(1-p_e)^{1-\omega(e)}\right\}q^{k(\omega)}, \quad \omega \in \Omega_\Lambda, \tag{2.3}$$

where $k(\omega)$ is the number of components of the graph $(\mathbb{Z}^d, \eta(\omega))$, and

$$Z_\Lambda = \sum_{\omega\in\Omega_\Lambda}\left\{\prod_{e\in\mathbb{E}_\Lambda} p_e^{\omega(e)}(1-p_e)^{1-\omega(e)}\right\}q^{k(\omega)}$$

is the appropriate normalizing constant. Note that the measure $\phi_{\Lambda,\mathbf{p},q}$ contains a 'product measure' term together with a 'Radon–Nikodym derivative' $q^{k(\omega)}/Z_\Lambda$. It is interesting to note that much of the contents of this paper may be generalized to situations in which this 'derivative' is replaced by one of a large family of non-constant positive functions.

If $q \ge 1$ then $\phi_{\Lambda,\mathbf{p},q}$ satisfies the FKG inequality (see Aizenman, Chayes, Chayes, and Newman (1988) for explanations of this and other general statements). The case $q = 1$ corresponds to percolation and has been considered in depth by Aizenman and Grimmett (1991). *We assume henceforth that* $q > 1$, and we shall occasionally suppress explicit reference to $q$. It may be seen that the limit measure

$$\phi_{\mathbf{p},q} = \lim_{\Lambda\uparrow\mathbb{Z}^d} \phi_{\Lambda,\mathbf{p},q}$$

exists in the weak sense, where the limit is understood to be taken along any increasing sequence of finite boxes $\Lambda$. For any such box $\Lambda$, let

$$\theta_\Lambda(\mathbf{p}, q) = \phi_{\Lambda,\mathbf{p},q}(0 \leftrightarrow \infty)$$

where $\{0 \leftrightarrow \infty\}$ is the event that the origin is in an infinite component; it follows from the definition of $\phi_{\Lambda,\mathbf{p},q}$ that $\theta_\Lambda(\mathbf{p}, q) = \phi_{\Lambda,\mathbf{p},q}(0 \leftrightarrow \partial\Lambda)$. It is the case that

$$\theta_\Lambda(\mathbf{p}, q) \downarrow \theta(\mathbf{p}, q) \quad \text{as } \Lambda \uparrow \mathbb{Z}^d$$

where

$$\theta(\mathbf{p}, q) = \phi_{\mathbf{p},q}(0 \leftrightarrow \infty).$$

The quantities $\theta_\Lambda(\mathbf{p}, q)$ and $\theta(\mathbf{p}, q)$ are non-decreasing functions of $\mathbf{p}$ when $q \geq 1$. We sometimes write $\theta(\cdot)$ for $\theta(\cdot, q)$.

Let $\Lambda(n) = [-n, n]^d$, and let $\theta_n(\mathbf{p}, q) = \theta_{\Lambda(n)}(\mathbf{p}, q)$. The following is our principal calculation.

**Theorem 1.** *Let $q > 1$. There exists a continuous function $\alpha : (0,1)^K \to (0, \infty)$ and a positive integer $N$ such that*

$$\frac{\partial \theta_n}{\partial p_x} \leq \alpha(\mathbf{p}) \frac{\partial \theta_n}{\partial p_y} \tag{2.4}$$

*for all $\mathbf{p} \in (0,1)^K$, all $n \geq N$, and all pairs $x, y \in K$.*

Note that $\alpha$ depends generally on the value of $q$. It is curious to note that our proof of Theorem 1 is invalid when $q = 1$ (the problem is that the forthcoming expression (3.15) is insufficient to make (3.14) strictly positive). Nevertheless, the *conclusion* of the theorem is valid if $q = 1$; see Aizenman and Grimmett (1991).

Theorem 1 has various implications for the nature of the phase transitions of random-cluster processes. The first of these concerns the monotonicity of the critical point. For $\mathbf{p} \in (0,1)^K$ let

$$\gamma_c(\mathbf{p}) = \sup\{\gamma \geq 0 : \theta(\gamma\mathbf{p}, q) = 0\}.$$

The set $\mathrm{C}_q = \{\mathbf{p} \in (0,1)^K : \gamma_c(\mathbf{p}) = 1\}$ is called the *critical surface* of the process, and the set $\mathrm{SP}_q = \{\mathbf{p} \in (0,1)^K : \gamma_c(\mathbf{p}) < 1\}$ the *supercritical* region. We shall prove that $\mathrm{C}_q$ has no 'flat regions' in the sense that, if $\mathbf{p} \in \mathrm{C}_q$ and $\mathbf{p} \leq \mathbf{p}'$ but $\mathbf{p} \neq \mathbf{p}'$, then $\theta(\mathbf{p}', q) > 0$. This, and slightly more, follows from the next theorem. Here and later, a vector $\mathbf{w} \in (0,1)^K$ is expressed as $\mathbf{w} = (w_x : x \in K)$; we write $\mathbf{p} \leq \mathbf{p}'$ if $p_x \leq p'_x$ for all $x \in K$, and $\mathbf{p} < \mathbf{p}'$ if $p_x < p'_x$ for all $x \in K$.

Let $U$ be the unit sphere of $\mathbb{R}^K$, with the topology induced from the euclidean topology on $\mathbb{R}^K$. We call an open subset $V$ of $U$ *full* if it contains the set of all points of $U$ having non-negative coordinates (*i.e.*, all points in the first orthant of $\mathbb{R}^K$).

**Theorem 2.** *Let $q > 1$ and $\mathbf{p} \in \mathrm{C}_q$. There exist positive constants $\delta$, $\nu$, and $\epsilon_0$, together with a full subset $V$ of $U$, such that*

$$\theta(\mathbf{p} + \nu\epsilon\mathbf{e}) \leq \theta(\mathbf{p} + \epsilon\mathbf{f}) \leq \theta(\mathbf{p} + \delta\epsilon\mathbf{e}) \tag{2.5}$$

*for all $0 < \epsilon < \epsilon_0$ and all $\mathbf{e}, \mathbf{f} \in V$.*

Note that $\delta$, $\nu$, $\epsilon_0$, and $V$ may depend on the value of $\mathbf{p}$. Conclusions similar to Theorem 2 have been derived for Ising models by Aizenman and Fernández (1986); see their Lemmas 2.2 and B.2.

It is easily seen from the definition of $\mathrm{C}_q$ that $\theta(\mathbf{p} + \epsilon\mathbf{f}) > 0$ when $\mathbf{p} \in \mathrm{C}_q$ and $\epsilon > 0$, $\mathbf{f} > \mathbf{0}$ (*i.e.*, $f_x > 0$ for all $x \in K$). The first consequence of Theorem 2 is therefore the existence, for each $\mathbf{p} \in \mathrm{C}_q$, of a positive real $\epsilon_1$ and a full subset $V$ of $U$ such that $\mathbf{p} + \epsilon V \subseteq \mathrm{SP}_q$ for $0 < \epsilon < \epsilon_1$. A result of the same general type was first obtained by Menshikov (1987) in the case of percolation.

Before stating the next theorem, we introduce one piece of terminology. A function $\mathbf{p} : \mathbb{R}^m \to \mathbb{R}^n$ is called *increasing* if $\mathbf{p}(\mathbf{x}) \le \mathbf{p}(\mathbf{y})$ whenever $\mathbf{x} \le \mathbf{y}$, and *strictly increasing* if, for all pairs $\mathbf{x}, \mathbf{y}$ with $\mathbf{x} \le \mathbf{y}$ but $\mathbf{x} \ne \mathbf{y}$, it is the case that $\mathbf{p}(\mathbf{x}) \le \mathbf{p}(\mathbf{y})$ but $\mathbf{p}(\mathbf{x}) \ne \mathbf{p}(\mathbf{y})$.

Theorem 2 has a further important consequence obtained by narrowing the field of application to a one-parameter family of processes. Let $\mathbf{J} = (J_x : x \in K) \in (0,\infty)^K$ and let $\beta > 0$; we think of $\mathbf{J}$ as describing the strengths of the bond-interactions, and $\beta$ as being the parameter (or 'inverse temperature') of the process. Suppose that $\mathbf{p}$ is a function of the pair $(\mathbf{J}, \beta)$, say $\mathbf{p} = \mathbf{p}(\mathbf{J}, \beta)$, a continuous increasing function of $(\mathbf{J}, \beta)$. The critical value of $\beta$ is given by

$$\beta_c(\mathbf{J}, q) = \sup\bigl\{\beta : \theta\bigl(\mathbf{p}(\mathbf{J}, \beta), q\bigr) = 0\bigr\}. \tag{2.6}$$

An example of particular importance is the case when

$$p_x = 1 - e^{-\beta J_x}, \quad \text{for } x \in K. \tag{2.7}$$

**Theorem 3.** *Suppose that the function $\mathbf{p}(\cdot,\cdot)$ is continuous and increasing, and that, for each $\beta > 0$, the function $\mathbf{p}(\cdot, \beta)$ is strictly increasing on $(0,\infty)^K$. If $q > 1$, the critical value $\beta_c(\mathbf{J}, q)$ is a strictly decreasing function of the pair-interactions $\mathbf{J}$.*

We now sketch the application of Theorem 3 to Potts models. Let $q$ be an integer satisfying $q \ge 2$, and let $\mathbf{J} = (J_x : x \in K)$ be a vector of positive reals. Write $\Sigma = \{1, 2, \dots, q\}^{\mathbb{Z}^d}$, and think of $\sigma \in \Sigma$ as being an allocation of a spin $\sigma(y)$ to each point $y \in \mathbb{Z}^d$, where $\sigma(y) \in \{1, 2, \dots, q\}$. Let $\Sigma_\Lambda$ be the set of spin-vectors $\sigma$ satisfying $\sigma(y) = 1$ if $y \notin \Lambda \setminus \partial\Lambda$. We consider the hamiltonian

$$H_\Lambda(\sigma) = \sum_{e \in \mathbb{E}} J_e\bigl(1 - \delta_e(\sigma)\bigr), \quad \sigma \in \Sigma_\Lambda,$$

where

$$\delta_{\langle u,v\rangle}(\sigma) = \begin{cases} 1 & \text{if } \sigma(u) = \sigma(v) \\ 0 & \text{otherwise,} \end{cases}$$

and $J_{\langle u,u+x\rangle} = J_x$ for $x \in K$. We are interested in the probability measure $\pi_{\Lambda,\mathbf{J},q}$ given by

$$\pi_{\Lambda,\mathbf{J},q}(\sigma) = \frac{1}{Y_\Lambda} \exp\bigl(-\beta H_\Lambda(\sigma)\bigr), \quad \sigma \in \Sigma_\Lambda,$$

where $\beta > 0$, and $Y_\Lambda$ is the appropriate normalizing factor. The weak limit

$$\pi^1_{\mathbf{J},q} = \lim_{\Lambda \uparrow \mathbb{Z}^d} \pi_{\Lambda,\mathbf{J},q}$$

exists, and is independent of the choice of increasing sequence of boxes $\Lambda$; see Aizenman, Chayes, Chayes, and Newman (1988) for details. The order parameter is given by

$$M(\mathbf{J}, \beta, q) = \frac{q}{q-1}\left(\pi^1_{\mathbf{J},q}\bigl(\{\sigma : \sigma(0) = 1\}\bigr) - \frac{1}{q}\right)$$

and the critical point by

$$\beta_c^{\mathrm{P}}(\mathbf{J}, q) = \sup\bigl\{\beta : M(\mathbf{J}, \beta, q) = 0\bigr\}.$$

Using the results of Fortuin and Kasteleyn (see Aizenman, Chayes, Chayes, and Newman (1988) and Edwards and Sokal (1988)), the order parameter of the Potts model is related to that of a corresponding random-cluster process by the equation

$$\theta\big(\mathbf{p}(\mathbf{J},\beta),q\big) = M(\mathbf{J},\beta,q) \tag{2.8}$$

where the function $\mathbf{p}$ is given by equation (2.7). It follows that $\beta_c^{\mathrm{P}}(\mathbf{J},q) = \beta_c(\mathbf{J},q)$, whence Theorem 3 implies that $\beta_c^{\mathrm{P}}(\mathbf{J},q)$ is strictly decreasing in $\mathbf{J}$. As in Aizenman and Grimmett (1991), the derivatives of $\beta_c^{\mathrm{P}}(\mathbf{J},q)$ with respect to the $J_x$ may be bounded above and below by continuous and strictly negative functions.

We turn now to the matter of critical exponents. The following discussion is based solely upon the inequalities of Theorem 1; it applies equally to any process of interaction whose order parameter $\theta$ is monotone in the strengths of pair-interactions and whose finite-box approximations satisfy the conclusions of that theorem. In the light of (2.8) and the results of Aizenman and Grimmett (1991), the discussion is valid for all percolation, Ising, and Potts models.

Suppose that $\mathbf{p} \in \mathrm{C}_q$, the critical surface of the random-cluster process with cluster-weighting factor $q$. The behaviour of $\theta(\cdot) = \theta(\cdot,q)$ in the neighbourhood of $\mathbf{p}$ may be described in terms of a critical exponent usually denoted by $\beta$. In the multiparameter setting, the value of this critical exponent depends in general upon the direction along which the point $\mathbf{p}$ is approached. That is to say, there is a family of exponents $B_{\mathbf{p}} = (\beta_{\mathbf{p}}(\mathbf{e}) : \mathbf{e} \in U)$ where $U$ is the set of unit vectors in $\mathbb{R}^K$; the exponent $\beta_{\mathbf{p}}(\mathbf{e})$ is defined by

$$\theta(\mathbf{p}+\epsilon\mathbf{e}) - \theta(\mathbf{p}) \approx \epsilon^{\beta_{\mathbf{p}}(\mathbf{e})} \quad \text{as } \epsilon \downarrow 0. \tag{2.9}$$

The asymptotic relation $f(\epsilon) \approx g(\epsilon)$ should be interpreted in some reasonable way, perhaps as $\log f(\epsilon)/\log g(\epsilon) \to 1$ as $\epsilon \downarrow 0$.

We restrict ourselves to directions $\mathbf{e}$ such that $\mathbf{p}+\epsilon\mathbf{e}$ lies in the supercritical regime for all small $\epsilon$. For such a pair $\mathbf{p}$, $\mathbf{e}$, it is expected that $\beta_{\mathbf{p}}(\mathbf{e})$ takes on a value which is independent of $\mathbf{e}$, say $\beta_{\mathbf{p}}(\mathbf{e}) = \beta_{\mathbf{p}}$. Theorem 2 helps us to go some way towards proving the latter statement. Inequality (2.5) implies that $\beta_{\mathbf{p}}(\mathbf{e}) = \beta_{\mathbf{p}}(\mathbf{f})$, and hence that $\beta_{\mathbf{p}}(\mathbf{e})$ is constant for $\mathbf{e}$ lying in the full set $V$ of unit vectors. In particular, since $V$ includes all unit vectors lying parallel to the axes (in the directions of increasing coordinate values), we obtain as a special case that the *axial* critical exponents (*i.e.*, those obtained by increasing exactly one $p_x$ for $x \in K$) are equal, at any given $\mathbf{p} \in \mathrm{C}_q$. Results of this kind have been obtained by Aizenman and Fernández (1986) for Ising models, and by Wierman (1992) for percolation processes.

Arguments similar to those of this paper are valid in principle for random-cluster processes with other boundary conditions such as 'free' boundary conditions, rather than the 'wired' conditions implicit in the definition of $\Omega_\Lambda$ above (2.3). Also, similar arguments are valid for certain other choices for the order parameter than $\theta(\cdot,q)$. In the case of percolation, one may obtain in like fashion results for the mean size of the open cluster at the origin, with implications for the family of critical exponents corresponding to the quantity usually denoted by $\gamma$. The method should be similarly applicable to random-cluster processes, but our current state of knowledge is inadequate for this.

Finally, we note that the results given above may be obtained also in the setting of many-body interactions, rather than pair-interactions, although the proofs are rather more complicated. Also, the methods of the next sections may be applied to graphs $(\mathbb{Z}^d, \mathbb{E})$ which are not vertex-transitive, but which are generated by periodic families of translations of edges belonging to some 'base set'. Such extensions will be the subject of the forthcoming paper of Grimmett (1992).

## 3. Russo's formula and time-evolutions for random-cluster processes

As before, $\Lambda$ is a finite box of $\mathbb{Z}^d$ and $\mathbb{E}_\Lambda$ is the set of edges (of $\mathbb{E}$) both of whose endpoints lie in $\Lambda$. On $(\Omega_\Lambda, \mathcal{F}_\Lambda)$ we construct a probability measure which is slightly more general than that defined by (2.3). Let $\widetilde{\mathbf{p}} = (\widetilde{p}_e : e \in \mathbb{E}_\Lambda)$ be a vector of reals each of which satisfies $0 < \widetilde{p}_e < 1$, and define the measure $\widetilde{\phi}$ on $(\Omega_\Lambda, \mathcal{F}_\Lambda)$ by

$$\widetilde{\phi}(\omega) = \frac{1}{Z}\Big\{\prod_{e\in\mathbb{E}_\Lambda} \widetilde{p}_e^{\,\omega(e)}(1-\widetilde{p}_e)^{1-\omega(e)}\Big\}q^{k(\omega)}, \quad \omega \in \Omega_\Lambda, \tag{3.1}$$

where $Z$ is the appropriate normalizing factor. The measure $\phi_{\Lambda,\mathbf{p},q}$ of (2.3) is retrieved by setting $\widetilde{p}_{\langle u,u+x\rangle} = p_x$ if $x \in K$. Thus, in (3.1), each edge $e$ has an associated independent variable $\widetilde{p}_e$.

We write $I_A$ for the indicator function of an event $A$; for $e \in \mathbb{E}_\Lambda$, we write $I_e$ for the indicator function of the event $\{\omega(e) = 1\}$, *i.e.*, $I_e(\omega) = \omega(e)$ for $\omega \in \Omega$.

**Proposition 4.** *Let $0 < q < \infty$, and let $A \in \mathcal{F}_\Lambda$. Then*

$$\frac{\partial}{\partial \widetilde{p}_e}\widetilde{\phi}(A) = \frac{1}{\widetilde{p}_e(1-\widetilde{p}_e)}\operatorname{cov}(I_e, I_A) \quad \textit{for all } e \in \mathbb{E}_\Lambda. \tag{3.2}$$

Here, 'cov' denotes covariance relative to the probability measure $\widetilde{\phi}$. In the case $q = 1$, equation (3.2) is a version of Russo's formula for percolation, but in a form first discovered by Barlow and Proschan (1965, p. 10); see Grimmett (1989, p. 38). It is notable that (3.2) is valid for all values of $q$. Indeed, (3.2) is valid for any measure of the form of (3.1) with the term $q^{k(\omega)}$ replaced by a general function of $\omega$ which does not depend on $\widetilde{p}_e$.

*Proof.* We may write

$$\widetilde{\phi}(A) = \frac{1}{Z}\sum_{\omega\in\Omega_\Lambda} I_A(\omega)\widetilde{p}_e^{\,\omega(e)}(1-\widetilde{p}_e)^{1-\omega(e)}\alpha(\omega), \tag{3.3}$$

where

$$\alpha(\omega) = \Big\{\prod_{f\in\mathbb{E}_\Lambda\setminus\{e\}} \widetilde{p}_f^{\,\omega(f)}(1-\widetilde{p}_f)^{1-\omega(f)}\Big\}q^{k(\omega)};$$

note that $\alpha$ is not a function of $\widetilde{p}_e$. Differentiate (3.3) to obtain

$$\begin{aligned}\frac{\partial}{\partial \widetilde{p}_e}\widetilde{\phi}(A) &= \frac{1}{Z}\sum_{\omega} I_A(\omega)\widetilde{p}_e^{\,\omega(e)}(1-\widetilde{p}_e)^{1-\omega(e)}\alpha(\omega)\left(\frac{\omega(e)}{\widetilde{p}_e} - \frac{1-\omega(e)}{1-\widetilde{p}_e}\right)\\ &\qquad - \frac{1}{Z^2}\frac{\partial Z}{\partial \widetilde{p}_e}\sum_{\omega} I_A(\omega)\widetilde{p}_e^{\,\omega(e)}(1-\widetilde{p}_e)^{1-\omega(e)}\alpha(\omega)\\ &= E\left(I_A\left(\frac{I_e - \widetilde{p}_e}{\widetilde{p}_e(1-\widetilde{p}_e)}\right)\right) - \frac{1}{Z}\frac{\partial Z}{\partial \widetilde{p}_e}\widetilde{\phi}(A),\end{aligned} \tag{3.4}$$

where $E$ denotes expectation relative to the probability measure $\widetilde{\phi}$. Set $A = \Omega_\Lambda$ to obtain that

$$0 = E\left(\frac{I_e - \widetilde{p}_e}{\widetilde{p}_e(1-\widetilde{p}_e)}\right) - \frac{1}{Z}\frac{\partial Z}{\partial \widetilde{p}_e}. \tag{3.5}$$

We eliminate $Z^{-1}\partial Z/\partial\widetilde{p}_e$ from (3.4) and (3.5), and we arrive at (3.2). □

Suppose that $q \geq 1$ and let $A$ be an increasing event of $(\Omega_\Lambda, \mathcal{F}_\Lambda)$ (that is, if $\omega \in A$ and $\omega \leq \omega'$ then $\omega' \in A$). It will be useful to have recourse to a representation for $\partial\widetilde{\phi}(A)/\partial\widetilde{p}_e$ containing only non-negative terms. Using (3.2), we may write

$$\frac{\partial}{\partial \widetilde{p}_e}\widetilde{\phi}(A) = \frac{\widetilde{\phi}(A)}{\widetilde{p}_e(1-\widetilde{p}_e)}\big\{\widetilde{\phi}(\omega(e)=1 \mid A) - \widetilde{\phi}(\omega(e)=1)\big\}. \tag{3.6}$$

The term in braces is non-negative (by the FKG inequality) and may be represented in terms of coupled time-evolutions of the random-cluster process. We show next how this may be done.

We shall construct certain Markov chains on the state space $\Omega_\Lambda$. Let $\mu$ be a probability measure on $(\Omega_\Lambda, \mathcal{F}_\Lambda)$, and let $G : \Omega_\Lambda \times \Omega_\Lambda \to [0,\infty)$ be the generator of a Markov chain (*i.e.*, for $\omega \neq \omega'$, $G(\omega,\omega')$ represents the jump-rate of the chain from $\omega$ to $\omega'$, while the diagonal element $G(\omega,\omega)$ is chosen in such a way that $\sum_{\omega'} G(\omega,\omega') = 0$). If the pair $(\mu, G)$ satisfies the balance equations

$$\mu(\omega)G(\omega,\omega') = \mu(\omega')G(\omega',\omega) \quad \text{for all } \omega,\omega' \in \Omega_\Lambda, \tag{3.7}$$

then the chain is reversible with stationary measure $\mu$.

For $\omega \in \Omega_\Lambda$ and $f \in \mathbb{E}_\Lambda$, we define the configurations $\omega_f$ and $\omega^f$ by

$$\omega_f(g) = \begin{cases} \omega(g) & \text{if } g \neq f \\ 0 & \text{if } g = f, \end{cases}$$
$$\omega^f(g) = \begin{cases} \omega(g) & \text{if } g \neq f \\ 1 & \text{if } g = f. \end{cases}$$

We denote by $D_f(\omega)$ the indicator function of the event that the endpoints of $f$ are in different components of the graph $(\mathbb{Z}^d, \eta(\omega_f))$, *i.e.*, the event that the endpoints of $f$ are not connected by any path of edges in $\eta(\omega) \setminus \{f\}$.

There are two instances of (3.7) of importance for (3.6). When $\mu = \widetilde{\phi}$, equations (3.7) are satisfied when we take for $G$ the function $H : \Omega_\Lambda \times \Omega_\Lambda \to [0,\infty)$ given by

$$H(\omega_f, \omega^f) = 1, \tag{3.8}$$

$$H(\omega^f, \omega_f) = \frac{\widetilde{\phi}(\omega_f)}{\widetilde{\phi}(\omega^f)} = \frac{1-\widetilde{p}_f}{\widetilde{p}_f} q^{D_f(\omega)} \tag{3.9}$$

for $\omega \in \Omega_\Lambda$ and $f \in \mathbb{E}_\Lambda$, where $H(\omega,\omega')$ is set to 0 for other pairs $\omega,\omega'$ with $\omega \neq \omega'$.

In the second instance, $\mu = \widetilde{\phi}(\cdot \mid A)$, and in this case (3.7) is satisfied by the generator $H^A$ given by

$$H^A(\omega,\omega') = H(\omega,\omega')I_A(\omega \wedge \omega') \quad \text{if } \omega \neq \omega' \tag{3.10}$$

where $\omega \wedge \omega'(g) = \min\{\omega(g), \omega'(g)\}$.

These two instances give rise to Markov chains which can be constructed on the same sample space in such a way that the first 'lies beneath' the second. To this end we argue as follows.

Let $\Theta_\Lambda$ be the set of all pairs $(\pi, \omega)$ of configurations in $\Omega_\Lambda$ satisfying $\pi \le \omega$. We define $J : \Theta_\Lambda \times \Theta_\Lambda \to [0, \infty)$ by

$$J(\pi_f, \omega; \pi^f, \omega^f) = 1, \tag{3.11}$$

$$J(\pi, \omega^f; \pi_f, \omega_f) = H^A(\omega^f, \omega_f), \tag{3.12}$$

$$J(\pi^f, \omega^f; \pi_f, \omega^f) = H(\pi^f, \pi_f) - H^A(\omega^f, \omega_f), \tag{3.13}$$

for all $(\pi, \omega) \in \Theta_\Lambda$ and $f \in \mathbb{E}_\Lambda$; all other off-diagonal values of $J$ are set to 0. Equation (3.11) specifies that, for $\pi \in \Omega_\Lambda$ and $f \in \mathbb{E}_\Lambda$, the edge $f$ is acquired by $\pi$ (if it does not already contain it) at rate 1; any edge so acquired is added also to $\omega$ if it does not already contain it. Equation (3.12) specifies that, for $\omega \in \Omega_\Lambda$ and $f \in \eta(\omega)$, the edge $f$ is removed from $\omega$ (and also from $\pi$ if $f \in \eta(\pi)$) at the rate given by (3.10). For $f \in \eta(\pi)$ ($\subseteq \eta(\omega)$), there is an additional rate at which $f$ is removed from $\pi$ but not from $\omega$. Note that this additional rate is indeed non-negative, since

$$H(\pi^f, \pi_f) - H^A(\omega^f, \omega_f) = \frac{1 - \widetilde{p}_f}{\widetilde{p}_f} \left\{ q^{D_f(\pi)} - q^{D_f(\omega)} I_A(\omega_f) \right\} \ge 0, \tag{3.14}$$

by (3.9) and (3.10); remember that $q \ge 1$ and $D_f(\omega) \le D_f(\pi)$ when $\pi \le \omega$. This additional rate is *strictly* positive if and only if

$$\text{either } \omega_f \notin A, \text{ or } \omega_f \in A,\ D_f(\pi) = 1, \text{ and } D_f(\omega) = 0. \tag{3.15}$$

We note also that the transitions referred to in (3.11)–(3.13) take pairs $(\pi, \omega)$ with $\pi \le \omega$ onto new pairs $(\pi', \omega')$ with $\pi' \le \omega'$. Therefore, the function $J$ is the generator of a Markov chain $(X_t, Y_t)_{t \ge 0}$ taking values in $\Theta_\Lambda$. We assume henceforth that $(X_t, Y_t)_{t \ge 0}$ is such a chain, and we set $(X_0, Y_0) = (\mathbf{0}, \mathbf{1})$ where $\mathbf{i}$ ($\in \Omega_\Lambda$) is the configuration which takes the value $i$ on every edge in $\mathbb{E}_\Lambda$. We write $P$ for the appropriate probability measure on the set of sample paths of this process.

By further examination of (3.11)–(3.13), it is easily seen that $X = (X_t)_{t \ge 0}$ is a Markov chain on $\Omega_\Lambda$ with generator $H$, and stationary measure $\widetilde{\phi}$. Similarly, $Y = (Y_t)_{t \ge 0}$ is a Markov chain on the subset $A$ of $\Omega_\Lambda$ with generator given by $H^A$ restricted to $A \times A$, and stationary measure $\widetilde{\phi}(\cdot \mid A)$. Both these chains are irreducible. This claim is trivial for $X$, since each edge has a strictly positive birth and death rate in this process. The claim is also true for $Y$, as the following argument indicates. If $\omega \in A$ and $Y_t = \omega$, then the process $Y$ can progress (with a strictly positive probability) from $\omega$ to the state $\mathbf{1}$; this is valid because edges are acquired by $Y$ at rate 1, and furthermore $\mathbf{1} \in A$ by virtue of the facts that $\omega \in A$ and $A$ is an increasing event. If $\omega' \in A$, then $Y$ can progress from $\mathbf{1}$ to $\omega'$ by the removal of exactly those edges $e$ for which $\omega'(e) = 0$; such transitions have strictly positive jump-rates since $A$ is an increasing event.

It follows from the irreducibility of $X$ and $Y$, and from the facts that $\widetilde{\phi}$ and $\widetilde{\phi}(\cdot \mid A)$ are stationary measures for $X$ and $Y$ respectively, that

$$\lim_{t \to \infty} P\big(X_t(e) = 1\big) = \widetilde{\phi}\big(\omega(e) = 1\big), \quad \text{and} \quad \lim_{t \to \infty} P\big(Y_t(e) = 1\big) = \widetilde{\phi}\big(\omega(e) = 1 \,\big|\, A\big),$$

(*cf.* Doob (1953, Thm VI.1.1)).

Finally we recall that $X_t \le Y_t$ for all $t$, in the light of which (3.6) may be written in the form given as follows.

**Proposition 5.** *If $A$ is an increasing event and $e \in \mathbb{E}_\Lambda$, then*

$$\begin{aligned}\frac{\partial}{\partial \widetilde{p}_e}\widetilde{\phi}(A) &= \frac{\widetilde{\phi}(A)}{\widetilde{p}_e(1-\widetilde{p}_e)}\lim_{t\to\infty}\bigl\{P\bigl(Y_t(e)=1\bigr)-P\bigl(X_t(e)=1\bigr)\bigr\} \\ &= \frac{\widetilde{\phi}(A)}{\widetilde{p}_e(1-\widetilde{p}_e)}\lim_{t\to\infty}\bigl\{P\bigl(X_t(e)=0, Y_t(e)=1\bigr)\bigr\},\end{aligned} \tag{3.16}$$

*where $(X_t, Y_t)_{t\ge 0}$ is a Markov chain on $\Omega_\Lambda \times A$ with generator given by* (3.11)–(3.13) *and satisfying* $(X_0, Y_0) = (\mathbf{0}, \mathbf{1})$.

Such couplings as that described above were used by Holley (1974) in his proof of the FKG inequality (see also Fortuin, Kasteleyn, and Ginibre (1971), and Preston (1974)). As shown by Holley's work, the coupled chain may be constructed (in such a way that the paired process $(X_t, Y_t)$ satisfies $X_t \le Y_t$ for all $t$, and has the desired marginals) precisely because $\widetilde{\phi}$ satisfies the inequality

$$\widetilde{\phi}(\omega \vee \omega')\widetilde{\phi}(\omega \wedge \omega') \ge \widetilde{\phi}(\omega)\widetilde{\phi}(\omega') \quad \text{for } \omega, \omega' \in \Omega_\Lambda, \tag{3.17}$$

where $\omega \vee \omega'(e) = \max\{\omega(e), \omega'(e)\}$. The product-measure component of $\widetilde{\phi}$ (see (3.1)) makes no essential contribution to (3.17); it is the factor $q^{k(\omega)}$ in (3.1) which leads to (3.17) when $q \ge 1$.

## 4. Remaining proofs

*Proof of Theorem 1.* We have that $\Lambda(n) = [-n, n]^d$, and

$$\frac{\partial \theta_n}{\partial p_x} = \sum_{u \in \Lambda(n)} \frac{\partial \theta_n}{\partial \widetilde{p}_{\langle u, u+x\rangle}}\bigg|_{\widetilde{\mathbf{p}}=\mathbf{p}} \tag{4.1}$$

where $\theta_n(\widetilde{\mathbf{p}}, q) = \widetilde{\phi}_n(0 \leftrightarrow \infty)$ and $\widetilde{\phi}_n$ is given by (3.1) with $\Lambda(n)$ in place of $\Lambda$. In writing $\widetilde{\mathbf{p}} = \mathbf{p}$, we mean the sum to be evaluated by setting $\widetilde{p}_e = p_e$ for all $e \in \mathbb{E}_{\Lambda(n)}$. The sum in (4.1) is over all points $u \in \Lambda(n)$ with the property that at least one of the points $u$ and $u+x$ lies in $\operatorname{int}(\Lambda(n)) = \Lambda(n)\setminus\partial\Lambda(n)$. For all other $u$ ($\in \mathbb{Z}^d$), the summand in (4.1) is 0, as the following argument shows. Certainly the summand is 0 if either $u$ or $u+x$ lies outside $\Lambda$, for then $\omega(\langle u, u+x\rangle) = 1$ for all $\omega \in \Omega_{\Lambda(n)}$. Suppose then that $u, u+x \in \partial\Lambda(n)$. We shall make use of the following claim:

(4.2) for any two vertices $v, w \in \partial\Lambda(n)$, there exists a path of $G$ from $v$ to $w$ using only edges in $\mathbb{E} \setminus \mathbb{E}_{\Lambda(n)}$.

If (4.2) is valid, then the value of $\omega(\langle u, u+x\rangle)$ influences neither $k(\omega)$ nor the occurrence of the event $\{0 \leftrightarrow \infty\}$, for $\omega \in \Omega_{\Lambda(n)}$. This holds since $u$ and $u+x$ are joined by a path of $G$ using only edges in $\mathbb{E} \setminus \mathbb{E}_{\Lambda(n)}$, and every such edge $e$ is such that $\omega(e) = 1$ for all $\omega \in \Omega_{\Lambda(n)}$. Therefore $\partial\theta_n/\partial\widetilde{p}_{\langle u, u+x\rangle} = 0$ when $u, u+x \in \partial\Lambda(n)$.

We turn now to the proof of (4.2). First we note that there exists a positive constant $D = D(K)$ such that any two vertices $v, w \in \mathbb{Z}^d$ with $|v - w| = 1$ are connected in $G$ by a path which stays within euclidean distance $D$ of $v$ (by virtue of (2.2)). It follows that, for any pair $v, w \in \mathbb{Z}^d$, and for any path $\pi$ from $v$ to $w$ on the graph $\mathbb{Z}^d$ with the usual nearest-neighbour relation, there exists a path $\widehat{\pi}$ of $G$ joining $v$ to $w$ which stays within distance $D$ of $\pi$. Now, if $v, w \in \partial\Lambda(n)$, there exist $z_1, z_2 \in \mathbb{Z}^d$ such that $v + z_1 \notin \Lambda(n)$, $w + z_2 \notin \Lambda(n)$, and either $z_i$ or $-z_i$ belongs to $K$ for each $i$. One may easily see then that $v + jz_1 \notin \Lambda(n)$ and $w + jz_2 \notin \Lambda(n)$ for $j \geq 1$, and that there exist $j_1$ and $j_2$ such that the euclidean distances from $\Lambda(n)$ to $v + j_1 z_1$ and to $w + j_2 z_2$ are each greater than $D$. Using the preceding observation, we may connect $v + j_1 z_1$ and $w + j_2 z_2$ by a path $\widehat{\pi}$ of $G$ lying outside $\Lambda(n)$. As a consequence, there exists a path from $v$ to $w$ using only vertices of the form $v + jz_1$ (for $0 \leq j \leq j_1$), $w + jz_2$ (for $0 \leq j \leq j_2$), together with the vertices of $\widehat{\pi}$; such a path uses only edges in $\mathbb{E} \setminus \mathbb{E}_{\Lambda(n)}$. This proves (4.2).

Our target now is to show that there exists a continuous function $\alpha : (0,1)^{\mathbb{E}_{\Lambda(n)}} \to (0, \infty)$ and a positive integer $N$ such that

$$\frac{\partial \theta_n}{\partial \widetilde{p}_{\langle u, u+x \rangle}} \leq \alpha(\widetilde{\mathbf{p}}) \left( \frac{\partial \theta_n}{\partial \widetilde{p}_{\langle u, u+y \rangle}} + \frac{\partial \theta_n}{\partial \widetilde{p}_{\langle u+x, u+x+y \rangle}} \right) \tag{4.3}$$

for all $n \geq N$, $x, y \in K$, and for all $u$ such that either $u$ or $u + x$ lies in $\mathrm{int}(\Lambda(n))$. If we are able to do this, then (2.4) follows immediately (with $\alpha$ replaced by $2\alpha$) by summing (4.3) over $u \in \Lambda(n)$.

Here is a sketch of the proof of (4.3). Take $A = \{0 \leftrightarrow \infty\}$, so that $\theta_n(\mathbf{p}, q) = \widetilde{\phi}_n(A)$. Let $(X_t, Y_t)$ be the Markov chain constructed in Section 3 for this event $A$. Fix $x, y \in K$ with $x \neq y$, and let $e = \langle u, u + x \rangle$ be an edge of $G$ with at least one endpoint lying in $\mathrm{int}(\Lambda(n))$. If $u \in \mathrm{int}(\Lambda(n))$, then set $f = \langle u, u + y \rangle$; if $u \in \partial\Lambda(n)$ but $u + x \in \mathrm{int}(\Lambda(n))$, then set $f = \langle u + x, u + x + y \rangle$. Then (4.3) will be implied by

$$\frac{\partial \theta_n}{\partial \widetilde{p}_e} \leq \alpha(\widetilde{\mathbf{p}}) \frac{\partial \theta_n}{\partial \widetilde{p}_f}. \tag{4.4}$$

By Proposition 5,

$$\frac{\partial \theta_n}{\partial \widetilde{p}_e} \bigg/ \frac{\partial \theta_n}{\partial \widetilde{p}_f} = \gamma(\widetilde{\mathbf{p}}) \frac{\lim_{t\to\infty} \{ P(X_t(e) = 0, Y_t(e) = 1) \}}{\lim_{t\to\infty} \{ P(X_t(f) = 0, Y_t(f) = 1) \}} \tag{4.5}$$

for some continuous $\gamma$ which is positive and finite when $\mathbf{0} < \widetilde{\mathbf{p}} < \mathbf{1}$. We shall show that there exists $\nu(\widetilde{\mathbf{p}})$, continuous and strictly positive when $\mathbf{0} < \widetilde{\mathbf{p}} < \mathbf{1}$, such that

$$P\big(X_{t+3}(f) = 0, Y_{t+3}(f) = 1\big) \geq \nu(\widetilde{\mathbf{p}})\, P\big(X_t(e) = 0, Y_t(e) = 1\big) \quad \text{for all } t. \tag{4.6}$$

This clearly implies that

$$\frac{\lim_{t\to\infty} \{ P(X_t(e) = 0, Y_t(e) = 1) \}}{\lim_{t\to\infty} \{ P(X_t(f) = 0, Y_t(f) = 1) \}} \leq \frac{1}{\nu(\widetilde{\mathbf{p}})}, \tag{4.7}$$

which in turn yields (4.4) with $\alpha(\widetilde{\mathbf{p}}) = \gamma(\widetilde{\mathbf{p}})/\nu(\widetilde{\mathbf{p}})$, via (4.5). Inequality (4.6) will be obtained by showing that

$$P\big(X_{t+3}(f) = 0, Y_{t+3}(f) = 1 \,\big|\, X_t(e) = 0, Y_t(e) = 1\big) \geq \nu(\widetilde{\mathbf{p}}). \tag{4.8}$$

We now prove (4.8). We introduce the $l^\infty$ norm on $\mathbb{Z}^d$, *i.e.*,

$$\|x\| = \max\{|x_i| : 1 \le i \le d\}, \quad \text{for } x \in \mathbb{Z}^d,$$

and we fix $m$ such that

$$m \ge 3 \max\big\{\|x\| : x \in K\big\}. \tag{4.9}$$

Let $\langle B_u \rangle$ be the collection of edges of $G$ having both endpoints in $(u+\Lambda(m)) \cap \Lambda(n)$; these are edges which are 'near' to $u$ but lying in $\Lambda(n)$. In particular $e, f \in \langle B_u \rangle$. We shall say that an edge $g$ is 'present in process $Z$ at time $s$' if $Z_s(g) = 1$. We introduce the following events:

$$V_t = \big\{X_t(e) = 0, Y_t(e) = 1\big\},$$
$$W_t = \big\{X_t(f) = 0, Y_t(f) = 1\big\}.$$

In addition we define the events $V^i$, $i = 1, 2, 3$, by

(i) $V^1$ is the event that: during the time-interval $(t, t+1]$, all edges in $\langle B_u \rangle$ which are present in $X_t$ are removed, and no edges in $\langle B_u \rangle$ are added to $X$; $e$ remains present in $Y$,

(ii) $V^2$ is the event that: during $(t+1, t+2]$, the edges $f$, $e'$, and $f'$ are added to $X$, but no other edges in $\langle B_u \rangle$ are added to $X$; $e$ remains present in $Y$,

where

$$\left.\begin{array}{l} e' = \langle u+y, u+x+y \rangle \\ f' = \langle u+x, u+x+y \rangle \end{array}\right\} \quad \text{if } u \in \operatorname{int}(\Lambda(n)),$$

and

$$\left.\begin{array}{l} e' = \langle u+y, u+x+y \rangle \\ f' = \langle u, u+y \rangle \end{array}\right\} \quad \text{if } u \in \partial\Lambda(n) \text{ and } u+x \in \operatorname{int}(\Lambda(n)).$$

Finally,

(iii) $V^3$ is the event that: during $(t+2, t+3]$, the edge $f$ is removed from $X$ but not from $Y$.

Note that, due to the coupling of $X$ and $Y$, the occurrence of $V^1$ may force the removal of some edges from $Y$ during $(t, t+1]$. However, on $V_t$, $e$ is absent from $X$ and present in $Y$ at time $t$. Thus, for $V^1$ to occur, $X_s(e)$ and $Y_s(e)$ must remain constant for $t < s \le t+1$. Similarly, the occurrence of $V^2$ will force the addition of $f$, $e'$, and $f'$ to $Y$, if not already present at time $t+1$, because $X_s \le Y_s$ (see also (3.11)–(3.13)). If either of the edges $e'$ or $f'$ lies outside $\mathbb{E}_{\Lambda(n)}$, then it is automatically present at all times in both $X$ and $Y$, and the requirement on such an edge under $V^2$ is vacuous. Note that none of the events $V_t$, $V^i$ for $1 \le i \le 3$, places any constraint on the states of edges outside $\langle B_u \rangle$.

Clearly $V_t \cap V^1 \cap V^2 \cap V^3 \subseteq W_{t+3}$. In order to prove (4.8) it therefore suffices to show that

$$P(V^1 \cap V^2 \cap V^3 \mid V_t) \ge \nu(\widetilde{\mathbf{p}}). \tag{4.10}$$

Fix $\mathbf{p} \in (0,1)^K$ and let $\pi(\mathbf{p}) = \min\{p_x, 1-p_x : x \in K\}$, and assume that all probabilities are calculated for a vector $\widetilde{\mathbf{p}}$ satisfying

$$\tfrac{1}{2}\pi(\mathbf{p}) < \widetilde{p}_g < 1 - \tfrac{1}{2}\pi(\mathbf{p}) \quad \text{for all } g \in \mathbb{E}_{\Lambda(n)}. \tag{4.11}$$

Assume now that $V_t$ occurs. According to (3.12) and (3.13), each edge $g$ ($\in \mathbb{E}_{\Lambda(n)}$) that is present in $X$ is lost at a rate which is bounded away from 0, uniformly in the states of all other edges (and uniformly in all $\widetilde{\mathbf{p}}$ satisfying (4.11)). Moreover, the total rate at which changes occur within $\langle B_u \rangle$ is no larger than the following upper bound for the sum of birth and death rates,

$$|\langle B_u \rangle| \left(1 + 2q \max\left\{\frac{1-\widetilde{p}_g}{\widetilde{p}_g} : g \in \mathbb{E}_{\Lambda(n)}\right\}\right) \le |\langle B_u \rangle| \left(1 + 2q\frac{2-\pi(\mathbf{p})}{\pi(\mathbf{p})}\right) = M \tag{4.12}$$

by (3.8)–(3.13); the bound $M$ does not depend on $n$ but only on $\mathbf{p}$. Therefore there exists $\nu_1(\widetilde{\mathbf{p}})$ such that

$$P(V^1 \mid V_t) \ge \nu_1(\widetilde{\mathbf{p}}) > 0; \tag{4.13}$$

the quantity $\nu_1(\widetilde{\mathbf{p}})$ (and quantities $\nu_i(\widetilde{\mathbf{p}})$ to be defined soon) may be chosen to be continuous and strictly positive when $\mathbf{0} < \widetilde{\mathbf{p}} < \mathbf{1}$, and not to depend on the values of $n$ and $u$.

Underlying the argument of the previous paragraph is a small lemma concerning Markov chains, which we choose not to state formally. Such a lemma amounts roughly to the following. If, for each edge $g$ belonging to some set $\mathcal{G}$ of bounded cardinality, the jump-rate of some transition of the state of $g$ is bounded away from 0, uniformly in the current states of all other edges, then the probability that the appropriate transitions take place on all the edges in $\mathcal{G}$ during a given time-interval is uniformly bounded away from 0. Similarly, if the total jump-rate on a collection $\mathcal{G}'$ of edges is uniformly bounded away from $\infty$, then the probability of any change at all on $\mathcal{G}'$ is uniformly bounded away from 1. A simple way of seeing this is to construct the chain in terms of 'exponential alarm clocks' sitting on the edges of $\mathcal{G}$ (or $\mathcal{G}'$), so that each edge examines its state whenever its alarm clock rings, and it changes its current state according to a probability distribution which may be a function of the current states of the other edges. The 'graphical representation' of Harris (1974, 1978) is such a representation for the contact process, and is one of the main techniques for its study.

Suppose that the event $V_t \cap V^1$ occurs, and consider the event $V^2$. Each edge $g$ ($\in \langle B_u \rangle$) is acquired by $X$ (and also by $Y$ if not already present there) at rate 1 (see (3.11)); therefore there exists $\nu_2(\widetilde{\mathbf{p}})$ of the required form such that

$$P(V^2 \mid V_t \cap V^1) \ge \nu_2(\widetilde{\mathbf{p}}) > 0. \tag{4.14}$$

Suppose that $V_t \cap V^1 \cap V^2$ occurs, and consider $V^3$. At time $t+2$, the only edges of $\langle B_u \rangle$ which are present in $X$ are $f$, $e'$, and $f'$. In general $Y_{t+2}$ contains more edges than does $X_{t+2}$, and certainly at least $e$, $f$, $e'$, and $f'$. We have therefore that $D_f(X_{t+2}) = 1$ and $D_f(Y_{t+2}) = 0$. It follows by (3.13)–(3.15) that the rate at which $f$ is removed from $Y$ *but not from* $X$ is at least

$$\frac{1-\widetilde{p}_f}{\widetilde{p}_f}(q-1) \ge \frac{\pi(\mathbf{p})}{2-\pi(\mathbf{p})}(q-1) > 0;$$

note that $f \in \mathbb{E}_{\Lambda(n)}$, since by construction $f$ has at least one endpoint in $\mathrm{int}(\Lambda(n))$. This implies that

$$P(V^3 \mid V_t \cap V^1 \cap V^2) \geq \nu_3(\widetilde{\mathbf{p}}) > 0 \tag{4.15}$$

for some suitable $\nu_3(\widetilde{\mathbf{p}})$. This in turn implies (4.10) with $\nu(\widetilde{\mathbf{p}}) = \nu_1(\widetilde{\mathbf{p}})\nu_2(\widetilde{\mathbf{p}})\nu_3(\widetilde{\mathbf{p}})$; as described before, (4.3) follows with $\alpha(\widetilde{\mathbf{p}}) = \gamma(\widetilde{\mathbf{p}})/\nu(\widetilde{\mathbf{p}})$. □

*Proof of Theorem 2.* It suffices to prove (2.5) in the special case when $\mathbf{f} = |K|^{-\frac{1}{2}}\mathbf{1}$, the unit vector with equal coordinates (here, $|K|$ denotes the cardinality of $K$). The case of general pairs of unit vectors follows by comparing each of the two directions with $\mathbf{f}$. Suppose then that $\mathbf{f}$ is given in this way.

Let $q > 1$, and let $\mathbf{p} \in \mathrm{C}_q$ and $\mathbf{e} \in U$. Certainly

$$\theta(\mathbf{p} + \alpha\mathbf{e}) \leq \theta(\mathbf{p} + \epsilon\mathbf{f}) \tag{4.16}$$

if $\mathbf{p} + \alpha\mathbf{e} \leq \mathbf{p} + \epsilon\mathbf{f}$, which is to say that

$$\alpha e_x \leq \epsilon|K|^{-\frac{1}{2}} \quad \text{for all } x \in K. \tag{4.17}$$

Now $e_x \leq 1$ for all $x \in K$, and therefore (4.17) holds with $\alpha = \epsilon|K|^{-\frac{1}{2}}$. This establishes the left-hand inequality of (2.5) for suitable $\nu$.

We turn to the other part of (2.5). Fix $N$ such that (2.4) holds for $n \geq N$. Choose in addition a real positive $\eta$ such that $2\eta \leq p_x \leq 1 - 2\eta$ for all $x \in K$, and find $a$ ($\in [1, \infty)$) such that

$$\alpha(\mathbf{p}') \leq a \quad \text{for all } \mathbf{p}' \text{ satisfying } \eta \leq p'_x \leq 1 - \eta \text{ for all } x \in K. \tag{4.18}$$

Finally, choose $\delta$ large and $\epsilon_0 \in (0, 1)$ small such that

$$\delta \geq 2(1 + a^2)|K|^{\frac{1}{2}}, \quad |K|\epsilon_0 \leq \tfrac{1}{4}, \quad \text{and} \quad (1 + \delta)\epsilon_0 \leq \eta, \tag{4.19}$$

and let

$$V = \{\mathbf{w} \in U : w_x > -\epsilon_0 \text{ for all } x \in K\}.$$

Clearly $V$ is open in $U$, and full. Let $\mathbf{e} \in V$, let $\epsilon$ satisfy $0 \leq \epsilon \leq \epsilon_0$, and define $\mathbf{p}'(t) = (p'_x(t) : x \in K) = \mathbf{p} + t\epsilon\mathbf{f} + (1 - t)\delta\epsilon\mathbf{e}$ for $0 \leq t \leq 1$. We have by (4.19) that $|p_x - p'_x(t)| \leq \eta$, and hence

$$\eta \leq p'_x(t) \leq 1 - \eta \quad \text{for } 0 \leq t \leq 1 \text{ and } x \in K.$$

It follows that $\alpha(\mathbf{p}'(t)) \leq a$ for $0 \leq t \leq 1$, by virtue of (4.18). From (2.4), we now obtain that $\theta_n(\cdot) = \theta_n(\cdot, q)$ satisfies

(4.20)

$$\begin{aligned}
\frac{d}{dt}\theta_n(\mathbf{p}'(t)) &= \sum_{x \in K} \epsilon(f_x - \delta e_x)\frac{\partial}{\partial p_x}\theta_n(\mathbf{p}'(t)) \\
&= \epsilon\left\{\sum_{x: f_x \geq \delta e_x} (f_x - \delta e_x)\frac{\partial}{\partial p_x}\theta_n(\mathbf{p}'(t)) + \sum_{x: f_x < \delta e_x} (f_x - \delta e_x)\frac{\partial}{\partial p_x}\theta_n(\mathbf{p}'(t))\right\} \\
&\leq \epsilon\left\{a \sum_{x: f_x \geq \delta e_x} (f_x - \delta e_x) + \frac{1}{a}\sum_{x: f_x < \delta e_x} (f_x - \delta e_x)\right\}\frac{\partial}{\partial p_u}\theta_n(\mathbf{p}'(t)) \\
&\leq \frac{\epsilon}{a}\left\{(a^2 + 1)\sum_{x \in K} f_x - \delta \sum_{x \in K} e_x\right\}\frac{\partial}{\partial p_u}\theta_n(\mathbf{p}'(t)),
\end{aligned}$$

where $u$ is any fixed member of $K$; in the last inequality, we used the assumption that $a \geq 1$, together with the fact that $\partial\theta_n/\partial p_u \geq 0$ (by the FKG inequality, as observed in Section 2). Moreover, since $e_x > -\epsilon_0$ for all $x \in K$, we have from (4.19) that

$$\begin{aligned}\sum_{x:e_x\geq 0} e_x &\geq \sum_{x:e_x\geq 0} e_x^2 = 1 - \sum_{x:e_x<0} e_x^2 \\ &\geq 1 - |K|\epsilon_0^2 \geq \tfrac{3}{4},\end{aligned}$$

whence (using (4.19) again)

$$\begin{aligned}\delta\sum_{x\in K} e_x &\geq \delta\big(\tfrac{3}{4} - |K|\epsilon_0^2\big) \geq \tfrac{1}{2}\delta \\ &\geq (1+a^2)|K|^{\frac{1}{2}} = (1+a^2)\sum_{x\in K} f_x.\end{aligned}$$

From these observations, it follows that the right-hand side of (4.20) is non-positive, and hence, after integrating with respect to $t$ from 0 to 1, that

$$\theta_n(\mathbf{p}+\epsilon\mathbf{f}) - \theta_n(\mathbf{p}+\epsilon\delta\mathbf{e}) = \theta_n(\mathbf{p}'(1)) - \theta_n(\mathbf{p}'(0)) \leq 0.$$

We take the limit as $n \to \infty$ to obtain the right-hand inequality of (2.5). □

*Proof of Theorem 3.* Let $q > 1$. It is evident that $\beta_c(\cdot, q)$ is non-increasing, since $\theta(\cdot, q)$ is non-decreasing. Suppose that $\beta_c(\mathbf{J}, q) = \beta_c(\mathbf{J}', q)$ $(= \beta_c$, say) for some $\mathbf{0} < \mathbf{J} \leq \mathbf{J}' < \boldsymbol{\infty}$. Then $\mathbf{p}(\mathbf{J}, \beta_c), \mathbf{p}(\mathbf{J}', \beta_c) \in \mathrm{C}_q$ and $\mathbf{p}(\mathbf{J}, \beta_c) \leq \mathbf{p}(\mathbf{J}', \beta_c)$; therefore $\mathbf{J} = \mathbf{J}'$ by Theorem 2 and the remark immediately afterwards. □

**Acknowledgement.** G.R.G. thanks Michael Aizenman for beneficial conversations enjoyed during the latter's stay in Voulangis during the summer of 1989.

## References

Aizenman, M. (1982). Geometric analysis of $\phi^4$ fields and Ising models. *Communications in Mathematical Physics* **86**, 1–48.

Aizenman, M., Chayes, J. T., Chayes, L., and Newman, C. M. (1988). Discontinuity of the magnetization in one-dimensional $1/|x-y|^2$ Ising and Potts models. *Journal of Statistical Physics* **50**, 1–40.

Aizenman, M. and Fernández, R. (1986). On the critical behavior of the magnetization in high-dimensional Ising models. *Journal of Statistical Physics* **44**, 393–454.

Aizenman, M. and Grimmett, G. R. (1991). Strict monotonicity for critical points in percolation and ferromagnetic models. *Journal of Statistical Physics* **63**, 817–835.

Barlow, R. N. and Proschan, F. (1965). *Mathematical Theory of Reliability.* Wiley, New York.

Doob, J. L. (1953). *Stochastic Processes.* Wiley, New York.

Edwards, R. G. and Sokal, A. D. (1988). Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm. *Physical Review* D **38**, 2009–2012.

Fortuin, C. M., Kasteleyn, P. W., and Ginibre, J. (1971). Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics* **22**, 89–103.

Grimmett, G. R. (1989). *Percolation.* Springer–Verlag, Berlin.

Grimmett, G. R. (1992). Potts models and random-cluster processes with many-body interactions. Preprint.

Grimmett, G. R. (1993). The random-cluster model. Preprint.

Harris, T. E. (1974). Contact interactions on a lattice. *Annals of Probability* **2**, 969–988.

Harris, T. E. (1978). Additive set-valued Markov processes and graphical methods. *Annals of Probability* **6**, 355–378.

Holley, R. (1974). Remarks on the FKG inequalities. *Communications in Mathematical Physics* **36**, 227–231.

Menshikov, M. (1987). Quantitative estimates and rigorous inequalities for critical points of a graph and its subgraphs. *Theory of Probability and its Applications* **32**, 544–547.

Preston, C. J. (1974). *Gibbs States on Countable Graphs.* Cambridge University Press, Cambridge.

Wierman, J. C. (1992). Equality of directional critical exponents in multiparameter percolation models. In preparation.

C. E. Bezuidenhout,
Mathematics Department, University of Rochester, Rochester NY 14627, USA

G. R. Grimmett,
Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, UK

H. Kesten,
Mathematics Department, Cornell University, Ithaca NY 14853–7901, USA