Random-cluster representations in the study of phase transitions

Olle Häggström
Chalmers University of Technology

September 3, 1997

Abstract

The Fortuin–Kasteleyn random-cluster model has, during the last 10 years, proved to be a highly useful probabilistic device for studying the phase transition behaviour of Ising and Potts models. In this survey paper, a detailed description is given of how this is accomplished. It is then shown how the same set of ideas can be used to study phase transitions in various other systems: the Ising model with random interactions, the Ashkin–Teller model, subshifts of finite type, and certain point processes. The tools used include coupling, stochastic domination, and percolation.

1 Introduction

There are several possible ways to extend the definition of a Markov chain to the case where the index set $\mathbf{Z}$ is replaced by $\mathbf{Z}^d$. Experience, however, indicates that the way of generalization which is by far the most fruitful, both from an applied and a theoretical point of view, is the **Markov random field**. The concept of a Markov random field gives rise to an interesting dichotomy between one dimension and higher: On one hand, an irreducible aperiodic finite state Markov chain (or, equivalently, an irreducible aperiodic finite state Markov random field indexed by $\mathbf{Z}$) always has a unique stationary distribution, and furthermore it is ergodic, has trivial tail, and so on. On the other hand, a finite state Markov random field in 2 or more dimensions satisfying the corresponding irreducibility and aperiodicity conditions may have more than one translation invariant distribution, and consequently all the usual mixing properties may fail. Measures satisfying the conditional distributions prescribed by the Markov random field are called **Gibbs measures**, and the nonuniqueness of such measures is referred to as a **phase transition**. The most well-known example of a Markov random field exhibiting a phase transition is the Ising model [61] [51] which was introduced in statistical mechanics as a microscopic model of a ferromagnet. The state space is $\{-1, 1\}$, and the classical interpretation of the model is that each site of $\mathbf{Z}^d$ is an atom, and that the states $+1$ and -1 represent 'spin up' and 'spin down', respectively. If the large-scale fraction of $+1$'s is $\frac{1}{2}$, then no magnetization has occurred, while if this fraction is different from $\frac{1}{2}$ so that the majority of spins point in the same direction, then the material is magnetized. It turns out that the latter scenario takes place if and only if the model exhibits a phase transition, and whether or not this happens depends on the so called reciprocal temperature parameter β. This is stated in the following well known and fundamental theorem; the precise definitions needed to fully understand the statement of the theorem will be given in Section 3.

Theorem 1.1: *For the Ising model on $\mathbf{Z}^d$ with $d \geq 2$, parameterized by the reciprocal temperature β, there exists a critical value $\beta_c = \beta_c(d) \in (0, \infty)$ such that for $\beta < \beta_c$ there is a unique Gibbs measure while for $\beta > \beta_c$ there is more than one Gibbs measure.*

The statement of this theorem can be broken down into three parts (each of which needs specific attention in a proof), namely:

(i) For sufficiently small $\beta > 0$, there is no phase transition.

(ii) For sufficiently large β, a phase transition occurs.

(iii) The occurrence of a phase transition is monotone increasing in β, i.e. if $\beta_1 < \beta_2$ and a phase transition occurs when $\beta = \beta_1$, then this is also the case when $\beta = \beta_2$.

Perhaps the most striking part (and the part which gives rise to the dichotomy between $d = 1$ and $d \geq 2$ referred to above) is (ii). The classical proof of this is a contour argument due to Peierls [70], Griffiths [33] and Dobrushin [18]. The argument is nice and physically intuitive, and appears in most textbooks on the subject (e.g. [22], [62], [29] and [40]). Other methods are required for (i) and (iii); the most common way is to use Dobrushin's uniqueness condition [20] for (i), and the Griffiths inequalities [34] for (iii).

Today, there exists a more or less unified approach to proving (i), (ii) and (iii). It uses the so called **random-cluster representation** of the Ising model, which was introduced by Fortuin and Kasteleyn [26] [24] [25] in the early 1970's, and brought into fashion some 15 years later through the papers by Swendsen and Wang [83], Edwards and Sokal [21] and Aizenman *et al.* [3] (see [37] for some further historical dicussion). The main point of this random-cluster representation is to translate apparently difficult questions about correlations into easier questions about percolation, i.e. about certain connectivity probabilities in a random graph. I believe that this may be the best approach to understanding Theorem 1.1, and that it will find its way into future textbooks in the field. Apart from simplicity and intuitive appeal, it also has the advantage of going through *mutatis mutandis* for the Potts model [76], i.e. for the generalization of the Ising model where the set $\{-1, 1\}$ of possible spins is replaced by $\{1, \ldots, q\}$ for some $q \geq 2$ ($q = 2$ is equivalent to the Ising model).

The present survey paper has two main purposes. The first is to give a detailed description of the random-cluster proof of Theorem 1.1 and the corresponding more general result for the Potts model. This is done in Sections 2 and 3. After briefly discussing the possibility of the random-cluster model itself exhibiting a phase transition in Section 4, I come to the second main purpose of the paper in Sections 5–8, which is to review some recent developments which show that the ideas of the random-cluster proof of Theorem 1.1 can be adapted in order to study phase transitions in many other models. I will focus on Ising models with random interactions (Section 5), the Ashkin–Teller model (Section 6), certain subshifts of finite type (Section 7), and models where the lattice structure $\mathbf{Z}^d$ is abandoned in favour of the continuum $\mathbf{R}^d$ (Section 8). For all the models considered, analogues of (i) and (ii) can be proved using the random-cluster approach, while (iii) appears to be feasible only for some of the models. Some concluding remarks are made in Section 9. The only new material in this paper are the random-cluster representations in Section 7. Sections 1–3 should be read first, whereas the remaining sections can be read in any order.

I can think of at least two good reasons why a probabilist or a statistician without any specialized interest in statistical mechanics might still want to consider taking the trouble to read this paper. The first, and most obvious, reason is that Ising, Potts and related models have found their way into an increasing number of applications outside of physics, e.g. in spatial statistics [49] and image analysis [28]. It is probably safe to conjecture that this trend will continue, and percolate into new fields. A question well worth asking in this context is whether the study of phase transitions is as important in these more recent fields of application as in statistical mechanics. A possible reason for being skeptical about this is that the phase transition phenomenon, as we define it here, only occurs for infinite systems, and that systems e.g. in image analysis are much smaller (typically 512×512 interacting components) than those in statistical mechanics (typically of the order 10^{23} components) and are thus less well approximated by infinite systems. However, even a system of size 512×512 turns out to be easily large enough for phase transition-like phenomena to be manifest. As a consequence of this, an understanding of phase transitions in infinite systems is likely to be important for a correct treatment of statistical and probabilistic aspects of such moderately sized systems.

The second reason that I have in mind for reading this paper is that the coupling and stochastic domination ideas, which will be used throughout, can facilitate the understanding of many other areas of probability theory. This I know from personal experience.

Readers who are content with a brief introduction to the random-cluster model may alternatively turn to Grimmett's survey [36]. Another introduction with a more combinatorial flavour is given by Welsh [86].

It should be pointed out that although I focus entirely on the most basic aspects of phase transition in this paper, the random-cluster representation has proved to be useful also for other problems concerning Ising and Potts models. Examples include studies of large deviation properties [75], correlation lengths [8], strong mixing properties [23], and various stochastic monotonicities [14] [45]. Probably just as important is the use of the random-cluster representation in the design of highly efficient Markov chain Monte Carlo algorithms [83] [64] [78].

I have found it convenient to build my presentation around Holley's Theorem [50] for stochastic domination between probability measures (see Theorem 3.5). In several instances where I have used Holley's Theorem, other authors have instead chosen to work with the FKG inequality [27]. This is mainly a matter of taste.

2 Ising, Potts, and random-cluster models

Ising and Potts models are easier to describe in a finite setting, i.e. when the lattice $\mathbf{Z}^d$ is replaced by a finite graph, so we begin with this case. Let G be a finite connected graph with vertex set V and edge set E, and let $\beta > 0$; β is called the reciprocal temperature parameter. A good example to have in mind is G being the nearest neighbour graph of a finite rectangular portion of $\mathbf{Z}^2$. For $v, w \in V$, write $v \sim w$ to denote the existence of an edge in G connecting v and w, and write $e(v, w)$ to denote the edge between v and

w. A spin configuration $\omega \in \{-1,1\}^V$ is said to have **energy**

$$H(\omega) = 2 \sum_{\substack{x \sim y \\ x,y \in V}} \mathbf{1}_{\{\omega(x) \neq \omega(y)\}} \tag{1}$$

where each pair $\{x,y\} \subseteq V$ having an edge in common is counted only once in the sum.

Definition 2.1: *The Gibbs measure ν_G^β on $\{-1,1\}^V$ for the* **Ising model** *at reciprocal temperature β is the measure for which*

$$\nu_G^\beta(\omega) = \frac{1}{Z_G^\beta} \exp\{-\beta H(\omega)\}$$

for all $\omega \in \{-1,1\}^V$. Here Z_G^β is a normalizing constant, making ν_G^β a probability measure.

The q-state Potts model is defined similarly:

Definition 2.2: *The Gibbs measure $\nu_G^{q,\beta}$ on $\{1,\ldots,q\}^V$ for the q-state* **Potts model** *at reciprocal temperature β is the probability measure for which*

$$\nu_G^{q,\beta}(\omega) = \frac{1}{Z_G^{q,\beta}} \exp\{-\beta H(\omega)\}$$

for all $\omega \in \{1,\ldots,q\}^V$. Here $H(\omega)$ is again given by (1), and $Z_G^{q,\beta}$ is a normalizing constant.

Identifying $\{-1,1\}$ with $\{1,2\}$, we see that the Ising model and the 2-state Potts model are the same. Gibbs measures for the Ising (resp. Potts) model will often simply be called Ising (resp. Potts) measures.

For a subset W of the vertex set V, define the **boundary** ∂W to be

$$\partial W = \{v \in V \setminus W : \exists w \in W \text{ such that } v \sim w\}.$$

Definition 2.3: *For a finite graph G with vertex set V and a finite set S, a random element X taking values in S^V is said to be a* **Markov random field** *on G if for each $W \subseteq V$, the conditional distribution of $X(W)$ given $X(V \backslash W)$ depends on $X(V \backslash W)$ only through its values on ∂W. More precisely, if P is the underlying probability measure, then X is a Markov random field if for all $\omega \in S^W$, $\omega' \in S^{V \backslash W}$ and $\omega'' \in S^{\partial W}$ such that ω'' is the restriction of ω' to ∂W and $P(X(V \setminus W) = \omega') > 0$, we have*

$$P(X(W) = \omega \mid X(V \setminus W) = \omega') = P(X(W) = \omega \mid X(\partial W) = \omega'').$$

If X is a $\{1,\ldots,q\}^V$-valued random element distributed according to the Potts measure $\nu_G^{q,\beta}$, then we have, for $\omega \in S^W$ and $\omega' \in S^{V \backslash W}$, that

$$\begin{aligned} &\nu_G^{q,\beta}(X(W) = \omega \mid X(V \setminus W) = \omega') \\ &= \frac{1}{Z_{\omega'}^{q,\beta}} \exp\left(-2\beta \sum_{\substack{x \sim y \\ x,y \in W}} \mathbf{1}_{\{\omega(x) \neq \omega(y)\}} - 2\beta \sum_{\substack{x \sim y \\ x \in W, y \in \partial W}} \mathbf{1}_{\{\omega(x) \neq \omega'(y)\}} \right) \end{aligned} \tag{2}$$

where $Z_{\omega'}^{q,\beta}$ is yet another normalizing constant (Z with various sub- and superscripts will always denote normalizing constants). The right hand side of (2) depends on ω' only through its values on ∂W, whence Ising and Potts measures define Markov random fields on G.

Let us now define the random-cluster model, and see how it relates to Ising and Potts models. The random-cluster model lives on the edges rather than on the vertices of G. Each edge is randomly declared to be present (state 1) or absent (state 0), so that a random element of $\{0,1\}^E$ is identified with a random subgraph of G.

Definition 2.4: *The* **random-cluster measure** $\mu_G^{p,q}$ *for G with parameters $p \in [0,1]$ and $q > 0$ is the probability measure on the set of subgraphs of G given by*

$$\mu_G^{p,q}(\eta) = \frac{1}{Z_G^{p,q}} \left\{ \prod_{e \in E} p^{\eta(e)}(1-p)^{1-\eta(e)} \right\} q^{k(\eta)}$$

for all $\eta \in \{0,1\}^E$. Here $k(\eta)$ is the number of connected components of η.

Note first that $q = 1$ corresponds to having the edges present or absent with probability p and $1-p$, independently of each other. All other values of q give rise to dependence between edges (as long as G is not a tree). The cases $q = 2, 3, \ldots$ correspond to the q-state Potts model, in a sense which we will now explain. Consider the probability measure $P_G^{p,q}$ on $\{1,\ldots,q\}^V \times \{0,1\}^E$ obtained through the following two-step procedure:

1. Assign each vertex a spin value chosen from $\{1,\ldots,q\}$ according to uniform distribution, assign each edge value 1 or 0 with respective probabilities p and $1-p$, and do this independently for all vertices and edges.

2. Condition on the event that no two vertices with different spins have an edge with value 1 connecting them.

More formally, $P_G^{p,q}$ is the measure which, to each element (ω, η) of $\{1,\ldots,q\}^V \times \{0,1\}^E$, assigns probability

$$\frac{1}{Z} \left\{ \prod_{e \in E} p^{\eta(e)}(1-p)^{1-\eta(e)} \right\} \prod_{\substack{x \sim y \\ x,y \in V}} \mathbf{1}_{\{(\omega(x)-\omega(y))\eta(e(x,y))=0\}}.$$

It turns out that the vertex and edge marginals of $P_G^{p,q}$ are distributed according to Potts and random-cluster measures, respectively, as stated in the following result:

Theorem 2.5: *Write $P_{G,vertex}^{p,q}$ (resp. $P_{G,edge}^{p,q}$) for the probability measure obtained by projecting $P_G^{p,q}$ on $\{1,\ldots,q\}^V$ (resp. $\{0,1\}^E$). Then*

$$P_{G,vertex}^{p,q} = \nu_G^{q,\beta} \tag{3}$$

where $\beta = -\frac{1}{2}\log(1-p)$, and

$$P_{G,edge}^{p,q} = \mu_G^{p,q}. \tag{4}$$

Proof: Fixing $\omega \in \{1,\ldots,q\}^V$ and summing over all $\eta \in \{0,1\}^E$ we get

$$\begin{aligned}
P^{p,q}_{G,vertex}(\omega) &= \sum_{\eta\in\{0,1\}^E} P^{p,q}_G(\omega,\eta) \\
&= \frac{1}{Z}\sum_{\eta\in\{0,1\}^E}\prod_{x\sim y} p^{\eta(e(x,y))}(1-p)^{1-\eta(e(x,y))}\mathbf{1}_{\{(\omega(x)-\omega(y))\eta(e(x,y))=0\}} \\
&= \frac{1}{Z}\prod_{x\sim y}(1-p)^{\mathbf{1}_{\{\omega(x)\neq\omega(y)\}}} \\
&= \frac{1}{Z}\exp\left(-2\beta\sum_{x\sim y}\mathbf{1}_{\{\omega(x)\neq\omega(y)\}}\right) \\
&= \nu^{q,\beta}_G(\omega)
\end{aligned}$$

since Z must equal $Z^{q,\beta}_G$ by normalization. Hence, (3) is proved. To prove (4) we proceed similarly, fixing $\eta \in \{0,1\}^E$ and summing over $\omega \in \{1,\ldots,q\}^V$. Note that, given η, there are exactly $q^{k(\eta)}$ spin configurations ω that are allowed, i.e. in which each pair of neighbouring vertices $x \sim y$ with $\eta(e(x,y)) = 1$ have the same spin. We get

$$\begin{aligned}
P^{p,q}_{G,edge}(\eta) &= \sum_{\omega\in\{1,\ldots,q\}^V} P^{p,q}_G(\omega,\eta) \\
&= q^{k(\eta)}\frac{1}{Z}\prod_{e\in E} p^{\eta(e)}(1-p)^{1-\eta(e)} \\
&= \mu^{p,q}_G(\eta)
\end{aligned}$$

again by normalization. □

Hence $P^{p,q}_G$ is a coupling of $\nu^{q,\beta}_G$ and $\mu^{p,q}_G$. (By a coupling, we here simply mean the joint construction of two random objects on the same probability space. Some authors reserve the term for a more specialized meaning. See [63] or [84] for general introductions to coupling methods in probability theory.) The coupling $P^{p,q}_G$ is the key to the use of the random-cluster model in studying Ising and Potts models. It was introduced implicitly by Swendsen and Wang [83] and explicitly by Edwards and Sokal [21]. The following two results are immediate consequences of Theorem 2.5 and the definition of $P^{p,q}_G$:

Corollary 2.6: *Let $p = 1 - e^{-2\beta}$, and suppose we pick a random spin configuration $X \in \{1,\ldots,q\}^V$ as follows:*

1. *Pick a random edge configuration $Y \in \{0,1\}^E$ according to the random-cluster measure $\mu^{p,q}_G$.*
2. *For each connected component $\mathcal{C}$ of Y, pick a spin at random (uniformly) from $\{1,\ldots,q\}$, assign this spin to every vertex of $\mathcal{C}$, and do this independently for different connected components.*

Then X is distributed according to the Potts measure $\nu^{q,\beta}_G$.

Corollary 2.7: *Let $p = 1 - e^{-2\beta}$, and suppose we pick a random edge configuration $Y \in \{0,1\}^E$ as follows:*

1. *Pick a random spin configuration $X \in \{1,\ldots,q\}^V$ according to the Potts measure $\nu_G^{q,\beta}$.*
2. *Conditional on X, assign each edge $e = e(x,y)$ independently value 1 with probability*
$$\begin{cases} p & \text{if } X(x) = X(y) \\ 0 & \text{if } X(x) \neq X(y), \end{cases}$$
and value 0 otherwise.

Then Y is distributed according to the random-cluster measure $\mu_G^{p,q}$.

The next result is a typical application of the random-cluster representation.

Corollary 2.8: *Let G be a finite graph, and pick a random spin configuration $X \in \{1,\ldots,q\}^V$ according to the Potts measure $\nu_G^{q,\beta}$. For $i \in \{1,\ldots,q\}$ and two vertices $u, v \in V$, the two events $\{X(u) = i\}$ and $\{X(v) = i\}$ are positively correlated, i.e.*
$$\nu_G^{q,\beta}(X(u) = i, X(v) = i) \geq \nu_G^{q,\beta}(X(u) = i)\nu_G^{q,\beta}(X(v) = i).$$

Proof: The measure $\nu_G^{q,\beta}$ is invariant under permutation of the spin set $\{1,\ldots,q\}$, so that
$$\nu_G^{q,\beta}(X(u) = i) = \nu_G^{q,\beta}(X(v) = i) = \frac{1}{q}.$$
It therefore suffices to show that
$$\nu_G^{q,\beta}(X(u) = i, X(v) = i) \geq \frac{1}{q^2}.$$
Now let us think of X as being obtained as in Corollary 2.6, i.e. by first picking an edge configuration $Y \in \{0,1\}$ according to the random-cluster measure $\mu_G^{p,q}$, and then assigning i.i.d. uniform spins to the connected components. Given Y, the conditional probability that $X(u) = X(v) = i$ is $\frac{1}{q}$ if u and v are in the same connected component of Y, and $\frac{1}{q^2}$ if they are in different connected components. Hence, for some $\alpha \in [0,1]$,
$$\nu_G^{q,\beta}(X(u) = i, X(v) = i) = \alpha\frac{1}{q} + (1-\alpha)\frac{1}{q^2} \geq \frac{1}{q^2}.$$
□

Remark: The same idea shows that if G is connected and $\beta > 0$, then the correlation between $\mathbf{1}_{\{X(u)=i\}}$ and $\mathbf{1}_{\{X(v)=i\}}$ is *strictly* positive.

We end this section with a very simple, but still useful, lemma about the random-cluster model.

Lemma 2.9: *For any edge $e \in E$, and any configuration $\eta \in \{0,1\}^{E\setminus\{e\}}$, we have that*
$$\mu_G^{p,q}(e \text{ is present} \mid \eta) = \begin{cases} p & \text{if the endvertices of } e \text{ are connected in } \eta \\ \frac{p}{p+(1-p)q} & \text{otherwise.} \end{cases}$$

Proof: Immediate from the definition of $\mu_G^{p,q}$. □

3 Phase transition in Ising and Potts models

In this section we consider Ising and Potts models on the cubic lattice $\mathbf{Z}^d$; with a slight abuse of notation we identify $\mathbf{Z}^d$ with its nearest neighbour graph (i.e. the graph in which two vertices $x, y \in \mathbf{Z}^d$ have an edge connecting them if and only if their Euclidean distance is 1). The definitions of Ising and Potts models on finite graphs do not carry over directly to infinite graphs, since e.g. the set of possible spin configurations becomes uncountable. On the other hand, the concept of a Markov random field does carry over in a straightforward manner:

Definition 3.1: *Let X be a random element taking values in $S^{\mathbf{Z}^d}$ (with S being a finite set) and write P for the underlying probability measure. We say that X is a Markov random field if P admits conditional probabilities such that for every finite $W \subset \mathbf{Z}^d$ and all $\omega \in S^W$, $\omega' \in S^{\mathbf{Z}^d \setminus W}$ and $\omega'' \in S^{\partial W}$ such that ω'' is the restriction of ω' to ∂W we have*

$$P(X(W) = \omega \mid X(\mathbf{Z}^d \setminus W) = \omega') = P(X(W) = \omega \mid X(\partial W) = \omega'').$$

Ising and Potts models for $\mathbf{Z}^d$ can now be defined in terms of conditional probabilities on finite sets – this is the so called Dobrushin–Lanford–Ruelle [19] [58] approach to infinite-volume Gibbs measures:

Definition 3.2: *A probability measure ν on $\{-1,1\}^{\mathbf{Z}^d}$ (resp. $\{1,\ldots,q\}^{\mathbf{Z}^d}$) is called an Ising (resp. Potts) measure at reciprocal temperature β if its corresponding random element X is a Markov random field satisfying*

$$\begin{aligned} &\nu(X(W) = \omega \mid X(\partial W) = \omega'') \\ &= \frac{1}{Z^{q,\beta}_{\omega''}} \exp\left(-2\beta \sum_{\substack{x\sim y \\ x,y\in W}} \mathbf{1}_{\{\omega(x)\neq\omega(y)\}} - 2\beta \sum_{\substack{x\sim y \\ x\in W, y\in\partial W}} \mathbf{1}_{\{\omega(x)\neq\omega''(y)\}} \right) \end{aligned} \qquad (5)$$

for every finite $W \subset \mathbf{Z}^d$ and all spin configurations ω and ω'' on W and ∂W.

A comparison with (2) suggests that this is a reasonable generalization of Definitions 2.1 and 2.2. The existence (for any q, β and d) of such measures can be deduced either by considering the models on larger and larger finite portions of $\mathbf{Z}^d$ and using compactness of $\{-1,1\}^{\mathbf{Z}^d}$ to deduce the existence of a subsequential weak limit, or via the monotonicity arguments that we will give below. As indicated in the introduction, the focus in this paper is on the question of (non-)uniqueness:

Given q, β and d, does there exist more than one Potts measure?

For $d = 1$, the answer is no, while in higher dimensions the answer depends on q and β. In the Ising case $q = 2$, this was stated already in Theorem 1.1, but the same phenomenon generalizes to arbitrary q:

Theorem 3.3: *For the q-state Potts model on $\mathbf{Z}^d$ with $d \geq 2$ and reciprocal temperature β, there exists a critical value $\beta_c = \beta_c(d,q) \in (0,\infty)$ such that for $\beta < \beta_c$ there is a unique Potts measure while for $\beta > \beta_c$ there is more than one Potts measure.*

We will prove this for the special case of the Ising model (Theorem 1.1) first; the reason for this is that the Ising model posseses certain stochastic monotonicities which make life a bit easier and which are not shared by the $q \geq 3$ Potts model. Note that the important and highly intricate question of what happens *at* the critical value $\beta = \beta_c$ is completely ignored here. This issue has received much attention elsewhere; see e.g. [57].

Let Λ_n denote the box $\{-n, \ldots, n\}^d \subset \mathbf{Z}^d$. Fix β, and let $\nu_{+,n}^{\beta}$ denote the probability measure on $\{-1,1\}^{\mathbf{Z}^d}$ which corresponds to picking a random spin configuration on $\mathbf{Z}^d$ as follows. First assign spin $+1$ to all sites outide the box Λ_n, and then pick the spins inside Λ_n according to (5), with $W = \Lambda_n$ and $\omega'' \equiv +1$. We call this the Ising measure for Λ_n with "plus" boundary condition. The significance of these measures is due to the following result, where the mode of convergence is weak convergence in the product topology for $\{-1,1\}^{\mathbf{Z}^d}$ (since the spin space is discrete, this means that the probabilities of all cylinder events converge; a cylinder event is an event which depends on the spins at finitely many vertices only).

Proposition 3.4: *The limiting probability measure*

$$\nu_+^{\beta} = \lim_{n \to \infty} \nu_{+,n}^{\beta}$$

on $\{-1,1\}^{\mathbf{Z}^d}$ exists and is an Ising measure.

In order to prove this result, we need to introduce the concept of **stochastic domination** and a result of Holley [50] which is central to more or less everything in this paper. Let V be a finite or countable set, and let P and P' be two probability measures on $\mathbf{R}^V$. A function $f : \mathbf{R}^V \to \mathbf{R}$ is said to be increasing if $f(\xi) \leq f(\eta)$ whenever $\xi \preceq \eta$. Here $\preceq$ is the usual (coordinatewise) partial order, i.e. $\xi \preceq \eta$ if $\xi(v) \leq \eta(v)$ for all $v \in V$. We say that P' dominates P, writing $P' \succeq_d P$ (or equivalently $P \preceq_d P'$) if

$$\int f \, dP \leq \int f \, dP'$$

for every bounded increasing f. We also write $X \preceq_d X'$ for two $\mathbf{R}^V$-valued random variables whose distributions are ordered in this way. By a well-known theorem of Strassen [82], this is equivalent to the existence of a coupling of X and X' such that $X \preceq X'$ a.s.; see also [63].

Theorem 3.5: Holley's Theorem. *Let V be finite and let S be a finite subset of $\mathbf{R}$. Let P and P' be two probability measures on S^V, assume that P' assigns positive probability to every element of S^V, and write X and X' for the corresponding random elements. Suppose that for every $v \in V$, every $s \in S$ and every $\xi, \eta \in S^{V \setminus \{v\}}$ such that $\xi \preceq \eta$ and $P(X(V \setminus \{v\}) = \xi) > 0$ we have*

$$P(X(v) \geq s \,|\, X(V \setminus \{v\}) = \xi) \;\leq\; P'(X'(v) \geq s \,|\, X'(V \setminus \{v\}) = \eta). \tag{6}$$

Then $P \preceq_d P'$.

Holley did in fact not state the result quite in this generality, but the following proof is a trivial extension of Holley's proof:

Proof of Theorem 3.5: Consider the Markov chain $\{X_k\}_{k=0}^{\infty}$ with state space S^V and transition probabilities defined by the following procedure. At each integer time $k \geq 1$, pick a site $v \in V$ at random (according to uniform distribution), let $X_k(w) = X_{k-1}(w)$ for each $w \in V \setminus \{v\}$, and select $X_k(v)$ according to the conditional distribution prescribed by P. This is a so called Gibbs sampler for P, and it is immediate that if the initial value X_0 is chosen according to P, then X_k has distribution P for each k.

Next define another Markov chain $\{X'_k\}_{k=0}^{\infty}$ with the same state space analogously, but with P replaced by P'. By the assumption that P' assigns positive probability to each element of S^V, it is easy to see that the $\{X'_k\}$ chain is irreducible.

Finally, define a coupling of $\{X_k\}_{k=0}^{\infty}$ and $\{X'_k\}_{k=0}^{\infty}$ as follows. First pick the initial values (X_0, X'_0) according to $P \times P'$. Then, for each k, pick a site $v \in V$ at random and let U_k be an independent random variable, uniformly distributed on the interval $[0,1]$. Let $X_k(w) = X_{k-1}(w)$ and $X'_k(w) = X'_{k-1}(w)$ for each site $w \neq v$, and update the values at site v by letting

$$X_k(v) = \max\{s \in S : P(X(v) \geq s \,|\, X(V \setminus \{v\}) = \xi) \geq U_k\}$$

and

$$X'_k(v) = \max\{s \in S : P'(X'(v) \geq s \,|\, X'(V \setminus \{v\}) = \eta) \geq U_k\}$$

where $\xi = X_{k-1}(V \setminus \{v\})$ and $\eta = X'_{k-1}(V \setminus \{v\})$. It is clear that this construction gives the correct marginal distributions of $\{X_k\}_{k=0}^{\infty}$ and $\{X'_k\}_{k=0}^{\infty}$. Furthermore, (6) implies that $X_k \preceq X'_k$ whenever $X_{k-1} \preceq X'_{k-1}$. Since the $\{X'_k\}$ chain is irreducible, it will a.s. hit the maximal state (with respect to $\preceq$) of S^V after finite time, and from this time on we will thus have $X_k \preceq X'_k$. Note furthermore that $\{X_k, X'_k\}_{k=0}^{\infty}$ is itself a finite state aperiodic Markov chain, so that (X_k, X'_k) has a limiting distribution as $k \to \infty$. Picking (X, X') according to this limiting distribution gives a coupling of X and X' such that $X \preceq X'$ a.s., whence $P \preceq_d P'$. $\Box$

A typical application of Holley's Theorem is the following lemma, which will also serve us in proving Proposition 3.4. For finite $W \subset \mathbf{Z}^d$ and a spin configuration $\omega'' \in \{-1,1\}^{\partial W}$, we define the Ising measure for W with boundary condition ω'' to be the probability measure on $\{-1,1\}^W$ defined by the right hand side of (5).

Lemma 3.6: *Let $W \subset \mathbf{Z}^d$ be finite, and let ω''_1 and ω''_2 be two spin configurations on ∂W satisfying $\omega''_1 \preceq \omega''_2$. Furthermore, let $\nu^{\beta}_{W,\omega''_1}$ and $\nu^{\beta}_{W,\omega''_2}$ be the Ising measures on W with boundary conditions ω''_1 and ω''_2. We then have*

$$\nu^{\beta}_{W,\omega''_1} \preceq_d \nu^{\beta}_{W,\omega''_2}.$$

Proof: The conditional probability of a plus spin at a single site $v \in W$ given everything else equals $(1 + e^{2\beta(2d-2\kappa)})^{-1}$, where κ is the number of nearest neighbours of v having plus spin. This is an increasing function of the spin configuration off v, whence Holley's Theorem can be invoked to give the desired conclusion. $\Box$

Proof of Proposition 3.4: We first show existence of the limit. This follows (by compactness of $\{-1,1\}^{\mathbf{Z}^d}$) if we can show that

$$\nu^{\beta}_{+,1} \succeq_d \nu^{\beta}_{+,2} \succeq_d \cdots. \tag{7}$$

To see this implication, note that (7) implies that the probabilities of events of the type "all sites in the (finite) set $A \subset \mathbf{Z}^d$ have spin $+1$" decrease, whence they converge, and by inclusion-exclusion the probabilities of all cylinder events converge. We proceed to show (7), i.e. to show that

$$\nu^{\beta}_{+,n} \succeq_d \nu^{\beta}_{+,n+1} \tag{8}$$

for any n. Consider the following way of picking a random element $X_{n+1} \in \{-1,1\}^{\mathbf{Z}^d}$ according to $\nu^\beta_{+,n+1}$. First let $X_{n+1} \equiv +1$ on $\mathbf{Z}^d \setminus \Lambda_{n+1}$. Then assign spins to $\Lambda_{n+1} \setminus \Lambda_n$ according to the projection of $\nu^\beta_{+,n+1}$ on $\{-1,1\}^{\Lambda_{n+1}\setminus\Lambda_n}$. Finally pick the spins on Λ_n according to the conditional distribution in Λ_n given the configuration on $\Lambda_{n+1} \setminus \Lambda_n$. By Lemma 3.6, this conditional distribution is stochastically dominated by the projection on $\{-1,1\}^{\Lambda_n}$ of $\nu^\beta_{+,n}$. Hence we can couple X_{n+1} and X_n so that $X_n \succeq X_{n+1}$ a.s., so we have (8), and existence of the limiting measure ν^β_+ follows. To see that this is an Ising measure, we need to check for any finite $W \subset \mathbf{Z}^d$ and any $\omega'' \in \{-1,1\}^{\partial W}$ that ν^β_+ satisfies (5). This, however, is immediate from the fact that the same property holds for $\nu^\beta_{+,n}$ for each n which is large enough for $W \cup \partial W$ to be contained in Λ_n. $\Box$

The measure ν^β_+ plays a special role, in that

$$\nu^\beta_+ \succeq_d \nu^\beta \tag{9}$$

for any other Ising measure ν^β at the same reciprocal temperature; this follows from the same use of Lemma 3.6 as in the proof of Proposition 3.4. By the ± 1 symmetry of the model, we of course have a similar measure ν^β_-, obtained with "minus" boundary conditions, for which

$$\nu^\beta_- \preceq_d \nu^\beta \tag{10}$$

for all Ising measures ν^β. One can check that ν^β_+ and ν^β_- are translation invariant. A key to analyzing the phase transition behaviour of the Ising model is the following lemma. Write $\mathbf{0}$ for the origin of $\mathbf{Z}^d$ and write, as before, X for a $\{-1,1\}^{\mathbf{Z}^d}$-valued random element.

Lemma 3.7: *There is a unique Ising measure for $\mathbf{Z}^d$ at reciprocal temperature β if and only if*

$$\nu^\beta_+(X(\mathbf{0}) = 1) = \frac{1}{2}. \tag{11}$$

Proof: Suppose first that $\nu^\beta_+(X(\mathbf{0}) = 1) \neq \frac{1}{2}$. Then, by (9), (10) and ± 1 symmetry, we have $\nu^\beta_-(X(\mathbf{0}) = 1) < \frac{1}{2} < \nu^\beta_+(X(\mathbf{0}) = 1)$, so $\nu^\beta_- \neq \nu^\beta_+$, and we have nonuniqueness of Ising measures. For the other direction, suppose that $\nu^\beta_+(X(\mathbf{0}) = 1) = \frac{1}{2}$. Then $\nu^\beta_-(X(\mathbf{0}) = 1) = \frac{1}{2}$, again by ± 1 symmetry. Translation invariance gives

$$\nu^\beta_-(X(x) = 1) = \nu^\beta_+(X(x) = 1) = \frac{1}{2}$$

for each $x \in \mathbf{Z}^d$. By (9) and Strassen's coupling theorem, there exists a pair X, X' of $\{-1,1\}^{\mathbf{Z}^d}$-valued random elements with distributions ν^β_+ and ν^β_- such that $X' \preceq X$ P-a.s., where P is the underlying probability measure of the coupling. This means that for each x,

$$\begin{aligned} P(X'(x) \neq X(x)) &= P(X'(x) = -1, X(x) = 1) \\ &= \nu^\beta_+(X(x) = 1) - \nu^\beta_-(X'(x) = 1) \\ &= 0 \end{aligned}$$

so that $X = X'$ P-a.s., whence $\nu^\beta_+ = \nu^\beta_-$. This, in conjunction with (9) and (10), implies uniqueness of Ising measures. $\Box$

We have thus reduced the question of uniqueness of Ising measures to that of whether (11) holds. The latter question is particularly amenable to random-cluster methods. For the remainder of this section, we shall stay within eyesight of the path staked out by Aizenman *et al.* [3]. We shall consider a close variant of the coupling between the random-cluster model and the Ising model in Section 2. Let G_n be the graph with vertex set $V_n = \Lambda_{n+1} = \{-n-1, \ldots, n+1\}^d$ and edge set E_n consisting of all nearest neighbour pairs in Λ_{n+1}. Write E_n as the disjoint union $E'_n \cup E''_n$, where E'_n consists of all edges that have at least one endvertex in Λ_n, and E''_n thus consists of those edges that have both endvertices in $\Lambda_{n+1} \setminus \Lambda_n$. For $p \in [0,1]$, let $P^p_{G_n}$ be the probability measure on $\{-1,1\}^{V_n} \times \{0,1\}^{E_n}$ defined by the following procedure:

1. Assign each vertex of $\Lambda_{n+1} \setminus \Lambda_n$ value 1 and also each edge of E''_n value 1 (this is because we are interested in the Ising model on Λ_n with plus boundary conditions).

2. Assign each vertex of Λ_n value 1 or -1 with probability $\frac{1}{2}$ each, assign each edge of E'_n value 1 or 0 with respective probabilities p and $1-p$, and do this independently for all edges and vertices.

3. Condition on the event that no two vertices with different spins have an edge connecting them.

It is now a simple modification of the proof of Theorem 2.5 to check that the vertex marginal of $P^p_{G_n}$ equals the projection on $\{-1,1\}^{\Lambda_n}$ of the Ising measure $\nu^\beta_{+,n}$, where $\beta = -\frac{1}{2}\log(1-p)$, and that the edge marginal equals $\tilde{\mu}^{p,2}_{G_n}$ which we define as the random-cluster measure $\mu^{p,2}_{G_n}$ conditioned on the event that all edges of E''_n are present. Write, as usual, X for a random spin configuration, and let $\mathbf{0} \leftrightarrow \partial\Lambda_n$ denote the event that there exists a path of open edges connecting the origin to $\partial\Lambda_n$. By a direct analogue of Corollary 2.6, we have that

$$\nu^\beta_{+,n}(X(\mathbf{0}) = 1) = \frac{1 + \tilde{\mu}^{p,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n)}{2}$$

whence

$$\begin{aligned} \nu^\beta_{+}(X(\mathbf{0}) = 1) &= \lim_{n\to\infty} \nu^\beta_{+,n}(X(\mathbf{0}) = 1) &(12)\\ &= \frac{1 + \lim_{n\to\infty} \tilde{\mu}^{p,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n)}{2} &(13)\end{aligned}$$

(existence of the limit in (13) follows from existence of the limit in (12)). The question of uniqueness of Ising measures, which we previously translated into that of whether (11) holds, can thus be further reduced to the question of whether or not for $p = 1 - e^{-2\beta}$ we have

$$\lim_{n\to\infty} \tilde{\mu}^{p,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n) = 0. \tag{14}$$

This question can be answered (partially) using Holley's Theorem together with the following result, which could possibly be dubbed "the fundamental theorem of percolation theory". In standard bond percolation, each edge of the nearest-neighbour graph on $\mathbf{Z}^d$ is independently present or absent with probability p resp. $1-p$. Let $\theta(p)$ denote the probability that, under this model, $\mathbf{0}$ is in an infinite connected component of present edges.

Theorem 3.8: *For bond percolation on $\mathbf{Z}^d$, $d \geq 2$, there exists a critical value $p_c = p_c(d) \in (0,1)$, such that*

$$\theta(p) \begin{cases} = 0 & \text{if } p < p_c \\ > 0 & \text{if } p > p_c. \end{cases}$$

See e.g. the standard monograph by Grimmett [35] for a simple proof; the substantial part is to show that p_c is *strictly* between 0 and 1. The result dates back some 40 years to the advent of percolation theory [10] [47] [48]. Recall that taking $q = 1$ in the random-cluster model yields independent edges with probability p of being present. Hence standard bond percolation can be seen as an instance of the random-cluster model. It follows that

$$\lim_{n\to\infty} \mu^{p,1}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n) \begin{cases} = 0 & \text{if } p < p_c \\ > 0 & \text{if } p > p_c \end{cases} \tag{15}$$

(the limit exists by the observation that $\mu^{p,1}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n)$ is decreasing in n). Analogously to Lemma 2.9, we have for each $e \in E'_n$, and each configuration η of edges off e that

$$\tilde{\mu}^{p,2}_{G_n}(e \text{ is present} \,|\, \eta) = \begin{cases} p & \text{if the endvertices of } e \text{ are connected in } \eta \\ \frac{p}{2-p} & \text{otherwise.} \end{cases} \tag{16}$$

Hence, we can apply Holley's Theorem to the projections on $\{0,1\}^{E'_n}$ of $\mu^{p,1}_{G_n}$ and $\tilde{\mu}^{p,2}_{G_n}$. We get that the former projection stochastically dominates the latter, so that in particular

$$\tilde{\mu}^{p,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n) \leq \mu^{p,1}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n).$$

Similarly,

$$\tilde{\mu}^{p,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n) \geq \mu^{\frac{p}{2-p},1}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n).$$

By (15), we thus have that (14) holds for $p < p_c$ and fails for $p > \frac{2p_c}{1+p_c}$. Hence

$$\lim_{n\to\infty} \nu^{\beta}_{+,n}(X(\mathbf{0}) = 1) \begin{cases} = \frac{1}{2} & \text{for } \beta < -\frac{1}{2}\log(1-p_c) \\ > \frac{1}{2} & \text{for } \beta > -\frac{1}{2}\log(1-\frac{2p_c}{1+p_c}) \end{cases}$$

and we have shown two thirds of Theorem 1.1. Since $\frac{2p_c}{1+p_c} > p_c$, it remains to show that the occurrence of a phase transition is monotone in β. By the reduction of the phase transition question to (14), it is sufficient to show that $\tilde{\mu}^{p,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n)$ is increasing in p for each n. A key observation now is that the right hand side of (16) is increasing in p as well as in η. Let $0 < p_1 < p_2 < 1$, and consider the measures $\tilde{\mu}^{p_1,2}_{G_n}$ and $\tilde{\mu}^{p_2,2}_{G_n}$. Holley's Theorem applied to their projections on $\{0,1\}^{E'_n}$ tells us that

$$\tilde{\mu}^{p_1,2}_{G_n} \preceq_d \tilde{\mu}^{p_2,2}_{G_n} \tag{17}$$

whence $\tilde{\mu}^{p_1,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n) \leq \tilde{\mu}^{p_2,2}_{G_n}(\mathbf{0} \leftrightarrow \partial\Lambda_n)$. This completes the proof of Theorem 1.1.

Remark: The problem of finding an exact expression for the critical value β_c in Theorem 1.1 is believed to be mathematically intractable, with the remarkable exception of $d = 2$, where it is known [68] [1] that $\beta_c = \frac{1}{2}\log(1+\sqrt{2})$. For the 2-dimensional q-state Potts model, it is believed that $\beta_c(q) = \frac{1}{2}\log(1+\sqrt{q})$; see e.g. [85] for some non-rigorous random-cluster arguments for this formula.

The remaining task in this section is to go from the Ising model to the Potts model, i.e. to extend the above proof of Theorem 1.1 into a proof of Theorem 3.3.

Consider the q-state Potts model on $\mathbf{Z}^d$, with $q \geq 3$. For a spin $i \in \{1, \ldots, q\}$, we can define $\nu_{i,n}^{q,\beta}$ to be the Potts measure for the box Λ_n with boundary condition consisting of spin i everywhere on $\mathbf{Z}^d \setminus \Lambda_n$. Let the probability measure $\tilde{\mu}_{G_n}^{p,q}$ on $\{0,1\}^{E_n}$ be given by the obvious generalization of $\tilde{\mu}_{G_n}^{p,2}$. The coupling of $\nu_{+,n}^{\beta}$ and $\tilde{\mu}_{G_n}^{p,2}$ generalizes directly to a completely analogous coupling of $\nu_{i,n}^{q,\beta}$ and $\tilde{\mu}_{G_n}^{p,q}$. Applying Holley's Theorem to the projections of $\tilde{\mu}_{G_n}^{p,q}$ and $\tilde{\mu}_{G_{n+1}}^{p,q}$ on $\{0,1\}^{E'_n}$, we get that the former projection dominates the latter stochastically. This implies that $\tilde{\mu}_{G_n}^{p,q}(\mathbf{0} \leftrightarrow \partial\Lambda_n)$ is decreasing in n, and hence has a limit as $n \to \infty$. Corollary 2.6 gives

$$\nu_{i,n}^{q,\beta}(X(\mathbf{0}) = i) = \frac{1 + (q-1)\tilde{\mu}_{G_n}^{p,q}(\mathbf{0} \leftrightarrow \partial\Lambda_n)}{q}$$

so that

$$\lim_{n\to\infty} \nu_{i,n}^{q,\beta}(X(\mathbf{0}) = i) = \frac{1 + (q-1)\lim_{n\to\infty}\tilde{\mu}_{G_n}^{p,q}(\mathbf{0} \leftrightarrow \partial\Lambda_n)}{q}.$$

Comparing with independent bond percolation as we did for the Ising case gives

$$\lim_{n\to\infty} \nu_{i,n}^{q,\beta}(X(\mathbf{0}) = 1) \begin{cases} = \frac{1}{q} & \text{for } \beta < -\frac{1}{2}\log(1-p_c) \\ > \frac{1}{q} & \text{for } \beta > -\frac{1}{2}\log(1-\frac{qp_c}{1+(q-1)p_c}). \end{cases} \tag{18}$$

One more application of Holley's Theorem tells us that $\tilde{\mu}_{G_n}^{p,q}(\mathbf{0} \leftrightarrow \partial\Lambda_n)$ is increasing in p, whence the same thing holds for the limit as $n \to \infty$. In conjunction with (18), this implies the following result:

Lemma 3.9: *There exists a critical value $\beta^* = \beta^*(q,d) \in (0,\infty)$ such that*

$$\lim_{n\to\infty} \nu_{i,n}^{q,\beta}(X(\mathbf{0}) = 1) \begin{cases} = \frac{1}{q} & \text{for } \beta < \beta^* \\ > \frac{1}{q} & \text{for } \beta > \beta^*. \end{cases}$$

Theorem 3.3 now follows, with $\beta_c = \beta^*$, once we can show:

Lemma 3.10: *There is a unique q-state Potts measure for $\mathbf{Z}^d$ at reciprocal temperature β if and only if*

$$\lim_{n\to\infty} \nu_{i,n}^{q,\beta}(X(\mathbf{0}) = i) = \frac{1}{q}.$$

For the Ising case $q = 2$, this result popped out rather easily (Lemma 3.7) due to the monotonicity relations (9) and (10). For the $q \geq 3$ Potts model some more work is required, in particular for the 'if' direction.

Proof of Lemma 3.10: By Lemma 3.9, we have that $\lim_{n\to\infty} \nu_{i,n}^{q,\beta}(X(\mathbf{0}) = i) \geq \frac{1}{q}$. Suppose first that $\lim_{n\to\infty} \nu_{i,n}^{q,\beta}(X(\mathbf{0}) = 1) > \frac{1}{q}$ and let $\nu_i^{q,\beta}$ be some subsequential weak limit of $\{\nu_{i,n}^{q,\beta}\}_{n=1}^{\infty}$. The measure $\nu_i^{q,\beta}$ is then a Potts measure for $\mathbf{Z}^d$ with the prescribed parameters. Since $\nu_i^{q,\beta}(X(\mathbf{0}) = i) > \frac{1}{q}$, there must be some other spin $j \in \{1, \ldots, q\}$ such that $\nu_i^{q,\beta}(X(\mathbf{0}) = j) < \frac{1}{q}$. Now construct another Potts measure $\nu_j^{q,\beta}$ by using the "all j" boundary condition instead of "all i" and taking a subsequential weak limit. We then have $\nu_j^{q,\beta}(X(\mathbf{0})=j) > \frac{1}{q}$, whence $\nu_j^{q,\beta} \neq \nu_i^{q,\beta}$. This proves the 'only if' direction.

For the 'if' direction, suppose that $\lim_{n\to\infty} \nu_{i,n}^{q,\beta}(X(\mathbf{0})=1) = \frac{1}{q}$. We then have that

$$\lim_{n\to\infty} \tilde{\mu}_{G_n}^{p,q}(\mathbf{0} \leftrightarrow \partial\Lambda_n) = 0. \tag{19}$$

Let $\nu^{q,\beta}$ be *any* Potts measure for $\mathbf{Z}^d$ with the given parameter values. Our task is to show that $\nu^{q,\beta} = \nu_i^{q,\beta}$, and for this it is sufficient to show that

$$\nu^{q,\beta}(C) = \nu_i^{q,\beta}(C) \tag{20}$$

for any cylinder event C. Fix C, and let n_1 be large enough so that C only depends on the spins inside Λ_{n_1}. For two sets $A, B \subset \mathbf{Z}^d$, write $A \leftrightarrow B$ for the event that there exists some $x \in A$ which is connected to some $y \in B$ via a path of present edges. Fix $\varepsilon > 0$, and pick n_2 so large that

$$\tilde{\mu}_{G_{n_2}}^{p,q}(\Lambda_{n_1} \leftrightarrow \partial\Lambda_{n_2}) < \varepsilon, \tag{21}$$

that this is possible follows from (19) and the observation that the conditional probability under $\tilde{\mu}_{G_n}^{p,q}$ that all edges in Λ_{n_1} are present given the edge configuration outside of Λ_{n_1} is bounded away from 0 uniformly in $n > n_1$ and the edge configuration off Λ_{n_1}. Similarly, we can pick n_3 so large that

$$\tilde{\mu}_{G_{n_3}}^{p,q}(\Lambda_{n_2} \leftrightarrow \partial\Lambda_{n_3}) < \varepsilon. \tag{22}$$

A random spin configuration $X \in \{1,\ldots,q\}^{\mathbf{Z}^d}$ distributed according to $\nu^{q,\beta}$ can be obtained as follows. First pick the configuration $X(\mathbf{Z}^d \setminus \Lambda_{n_3})$ of spins outside Λ_{n_3} according to the appropriate projection of $\nu^{q,\beta}$, and then pick $X(\Lambda_{n_3})$ according to the right conditional distribution. The choice of $X(\Lambda_{n_3})$ will be made by a special procedure: First define $G_{n_3}^*$ to be the graph with vertex set

$$V_{n_3}^* = \Lambda_{n_3} \cup \partial\Lambda_{n_3} \cup \{\Delta_1,\ldots,\Delta_q\},$$

where $\Delta_1,\ldots,\Delta_q$ are q auxiliary vertices, and edge set

$$E_{n_3}^* = E_{n_3} \cup \{(x,y) : x \in \partial\Lambda_{n_3}, y \in \{\Delta_1,\ldots,\Delta_q\}\}.$$

Write $\omega \in \{1,\ldots,q\}^{\partial\Lambda_{n_3}}$ for the spin configuration on $\partial\Lambda_{n_3}$ obtained as the restriction to $\partial\Lambda_{n_3}$ of $X(\mathbf{Z}^d \setminus \Lambda_{n_3})$. Recall the definitions of E'_n and E''_n, and pick a random edge configuration $Y \in \{0,1\}^{E_{n_3}^*}$ according to $\tilde{\mu}_{G_{n_3}^*}^{p,q}$ defined as the random-cluster measure $\mu_{G_{n_3}^*}^{p,q}$ conditioned on the event that

(i) for all $j = 1,\ldots,q$ and $x \in \partial\Lambda_{n_3}$, the edge between x and Δ_j is present if and only if $\omega(x) = j$,

(ii) no edge of E''_{n_3} is present, and

(iii) no two vertices $x, y \in \partial\Lambda_{n_3}$ with $\omega(x) \neq \omega(y)$ have a path of present edges in E'_{n_3} connecting them.

Once the edge configuration Y has been chosen, pick $X(\Lambda_{n_3})$ by assigning spins to the connected components of Y in such a way that

(a) for $j = 1, \ldots, q$, the connected component containing Δ_j gets spin j, and

(b) each connected component not containing any of the Δ_j's gets a spin chosen uniformly from $\{1, \ldots, q\}$, with different connected components independent.

Again, it is just a simple modification of the proof of Theorem 2.5 to check that this gives the right distribution of the spin configuration $X(\Lambda_{n_3})$. For any $e \in E'_{n_3}$ and any configuration η of edges off e, we see that

$$\tilde{\nu}^{p,q}_{G^*_{n_3}}(e \text{ is present} \mid \eta) = \begin{cases} p, \text{ if the endvertices of } e \text{ are connected in } \eta \\ 0, \text{ if for two distinct } j_1, j_2 \in \{1, \ldots, q\}, \text{ the endvertices} \\ \quad \text{of } e \text{ are connected in } \eta \text{ to, respectively, } \Delta_{j_1} \text{ and } \Delta_{j_2} \\ \frac{p}{p+(1-p)q}, \text{ otherwise.} \end{cases}$$

Applying Holley's Theorem we get that the projection on $\{0,1\}^{E'_{n_3}}$ of $\tilde{\nu}^{p,q}_{G_{n_3}}$ stochastically dominates the projection on $\{0,1\}^{E'_{n_3}}$ of $\tilde{\nu}^{p,q}_{G^*_{n_3}}$. We shall now describe a specific coupling of two $\{0,1\}^{E'_{n_3}}$-valued random elements Y and Y^* with these distributions, and for which $Y \succeq Y^*$ a.s. This coupling involves a kind of sequential construction called "disagreement percolation", which is highly useful also in a number of other situations (see [6]). Let $e_1, e_2, \ldots$ be some arbitrary enumeration of the edges of E'_{n_3}. At each step k of the sequential construction, select the edge with the smallest index i satisfying

(1) e_i has not been selected in any previous step, and

(2) e_i is incident either to some vertex of $\partial\Lambda_{n_3}$ or to some previously selected edge which has been assigned value 1 for the Y realization.

Let U_k be a uniform $[0,1]$ random variable independent of everything else so far, and let

$$Y(e_i) = \max\{a \in \{0,1\} : \tilde{\nu}^{p,q}_{G_{n_3}}(Y(e_i \geq a) \mid \text{the } Y \text{ values determined so far}) \geq U_k\}$$

and

$$Y^*(e_i) = \max\{a \in \{0,1\} : \tilde{\nu}^{p,q}_{G^*_{n_3}}(Y^*(e_i \geq a) \mid \text{the } Y^* \text{ values determined so far}) \geq U_k\}.$$

This guarantees that $Y(e_i) \geq Y^*(e_i)$. Usually, it will happen at some stage before the construction is complete that no edge satisfying (1) and (2) can be found. Write $\bar{E}$ for the (random) subset of E'_{n_3} consisting of those edges whose Y and Y^* values are yet to be determined when this happens. Clearly, all edges of $E'_{n_3} \setminus \bar{E}$ that are incident to some edge of $\bar{E}$ have value 0 for both Y and Y^*. Therefore, the conditional distributions of $Y(\bar{E})$ and $Y^*(\bar{E})$ coincide and are equal to the random-cluster measure with parameters p and q on a graph with edge set $\bar{E}$ and the appropriate vertex set. We end the sequential construction by taking $Y(\bar{E})$ and $Y^*(\bar{E})$ to be identical, the realization chosen according to this random-cluster measure.

We can then obtain spin configurations distributed according to $\nu^{q,\beta}_{i,n_3}$ and the projection on Λ_{n_3} of $\nu^{q,\beta}$ by assigning spins to the connected components of $Y(E'_{n_3})$ and $Y^*(E'_{n_3})$ as described above. Let A be the event that $Y(E'_{n_3})$ and $Y^*(E'_{n_3})$ coincide inside the box Λ_{n_2} and that furthermore all connected components which intersect Λ_{n_1} are contained in Λ_{n_2}. When A occurs, the spin choices can be made in such a way that the two spin configurations are identical on Λ_{n_1}. By (21) and (22), A has probability at least $1-2\varepsilon$, whence

$$|\nu^{q,\beta}(C)-\nu^{q,\beta}_{i,n_3}(C)|<2\varepsilon.$$

Since n_3 could be taken arbitrarily large, we have

$$|\nu^{q,\beta}(C)-\nu^{q,\beta}_{i}(C)|<2\varepsilon,$$

and since ε was arbitrary, we have (20) and the proof is complete. □

4 The infinite-volume random-cluster model

We saw in the previous section that Ising and Potts models could be generalized from finite graphs to the infinite lattice $\mathbf{Z}^d$, using the Dobrushin–Lanford–Ruelle equation (5). Note that the use of random-cluster representations for studying these infinite systems only involved random-cluster measures on finite graphs; the corresponding thing holds for the various other systems that we will study in Sections 5–8. Nevertheless, it is natural to ask whether the random-cluster model itself can be generalized from finite to infinite graphs in the Dobrushin–Lanford–Ruelle spirit. This was recently done by Grimmett [39], and independently by Pfister and Vande Velde [73] and Borgs and Chayes [8]. The definition of a random-cluster measure for $\mathbf{Z}^d$ can be formulated as follows.

Write $\mathbf{E}^d$ for the edge set of the nearest neighbour graph on $\mathbf{Z}^d$. Pick p and q satisfying $0\le p\le 1$ and $q>0$. For a finite set $S\subset \mathbf{E}^d$, let S' denote the set $\{v\in\mathbf{Z}^d : \exists e\in S \text{ such that } e \text{ is incident to } v\}$. For a configuration $\xi\in\{0,1\}^{\mathbf{E}^d\setminus S}$ of edges off S, let the random-cluster measure $\mu_S^{p,q}$ "on S with boundary condition ξ" be given by

$$\mu^{p,q}_{S,\xi}(\eta)=\frac{1}{Z^{p,q}_{S,\xi}}\left\{\prod_{e\in S}p^{\eta(e)}(1-p)^{1-\eta(e)}\right\}q^{k(\eta,\xi)} \tag{23}$$

for all $\eta\in\{0,1\}^S$. Here $k(\eta,\xi)$ is the number of connected components which intersect S' in the configuration which agrees with η on S and with ξ on $\mathbf{E}^d\setminus S$. As before, we denote a random edge-configuration by Y.

Definition 4.1: *A probability measure μ on $\{0,1\}^{\mathbf{E}^d}$ is called a random-cluster measure with parameters p and q if its conditional probabilities satisfy*

$$\mu(Y(S)=\eta\,|\,Y(\mathbf{E}^d\setminus S)=\xi)=\mu^{p,q}_{S,\xi}(\eta) \tag{24}$$

for all finite $S\subset\mathbf{E}^d$, all $\eta\in\{0,1\}^S$ and μ-a.e. $\xi\in\{0,1\}^{\mathbf{E}^d\setminus S}$.

This gives rise to a consistent set of conditional probabilities. To argue that it is a natural extension of the finite graph case, note that random-cluster measures on finite graphs have conditional probabilities of the same form.

The existence of measures having the prescribed conditional probabilities is not quite as easily proved for the random-cluster measures as for Ising and Potts models. The difficulty in the random-cluster case is that, whereas in Ising and Potts models the conditional distribution of the spin at a site given the spins at all other sites depends on nearest neighbours only, for random-cluster measures the conditional probability that an edge is present given everything else may depend on the status of edges arbitrarily far away (see Lemma 2.9). Nevertheless, random-cluster measures for $\mathbf{Z}^d$ do exist for any choice of p and q, as stated in the following result due to Grimmett:

Proposition 4.2: *For any $p \in [0,1]$ and $q > 0$, there exists at least one random-cluster measure for $\mathbf{Z}^d$ with parameters p and q.*

Proof: We may assume that $p \in (0,1)$, since otherwise the conclusion is trivial. Suppose first that $q \geq 1$. We saw in Section 3 that the measures $\tilde{\mu}_{G_n}^{p,q}$ form a sequence of probability measures which is stochastically decreasing (there we only considered integer q, but the arguments extend verbatim to noninteger $q \geq 1$). Therefore, these measures converge weakly to some probability measure $\mu_{wired}^{p,q}$ on $\{0,1\}^{\mathbf{E}^d}$ (we call it $\mu_{wired}^{p,q}$ because it is a limit of random-cluster measures with "wired" boundary conditions; recall that all edges in E_n'' are present $\tilde{\mu}_{G_n}^{p,q}$-a.s.). This measure is our candidate for a random-cluster measure, and what we need to do is to check that it satisfies (24) for every finite $S \subset \mathbf{E}^d$, every $\eta \in \{0,1\}^S$ and $\mu_{wired}^{p,q}$-a.e. $\xi \in \{0,1\}^{\mathbf{E}^d \setminus S}$. It is easy to see, using monotonicity, that $\mu_{wired}^{p,q}$ is translation invariant. Furthermore, the conditional probability under $\mu_{wired}^{p,q}$ that an edge is present given everything else is always in the interval $[\frac{p}{p+(1-p)q}, p]$; this follows from the observation that the same holds for $\tilde{\mu}_{G_n}^{p,q}$ for each n. Hence $\mu_{wired}^{p,q}$ satisfies the so called "finite energy condition" which is the property that all conditional probabilities on finite sets are strictly positive. We can thus apply the famous Burton–Keane Uniqueness Theorem [11], which states that if a probability measure on $\{0,1\}^{\mathbf{E}^d}$ satisfies both translation invariance and finite energy, then it is concentrated on the event that the edge configuration contains at most one infinite connected component. For any finite $S \subset \mathbf{Z}^d$, we also get that $\mu_{wired}^{p,q}(A_S) = 1$, where A_S is the event that we would have at most one infinite connected component if all edges in S were switched to value 0. On the event A_S, we can always find, given the configuration ξ of edges off S, a box Λ_{n_1} large enough so that we can tell which vertices of S' are connected to each other via edges present in ξ by looking only at edges inside Λ_{n_1}. For all $n > n_1$ the conditional distribution under $\tilde{\mu}_{G_n}^{p,q}$ of the configuration on S given ξ satisfies (24) and is independent of what happens outside Λ_{n_1}. Hence the same thing holds for the limiting measure $\mu_{wired}^{p,q}$. This proves the result for $q \geq 1$.

The proof for the $q < 1$ case is similar, except that one has to be a bit less explicit in the construction of a candidate measure (the monotonicity used for $q \geq 1$ does not work for $q < 1$, due to the fact that the conditional probability that an edge e is present given the configuration ξ of other edges fails to be increasing in ξ). As a candidate, we consider some subsequential weak limit of random-cluster measures on finite boxes increasing to $\mathbf{Z}^d$. Translation invariance of the limit can be ensured by using torus (wrap-around) boundary conditions for the boxes. The Burton–Keane Uniqueness Theorem then finishes the proof as for the $q \geq 1$ case. □

Now that the existence of random-cluster measures for $\mathbf{Z}^d$ is established, an obvious next question (in analogy with phase transition considerations for Ising and Potts models) is whether there is a *unique* random-cluster measure with given parameters. I will only give a short introduction to this highly intricate issue here, and refer to Grimmett [39] for a more thorough discussion. The following paragraph is a slight digression.

For $q = 2, 3 \ldots$, one might at first think that the question of uniqueness of random-cluster measures would be intimately connected with uniqueness of Potts measures with the corresponding parameter values. This, however, is the wrong intuition. Instead, we have the following:

> *Phase transition for the Potts model corresponds to the occurrence of infinite clusters in the random-cluster model.*

More precisely, the q-state Potts model for $\mathbf{Z}^d$ at reciprocal temperature β has more than one Gibbs measure if and only if $\mu^{p,q}_{wired}$ (with $p = 1 - e^{-2\beta}$) assigns positive probability to the existence of infinite clusters. This follows from arguments in Section 3.

Now, returning to the issue of uniqueness of random-cluster measures, let us first restrict to the case $q \geq 1$. We can define the probability measure $\mu^{p,q}_{free}$ on $\{0,1\}^{\mathbf{E}^d}$ as a weak limit of the measures $\mu^{p,q}_{G_n}$ as $n \to \infty$. In other words, $\mu^{p,q}_{free}$ is obtained as $\mu^{p,q}_{wired}$ except that we do not at the finite stages condition on the boundary being wired. Again, existence of the limit follows by monotonicity (this time, though, the measures $\mu^{p,q}_{G_n}$ are stochastically *increasing*). By the same arguments as those used for the Ising model in Section 3, we have that

$$\mu^{p,q}_{free} \preceq \mu^{p,q} \preceq \mu^{p,q}_{wired}$$

for any random-cluster measure $\mu^{p,q}$ for $\mathbf{Z}^d$ with the given parameters. Hence, uniqueness of random-cluster measures is equivalent to having $\mu^{p,q}_{free} = \mu^{p,q}_{wired}$.

In order to state the generally agreed belief concerning uniqueness of random-cluster measures when $q \geq 1$, we first need to define the percolation probability

$$\theta(p,q) = \mu^{p,q}_{wired}(\text{the origin is in an infinite cluster}).$$

By now familiar monotonicity arguments also imply that for $p_1 \leq p_2$ we have

$$\mu^{p_1,q}_{wired} \preceq \mu^{p_2,q}_{wired}$$

whence $\theta(p,q)$ for fixed q is an increasing function of p. Let

$$p_c(q) = \inf\{p : \theta(p,q) > 0\},$$

i.e. $p_c(q)$ is the critical value of p needed to get infinite clusters in $\mu^{p,q}_{wired}$. The following is conjectured in [39] and partly in [8]. Write $\mathcal{P}(q)$ for the set of p-values for which there is more than one random-cluster measure with parameters p and q.

Conjecture 4.3: *For the random-cluster model on $\mathbf{Z}^d$, $d \geq 2$, there exists $Q = Q(d) \geq 2$ such that*

$$\mathcal{P}(q) = \begin{cases} \emptyset & \text{if } 1 \leq q < Q \\ \{p_c(q)\} & \text{if } q > Q. \end{cases}$$

In other words, phase transitions for the random-cluster model should occur very rarely in the parameter space: only for large q and with p equal to the percolation threshold $p_c(q)$. Some rigorous progress has been made towards a proof of this conjecture. The arguments due to Aizenman *et al.* [3] used in the proof of Lemma 3.10 also show that there is a unique random-cluster measure whenever $p < p_c(q)$. Grimmett [39] showed that for fixed $q \geq 1$, there is a unique random-cluster measure for all but at most countably many values of p. The results in [55] and [57] imply that there is more than one random-cluster measure for large enough integer q and $p = p_c(q)$; a distinguishing feature for such (p,q) is that there exist (a.s.) infinite clusters under $\mu^{p,q}_{wired}$ but not under $\mu^{p,q}_{free}$.

In contrast, hardly anything is known about the phase transition behaviour for $q < 1$, due to the lack of monotonicity arguments in this part of the parameter space. However, it is hard to imagine any mechanism that would boost a phase transition in this case, so maybe one should conjecture the absence of phase transition throughout the $q < 1$ regime of the parameter space.

There are of course other infinite graphs, apart from $\mathbf{Z}^d$, on which the random-cluster model can be defined and studied. One interesting case is the regular tree of order $n \geq 2$, denoted by $\mathbf{T}_n$ and defined as the (unique) infinite connected graph which has no circuits and in which each vertex is incident to exactly $n+1$ edges. The random-cluster model on $\mathbf{T}_n$ is studied in [43] and [44], motivated by previous studies of Ising and Potts models on $\mathbf{T}_n$ (see e.g. [29]), and also by the general philosophy that when a question (such as Conjecture 4.3) about a $\mathbf{Z}^d$-indexed random process appears to be "too difficult", then one can sometimes shed light on it by studying the corresponding $\mathbf{T}_n$-indexed random process.

It is not immediately clear how the random-cluster model on $\mathbf{T}_n$ should be defined. For finite trees, the definition of finite graph random-cluster measures leads simply to i.i.d. edges with probability $\frac{p}{p+(1-p)q}$ of being present; this is a direct consequence of Lemma 2.9. The i.i.d. behaviour carries over to $\mathbf{T}_n$ if Definition 4.1 is brought over directly to the $\mathbf{T}_n$ case, so such a definition is nothing but a complicated way of defining standard bond percolation on a tree. A different definition is proposed in [43] and [44], where Definition 4.1 is brought over to the tree case with the important modification that the exponent $k(\eta, \xi)$ in (23) is defined as the number of *finite* connected components that intersect S'. It is only with this alternative definition that the correspondence between phase transition in the Potts model and infinite clusters in the random-cluster model carries over to $\mathbf{T}_n$. The two different ways of counting connected components are (trivially) equivalent for finite graphs, but probably also on $\mathbf{Z}^d$, since it is believed that random-cluster measures for $\mathbf{Z}^d$ give rise a.s. to at most one infinite cluster (if we restrict to translation invariant measures, then this is certainly the case due to the Burton–Keane Uniqueness Theorem). The basic existence and uniqueness questions are studied in [43]: The phase transition behaviour on $\mathbf{T}_n$ (for the version of the model where only finite clusters count) turns out to be a bit different from that on $\mathbf{Z}^d$, in that for $q > 2$ the set $\mathcal{P}(q)$ contains an entire interval of p-values. The existence result (Proposition 4.2) carries over to $\mathbf{T}_n$, although it requires a different (and more constructive) proof, due to the failure of the "uniqueness of the infinite cluster" property on $\mathbf{T}_n$.

5 Ising model with random interactions

In this and the following three sections, we shall see examples of how the random-cluster methods of Sections 2 and 3 can be adapted in order to illuminate the phase transition behaviour of systems other than the standard Ising and Potts models. This section is devoted to **spin glasses**, i.e. to Ising models with random interactions. The study of such systems is a highly active research area in statistical mechanics; see Petritis [71] for a general survey. The arguments of this section are taken from Aizenman *et al.* [2] and Newman [67], and go through also for $q \geq 3$ Potts models with random interactions, but for simplicity and tradition we stick to the Ising case $q = 2$.

Consider first the following "inhomogeneous" version of the Ising model on a finite graph. Let G be a finite graph with vertex set V and edge set E, and let $\{J_e\}_{e \in E}$ be non-negative real numbers, called the coupling constants.

Definition 5.1: *The Gibbs measure* $\nu_G^{\{J_e\},\beta}$ *on* $\{-1,1\}^V$ *for the Ising model on* G *with coupling constants* $\{J_e\}_{e\in E}$ *and reciprocal temperature* β *is the measure for which*

$$\nu_G^{\beta}(\omega) = \frac{1}{Z_G^{\{J_e\},\beta}} \exp\{-H\beta(\omega)\}$$

for all $\omega \in \{-1,1\}^V$. *Here*

$$H(\omega) = 2 \sum_{\substack{x\sim y \\ x,y\in V}} J_{e(x,y)} \mathbf{1}_{\{\omega(x)\neq\omega(y)\}}.$$

Obviously, taking $J_e = 1$ for each $e \in E$ gives the standard Ising model. Having different coupling constants for different edges corresponds to having different strengths of interaction between different pairs of neighbouring vertices.

The corresponding "inhomogenisation" of the random-cluster model is to have different values of p for different edges:

Definition 5.2: *The random-cluster measure* $\mu_G^{\{p_e\},q}$ *for* G *with parameters* $q > 0$ *and* $\{p_e\}_{e\in E}$ *taking values in* $[0,1]$ *is the probability measure on the set of subgraphs of* G *given by*

$$\mu_G^{\{p_e\},q}(\eta) = \frac{1}{Z_G^{\{p_e\},q}} \left\{ \prod_{e\in E} p_e^{\eta(e)}(1-p_e)^{1-\eta(e)} \right\} q^{k(\eta)}$$

for all $\eta \in \{0,1\}^E$, *where as before* $k(\eta)$ *is the number of connected components of* η.

The proof of Theorem 2.5 carries over directly to the inhomogeneous case, and as a consequence we get the following analogue of Corollary 2.6:

Proposition 5.3: *A random spin configuration* $X \in \{-1,1\}^V$ *distributed according to the Ising measure* $\nu_G^{\{J_e\},\beta}$ *can be obtained as follows. For each* $e \in E$, *let* $p_e = 1-\exp(-2\beta J_e)$, *and pick an edge configuration* $Y \in \{0,1\}^E$ *according to the random-cluster measure* $\mu_G^{\{p_e\},2}$. *Then, for each connected component* $\mathcal{C}$ *of* Y, *pick a spin at random (uniformly) from* $\{-1,1\}$, *assign this spin to every vertex of* $\mathcal{C}$, *and do this independently for different connected components.*

Next, the inhomogeneous Ising model can of course be generalised to a $\mathbf{Z}^d$ setting in the exact same way as the standard Ising model (Definition 3.2). The spin glass model is simply an inhomogeneous Ising model on $\mathbf{Z}^d$ where the coupling "constants" $\{J_e\}_{e\in\mathbf{E}_n}$ are chosen *at random*. Thus, the spin glass model contains randomness on two levels:

1. First, the coupling constants $\{J_e\}_{e\in\mathbf{E}_n}$ are chosen according to some probability measure Π. Here, and throughout most of the spin glass literature, it is assumed that the coupling constants are i.i.d. We write π for the distribution of a single J_e.

2. Given $\{J_e\}_{e\in\mathbf{E}_n}$, the spins are chosen according to some Gibbs measure for the ingomogeneous Ising model on $\mathbf{Z}^d$ with coupling constants $\{J_e\}_{e\in\mathbf{E}_n}$ and reciprocal temperature β.

The basic question is whether there is a unique Gibbs measure in Step 2 of this procedure. It is easy to see that the answer to this question is independent of $\{J_e\}_{e\in E}$ for any finite edge set $E \subset \mathbf{E}^d$, so by Kolmogorov's 0-1 law the answer must be either yes Π-a.s. or no Π-a.s. The arguments used for the standard Ising model in Section 3 imply the existence of two particular Gibbs measures $\nu_+^{\{J_e\},\beta}$ and $\nu_-^{\{J_e\},\beta}$, obtained as weak limits with plus and minus boundary conditions, respectively. Just as for the standard Ising model, uniqueness of Gibbs measures is equivalent to $\nu_+^{\{J_e\},\beta}$ and $\nu_-^{\{J_e\},\beta}$ being identical. The following result, in which p_c as before denotes the critical value for standard bond percolation on $\mathbf{Z}^d$, provides a partial answer. The proof is so similar to the proof of Theorem 1.1 that a short sketch will suffice; the reader (having come this far) will have no trouble filling in the details.

Theorem 5.4: *Consider the spin glass model on $\mathbf{Z}^d$, $d \geq 2$, with non-negative coupling constant distribution π.*

(i) If $\pi(J_e = 0) > 1 - p_c$, then there is Π-a.s. a unique Gibbs measure for any β.

(ii) If on the other hand $\pi(J_e = 0) < 1 - p_c$, then there exists $\beta_c = \beta_c(d, \pi)$ such that there is Π-a.s. a unique Gibbs measure when $\beta < \beta_c$, and Π-a.s. more than one Gibbs measure when $\beta > \beta_c$.

Sketch proof: Let $p_e = 1 - \exp(-2\beta J_e)$ for each $e \in \mathbf{E}_n$. For (i), note that $\pi(J_e = 0) > 1 - p_c$ implies that $\Pi(p_e = 0) > 1 - p_c$. Hence, the edge set $\{e \in \mathbf{E}^d : p_e > 0\}$ does not form any infinite cluster Π-a.s. Using Proposition 5.3, we thus get that the spin at a given $x \in \mathbf{Z}^d$ is asymptotically independent of the spins on $\partial\Lambda_n$ as $n \to \infty$. Uniqueness now follows as in the proof of Lemma 3.7.

To get (ii), write $p'_e = \frac{p_e}{2 - p_e}$, note that

$$\mu_G^{\{p_e\},2}(e \text{ is present} \mid \text{all other edges}) = \begin{cases} p_e & \text{if } e\text{'s endvertices are already connected} \\ p'_e & \text{otherwise,} \end{cases}$$

and use Holley's Theorem to deduce that

$$\mu_G^{\{p'_e\},1} \preceq_d \mu_G^{\{p_e\},2} \preceq_d \mu_G^{\{p_e\},1} \tag{25}$$

for any G. Write m (resp. m') for the expected value of p_e (resp. p'_e) under Π. Since the p_e's are i.i.d., we have that the unconditional distribution (i.e. averaged over Π) of the edge configuration is i.i.d. with edge probability m (resp. m') under $\mu_G^{\{p_e\},1}$ (resp. $\mu_G^{\{p'_e\},1}$). Uniqueness now follows as for (i) when β is small using the second inequality in (25) and the observation that m tends to 0 as $\beta \to 0$. Phase transition for large β follows similarly, using the first inequality of (25) and $\lim_{\beta\to\infty} m = 1 - \pi(J_e = 0) > p_c$. Finally, monotonicity in β follows from a translation of (17) to the inhomogeneous case. $\Box$

Let me end this section by pointing out that the restriction to non-negative coupling constants can be described as "spin glasses for beginners". The really difficult questions arise for systems where the coupling constants can take both positive and negative values. Newman [67] shows how the approach sketched above can be carried further, so as to actually shed some light on this more difficult case.

6 The Ashkin–Teller model

The **Ashkin–Teller model** [4] is a spin system which is of interest because it exhibits a double critical phenomenon: it has two critical values $\beta_1 < \beta_2$ such that the behaviour of the model is qualitatively different for β chosen in the high temperature regime $\beta < \beta_1$, the low temperature regime $\beta > \beta_2$ and the intermediate regime $\beta \in (\beta_1, \beta_2)$. Each site $x \in \mathbf{Z}^d$ takes a value from the spin set $\{A, B, C, D\}$. The spin set should be thought of as hierarchical: A and B are closely related to each other, as are C and D, whereas the relationship between $\{A, B\}$ and $\{C, D\}$ is more distant. Fix two coupling constants J_1 and J_2 such that $0 \leq J_1 \leq J_2$, and define the symmetric pair interaction function $h : \{A, B, C, D\}^2 \to \mathbf{R}$ as

$$\begin{aligned}
&h(A, A) = h(B, B) = h(C, C) = h(D, D) = 0 \\
&h(A, B) = h(C, D) = J_1 \\
&h(A, C) = h(A, D) = h(B, C) = h(B, D) = J_2.
\end{aligned}$$

As usual, we first define the model on a finite graph.

Definition 6.1: *The Gibbs measure $\nu_G^{J_1,J_2,\beta}$ on $\{A, B, C, D\}^V$ for the Ashkin–Teller model on G at reciprocal temperature β is the probability measure on $\{A, B, C, D\}^V$ which assigns probability*

$$\nu_G^{J_1,J_2,\beta}(\omega) = \frac{1}{Z_G^{J_1,J_2,\beta}} \exp\{-\beta H(\omega)\}$$

to each $\omega \in \{A, B, C, D\}^V$. Here

$$H(\omega) = 2 \sum_{\substack{x \sim y \\ x,y \in V}} h(\omega(x), \omega(y)).$$

This means e.g. that an A prefers to sit next to A's, likes B's a bit worse, and hates to sit next to C's and D's. Since $H(\omega)$ only contains terms involving nearest neighbours, this defines a Markov random field on G, and it generalizes in the obvious way (analogously to Definition 3.2) to a $\mathbf{Z}^d$ setting. The model can be seen as a kind of interpolation between the Ising model and the 4-state Potts model: taking $J_1 = 0$ and identifying $(\{A, B\}, \{C, D\})$ with $(-1, 1)$ gives the Ising model, whereas for $J_1 = J_2$ the hierarchical structure of the model vanishes and the $q = 4$ Potts model is obtained.

The phase transition behaviour mentioned above is made precise in the following theorem. Pfister [72] proved it using Griffiths inequalities; here we shall indicate how it can be proved using a random-cluster representation.

Theorem 6.2: *Fix J_1 and J_2 such that $0 < J_1 < J_2$. If the ratio $\frac{J_2}{J_1}$ is sufficiently large, then the Ashkin–Teller model on $\mathbf{Z}^d$, $d \geq 2$, with coupling constants J_1 and J_2 has two critical values β_1 and β_2 such that $0 < \beta_1 < \beta_2 < \infty$ and the following holds.*

(a) *At reciprocal temperature $\beta < \beta_1$, there is a unique Gibbs measure.*

(b) *For $\beta \in (\beta_1, \beta_2)$, there is more than one Gibbs measure, but all Gibbs measures are invariant under permutations of $\{A, B\}$ and of $\{C, D\}$.*

(c) *For $\beta > \beta_2$ there exist Gibbs measures in which a single spin dominates. For instance, there exists a Gibbs measure $\nu^{J_1,J_2,\beta}$ such that for each $x \in \mathbf{Z}^d$ the probability that site x takes value A is strictly greater than $\frac{1}{4}$ whereas the corresponding probabilities for spins B, C, and D are strictly smaller than $\frac{1}{4}$.*

In other words, β_1 is the critical value for breaking the symmetry between $\{A, B\}$ and $\{C, D\}$, whereas β_2 is the critical value for breaking the symmetry within $\{A, B\}$ or $\{C, D\}$.

Random-cluster representations for a somewhat more general version of the Ashkin–Teller model appear in papers by Wiseman and Domany [88], Salas and Sokal [80], and Pfister and Velenik [74]. The random-cluster representation given below is a simplified special case of those in [88], [80] and [74]. We call this variant of the random-cluster model the ATRC (Ashkin–Teller random-cluster) model. In contrast to the random-cluster models of the previous sections, the ATRC model allows an edge to be in three different states: 0 (absent), 1 (weakly present), and 2 (strongly present).

Definition 6.3: *The ATRC measure $\mu_G^{(p_0,p_1,p_2)}$ with parameters $p_0, p_1, p_2 \in [0,1]$ satisfying $p_0 + p_1 + p_2 = 1$ is the probability measure on $\{0,1,2\}^E$ which to each element $\eta \in \{0,1,2\}^E$ assigns probability*

$$\mu_G^{(p_0,p_1,p_2)} = \frac{1}{Z_G^{(p_0,p_1,p_2)}} \left\{ \prod_{e \in E} p_0^{\mathbf{1}\{\eta(e)=0\}} p_1^{\mathbf{1}\{\eta(e)=1\}} p_2^{\mathbf{1}\{\eta(e)=2\}} \right\} 2^{k_1(\eta)+k_2(\eta)}.$$

Here $k_1(\eta)$ is the number of connected components of the edge set $\{e \in E : \eta(E) \in \{1,2\}\}$ and $k_2(\eta)$ is the number of connected components of $\{e \in E : \eta(E) = 2\}$.

Let us now see how the Ashkin–Teller model and the ATRC model can be coupled analogously to the coupling in Section 2 of the Potts model and the standard random-cluster model. Let $P_G^{(p_0,p_1,p_2)}$ be the probability measure on $\{A,B,C,D\}^V \times \{0,1,2\}^E$ corresponding to the following procedure:

1. Assign each vertex a spin value chosen uniformly from $\{A, B, C, D\}$ and each edge a value chosen from $\{0, 1, 2\}$ according to the probability vector (p_0, p_1, p_2), and do this independently for all vertices and edges.

2. Condition on the event that

   (i) for each edge e with value 1, the two endvertices of e have spins which are either both in $\{A, B\}$ or both in $\{C, D\}$, and

   (ii) for each edge e with value 2, the two endvertices of e have the same spin.

The following result, which is analogous to Theorem 2.5 and which is easily proved by the same kind of counting argument as in the proof of Theorem 2.5, tells us that $P_G^{(p_0,p_1,p_2)}$ really is a coupling of the Ashkin–Teller model and the ATRC model.

Proposition 6.4: *Given β, J_1 and J_2, let $p_0 = e^{-2\beta J_2}$, $p_1 = e^{-2\beta J_1} - e^{-2\beta J_2}$ and $p_2 = 1 - e^{-2\beta J_1}$. Write $P_{G,vertex}^{(p_0,p_1,p_2)}$ (resp. $P_{G,edge}^{(p_0,p_1,p_2)}$) for the probability measure obtained by projecting $P_G^{(p_0,p_1,p_2)}$ on $\{A,B,C,D\}^V$ (resp. $\{0,1,2\}^E$). Then*

$$P_{G,vertex}^{(p_0,p_1,p_2)} = \nu_G^{J_1,J_2,\beta}$$

and

$$P_{G,edge}^{(p_0,p_1,p_2)} = \mu_G^{(p_0,p_1,p_2)}.$$

The following analogue of Corollary 2.6 is an immediate consequence:

Corollary 6.5: *Given β, J_1 and J_2, let p_0, p_1 and p_2 be as in Proposition 6.4. Suppose we pick a random spin configuration $X \in \{A,B,C,D\}^V$ as follows:*

1. *First, pick a random edge configuration $Y \in \{0,1,2\}^E$ according to the ATRC measure $\mu_G^{(p_0,p_1,p_2)}$.*
2. *Then, for each connected component $\mathcal{C}$ of the edge set $\{e \in E : Y(e) \in \{1,2\}\}$, flip a fair coin to determine whether the spins in $\mathcal{C}$ will be taken from $\{A,B\}$ or from $\{C,D\}$, and do this independently for all $\mathcal{C}$.*
3. *Finally, for each connected component $\mathcal{C}'$ of the edge set $\{e \in E : Y(e) = 2\}$, flip a fair coin to determine which of the two possible spins (consistent with Step 2) the vertices of $\mathcal{C}'$ will get, and do this independently for all $\mathcal{C}'$.*

Then X is distributed according to the Ashkin–Teller measure $\nu_G^{J_1,J_2,\beta}$.

Theorem 6.2 can now be proved by a straightforward modification of the proof of Theorem 1.1. The behaviours in (a) and (c) follow exactly as in the proof of Theorem 1.1; the (a) scenario corresponds to having no infinite clusters of edges taking values in $\{1,2\}$, and the (c) scenario corresponds to having infinite clusters of edges taking value 2. To show also that $\beta_1 < \beta_2$ so that the behaviour in (b) also occurs for some parameter values, the following key observation is needed. If $\frac{J_2}{J_1}$ is sufficiently large, then β can be chosen in such a way that the p_1 parameter of the corresponding ATRC model is large enough to ensure that edges taking values in $\{1,2\}$ dominate i.i.d. edges with edge probability greater the critical value p_c of standard bond percolation on $\mathbf{Z}^d$, whereas edges taking value 2 are dominated by some subcritical percolation process.

We end this section with an open problem: It seems reasonable to expect that the proviso that $\frac{J_2}{J_1}$ is large can be dropped in Theorem 6.2. It would be interesting to see a proof of this, either by some refinement of the above random-cluster argument or by some other approach.

7 Subshifts of finite type

Let S be a finite set. A d-dimensional **subshift of finite type** $\mathcal{X}$ is defined as the subset of $S^{\mathbf{Z}^d}$ where certain prespecified finite patterns ("forbidden words") do not occur. Only finitely many disallowed configurations may be specified. Perhaps the simplest nontrivial example of a subshift of finite type is the hard core model, where $\mathcal{X}$ consists of those elements of $\{0,1\}^{\mathbf{Z}^d}$ in which no two 1's sit next to each other anywhere in the lattice.

Subshifts of finite type have been studied mainly in ergodic theory; see e.g. [81] for a general discussion. We shall be concerned with probability measures on $\mathcal{X}$; of particular interest are so called measures of maximal entropy. Assuming translation invariance and a certain regularity condition on $\mathcal{X}$ called strong irreducibility (see Definition 7.1 below), these are exactly the measures where the conditional distribution on any finite set $W \subset \mathbf{Z}^d$ given the configuration off W is uniform over all elements of S^W that do not give rise to any forbidden word. We call measures on $\mathcal{X}$ satisfying this uniform conditional probability property **maximal**.

As long as a subshift of finite type $\mathcal{X}$ is nonempty, there has to exist at least one maximal probability measure on $\mathcal{X}$; this follows from standard compactness arguments. In analogy with the phase transition questions of the previous sections, the issue here is whether there can exist more than one maximal measure.

To see the need for some irreducibility condition, consider the subshift of finite type $\mathcal{X}$ consisting of those elements of $\{0,1\}^{\mathbf{Z}^d}$ in which no 1 ever sits next to a 0. There are exactly two elements of $\mathcal{X}$, namely the configuration with all 1's and the configuration with all 0's. There is a continuum of probability measures on $\mathcal{X}$, parameterized by the probability of having all 1's, and it is obvious that these measures are all maximal. This, however, is a totally uninteresting example.

If $\mathcal{X}$ is a subshift of finite type, $W \subset \mathbf{Z}^d$ is finite, and η is a configuration on W (i.e. an element of S^W), we say that η is **compatible** (with $\mathcal{X}$) if there exists an element of $\mathcal{X}$ whose projection on S^W is η.

Definition 7.1: *The subshift of finite type $\mathcal{X}$ is said to be* **strongly irreducible** *if there exists an $r \geq 0$ such that whenever W_1 and W_2 are finite subsets of $\mathbf{Z}^d$ that are separated by distance at least r and both $\eta_1 \in S^{W_1}$ and $\eta_2 \in S^{W_2}$ are compatible configurations, then there exists some element of $\mathcal{X}$ whose projection on S^{W_1} is η_1 and whose projection on S^{W_2} is η_2.*

If we restrict to strongly irreducible subshifts of finite type, then we have the following dichotomy: A 1-dimensional strongly irreducible subshift of finite type has a unique maximal measure (as shown by Parry [69]), whereas for $d \geq 2$ there exist d-dimensional strongly irreducible subshifts of finite type that have more than one maximal measure. The analogy with Markov random fields is obvious, and we refer to the nonuniqueness of maximal measures as a phase transition.

The following two examples of subshifts of finite type in dimensions $d \geq 2$ exhibiting phase transition are due to Burton and Steif [12] [13].

Example 7.2: The beach model. Let M_1 and M_2 be positive integers such that $M_1 < M_2$, and let the symbol set be

$$F = F_1 \cup F_2 \cup F_3 \cup F_4 \tag{26}$$

where

$$\begin{aligned} F_1 &= \{-M_2, -M_2+1, \ldots, -M_1-1\} \\ F_2 &= \{-M_1, -M_1+1, \ldots, -1\} \\ F_3 &= \{1, 2, \ldots, M_1\} \\ F_4 &= \{M_1+1, M_1+2, \ldots, M_2\}. \end{aligned}$$

Call a symbol $f \in F$

$$\begin{cases} \text{negative} & \text{if} \quad f \in F_1 \cup F_2 \\ \text{positive} & \text{if} \quad f \in F_3 \cup F_4 \\ \text{unprivileged} & \text{if} \quad f \in F_1 \cup F_4 \\ \text{privileged} & \text{if} \quad f \in F_2 \cup F_3, \end{cases}$$

and consider the d-dimensional subshift of finite type where a negative may not sit next to a positive unless they are both privileged. Write M for the ratio $\frac{M_2}{M_1}$. The name "beach model" comes from the interpretation in two dimensions that if a symbol represents altitude above sea level, then the rules of the model prevent shores from being too steep. Burton and Steif [12] showed that for large enough M, the model exhibits a phase transition. It was then shown in [42] that there is a unique maximal measure for M sufficiently close to 1, and furthermore that the occurrence of a phase transition is monotone in M, so that there is a critical M_c below which there is uniqueness and above which there in nonuniqueness of maximal measures.

Example 7.3: The iceberg model. Let N_1 and N_2 be positive integers, and define the symbol set $H = H_- \cup H_0 \cup H_+$, where

$$\begin{aligned} H_- &= \{-N_2, -N_2+1, \ldots, -1\} \\ H_0 &= \{0_1, 0_2, \ldots, 0_{N_1}\} \\ H_+ &= \{1, 2, \ldots, N_2\}. \end{aligned}$$

The iceberg model in $d \geq 2$ dimensions is the subshift of finite type consisting of all configurations $\omega \in H^{\mathbf{Z}^d}$ in which no element of H_- sits next to an element of H_+. This model can also be viewed as a variant of the lattice Widom–Rowlinson model, which was first studied in [59] and which is a lattice analogue of the point process to be discussed in Section 8. In [13], it is noted that the methods used in [12] for the beach model can be adapted to prove phase transition in the iceberg model for large values of $N = \frac{N_2}{N_1}$. The uniqueness of maximal measures for sufficiently small N is also established in [13].

In addition to these examples of subshifts of finite type exhibiting phase transition, there is now available a general scheme for the construction of subshifts of finite type whose properties mimic those of a wide variety of Gibbs systems, such as the Ising, Potts, and Ashkin–Teller models; see [41].

The following two theorems are the closest analogues to Theorem 1.1 that current knowledge of the phase transition behaviour of beach and iceberg models permit.

Theorem 7.4: *For the d-dimensional beach model, $d \geq 2$, there exists an $M_c = M_c(d) \in (1, \infty)$ such that we have a phase transition whenever $M > M_c$, and a unique maximal measure whenever $M < M_c$.*

Theorem 7.5: *For the d-dimensional iceberg model, $d \geq 2$, there exist $N_c' = N_c'(d) > 0$ and $N_c'' = N_c''(d) < \infty$ such that we have a phase transition whenever $N > N_c''$ and a unique maximal measure whenever $N < N_c'$.*

Note that the beach model is better understood than the iceberg model, in that monotonicity in the parameter (M resp. N) is known for the beach model but not for the iceberg model. Nevertheless, it is intuitively believable that the occurrence of a phase transition should be monotone in the parameter also for the iceberg model, i.e. that one should be able to take $N_c' = N_c''$ in Theorem 7.5.

We shall now consider a random-cluster approach to phase transitions in beach and iceberg models. It turns out that this approach yields simple new proofs of Theorems

7.4 and 7.5, whereas like previous methods it fails to resolve the monotonicity-in-N question for the iceberg model. The two models look, at first sight, so similar that it is surprising that monotonicity question is resolvable for one of them but (so far) not for the other. For the random-cluster approach, we will soon see exactly what it is that leads to this difference.

The method of proof is essentially the same as that used for the Ising model in Section 3, so in particular we will only use random-cluster representations on finite graphs. To this end, we first need to define beach and iceberg models on finite graphs. As usual, G denotes a finite graph with vertex set V and edge set E. The maximal measure for the beach model on G with parameters M_1 and M_2 is simply the probability measure on F^V (where F is defined as in (26)) which is uniformly distributed over those elements of F^V in which no two vertices that have an edge in common take values with opposite sign unless they are both privileged. The maximal measure for the iceberg model on G is defined similarly.

The random-cluster representations of these models differ from those of previous sections by living on the vertices of a graph rather than on the edges. Let us begin with a random-cluster representation of the beach model.

Definition 7.6: *For $p \in [0,1]$, the BRC (beach random-cluster) measure $\mu^p_{G,BRC}$ for G is the probability measure on $\{0,1\}^V$ which to each $\eta \in \{0,1\}^V$ assigns probability*

$$\mu^p_{G,BRC}(\eta) = \frac{1}{Z^p_{G,BRC}} \left\{ \prod_{v \in V} p^{\eta(v)}(1-p)^{1-\eta(v)} \right\} 2^{k_{BRC}(\eta)}.$$

Here $k_{BRC}(\eta)$ is the number of connected components in the subgraph of G obtained by letting each edge $e \in E$ be present if and only if it is incident to some vertex v for which $\eta(v) = 1$.

Now fix parameters M_1 and M_2 for the beach model, and let $M = \frac{M_2}{M_1}$ and $p = \frac{M-1}{M}$. Consider the following way of picking a random element $X \in F^V$. First pick a random-cluster configuration $Y \in \{0,1\}^V$ according to the BRC measure $\mu^p_{G,BRC}$. For each connected component $\mathcal{C}$ of the subgraph of G used to define k_{BRC}, flip a fair coin to determine whether the vertices of $\mathcal{C}$ should take all positive or all negative values in X, and do this independently for each $\mathcal{C}$. It remains to determine the absolute values in X. For each $v \in V$ independently, we pick $|X(v)|$ uniformly from

$$\begin{cases} \{1,\ldots,M_1\} & \text{if } Y(v) = 0 \\ \{M_1+1,\ldots,M_2\} & \text{if } Y(v) = 1. \end{cases}$$

We now claim that X obtained in this way is distributed according to the maximal measure for the beach model on G with parameters M_1 and M_2; this follows from a simple counting argument analogous to the proof of Theorem 2.5.

Sketch proof of Theorem 7.4: The beach model on $\mathbf{Z}^d$ with parameters M_1 and M_2 has two particular maximal measures $\nu_+^{M_1,M_2}$ and $\nu_-^{M_1,M_2}$, obtained as weak limits of beach models on finite boxes with $+M_2$ resp. $-M_2$ boundary conditions, satisfying

$$\nu_-^{M_1,M_2} \preceq_d \nu^{M_1,M_2} \preceq_d \nu_+^{M_1,M_2}$$

for every maximal measure ν^{M_1,M_2}. This follows from the arguments used for the Ising model in Section 3, together with a straightforward modification of the proof of Holley's Theorem. By further copying of the arguments of Section 3, we get that phase transition for the beach model is equivalent to having infinite connected components of 1's in the "wired" limit of BRC models on finite boxes tending to $\mathbf{Z}^d$. Here, "connected components" should refer to the graph $\tilde{\mathbf{Z}}^d$ defined to have vertex set $\mathbf{Z}^d$ and edge set consisting of pairs of vertices within L_1-distance at most 2 from each other. To prove Theorem 7.4, we first note that $p \to 0$ as $M \to 1$, and that $p \to 1$ as $M \to \infty$. Next, we need nontriviality of the critical probability $\tilde{p}_c$ for independent site precolation (site percolation is defined similarly as bond percolation, except that it is the vertices rather than the edges that can be present or absent) on $\tilde{\mathbf{Z}}^d$. This is analogous to Theorem 3.8, and follows from standard percolation-theoretic arguments [35] [65]. The proof of Theorem 7.4 can then be completed with stochastic domination arguments: The conditional probability under $\mu^p_{G,BRC}$ that $Y(v) = 1$ given the configuration $Y(V \setminus \{v\}) = \eta$ equals

$$\mu^p_{G,BRC}(Y(v) = 1 \,|\, Y(V \setminus \{v\}) = \eta) = \frac{p2^{1-\kappa(\eta)}}{p2^{1-\kappa(\eta)} + 1 - p} \tag{27}$$

where $\kappa(\eta)$ is the number of connected components of the edge configuration corresponding to η (as in Definition 7.6, taking $\eta(v) = 0$) that contain v or some vertex incident to v. When G is a portion of $\mathbf{Z}^d$, we have that $\kappa(\eta)$ is between 1 and $2d+1$ for all η. Hence, the right hand side of (27) can be made arbitrarily close to 1, uniformly in η, by picking p sufficiently close to 1. Using Holley's Theorem to compare with independent site percolation, we get that the limiting BRC measure for $\mathbf{Z}^d$ yields infinite connected components for p sufficiently close to 1, so we have phase transition for the beach model with large enough M. Absence of phase transition for M close to 1 follows similarly. It remains to show that the occurrence of phase transition is increasing in M, and this follows as for the Ising model using the fact that

$$\mu^{p_1}_{G,BRC} \preceq_d \mu^{p_2}_{G,BRC} \tag{28}$$

whenever $p_1 \leq p_2$; this in turn follows from Holley's Theorem using the crucial observation that the right hand side of (27) is increasing in η. $\Box$

Next, we consider a random-cluster representation of the iceberg model.

Definition 7.7: *For $p \in [0,1]$, the IRC (iceberg random-cluster) measure $\mu^p_{G,IRC}$ for G is the probability measure on $\{0,1\}^V$ which to each $\eta \in \{0,1\}^V$ assigns probability*

$$\mu^p_{G,IRC}(\eta) = \frac{1}{Z^p_{G,IRC}} \Big\{ \prod_{v \in V} p^{\eta(v)}(1-p)^{1-\eta(v)} \Big\} 2^{k_{IRC}(\eta)}.$$

Here $k_{IRC}(\eta)$ is the number of connected components in the subgraph of G obtained by including only those vertices v for which $\eta(v) = 1$, and only those edges both of whose endvertices have been included.

A configuration $X \in H^V$ distributed according to the maximal measure for the iceberg model can be obtained in the following manner, analogous to how the beach model could be obtained from the BRC model: First pick $Y \in \{0,1\}^V$ according to the IRC measure $\mu^p_{G,IRC}$, with $p = \frac{N_2}{N_1+N_2} = \frac{N}{N+1}$. For each $v \in V$ for which $Y(v) = 0$, let $X(v)$ take a value chosen uniformly from $\{0_1, \ldots, 0_{N_1}\}$, independently. For each connected component $\mathcal{C}$ of $\{v \in V : Y(v) = 1\}$, flip an independent fair coin to determine whether the X-values of the vertices of $\mathcal{C}$ should be taken from $\{-N_2, \ldots, 1\}$ or from $\{1, \ldots, N_2\}$, and for each v in $\mathcal{C}$ independently pick the precise value of $X(v)$ uniformly from the chosen set. This defines X, and the usual counting argument shows that X is distributed according to the maximal measure for the iceberg model on G.

The proof of Theorem 7.5 is now more or less identical to that of Theorem 7.4, except that the IRC model replaces the BRC model. However, if we try to prove monotonicity in N of the occurrence of phase transition by the same approach, we run into the following obstacle. The conditional probability

$$\mu^p_{G,IRC}(Y(v) = 1 \mid Y(V \setminus \{v\}) = \eta) = \frac{p 2^{1-\kappa(\eta)}}{p 2^{1-\kappa(\eta)} + 1 - p} \qquad (29)$$

where $\kappa(\eta)$ this time is the number of connected components of the random subgraph of G used in Definition 7.7 that contain some vertex incident to v. Unlike the BRC counterpart (27), the right hand side of (29) is *not* increasing in η, and therefore the use of Holley's Theorem to prove the stochastic domination (28) cannot be translated to the iceberg setting. The property causing the right hand side of (27) but not the right hand side of (29) to be increasing is the following. The way we count connected components for the BRC model, all vertices of G belong to some connected component, so that increasing η can only cause old connected components to become connected to each other, thus reducing the number of connected components. On the other hand, only vertices v with $\eta(v) = 1$ are considered when we count connected components in the IRC model, and therefore an increment in η can increase as well as decrease the number of connected components.

Monotonicity in N of the occurrence of phase transition for the iceberg model on $\mathbf{Z}^d$ thus remains an open problem. Very recently, Brightwell and Winkler [9] gave a proof of this monotonicity for the iceberg model on the regular tree $\mathbf{T}_2$.

8 Continuum models

All the models considered so far live on a lattice. In many physical applications, it is more realistic to have a model which lives in continuous space. In this section, we consider the **Widom–Rowlinson model** [87] and some of its generalizations. The Widom–Rowlinson model is the most well-known example of a continuum model exhibiting a phase transition phenomenon. It is a spatial point process with particles of two types A and B. The particles are distributed in space according to two Poisson processes subject to the condition that no two points of different type are within unit distance from each other, so that in other words we have a hard core repulsion between points of type A and points of type B.

Let S be a compact subset of $\mathbf{R}^d$, $d \geq 2$, and let Ω_S be the set of all finite subsets of S. We should think of Ω_S as being the set of all possible realizations of a finite point process on S. An element of Ω_S is denoted $\mathbf{x} = \{x_1, \ldots, x_n\}$ if it consists of n points. Let π^λ_S be the probability measure on Ω_S which yields a Poisson process on S with intensity λ. In other words, let l denote d-dimensional Lebesgue measure on S, and define π^λ_S by letting

$$\pi^\lambda_S(F) = \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda l(S)}}{n!} \int \cdots \int \mathbf{1}_{[\{x_1,\ldots,x_n\} \in F]} dl(x_1) \cdots dl(x_n)$$

for all $F \in \mathcal{F}_S$, where $\mathcal{F}_S$ is the smallest σ-algebra which allows us to count the number of points in each Borel subset of S. A configuration of particles of types A and B on S is identified with an element $(\mathbf{x}, \mathbf{y})$ of $\Omega_S \times \Omega_S$. The Widom–Rowlinson measure ν^λ_S

on $\Omega_S \times \Omega_S$ is obtained by picking an element $(\mathbf{x}, \mathbf{y}) = (\{x_1, \ldots, x_n\}, \{y_1, \ldots, y_m\})$ according to $\pi_S^\lambda \times \pi_S^\lambda$ conditional on the event that

$$\inf_{i,j} |x_i - y_j| \geq 1$$

where $|\cdot|$ is Euclidean norm. A useful picture to have in mind is to have a closed sphere of radius $\frac{1}{2}$ around each point; the hard core repulsion then prevents balls centered at different types of particles from overlapping each other. By scaling S, we see that there is no loss in generality in having the hard core radius equal to $\frac{1}{2}$.

Besides physical motivation [87], the Widom–Rowlinson model is also of potential interest in spatial statistics, both in itself and via its close relation to the so called area-interaction process [5] which arises if we pick a particle configuration according to the Widom–Rowlinson model, and then disregard one of the particle types (see also [46]).

Taking S to be $[-R, R]^d$ for some large R, the phase transition behaviour of the Widom–Rowlinson model manifests itself as follows: If λ is small, then with high probability approximately half of the particles will be of type A and the others of type B, whereas for λ large there will with high probability be a large majority of either A- or B-particles. For fixed large or small λ, this holds uniformly for sufficiently large box sizes R.

To give a more precise definition of phase transition, we need to define the Widom–Rowlinson model on $\mathbf{R}^d$: Let $\Omega_{\mathbf{R}^d}$ be the set of locally finite subsets of $\mathbf{R}^d$. A measure μ on $\Omega_{\mathbf{R}^d} \times \Omega_{\mathbf{R}^d}$ is called a Gibbs measure for the Widom–Rowlinson model on $\mathbf{R}^d$ at intensity λ if for every compact $S \subset \mathbf{R}^d$, and μ-a.e. configuration $(\mathbf{x}, \mathbf{y})$ of points outside S the conditional distribution of what happens inside S given $(\mathbf{x}, \mathbf{y})$ is $\pi_S^\lambda \times \pi_S^\lambda$ conditioned on no two points of opposite type coming within unit distance from each other.

Theorem 8.1: *For the Widom–Rowlinson on $\mathbf{R}^d$, $d \geq 2$, there exist $\lambda' = \lambda'(d) > 0$ and $\lambda'' = \lambda''(d) < \infty$ such that there is a unique Gibbs measure when $\lambda < \lambda'$, while for $\lambda > \lambda''$ there is more than one Gibbs measure.*

The phase transition behaviour was first proved by Ruelle [79]. Recently, Chayes *et al.* [15] gave an alternative proof based on a random-cluster representation which we now go on to describe (see also [32] for some related ideas). Let $S_{x,r}$ denote the closed sphere of radius $r > 0$ centered at $x \in \mathbf{R}^d$.

Definition 8.2: *The* **continuum random-cluster measure** μ_S^λ *for $S \subset \mathbf{R}^d$ at intensity λ is the probability measure on Ω_S which has density $f(\mathbf{x})$ with respect to the unit intensity Poisson process π_S^1 given by*

$$f(\mathbf{x}) = \frac{1}{Z_S^\lambda} \lambda^n 2^{k(\mathbf{x})} \tag{30}$$

where $k(\mathbf{x})$ is the number of connected components of the set $\bigcup_{x \in \mathbf{x}} S_{x,\frac{1}{2}}$.

The continuum random-cluster model was first introduced by Klein [53]. Note that without the factor $2^{k(\mathbf{x})}$ in (30), we would simply get the Poisson process π_S^λ. The factor $2^{k(\mathbf{x})}$ plays the same role here as the factor $q^{k(\eta)}$ does in the definition of the Fortuin–Kasteleyn random-cluster model: they are both weightings of a "completely random" measure (the Poisson process resp. product measure on $\{0,1\}^E$) depending on the number of connected components in the random object which arises. One can of course generalize Definition 8.2 by replacing 2 by arbitrary $q > 0$ in (30). The relation between the continuum random-cluster model and the Widom–Rowlinson model is given in the following result.

Proposition 8.3:

(i) *If a random point configuration $X \in \Omega_S$ is obtained by first picking an element of $\Omega_S \times \Omega_S$ according to the Widom–Rowlinson measure ν_S^λ and then simply disregarding the type of each point, then X is distributed according to the continuum random-cluster measure μ_S^λ.*

(ii) *Conversely, an element of $\Omega_S \times \Omega_S$ with distribution ν_S^λ can be obtained by first picking $X \in \Omega_S$ according to μ_S^λ and then for each connected component $\mathcal{C}$ of $\bigcup_{x\in X} S_{x,\frac{1}{2}}$ independently tossing a fair coin to determine whether the points $x \in X \cap \mathcal{C}$ should be of type A or of type B.*

To see why this should be true, note that there are exactly $2^{k(X)}$ allowed elements of $\Omega_S \times \Omega_S$ which correspond to $X \in \Omega_S$. See [15] for a detailed proof.

The Widom–Rowlinson model on $\mathbf{R}^d$ has, analogously to the Ising model, two particular Gibbs measures ν_A^λ and ν_B^λ, where ν_A^λ is obtained as a weak limit of Widom–Rowlinson measures on finite boxes (tending to $\mathbf{R}^d$) with boundary condition consisting of a dense crowd of points of type A, and ν_B^λ is obtained similarly. Let us define the partial order $\preceq$ on $\Omega_S \times \Omega_S$ by

$$(\mathbf{x},\mathbf{y}) \preceq (\mathbf{x}',\mathbf{y}') \quad \text{if} \quad \mathbf{x} \subseteq \mathbf{x}' \text{ and } \mathbf{y} \supseteq \mathbf{y}'$$

so that in other words a configuration increases with respect to $\preceq$ if points of type A are added and points of type B are deleted. We then have

$$\nu_B^\lambda \preceq_d \nu^\lambda \preceq_d \nu_A^\lambda$$

for any other Gibbs measure ν^λ (this is analogous to the relations (9) and (10) for the Ising model), so uniqueness of Gibbs measures is equivalent to having

$$\nu_A^\lambda = \nu_B^\lambda .$$

This, in turn, is equivalent to the probability, in the continuum random-cluster model on $[-R,R]^d$ with appropriate boundary conditions, of having the origin in a connected component of $\bigcup_{x\in\mathbf{x}} S_{x,\frac{1}{2}}$ which intresects the boundary of $[-R,R]^d$, tending to 0 as $R \to \infty$. This issue can, like the corresponding issues for lattice models, be resolved (at least partially) using percolation and stochastic domination.

The percolation result needed is the following continuum analogue of Theorem 3.8; see [65] for a proof and for a general survey of continuum percolation.

Theorem 8.4: *Pick a random point configuration $X \in \Omega_{\mathbf{R}^d}$, $d \geq 2$, according to a Poisson process on $\mathbf{R}^d$ with intensity λ, and let $\bar{X} = \cup_{x\in X} S_{x,\frac{1}{2}}$. Furthermore let $\theta(\lambda)$ denote the probability that the origin belongs to an unbounded connected of $\bar{X}$. Then there exists a $\lambda_c = \lambda_c(d) \in (0,\infty)$ such that*

$$\theta(\lambda) \begin{cases} = 0 & \text{if } \lambda < \lambda_c \\ > 0 & \text{if } \lambda > \lambda_c . \end{cases}$$

Meester and Roy [65] call this particular model of continuum percolation the "Boolean model"; others have used the terms "Poisson blob model" and the more poetic "lily pad model".

Theorem 8.1 now follows if we can show for some $\lambda_1 > \lambda_c$ that the continuum random-cluster model with intensity λ stochastically dominates an intensity λ_1 Poisson process when λ is sufficiently large, and that it for some $\lambda_2 < \lambda_c$ is stochastically dominated by an intensity λ_2 Poisson process when λ is sufficiently small. To this end, we first need some point process analogue of Holley's Theorem, and for this we need to introduce the concept of **Papangelou conditional intensities** for point processes. Suppose that P is a probability measure on Ω_S which is absolutely continuous with respect to the unit intensity Poisson process π_S^1, and write $f(\mathbf{x})$ for the density of P with respect to π_S^1. For $x \in S$ and a point configuration $\mathbf{x} \in \Omega_S$ not containing x, the Papangelou conditional intensity under P of a point at x given $\mathbf{x}$ is, if it exists,

$$\lambda^*(x|\mathbf{x}) = \frac{f(\mathbf{x} \cup x)}{f(\mathbf{x})}. \tag{31}$$

Heuristically, $\lambda^*(x|\mathbf{x})du$ can be interpreted as the probability of finding a point inside a small region of volume du given that the point configuration outside this region is $\mathbf{x}$ (see [52] or [17]).

The following point process analogue of Holley's Theorem is due to Preston [77] who proved it using a coupling of so called spatial birth-and-death processes similar to the coupling used in the proof of Holley's Theorem. An alternative proof, based on a discretization argument, can be found in [31].

Theorem 8.5: *Suppose that P and $\tilde{P}$ are probability measures on Ω_S with Papangelou conditional intensities λ^* and $\tilde{\lambda}^*$ satisfying*

$$\lambda^*(x|\mathbf{x}) \leq \tilde{\lambda}^*(x|\tilde{\mathbf{x}})$$

for all $x \in S$ and all $\mathbf{x}, \tilde{\mathbf{x}} \in \Omega_S$ such that $\mathbf{x} \subseteq \tilde{\mathbf{x}}$. Then $P \preceq_d \tilde{P}$, in the sense that there exists a coupling $(X, \tilde{X})$ of P and $\tilde{P}$ such that $X \subseteq \tilde{X}$ a.s.

(Actually, Preston's proof does not quite work in the generality stated here, but the Georgii–Küneth proof [31] does.)

By inserting (30) into (31), we get that the Papangelou conditional intensity under the continuum random-cluster measure μ_S^λ of a point at x given $\mathbf{x}$ is

$$\lambda^*(x|\mathbf{x}) = \lambda 2^{1-\kappa(x,\mathbf{x})} \tag{32}$$

where $\kappa(x, \mathbf{x})$ is the number of connected components of $\bigcup_{y \in \mathbf{x}} S_{y,\frac{1}{2}}$ which intersect $S_{x,\frac{1}{2}}$. We then use the simple geometric fact that there exists a $\kappa_{max} = \kappa_{max}(d) < \infty$ such that $\kappa(x, \mathbf{x}) \leq \kappa_{max}$ for all x and $\mathbf{x}$ (for $d = 2$ we may for instance take $\kappa_{max} = 5$). It follows that

$$\lambda 2^{1-\kappa_{max}} \leq \lambda^*(x|\mathbf{x}) \leq 2\lambda \tag{33}$$

for all x and $\mathbf{x}$. Hence, we have by Theorem 8.5 that taking $\lambda < \frac{\lambda_c}{2}$ yields absence of unbouded connected components of $\bigcup_{y \in \mathbf{x}} S_{y,\frac{1}{2}}$ in the $S \to \mathbf{R}^d$ limit of continuum random-cluster measures, while taking $\lambda > \lambda_c 2^{\kappa_{max}-1}$ yields presence of unbounded connected components in the same limit. Theorem 8.1 follows.

Note that this approach fails in proving that the occurrence of phase transition is increasing in λ. The reason is (similarly to what we saw for the iceberg model in Section 7) that the right hand side in (32) fails to be increasing in $\mathbf{x}$.

There are several interesting directions in which the Widom–Rowlinson model can be generalized. An obvious direction is to increase the number of particle types to an arbitrary integer $q \geq 3$. This corresponds to replacing $2^{k(\mathbf{x})}$ by $q^{k(\mathbf{x})}$ in (30), and it is not hard to extend the above arguments to prove an analogue of Theorem 8.1 for this case.

Chayes and Kotecký [16] consider a related model with four types of particles $\{A, B, C, D\}$ and a symmetry of Ashkin–Teller type: For some $0 < r < R < \infty$, there is a hard core repulsion of radius r between A and B and between C and D, and a hard core repulsion of radius R between $\{A, B\}$ and $\{C, D\}$. By an extension of the above random-cluster approach, they establish for sufficiently large $\frac{R}{r}$ a phase transition behaviour similar to that in the Ashkin–Teller model (Theorem 6.2).

Yet another direction of generalization is to relax the hard core repulsion into a "soft core" repulsion of finite strength. For a repulsion function $g : [0, R] \to \mathbf{R}_+$, consider the measure $\mu_S^{g,\lambda}$ on $\Omega_S \times \Omega_S$ whose density $f(\mathbf{x}, \mathbf{y})$ with respect to $\pi_S^\lambda \times \pi_S^\lambda$ is

$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{Z_S^{g,\lambda}} \exp\left(- \sum_{\substack{x \in \mathbf{x}, y \in \mathbf{y} \\ |x-y| \leq R}} g(|x-y|) \right).$$

This means that configurations with plently of particles of different type sitting close to each other tend to by less likely (but not impossible, as in the Widom–Rowlinson model) than those where particles of different type keep far apart. Phase transition for large λ under a certain condition on g was first proved by Lebowitz and Lieb [60]. Georgii and Häggström [30] proved phase transition in a larger class of systems using a random-cluster arguments. This leads to a kind of generalized continuum random-cluster model (also considered in [53]), consisiting of a random point configuration $X \in \Omega_S$ together with edges that are included in a certain random fashion between points $x_1, x_2 \in X$ that are within distance R from each other. This model relates to the so called random connection model of continuum percolation [65] in the same way as the continuum random-cluster model of Definition 8.2 relates to the percolation model in Theorem 8.4. The proof of phase transition in [30] contains the usual ingredients of percolation and stochastic domination: For the percolation part it suffices to refer to an analogue of Theorem 8.4 for the random connection model (see [65]), whereas the stochastic domination arguments are, due to the extra randomness of the edges, somewhat more involved than for the Widom–Rowlinson model (in particular, more work is needed to prove an analogue of (33) since we no longer have any deterministic bound on how much the number of connected components can decrease when a point is added to the point configuration).

9 Concluding remarks

We have seen that random-cluster representations are useful for studying the phase transition behaviour in a wide range of models, none of which was originally proposed with a random-cluster representation in mind. This is, in my opinion, a clear case in favour of the importance of random-cluster ideas. I also hope that if anyone has the feeling that the relationship between Ising and Potts models on one hand, and the random-cluster model on the other, has a somewhat "coincidental" or "miraculous" aspect, then this feeling is lessened after having seen these further examples of random-cluster representations.

It has to be admitted, however, that all systems studied in this paper feature a certain symmetry (such as the ± 1 symmetry of the Ising model) and it seems to be difficult to apply the random-cluster approach to the study of phase transitions in systems where no such symmetry is present. For such systems, so called Pirogov–Sinai theory has proved to be more successful. In fact, Pirogov–Sinai theory has been used to study various properties of the infinite-volume random-cluster model; see e.g. [56], [54] and [23]. These developments, however, go beyond the scope of the present paper.

We saw for some models (the iceberg model in Section 7 and the continuum models in Section 8) that the random-cluster approach failed to establish the existence of a critical parameter value above which there is phase transition and below which there is not. For each of these models, it nevertheless seems intuitively plausible that such a critical value should exist, and it is in my opinion an important class of open problems to prove (or disprove!) this. The analogous problem for the so called hard core lattice gas model in two or more dimensions [7] is also open. The discussion at the end of Section 7 shows why some new idea is needed to take care of this class of problems.

It certainly seems worthwhile to study the various random-cluster models that we have seen in this paper in their own right, and not only as representations of other systems. For the original Fortuin–Kasteleyn random-cluster model we have seen examples of work in this direction in Section 4. Perhaps some of the models can be given an interesting unified treatment. The setups in [38], [66] and [74] are all strict generalizations of one or more of the random-cluster models studied here, and they all deserve attention.

Acknowledgement. This work was supported by a grant from the Swedish Natural Science Research Council. I am indebted to Aernout van Enter, Geoffrey Grimmett, Peter Jagers, Johan Jonasson, Jens Lund, Jesper Møller, Karin Nelander, Jim Propp, Timo Seppäläinen, Jeff Steif and Peter Winkler for comments, corrections and constructive criticism on earlier verisons of this paper.

References

[1] Abraham, D.B. and Martin-Löf, A. (1973) The transfer matrix for a pure phase in the two-dimensional Ising model, *Commun. Math. Phys.* **32**, 245–268.

[2] Aizenman, M., Chayes, J.T., Chayes, L. and Newman, C.M. (1987) The phase boundary in dilute and random Ising and Potts ferromagnets, *J. Phys. A* **20**, L313–L318.

[3] 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, *J. Statist. Phys.* **50**, 1–40.

[4] Ashkin, J. and Teller, E. (1943) Statistics of two-dimensional lattices with four components, *Phys. Rev.* **64**, 178–184.

[5] Baddeley, A.J. and van Lieshout, M.N.M. (1995) Area-interaction point processes, *Ann. Inst. Statist. Math.* **46**, 601–619.

[6] van den Berg, J. and Maes, C. (1994) Disagreement percolation in the study of Markov fields, *Ann. Probab.* **22**, 749–763.

[7] van den Berg, J. and Steif, J. (1994) Percolation and the hard core lattice gas model, *Stoch. Proc. Appl.* **49**, 179–197.

[8] Borgs, C. and Chayes, J.T. (1996) The covariance matrix of the Potts model: a random cluster analysis, *J. Statist. Phys.* **82**, 1235–1297.

[9] Brightwell, G. and Winkler, P. (1997) Graph homomorphisms and phase transitions, preprint.

[10] Broadbent, S.R. and Hammersley, J.M. (1957) Percolation processes I. Crystals and mazes, *Proc. Cambridge Phil. Soc.* **53**, 629–641.

[11] Burton, R. and Keane, M. (1989) Density and uniqueness in percolation, *Commun. Math. Phys.* **121**, 501–505.

[12] Burton, R. and Steif, J. (1994) Nonuniqueness of measures of maximal entropy for subshifts of finite type, *Ergod. Th. Dynam. Sys.* **14**, 213–236.

[13] Burton, R. and Steif, J. (1995) New results on measures of maximal entropy, *Israel J. Math.* **89**, 275–300.

[14] Cammarota, C. (1993) Stochastic order and monotonicity in temperature for Gibbs measures, *Lett. Math. Phys.* **29**, 287–295.

[15] Chayes, J.T., Chayes, L. and Kotecký, R. (1995) The analysis of the Widom-Rowlinson model by stochastic geometric methods, *Commun. Math. Phys.* **172**, 551–569.

[16] Chayes, L. and Kotecký, R. (1996) Intermediate phase for a classical continuum model, *Phys. Rev. B* **54**, 9221–9224.

[17] Daley, D.J. and Vere-Jones, D. (1988) *An Introduction to the Theory of Point Processes*, Springer, New York.

[18] Dobrushin, R.L. (1965) Existence of a phase transition in two and three dimensional Ising models, *Th. Probab. Appl.* **10**, 193–213.

[19] Dobrushin, R.L. (1968) Gibbsian random fields for lattice systems with pairwise interactions, *Funct. Anal. Appl.* **2**, 292–301.

[20] Dobrushin, R.L. (1968) The problem of uniqueness of a Gibbsian random field and the problem of phase transition, *Funct. Anal. Appl.* **2**, 302–312.

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

[22] Ellis, R.S. (1985) *Entropy, Large Deviations and Statistical Mechanics*, Springer, New York.

[23] van Enter, A.C.D., Fernández, R., Schonmann, R.A. and Shlosman, S.B. (1997) Complete analyticity of the 2D Potts model above the critical temperature, *Commun. Math. Phys.*, to appear.

[24] Fortuin, C.M. (1972) On the random-cluster model. II. The percolation model, *Physica* **58**, 393–418.

[25] Fortuin, C.M. (1972) On the random-cluster model. III. The simple random-cluster process, *Physica* **59**, 545–570.

[26] Fortuin, C.M. and Kasteleyn, P.W. (1972) On the random-cluster model. I. Introduction and relation to other models, *Physica* **57**, 536–564.

[27] Fortuin, C.M., Kasteleyn, P.W. and Ginibre, J. (1971) Correlation inequalities on some partially ordered sets, *Commun. Math. Phys* **22**, 89–103.

[28] Geman, D. (1990) Random fields and inverse problems in imaging, in *École d'été de probabilités de Saint-Flour XVIII*, Lecture Notes in Math. **1427**, Springer, New York.

[29] Georgii, H.-O. (1988) *Gibbs Measures and Phase Transitions*, de Gruyter, New York.

[30] Georgii, H.-O. and Häggström, O. (1996) Phase transition in continuum Potts models, *Commun. Math. Phys.* **181**, 507–528.

[31] Georgii, H.-O. and Küneth, T. (1997) Stochastic comparison of point random fields, *J. Appl. Probab.*, to appear.

[32] Giacomin, G., Lebowitz, J.L. and Maes, C. (1995) Agreement percolation and phase coexistence in some Gibbs systems, *J. Statist. Phys.* **80**, 1379–1403.

[33] Griffiths, R.B. (1964) Peierls' proof of spontaneous magnetization in a two-dimensional Ising ferromagnet. *Phys. Rev.* **136A**, 437–439.

[34] Griffiths, R.B. (1967) Correlations in Ising ferromagnets, *J. Math. Phys.* **8**, 478–489.

[35] Grimmett, G. (1989) *Percolation*, Springer, New York.

[36] Grimmett, G. (1994) Percolative problems, *Probability and Phase Transition* (ed. G. Grimmett), Kluwer, Dordrecht, 69–86.

[37] Grimmett, G. (1994) The random-cluster model, *Probability, Statistics and Optimization* (ed. F.P. Kelly), Wiley, Chichester, 49–63.

[38] Grimmett, G. (1994) Potts models and random-cluster processes with many-body interactions, *J. Statist. Phys* **75**, 67–121.

[39] Grimmett, G. (1995) The stochastic random-cluster process, and the uniqueness of random-cluster measures, *Ann. Probab.* **23**, 1461–1510.

[40] Guttorp, P. (1995) *Stochastic Modeling of Scientific Data*, Chapman and Hall, London.

[41] Häggström, O. (1995) On the relation between finite range potentials and subshifts of finite type, *Probab. Th. Rel. Fields* **101**, 469–478.

[42] Häggström, O. (1996) On phase transitions for subshifts of finite type, *Israel J. Math.* **94**, 319–352.

[43] Häggström, O. (1996) The random-cluster model on a homogeneous tree, *Probab. Th. Rel. Fields* **104**, 231–253.

[44] Häggström, O. (1996) Almost sure quasilocality fails for the random-cluster model on a tree, *J. Statist. Phys.* **84**, 1351–1361.

[45] Häggström, O. (1996) A note on (non-)monotonicity in temperature for the Ising model, *Markov Proc. Rel. Fields* **2**, 529–537.

[46] Häggström, O., van Lieshout, M.N.M. and Møller, J. (1996) Characterisation results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes, preprint.

[47] Hammersley, J.M. (1957) Percolation processes. Lower bounds for the critical probability, *Ann. Math. Statist.* **28**, 790–795.

[48] Hammersley, J.M. (1957) Percolation processes II. The connective constant, *Proc. Cambridge Phil. Soc.* **53**, 642–645.

[49] Hjort, N.L. and Omre, H. (1994) Topics in spatial statistics (with discussion), *Scand. J. Statist.* **21**, 289–357.

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

[51] Ising, E. (1925) Beitrag zur Theorie der Ferromagnetismus, *Z. Phys.* **31**, 253–258.

[52] Kallenberg, O. (1984) An informal guide to the theory of conditioning in point processes, *Int. Statist. Rev.* **52**, 21–43.

[53] Klein, W. (1982) Potts-model formulation of continuum percolation, *Phys. Rev. B* **26**, 2677-2678.

[54] Kotecký, R. (1994) Geometric representation of lattice models and large volume asymptotics, *Probability and Phase Transition* (ed. G. Grimmett), Kluwer, Dordrecht, 153–176.

[55] Kotecký, R. and Shlosman, S. (1982) First-order phase transitions in large entropy lattice systems, *Commun. Math. Phys.* **83**, 493–550.

[56] Laanait, L., Messager, A., Miracle-Sole, S., Ruiz, J. and Shlosman, S. (1991) Interfaces in the Potts model I: Pirogov–Sinai theory of the Fortuin–Kasteleyn representation, *Commun. Math. Phys.* **140**, 81–91.

[57] Laanait, L., Messager, A. and Ruiz, J. (1986) Phase coexistence and surface tensions for the Potts model, *Commun. Math. Phys.* **105**, 527–545.

[58] Lanford, O. and Ruelle, D. (1969) Observables at infinity and states with short range correlations in statistical mechanics, *Commun. Math. Phys.* **13**, 194–215.

[59] Lebowitz, J.L. and Gallavotti, G. (1971) Phase transitions in binary lattice gases, *J. Math. Phys.* **12**, 1129–1133.

[60] Lebowitz, J.L. and Lieb, E.H. (1972) Phase transition in a continuum classical system with finite interactions, *Phys. Lett.* **39A**, 98–100.

[61] Lenz, W. (1920) Beitrag zum verständnis der magnetischen Erscheinungen in festen Körpern, *Phys. Zeitschr.* **21**, 613–615.

[62] Liggett, T.M. (1985) *Interacting Particle Systems*, Springer, New York.

[63] Lindvall, T. (1992) *Lectures on the Coupling Method*, Wiley, New York.

[64] Machta, J., Choi, Y.S., Lucke, A., Schweizer, T. and Chayes, L. (1995) Invaded cluster algorithm for equilibrium critical points, *Phys. Rev. Lett.* **75**, 2792–2795.

[65] Meester, R. and Roy, R. (1996) *Continuum Percolation*, Cambridge University Press.

[66] Møller, J. and Waagepetersen, R. (1997) Markov connected component fields, *Adv. Appl. Probab.*, to appear.

[67] Newman, C.M. (1994) Disordered Ising systems and random cluster representations, *Probability and Phase Transition* (ed. G. Grimmett), Kluwer, Dordrecht, 247–260.

[68] Onsager, L. (1944) Crystal statistics, I. A two-dimensional model with an order-disorder transformation, *Phys. Rev.* **65**, 117–149.

[69] Parry, W. (1964) Intrinsic Markov Chains, *Trans. Amer. Math. Soc.* **112**, 55–66.

[70] Peierls, R. (1936) On Ising's model of ferromagnetism, *Proc. Cambridge Phil. Soc.* **32**, 477–481.

[71] Petritis, D. (1996) Equilibrium statistical mechanics of frustrated spin glasses: a survey of mathematical results, *Ann. Inst. H. Poincaré Phys. Theorique* **64**, 255–288.

[72] Pfister, C.-E. (1982) Phase transitions in the Ashkin–Teller model, *J. Statist. Phys.* **29**, 113–116.

[73] Pfister, C.-E. and Vande Velde, K. (1995) Almost sure quasilocality in the random cluster model, *J. Statist. Phys.* **79**, 765–774.

[74] Pfister, C.-E. and Velenik, Y. (1996) Random-cluster representation of the Ashkin–Teller model, preprint.

[75] Pisztora, A. (1996) Surface order large deviations for Ising, Potts, and percolation models, *Probab. Th. Rel. Fields* **104**, 427–466.

[76] Potts, R.B. (1952) Some generalized order-disorder transformations, *Proc. Cambridge Phil. Soc.* **48**, 106–109.

[77] Preston, C.J. (1977) Spatial birth-and-death processes, *Bull. Int. Statist. Inst.* **46**(2), 371–391.

[78] Propp, J.G. and Wilson, D.B. (1996) Exact sampling with coupled Markov chains and applications to statistical mechanics, *Random Structures Algorithms* **9**, 223–252.

[79] Ruelle, D. (1971) Existence of a phase transition in a continuous classical system, *Phys. Rev. Lett.* **27**, 1040–1041.

[80] Salas, J. and Sokal, A.D. (1996) Dynamic critical behavior of a Swendsen–Wang-type algorithm for the Ashkin–Teller model, *J. Statist. Phys.* **85**, 297–361.

[81] Schmidt, K. (1990) *Algebraic Ideas in Ergodic Theory*, American Mathematical Society, Providence, Rhode Island.

[82] Strassen, V. (1965) The existence of probability measures with given marginals, *Ann. Math. Statist.* **36**, 423–439.

[83] Swendsen, R.H. and Wang, J.-S. (1987) Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58**, 86–88.

[84] Thorisson, H. (1995) Coupling methods in probability theory, *Scand. J. Statist.* **22**, 159–182.

[85] Welsh, D.J.A. (1993) Percolation in the random cluster process and the Q-state Potts model, *J. Phys. A*, **26**, 2471–2483.

[86] Welsh, D.J.A. (1994) The random cluster process, *Discr. Math.* **136**, 373–390.

[87] Widom, B. and Rowlinson, J.S. (1970) New model for the study of liquid-vapor phase transition, *J. Chem. Phys.* **52**, 1670–1684.

[88] Wiseman, S. and Domany, E. (1993) Cluster method for the Ashkin–Teller model, *Phys. Rev. E* **48**, 4080–4090.