Markov random fields and percolation on general graphs

Olle Häggström*†

February 17, 1999

Abstract

Let G be an infinite, locally finite, connected graph with bounded degree. We show that G supports phase transition in all or none of the following five models: bond percolation, site percolation, the Ising model, the Widom–Rowlinson model and the beach model. Some, but not all, of these implications hold without the bounded degree assumption. We finally give two examples of (random) unbounded degree graphs in which phase transition in all five models can be established: supercritical Galton–Watson trees, and Poisson–Voronoi tessellations of $\mathbf{R}^d$ for $d \geq 2$.

Keywords: Percolation, Ising model, Widom–Rowlinson model, beach model, Galton–Watson tree, Poisson–Voronoi tessellation.

AMS Subject Classification: Primary 60K35, Secondary 82B20, 82B43

1 Introduction

Over the last few decades, it has become increasingly clear that there are important connections between percolation theory on one hand, and the issue of Gibbs state multiplicity in Markov random fields on the other. Examples of such connections are the Fortuin–Kasteleyn representation of Ising and Potts models [16, 1, 20, 23], the disagreement percolation technique for establishing Gibbsian uniqueness [5, 6], and the equivalence between spin percolation and Gibbs state multiplicity for Ising and Potts models on the square lattice [15, 13]; see also [18] for a general introduction to such ideas. Here we shall focus on the two basic percolation models (**bond percolation** and **site percolation**) and on three different Markov random field models (the **Ising model**, the **Widom–Rowlinson model**, and the **beach model**). Suppose for the moment that the graph structure on which the models live is taken to be the integer lattice $\mathbf{Z}^d$, with edges connecting (Euclidean) nearest neighbors. It is well known that all five models exhibit interesting phase transition phenomena for $d \geq 2$, whereas none of them do for $d = 1$. The purpose of this paper is to investigate to what extent such a dichotomy can be extended to the setting of general graphs.

Let $\mathcal{G}$ denote the class of all infinite, locally finite, connected graphs, and let $\mathcal{G}^b$ be the class of all such graphs with bounded degree. Let $\mathcal{G}_{BP}$ (resp. $\mathcal{G}_{SP}$) be the class of graphs in $\mathcal{G}$ whose critical value p_c^{bond} (resp. p_c^{site}) for bond (resp. site) percolation is less than 1; careful definitions will be given in the next section. Furthermore write $\mathcal{G}_I$ for the class of graphs $G \in \mathcal{G}$ which exhibit phase transition in the Ising model, in the sense

*Dept of Mathematics, Chalmers University of Technology, 412 96 Göteborg, Sweden

†Research supported by the Swedish Natural Science Research Council

that there exists some parameter values for which the Ising model on G has more than one Gibbs measure. Similarly, we write $\mathcal{G}_{WR}$ (resp. $\mathcal{G}_{BM}$) for the class of graphs in $\mathcal{G}$ which exhibit phase transition for the Widom–Rowlinson (resp. beach) model. Finally, set $\mathcal{G}^b_{BP} = \mathcal{G}^b \cap \mathcal{G}_{BP}$, and define $\mathcal{G}^b_{SP}$, $\mathcal{G}^b_I$, $\mathcal{G}^b_{WR}$ and $\mathcal{G}^b_{BM}$ analogously.

Our first main result says that if G is a bounded degree graph, then either

(i) G has $p_c^{bond} < 1$, $p_c^{site} < 1$, and exhibits phase transition for all three Markov random fields models, or

(ii) G has $p_c^{bond} = p_c^{site} = 1$, and does not exhibit phase transition for any of the three Markov random field models.

This admits a slick formulation as follows.

Theorem 1.1 $\quad \mathcal{G}^b_{BP} = \mathcal{G}^b_{SP} = \mathcal{G}^b_I = \mathcal{G}^b_{WR} = \mathcal{G}^b_{BM}$

Without the bounded degree assumption, the situation is less clearcut, as indicated in the next theorem.

Theorem 1.2

(a) $\quad \mathcal{G}_{WR} \subset \mathcal{G}_{SP} \subset \mathcal{G}_{BP} = \mathcal{G}_I$

(b) $\quad \mathcal{G}_{BM} \not\subseteq \mathcal{G}_{SP}$

(c) $\quad \mathcal{G}_{SP} \not\subseteq \mathcal{G}_{BM}$

Perhaps this result is best visualized in the diagram of implications (and nonimplications) given in Figure 1.

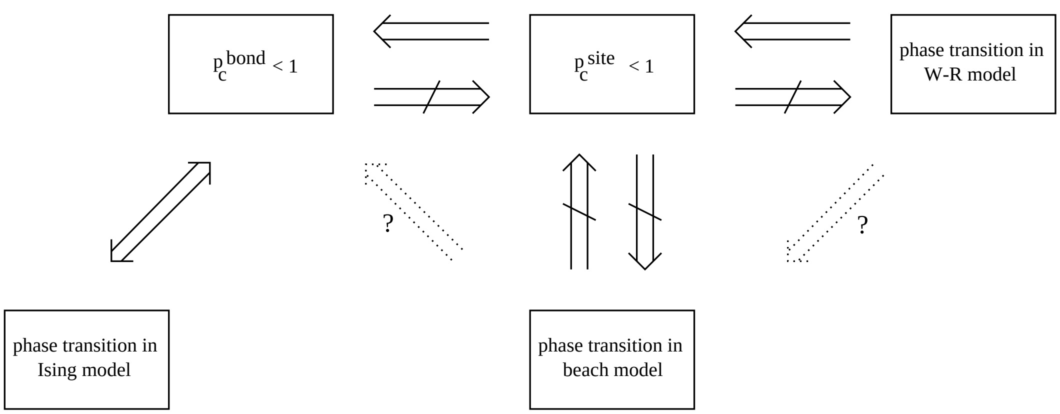

Figure 1: Implications and nonimplications for graphs in the class $\mathcal{G}$. Dotted arrows are the remaining implications which we have neither been able to prove nor disprove.

Some of the implications in the above results are known from previous work. The inclusion $\mathcal{G}_{SP} \subseteq \mathcal{G}_{BP}$ is immediate from the relation $p_c^{bond} \le p_c^{site}$, which is due to Hammersley [28]. The equivalence $\mathcal{G}_{BP} = \mathcal{G}_I$ is clear from the work of Aizenman et al. [1], although they stated their results only in a $\mathbf{Z}^d$ setting. An example of a graph which is in $\mathcal{G}_{SP}$ but not in $\mathcal{G}_{WR}$ can be found in Brightwell, Häggström and Winkler [8].

It would of course be desirable to obtain a more explicit structural chararcterization, e.g. of graphs with $p_c^{bond} < 1$. However, a general result of this kind appears to be fairly remote. For trees, $p_c^{bond} = p_c^{site}$, and Lyons [36] characterized the critical value in terms of a rather explicit quantity known as the branching number of the tree; in particular, $p_c^{site} < 1$ if and only if the branching number is strictly greater than 1. Benjamini and Schramm [3] conjectured that a Cayley graph of an infinite finitely generated group has $p_c^{site} < 1$ unless it is a finite extension of $\mathbf{Z}$. Theorem 1.1 stresses the importance of this conjecture.

The rest of this paper is organized as follows. In Section 2 we give careful definitions of all models under consideration. In Section 3 we recall the random-cluster representations of the Markov random field models; these will play a key role in the following sections. The proof of Theorem 1.2 is split into two sections: the positive implications are proved in Section 4, whereas the nonimplications are demonstrated in Section 5. Theorem 1.1 is proved in Section 6. In Section 7 we provide a sufficient condition on unbounded degree graphs for the conclusions of Theorem 1.1, and give two examples of (random) graphs which satisfy this condition: supercritical Galton–Watson trees, and Poisson–Voronoi tessellations of $\mathbf{R}^d$ for $d \geq 2$. Some final remarks, concerning possible extensions to other models, are given in Section 8.

2 The models

2.1 Bond percolation

In standard bond percolation on a graph $G = (V, E) \in \mathcal{G}$ with parameter $p \in [0,1]$, each edge $e \in E$ is independently assigned value 1 (open) with probability p, or value 0 (closed) with probability $1-p$. We write μ_{BP}^p for the corresponding product probability measure on $\{0,1\}^E$ (note that we are supressing the dependence on G in the notation). A **cluster** is a (maximal) connected component of open edges. The primary focus of percolation theory is on the possible occurence of infinite clusters. The existence of at least one infinite cluster is not influenced by changing the status of any finite set of edges, so by Kolmogorov's 0-1 law the μ_{BP}^p-probability of having some infinite cluster must be 0 or 1. An obvious coupling argument shows that the probability of having some infinite cluster cannot decrease as p increases. Combining these two observations, we have the existence of a critical value $p_c^{bond} = p_c^{bond}(G) \in [0,1]$ such that

$$\mu_{BP}^p(\exists \text{ some infinite cluster}) = \begin{cases} 0 & \text{if } p < p_c^{bond} \\ 1 & \text{if } p > p_c^{bond} \end{cases}.$$

(At $p = p_c^{bond}$, the probability of existence of some infinite cluster may be either 0 or 1, depending on the choice of G.) We define

$$\mathcal{G}_{BP} = \{G \in \mathcal{G} : p_c^{bond}(G) < 1\}.$$

By far the most studied choice of G is the integer lattice $\mathbf{Z}^d$ in $d \geq 2$ dimensions; see [19] for an introduction to percolation theory with emphasis on the $\mathbf{Z}^d$ case. Recently, there has been an upsurge of interest in percolation beyond this setting; see e.g. [3, 2, 26], where the focus is mainly on Cayley graphs and other quasi-transitive graphs, which still have some structure that can be exploited in various ways. In this paper we basically drop all such structure. The obvious cost of doing so is that we are able to say much less about the percolation behavior.

2.2 Site percolation

Site percolation on $G = (V, E) \in \mathcal{G}$ with parameter $p \in [0, 1]$ is similar to bond percolation, except that the randomness is in the vertices rather than the edges: each vertex $v \in V$ is independently assigned value 1 (open) with probability p, or 0 (closed) with probability $1 - p$. Write μ_{SP}^p for the resulting probability measure on $\{0, 1\}^V$. Clusters are defined similarly as for bond percolation, and in the same way as in the bond case we get the existence of a critical value $p_c^{site} = p_c^{site}(G)$ such that

$$\mu_{SP}^p(\exists \text{ some infinite cluster}) = \begin{cases} 0 & \text{if } p < p_c^{site} \\ 1 & \text{if } p > p_c^{site} \, . \end{cases}$$

We set

$$\mathcal{G}_{SP} = \{G \in \mathcal{G} : \, p_c^{site}(G) < 1\} \, .$$

2.3 Markov random fields

Let S be a finite set, and let $G = (V, E)$ be some finite or infinite graph. Let X be some S^V-valued random object, and let π be the corresponding probability measure on S^V. For a vertex set $\Lambda \subset V$, we define its boundary $\partial \Lambda$ as

$$\partial \Lambda = \{x \in V \setminus \Lambda : \, \exists y \in \Lambda \text{ such that } x \sim y\} \, ,$$

where $x \sim y$ denotes the existence of an edge $e \in E$ connecting x and y. The random object X (or the measure π) is said to be a **Markov random field** if π admits conditional probabilities such that for all finite $\Lambda \subset V$, all $\xi \in S^\Lambda$, and all $\eta \in S^{V \setminus \Lambda}$ we have

$$\mu(X(\Lambda) = \xi \,|\, X(V \setminus \Lambda) = \eta) = \mu(X(\Lambda) = \xi \,|\, X(\partial \Lambda) = \eta(\partial \Lambda)) \, .$$

In other words, the Markov random field property says that the conditional distribution of what we see on Λ given everything else only depends on what we see on the boundary $\partial \Lambda$.

Now take $G \in \mathcal{G}$. A consistent set of conditional distributions for all finite Λ and all boundary conditions η as above is called a **specification**, denoted $\mathcal{Q}$. The specification is said to be Markovian if

$$\mathcal{Q}(X(\Lambda) = \xi \,|\, X(V \setminus \Lambda) = \eta) = \mathcal{Q}(X(\Lambda) = \xi \,|\, X(V \setminus \Lambda) = \eta')$$

for all Λ, all $\xi \in S^\Lambda$ and all $\eta, \eta' \in S^{V \setminus \Lambda}$ such that $\eta(\partial \Lambda) = \eta'(\partial \Lambda)$. A probability measure μ on S^V satisfying the prescribed conditional distributions for such a specification $\mathcal{Q}$ is called a Gibbs measure for $\mathcal{Q}$. Such measures are automatically Markov random fields, and the existence of Gibbs measures for a given such specification follows from a standard compactness argument. In contrast, uniqueness does not alway hold. This possible nonuniqueness of Gibbs measure is of central interest in statistical mechanics, and is also of primary interest in this paper. All three Markov random field models to be discussed here (Ising, Widom–Rowlinson and beach) exhibit nonuniqueness of Gibbs measures for certain G and certain parameter values. For such a parameterized Markov random field model and a given graph G, we reserve the term **phase transition** to denote the existence of some parameter values for which the model, living on G, has more than one Gibbs measure. (For bond or site percolation, phase transition means simply that the critical value is strictly less than one.) The three Markov random field models considered here all possess a ± 1 symmetry, and share the feature that phase transition is characterized by a breaking of this symmetry; see Propositions 2.1, 2.3 and 2.4 below.

2.4 Ising model

The Ising model on a graph G is a certain random assignment of $+1$'s and -1's to the vertices of G. It was introduced in the 1920's as a model for ferromagnetism, and is today the most studied of all Markov random field models; see e.g. [31, 17] for introductions and some history. Take $G = (V, E) \in \mathcal{G}$. A probability measure π on $\{-1, 1\}^V$ is said to be a Gibbs measure for the (ferromagnetic) Ising model on G at inverse temperature $\beta > 0$ if it is Markov and for all finite $\Lambda \subset V$ and all $\eta \in \{-1, 1\}^{\partial\Lambda}$, $\xi \in \{-1, 1\}^{\Lambda}$ we have

$$\pi(X(\Lambda) = \xi \mid X(\partial\Lambda) = \eta) = Z^{-1} \exp\left[\beta\left(\sum_{\substack{\langle x,y\rangle \\ x,y\in\Lambda}} \xi(x)\xi(y) + \sum_{\substack{\langle x,y\rangle \\ x\in\Lambda, y\in\partial\Lambda}} \xi(x)\eta(y)\right)\right]. \quad (1)$$

Here Z is a normalizing constant which depends on β, Λ and η but not on ξ, and $\langle x, y\rangle$ means that we sum only over x and y that have an edge in common, and that each such nearest-neighbor pair is counted only once.

The study of phase transition in the Ising model is greatly facilitated by the existence of two particular probability measures $\pi_I^{\beta,+}$ and $\pi_I^{\beta,-}$ that are extreme in the sense of stochastic ordering (the precise meaning of this is given in Section 3.4 below). The "plus measure" $\pi_I^{\beta,+}$ is constructed as follows; the "minus measure" is obtained analogously.

Let $\{\Lambda_n\}_{n=1}^{\infty}$ be a increasing sequence of finite subsets of V converging to V in the sense that each $v \in V$ is in all but finitely many of the Λ_n's. We refer to such a sequence as an exhaustion of G. Fix a vertex $o \in \Lambda_1$ called the origin. For each n, let $\pi_I^{\beta,n,+}$ be the probability measure on $\{-1, 1\}^V$ corresponding to picking $X \in \{-1, 1\}^V$ by setting $X(V \setminus \Lambda_n) \equiv +1$ and picking $X(\Lambda_n)$ according to (1) with $\Lambda = \Lambda_n$ and $\eta \equiv +1$. Standard monotonicity arguments based on Holley's Theorem (Theorem 3.7 below) show that the measures $\pi_I^{\beta,n,+}$ converge to a Gibbs measure $\pi_I^{\beta,+}$ which is independent of the choice of exhaustion, and that the following result holds; see e.g. [23] or [18] for details.

Proposition 2.1 *For the Ising model on $G \in \mathcal{G}$ at inverse temperature β, the following statements are equivalent.*

(i) *There is more than one Gibbs measure.*

(ii) $\pi_I^{\beta,+} \neq \pi_I^{\beta,-}$

(iii) $\pi_I^{\beta,+}(X(o) = +1) > \frac{1}{2}$

(iv) $\exists \varepsilon > 0$ *such that* $\pi_I^{\beta,n,+}(X(o) = +1) \geq \frac{1}{2} + \varepsilon$ *for all* n.

It is also well known that the existence of more than one Gibbs measure is increasing in β. This was originally proved using so-called Griffiths inequalities (see e.g. [33]); the modern approach based on the random-cluster model will be indicated in Section 3.4. The following result is an immediate consequence.

Theorem 2.2 *For any $G \in \mathcal{G}$, there exists a critical value $\beta_c = \beta_c(G) \in [0, \infty]$ such that for $\beta < \beta_c$ we have that the Ising model on G at inverse temperature β has a unique Gibbs measure whereas for $\beta > \beta_c$ there are multiple Gibbs measures.*

The graph G supports phase transition for the Ising model if and only if $\beta_c < \infty$, and we define

$$\mathcal{G}_I = \{G \in \mathcal{G} : \beta_c(G) < \infty\}. \tag{2}$$

2.5 Widom–Rowlinson model

The Widom–Rowlinson [44] model was originally introduced as a (two-type) point process in $\mathbf{R}^d$. The following discrete variant was soon thereafter studied in [43] and [32]. Each vertex of a graph $G = (V, E)$ takes values in $\{-1, 0, 1\}$, where -1 and 1 should be thought of as two types of particles with a mutual hard-core exclusion, and 0 as an empty location. For $\Lambda \subseteq V$, we say that a configuration $\eta \in \{-1, 0, 1\}^\Lambda$ is **WR-feasible** if, for all nearest-neighbor pairs $\langle x, y \rangle$ in Λ we have $\eta(x)\eta(y) \neq -1$. For disjoint vertex sets Λ and Λ' and two configurations $\eta \in \{-1, 0, 1\}^\Lambda$ and $\eta' \in \{-1, 0, 1\}^{\Lambda'}$, we write $(\eta \vee \eta')$ for the configuration on $\Lambda \cup \Lambda'$ which agrees with η on Λ and with η' on Λ'.

Take $G = (V, E) \in \mathcal{G}$, and fix the so-called activity parameter $\lambda > 0$. A probability measure π on $\{-1, 0, 1\}^V$ is said to be a Gibbs measure for the Widom–Rowlinson model on G with activity λ, if it is Markov and for all finite $\Lambda \subset V$, all WR-feasible $\eta \in \{-1, 0, 1\}^{\partial\Lambda}$ and all $\xi \in \{-1, 0, 1\}^\Lambda$ we have

$$\pi(X(\Lambda) = \xi \mid X(\partial\Lambda) = \eta) = Z^{-1} \lambda^{n_-(\xi) + n_+(\xi)} I_{\{(\xi \vee \eta) \text{ WR-feasible}\}}$$

where $n_-(\xi)$ and $n_+(\xi)$ are the number of -1's and $+1$'s in ξ, and $I_{\{(\xi \vee \eta) \text{ WR-feasible}\}}$ is the indicator function of the event that $(\xi \vee \eta)$ is a WR-feasible configuration.

Let $\{\Lambda_n\}_{n=1}^\infty$ be an exhaustion of G, and fix $o \in \Lambda_1$. Define the measures $\pi_{WR}^{\lambda,n,+}$ on $\{-1, 0, 1\}^V$ analogously to $\pi_I^{\beta,n,+}$ for the Ising model. The same arguments as for the Ising model give existence of the limiting Gibbs measure

$$\pi_{WR}^{\lambda,n} = \lim_{n \to \infty} \pi_{WR}^{\lambda,n,+}$$

and the analogous measure $\pi_{WR}^{\lambda,-}$ obtained with minus instead of plus boundary conditions. We also get the following analogue of Proposition 2.1.

Proposition 2.3 *For the Widom–Rowlinson model on $G \in \mathcal{G}$ with activity parameter λ, the following statements are equivalent.*

(i) *There is more than one Gibbs measure.*

(ii) $\pi_{WR}^{\lambda,+} \neq \pi_{WR}^{\lambda,-}$

(iii) $\pi_{WR}^{\lambda,+}(X(o) = +1) > \pi_{WR}^{\lambda,+}(X(o) = -1)$

(iv) $\exists \varepsilon > 0$ *such that* $\pi_{WR}^{\lambda,n,+}(X(o) = +1) \geq \pi_{WR}^{\lambda,n,+}(X(o) = -1) + \varepsilon$ *for all n.*

On the other hand, the existence of more than one Gibbs measure may fail to be increasing in λ, so there is no Widom–Rowlinson analogue of Theorem 2.2 (see Section 3.4). We therefore have to settle for a definition of $\mathcal{G}_{WR}$ which is slightly less elegant than (2): $\mathcal{G}_{WR}$ is the set of graphs $G \in \mathcal{G}$ with the property that there exists some $\lambda > 0$ for which the Widom–Rowlinson model on G has more than one Gibbs measure.

2.6 Beach model

The beach model was introduced by Burton and Steif [9] as an example of a so-called subshift of finite type which has more than one measure of maximal entropy despite having strong irreducibility properties. The following formulation is slightly different from, but (essentially) equivalent to that of Burton and Steif.

Take $G = (V, E) \in \mathcal{G}$. Each vertex will be assigned a value from $\{-2, -1, 1, 2\}$. A configuration $\eta \in \{-2, -1, 1, 2\}^\Lambda$ with $\Lambda \subseteq V$ is said to be **BM-feasible** if for each nearest neighbor pair $\langle x, y \rangle$ we have $\eta(x)\eta(y) \geq -1$. In other words, two spins with different signs may not sit next to each other unless they are both ± 1. A probability measure π on $\{-2, -1, 1, 2\}^V$ is said to be a Gibbs measure for the beach model on G with parameter $M > 1$ if for all finite $\Lambda \subset V$, π-a.e. $\eta \in \{-2, -1, 1, 2\}^{\partial\Lambda}$ and all $\xi \in \{-2, -1, 1, 2\}^\Lambda$ we have

$$\pi(X(\Lambda) = \xi \mid X(\partial\Lambda) = \eta) = Z^{-1}(M-1)^{n_{-2}(\xi)+n_{+2}(\xi)} I_{\{(\xi \vee \eta) \text{ BM-feasible}\}} . \qquad (3)$$

Here $n_{-2}(\xi)$ and $n_{+2}(\xi)$ are the number of -2's and $+2$'s in ξ. The reason why we use the quantifier "π-a.e." rather than "all BM-feasible" for the set of bondary conditions is that certain BM-feasible boundary conditions η may cause $(\xi \vee \eta)$ to be not BM-feasible for all $\xi \in \{-2, -1, 1, 2\}^\Lambda$. (Compared to the original formulation of the model in [9], the present formulation has the advantages of a smaller state space and a real-valued (rather than integer-valued) parameter.)

Again fix an exhaustion $\{\Lambda_n\}_{n=1}^\infty$ and $o \in \Lambda_1$. Let $\pi_{BM}^{M,n,+}$ be the probability measure on $\{-2, -1, 1, 2\}^V$ corresponding to taking $X(V \setminus \Lambda_n) \equiv 2$ and picking $X(\Lambda_n)$ according to (3) with $\Lambda = \Lambda_n$ and $\eta \equiv 2$. As usual, we get a limiting "plus measure"

$$\pi_{BM}^{M,+} = \lim_{n\to\infty} \pi_{BM}^{M,n,+} ,$$

an analogous "minus measure" $\pi_{BM}^{M,-}$ (obtained with boundary condition -2 rather than 2), and the following result:

Proposition 2.4 *For the beach model on $G \in \mathcal{G}$ with parameter M, the following statements are equivalent.*

(i) *There is more than one Gibbs measure.*

(ii) $\pi_{BM}^{M,+} \neq \pi_{BM}^{M,-}$

(iii) $\pi_{BM}^{M,+}(X(o) \geq 1) > \frac{1}{2}$

(iv) $\exists \varepsilon > 0$ *such that* $\pi_{BM}^{M,n,+}(X(o) \geq 1) \geq \frac{1}{2} + \varepsilon$ *for all n.*

Unlike the Widom–Rowlinson model, the beach model has sufficient monotonicity properties to imply the following analogue of Theorem 2.2:

Theorem 2.5 *For any $G \in \mathcal{G}$, there exists a critical value $M_c = M_c(G) \in [1, \infty]$ such that for $M < M_c$ we have that the beach model on G with parameter M has a unique Gibbs measure whereas for $M > M_c$ there are multiple Gibbs measures.*

For $G = \mathbf{Z}^d$, this result was first obtained by Häggström [21]; in Section 3.4 we shall prove the full result using the random-cluster approach. We define

$$\mathcal{G}_{BM} = \{G \in \mathcal{G} : M_c(G) < \infty\} .$$

In this language, the main result in [9] says that $\mathbf{Z}^d \in \mathcal{G}_{BM}$ for $d \geq 2$. Alternative proofs of this result were later given in [21] and in [23].

2.7 The other end of the parameter space

Our main results (Theorems 1.1 and 1.2) concern the "top end" of the parameter space for the five models under consideration, i.e. what happens when p, β, λ and M are sufficiently large. One can of course consider the analogous problem at the other end of the parameter space, where p, β, λ and M are small. For instance, does there exist a $p > 0$ such that $\mu_{BP}^{p}(\exists$ some infinite cluster$) = 0$?

For bounded degree graphs, the answer is yes, because it is well known (and easy to prove) that $p_c^{bond}(G) \geq \frac{1}{\Delta - 1}$ when all vertices in G have degree at most Δ, and the same bound holds for site percolation. The random-cluster methods in Sections 3, 4 and 6 can then be used to show that $\beta_c(G) > 0$ and $M_c(G) > 1$, and also that the Widom–Rowlinson model on G has a unique Gibbs measure for sufficiently small $\lambda > 0$.

For unbounded degree graphs the situation is less clear-cut. If $p_c^{site}(G) > 0$, then we can use random-cluster methods (or disagreement percolation [6]) to show that $\beta_c(G) > 0$ and that the Widom–Rowlinson model on G has Gibbsian uniqueness for small λ. We do not see how to similarly conclude that $M_c(G) > 1$ in this situation.

3 The random-cluster representations

In this section we recall the random-cluster representations of the three Markov random field models under consideration. In the final subsection we also recall a key result on stochastic domination (Holley's Theorem), and use it to demonstrate the monotonicities needed in Theorems 2.2 and 2.5.

3.1 FK representation of the Ising model

It is today widely recognized that the random-cluster model, originally introduced by Fortuin and Kasteleyn [16], is one of the most important tools for studying the Ising model. It is customary to start by defining the random-cluster model on a finite graph, but we shall go directly to the context of an infinite graph $G \in \mathcal{G}$ with an exhaustion $\{\Lambda_n\}_{n=1}^{\infty}$. (For gentler introductions, the reader may turn to [20] or [23].)

Define $E_{\Lambda_n} \subset E$ as the set of edges that have at least one endpoint in $V \setminus \Lambda_n$. The "wired" random-cluster measure $\phi_I^{n,p,q}$ for Λ_n with parameters $p \in [0,1]$ and $q > 0$ is defined as the probability measure on $\{0,1\}^E$ which to each $\xi \in \{0,1\}^E$ assigns probability

$$\phi_I^{n,p,q}(\xi) = Z^{-1} p^{n_1(\xi)} (1-p)^{n_0(\xi)} q^{k(\xi)} I_{\{\xi(e)=1 \text{ for all } e \in E \setminus E_{\Lambda_n}\}},$$

where n_0 (resp. n_1) is the number of edges in E_{Λ_n} taking value 0 (resp. 1), and $k(\xi)$ is the number of connected components in ξ (including isolated vertices) that do not intersect $V \setminus \Lambda_n$. Note that $I_{\{\xi(e)=1 \text{ for all } e \in E \setminus E_{\Lambda_n}\}}$ takes value one for only finitely many ξ, so that in particular the normalizing factor Z^{-1} is well-defined.

The usefulness of the random-cluster model for studying the Ising model should be clear from the following two results.

Proposition 3.1 *Let X be the $\{-1,1\}^V$-valued random spin configuration defined as follows. First, pick an edge configuration $Y \in \{0,1\}^E$ according to the random-cluster measure $\phi_I^{n,p,q}$ with $p = 1 - e^{-2\beta}$ and $q = 2$. Second, obtain X from Y by assigning spins to V in such a way that*

(i) *two vertices in the same connected component of Y always get the same spin,*

(ii) *any connected component of Y intersecting $V \setminus \Lambda_n$ gets spin $+1$, and*

(iii) *all other connected components independently get spin -1 or $+1$ with probability $1/2$ each.*

The distribution of X is then given by $\pi_I^{\beta,n,+}$.

Proof: This follows from a standard counting argument; see e.g. [23]. □

Let $(o \leftrightarrow V \setminus \Lambda_n)$ denote the event that there is some path of open edges connecting o to some vertex in $V \setminus \Lambda_n$.

Corollary 3.2 *With $p = 1 - e^{-2\beta}$, we have*

$$\pi_I^{\beta,n,+}(X(o) = +1) = \frac{1}{2}(1 + \phi_I^{n,p,2}(o \leftrightarrow V \setminus \Lambda_n)).$$

Proof: Immediate from Proposition 3.1. □

3.2 Random-cluster representation of the Widom–Rowlinson model

The so-called site-random-cluster model, discussed e.g. in [18], plays a similar role with respect to the Widom–Rolwinson model as the Fortuin–Kasteleyn random-cluster model does with respect to the Ising model. The site-random-cluster can also be seen as a lattice analogue of the (somewhat less elementary) continuum random-cluster model which was discovered independently by several different research groups and which arises as a random-cluster representation of the original continuum model of Widom and Rowlinson [44]; see e.g. Chayes et al. [12].

Let $G = (V, E)$ and $\{\Lambda_n\}_{n=1}^{\infty}$ be as before. The wired site-random-cluster model $\phi_{WR}^{n,p,q}$ for Λ_n with parameters $p \in [0,1]$ and $q > 0$ is defined as the probability measure on $\{0,1\}^V$ which to each $\xi \in \{0,1\}^V$ assigns probability

$$\phi_{WR}^{n,p,q}(\xi) = Z^{-1} p^{n_1(\xi)}(1-p)^{n_0(\xi)} q^{k(\xi)} I_{\{\xi(e)=1 \text{ for all } v \in V \setminus \Lambda_n\}},$$

where n_0 (resp. n_1) is the number of vertices in Λ_n taking value 0 (resp. 1), and $k(\xi)$ is again the number of connected components in ξ that do not intersect $V \setminus \Lambda_n$.

The following analogues of Proposition 3.1 and Corollary 3.2 are well-known and easy to prove; see e.g. [18].

Proposition 3.3 *Let X be the $\{-1,0,1\}^V$-valued random object defined as follows. First, pick $Y \in \{0,1\}^V$ according to the site-random-cluster measure $\phi_{WR}^{n,p,q}$ with $p = \frac{\lambda}{\lambda+1}$ and $q = 2$. Second, obtain X from Y by letting $X(v) = 0$ for each v such that $Y(v) = 0$, and assigning $+1$'s and -1's to the connected components of 1's in Y in such a way that*

(i) *two vertices in the same connected component of Y always get the same spin,*

(ii) *any connected component of Y intersecting $V \setminus \Lambda_n$ gets spin $+1$, and*

(iii) *all other connected components independently get spin -1 or $+1$ with probability $1/2$ each.*

The distribution of X is then given by $\pi_{WR}^{\lambda,n,+}$.

Corollary 3.4 *With $p = \frac{\lambda}{\lambda+1}$, we have*

$$\pi_{WR}^{\lambda,n,+}(X(o) = +1) - \pi_{WR}^{\lambda,n,+}(X(o) = -1) = \phi_{WR}^{n,p,2}(o \leftrightarrow V \setminus \Lambda_n)).$$

3.3 Random-cluster representation of the beach model

Also the beach model has a random-cluster representation, introduced in [23]. For lack of a better name, we call it the beach-random-cluster model.

The random-cluster representation is defined as follows. Again, let $G = (V, E)$ and $\{\Lambda_n\}_{n=1}^{\infty}$ be as before. For a site configuration $\xi \in \{0,1\}^V$, define the bond configuration $\xi^* \in \{0,1\}^E$ by letting

$$\xi^*(e) = \begin{cases} 1 & \text{if at least one of its endpoints take value 1 in } \xi \\ 0 & \text{otherwise,} \end{cases}$$

for each $e \in E$. The wired beach-random-cluster model $\phi_{BM}^{n,p,q}$ for Λ_n with parameters $p \in [0,1]$ and $q > 0$ is defined as the probability measure on $\{0,1\}^V$ which to each $\xi \in \{0,1\}^V$ assigns probability

$$\phi_{BM}^{n,p,q}(\xi) = Z^{-1} p^{n_1(\xi)} (1-p)^{n_0(\xi)} q^{k^*(\xi)} I_{\{\xi(e)=1 \text{ for all } v \in V \setminus \Lambda_n\}},$$

where n_0 (resp. n_1) is the number of vertices in Λ_n taking value 0 (resp. 1), and $k^*(\xi)$ is the number of connected components in ξ^* (including isolated vertices) that do not intersect $V \setminus \Lambda_n$.

Similarly as in the previous subsections, the following two results are easily established.

Proposition 3.5 *Let X be the $\{-2,-1,1,2\}^V$-valued random object defined as follows. First, pick $Y \in \{0,1\}^V$ according to the beach-random-cluster measure $\phi_{BM}^{n,p,q}$ with $p = \frac{M-1}{M}$ and $q = 2$, and let $Y^* \in \{0,1\}^E$ be the corresponding edge configuration. Second, obtain the absolute values of X from Y by letting $|X(v)| = Y(v) + 1$ for each $v \in V$. Third, assign signs ($+$ or $-$) to X in such a way that*

(i) *two vertices in the same connected component of Y^* always get the same sign,*

(ii) *any connected component of Y^* intersecting $V \setminus \Lambda_n$ gets sign $+$, and*

(iii) *all other connected components independently get sign $-$ or $+$ with probability 1/2 each.*

The distribution of X is then given by $\pi_{BM}^{M,n,+}$.

Corollary 3.6 *With $p = \frac{M-1}{M}$, we have*

$$\pi_{BM}^{M,n,+}(X(o) \geq 1) = \frac{1}{2}(1 + \phi_{BM}^{n,p,q}(o \stackrel{*}{\longleftrightarrow} V \setminus \Lambda_n)).$$

Here $\stackrel{*}{\longleftrightarrow}$ refers to connectivity in the edge configuration Y^*.

3.4 Stochastic domination

For a (finite or infinite) set V and a finite set S of reals, we equip S^V with the usual coordinatewise partial order, denoted $\preceq$. A function $f : S^V \to \mathbf{R}$ is said to be increasing if $f(\xi) \leq f(\eta)$ whenever $\xi \preceq \eta$. For two probability measures μ and μ' on S^V, we say that μ is **stochastically dominated by** μ', writing $\mu \preceq_{\mathcal{D}} \mu'$, if for all increasing f we have

$$\int f \, d\mu \leq \int f d\mu' .$$

If $\mu \preceq_{\mathcal{D}} \mu'$ and X and X' are S^V-valued random elements with distributions μ and μ', then we also write $X \preceq_{\mathcal{D}} X'$. By Strassen's Theorem (see e.g. [33]), this is equivalent to the existence of a coupling P of X and X' such that $P(X \preceq X') = 1$.

A probability measure μ on S^V is said to be **irreducible** if, for any $\xi, \eta \in S^V$ such that both ξ and η have positive μ-probability, we can move from ξ to η through single-site flips without passing through any element of zero μ-probability.

The following result, essentially due to Holley [29], will play a key role in most of the rest of this paper. The proof is the same as Holley's original proof (which he gave under slightly stronger conditions); see e.g. [18].

Theorem 3.7 (Holley) *Let X and X' be S^V-valued random elements with irreducible distributions μ and μ', and assume that μ' assigns positive probability to the maximal element of S^V. If for all $v \in V$, all $s \in S$, μ-a.e. $\xi \in S^{V\setminus\{v\}}$ and μ'-a.e. $\eta \in S^{V\setminus\{v\}}$ such that $\xi \preceq \eta$ we have*

$$\mu(X(v) \geq s \,|\, X(V \setminus \{v\}) = \xi) \leq \mu'(X'(v) \geq s \,|\, X'(V \setminus \{v\}) = \eta)\,,$$

then $\mu \preceq_{\mathcal{D}} \mu'$.

The power of the random-cluster method in conjunction with Holley's Theorem is well illustrated by considering the problems of monotonicity in parameter values mentioned in Sections 2.4–2.6.

Consider first the Ising model on $G \in \mathcal{G}$ at two different parameter values $\beta_1 < \beta_2$. We want to show that if the Ising model on G at inverse temperature β_1 has multiple Gibbs measures, then the same is true for the Ising model on G at inverse temperature β_2. To do this, we use the random-cluster approach devised in [1]. Set $p_1 = 1 - e^{-2\beta_1}$ and $p_2 = 1 - e^{-2\beta_2}$, and consider the random-cluster measures $\phi_I^{n,p_1,2}$ and $\phi_I^{n,p_2,2}$. A direct calculation shows that the conditional $\phi_I^{n,p,2}$-probability that an edge $e \in E_{\Lambda_n}$ is open given everything else, is given by

$$\phi_I^{n,p,2}(Y(e) = 1 | Y(E\setminus\{e\} = \xi)) = \begin{cases} p & \text{if the endpoints } x \text{ and } y \text{ of } e \text{ are either} \\ & \text{connected by an open path in } \xi\text{, or are both} \\ & \text{in open clusters of } \xi \text{ that intersect } V \setminus \Lambda_n, \\ & \\ \frac{p}{2-p} & \text{otherwise.} \end{cases} \tag{4}$$

Since this conditional probability is increasing both in p and in ξ, we may directly apply Theorem 3.7 to $\phi_I^{n,p_1,2}$ and $\phi_I^{n,p_2,2}$, and deduce that

$$\phi_I^{n,p_1,2} \preceq_{\mathcal{D}} \phi_I^{n,p_2,2}\,. \tag{5}$$

[More precisely, we apply Theorem 3.7 to the projections of $\phi_I^{n,p_1,2}$ and $\phi_I^{n,p_2,2}$ on $\{0,1\}^{E_{\Lambda_n}}$, to get stochastic domination between the projected measures. The full stochastic domination (5) follows easily. We will frequently commit this kind of language abuse.] In particular,

$$\phi_I^{n,p_1,2}(o \leftrightarrow V \setminus \Lambda_n) \leq \phi_I^{n,p_2,2}(o \leftrightarrow V \setminus \Lambda_n)\,,$$

whence, by Corollary 3.2,

$$\pi_I^{\beta_1,n,+}(X(o) = +1) \leq \pi_I^{\beta_2,n,+}(X(o) = +1)\,.$$

Hence, if the statement in Proposition 2.1(iv) holds with $\beta = \beta_1$, then the same is true with $\beta = \beta_2$. Proposition 2.1 then gives the desired conclusion: if there are multiple Gibbs measures at $\beta = \beta_1$, then the same thing holds at $\beta = \beta_2$.

Let us next use the same technique to prove the monotonicity claim for the beach model made in Theorem 2.5. We need to show, for $M_1 < M_2$, that if the beach model on G with parameter M_1 has multiple Gibbs measures, then the same is true for the beach model on G with parameter M_2. This we do by comparison between the beach-random-cluster measures $\phi_{BM}^{n,p_1,2}$ and $\phi_{BM}^{n,p_2,2}$, with $p_1 = \frac{M_1-1}{M_1}$ and $p_2 = \frac{M_2-1}{M_2}$. Single-site conditional probabilities for $v \in \Lambda_n$ under $\phi_{BM}^{n,p,2}$ can be calcualted directly from the definition of beach-random-cluster measures. We get

$$\phi_{BM}^{n,p,2}(Y(v) = 1 \,|\, Y(V \setminus \{v\}) = \xi) = \frac{p 2^{1-\kappa^*(v,\xi)}}{p 2^{1-\kappa^*(v,\xi)} + 1 - p} \tag{6}$$

where $\kappa^*(v,\xi)$ is the number of connected components contaning either v or some vertex incident to v, in the edge configuration ξ^* corresponding to ξ. A moment's thought reveals that this conditional distribution is increasing both in p and in ξ, whence by Theorem 3.7 we have

$$\phi_{BM}^{n,p_1,2} \preceq_{\mathcal{D}} \phi_{BM}^{n,p_2,2}. \tag{7}$$

The desired conclusion now follows using Corollary 3.6 and Proposition 2.4 (in the same way that we applied Corollary 3.2 and Proposition 2.1 to the Ising model above).

Finally, we may try to apply the same method to the Widom–Rowlinson model. Single-site conditional probabilities for $v \in \Lambda_n$ under the site-random-cluster measure $\phi_{WR}^{n,p,2}$ turn out to be

$$\phi_{WR}^{n,p,2}(Y(v) = 1 \,|\, Y(V \setminus \{v\}) = \xi) = \frac{p 2^{1-\kappa(v,\xi)}}{p 2^{1-\kappa(v,\xi)} + 1 - p} \tag{8}$$

where $\kappa(v,\xi)$ is the number of connected components in ξ that intersect the neighborhood of v. This conditional probability is increasing in p *but not in* ξ, so the use of Theorem 3.7 to obtain a Widom–Rowlinson analogue of (5) and (7) is unwarranted. The crucial difference between the models is that in both the FK random-cluster model and the beach-random-cluster model, the number of connected components is decreasing in ξ, whereas in the site-random-cluster model it is not. This, in turn, is a reflection of the fact that the former two models count the number of connected components in an edge configuration, while the latter deals with a site configuration: adding an open edge can never increase the number of connected components, but adding an open vertex can.

One may ask whether this reflects a fundamental difference between the Widom–Rowlinson model and the other models, or just a shortcoming of the random-cluster approach. The (perhaps surprising) answer is that the desired monotonicity of the Gibbs measure multiplicity phenomenon *fails* for the Widom–Rowlinson model on certain graphs. Examples of graphs where the existence of more than one Gibbs measure varies nonmonotonically in λ are given in [8].

4 Proof of Theorem 1.2 – positive implications

In this section we prove the positive parts of Theorem 1.2. These are:

Lemma 4.1 $\mathcal{G}_{SP} \subseteq \mathcal{G}_{BP}$

Lemma 4.2 $\mathcal{G}_{BP} \subseteq \mathcal{G}_{I}$

Lemma 4.3 $\mathcal{G}_{I} \subseteq \mathcal{G}_{BP}$

Lemma 4.4 $\mathcal{G}_{WR} \subseteq \mathcal{G}_{SP}$

In Section 5, we will then prove the following negative results:

Lemma 4.5 *There exists a graph G which is in $\mathcal{G}_{BP}$ but not in $\mathcal{G}_{SP}$.*

Lemma 4.6 *There exists a graph G which is in $\mathcal{G}_{SP}$ but not in $\mathcal{G}_{WR}$.*

Lemma 4.7 *There exists a graph G which is in $\mathcal{G}_{SP}$ but not in $\mathcal{G}_{BM}$.*

Lemma 4.8 *There exists a graph G which is in $\mathcal{G}_{BM}$ but not in $\mathcal{G}_{SP}$.*

It is clear that Theorem 1.2 follows once we have proved Lemmas 4.1–4.8.

Proof of Lemma 4.1: As mentioned already in the introduction, this is immediate from the old result of Hammersley [28] that $p_c^{bond}(G) \leq p_c^{site}(G)$ for any $G \in \mathcal{G}$. □

Proof of Lemma 4.2: Suppose that $G \in \mathcal{G}_{BP}$, and pick $p' \in (p_c^{bond}(G), 1)$. Let $(o \leftrightarrow \infty)$ denote the event that o is in an infinite cluster. Set $\alpha = \mu_{BP}^{p'}(o \leftrightarrow \infty)$, and note that $\alpha > 0$ by the choice of p'. Set $p = \frac{2p'}{1+p'}$ so that $p' = \frac{p}{2-p}$. Consider the random-cluster measure $\phi_I^{n,p,2}$, and note that its single-edge conditional probabilities in (4) are bounded from below by p'. Theorem 3.7 therefore implies that

$$\phi_I^{n,p,2} \succeq_{\mathcal{D}} \mu_{BP}^{p'}$$

whence

$$\begin{aligned} \phi_I^{n,p,2}(o \leftrightarrow V \setminus \Lambda_n) &\geq \mu_{BP}^{p'}(o \leftrightarrow V \setminus \Lambda_n) \\ &\geq \mu_{BP}^{p'}(o \leftrightarrow \infty) \\ &= \alpha\,, \end{aligned}$$

so that by Corollary 3.2 we have

$$\pi_I^{\beta,n,+}(X(o) = +1) \geq \frac{1+\alpha}{2}$$

for $\beta = -\frac{1}{2}\log(1-p)$. Since this bound is independent of n, we have that the statement in Proposition 2.1(iv) holds with $\varepsilon = \alpha$. The Ising model on G with parameter β therefore has more than one Gibbs measure by Proposition 2.1, so $G \in \mathcal{G}_I$ as needed. □

Proof of Lemma 4.3: Suppose that $G \in \mathcal{G} \setminus \mathcal{G}_{BP}$; we need to show that $G \in \mathcal{G} \setminus \mathcal{G}_I$. Let $\beta > 0$ be arbitrary, and set $p = 1 - e^{-2\beta}$. Since $G \in \mathcal{G} \setminus \mathcal{G}_{BP}$, we have $\mu_{BP}^{p}(o \leftrightarrow \infty) = 0$, whence

$$\lim_{n\to\infty} \mu_{BP}^{p}(o \leftrightarrow V \setminus \Lambda_n) = 0\,. \tag{9}$$

By Theorem 3.7 and the fact that the conditional $\phi_I^{n,p,2}$-probability in (4) is bounded above by p, we have that the projection on $\{0,1\}^{E_{\Lambda_n}}$ of $\phi_I^{n,p,2}$ is stochastically dominated by the projection on $\{0,1\}^{E_{\Lambda_n}}$ of μ_{BP}^p. In particular,

$$\phi_I^{n,p,2}(o \leftrightarrow V \setminus \Lambda_n) \leq \mu_{BP}^p(o \leftrightarrow V \setminus \Lambda_n)\,,$$

so that by (9) we have

$$\lim_{n\to\infty} \phi_I^{n,p,2}(o \leftrightarrow V \setminus \Lambda_n) = 0\,.$$

Corollary 3.2 then implies that

$$\lim_{n\to\infty} \pi_I^{\beta,n,+}(X(o) = +1) = \frac{1}{2}$$

so that, by Proposition 2.1, the Ising model on G with parameter β has a unique Gibbs measure. But β was arbitrary, so $G \in \mathcal{G} \setminus \mathcal{G}_I$. □

Proof of Lemma 4.4: We proceed similarly as in the proof of Lemma 4.3: we assume that $G \in \mathcal{G} \setminus \mathcal{G}_{SP}$, and need to show that $G \in \mathcal{G} \setminus \mathcal{G}_{WR}$. Let $\lambda > 0$ be arbitrary, set $p = \frac{\lambda}{\lambda+1}$ and $p' = \frac{2\lambda}{2\lambda+1}$, and note that $\frac{2p}{p+1} = p'$. By the choice of G, we have $\mu_{SP}^{p'}(o \leftrightarrow \infty) = 0$, so that

$$\lim_{n\to\infty} \mu_{SP}^{p'}(o \leftrightarrow V \setminus \Lambda_n) = 0\,. \tag{10}$$

The conditional $\phi_{WR}^{n,p,2}$-probability in (8) is bounded above by $\frac{2p}{p+1}$, i.e. by p'. Theorem 3.7 therefore implies that the projection on $\{0,1\}^{E_{\Lambda_n}}$ of $\phi_{WR}^{n,p,2}$ is stochastically dominated by the projection on $\{0,1\}^{E_{\Lambda_n}}$ of $\mu_{SP}^{p'}$. Hence,

$$\phi_{WR}^{n,p,2}(o \leftrightarrow V \setminus \Lambda_n) \leq \mu_{SP}^{p'}(o \leftrightarrow V \setminus \Lambda_n)\,,$$

which in conjunction with (10) implies that

$$\lim_{n\to\infty} \phi_{WR}^{n,p,2}(o \leftrightarrow V \setminus \Lambda_n) = 0\,.$$

Corollary 3.4 then implies that

$$\lim_{n\to\infty} \left(\pi_{WR}^{\lambda,n,+}(X(o) = +1) - \pi_{WR}^{\lambda,n,+}(X(o) = -1)\right) = 0$$

whence by Proposition 2.3 the Widom–Rowlinson model on G with parameter λ has a unique Gibbs measure. Since λ was arbitrary, we have $G \in \mathcal{G} \setminus \mathcal{G}_{WR}$ as desired. □

5 Proof of Theorem 1.2 – counterexamples

In this section we finish the proof of Theorem 1.2 by providing the counterexamples needed to demonstrate Lemmas 4.5–4.8. All examples will explicitly contain a sequence of vertices whose degree tend to ∞ (this is no surprise in view of Theorem 1.1).

Proof of Lemma 4.5: Construct $G = (V, E)$ as follows. For some sequence $\{k_i\}_{i=0}^{\infty}$ of positive integers, let

$$V = \{v_0, v_1, \ldots\} \cup \{v_{i,j} : i \in \{0, 1, \ldots\}, j \in \{1, \ldots, k_i\}\}$$

and

$$\begin{aligned} E &= \{\{v_i, v_{i,j}\} : i \in \{0,1,\ldots\}, j \in \{1,\ldots,k_i\}\} \\ &\quad \cup \{\{v_{i,j}, v_{i+1}\} : i \in \{0,1,\ldots\}, j \in \{1,\ldots,k_i\}\}. \end{aligned}$$

Pictorially, G consists of the vertex sequence $v_0, v_1, \ldots$ in which each pair (v_i, v_{i+1}) is connected by k_i parallel paths of length 2. For site percolation on G, it is clear that v_0 fails to be in an infinite cluster if any of the vertices $v_0, v_1, \ldots$ is closed. Hence, $\mu^p_{SP}(v_0 \leftrightarrow \infty) = 0$ for any $p < 1$, so $G \in \mathcal{G} \setminus \mathcal{G}_{SP}$ (regardless of the choice of $\{k_i\}_{i=0}^{\infty}$).

On the other hand, consider bond percolation on G. We have

$$\mu^p_{BP}(v_i \not\leftrightarrow v_{i+1}) = (1-p^2)^{k_i}$$

so that

$$\begin{aligned} \mu^p_{BP}(v_0 \leftrightarrow \infty) &= \mu^p_{BP}(\cap_{i=0}^{\infty}\{v_i \leftrightarrow v_{i+1}\}) \\ &= \prod_{i=0}^{\infty} \mu^p_{BP}(v_i \leftrightarrow v_{i+1}) \end{aligned}$$

which is positive if and only if

$$\sum_{i=0}^{\infty}(1-p^2)^{k_i} < \infty .$$

For instance, taking $k_i = \frac{\log i}{\log 2}$ rounded up to the nearest integer, we get $p_c^{bond}(G) = 1/\sqrt{2}$, so that $G \in \mathcal{G}_{BP}$. □

Proof of Lemma 4.6: The counterexample $G \in \mathcal{G}_{SP} \setminus \mathcal{G}_{WR}$ will be constructed by "decorating" another graph G' with "dead ends". G' can be taken to be any graph whose critical value for site percolation is strictly between 0 and 1; for concreteness we take $G' = (V', E')$ to be the usual square lattice, i.e. $V' = \mathbf{Z}^2$ and E' consists of all pairs of vertices at Euclidean distance 1 from each other. It is well known that $p_c^{site}(G')$ is strictly between 0 and 1; see e.g. [19].

To obtain $G = (V, E)$ from G', we add a number of edges coming out of each $v \in V'$, and each of these edges end up in a single new vertex which has no further edges incident to it. More precisely, let $\{k_v\}_{v \in V'}$ be positive integers associated with the vertices in G', and set

$$V = V' \cup \{w_v^i : v \in V', i \in \{1,\ldots,k_v\}\}$$

and

$$E = E' \cup \{\{v, w_v^i\} : v \in V', i \in \{1,\ldots,k_v\}\}.$$

The only assumption we need to make on $\{k_v\}_{v \in V'}$ is that k_v tends to infinity as we move away from the origin in V', i.e. that for any M there are at most finitely many $v \in V'$ with $k_v \le M$. To show that $G \in \mathcal{G}_{SP} \setminus \mathcal{G}_{WR}$, what we need to do is to show

(i) $p_c^{site}(G) < 1$, and

(ii) for any $\lambda > 0$, the Widom–Rowlinson model on G with parameter λ has a unique Gibbs measure.

Since G' is a subgraph of G, we obviously have

$$p_c^{site}(G) \leq p_c^{site}(G') , \tag{11}$$

so (i) is immediate (in fact, (11) holds with equality).

It remains to prove (ii). Fix $\lambda > 0$. Let $\{\Lambda_n\}_{n=1}^{\infty}$ be an exhaustion of G with the property that for all w_v^i and all n we have $w_v^i \in \Lambda_n$ if and only if $v \in \Lambda_n$. Also set $\Lambda'_n = \Lambda_n \cap V'$ for each n. Consider the measure $\pi_{WR}^{\lambda,n,+}$ and its projection on $\{-1,0,1\}^{\Lambda'}$ (although here and throughout the proof, all processes live on G rather than on G'). For any $v \in \Lambda'_n$, we have

$$\frac{\pi_{WR}^{\lambda,n,+}(X(v) = 1 \,|\, X(\Lambda'_n \setminus \{v\}) = \xi)}{\pi_{WR}^{\lambda,n,+}(X(v) = 0 \,|\, X(\Lambda'_n \setminus \{v\}) = \xi)} = \begin{cases} 0 & \text{if } \xi(u) = -1 \text{ for some} \\ & u \in \Lambda' \text{ with } u \sim v \\ & \\ \lambda(\frac{\lambda+1}{2\lambda+1})^{k_v} & \text{otherwise;} \end{cases}$$

this follows easily by summing the conditional probabilities of all possible configurations $\eta \in \{-1,0,1\}^{\{v,w_v^1,\ldots,w_v^{k_v}\}}$. A similar formula holds for $\frac{\pi_{WR}^{\lambda,n,+}(X(v)=-1 \,|\, X(\Lambda'_n \setminus\{v\})=\xi)}{\pi_{WR}^{\lambda,n,+}(X(v)=0 \,|\, X(\Lambda'_n \setminus\{v\})=\xi)}$, and it follows that

$$\pi_{WR}^{\lambda,n,+}(X(v) \in \{-1,1\} \,|\, X(\Lambda'_n \setminus \{v\}) = \xi) \leq 2\lambda \left(\frac{\lambda+1}{2\lambda+1}\right)^{k_v} . \tag{12}$$

Translating this to the random-cluster representation using Proposition 3.3, we get (with $p = \frac{\lambda}{\lambda+1}$)

$$\phi_{WR}^{n,p,2}(Y(v) = 1 \,|\, Y(\Lambda'_n \setminus \{v\}) = \eta) \leq 2\lambda \left(\frac{\lambda+1}{2\lambda+1}\right)^{k_v} . \tag{13}$$

Fix $p' \in (0, p_c^{site}(G'))$. Note that the right hand side of (13) tends to 0 as $k_v \to \infty$. By Theorem 3.7 and the assumption made on $\{k_v\}_{v \in V'}$, we can therefore find an m such that for any $n > m$, the projection on $\{0,1\}^{\Lambda'_n \setminus \Lambda'_m}$ of $\phi_{WR}^{n,p,2}$ is stochastically dominated by the corresponding projection of $\mu_{SP}^{p'}$. Since $p' < p_c^{site}(G')$, we have

$$\lim_{n\to\infty} \mu_{SP}^{p'}(\Lambda'_m \leftrightarrow V' \setminus \Lambda'_n) = 0$$

so that

$$\begin{aligned} \lim_{n\to\infty} \phi_{WR}^{n,p,2}(o \leftrightarrow V' \setminus \Lambda'_n) &\leq \lim_{n\to\infty} \phi_{WR}^{n,p,2}(\Lambda'_m \leftrightarrow V' \setminus \Lambda'_n) \\ &\leq \mu_{SP}^{p'}(\Lambda'_m \leftrightarrow V' \setminus \Lambda'_n) \\ &= 0 . \end{aligned} \tag{14}$$

By Corollary 3.4 and Proposition 2.3, we have that the Widom–Rowlinson model on G with parameter λ has a unique Gibbs measure. Since λ was arbitrary, we have (ii), and the proof is complete. $\Box$

Remark: The example in the above proof is essentially the same as the one used by Brightwell, Häggström and Winkler [8] to show that for two graphs G and G', we can have that G is "larger" than another graph G', yet $G \in \mathcal{G} \setminus \mathcal{G}_{WR}$, $G' \in \mathcal{G}_{WR}$. The significance of graphs with this kind of "dead end decorations" was first discovered by Schonmann and Tanaka [42].

Proof of Lemma 4.7: Let G be as in the proof of Lemma 4.6; we need to show that for any M the beach model on G with parameter M has a unique Gibbs measure. The proof proceeds similarly as the proof of Lemma 4.6: A calculation shows that in place of (12), we get

$$\pi_{BM}^{\lambda,n,+}(X(v) \in \{-2,2\} \,|\, X(\Lambda_n' \setminus \{v\}) = \xi) \leq (M-1)\left(\frac{M}{M+1}\right)^{k_v}$$

so that in place of (13) we get (with $p = \frac{M-1}{M}$)

$$\phi_{BM}^{n,p,2}(Y(v) = 1 \,|\, Y(\Lambda_n' \setminus \{v\}) = \eta) \leq (M-1)\left(\frac{M}{M+1}\right)^{k_v}$$

which tends to 0 as $k_v \to \infty$. In the remainder of the proof we just need to make one more modification compared to the Widom–Rowlinson case: The relevant connectivity in the beach model analogue of (14) is $\stackrel{*}{\longleftrightarrow}$ rather than $\leftrightarrow$. We therefore have to pick

$$p' \in (0, p_c^{site}(G'')) \tag{15}$$

where G'' is the graph obtained from G' by adding an edge between any pair of vertices $u, v \in \mathbf{Z}^2$ whose graph-theoretic distance in G' is 2. The proof then goes through as for the Widom–Rowlinson case, although it remains to show that we can pick p' as in (15), i.e. to show that $p_c^{site}(G'') > 0$. This, however, is immediate from Hammersley's [27] result that any graph whose degree is bounded by some Δ has $p_c^{site} \geq \frac{1}{\Delta-1}$; we get $p_c^{site} \geq \frac{1}{11}$. □

Proof of Lemma 4.8: Let $G = (V, E)$ be as in the proof of Lemma 4.5, with $k_i = \frac{\log i}{\log 2}$ rounded up to the nearest integer. We know from that proof that $p_c^{site}(G) = 1$, so we are done if we can find an M such that the beach model on G with parameter M has multiple Gibbs measures. Take $M = 6$ for concreteness (actually, any $M > 5$ suffices for our argument), so that the corresponding parameter $p = \frac{M-1}{M}$ for the beach-random-cluster model is given by $p = 5/6$. Define the exhaustion $\{\Lambda_n\}_{n=1}^{\infty}$ of G by setting

$$\Lambda_n = \{v_0, \ldots, v_n\} \cup \{v_{i,j} : i \in \{0, \ldots, n-1\}, j \in \{1, \ldots, k_i\}\}.$$

Each $v_{i,j}$ has just two nearest neighbors, whence for $v_{i,j} \in \Lambda_n$ we have that the conditional $\phi_{BM}^{n,p,2}$-probability in (6) is at least $\frac{p2^{-2}}{p2^{-2}+1-p} = \frac{5}{9}$. We can therefore apply Theorem 3.7 to get that

$$\phi_{BM}^{n,p,2} \succeq_{\mathcal{D}} \tilde{\mu}$$

for each n, where $\tilde{\mu}$ is the probability measure on $\{0,1\}^V$ where each v_i takes value 0 a.s., and each $v_{i,j}$ independently takes value 1 with probability 5/9, and 0 with probability 4/9. To have $(v_0 \stackrel{*}{\longleftrightarrow} V \setminus \Lambda_n)$, it is enough that for all $i \in \{0, \ldots, n-1\}$, there is some $j \in \{1, \ldots, k_i\}$ such that $v_{i,j}$ takes value 1. Since $\sum_{i=1}^{\infty}(1-\frac{5}{9})^{k_i} < \infty$ by the choice of $\{k_i\}_{i=0}^{\infty}$, we have

$$\begin{aligned} \phi_{BM}^{n,p,2}(v_0 \stackrel{*}{\longleftrightarrow} V \setminus \Lambda_n) &\geq \tilde{\mu}(v_0 \stackrel{*}{\longleftrightarrow} V \setminus \Lambda_n) \geq \tilde{\mu}(v_0 \stackrel{*}{\longleftrightarrow} \infty) \\ &= \prod_{i=0}^{\infty} \tilde{\mu}(v_i \stackrel{*}{\longleftrightarrow} v_{i+1}) \geq \prod_{i=0}^{\infty} (\tfrac{5}{9})^{k_i} > 0. \end{aligned}$$

This shows that $\phi_{BM}^{n,p,2}(v_0 \stackrel{*}{\longleftrightarrow} V \setminus \Lambda_n)$ is bounded away from 0 uniformly in n, so we can apply Corollary 3.6 and Proposition 2.4 to deduce that the beach model on G with parameter $M = 6$ has multiple Gibbs measures, as desired. □

6 Proof of Theorem 1.1

This section is devoted to the proof of Theorem 1.1. In view of Theorem 1.2, it is enough to prove the following four lemmas:

Lemma 6.1 $\mathcal{G}^b_{BP} \subseteq \mathcal{G}^b_{SP}$

Lemma 6.2 $\mathcal{G}^b_{SP} \subseteq \mathcal{G}^b_{WR}$

Lemma 6.3 $\mathcal{G}^b_{SP} \subseteq \mathcal{G}^b_{BM}$

Lemma 6.4 $\mathcal{G}^b_{BM} \subseteq \mathcal{G}^b_{BP}$

The proofs of Lemmas 6.1 and 6.4 use the following result of Liggett, Schonmann and Stacey [34]. For $G = (V, E) \in \mathcal{G}$, consider some probability measure ν on $\{0,1\}^V$, and a corresponding $\{0,1\}^V$-valued random element X. We say that ν is **1-dependent** if $X(A)$ and $X(B)$ are independent for all finite disjoint edge sets $A, B \subset E$ such that no pair (v_1, v_2) with $v_1 \in A$, $v_2 \in B$, share an edge.

Theorem 6.5 (Liggett, Schonmann and Stacey) *Suppose that $G = (V, E) \in \mathcal{G}^b$. For each $p < 1$, there exists some $p' < 1$ such that the following holds: For all 1-dependent probability measures ν on $\{0,1\}^V$ with the property that*

$$\nu(X(v) = 1) \geq p' \quad \textit{for all } v \in V\,,$$

we have $\nu \succeq_D \mu^p_{SP}$.

Next, let ν be a probability measure on $\{0,1\}^E$, and let Y be the corresponding $\{0,1\}^E$-valued random element. We say that ν is 1-dependent if $X(A)$ and $X(B)$ are independent for all finite disjoint edge sets $A, B \subset E$ with the property that no pair (e_1, e_2) with $e_1 \in A$, $e_2 \in B$, share an endpoint. The following edge version of Theorem 6.5 follows immediately from the original vertex result by applying it to the graph $G' = (V', E')$ with $V' = E$ and

$$E' = \{\{e_1, e_2\} \in E \times E : e_1 \text{ and } e_2 \text{ share an endpoint in } V\}$$

(note that if G has bounded degree, then so has G').

Corollary 6.6 *Suppose that $G = (V, E) \in \mathcal{G}^b$. For each $p < 1$, there exists some $p' < 1$ such that the following holds. For all 1-dependent probability measures ν on $\{0,1\}^E$ with the property that*

$$\nu(X(e) = 1) \geq p' \quad \textit{for all } e \in E\,,$$

we have $\nu \succeq_D \mu^p_{BP}$.

Proof of Lemma 6.1: Fix $G \in \mathcal{G}^b_{BP}$ and $p \in (p_c^{bond}(G), 1)$. Then pick $p' < 1$ as in Corollary 6.6, and set $p'' = \sqrt{p'}$. Let X be a $\{0,1\}^V$-valued random element with distribution $\mu^{p''}_{SP}$, and let Y be the $\{0,1\}^E$-valued random element obtained by setting

$$Y(e) = \begin{cases} 1 & \text{if } X(u) = X(v) = 1 \text{ for the two endpoints } u \text{ and } v \text{ of } e \\ 0 & \text{otherwise,} \end{cases}$$

and write ν for the distribution of Y. It is clear that ν is 1-dependent, and moreover

$$\nu(Y(e)=1)=(p'')^2=p'$$

for each $e \in E$. By the choice of p', we have from Theorem 6.5 that $\nu \succeq_D \mu^p_{BP}$, so that by the choice of p we have $\nu(o \leftrightarrow \infty) > 0$. However, it is easy to see that the set of vertices that can be reached from o through open paths in X coincides with the set of vertices that can be reached from o through open paths in Y. Hence $\mu^{p''}_{SP}(o \leftrightarrow \infty) > 0$, so that $p_c^{site}(G) \le p''$ and $G \in \mathcal{G}^b_{SP}$. □

In the remaining proofs in this section, $\Delta(G)$ will denote the maximum degree of vertices in the graph $G \in \mathcal{G}^b$.

Proof of Lemma 6.2: Fix $G \in \mathcal{G}^b_{SP}$ and $p' \in (p_c^{site}(G), 1)$. Set $p = \frac{p'}{p'+2^{1-\Delta(G)}(1-p')}$. The quantity $\kappa(v,\xi)$ in (8) is bounded by $\Delta(G)$, and therefore the conditional $\phi^{n,p,2}_{WR}$-probability in (8) is bounded from below by $\frac{p2^{1-\Delta(G)}}{p2^{1-\Delta(G)}+1-p} = p'$. Theorem 3.7 thus ensures that

$$\phi^{n,p,2}_{WR} \succeq_{\mathcal{D}} \mu^{p'}_{SP}\,,$$

so that

$$\liminf_{n\to\infty} \phi^{n,p,2}_{WR}(o \leftrightarrow V \setminus \Lambda_n) \ge \lim_{n\to\infty} \mu^{p'}_{SP}(o \leftrightarrow V \setminus \Lambda_n) > 0\,. \tag{16}$$

By applying Corollary 3.4 and Proposition 2.3 in the usual way, we get that the Widom–Rowlinson model on G with parameter $\lambda = \frac{p}{1-p}$ has multiple Gibbs measures, so that $G \in \mathcal{G}^b_{WR}$. □

Proof of Lemma 6.3: This is similar to the proof of Lemma 6.2. Fix $G \in \mathcal{G}^b_{SP}$ and $p' \in (p_c^{site}(G), 1)$, and set $p = \frac{p'}{p'+2^{1-\Delta(G)}(1-p')}$. The conditional $\phi^{n,p,2}_{BM}$-probability in (6) is always at least $\frac{p2^{1-\Delta(G)}}{p2^{1-\Delta(G)}+1-p} = p'$, so that by Theorem 3.7 we have

$$\phi^{n,p,2}_{BM} \succeq_{\mathcal{D}} \mu^{p'}_{SP}\,.$$

Hence (16) holds with $\phi^{n,p,2}_{WR}$ replaced by $\phi^{n,p,2}_{BM}$. Corollary 3.6 and Proposition 2.4 then allow us to deduce that the beach model on G with parameter $M = \frac{1}{1-p}$ has multiple Gibbs measures, and $G \in \mathcal{G}^b_{BM}$. □

Proof of Lemma 6.4: Fix $G \in \mathcal{G}^b_{BM}$ and $M > M_c(G)$, and set $p = \frac{M-1}{M}$. The single-site conditional $\phi^{n,p,2}_{BM}$-probability in (6) is bounded by $\frac{2p}{p+1}$. Hence, Theorem 3.7 implies that

the projection on Λ_n of $\phi^{n,p,2}_{BM}$ is stochastically dominated by the corresponding projection of $\mu^{p'}_{SP}$, (17)

where $p' = \frac{2p}{p+1}$.

For some (not yet specified) $\tilde{p} < 1$, consider the $\{0,1\}^V$-valued random element X obtained by picking $Y \in \{0,1\}^E$ according to $\mu^{\tilde{p}}_{BP}$, and then setting

$$X(v) = \begin{cases} 1 & \text{if } Y(e)=1 \text{ for all edges } e \text{ incident to } v \\ 0 & \text{otherwise} \end{cases}$$

for each $v \in V$. Write $\nu^{\tilde{p}}$ for the distribution of X. Clearly, $\nu^{\tilde{p}}$ is 1-dependent, and satisfies

$$\nu^{\tilde{p}}(X(v)=1) \ge \tilde{p}^{\Delta(G)} \tag{18}$$

for each $v \in V$. Note that the right hand side of (18) tends to 1 as $\tilde{p} \to 1$, so that Theorem 6.5 guarantees that

$$\nu^{\tilde{p}} \succeq_D \mu_{SP}^{p'} \tag{19}$$

if $\tilde{p}$ is sufficiently close to 1. We choose $\tilde{p}$ in such a way that (19) holds. By combining the stochastic dominations in (17) and (19), we get that

$$\mu_{BP}^{\tilde{p}}(o \leftrightarrow V \setminus \Lambda_n) \geq \phi_{BM}^{n,p,2}(o \overset{*}{\longleftrightarrow} V \setminus \Lambda_n)\,. \tag{20}$$

Using Proposition 2.4, Corollary 3.6 and the choice of M, we get that the right hand side of (20) is bounded away from 0 as $n \to \infty$. Hence,

$$\mu_{BM}^{\tilde{p}}(o \leftrightarrow \infty) > 0$$

and $G \in \mathcal{G}_{BP}^b$ as desired. □

7 A sufficient condition for phase transition in unbounded degree graphs

The counterexamples in Section 5 used to show that Theorem 1.1 does not extend to arbitrary unbounded degree graphs, are all somewhat artificial. This raises the question of whether there are natural ways to weaken the bounded degree assumption and still have equivalence between phase transition in the five models under consideration. In this section we provide such a condition, and then demonstrate its usefulness by applying it to supercritical Galton–Watson trees and to Poisson–Voronoi tessellations.

For a graph $G = (V, E) \in \mathcal{G}$ and a vertex $v \in V$, we write $\deg_G(v)$ for the the number of nearest neighbors of v in G. By a **subgraph** of G, we here mean a connected graph $H = (V_H, E_H)$ whose vertex set V_H is a subset of V, and whose edge set E_H is the set of edges in E that have both endpoints in V_H.

Theorem 7.1 *Suppose that the graph $G \in \mathcal{G}$ has a subgraph $H = (V_H, E_H)$ satisfying*

(i) $p_c^{bond}(H) < 1$, *and*

(ii) $\sup_{v \in V_H} \deg_G(v) < \infty$.

Then $G \in \mathcal{G}_{BP} \cap \mathcal{G}_{SP} \cap \mathcal{G}_I \cap \mathcal{G}_{WR} \cap \mathcal{G}_{BM}$.

(Note that condition (ii) is stronger than just assuming that H has bounded degree, because we are taking the supremum of $\deg_G(v)$ rather than $\deg_H(v)$.)

Proof: Condition (i) says that $H \in \mathcal{G}_{BP}$. Since H has bounded degree, we have by Theorem 1.1 that $H \in \mathcal{G}_{SP}$. Since p_c^{bond} and p_c^{site} are nonincreasing in the graph structure (in the obvious sense) we get that $G \in \mathcal{G}_{BP}$ and $G \in \mathcal{G}_{SP}$. By Theorem 1.2, we then also have $G \in \mathcal{G}_I$. It remains to show that $G \in \mathcal{G}_{WR}$ and $G \in \mathcal{G}_{BM}$; we will show only $G \in \mathcal{G}_{WR}$, as the proof of $G \in \mathcal{G}_{BM}$ is completely analogous.

Pick $o \in V_H$ and let $\{\Lambda_n\}_{n=1}^{\infty}$ be an exhaustion of G such that $o \in \Lambda_1$. Also set $\Lambda_n' = \Lambda_n \cap V_H$ for each n. Pick $p' > p_c^{site}(H)$, and consider site percolation on G with parameter p'. By the choice of p', we have that

$$\lim_{n\to\infty} \mu_{SP}^{p'}(o \overset{H}{\longleftrightarrow} V \setminus \Lambda_n) > 0$$

where $(o \overset{H}{\longleftrightarrow} V \setminus \Lambda_n)$ denotes the existence of an open path from o to some vertex in $V \setminus \Lambda_n$ using only vertices in V_H. Write Δ for the supremum in (ii). Set $p = \frac{p'}{p'+2^{1-\Delta}(1-p')}$, and consider the site-random-cluster model on G. For $v \in \Lambda'_n$, we have as in the proof of Lemma 6.2 that the conditional $\phi^{n,p,2}_{WR}$-probability in (8) is bounded from below by p'. By Theorem 3.7, we therefore have the projection of $\phi^{n,p,2}_{WR}$ on $\{0,1\}^{\Lambda'_n}$ stochastically dominates the same projection of $\mu^{p'}_{SP}$. Hence,

$$\begin{aligned}
\lim_{n\to\infty} \phi^{n,p,2}_{WR}(o \leftrightarrow V \setminus \Lambda_n) &\geq \lim_{n\to\infty} \phi^{n,p,2}_{WR}(o \overset{H}{\longleftrightarrow} V \setminus \Lambda_n) \\
&\geq \lim_{n\to\infty} \mu^{p'}_{SP}(o \overset{H}{\longleftrightarrow} V \setminus \Lambda_n) \\
&> 0\,.
\end{aligned}$$

The proof is now finished by the same application of Corollary 3.4 and Proposition 2.3 as in the proof of Lemma 6.2. □

7.1 Galton–Watson trees

A tree is a graph with no cycles. A Galton–Watson tree is a random tree $T = (V_T, E_T)$ obtained as follows (see e.g. [40, 37] for more extensive discussions). Let L be a non-negative integer-valued random variable, and set $p_k = \mathbf{P}(L = k)$ for each k. We call $\{p_k\}_{k=0}^{\infty}$ the offspring distribution for T. The so-called root $\rho \in V_T$ has L neighbors $v_1, \ldots, v_L$ called the children of ρ. For each $v \in \{v_1, \ldots, v_L\}$, we let L_v be an independent copy of L, and introduce L_v vertices, adjacent to v, called the children of v, and continue inductively in the same way. We consider the supercritical case, i.e. the case where $\mathbf{E}[L] = \sum_{k=0}^{\infty} k p_k > 1$; except for the trivial case $p_1 = 1$, it is precisely for supercritical Galton–Watson trees that we have $\mathbf{P}(T \text{ is infinite}) > 0$. Using Theorem 7.1, we shall prove the following.

Theorem 7.2 *If T is a supercritical Galton–Watson tree, then*

$$\mathbf{P}(T \in \mathcal{G}_{BP} \cap \mathcal{G}_{SP} \cap \mathcal{G}_I \cap \mathcal{G}_{WR} \cap \mathcal{G}_{BM} \,|\, T \textit{ is infinite}) = 1\,.$$

Note that if L is unbounded, then $\mathbf{P}(T \text{ has unbounded degree} \,|\, T \text{ in infinite}) = 1$; this is an easy application of [37, Chapter 3, Prop. 6].

We remark that it is only the inclusions $T \in \mathcal{G}_{WR}$ and $T \in \mathcal{G}_{BM}$ that are new results; the other inclusions $T \in \mathcal{G}_{BP}$, $T \in \mathcal{G}_{SP}$ and $T \in \mathcal{G}_I$ are clear from the work of Lyons [35, 36, 37].

One may ask whether $T \in \mathcal{G}_{BP} \cap \mathcal{G}_{SP} \cap \mathcal{G}_I \cap \mathcal{G}_{WR} \cap \mathcal{G}_{BM}$ holds for any tree T with branching number greater than one (see any of [36, 40, 37] for the definition of branching number), but the answer is no. A counterexample arises e.g. by replacing G' by an infinite binary tree in the proof of Lemma 4.6.

Proof of Theorem 7.2: Let $\mathcal{G}_{good}$ be the set of graphs G that contain a subgraph H satisfying (i) and (ii) in Theorem 7.1. We are done if we can show that

$$\mathbf{P}(T \in \mathcal{G}_{good} \,|\, T \text{ is infinite}) = 1\,.$$

A tree may or may not be an element of $\mathcal{G}_{good}$, and the property of not being an element of $\mathcal{G}_{good}$ is *inherited*, in the sense of Lyons and Peres [37]. By [37, Chapter 3, Prop. 6], we therefore have that

$$\mathbf{P}(T \in \mathcal{G}_{good} \,|\, T \text{ is infinite}) \in \{0, 1\}\,,$$

so it suffices to show that

$$\mathbf{P}(T \in \mathcal{G}_{good}) > 0\,. \tag{21}$$

Since $\sum_{k=0}^{\infty} kp_k > 1$, we can find an $n < \infty$ such that $\sum_{k=0}^{n} kp_k > 1$. Define $\{\tilde{p}_k\}_{k=0}^{\infty}$ by setting

$$\tilde{p}_k = \begin{cases} p_0 + \displaystyle\sum_{j=n+1}^{\infty} p_j & \text{for } k = 0 \\ p_k & \text{for } k = 1, \ldots, n \\ 0 & \text{for } k > n\,. \end{cases}$$

Define a subgraph $\tilde{T}$ of T by deleting all vertices with at least n siblings, and all their ancestors. Also define the subgraph $\tilde{T}^*$ by further deleting all vertices that had at least $n+1$ children in T. A moment's reflection reveals that the distribution of $\tilde{T}$ is that of a Galton–Watson tree with offspring distribution $\{\tilde{p}_k\}_{k=0}^{\infty}$. Hence $\mathbf{P}(\tilde{T}$ is infinite$) > 0$, and furthermore $p_c^{site}(\tilde{T}) < 1$ a.s. on the event that $\tilde{T}$ is infinite (this is a general property of supercritical Galton–Watson trees; see e.g. [36, 37]). Next, note that $\tilde{T}$ only differs from $\tilde{T}^*$ by having a (possibly infinite) number of dead ends. These dead ends influence neither the (in)finiteness of the tree, nor the site percolation critical value. Hence,

$$\mathbf{P}(\tilde{T}^* \text{ is infinite and has } p_c^{site}(\tilde{T}^*) < 1) > 0\,.$$

Note also that

$$\sup_{v \in \tilde{T}^*} \deg_T(v) \leq n+1$$

by construction. In summary, we have positive probability for the event that $\tilde{T}^*$ satisfies (i) and (ii) in Theorem 7.1, and (21) follows. □

7.2 Poisson–Voronoi tessellations

Here we consider a random graph related to Poisson–Voronoi tessellations in $\mathbf{R}^d$ with $d \geq 2$. We restrict the discussion to a quick definition and our main result. An extensive discussion of Poisson–Voronoi tessellations can be found in [38]; see also [4] for a treatment of percolation on such tessellations.

Let $X_1, X_2, \ldots$ be the points of a homogeneous Poisson process in $\mathbf{R}^d$ with intensity $\lambda > 0$. These points will be the nuclei in a Voronoi tessellation of $\mathbf{R}^d$. More precisely: Each X_i is associated with a tile τ_i defined as

$$\tau_i = \{x \in \mathbf{R}^d : |x - X_i| < |x - X_j| \text{ for all } j \neq i\}$$

where $|\cdot|$ denotes Euclidean distance. Boundaries between different tiles are flat $(d-1)$-dimensional surfaces. We let G be the random graph with vertex set $V = \{X_1, X_2, \ldots\}$ and edge set E consisting of the pairs of points (nuclei) in V whose tiles share such a boundary. Equivalently, we may define

$$E = \{\{X_i, X_j\} : \exists x \in \mathbf{R}^d \text{ such that } |x - X_i| = |x - X_j| < |x - X_k| \text{ for all } k \notin \{i, j\}\}\,.$$

We remark that for $d = 2$, a straight-line representation of this graph would form a so-called Delaunay triangulation; see [38]. For any $d \geq 2$, it is easy too see that G a.s. has unbounded degree.

Theorem 7.3 *If G is the random graph constructed as above from a homogeneous Poisson process in $\mathbf{R}^d$, $d \geq 2$, with intensity $\lambda > 0$, then*

$$\mathbf{P}(G \in \mathcal{G}_{BP} \cap \mathcal{G}_{SP} \cap \mathcal{G}_I \cap \mathcal{G}_{WR} \cap \mathcal{G}_{BM}) = 1\,. \tag{22}$$

Proof: For simplicity and concreteness, we give the proof for $d = 2$ only; it will be evident how to generalize it to higher dimensions. The key ingredient (besides Theorem 7.1) of the proof is a simple renormalization argument, similar to one used by Häggström and Meester [24] in a different context. Note first that by scaling, the probability in (22) is independent of the choice of λ, so we are free to choose $\lambda > 0$ as we wish. Set $\varepsilon = \frac{1-p_c(G')}{3}$, where G' is the square lattice considered in the proof of Lemma 4.6, and note that since $p_c(G') < 1$, we have $\varepsilon > 0$. Pick the Poisson intensity λ large enough so that the probability of seeing no nucleus in the square $[0, \frac{1}{9}]^2$ satisfies

$$\mathbf{P}\left(V \cap \left[0, \frac{1}{9}\right]^2 = \emptyset\right) \leq \frac{\varepsilon}{162} \tag{23}$$

and then pick M large enough so that

$$\mathbf{P}\left(\#\left\{V \cap \left[0, \frac{1}{9}\right]^2\right\} > M\right) \leq \frac{\varepsilon}{162}\,. \tag{24}$$

Let $A_{0,0}$ be the event that for $i, j = 0, \ldots, 8$ we have at least one and at most M nuclei in the square $[\frac{i}{9}, \frac{i+1}{9}] \times [\frac{j}{9}, \frac{j+1}{9}]$. Note that $A_{0,0}$ depends only on the nuclei in the unit square $[0,1]^2$. For $k, l \in \mathbf{Z}$, let $A_{k,l}$ be the obvious analogous event concerning the square $[k, k+1] \times [l, l+1]$. The events $\{A_{k,l}\}_{k,l \in \mathbf{Z}}$ are clearly i.i.d., and by (23) and (24) we have

$$\mathbf{P}(A_{k,l}) \geq 1 - 81\frac{\varepsilon}{162} - 81\frac{\varepsilon}{162} = 1 - \varepsilon\,.$$

Next, suppose that the event $A_{0,0}$ happens. By simple geometric considerations, we see that the nucleus of any Voronoi tile intersecting $[\frac{4}{9}, \frac{5}{9}]^2$ must then be contained in $[\frac{2}{9}, \frac{7}{9}]^2$, and furthermore that the nucleus of any neighboring Voronoi tile must be contained in $[0,1]^2$. Hence, a nucleus v of a Voronoi tile intersecting $[\frac{4}{9}, \frac{5}{9}]^2$ has $\deg_G(v) \leq 81M$. Similarly, if the events $A_{k,l}$ and $A_{k+1,l}$ occur, then all nuclei v whose Voronoi tiles intersect $[k + \frac{4}{9}, k+1+\frac{5}{9}] \times [l + \frac{4}{9}, l + \frac{5}{9}]$ have $\deg_G(v) \leq 81M$.

Now define the subgraph $H = (V_H, E_H)$ of G by letting V_H be as follows. For all $k, l \in \mathbf{Z}$, whenever the two events $A_{k,l}$ and $A_{k+1,l}$ occur, we include all nuclei v whose Voronoi tiles intersect $[k + \frac{4}{9}, k+1+\frac{5}{9}] \times [l + \frac{4}{9}, l + \frac{5}{9}]$. Similarly, whenever the two events $A_{k,l}$ and $A_{k,l+1}$ occur, we include all nuclei v whose Voronoi tiles intersect $[k + \frac{4}{9}, k + \frac{5}{9}] \times [l + \frac{4}{9}, l+1+\frac{5}{9}]$. No other nuclei are in V_H.

It follows from the construction that $\sup_{v \in V_H} \deg_G(v) \leq 81M$. The set of all $(k,l) \in \mathbf{Z}^2$ for which $A_{k,l}$ happens can be seen as a site percolation process on G', and by the choice of ε we have that this percolation process is supercritical. It follows that H contains an infinite connected component with probability 1. By Theorem 7.1, we are done if we can show that $p_c^{site}(H) < 1$ with probability 1. Set $p = 1 - \frac{\varepsilon}{81M}$, and do site percolation on G with parameter p. Let $B_{k,l}$ be the event that all nuclei in $[k, k+1] \times [l, l+1]$ are open, and also define $C_{k,l} = A_{k,l} \cap B_{k,l}$. We see that the events $\{C_{k,l}\}_{k,l \in \mathbf{Z}}$ are i.i.d., with each $C_{k,l}$ having probability

$$\begin{aligned} \mathbf{P}(C_{k,l}) &= \mathbf{P}(A_{k,l})\mathbf{P}(B_{k,l} \mid A_{k,l}) \\ &\geq (1-\varepsilon)p^{81M} \\ &> 1 - 2\varepsilon \;>\; p_c(G')\,. \end{aligned}$$

Hence the set of $(k,l) \in \mathbf{Z}^2$ for which $C_{k,l}$ happens, contains an infinite cluster with probability 1, when viewed as a site percolation process on G'. A moment's thought reveals that the existence of such an infinite cluster implies the existence of an infinite cluster in the site percolation process on G restricted to H. Hence $p_c^{site}(H) \leq 1 - \frac{\varepsilon}{81M}$ with probability 1, and we are done. □

8 Other models

There are of course many other Markov random field models, besides those discussed in the previous sections, for which it is of interest to ask what kinds of extensions of Theorems 1.1 and 1.2 hold. In this section we will make some brief comments about such models.

Potts model. The Potts model (see e.g. [17, 1, 23]) is a much-studied generalization of the Ising model, where each vertex can be in one of q different states (taking q yields the Ising model), and the exponent in (1) is rewritten as 2β times the number of nearest neighbor pairs with different spins. The FK representation of the Potts model arises by replacing the random-cluster measure $\phi_I^{n,p,2}$ by $\phi_I^{n,p,q}$, and everything we did for the Ising model goes through for the Potts model, to show that for any $q \geq 2$, phase transition in the q-state Potts model on $G \in \mathcal{G}$ is equivalent to having $p_c^{bond}(G) < 1$. See e.g. the arguments in [1] and [23].

Ising model with external field. Another important generalization of the Ising model setup in Section 2.4 arises by introducing an external field parameter $h \in \mathbf{R}$ and an extra term $\beta h \sum_x \xi(x)$ in the exponent in (1). Of course, taking $h = 0$ takes us back to the setup in Section 2.4. One may ask for which graphs there exist (β, h) with $h \neq 0$ giving multiple Gibbs measures; let us write $\mathcal{G}_{I,h}$ for the class of such graphs. A result of Schonmann and Tanaka [42] shows that uniqueness of Gibbs measures for $h = 0$ implies uniqueness for all $h \neq 0$, so that $\mathcal{G}_{I,h} \subseteq \mathcal{G}_I$. This inclusion is strict, also for bounded degree graphs, as examplified e.g. by the usual $\mathbf{Z}^d$ lattice with $d \geq 2$. Jonasson and Steif [30] make interesting progress towards characterizing $\mathcal{G}_{I,h}$.

Multitype Widom–Rowlinson model. Similarly to the Potts generalization of the Ising model, the Widom–Rowlinson model has been extended to a model with state space $\{0, \ldots, q\}$, where $1, \ldots, q$ may be thought of as q particle types with hard-core interaction between different types, and 0's are empty locations. A random-cluster representation arises by replacing 2 by q in the site-random-cluster model. An interesting difference between this extended Widom–Rowlinson model and the other models considered so far is that the analogue of (i) in Proposition 2.3 does *not* imply the corresponding analogues of (ii)–(iv); this is due to the existence of a "staggered" regime of the parameter space for the q-type Widom–Rowlinson model on $\mathbf{Z}^d$ in which nonuniqueness of Gibbs measures corresponds, not to a breakdown of the particle symmetry, but rather to a breakdown of an even-odd lattice symmetry; see e.g. [41] and [14]. This is similar to what happens in the hard-core model discussed below.

Nevertheless, for bounded or unbounded degree graphs, the disagreement percolation condition of van den Berg [5] can be used to show that $p_c^{site}(G) = 1$ implies absence of phase transition for the q-type Widom–Rowlinson model. The random-cluster representation may then be exploited to show that for bounded degree graphs, phase transition in the q-type Widom–Rowlinson model is equivalent to the five other properties in Theorem 1.1.

Multitype beach model. A q-type variant of the beach model, analogous to the Potts model and the q-type Widom–Rowlinson model, was introduced and studied in [10]. We get a random-cluster representation of the extended model by replacing 2 by q in the beach-random-cluster model. Using this, and following the arguments in Sections 3.4, 4 and 6, we get the following for any $q \geq 3$: First, for bounded degree graphs, phase transition in the q-type beach model is equivalent to the other properties in Theorem 1.1. Second, for unbounded degree graphs, phase transition in the q-type beach model implies phase transition in the usual ($q = 2$) beach model.

Hard-core model. The hard-core lattice gas model (hard-core model, for short) with activity parameter $a > 0$, can informally be described as letting all vertices independently take value 0 or 1 with respective probabilities $\frac{1}{a+1}$ and $\frac{a}{a+1}$, and then conditioning on the event that no two 1's appear on adjacent vertices. The most famous example of nonuniqueness of Gibbs measures in the hard-core model is the $\mathbf{Z}^d$ lattice for $d \geq 2$. In this case, the nonuniqueness manifests itself as a breaking of the odd-even symmetry of the lattice. Results in [7] and [22] suggest that the phase transtion phenomenon is rather non-robust under modifications of the graph structure, in the sense that even a relatively minor perturbation of this lattice symmetry will be enough to remove the phase transition. It also follows from a result in [7] that $p_c^{site}(G) < 1$ is a necessary condition for the hard-core model on G to have phase transition; this is true for bounded as well as unbounded degree graphs. An example of a bounded degree graph with $p_c^{site} < 1$ but no phase transition for the hard-core model can be found in [8].

Rotor model. Let us finally mention a model with continuous state space: the rotor model. Here the vertices take values in the $[0, 2\pi]$, and configurations appear with densities with respect to Lebesgue measure that are proportional to $\exp(\beta \sum \cos(\xi(u) - \xi(v))$ where $\beta > 0$ is the inverse temperature parameter, and the sum is as usual over nearest neighbor pairs. Phase transition in this model appears to be less common than e.g. in the Ising model. For bounded degree graphs, it is strongly believed that transience of simple random walk is necessary for phase transition in the rotor model; see e.g. [11]. Furthermore, still assuming bounded degree, the disagreement percolation technique of van den Berg and Maes [6] can be exploited to show that $p_c^{site} < 1$ is another necessary condition. Regarding sufficient conditions, one may ask whether $p_c^{site} < 1$ and transience of simple random walk together form a sufficient condition, but this is probably not the case; the graph considered in the final section of [25] is almost certainly a counterexample. Y. Peres has conjectured that a somewhat stronger property is necessary and sufficient, namely that there exists a $p < 1$ such that bond percolation with parameter p produces an infinite cluster on which simple random walk is transient; see [39].

Acknowledgement. I thank Johan Jonasson, Russ Lyons, Roberto Schonmann and Oded Schramm for helpful comments and corrections.

References

[1] 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.

[2] Benjamini, I., Lyons, R., Peres, Y. and Schramm, O. (1998) Critical percolation on any nonamenable group has no infinite clusters, *Ann. Probab.*, to appear.

[3] Benjamini, I. and Schramm, O. (1996) Percolation beyond $\mathbf{Z}^d$, many questions and a few answers, *Electr. Commun. Probab.* **1**, 71–82.

[4] Benjamini, I. and Schramm, O. (1998) Conformal invariance of Voronoi percolation, *Commun. Math. Phys.* **197**, 75–107.

[5] van den Berg, J. (1993) A uniqueness condition for Gibbs measures, with application to the 2–dimensional Ising antiferromagnet, *Commun. Math. Phys.* **152**, 161–166.

[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] Brightwell, G.R., Häggström, O. and Winkler, P. (1998) Nonmonotonic behavior in hard-core and Widom–Rowlinson models, *J. Statist. Phys.*, to appear.

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

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

[11] Cassi, D. (1992) Phase transition and random walks on graphs: a generalization of the Marmin–Wagner theorem to disordered lattices, fractals, and other discrete structures, *Phys. Rev. Lett.* **68**, 3631–3634.

[12] 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.

[13] Chayes, L. (1996) Percolation and ferromagnetism on $\mathbf{Z}^2$: the q-state Potts cases, *Stoch. Proc. Appl.* **65**, 209–216.

[14] Chayes, L., Kotecký, R. and Shlosman, S.B. (1997) Staggered phases in diluted systems with continuous spins, *Commun. Math. Phys.* **189**, 631–640.

[15] Coniglio, A., Nappi, C.R., Peruggi, F. and Russo, L. (1976) Percolation and phase transitions in the Ising model, *Commun. Math. Phys.* **51**, 315–323.

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

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

[18] Georgii, H.-O., Häggström, O. and Maes, C. (1998) The random geometry of equilibrium phases, preprint.

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

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

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

[22] Häggström, O. (1997) Ergodicity of the hard core model on $\mathbf{Z}^2$ with parity-dependent activities, *Ark. Mat.* **35**, 171–184.

[23] Häggström, O. (1998) Random-cluster representations in the study of phase transitions, *Markov Proc. Relat. Fields* **4**, 275–321.

[24] Häggström, O. and Meester, R. (1996) Nearest neighbor and hard sphere models in continuum percolation, *Rand. Struct. Algorithms* **9**, 295–315.

[25] Häggström, O. and Mossel, E. (1998) Nearest-neighbor walks with low predictability profile and percolation in $2+\varepsilon$ dimensions, *Ann. Probab.* **26**, 1212–1231.

[26] Häggström, O., Peres, Y. and Schonmann, R.H. (1998) Percolation on transitive graphs as a coalescent process: relentless merging followed by simultaneous uniqueness, *Perplexing Probability Problems: Papers in Honor of Harry Kesten* (M. Bramson and R. Durrett, eds), Birkhäuser, Boston, to appear.

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

[28] Hammersley, J.M. (1961) Comparison of atom and bond percolation, *J. Math. Phys.* **2**, 728–733.

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

[30] Jonasson, J. and Steif, J. (1998) Amenability and phase transition in the Ising model, *J. Theor. Probab.*, to appear.

[31] Kindermann, R. and Snell J.L. (1980) *Markov Random Fields and their Applications*, American Mathematical Society, Providence, Rhode Island.

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

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

[34] Liggett, T.M., Schonmann, R.H. and Stacey, A.M. (1997) Domination by product measures, *Ann. Probab.* **25**, 71–95.

[35] Lyons, R. (1989) The Ising model and percolation on trees and tree-like graphs, *Commun. Math. Phys.* **125**, 337–353.

[36] Lyons, R. (1990) Random walks and percolation on trees, *Ann. Probab.* **18**, 931–958.

[37] Lyons, R. and Peres, Y. (1998) *Probability on Trees and Networks*, book in preparation, draft available at `http://php.indiana.edu/~rdlyons`

[38] Møller, J. (1994) *Lectures on Random Voronoi Tessellations*, Springer, New York.

[39] Pemantle, R. and Steif, J. (1998) Robust phase transition for spherical and other models on general trees, *Ann. Probab.*, to appear.

[40] Peres, Y. (1998) Probability on trees: an introductory climb, in *École d'été de probabilités de Saint-Flour 1997*, Springer, New York, to appear.

[41] Runnels, L.K. and Lebowitz, J.L. (1974) Phase transitions of a multicomponent Widom–Rowlinson model, *J. Math. Phys.* **15**, 1712–1717.

[42] Schonmann, R.H. and Tanaka, N. (1998) Lack of monotonicity in ferromagnetic Ising model phase diagrams, *Ann. Appl. Probab.* **8**, 234–245.

[43] Wheeler, J.C. and Widom, B. (1970) Phase equilibrium and critical behavior in a two-component Bethe-lattice gas or three-component Bethe-lattice solution, *J. Chem. Phys.* **52**, 5334–5343.

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