The random geometry of equilibrium phases

Hans-Otto Georgii, Mathematisches Institut,
Ludwig-Maximilians-Universität, D-80333 München,
Germany.
Email: georgii@rz.mathematik.uni-muenchen.de

Olle Häggström, Department of Mathematics,
Chalmers University of Technology, S-412 96 Göteborg,
Sweden.
Email: olleh@math.chalmers.se

Christian Maes, Instituut voor Theoretische Fysica,
K.U.Leuven, B-3001 Leuven,
Belgium.
Email: Christian.Maes@fys.kuleuven.ac.be

Contents

1 Introduction

Equilibrium statistical mechanics intends to describe and explain the macroscopic behavior of systems in thermal equilibrium in terms of the microscopic interaction between their great many constituents. As a typical example, let us take some ferromagnetic material like iron; the constituents are then the spins of elementary magnets at the sites of some crystal lattice. Or we may think of a lattice approximation to a real gas, in which case the constituents are the particle numbers in the elementary cells of any partition of space. The central object is the Hamiltonian describing the interaction between these constituents. This interaction determines the relative energies between configurations that differ only microscopically. The equilibrium states with respect to the given interaction are described by the associated Gibbs measures. These are probability measures on the space of configurations which have prescribed conditional probabilities with respect to fixed configurations outside of finite regions. These conditional probabilities are given by the Boltzmann factor, the exponential of the inverse temperature times the relative energy. This allows one to compute, at least in principle, equilibrium expectations and spatial correlation functions following the standard Gibbs formalism. Most important are the so called extremal Gibbs measures since they describe the possible macrostates of our physical system. In such a state, macroscopic observables do not fluctuate while the correlation between local observations made far apart from each other decays to zero.

Since the early days of statistical mechanics, geometric notions have played a role in elucidating certain aspects of the theory. This has taken many different forms. Arguably, the thermodynamic formalism, as first developed by Gibbs, already admits some geometric interpretations primarily related to convexity. For example, entropy is a concave function of the specific energy, the pressure is convex as a function of the interaction potential, the Legendre–Fenchel transformation relates various fundamental thermodynamic quantities to each other, and the set of Gibbs measures for an interaction is a simplex with vertices corresponding to the physically realized macrostates, the equilibrium phases.

Here, however, we will not be concerned with this kind of convex geometry which is described in detail e.g. in the books by Israel [138] and Georgii [96]. Rather, the geometry considered here is a way of visualizing the structure in the typical realizations of the system's constituents. To be more specific let us consider for a moment the case of the standard ferromagnetic Ising model on the square lattice. At each site we have a spin variable taking only two possible values, $+1$ and -1. The interaction is nearest-neighbor and tends to align neighboring spins in the same direction. By the ingenious arguments first formulated in 1936 by Peierls [187] (see also [63, 213, 96]), the phase transition in this model can be understood from looking at the typical configurations of contours, i.e., the broken lines separating the domains with plus resp. minus spins. The plus phase (the positively magnetized phase) is realized by an infinite ocean of plus spins with finite islands of minus spins (which in turn may contain lakes of plus spins, and so on). On the other hand, above the Curie temperature (first computed by Onsager) there is no infinite path joining nearest neighbors with the same spin value. So, for this model the geometric picture is rather complete (as we will show later). In general, however, much less is known, and much less is true. Still, certain aspects of this geometric analysis have wide applications, at least in certain regimes of the phase diagram. These 'certain regimes' are, on the one hand, the high-temperature (or, in a lattice gas setting, low-density) regime and, to the other extreme, the low temperature behavior.

At high temperatures, all thermodynamic considerations are based on the fact that entropy dominates over energy. That is, the interaction between the constituents is not effective enough to enforce a macroscopic ordering of the system. As a result, every constituent is more or less free to behave at random, not much influenced by other constituents which are far apart. So, the system's behavior is almost like that of a free system with independent components. This means, in particular, that in the center of a large box we will typically encounter more or less the same configurations no matter what boundary conditions outside this box are imposed. That is, if we compare two independent realizations of the system in the box with different boundary conditions outside then, still at high temperatures, the difference between the boundary conditions cannot be felt by the spins in the center of the box; specifically, there should not exist any path from the boundary to the central part of the box along which the spins of the two realizations disagree. This picture is rather robust and can for example also be applied when the interaction is random; see Sections 7 and 9.

At low temperatures, or large densities (when the interaction is sufficiently strong), the picture above no longer holds. Rather, the specific characteristics of the interaction will come into play and determine the specific features of the low temperature phase. In many cases, the low temperature behavior can be described as a random perturbation of a ground state, i.e., of a fixed configuration of minimal energy. Then we can expect that at low temperatures, and sometimes even up to the critical temperature, the equilibrium phases are realized as a deterministic ground state configuration, perturbed by finite random islands on which the configuration disagrees with the ground state. This means that the ground state pattern can percolate through the space to infinity. One prominent way of confirming this picture is provided by the so called Pirogov–Sinai theory which is described in detail e.g. in [230]. In Section 8 we will discuss some other techniques of establishing the same geometric picture.

It is evident from the above that percolation theory will play an important role in this text. In fact, we will mainly be concerned with dependent percolation, but one can say that independent percolation stands as a prototype for the study of statistical equilibrium properties in geometric terms. In independent percolation, the model is extremely simple: the components are binary-valued and independent from each other. What is hard is the type of question one asks, namely the question of existence of infinite paths of 1's and their geometry. We will introduce percolation below but refer to other publications (such as the book by Grimmett [108]) for a systematic account of the theory.

Percolation will come into play here on various levels. Its concepts like clusters, open paths, connectedness etc. will be useful for describing certain geometric features of equilibrium phases, allowing characterizations of phases in percolation terms. Examples will be presented where the (thermal) phase transition goes hand in hand with a phase transition in an associated percolation process. Next, percolation techniques can be used to obtain specific information about the phase diagram of the system. For example, equilibrium correlation functions are sometimes dominated by connectivity functions in an associated percolation problem which is easier to investigate. Finally, representations in terms of percolation models will yield explicit relations between certain observables in equilibrium models and some corresponding percolation quantities. In fact, the resulting percolation models, like the random-cluster model, have some interest in their own right and will also be studied in some detail.

This text is supposed to be self-contained. Therefore we need to introduce various concepts and techniques which are well-known to some readers. On the other hand, important related issues will not be discussed when they are not explicitly needed. For more complete discussions on the introductory material we will refer the interested reader to other sources. More seriously, we will not include here a discussion of some important geometric concepts developed in the 1980's for the study of critical behavior in statistical mechanical systems, namely random walk expansions or random current representations. Fortunately, we can refer to an excellent book [77] where the interested reader will find all the relevant results and references. Important steps in this context include [5, 6, 7, 11, 40, 41] and the references therein.

Finally, to avoid misunderstanding, the random geometry in the title of this work should not be confused with stochastic geometry (or geometric probability) which, as a branch of integral geometry, provides a very interesting tool-box for the discussion of morphological characteristics of random fields appearing in statistical physics and beyond, see [145, 172, 2].

Acknowledgements: It is a pleasure to thank J. Lebowitz for suggesting (and insisting) that we should write this report. We are also grateful to L. Chayes, A.C.D. van Enter and J. Lőrinczi who looked at parts of the manuscript and made numerous suggestions, and to Y. Higuchi for discussions on Proposition 8.5.

2 Equilibrium phases

2.1 The lattice

Our object of study are physical systems with many constituents, spins or particles, which will be located at the sites of a crystal lattice $\mathcal{L}$. The standard case is when $\mathcal{L} = \mathbf{Z}^d$, the d-dimensional hypercubic lattice. In general, we shall assume that $\mathcal{L}$ is the vertex set of a countable graph. That is, $\mathcal{L}$ is an at most countable set and comes equipped with a (symmetric) adjacency relation. Namely, we write $x \sim y$ if the vertices $x, y \in \mathcal{L}$ are adjacent, and this is visualized by drawing an edge between x and y. In this case, x and y are also called neighbors, and the edge (or bond) between x and y is denoted by $\langle xy \rangle$. We write $\mathcal{B}$ for the set of all edges (bonds) in $\mathcal{L}$. A complete description of the graph is thus given by the pairs $(\mathcal{L}, \sim)$ or $(\mathcal{L}, \mathcal{B})$.

In the case $\mathcal{L} = \mathbf{Z}^d$, the edges will usually be drawn between lattice sites of distance one; hence $x \sim y$ whenever $|x - y| = 1$. Here, $|\cdot|$ stands for the sum-norm, i.e. $|x| = \sum_{i=1}^d |x_i|$ whenever $x = (x_1, \ldots, x_d) \in \mathbf{Z}^d$. This choice is natural because then $|x - y|$ coincides with the graph-theoretical (or lattice) distance, viz. the length of the shortest path (consisting of consecutive edges) connecting x and y. For convenience, we sometimes use the same notation in the case of a general graph. On $\mathbf{Z}^d$ we will occasionally distinguish between the standard metrics $d_1(x, y) = \sum_i |x_i - y_i|$, $d_2(x, y) = [\sum_{i=1}^d (x_i - y_i)^2)]^{1/2}$ and $d_\infty(x, y) = \max_i |x_i - y_i|$. Given any metric d on $\mathcal{L}$, we write $d(\Lambda, \Delta) = \inf_{x \in \Lambda, y \in \Delta} d(x, y)$ for the distance of two subsets $\Lambda, \Delta \subset \mathcal{L}$.

We will always assume that the graph $(\mathcal{L}, \sim)$ is *locally finite*, which means that each $x \in \mathcal{L}$ has only a finite number N_x of nearest neighbors. In other words, N_x is the number of edges emanating from x. N_x is also called the *degree* of the graph at x. In many cases we will even assume that $(\mathcal{L}, \sim)$ is of bounded degree, which means that $N = \sup_{x \in \mathcal{L}} N_x < \infty$. Common examples of such graphs, besides $\mathbf{Z}^d$, are the triangular lattice in two dimensions, and the regular tree $\mathbf{T}_d$ (also known as the Cayley tree or the Bethe lattice), which is defined as the (unique) infinite connected graph containing no circuits in which every vertex has exactly $d + 1$ nearest neighbors.

A region of the lattice, that is a subset $\Lambda \subset \mathcal{L}$, is called finite if its cardinality $|\Lambda|$ is finite. We write $\mathcal{E}$ for the collection of all finite regions. The complement of a region Λ will be denoted by $\Lambda^c = \mathcal{L} \setminus \Lambda$. The boundary $\partial\Lambda$ of Λ is the set of all sites (vertices) in Λ^c which are adjacent to some site of Λ.

At some occasions we will need the notion of thermodynamic (or infinite volume) limit, and we need to describe in what sense a region $\Lambda \in \mathcal{E}$ grows to the full lattice $\mathcal{L}$. For our purposes, it will in general be sufficient to take an arbitrary increasing sequence (Λ_n) with $\bigcup_{n \geq 1} \Lambda_n = \mathcal{L}$. In the case $\mathcal{L} = \mathbf{Z}^d$, we will often make the standard choice $\Lambda_n = [-n, n]^d \cap \mathbf{Z}^d$, the lattice cubes centered around the origin. As $\mathcal{E}$ is a directed set ordered by inclusion, we will occasionally also consider the limit along $\mathcal{E}$. In each of these cases we will use the notation $\Lambda \uparrow \mathcal{L}$.

2.2 Configurations

The constituents of our systems are the spins or particles at the lattice sites. So, at each site $x \in \mathcal{L}$ we have a variable $\sigma(x)$ taking values in a non-empty set S, the *state space* or *single-spin space*. In a magnetic set-up (to which we mostly adhere for simplicity), $\sigma(x)$ is interpreted as the spin of an elementary magnet at x. In a lattice gas interpretation, there is a distinguished vacuum state $0 \in S$ representing the absence of any particle, and the remaining elements correspond to the types and/or the number of the particles at x. Unless stated otherwise, we will always assume that S is finite. Elements of S will typically be denoted by $a, b, \ldots$.

A *configuration* is a function $\sigma : \mathcal{L} \to S$ which assigns to each vertex $x \in \mathcal{L}$ a spin value $\sigma(x) \in S$. In other words, a configuration σ is an element of the product space $\Omega = S^{\mathcal{L}}$. Ω is called the configuration space and its elements are in general written as $\sigma, \eta, \xi, \ldots$. (It is sometimes useful to visualize $a, b \ldots$ as colors. A configuration is then a coloring of the lattice.) A configuration σ is constant if for some $a \in S$, $\sigma(x) = a$ for all $x \in \mathcal{L}$. Two configurations σ and η are said to agree on a region $\Lambda \subset \mathcal{L}$, written as "$\sigma \equiv \eta$ on Λ", if $\sigma(x) = \eta(x)$ for all $x \in \Lambda$. Similarly, we write "$\sigma \equiv \eta$ off Λ" if $\sigma(x) = \eta(x)$ for all $x \notin \Lambda$.

We also consider configurations in regions $\Lambda \subset \mathcal{L}$. These are elements of S^Λ, again denoted by letters like $\sigma, \eta, \xi, \ldots$. Given $\sigma, \eta \in \Omega$, we write $\sigma_\Lambda \eta_{\Lambda^c}$ for the configuration $\xi \in \Omega$ with $\xi(x) = \sigma(x)$ for $x \in \Lambda$ and $\xi(x) = \eta(x)$ for $x \in \Lambda^c$. Then, obviously, $\xi \equiv \sigma$ on Λ. The cylinder sets

$$\mathcal{N}_\Lambda(\sigma) = \{\xi \in \Omega : \ \xi \equiv \sigma \text{ on } \Lambda\} ,$$

$\Lambda \in \mathcal{E}$, form a countable neighborhood basis of $\sigma \in \Omega$; they generate the product topology on Ω. Hence, two configurations are close to each other if they agree on some large finite region, and a diagonal-sequence argument shows that Ω is a compact in this topology.

We will often change a configuration $\sigma \in \Omega$ at just one site $x \in \mathcal{L}$. Changing $\sigma(x)$ into a prescribed value $a \in S$ we obtain a new configuration written $\sigma^{x,a}$. In particular, for $S = \{-1, +1\}$ we write

$$\sigma^x(y) = \begin{cases} \sigma(y) & \text{for } y \neq x \\ -\sigma(x) & \text{for } y = x \end{cases}$$

for the configuration resulting from flipping the spin at x.

We will also deal with automorphisms of the underlying lattice $(\mathcal{L}, \sim)$. Each such automorphism defines a measurable transformation of the configuration space Ω. The most interesting automorphisms are the translations of the integer lattice $\mathcal{L} = \mathbf{Z}^d$; the associated translation group acting on Ω is given by $\theta_x \sigma(y) = \sigma(x + y)$, $y \in \mathbf{Z}^d$. In particular, any constant configuration is translation invariant. Similarly, we can speak about periodic configurations which are invariant under θ_x with x in some sublattice of $\mathbf{Z}^d$.

Later on, we will also consider configurations which refer to the lattice bonds rather than the vertices. These are elements η of the product space $\{0, 1\}^{\mathcal{B}}$, and a bond $b \in \mathcal{B}$ will be called open if $\eta(b) = 1$, and otherwise closed. The above notations apply to this situation as well.

2.3 Observables

An observable is a real function on the configuration space which may be thought of as the numerical outcome of some physical measurement. Mathematically, it is a measurable real function on Ω. Here, the natural underlying σ-field of measurable events in Ω is the product σ-algebra $\mathcal{F} = (\mathcal{F}_0)^{\mathcal{L}}$, where $\mathcal{F}_0$ is the set of all subsets of S. $\mathcal{F}$ is defined as the smallest σ-algebra on Ω for which all projections $X(x) : \Omega \to S$,

$X(x)(\sigma) = \sigma(x)$ with $\sigma \in \Omega$ and $x \in \mathcal{L}$, are measurable. It coincides, in fact, with the Borel σ-algebra for the product topology on Ω.

We also consider events and observables depending only on some region $\Lambda \subset \mathcal{L}$. We let $\mathcal{F}_\Lambda$ denote the smallest sub-σ-field of $\mathcal{F}$ containing the events $\mathcal{N}_\Delta(\sigma)$ for $\sigma \in S^\Delta$ and $\Delta \in \mathcal{E}$ with $\Delta \subset \Lambda$. Equivalently, $\mathcal{F}_\Lambda$ is the σ-algebra generated by the projections $X(x)$ with $x \in \Lambda$. $\mathcal{F}_\Lambda$ is the σ-algebra of events occurring in Λ.

An event A is called *local* if it occurs in some finite region, which means that $A \in \mathcal{F}_\Lambda$ for some $\Lambda \in \mathcal{E}$. Similarly, an observable $f : \Omega \to \mathbf{R}$ is called *local* if it depends on only finitely many spins, meaning that f is measurable with respect to $\mathcal{F}_\Lambda$ for some $\Lambda \in \mathcal{E}$. More generally, an observable f is called *quasilocal* if it is (uniformly) continuous, i.e., if for all $\epsilon > 0$ there is some $\Lambda \in \mathcal{E}$ such that $|f(\sigma) - f(\xi)| < \epsilon$ whenever $\xi \equiv \sigma$ on Λ. The set $C(\Omega)$ of continuous observables is a Banach space for the supremum norm $||f|| = \sup_\sigma |f(\sigma)|$, and the local observables are dense in it.

The local events and observables should be viewed as microscopic quantities. On the other side we have the macroscopic quantities which only depend on the collective behavior of all spins, but not on the values of any finite set of spins. They are defined in terms of the *tail σ-algebra* $\mathcal{T} = \bigcap_{\Lambda \in \mathcal{E}} \mathcal{F}_{\Lambda^c}$, which is also called the σ-algebra of all events at infinity. Any tail event $A \in \mathcal{T}$ and any $\mathcal{T}$-measurable observable is called *macroscopic.*

As a final piece of notation we introduce the *indicator function* I_A of an event A; it takes the value 1 if the event occurs ($I_A(\sigma) = 1$ if $\sigma \in A$) and is zero otherwise.

2.4 Random fields

As the spins of the system are supposed to be random, we will consider suitable probability measures μ on $(\Omega, \mathcal{F})$. Each such μ is called a *random field.* Equivalently, the family $X = (X(x), x \in \mathcal{L})$ of random variables on the probability space $(\Omega, \mathcal{F}, \mu)$ which describe the spins at all sites is called a random field.

Here are some standard notations concerning probability measures. The expectation of an observable f with respect to μ is written as $\mu(f) = \int f d\mu$. The probability of an event A is $\mu(A) = \mu(I_A) = \int_A d\mu$, and we omit the set braces when A is given explicitly. For example, given any $x \in \mathcal{L}$ and $a \in S$ we write $\mu(X(x) = a)$ for $\mu(A)$ with $A = \{\sigma \in \Omega : \sigma(x) = a\}$. Covariances are abbreviated as $\mu(f; g) = \mu(fg) - \mu(f)\mu(g)$.

Whenever we need a topology on probability measures on Ω, we shall take the *weak topology.* In this (metrizable) topology, a sequence of probability measures μ_n converges to μ, denoted by $\mu_n \to \mu$, if $\mu_n(A) \to \mu(A)$ for all local events $A \in \bigcup_{\Lambda \in \mathcal{E}} \mathcal{F}_\Lambda$. This holds if and only if $\mu_n(f) \to \mu(f)$ for all local, or equivalently, all quasilocal functions f. In applications, μ_n will often be an equilibrium state in a finite box Λ_n tending to $\mathcal{L}$ as $n \to \infty$, and we are interested in whether the probabilities of events occurring in some fixed finite volume have a well-defined thermodynamic (or bulk) limit. That is, we observe what happens around the origin (via the local function f) while the boundary of the box in which we realize the equilibrium state receeds to infinity. As there are only countably many local events, one can easily see by a diagonal-sequence argument that the set of all probability measures on Ω is compact in the weak topology.

2.5 The Hamiltonian

We will be concerned with systems of *interacting* spins. As usual, the interaction is described by a Hamiltonian. As the spins are located at the sites of a graph $(\mathcal{L}, \sim)$, it is natural to consider the case of homogeneous neighbor potentials. (We will deviate from homogeneity in Section 9 when considering random interactions.) The Hamiltonian H then takes the form

$$H(\sigma) = \sum_{x \sim y} U(\sigma(x), \sigma(y)) + \sum_{x} V(\sigma(x)) \tag{1}$$

with a symmetric function $U : S \times S \to \mathbf{R} \cup \{\infty\}$, the *neighbor-interaction*, and a *self-energy* $V : S \to \mathbf{R}$. The infinite sums are formal; the summation index $x \sim y$ means that the sum extends over all bonds $\langle xy \rangle \in \mathcal{B}$ of the lattice. U thus describes the interaction between spins at neighboring sites, while V might come from the action of an external magnetic field. In a lattice gas interpretation when $S = \{0, 1\}$ (the value 1 being assigned to sites which are occupied by a particle), V corresponds to a chemical potential.

To make sense of the formal sums in (1) we compare the Hamiltonian for two different configurations $\sigma, \eta \in \Omega$ which *differ only locally* (or are "local perturbations" or "excitations" of each other), in that $\sigma \equiv \eta$ off some $\Lambda \in \mathcal{E}$. For such configurations we can define the *relative Hamiltonian*

$$H(\sigma|\eta) = \sum_{x \sim y} [U(\sigma(x), \sigma(y)) - U(\eta(x), \eta(y))] + \sum_{x} [V(\sigma(x)) - V(\eta(x))] \tag{2}$$

in which the sums now contain only finitely man non-zero terms: the first part is over those neighbor pairs $\langle xy \rangle$ for which at least one of the sites belongs to Λ, and the second part is over all $x \in \Lambda$.

2.6 Gibbs measures

Gibbs measures are random fields which describe our physical spin system when it is in *macroscopic equilibrium with respect to the given microscopic interaction* at a fixed temperature. Here, macroscopic equilibrium means that all parts of the system are in equilibrium with their exterior relative to the prescribed interaction and temperature. So it is natural to define Gibbs measures in terms of conditional probabilities.

Definition 2.1 *A probability measure μ on the configuration space Ω is called a* **Gibbs measure** *for the Hamiltonian H in (1) or (2) at inverse temperature $\beta \sim 1/T$ if for all $\Lambda \in \mathcal{E}$ and all $\sigma \in \Omega$,*

$$\mu(X \equiv \sigma \text{ on } \Lambda \,|\, X \equiv \eta \text{ off } \Lambda) = \mu^{\eta}_{\beta,\Lambda}(\sigma) \tag{3}$$

for μ-almost all $\eta \in \Omega$. In the above, $\mu^{\eta}_{\beta,\Lambda}(\sigma)$ is the **Boltzmann–Gibbs distribution** *in Λ for β and H, which is given by*

$$\mu^{\eta}_{\beta,\Lambda}(\sigma) = \frac{I_{\{\sigma \equiv \eta \text{ off } \Lambda\}}}{Z_\Lambda(\beta, \eta)} \exp[-\beta H(\sigma|\eta)] \,. \tag{4}$$

Here, $Z_\Lambda(\beta, \eta)$ is a normalization constant making $\mu^{\eta}_{\beta,\Lambda}$ a probability measure, and the constraint that σ has to coincide with η outside Λ is added because we want to realize these probability measures immediately on the infinite lattice. Note that $\mu^{\eta}_{\beta,\Lambda}$ in fact only depends on the restriction of η to Λ^c.

So, μ is a Gibbs measure if it has prescribed conditional distributions *inside* some finite set of vertices, given that the configuration is held fixed *outside*, and these conditional distributions are given by the usual Boltzmann–Gibbs formalism. This definition goes back to the work of Dobrushin [64] and Lanford and Ruelle [151] in the late 1960's, whence Gibbs measures are often called *DLR-states.* By this work, equilibrium statistical physics and the study of phase transitions made firm contact with probability theory and the study of random fields. A thermodynamic justification of this definition can be given by the variational principle, which states that (in the case $\mathcal{L} = \mathbf{Z}^d$) the translation invariant Gibbs measures are precisely those translation invariant random fields which minimize the free energy density, cf. [151, 96]. For better distinction, the Gibbs distributions $\mu^{\eta}_{\beta,\Lambda}$ are often called *finite volume* Gibbs distributions, whereas the Gibbs measures are sometimes specified as *infinite volume* Gibbs measures.

We write $\mathcal{G}(\beta H)$ for the set of all Gibbs measures with given Hamiltonian H and inverse temperature β. In the special case $U \equiv 0$ of no interaction, there is only one Gibbs measure, namely the product measure with one-site marginals $\mu(X(x) = a) = e^{-\beta V(a)} / \sum_{b \in S} e^{-\beta V(b)}$. In general, several Gibbs measures for the same interaction and temperature can coexist. This is the fundamental phenomenon of nonuniqueness of phases which is one of our main subjects; we return to this point in Section 2.7 below.

First we want to emphasize an important consequence of our assumption that the underlying interaction U involves only neighbor spins. Due to this assumption, the Gibbs distribution $\mu^{\eta}_{\beta,\Lambda}$ only depends on the restriction $\eta_{\partial\Lambda}$ of η to the boundary $\partial\Lambda$ of Λ, and this implies that each Gibbs measure $\mu \in \mathcal{G}(\beta H)$ is a *Markov random field.* By definition, this means that for each $\Lambda \in \mathcal{E}$ and $\sigma \in S^{\Lambda}$

$$\mu(X \equiv \sigma \text{ on } \Lambda \,|\, \mathcal{F}_{\Lambda^c}) = \mu(X \equiv \sigma \text{ on } \Lambda \,|\, \mathcal{F}_{\partial\Lambda}), \tag{5}$$

$\mu-$almost surely. This Markov property will be an essential tool in the geometric arguments to be discussed in this review. There is in fact an equivalence between Markov random fields and Gibbs measures for nearest neighbor potentials, see e.g. Averintsev [17], Grimmett [107] or Georgii [96].

As an aside, let us comment on the case when the interaction of spins is not nearest-neighbor but only decays sufficiently fast with their distance. The Boltzmann-Gibbs distributions in (4), and therefore also Gibbs measures, can then still be defined, but the Gibbs measures fail to possess the Markov property (5). Rather their local conditional distributions $\mu^{\eta}_{\beta,\Lambda}$ satisfy a weakening of the Markov property called *quasilocality* or *almost-Markov property*: for every $\Lambda \in \mathcal{E}$ and $A \in \mathcal{F}_\Lambda$, $\mu^{\eta}_{\beta,\Lambda}(A)$ is a continuous function of η. So, in this case, Gibbs measures have prescribed continuous versions of their local conditional probabilities. To obtain a sufficiently general definition of Gibbs measures including this and other cases, one introduces the concept of a specification $G = (G_\Lambda,\ \Lambda \in \mathcal{E})$. This is a family of probability kernels G_Λ from $(\Omega, \mathcal{F}_{\Lambda^c})$ to $(\Omega, \mathcal{F})$. $G_\Lambda(\cdot, \eta)$ stands for any distribution of spins with fixed configuration $\eta_{\Lambda^c} \in S^{\mathcal{L}\setminus\Lambda}$ outside Λ; the standard case is the Gibbs specification $G_\Lambda(\cdot, \eta) = \mu^{\eta}_{\beta,\Lambda}$. A Gibbs measure is then a probability measure μ on Ω satisfying $\mu(A \,|\, \mathcal{F}_{\Lambda^c}) = G_\Lambda(A, \cdot)$ μ-almost surely for all $\Lambda \in \mathcal{E}$ and $A \in \mathcal{F}$; this property can be expressed in a condensed form by the invariance equation $\mu G_\Lambda = \mu$. In order for this definition to make sense the specification G needs to satisfy a natural compatibility condition for pairs of volumes $\Lambda \subset \Lambda'$ expressing the fact that if the system in Λ' is in equilibrium with its exterior, then the subsystem in Λ is also in equilibrium with its own exterior. It is easy to see that the Gibbs distributions in (4) are compatible in this sense. Details and further discussion can be found in many books and articles dealing with mathematical results in equilibrium statistical mechanics, including [138, 206, 96, 195, 76, 213]. In [160], the relation between Gibbs measures and the condition of detailed balance (reversibility) in certain stochastic dynamics is explained.

Finally, we mention an alternative and constructive approach to the concept of Gibbs measures. Starting from the finite-volume Gibbs distributions $\mu^{\eta}_{\beta,\Lambda}$, one might ask what kind of limits could be obtained if η is randomly chosen and Λ increases to the whole lattice $\mathcal{L}$. (This slightly older but still important approach was suggested by Minlos [179].) To make this precise we consider the measures $\mu^{\rho}_{\beta,\Lambda} = \int \mu^{\eta}_{\beta,\Lambda}\,\rho(d\eta)$, where ρ is any probability measure on Ω describing a "stochastic boundary condition". Any such $\mu^{\rho}_{\beta,\Lambda}$ is called a (finite volume) Gibbs distribution with respect to H at inverse temperature β, and their collection is denoted by $\mathcal{G}_\Lambda(\beta H)$. The set of all (infinite volume) Gibbs measures is then equal to

$$\mathcal{G}(\beta H) = \bigcap_{\Lambda \in \mathcal{E}} \mathcal{G}_\Lambda(\beta H).$$

Equivalently, a probability measure μ on Ω is a Gibbs measure for the Hamiltonian βH if it belongs to the closed convex hull of the set of limit points of $\mu^{\eta}_{\beta,\Lambda}$ as $\Lambda \uparrow \mathcal{L}$.

One important consequence is that $\mathcal{G}(\beta H) \neq \emptyset$. This is because each $\mathcal{G}_\Lambda(\beta H)$ is obviously non-empty and compact. Equivalently, to obtain an infinite volume Gibbs measure one can fix a particular configuration η and take it as boundary condition. By compactness, we obtain an infinite volume Gibbs measure μ^{η}_{β} by taking the limit of (4) as $\Lambda \uparrow \mathcal{L}$, at least along suitable subsequences; for details see e.g. Preston [195] or Georgii [96]. We remark that in general there is no unique limiting measure μ^{η}_{β}; rather there may be several such limiting measures obtained as limits along different subsequences. Fortunately, however, this is not the case for a wide class of models, either at low temperatures (β large) when η is a ground state configuration (in the realm of the Pirogov–Sinai theory), or at high temperatures when β is small.

We conclude this subsection with a general remark. As all systems in nature are finite, one may wonder why we consider here systems with infinitely many constituents. The answer is that sharp results for bulk quantities can only be obtained when we make the idealization to an infinite system. The thermodynamic limit eliminates finite size effects (which are always present but which are not always relevant for certain phenomena) and it is only in the thermodynamic limit of infinite volume that we can get a clean and precise picture of realistic phenomena such as phase transitions or phase coexistence. This is a consequence of the general probabilistic principle of large numbers. In this sense, infinite systems serve as an idealized approximation to very large finite systems.

2.7 Phase transition and phases

As pointed out above, in general there may exist several solutions μ to the DLR-equation (3) for given U, V and β, which means that multiple Gibbs measures exist. The system can then choose between several equilibrium states. (In a dynamical theory this choice would depend on the past; but here we are in a pure equilibrium setting.) *The phenomenon of non-uniqueness therefore corresponds to a phase transition.* In fact, it is then possible to construct different Gibbs measures as infinite volume limits of Gibbs distributions with different choices of boundary conditions [91, 96]. Since any two Gibbs measures can be distinguished by a suitable local observable, a phase transition can be detected by looking at such a local observable which is then called an *order parameter*. Varying the external parameters such as temperature or an external magnetic field (which can be tuned by the experimenter via some heatbath or reservoir) one will observe different scenarios; these are collected in the so called *phase diagram* of the considered system.

As we have indicated in the introduction, the phase transition phenomenon is of central interest in equilibrium statistical mechanics. When phase transitions occur and when they do not is also one of the primary questions (although we will encounter many others) that we will try to answer with the geometric methods to be developed in subsequent sections.

If multiple Gibbs measures for a given interaction exist, the structure of the set $\mathcal{G}(\beta H)$ of all Gibbs measures becomes relevant. We only state here the most basic results; a detailed exposition can be found in [96], for example. The basic observation is that $\mathcal{G}(\beta H)$ is a convex set. Its extremal elements, the extremal Gibbs measures, have a trivial tail σ-field $\mathcal{T}$ (which means that all events in $\mathcal{T}$ have probability 0 or 1). Equivalently, all macroscopic observables are almost surely constant. In addition, the tail triviality can be characterized by an asymptotic independence (or mixing) property. On the other hand, any Gibbs measure μ can be decomposed into extremal Gibbs measures; therefore every configuration which is typical for μ is in fact typical for some extremal Gibbs measure. This shows that the extremal Gibbs measures correspond to what one can really see in nature as far as large systems in equilibrium are concerned. The extremal Gibbs measures therefore correspond to the physical macrostates, whereas non-extremal Gibbs measures only provide a limited description when the system's precise state is unknown. For all these reasons, the extremal Gibbs measures are called (equilibrium) *phases*. The central subject of this review is the geometric analysis of their typical configurations, and thereby the analysis of the phase diagram giving the variation in the number and the nature of the phases as one changes various control parameters (coupling, temperature, external fields, etc.).

Often it is natural to consider automorphisms of the graph $(\mathcal{L}, \sim)$. For example, if $\mathcal{L} = \mathbf{Z}^d$ we consider the translation group $(\theta_x)_{x\in\mathbf{Z}^d}$. A homogeneous phase is then an extremal Gibbs measure which is also translation invariant. On the other hand, we can regard the extremal points of the convex set of all translation invariant Gibbs measures. These are ergodic, which means that they cannot be decomposed into distinct translation invariant probability measures, and are trivial on the σ-algebra of all translation invariant events. However, these extremal translation invariant Gibbs measures need not be homogeneous phases; they are *only* ergodic. Yet, ergodic measures μ satisfy a law of large numbers: for any observable f and any sequence of increasing cubes Λ,

$$\lim_{\Lambda\uparrow\mathbf{Z}^d} \frac{1}{|\Lambda|} \sum_{x\in\Lambda} f \circ \theta_x = \mu(f) \quad \mu\text{-almost surely.}$$

Hence, ergodic Gibbs measures are suitable for modelling macrostates in equilibrium if one limits oneself to measuring certain bulk observables or macroscopic quantities with additivity properties. Notice, however, that there exists a certain non-uniformity in the literature concerning the nomenclature. Sometimes these ergodic Gibbs measures are called (pure) phases. It is then argued that it might happen that two phases (as defined above) for a system can by no means be macroscopically distinguished (for example if one is a translation of the other). We do not wish to enter into a detailed discussion of these points.

3 Some models

In this section we discuss briefly the phase transition behaviour of some prototypical examples of Gibbs systems. Although these examples are fairly standard and well-known to most of our readers, we need to include them here to set up the stage. They will be studied in detail in the later sections. An account of phase transition phenomena in more general lattice models can be found in many other sources, including [147, 96, 212, 213, 180].

3.1 The ferromagnetic Ising model

The Ising model was introduced in the 1920's by Wilhelm Lenz [158] and his student Ernst Ising [137] as a simple model for magnetism and, in particular, ferromagnetic phase transitions. Each site $x \in \mathcal{L}$ can take either of two spin values, $+1$ ("spin up") and -1 ("spin down"), so that the state space is equal to $S = \{-1, +1\}$. The Hamiltonian is given by (1) with $U(\sigma(x), \sigma(y)) = -\sigma(x)\,\sigma(y)$ and $V(\sigma(x)) = -h\,\sigma(x)$. The parameter $h \in \mathbf{R}$ describes an external field. The finite volume Gibbs distribution in a box Λ with external field h at inverse temperature $\beta > 0$ with boundary condition η is thus the probability measure $\mu^{\eta}_{h,\beta,\Lambda}$ on $\Omega = \{-1, +1\}^{\mathcal{L}}$ which to each $\sigma \in \Omega$ assigns probability proportional to

$$I_{\{\sigma \equiv \eta \text{ on } \Lambda^c\}} \; \exp\left[\beta\Big(\sum_{\substack{x \sim y \\ x \in \Lambda \text{ or } y \in \Lambda}} \sigma(x)\,\sigma(y) + h \sum_{x \in \Lambda} \sigma(x)\Big)\right].$$

For $\beta = 0$ ("infinite temperature") the spin variables are independent under $\mu^{\eta}_{h,\beta,\Lambda}$, but as soon as $\beta > 0$ the probability distribution starts to favour configurations with many neighbor pairs of aligned spins. This tendency becomes stronger and stronger as β increases.

In the case $h = 0$ of no external field, the model is symmetric under interchange of the spin values -1 and $+1$, so that there is an equal chance of having many pairs of plus spins or having many pairs of minus spins. This dichotomy gives rise to the following interesting behavior. Suppose that $\mathcal{L} = \mathbf{Z}^d$, $d \geq 2$. If β is sufficiently small (i.e., in the high temperature regime), the interaction is not strong enough to produce any long range order, so that the boundary conditions become irrelevant in the infinite volume limit and the Gibbs measure is uniquely determined. By ergodicity and the $\pm$ symmetry, the limiting fraction of plus spins will almost surely be $1/2$ under this unique Gibbs measure. In contrast, when β is sufficiently large (in the low temperature regime), the interaction becomes so strong that a long range order appears: the bias towards neighbor pairs of equal spin then implies that Gibbs measures prefer configurations with either a vast majority of plus spins or a vast majority of minus spins, and this preference even survives in the infinite volume limit. The system thus undergoes a phase transition which manifests itself in a non-uniqueness of Gibbs measures. Specifically, there exist two particular Gibbs measures μ^+ and μ^-, obtained as infinite volume limits with respective boundary conditions $\eta \equiv +1$ and $\eta \equiv -1$, which can be distinguished by their overall density of $+1$'s: the density is greater than $1/2$ under μ^+ and (by symmetry) less than $1/2$ under μ^-. This is the spontaneous magnetization phenomenon that Lenz and Ising were looking for but were discouraged by not finding it in one dimension. In higher dimensions, the uniqueness regime and the phase transition regime are separated by a sharp critical value β_c, as is summarized in the following classical theorem [187, 63, 65]:

Theorem 3.1 *For the ferromagnetic Ising model on the integer lattice $\mathbf{Z}^d$ of dimension $d \geq 2$ at zero external field, there exists a critical inverse temperature $\beta_c \in (0, \infty)$ (depending on d) such that for $\beta < \beta_c$ the model has a unique Gibbs measure while for $\beta > \beta_c$ there are multiple Gibbs measures.*

A stochastic-geometric proof of this result will be given in Section 6. In fact, the result (as well as its proof) holds for any graph $(\mathcal{L}, \sim)$ in place of $\mathbf{Z}^d$, except that β_c may then take the values 0 or ∞. For instance, on the one-dimensional lattice $\mathbf{Z}^1$ we have $\beta_c = \infty$, which means that there is a unique Gibbs measure for all β. For $\mathbf{Z}^2$, the critical value has been found to be $\beta_c = \frac{1}{2}\log(1+\sqrt{2})$. This calculation is a remarkable achievement which began with Onsager [186] in 1944. An account of various (algebraic and/or combinatorial) methods can be found e.g. in [220, 171]. Let us also mention the work done in 1973 by Abraham and Martin-Löf [1] relating these exact computations to the real magnetization in the appropriate Gibbs measures; it also gives the result that there is a unique Gibbs measure *at* the critical value $\beta = \beta_c$. A rigorous calculation of the critical value in higher dimensions is beyond current knowledge. It is believed that uniqueness holds at criticality in all dimensions $d \geq 2$, but so far this is only known for $d = 2$ and $d \geq 4$ [11].

The case of a nonzero external field $h \neq 0$ is less interesting, in that one finds a unique Gibbs measure for all β and d. The intuitive explanation is that for $h \neq 0$ there is no $\pm$ symmetry which could be broken; depending on the sign of h, the system is forced to prefer either $+1$'s or -1's. This comes from the fact that the magnetic field acts on the whole volume, whereas the influence of a boundary condition is of smaller order as the volume increases. In contrast, a phase transition for $h \neq 0$ does occur when $\mathbf{Z}^d$ is replaced by certain nonamenable graph structures for which the boundary of a volume is of the same order of magnitude as the volume itself (which makes them physically perhaps less realistic) – an example is the regular tree $\mathbf{T}_d$ with $d \geq 2$; we refer to [215, 96, 140]. A phase transition can also occur for a non-zero external field for the Ising model on a half-space where it is due to the so-called Basuev phenomenon [19, 20].

Because of the simplicity of its model assumptions, the standard Ising model has inspired a variety of techniques for analyzing interacting random fields. Its ferromagnetic structure suggests various monotonicity properties which can be checked by the coupling methods to be described in Section 4, and the assumption of neigbor interaction implies the spatial Markov property (5) which plays a fundamental role in the geometric analysis of typical configurations. Many techniques which were developed on this testing ground turned out to be fruitful also in more general cases.

3.2 The antiferromagnetic Ising model

The Ising antiferromagnet is defined quite similarly to the ferromagnetic case, except that $U(\sigma(x), \sigma(y))$ is taken to be $+\sigma(x)\,\sigma(y)$ rather than $-\sigma(x)\,\sigma(y)$. This means that neighboring sites now prefer to take *opposite* spins.

Suppose that $h = 0$ and that the underlying graph is bipartite; this means that $\mathcal{L}$ can be partitioned into two sets $\mathcal{L}_{even}$ and $\mathcal{L}_{odd}$ such that sites in $\mathcal{L}_{even}$ only have edges to sites in $\mathcal{L}_{odd}$, and vice versa. Clearly, $\mathbf{Z}^d$ is an example of a bipartite graph. In this situation, we can reduce the antiferromagnetic Ising model to the ferromagnetic case by a simple spin-flipping trick: The bijection $\sigma \leftrightarrow \tilde{\sigma}$ of Ω defined by

$$\tilde{\sigma}(x) = \begin{cases} \sigma(x) & \text{if } x \in \mathcal{L}_{even}\,, \\ -\sigma(x) & \text{if } x \in \mathcal{L}_{odd} \end{cases} \tag{6}$$

maps any Gibbs measure for the antiferromagnetic Ising model to a Gibbs measure for the ferromagnetic Ising model with the same parameters, and vice versa. As a consequence, a phase transition in the antiferromagnetic model is equivalent to a phase transition in the ferromagnetic model with the same parameters. Hence, Theorem 3.1 immediately carries over to the antiferromagnetic case.

The model becomes more interesting (or, at least, more genuinely antiferromagnetic) if either $h \neq 0$ or the graph is taken to be non-bipartite. Suppose first that $h \neq 0$ but still $\mathcal{L} = \mathbf{Z}^d$. If $|h|$ is small and β sufficiently large, we have the same picture as in the case $h = 0$: there exist two distinct phases, one having a majority of plus spins on the even sublattice and a majority of minus spins on the odd sublattice, the other one having a majority of plus spins on the odd sublattice and a majority of minus spins on the even sublattice. We will show this in Section 8.5, Example 8.17; see also [65, 96]. Note that this phase transition is somewhat different in flavor compared to that in the ferromagnetic Ising model: whereas in the Ising ferromagnet the phase transition produces a breaking of a state-space symmetry, the phase transition in the Ising antiferromagnet instead breaks the translation symmetry between the sublattices $\mathcal{L}_{even}$ and $\mathcal{L}_{odd}$.

To see what happens in the case of a non-bipartite graph we consider the triangular lattice which can be obtained by taking the usual square lattice $\mathbf{Z}^2$ and adding an edge between each vertex x and its north-east neighbor $x + (1,1)$. In this case, one expects uniqueness when $h = 0$, and existence of three distinct phases when $|h| \neq 0$ is small and β is large. Phase transitions in these models were studied in [65, 129], for example.

3.3 The Potts model

A natural generalization of the ferromagnetic Ising model is the (ferromagnetic) Potts model [194], in which each spin may take $q \geq 2$ (rather than only two) different values. The state space is then $S = \{1, 2, \ldots, q\}$, and the pair interaction is given by

$$U(\sigma(x), \sigma(y)) = 1 - 2I_{\{\sigma(x)=\sigma(y)\}}\,.$$

We confine ourselves to the case of zero external field, so that $V(\sigma(x)) \equiv 0$. Taking $q = 2$ and identifying the state space $\{1, 2\}$ with $\{-1, +1\}$ we reobtain the ferromagnetic Ising model with zero external field. Just as in the latter case, the Potts interaction favours configurations where many neighbor pairs agree, and Theorem 3.1 can be extended to the Potts model as follows, as we will show in Section 6.3.

Theorem 3.2 *For the q-state Potts model on $\mathbf{Z}^d$, $d \geq 2$, there exists a critical inverse temperature $\beta_c \in (0, \infty)$ (depending on d and q) such that for $\beta < \beta_c$ the model has a unique Gibbs measure while for $\beta > \beta_c$ there exist q mutually singular Gibbs measures.*

In the same way as Theorem 3.1, this theorem also holds on general graphs provided we allow β_c to be 0 or ∞. The Potts model differs from the Ising model in that, for q large enough, there are multiple Gibbs measures also at the critical value $\beta = \beta_c$, as demonstrated by Kotecký and Shlosman [148]; an outline of a proof will be given in Example 8.21. The Onsager critical value for the two-dimensional Ising model is believed to extend to the Potts model on $\mathbf{Z}^2$ through the formula $\beta_c(q) = \frac{1}{2}\log(1+\sqrt{q})$; see Welsh [221], for example. This has so far only been established when q is sufficiently large [149].

3.4 The hard-core lattice gas model

The hard-core lattice gas model (or hard-core model for short) describes a gas of particles which can only sit on the lattice sites but are so large that adjacent sites cannot be occupied simultaneously. The state space is $S = \{0, 1\}$, the pair interaction

$$U(\sigma(x), \sigma(y)) = \begin{cases} \infty & \text{if } \sigma(x) = \sigma(y) = 1 \\ 0 & \text{otherwise,} \end{cases}$$

describes the hard core of the particles, and the chemical potential is $V(\sigma(x)) = -(\log\lambda)\ \sigma(x)$. Here $\lambda > 0$ is the so-called activity parameter. The hard-core model shows some similarities to the Ising antiferromagnet in an external field and can, in fact, obtained from it by a limiting procedure ($\beta \to \infty$, $h \to 2d$, $\beta(h - 2d) = \text{const}$, [68]). Since U is either 0 or ∞, the inverse temperature β is irrelevant and will thus be fixed as 1, and we can vary only the parameter λ. Finite volume Gibbs distributions can then be thought of as first letting all spins be independent, taking values 0 and 1 with respective probabilities $\frac{1}{1+\lambda}$ and $\frac{\lambda}{1+\lambda}$, and then conditioning on the event that no two 1's sit next to each other anywhere on the lattice.

The phase transition behavior of the hard-core model on $\mathbf{Z}^d$, $d \geq 2$, is as follows. For λ sufficiently close to 0, the particles are spread out rather sparsely on the lattice, and we get a unique Gibbs measure, just as in the Ising antiferromagnet at high temperatures. When λ increases, the particle density also increases, and the system finally starts looking for optimal packings of particles. There are two such optimal packings, one where all sites in $\mathcal{L}_{even}$ are occupied and those in $\mathcal{L}_{odd}$ are empty, and one vice versa; we denote these configurations by η_{even} and η_{odd}, respectively. (These chessboard configurations look similar to those favoured in the Ising antiferromagnet.) For sufficiently large λ, the infinite volume construction of Gibbs measures with these two choices of boundary condition produces different Gibbs measures, so we get a phase transition [65].

Theorem 3.3 *For the hard-core model on $\mathbf{Z}^d$, $d \geq 2$, there exist two constants $0 < \lambda_c \leq \lambda'_c < \infty$ (depending on d) such that for $\lambda < \lambda_c$ the model has a unique Gibbs measure while for $\lambda > \lambda'_c$ there are multiple Gibbs measures.*

This result will be proved in Section 6.7. From a computer-assisted proof [201] we know that $\lambda_c \geq 1.50762$. It is widely believed that one should be able to take $\lambda_c = \lambda'_c$ in this result, which would mean that the occurrence of phase transition is increasing in λ. Such a result, however, would (unlike Theorems 3.1 and 3.2) *not* extend to arbitrary graph structures; some counterexamples were recently provided by Brightwell, Häggström and Winkler [38].

The hard-core model analogue of introducing an external field in the Ising model on $\mathbf{Z}^d$ is obtained by replacing the single activity parameter λ by two different activities λ_{even} and λ_{odd}, one for sites in $\mathcal{L}_{even}$ and the other for sites in $\mathcal{L}_{odd}$. By analogy with the Ising model, one would expect to have a unique Gibbs measure as soon as $\lambda_{even} \neq \lambda_{odd}$; this was conjectured by Van den Berg and Steif [27] and proved for the case $d = 2$ by Häggström [117].

3.5 The Widom–Rowlinson lattice model

The Widom–Rowlinson model is another lattice gas model, where this time there are two types of particles, and two particles are allowed to sit on neighboring sites only if they are of the same type. Actually, Widom and Rowlinson [223] originally introduced it as a continuum model of particles living in $\mathbf{R}^d$; see Section 10.2 below. The lattice variant described here was first studied by Lebowitz and Gallavotti [153]. The state space is $S = \{-1, 0, +1\}$, where -1 and $+1$ are the two particle types, and 0's correspond to empty sites. The pair interaction is given by

$$U(\sigma(x), \sigma(y)) = \begin{cases} \infty & \text{if } \sigma(x)\,\sigma(y) = -1\,, \\ 0 & \text{otherwise,} \end{cases}$$

and the chemical potential by

$$V(\sigma(x)) = \begin{cases} -\log\lambda_- & \text{if} \quad \sigma(x) = -1\,, \\ 0 & \text{if} \quad \sigma(x) = 0\,, \\ -\log\lambda_+ & \text{if} \quad \sigma(x) = +1\,. \end{cases}$$

Here λ_-, $\lambda_+ > 0$ are the activity parameters for the two particle types -1 and $+1$. As in the hard-core model, we fix the inverse tempreature $\beta = 1$ and only vary the activity parameters. Gibbs measures can then be thought of as first picking all spins independently, taking values -1, 0 or $+1$ with probabilities proportional to λ_-, 1, and λ_+, and then conditioning on the event that no two particles of different type sit next to each other in the lattice. We are mainly interested in the symmetric case $\lambda_- = \lambda_+ = \lambda$, where the phase transition behavior on $\mathbf{Z}^d$, $d \geq 2$ is similar to the Ising model: For λ small, there is a unique Gibbs measure in which the overall density of plus-particles is almost surely equal to that of the minus-particles. For λ sufficiently large, the system wants to pack the particles so densely that the ± 1 symmetry is broken. As for the Ising model, one can construct two particular Gibbs measures μ^+ and μ^- using boundary conditions $\eta \equiv +1$ and $\eta \equiv -1$; for small λ we get $\mu_+ = \mu_-$ whereas for large λ the two measures are different (and distinguishable through the densities of the two particle types), producing a phase transition.

Theorem 3.4 *For the Widom–Rowlinson model on $\mathbf{Z}^d$, $d \geq 2$, with activities $\lambda_- = \lambda_+ = \lambda$, there exist $0 < \lambda_c \leq \lambda'_c < \infty$ (depending on d) such that for $\lambda < \lambda_c$ the model has a unique Gibbs measure while for $\lambda > \lambda'_c$ there are multiple Gibbs measures.*

As in the hard-core model, we expect that one should be able to take $\lambda_c = \lambda'_c$, but such a monotonicity is not known. Examples of graph structures where the desired monotonicity fails can be found in [38].

We furthermore expect that the asymmetric Widom–Rowlinson model on $\mathbf{Z}^d$ with $\lambda_- \neq \lambda_+$ always has a unique Gibbs measure (similarly to the Ising model with a nonzero external field), but this also is not rigorously known.

4 Coupling and stochastic domination

Geometry alone will not be sufficient for our analysis of equilibrium phases. We also need some probabilistic tools which allow us to compare different configurations and different probability measures. So we need to include another preparatory section describing these tools and their basic applications to our setting.

Coupling is a probabilistic technique which has turned out to be immensely useful in virtually all areas of probability theory, and especially in its applications to statistical mechanics. The basic idea is to define two (or more) stochastic processes jointly on the same probability space so that they can be compared realizationwise. This direct comparison often leads to conclusions which would not be easily available by considering the processes separately. Although an independent coupling is sometimes already quite useful (as we will see in Section 7.3, for example), it is usually more efficient to introduce a dependence which relates the two processes in an efficient way. One such particularly nice relationship is that one process is pointwise smaller than the other in some partial order. This case is related to the central concept of *stochastic domination*, via Strassen's Theorem (Theorem 4.6) below. We will confine ourselves to those parts of coupling theory that are needed for our applications; a more general account can for example be found in the monograph by Lindvall [162].

4.1 The coupling inequality

In this and the next subsection of general character, $\mathcal{L}$ will be an arbitrary finite or countably infinite set. As the notation indicates, we think of the standard case that $\mathcal{L}$ is the lattice introduced in Section 2.1, but the following results will also be applied to the case when $\mathcal{L}$ is replaced by its set $\mathcal{B}$ of bonds. We consider again the product space $\Omega = S^{\mathcal{L}}$, where for the moment S is an arbitrary measurable space.

Suppose X and X' are random elements of Ω, and let μ and μ' be their respective distributions. We define the (half) *total variation distance* $\|\mu - \mu'\|$ between μ and μ' by

$$\|\mu - \mu'\| = \sup_{A \subset \Omega} |\mu(A) - \mu'(A)| \tag{7}$$

where A ranges over all measurable subsets of Ω. The coupling inequality below provides us with a convenient upper bound on this distance. To state it we first need to define what we mean by a coupling of X and X'.

Definition 4.1 *A* **coupling** P *of two* Ω*-valued random variables* X *and* X'*, or of their distributions* μ *and* μ'*, is a probability measure on* $\Omega \times \Omega$ *having marginals* μ *and* μ'*, in that for every event* $A \subset \Omega$

$$P((\xi, \xi') : \xi \in A) = \mu(A) \tag{8}$$

and

$$P((\xi, \xi') : \xi' \in A) = \mu'(A) \; . \tag{9}$$

We think of a coupling as a redefinition of the random variables X and X' on a new common probability space such that their distributions are preserved. Sometimes it will be convenient to keep the underlying probability space implicit, but in general, as in (8) and (9), we make the canonical choice, which is the product space $\Omega \times \Omega$; X and X' are then simply the projections on the two coordinate spaces. With this in mind, we write $P(X \in A)$ and $P(X' \in A)$ for the left hand sides of (8) and (9), respectively. In the same spirit, $P(X = X')$ is a short-hand for $P((\xi, \xi') : \xi = \xi')$.

Proposition 4.2 (The coupling inequality) *Let P be a coupling of two Ω-valued random variables X and X', with distributions μ and μ'. Then*

$$\|\mu - \mu'\| \leq P(X \neq X'). \tag{10}$$

Proof: For any $A \in \Omega$, we have

$$\begin{aligned} \mu(A) - \mu'(A) &= P(X \in A) - P(X' \in A) \\ &= P(X \in A, X' \notin A) - P(X \notin A, X' \in A) \\ &\leq P(X \in A, X' \notin A) \\ &\leq P(X \neq X')\,, \end{aligned}$$

whence (10) follows by symmetry. □

The next result states that there always exists some coupling which achieves equality in (10). We call such a coupling *optimal*.

Definition 4.3 *A coupling P of two Ω-valued random variables X and X', with distributions μ and μ', is said to be an* **optimal coupling** *if*

$$\|\mu - \mu'\| = P(X \neq X')\,.$$

Proposition 4.4 *For any two Ω-valued random variables X and X', there exists an optimal coupling of X and X'.*

To construct an optimal coupling, one simply puts the common mass $\mu \wedge \mu'$ of μ and μ' on the diagonal of $\Omega \times \Omega$ and adds any measure with marginals $\mu - \mu \wedge \mu'$ and $\mu' - \mu \wedge \mu'$; the simplest choice of such a measure is the appropriately scaled product measure. For details we refer to [162], where such a coupling is called the γ-coupling of μ and μ'. This construction shows, in particular, that the optimal coupling is in general not unique. Applications of optimal couplings to interacting particle systems can be found in [160, 166], for example.

4.2 Stochastic domination

Suppose now that S is a closed subset of $\mathbf{R}$, so that S is linearly ordered. The product space Ω is then equipped with a natural partial order $\preceq$ which is defined coordinatewise: For $\xi, \xi' \in \Omega$, we write $\xi \preceq \xi'$ (or $\xi' \succeq \xi$) if $\xi(x) \leq \xi'(x)$ for every $x \in \mathcal{L}$. A function $f : \Omega \to \mathbf{R}$ is said to be *increasing* (or, non-decreasing) if $f(\xi) \leq f(\xi')$ whenever $\xi \preceq \xi'$. An event A is said to be increasing if its indicator function I_A is increasing. The following standard definition of stochastic domination expresses the fact that μ' prefers larger elements of Ω than μ.

Definition 4.5 *Let μ and μ' be two probability measures on Ω. We say that μ is* **stochastically dominated** *by μ', or μ' is stochastically larger than μ, writing $\mu \preceq_{\mathcal{D}} \mu'$, if for every bounded increasing observable $f : \Omega \to \mathbf{R}$ we have $\mu(f) \leq \mu'(f)$.*

In the one-dimensional case when $|\mathcal{L}| = 1$ and $\Omega = S \subset \mathbf{R}$, this definition is equivalent to the classical requirement that $\mu([r, \infty)) \le \mu'([r, \infty))$ for all $r \in \mathbf{R}$. The following fundamental result of Strassen [216] characterizes stochastic domination in coupling terms.

Theorem 4.6 (Strassen) *For any two probability measures μ and μ' on Ω, the following statements are equivalent.*

(i) $\mu \preceq_{\mathcal{D}} \mu'$

(ii) *For all* continuous *bounded increasing functions $f : \Omega \to \mathbf{R}$, $\mu(f) \le \mu'(f)$.*

(iii) *There exists a coupling P of μ and μ' such that $P(X \preceq X') = 1$.*

Sketch of proof: While the implications (i) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (i) are trivial, the assertion (ii) $\Rightarrow$ (iii) is too deep to be explained here in detail. To start one should note that μ and μ' may be considered as measures on the compact space $[-\infty, \infty]^{\mathcal{L}}$, and one can then follow the arguments outlined in [160], pp. 72 ff., or [162]. Further discussion of Strassen's theorem can be found in [162, 141]. □

The equivalence (i) $\Leftrightarrow$ (ii) in Theorem 4.6 readily implies the following corollary.

Corollary 4.7 *The relation $\preceq_{\mathcal{D}}$ of stochastic domination is preserved under weak limits.*

Next we recall a famous sufficient condition for stochastic domination. This condition is (essentially) due to Holley [134] and refers to the finite-dimensional case when $|\mathcal{L}| < \infty$. We also assume for simplicity that $S \subset \mathbf{R}$ is finite. Hence Ω is finite. In this case, a probability measure μ on Ω is called *irreducible* if the set $\{\eta \in \Omega : \mu(\eta) > 0\}$ is connected in the sense that any element of Ω with positive μ-probability can be reached from any other via successive coordinate changes without passing through elements with zero μ-probability.

Theorem 4.8 (Holley) *Let $\mathcal{L}$ be finite, and let S be a finite subset of* $\mathbf{R}$. *Let μ and μ' be probability measures on Ω. Assume that μ' is irreducible and assigns positive probability to the maximal element of Ω (with respect to $\preceq$). Suppose further that*

$$\mu(X(x) \ge a \,|\, X = \xi \text{ off } x) \le \mu'(X(x) \ge a \,|\, X = \eta \text{ off } x) \tag{11}$$

whenever $x \in \mathcal{L}$, $a \in S$, and $\xi, \eta \in S^{\mathcal{L}\setminus\{x\}}$ are such that $\xi \preceq \eta$, $\mu(X = \xi \text{ off } x) > 0$ and $\mu'(X = \eta \text{ off } x) > 0$. Then $\mu \preceq_{\mathcal{D}} \mu'$.

Proof: Consider a Markov chain $(X_k)_{k=0}^{\infty}$ with state space Ω and transition probabilities defined as follows. At each integer time $k \ge 1$, pick a random site $x \in \mathcal{L}$ according to the uniform distribution. Let $X_k = X_{k-1}$ on $\mathcal{L} \setminus \{x\}$, and select $X_k(x)$ according to the single-site conditional distribution prescribed by μ. This is a so-called Gibbs sampler for μ, and it is immediate that if the initial configuration X_0 is chosen according to μ, then X_k has distribution μ for each k. Define a similar Markov chain $(X'_k)_{k=0}^{\infty}$ with μ replaced by μ'.

Next, 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 the product measure $\mu \times \mu'$. Then, for each k, pick a site $x \in \mathcal{L}$ at random and let U_k be an independent random variable, uniformly distributed on the interval $[0,1]$. Let $X_k(y) = X_{k-1}(y)$ and $X'_k(y) = X'_{k-1}(y)$ for each site $y \neq x$, and update the values at site x by letting

$$X_k(x) = \max\{a \in S : \mu(X(x) \geq a \,|\, X = \xi \text{ off } x) \geq U_k\}$$

and

$$X'_k(x) = \max\{a \in S : \mu'(X'(x) \geq a \,|\, X' = \eta \text{ off } x) \geq U_k\}$$

where $\xi = X_{k-1}(\mathcal{L} \setminus \{x\})$ and $\eta = X'_{k-1}(\mathcal{L} \setminus \{x\})$. It is clear that this construction gives the correct marginal behaviors of $(X_k)_{k=0}^{\infty}$ and $(X'_k)_{k=0}^{\infty}$. The assumption (11) implies that $X_k \preceq X'_k$ whenever $X_{k-1} \preceq X'_{k-1}$. By the irreducibility of μ', the chain $(X'_k)_{k=0}^{\infty}$ will almost surely hit the maximal state of Ω at some finite (random) time, and from this time on we will thus have $X_k \preceq X'_k$. Since the coupled chain $(X_k, X'_k)_{k=0}^{\infty}$ is a finite state aperiodic Markov chain, (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 $X \preceq X'$ almost surely, whence $\mu \preceq_{\mathcal{D}} \mu'$ by Theorem 4.6. □

A non-dynamical proof of Holley's inequality by induction on $|\mathcal{L}|$, together with an extension to non-finite S, was given by Preston [197]; the simplest induction proof of an even more general result can be found in [21].

As a consequence of Holley's inequality we obtain the celebrated FKG inequality (Theorem 4.11 below) of Fortuin, Kasteleyn and Ginibre [83], who stated it under slightly different conditions. It concerns the correlation structure in a single probability measure rather than a comparison between two probability measures.

Definition 4.9 *A probability measure μ on Ω is called* **monotone** *if*

$$\mu(X(x) \geq a \,|\, X = \xi \text{ off } x) \leq \mu(X(x) \geq a \,|\, X = \eta \text{ off } x) \tag{12}$$

whenever $x \in \mathcal{L}$, $a \in S$, and $\xi, \eta \in S^{\mathcal{L} \setminus \{x\}}$ are such that $\xi \preceq \eta$, $\mu(X = \xi \text{ off } x) > 0$ and $\mu(X = \eta \text{ off } x) > 0$.

Intuitively, μ is monotone if the spin at a site x prefers to take large values whenever its surrounding sites do.

Definition 4.10 *A probability measure μ on Ω is said to have* **positive correlations** *if for all bounded increasing functions $f, g : \Omega \to \mathbf{R}$ we have*

$$\mu(fg) \geq \mu(f)\,\mu(g)\,.$$

Since the preceding inequality is preserved under rescaling and addition of constants to f and g, μ has positive correlations whenever $\mu \preceq_{\mathcal{D}} \mu'$ for any probability measure μ' with bounded increasing Radon–Nikodym density relative to μ. Theorem 4.6 thus shows that μ has positive correlations whenever $\mu(fg) \geq \mu(f)\,\mu(g)$ for all *continuous* bounded increasing functions f and g. Hence, the property of positive correlations is also preserved under weak limits.

Theorem 4.11 (The FKG inequality) *Let $\mathcal{L}$ be finite, S a finite subset of $\mathbf{R}$, and μ a probability measure on Ω which is irreducible and assigns positive probability to the maximal element of Ω (relative to $\preceq$). If μ is monotone, it also has positive correlations.*

Proof: Suppose μ' is a second probability measure on Ω such that $\mu'(\eta) = \mu(\eta)g(\eta)$ for all $\eta \in \Omega$ and some positive increasing function g. For $x \in \mathcal{L}$, $a \in S$ and $\xi \in S^{\mathcal{L}\setminus\{x\}}$ such that $\mu(X = \xi \text{ off } x) > 0$ we write $q_x(a,\xi) = \mu(X(x) \geq a \,|\, X = \xi \text{ off } x)$ and define $q'_x(a,\xi)$ similarly in terms of μ'. Then

$$\begin{aligned} q'_x(a,\xi) \,/\, (1 - q'_x(a,\xi)) &= \sum_{s\geq a} \mu(\xi^{x,s})\, g(\xi^{x,s}) \Big/ \sum_{s<a} \mu(\xi^{x,s})\, g(\xi^{x,s}) \\ &\geq \sum_{s\geq a} \mu(\xi^{x,s}) \Big/ \sum_{s<a} \mu(\xi^{x,s}) \\ &= q_x(a,\xi) \,/\, (1 - q_x(a,\xi)). \end{aligned}$$

Together with assumption (12), this implies that μ and μ' satisfy (11). Theorem 4.8 thus implies that $\mu \preceq_{\mathcal{D}} \mu'$, and the corollary follows. $\square$

Finally we state a simple observation showing that, under the condition of stochastic domination, the equality of the single-site marginal distributions already implies the equality of the whole probability measures.

Proposition 4.12 *Let $\mathcal{L}$ be finite or countable, and let μ and μ' be two probability measures on $\Omega = \mathbf{R}^{\mathcal{L}}$ satisfying $\mu \preceq_{\mathcal{D}} \mu'$. If, in addition, $\mu(X(x) \leq r) = \mu'(X(x) \leq r)$ for all $x \in \mathcal{L}$ and $r \in \mathbf{R}$ then $\mu = \mu'$.*

Proof: Let P be a coupling of μ and μ' such that $P(X \preceq X') = 1$ which exists by Theorem 4.6. Writing $\mathbf{Q}$ for the set of rational numbers, we have for each $x \in \mathcal{L}$

$$\begin{aligned} P(X(x) \neq X'(x)) &= P(X(x) < X'(x)) \\ &\leq \sum_{r\in\mathbf{Q}} P(X(x) \leq r,\ X'(x) > r) \\ &= \sum_{r\in\mathbf{Q}} \Big(P(X(x) \leq r) - P(X'(x) \leq r) \Big) \\ &= 0. \end{aligned}$$

Summing over all $x \in \mathcal{L}$ we get $P(X \neq X') = 0$, whence $\mu = \mu'$ by (10). $\square$

4.3 Applications to the Ising model

We will now apply the results of the previous subsection to the ferromagnetic Ising model. Let $(\mathcal{L}, \sim)$ be any infinite locally finite graph. For definiteness, one may think of the case $\mathcal{L} = \mathbf{Z}^d$; the arguments are, however, independent of the particular graph structure.

As in Section 3.1, we write $\mu^{\eta}_{h,\beta,\Lambda}$ for the Gibbs distribution in a finite region Λ with boundary condition $\eta \in \Omega$ and external field $h \in \mathbf{R}$ at inverse temperature $\beta > 0$. Our first application of Holley's theorem asserts that if one boundary condition dominates another, then we also have stochastic domination between the corresponding finite volume Gibbs distributions.

Lemma 4.13 *If the boundary conditions $\xi, \eta \in \Omega$ satisfy $\xi \preceq \eta$, then*

$$\mu^{\xi}_{h,\beta,\Lambda} \preceq_{\mathcal{D}} \mu^{\eta}_{h,\beta,\Lambda}\,.$$

Also, each $\mu^{\xi}_{h,\beta,\Lambda}$ has positive correlations.

Proof: The conditional probability of having a plus spin at a given site x given the configuration ξ everywhere else is equal to $(1 + \exp[-2\beta(h + \sum_{y:y\sim x} \xi(y))])^{-1}$, which is an increasing function of ξ. Theorem 4.8 and Theorem 4.11 thus imply stochastic domination between the projections of the Gibbs distributions to S^Λ and the positive correlations property. As their behavior outside Λ is deterministic, the lemma follows. □

We write $\mu^+_{h,\beta,\Lambda}$ and $\mu^-_{h,\beta,\Lambda}$ for the finite volume Gibbs distributions obtained with respective boundary conditions $\eta \equiv +1$ and $\eta \equiv -1$. Lemma 4.13 then shows that

$$\mu^-_{h,\beta,\Lambda} \preceq_{\mathcal{D}} \mu^\eta_{h,\beta,\Lambda} \preceq_{\mathcal{D}} \mu^+_{h,\beta,\Lambda} \tag{13}$$

for any $\eta \in \Omega$. This sandwich inequality reveals the special role played by the "all plus" and "all minus" boundary conditions. Next we establish the existence of the limiting "plus measure" discussed in Section 3.1. We will say that a measure μ on Ω is *homogeneous* if it is invariant under all graph automorphisms of $(\mathcal{L}, \sim)$. In the case $\mathcal{L} = \mathbf{Z}^d$, a homogeneous measure is thus invariant under translations, lattice rotations, and reflections in the axes. We write $\Lambda \uparrow \mathcal{L}$ for the limit along an arbitrary increasing sequence of finite regions which exhaust the full graph $\mathcal{L}$.

Proposition 4.14 *The limiting probability measure*

$$\mu^+_{h,\beta} = \lim_{\Lambda\uparrow\mathcal{L}} \mu^+_{h,\beta,\Lambda}$$

exists. $\mu^+_{h,\beta}$ is a homogeneous Gibbs measure for the Ising model on $(\mathcal{L}, \sim)$ with external field h and inverse temperature β and has positive correlations.

Proof: By the general theory in Section 2.6, the limit is a Gibbs measure whenever it exists. Also, by Lemma 4.13 the limit must have positive correlations. To show the existence of the limit we note that

$$\mu^+_{h,\beta,\Lambda} \succeq_{\mathcal{D}} \mu^+_{h,\beta,\Delta} \quad \text{whenever } \Lambda \subset \Delta\ . \tag{14}$$

This follows from Lemma 4.13 because $\mu^+_{h,\beta,\Lambda}$ is obtained from $\mu^+_{h,\beta,\Delta}$ by conditioning on the increasing event that $X \equiv +1$ on $\Delta \setminus \Lambda$.

Now, for any finite $A \subset \Lambda$, if Λ increases, then, by (14), $\mu^+_{h,\beta,\Lambda}(X \equiv +1 \text{ on } A)$ decreases and therefore converges to $\inf_\Delta \mu^+_{h,\beta,\Delta}(X \equiv +1 \text{ on } A)$. Note that this limit is obviously invariant under any automorphism of $(\mathcal{L}, \sim)$. By inclusion-exclusion it follows that, for any local observable f, $\mu^+_{h,\beta,\Lambda}(f)$ converges to an automorphism invariant limit as $\Lambda \uparrow \mathcal{L}$. These limits determine a unique homogeneous probability measure $\mu^+_{h,\beta}$ which, as a weak limit of finite volume Gibbs distributions, is a Gibbs measure. The lemma is thus proved. □

Obviously, replacing the "all plus" boundary condition by the "all minus" boundary condition, we obtain in the same way an automorphism invariant infinite volume Gibbs measure $\mu^-_{h,\beta}$ with positive correlations. In the same way, Lemma 4.13 shows that *any extremal Gibbs measure has positive correlations* (since it is a weak limit of $\mu^\eta_{h,\beta,\Lambda}$ for suitable η). However, positive correlations may fail for suitable $(\mathcal{L}, \sim)$ and some particular non-extremal Gibbs measures. For instance, when $\mathcal{L} = \mathbf{Z}^3$ and β is sufficiently large, one can take a convex combination of two different so-called Dobrushin states; see [66, 22].

We now take the limit in the sandwich inequality (13). Let μ be any Gibbs measure for the Ising model with parameters h and β. Taking the mean $\int \mu(d\eta)$ in (13), we obtain that $\mu^-_{h,\beta,\Lambda} \preceq_{\mathcal{D}} \mu \preceq_{\mathcal{D}} \mu^+_{h,\beta,\Lambda}$, and since stochastic domination is preserved under weak limits, we end up with

$$\mu^-_{h,\beta} \preceq_{\mathcal{D}} \mu \preceq_{\mathcal{D}} \mu^+_{h,\beta} \tag{15}$$

when μ is any Ising-model Gibbs measure for β, h. On the one hand, this shows that $\mu^-_{h,\beta}$ and $\mu^+_{h,\beta}$ are extremal, and thus equilibrium phases in the sense of Section 2.7. On the other hand, we obtain an efficient criterion for the existence of a phase transition which was first observed by Lebowitz and Martin-Löf [155] and Ruelle [205].

Theorem 4.15 *For the Ising model on an infinite locally finite graph $(\mathcal{L}, \sim)$ with external field $h \in \mathbf{R}$ and inverse temperature β, the following statements are equivalent.*

(i) *There is a unique infinite volume Gibbs measure.*

(ii) $\mu^-_{h,\beta} = \mu^+_{h,\beta}$

(iii) $\mu^-_{h,\beta}(X(x) = +1) = \mu^+_{h,\beta}(X(x) = +1)$ *for all $x \in \mathcal{L}$.*

Proof: The implications (i) ⇒ (ii) ⇒ (iii) are immediate. (iii) ⇒ (ii) follows directly from (15) and Proposition 4.12, and (ii) ⇒ (i) from (15). □

Remarks: (a) In the case $h = 0$ of no external field, assertion (iii) is equivalent to $\mu^+_{h,\beta}(X(x) = +1) = 1/2$ for all $x \in \mathcal{L}$, by the $\pm$ symmetry of the model. An extension of Theorem 4.15 in this case to the q-state Potts model will be given in Theorem 6.10.

(b) If the graph automorphisms act transitively on $(\mathcal{L}, \sim)$ then, by homogeneity, assertion (iii) is equivalent to having the equation only for *some* $x \in \mathcal{L}$. Using for example the random-cluster methods of Section 6 one can obtain the same equivalence also in the general case, assuming only that $\mathcal{L}$ is connected.

(c) If $\mathcal{L} = \mathbf{Z}^d$, (iii) is equivalent to the condition that the free energy density is differentiable with respect to h at the given values of h and β [155]. By the celebrated Lee-Yang circle theorem [203], this is the case whenever $h \neq 0$. Alternatively, one can use the so-called GHS inequality to establish (iii) for $h \neq 0$ [196]. Hence, for non-zero external field the Ising model on $\mathbf{Z}^d$ does not exhibit a phase transition.

We conclude this subsection comparing the Ising-model "plus" measures for different values of the parameters. (For similar results in a lattice gas setting see [156].)

Proposition 4.16 *Consider the Ising model on an arbitrary graph $(\mathcal{L}, \sim)$ at two inverse temperatures β_1, β_2 and two external fields h_1, h_2. Suppose that either $\beta_1 = \beta_2$ and $h_1 \leq h_2$, or $(\mathcal{L}, \sim)$ is of bounded degree $N = \sup_{x \in \mathcal{L}} N_x$ and $\beta_2 h_2 \geq \beta_1 h_1 + N|\beta_1 - \beta_2|$. Then*

$$\mu^+_{h_1,\beta_1} \preceq_{\mathcal{D}} \mu^+_{h_2,\beta_2} \,.$$

Proof: The stated conditions imply that $\mu^\xi_{h_1,\beta_1,\{x\}}(+1) \leq \mu^\eta_{h_2,\beta_2,\{x\}}(+1)$ whenever $\xi \preceq \eta$. Hence, by Theorem 4.8, $\mu^+_{h_1,\beta,\Lambda} \preceq_{\mathcal{D}} \mu^+_{h_2,\beta,\Lambda}$ for all Λ, and the proposition follows by letting $\Lambda \uparrow \mathcal{L}$. □

If $\mathcal{L}$ has bounded degree N, we obtain in particular a comparison of Ising and Bernoulli measures. Let ψ_p denote the Bernoulli measure on $\{-1, 1\}^{\mathcal{L}}$ with density p. Then

$\mu^+_{h,\beta} \to \psi_p$ as $\beta \to 0$ and $\beta h \to \frac{1}{2}\log\frac{p}{1-p}$. The preceding proposition and Corollary 4.7 thus show that

$$\mu^+_{h,\beta} \succeq_{\mathcal{D}} \psi_p \quad \text{for} \quad h \geq \frac{1}{2\beta}\log\frac{p}{1-p} + N$$

and

$$\mu^+_{h,\beta} \preceq_{\mathcal{D}} \psi_p \quad \text{for} \quad h \leq \frac{1}{2\beta}\log\frac{p}{1-p} - N\ .$$

4.4 Application to other models

Do the arguments of the previous subsection extend to the other models of Section 3? The answer to this question is different for the different models. For the *Potts model* with $q \geq 3$, Lemma 4.13 fails because the conditional probability that the spin at a site x takes a large value is not increasing in the surrounding spin configuration. Nevertheless, the Potts model admits some analogues of Proposition 4.14 and Theorem 4.15. These results are deeper than their Ising counterparts, and will be demonstrated by random-cluster arguments in Section 6.

The *Widom–Rowlinson model* exhibits the same monotonicity properties as the Ising model. We thus obtain Widom–Rowlinson analogues of Lemma 4.13 and Proposition 4.14. Fixing any two activity parameters λ_+, $\lambda_- > 0$ and writing μ^+ (resp. μ^-) for the associated limiting Gibbs measures with "all plus" (resp. "all minus") boundary conditions, we find that any other Gibbs measure for the same parameters is sandwiched (in the sense of (15)) between μ^+ and μ^-. The analogue of Theorem 4.15 reads as follows.

Theorem 4.17 *For the Widom–Rowlinson model on an infinite locally finite graph* $(\mathcal{L}, \sim)$ *with activity parameters* λ_+, $\lambda_- > 0$*, the following statements are equivalent.*

(i) *There is a unique infinite volume Gibbs measure.*

(ii) $\mu^+ = \mu^-$

(iii) $\mu^+(X(x) = +1) = \mu^-(X(x) = +1)$ *for all* $x \in \mathcal{L}$.

The *Ising antiferromagnet* and the *hard-core lattice gas model* are far from satisfying the monotonicity properties needed for the arguments in Section 4.3; the conditional probability that a site x takes the value $+1$ is *decreasing*, rather than increasing, in the surrounding configuration. However, when $(\mathcal{L}, \sim)$ is bipartite (as in the case $\mathcal{L} = \mathbf{Z}^d$) we can use again the trick of (6) to flip all spins on the odd sublattice. The Ising antiferromagnet is then mapped onto the Ising ferromagnet with a staggered external field (having alternating signs on the even and the odd sublattices). Similarly, the hard-core model is mapped into a model which also exhibits the necessary monotonicity properties. We thus obtain analogous results which we spell out only for the hard-core model with activity $\lambda > 0$: There exist two particular Gibbs measures μ^{even}_λ and μ^{odd}_λ, obtained as infinite volume limits of finite volume Gibbs distributions with respective boundary conditions η_{even} and η_{odd}, defined in Section 3.4. In terms of these two Gibbs measures, the existence of a phase transition can be characterized as follows.

Theorem 4.18 *For the hard-core model on an infinite locally finite bipartite graph* $(\mathcal{L}, \sim)$ *with activity parameter* λ*, the following are equivalent.*

(i) *There is a unique infinite volume Gibbs measure.*

(ii) $\mu_\lambda^{even} = \mu_\lambda^{odd}$

(iii) $\mu_\lambda^{even}(X(x) = 1) = \mu_\lambda^{odd}(X(x) = 1)$ *for all* $x \in \mathcal{L}$*.*

5 Percolation

We will now introduce the ideas of random geometry we referred to in the title of this paper. As these were developed first in the framework of percolation theory, we devote this section to a description of this subject. We will start with the classical case of independent, or Bernoulli, percolation, and will then proceed to the case of dependent percolation. In the subsequent sections we will see how these results and ideas can be used for the geometric analysis of equilibrium phases.

5.1 Bernoulli percolation

Bernoulli percolation was introduced in the 1950's in papers by Broadbent and Hammersley [39, 124, 125] as a model for the passage of a fluid through a porous medium. In fact, the model has appeared first in [79] in the context of polymerization. We give here only a brief introduction; a thorough treatment can be found in the the books and lectures by Grimmett [108, 110], Chayes and Chayes [52], and Kesten [143].

The porous medium is modelled by a graph $(\mathcal{L}, \sim)$, and either the sites or the bonds of this graph are considered to be randomly open or closed (blocked). We begin with the case of site percolation; the alternative case of random bonds will be discussed at the end of this subsection.

The basic question of percolation theory is how a fluid can spread through the medium. This involves the connectivity properties of the set of open vertices. To describe this we introduce some terminology. A finite *path* is a sequence $(v_0, e_1, v_1, e_2, \dots, e_k, v_k)$, where $v_0, \dots, v_k \in \mathcal{L}$ are pairwise distinct vertices and $e_1, \dots, e_k \in \mathcal{B}$ are pairwise distinct edges such that, for each $i \in \{1, \dots, k\}$, the edge e_i connects the vertices v_{i-1} and v_i. Obviously, a path is equivalently described by its sequence $(v_0, v_1, \dots, v_k)$ of vertices or its sequence $(e_1, e_2, \dots, e_k)$ of edges. The number k is called the *length* of the path. In the same way, we can also speak of infinite paths $(v_0, e_1, v_1, e_2, \dots)$ and doubly infinite paths $(\dots, v_{-1}, e_0, v_0, e_1, v_1, \dots)$. A region $C \subset \mathcal{L}$ is called *connected* if for any $x, y \in C$ there exists a path which starts at x, ends at y, and which only contains vertices in C.

An *open path* is a path on which all vertices are open. An *open cluster* is a maximal connected set C in which all vertices are open; here maximal means that there is no larger region $C' \supset C$ which is connected and only contains open vertices. An infinite open cluster (or infinite cluster, for short) is an open cluster containing infinitely many vertices. Using these terms, we may say that the existence of an infinite open cluster is equivalent to the fact that a fluid can wet a macroscopic part of the medium.

We now turn to the classical case of Bernoulli site percolation with retention parameter $p \in [0, 1]$. In this case, each vertex of $\mathcal{L}$, *independently of all others*, is declared to be open (and represented by the value 1) with probability p and closed (with value 0) with probability $1 - p$. We write ψ_p for the associated (Bernoulli) probability measure on the configuration space $\{0, 1\}^{\mathcal{L}}$.

The first question to be asked is whether or not infinite clusters can exist. This depends, of course, on both the graph $(\mathcal{L}, \sim)$ and the parameter p. The basic observation is the following.

Proposition 5.1 *For Bernoulli site percolation on an infinite locally finite graph*

$(\mathcal{L}, \sim)$*, there exists a critical value* $p_c \in [0,1]$ *such that*

$$\psi_p(\exists \text{ an infinite open cluster}) = \begin{cases} 0 & \text{if } p < p_c \\ 1 & \text{if } p > p_c. \end{cases}$$

At the critical value $p = p_c$*, the* ψ_p*-probability of having an infinite open cluster is either* 0 *or* 1.

Proof: A moment's thought reveals that the existence of infinite clusters is invariant under a change of the status of finitely many vertices. By Kolmogorov's zero-one law, the ψ_p-probability of having infinite clusters is therefore either 0 or 1. It remains to show that this probability is increasing in p. The existence of an infinite open cluster is obviously an increasing event, so we are done if we can show that

$$\psi_{p_1} \preceq_{\mathcal{D}} \psi_{p_2} \quad \text{whenever } p_1 \leq p_2. \tag{16}$$

This is intuitively obvious and can be proved by the following elementary coupling argument. Let $Y = (Y(x))_{x \in \mathcal{L}}$ be a family of i.i.d. random variables with uniform distribution on $[0,1]$, and for $p \in [0,1]$ let $X_p = (X_p(x))_{x \in \mathcal{L}}$ be defined by $X_p(x) = I_{\{Y(x) \leq p\}}$. It is then clear that X_p has distribution ψ_p and $X_{p_1} \preceq X_{p_2}$ whenever $p_1 \leq p_2$. This implies (16) by (the trivial part of) Theorem 4.6. (Note that we have in fact constructed a simultaneous coupling of all ψ_p's, and that this construction would be used in Monte Carlo simulations of Bernoulli percolation.) □

We write $\{x \leftrightarrow \infty\}$ for the event that $x \in \mathcal{L}$ belongs to an infinite cluster, and set $\theta_x(p) = \psi_p(x \leftrightarrow \infty)$. For homogenous graphs such as $\mathcal{L} = \mathbf{Z}^d$, $\theta_x(p)$ does not depend on x, and then we write simply $\theta(p)$. Equation (16) shows that $\theta_x(p)$ is increasing in p. We also make the following observation.

Proposition 5.2 *For any infinite locally finite connected graph* $(\mathcal{L}, \sim)$*, any* $x \in \mathcal{L}$ *and any* $p \in (0,1)$*, we have* $\theta_x(p) > 0$ *if and only if* $\psi_p(\exists \text{ an infinite open cluster}) = 1$.

Proof: From Proposition 5.1 we know that an infinite cluster exists with probability 0 or 1. The implication "only if" is therefore immediate. For the "if" part, we note that if an infinite cluster exists with positive probability, then there is some N such that $\psi_p(A_N) > 0$, where A_N is the event that some vertex within distance N from x belongs to an infinite cluster. On the other hand we have $\psi_p(B_N) > 0$, where B_N is the event that all vertices within distance $N-1$ from x are open. The event $A_N \cap B_N$ implies that x belongs to an infinite cluster. But A_N and B_N are increasing events (see Section 4.2), so that we can apply Theorem 4.11 to obtain $\theta_x(p) \geq \psi_p(A_N \cap B_N) \geq \psi_p(A_N)\psi_p(B_N) > 0$. □

Note that the proof above applies to the much broader class of all measures with positive correlations (recall Definition 4.10), rather than only the Bernoulli measures.

Next we ask whether both possibilities in Proposition 5.2 really occur, that is, if $0 < p_c < 1$. For, only in this case we really have a *nontrivial* critical phenomenon at p_c. The answer depends on the graph. For $\mathcal{L} = \mathbf{Z}^d$ with dimension $d \geq 2$, the threshold p_c is indeed nontrivial, as is stated in the theorem below. This nontriviality of p_c is a fundamental ingredient of many of the stochastic-geometric arguments employed later on. On the other hand, it is easy to see that $p_c = 1$ for $\mathcal{L} = \mathbf{Z}^1$.

Theorem 5.3 *The critical value $p_c = p_c(d)$ for site percolation on $\mathcal{L} = \mathbf{Z}^d$, $d \geq 2$, satisfies the inequalities*

$$\frac{1}{2d-1} \leq p_c \leq \frac{6}{7}. \tag{17}$$

Proof: We begin with the lower bound on p_c. For $k = 1, 2, \ldots$, we write N_k for the random number of open paths of length k starting at 0. On the event $\{0 \leftrightarrow \infty\}$ we have $N_k \geq 1$ for each k, whence

$$\theta(p) \leq \psi_p(N_k)\,. \tag{18}$$

The number of all paths of length k starting at 0 is at most $2d(2d-1)^{k-1}$, and each path is open with probability p^k. Hence

$$\psi_p(N_k) \leq 2d(2d-1)^{k-1}p^k$$

which tends to 0 as $k \to \infty$ whenever $p < 1/(2d-1)$. In combination with (18) this implies that $\theta(p) = 0$ for $p < 1/(2d-1)$, and the first half of (17) is established.

The second half of (17) only needs to be proved for $d = 2$; this is because $\mathbf{Z}^2$ can be embedded into $\mathbf{Z}^d$ for any $d \geq 2$, so that $p_c(d) \leq p_c(2)$. So let $d = 2$. We first need some additional terminology. A $*$-path in $\mathbf{Z}^2$ is a sequence $(v_0, v_1, \ldots, v_k)$ of distinct vertices such that $d_\infty(v_{j-1}, v_j) = 1$ for $j = 1, \ldots k$. Note that two consecutive vertices in a $*$-path need not be nearest-neighbors; they may also be "diagonal neighbors". A *$*$-circuit* is a sequence $(v_0, v_1, \ldots, v_k, v_0)$ such that $(v_0, v_1, \ldots, v_k)$ is a $*$-path and $d_\infty(v_k, v_0) = 1$. Informally, a $*$-circuit is a $*$-path which ends where it starts. $*$-circuits with the same set of sites are identified. A *closed $*$-circuit* is a $*$-circuit in which all vertices are closed.

Now let M be the number of closed $*$-circuits that surround the origin 0. As $\mathbf{Z}^2$ is a planar graph, the "outer boundary" of a finite open cluster containing 0 defines a closed $*$-circuit around 0. Hence, the event $\{0 \leftrightarrow \infty\}$ occurs if and only if $M = 0$.

The number of *all* (not necessarily closed) $*$-circuits of a given length k surrounding 0 allows the following crude estimate. Consider the leftmost crossing of the x-axis of such a circuit; the location of such a crossing is at distance at most k from the origin, so there are at most k such locations to choose from. Starting at this location, we may trace the $*$-circuit clockwise (say), and at each step we have at most 7 d_∞-neighbors to choose from. Hence, the number of $*$-circuits of length k around 0 is at most $k\,7^{k-1}$. Each one is closed with probability $(1-p)^k$, so

$$\psi_p(M) \leq \sum_{k=1}^{\infty} k\,7^{k-1}(1-p)^k.$$

The last sum is finite for $p > 6/7$. Hence, for such a p and n large enough, there is a positive probability for having no closed $*$-circuit around 0 which contains a site of distance at least n from 0. By the argument in the proof of Proposition 5.2, it follows that $\psi_p(M = 0) > 0$ for such p. Hence $\theta(p) > 0$ for $p > 6/7$, and the second half of (17) follows. □

The exact value of the percolation threshold $p_c(d)$ of $\mathbf{Z}^d$ is not known for any $d \geq 2$. The best rigorous bounds for $d = 2$ are presently

$$0.556 < p_c(2) < 0.680 \tag{19}$$

where the first inequality is due to van den Berg and Ermakov [26] (inspired by [176]) and the second to Wierman [224]. For high dimensions, it is known that

$$\lim_{d\to\infty} 2d\, p_c(d) = 1\,, \tag{20}$$

see Kesten [144], Gordon [104] and Hara and Slade [127] for this and finer asymptotics. On some particular graphs, p_c can be determined exactly. For example, for $\mathcal{L} = \mathbf{T}_d$, the regular (Cayley or Bethe) tree, branching-process arguments immediately show that $p_c(\mathbf{T}_d) = 1/d$, and for the triangular lattice it follows from planar duality that $p_c = 1/2$ [143].

The preceding considerations do not tell us what happens *at* the critical value p_c. It is believed that, for the integer lattice $\mathbf{Z}^d$ of any dimension $d \geq 2$, there is ψ_{p_c}-a.s. no infinite cluster, which means that $\theta(p_c) = 0$; so far this is only known for $d = 2$ and $d \geq 19$, see Russo [209] and Hara and Slade [126]. One can show that the relation $\theta(p_c) = 0$ implies continuity of $\theta(p)$ at $p = p_c$, so in combination with the trivial continuity of $\theta(p)$ in the subcritical regime and the following more interesting result (which can be found e.g. in [108]) we get continuity of $\theta(p)$ throughout $[0, 1]$ as soon as absence of infinite clusters at criticality is established.

Theorem 5.4 *For Bernoulli site percolation on $\mathbf{Z}^d$, $d \geq 2$, the function $\theta(p)$ is continuous throughout the supercritical regime $(p_c, 1]$.*

So far we were interested in the existence of infinite clusters. In the subcritical regime $p < p_c$ when no infinite cluster exists, one may ask for the size of a typical cluster. Let $|C_0|$ be the random number of vertices in the open cluster containing the origin; we set $|C_0| = 0$ if the origin is closed. By the definition of p_c, $\psi_p(|C_0| = \infty) = 0$. For $\mathcal{L} = \mathbf{Z}^d$, we even have the stronger statement that the expected value of $|C_0|$ is finite.

Theorem 5.5 *For Bernoulli site percolation on $\mathbf{Z}^d$ with retention parameter $p < p_c$, we have $\psi_p(|C_0|) < \infty$.*

This was proved independently by Menshikov [174] and by Aizenman and Barsky [6]. The proofs are rather involved, so we refer the reader to the original articles and [108]. It is worth noting that Theorem 5.5 fails in the setting of general graphs; a striking counterexample is the "three-one-tree" discussed on p. 936 of Lyons [164].

Menshikov even showed that the distribution of the radius of the open cluster containing the origin decays exponentially. The following even stronger result states that the same is true for the distribution of $|C_0|$; see Grimmett [108] for a proof (in the case of bond percolation).

Theorem 5.6 *For Bernoulli site percolation on $\mathbf{Z}^d$ with retention parameter $p < p_c$, there exists a constant c (depending on p) such that*

$$\psi_p(|C_0| \geq n) \leq e^{-c\,n}$$

for all n.

Looking at a fixed path of length n starting at 0, we immediately obtain the lower bound $\psi_p(|C_0| \geq n) \geq p^n$. So the preceding upper bound is best possible, except that the optimal constant c is unknown.

We conclude this section with some remarks on *Bernoulli bond percolation.* The model is similar, except that now the edges rather than the vertices in $(\mathcal{L}, \sim)$ are independently open (described by the value 1) or closed (with value 0) with respective probabilities p and $1-p$. The associated configuration space is thus $\{0,1\}^{\mathcal{B}}$. We write ϕ_p for the associated (Bernoulli) probability measure on $\{0,1\}^{\mathcal{B}}$. In the present context of bond percolation, an *open path* is a path in which all edges are open, and an *open cluster* is a maximal region $C \subset \mathcal{L}$ which is connected, in that for any $x, y \in C$ there is an open path in C from x to y.

All results for site percolation discussed so far extend to the bond percolation set-up. This is no surprise because bond percolation is equivalent to site percolation on the so-called covering graph for which $\mathcal{B}$ is taken as set of vertices, and edges are drawn between any two coincident elements of $\mathcal{B}$. In particular, there exists again a critical value p_c for the occurrence of infinite open clusters, and Propositions 5.1 and 5.2, Theorems 5.3, 5.5, 5.6 and 5.4, and the asymptotic formula (20) are still true in the case of bond percolation. What is generally different, are the critical values for site and bond Bernoulli percolation on a given graph. One remarkable case is that of bond percolation on $\mathbf{Z}^2$, where (again by planar duality) $p_c = 1/2$; this is a famous result of Kesten [142]. In the specific case of trees, however, site and bond percolation are equivalent. In particular, for $\mathcal{L} = \mathbf{T}_d$ we have $p_c = 1/d$ for both site and bond percolation; for more general trees a formula for p_c was given by Lyons [164].

5.2 Dependent percolation: the role of the density

Our main subject is the analysis of equilibrium phases by means of percolation methods. In this case, a site will be considered as open if, for example, the configuration in a neighborhood of this site shows a specified pattern, and the events "site x is open" with $x \in \mathcal{L}$ are then far from being independent. This leads us to considering the case of dependent percolation. In this subsection we do some first steps in this direction.

Our starting point is the following question. In the case of Bernoulli percolation, there is a unique parameter, the occupation probability or density p, which governs the phase diagram and allows to distinguish between subcritical ("no infinite cluster") and supercritical ("at least one infinite cluster") behavior. Does this also hold in general? Specifically, is it true that for any translation invariant probability measure μ on $\{0,1\}^{\mathbf{Z}^d}$, the occurrence of an infinite cluster only depends on the density $p(\mu) = \mu(X(x) = 1)$ of open sites $x \in \mathbf{Z}^d$? In general, the answer is obviously "no", as we will now show by two simple examples: there exist translation invariant measures μ on $\{0,1\}^{\mathbf{Z}^d}$ with arbitrarily small densities such that infinite clusters exist almost surely, and also translation invariant measures with densities arbitrarily close to 1 for which no infinite clusters exist with probability 1.

Example 5.7 For $q \in (0,1)$, let $(Y(x),\, x \in \mathbf{Z})$ be i.i.d. random variables taking values 0 and 1 with probability $1-q$ resp. q. We define a translation invariant random field $(X(x),\, x \in \mathbf{Z}^d)$, $d \geq 2$, by setting $X(x) = X(x_1, \ldots, x_d) = Y(x_1)$ for each $x \in \mathbf{Z}^d$. Writing μ for the distribution of $(X(x),\, x \in \mathbf{Z}^d)$, we have $p(\mu) = q$, but with μ-probability 1 there exist infinitely many infinite clusters, even if q is arbitrarily small.

Example 5.8 Again let $q \in (0,1)$, $d \geq 2$, and $(Y(x,i),\, x \in \mathbf{Z},\, i \in \{1, \ldots, d\})$ be i.i.d. random variables taking values 0 and 1 with probabilities $1-q$ and q. We define a translation invariant random field $(X(x),\ x \in \mathbf{Z}^d)$ by setting

$$X(x) = X(x_1, \ldots, x_d) = \prod_{i=1}^{d} Y(x_i, i).$$

A moment's thought reveals that μ-a.s. there exist no infinite clusters, despite the fact that $p(\mu) = q^d$ may be arbitrarily close to 1.

These examples suggest to look for additional assumptions under which high (resp. low) density guarantees existence (resp. nonexistence) of infinite clusters. Positive correlations (in the sense of Definition 4.10) does not suffice, because both examples above obviously have positive correlations.

An alternative might be to assume R-independence in the sense that $X(\Lambda)$ and $X(\Delta)$ are independent for any two finite regions $\Lambda, \Delta \subset \mathcal{L}$ such that

$$\min_{x\in\Lambda, y\in\Delta} |x - y| > R$$

for some given R. For $\mathbf{Z}^d, d \geq 2$, this gives nontrivial thresholds $0 < p_1 < p_2 < 1$ (depending on R) such that existence (resp. non-existence) of infinite clusters is guaranteed as long as $p(\mu) > p_2$ (resp. $p(\mu) < p_1$); see e.g. Liggett, Schonmann and Stacey [161]. However, R-independence rarely holds in Gibbs models. For instance, for plus phase of the Ising model with vanishing external field and inverse temperature $\beta > 0$, the spins at any two vertices are always strictly positively correlated no matter how far apart they are (although the correlation does tend to 0 in the distance).

However, in contrast to what we just saw in the case of the cubic lattices, the density does play a significant role for the regular trees $\mathbf{T}_d$. To show this we consider a probability measure μ on $\{0,1\}^{\mathcal{L}}$, where now $\mathcal{L} = \mathbf{T}_d$ with $d \geq 2$. The natural analogue of translation invariance in this setting is *automorphism invariance* of μ, which means that μ inherits all the symmetries of $\mathbf{T}_d$. In particular, this implies that $\mu(X(x) = 1)$ is independent of x, so that the density $p(\mu)$ is well-defined. As opposed to the $\mathbf{Z}^d$ case, having $p(\mu)$ sufficiently close to 1 now does guarantee that an infinite cluster exists with positive probability. This is also true in the bond percolation case, where $p(\mu)$ is defined as the probability that a given edge is open. The following result is due to Häggström [118].

Theorem 5.9 *For any automorphism invariant site percolation model μ on $\mathbf{T}_d$ with density $p(\mu) \geq \frac{d+1}{2d}$, we have $\mu(\exists \text{ an infinite open cluster}) > 0$. The same is true for bond percolation on $\mathbf{T}_d$ with density $p(\mu) \geq \frac{2}{d+1}$.*

These bounds are in fact sharp, in that for any $p < \frac{d+1}{2d}$ there exists some automorphism invariant probability measure on $\{0,1\}^{\mathbf{T}_d}$ with density p, which does not allow an infinite cluster with probability 1, and similarly for the case of bond percolation; see [118]. It follows from Example 5.8 that the corresponding threshold for $\mathbf{Z}^d$ is trivial: only density 1 is enough to rule out the nonexistence of infinite clusters. The intuitive reason is the following. On $\mathbf{Z}^d$, one can find finite regions $\Lambda \subset \mathcal{L}$ with arbitrarily small surface-to-volume ratio, which means that a vast majority of sites is not adjacent to a vertex outside Λ; we can simply take $\Lambda = \Lambda_n = [-n, n]^d \cap \mathbf{Z}^d$ with large n; this property of $\mathbf{Z}^d$ is known as *amenability.* Hence, a relatively small number of closed vertices may easily "surround" a large number of open sites. In contrast, every region in $\mathbf{T}_d$ has a surface of the same order of magnitude as its volume; this makes it impossible for a small minority of closed vertices to surround a large number of open vertices. This intuition can be turned into a proof using the so called mass-transport method sketched below. Benjamini, Lyons, Peres and Schramm [23] have recently extended this method to derive a similar dichotomy for a large class of graphs, including Cayley graphs of finitely generated groups.

Sketch proof of Theorem 5.9: For simplicity we confine ourselves to the case of bond percolation on $\mathbf{T}_2$. We want to show that if $p(\mu) \geq 2/3$, then an infinite cluster exists with positive μ-probability. Imagine the following allocation of mass to the edges of $\mathbf{T}_2$. Originally every edge receives mass 1. Then the mass is redistributed, or transported, as follows. If an edge e is open and is contained in a finite open cluster, then it distributes all its mass equally among those closed edges that are adjacent to the open cluster containing e. If e is open and contained in an infinite open cluster, then it keeps its mass. Closed edges, finally, keep their own mass and happily accept any mass that open edges decide to send them. The expected mass at each edge before transport is obviously 1, and one can show — this is an instance of the mass-transport principle [23] — that the expected mass at a given edge is 1 also after the transport. Suppose now, for contradiction, that $p(\mu) \geq 2/3$ and that all open clusters are finite μ-a.s. Then all open edges have mass 0 after transport. Furthermore, since each open cluster containing exactly n edges has exactly $n+3$ adjacent closed edges (as is easily shown by induction — it is here that the tree structure is used), the mass after transport at a closed edge adjacent to two open clusters of sizes n_1 and n_2 has mass

$$1 + \frac{n_1}{n_1+3} + \frac{n_2}{n_2+3} < 3$$

Hence the expected mass after transport at a given edge e is strictly less than

$$3\,\mu(X(e) = 0) = 3(1 - p(\mu)) \leq 1,$$

contradicting the mass-transport principle. □

5.3 Examples of dependent percolation

¿From the previous subsection the reader might get a rather pessimistic view of the possibilities of establishing existence (or non-existence) of infinite clusters for dependent percolation models on $\mathbf{Z}^d$. This is certainly *not* the case, and a lot can be done. One standard way of determining the percolation behavior of a dependent model is by *stochastic comparison with a suitable Bernoulli percolation model*: For the existence of infinite clusters, it is sufficient to show that the given dependent model is stochastically larger than the Bernoulli model for some parameter $p > p_c$, and the absence of infinite clusters will follow if the model at hand is stochastically dominated by the Bernoulli model for some $p < p_c$. Let us demonstrate this technique for the Ising model on $\mathbf{Z}^d$.

Consider percolation of plus spins in the plus measure $\mu^+_{h,\beta}$, defined in Section 4.3. If we keep β fixed then Proposition 4.16 tells us that $\mu^+_{h,\beta}$ is stochastically increasing in h. Consequently, both the probability of having an infinite cluster of plus spins, as well as the probability that a given vertex is in such an infinite cluster, are increasing in h. Furthermore, as $\mathbf{Z}^d$ is of bounded degree $N = 2d$, the remarks after the same proposition imply that, for any given $p \in (0,1)$ and β, the Ising measure $\mu^+_{h,\beta}$ stochastically dominates the Bernoulli measure ψ_p when h is large enough, and is dominated by ψ_p for h below some bound. We may combine this observation with Proposition 5.1 to deduce the following critical phenomenon:

Theorem 5.10 *For the Ising model on $\mathbf{Z}^d$, $d \geq 2$, at a fixed temperature β, there exists a critical value $h_c \in \mathbf{R}$ (depending on d and β) for the external field, such that*

$$\mu^+_{h,\beta}(\exists \text{ an infinite cluster of plus spins}) = \begin{cases} 0 & \text{if } h < h_c \\ 1 & \text{if } h > h_c \,. \end{cases}$$

As we shall see later in Theorem 8.2, we have $h_c = 0$ when $d \geq 2$ and $\beta > \beta_c$. Higuchi [133] has shown that the percolation transition at h_c is sharp, in that the connectivity function decays exponentially when $h < h_c$, and that the percolation probability is continuous in (β, h), except on the critical half-line $h = 0$, $\beta > \beta_c$. In Section 6 below, we will make a similar use of stochastic comparison arguments for random-cluster measures, cf. Proposition 6.11. The stochastic domination approach works also in the framework of lattice gases with attractive potential; see Lebowitz and Penrose [156].

In the rest of this subsection we shall give some examples of strongly dependent systems where other approaches to the question of percolation are needed. Typically in these examples, the probability that all vertices in a finite region Λ are open (or closed) fails to decay exponentially in the volume of Λ, and as a consequence, the random field neither dominates nor is stochastically dominated by any nontrivial Bernoulli model.

The geometry of level heights of a random field forms an important object of study both from the theoretical and the applied side. For example, it relates to the presence of hills and valleys on a rough surface, or to the random location of potential barriers in a doped semi-conductor. To fix the ideas we consider a random field $X = (X(x), x \in \mathbf{Z}^d)$ with values $X(x) \in S \subset \mathbf{R}$ which are not necessarily discrete. It is often interesting to divide S into two parts S_1 and S_0 and to define a new discrete random field Y via $Y(x) = I_{\{X(x) \in S_1\}}$. For $S = \mathbf{R}$ one typically considers $S_1 = [\ell, \infty)$ for some level $\ell \in \mathbf{R}$. In this way we obtain a coarse-grained description of a system of continuous spins. One question is to which extent one can reconstruct the complete image from this information. We consider here a different question: what is the geometry of the random set $\{x \in \mathbf{Z}^d : Y(x) = 1\}$? This set is called the excursion or exceedance set when it corresponds, as in the example above, to the set on which the original random field exceeds a given level. For a recent review of this subject we refer to [2].

We now give four examples of equilibrium systems with continuous spins where one can show (the absence of) percolation of an excursion (exceedance) set. Here we only state the results. Some hints on the proofs will be given later in Section 8 via Theorem 8.1. Details can be found in the paper by Bricmont, Lebowitz and Maes [36].

Example 5.11 Consider a general model of real-valued spins $(\sigma(x), x \in \mathbf{Z}^d)$ with ferromagnetic nearest-neighbor interaction. The formal Hamiltonian is given by

$$H(\sigma) = -\sum_{x \sim y} \sigma(x)\sigma(y) \,. \tag{21}$$

The reference (or single-spin) measure $\lambda \neq \delta_0$ on $\mathbf{R}$ is assumed to be even and to decay fast enough at $\pm\infty$ so that the model is well defined. Then, for any Gibbs measure μ relative to (21) with $\mu(\text{sgn}(\sigma(0))) > 0$, there will be percolation of all sites x with $\sigma(x) \geq 0$. Such Gibbs measures always exist at sufficiently low temperatures when $d \geq 2$.

Example 5.12 Consider again a spin system ($\sigma(x)$, $x \in \mathbf{Z}^d$), where now the Hamiltonian has the 'massless' form

$$H(\sigma) = -\sum_{x \sim y} \psi(\sigma(x) - \sigma(y))$$

with $\sigma(x) \in \mathbf{R}$ or $\mathbf{Z}$ and ψ an even convex function. The single-spin measure λ is either Lebesgue measure on $\mathbf{R}$ or counting measure on $\mathbf{Z}$. The case $\sigma(x) \in \mathbf{Z}$ and $\psi(t) = |t|$ corresponds to the so-called solid-on-solid (SOS) model of a d-dimensional surface in $\mathbf{Z}^{d+1}$; the choice $\sigma(x) \in \mathbf{R}$ and $\psi(t) = t^2$ gives the harmonic crystal. Let μ be a Gibbs measure which is obtained as infinite volume limit of finite volume Gibbs distributions with zero boundary condition. (In the continuous-spin case, such Gibbs measures exist for any temperature when $d \geq 3$ and $\psi(t) = \alpha t^2 + \phi(t)$, where $\alpha > 0$ and ϕ is convex [33].) Then, for any $\ell < 0$, there is percolation of the sites $x \in \mathbf{Z}^d$ with $\sigma(x) \geq \ell$.

Example 5.13 Consider next a model of two-component spins $\sigma(x) \in \mathbf{R}^2$, $x \in \mathbf{Z}^d$, $\sigma(x) = (r_x \cos \phi_x, r_x \sin \phi_x)$, with formal Hamiltonian

$$H(\sigma) = -\sum_{x \sim y} \sigma(x) \cdot \sigma(y)$$

and some rotation-invariant and suitably decaying reference measure λ on $\mathbf{R}^2$. Then, for any Gibbs measure μ with $\mu(\cos \phi_0) > 0$, there is percolation of the sites $x \in \mathbf{Z}^d$ with $\cos \phi_x \geq 0$. Such Gibbs measures exist at low temperatures if $d \geq 3$.

Example 5.14 Consider again the massless harmonic crystal of Example 5.12 above (with $\psi(t) = t^2$) in $d = 3$ dimensions. There exists a value $\ell_c < \infty$ so that for all $\ell \geq \ell_c$ there is *no* percolation of sites $x \in \mathbf{Z}^d$ with $\sigma(x) \geq \ell$.

Finally, we give an example of a strongly correlated system, sharing some properties with the harmonic crystal of Example 5.14, where at present there is no proof of a percolation transition. The model is one of the simplest examples of an interacting particle system. What makes the problem difficult is that the random field is not Markov (not even Gibbsian) and not explicitly described in terms of a family of local conditional distributions.

Example 5.15 The voter model is a stochastic dynamics in which individuals (voters) sitting at the vertices of a graph update their position (yes/no) by randomly selecting a neighboring vertex and adopting its position, see Liggett [160] for an introduction. Using spin language and putting ourselves on $\mathbf{Z}^3$, the time evolution of this voter model is specified by giving the rate $c(x, \sigma)$ for a spin flip at the site x when the spin configuration is $\sigma \in \{+1, -1\}^{\mathbf{Z}^3}$,

$$c(x, \sigma) = \frac{1}{6} \sum_{y \sim x} \Big(1 - \sigma(x)\sigma(y)\Big) .$$

There is a one-parameter family of extremal invariant measures μ_p each obtained asymptotically (in time) from taking the Bernoulli measure ψ_p with density p as initial condition. These stationary states μ_p are strongly correlating. The spin-spin correlations decay as the inverse $1/r$ of the spin-distance r on $\mathbf{Z}^3$. It is an open question whether for p sufficiently close to 1 the plus spins percolate, and whether for sufficiently small p

there is no percolation. Simulations by Lebowitz and Saleur [157] indicate that there is indeed a non-trival percolation transition with critical value $p_c \approx 0.16$.

The same problem may be considered for $d \geq 4$. For $d = 1$, 2, however, the problem is not interesting because in these cases μ_p is known to put mass p on the "all +1" configuration and mass $1-p$ on the "all -1" configuration. Alternatively, one may consider the same model with $\mathbf{Z}^3$ replaced by $\mathbf{T}_d$; Theorem 5.9 can then be applied to show that the plus spins do percolate for $p \geq \frac{d+1}{2d}$.

5.4 The number of infinite clusters

Once infinite clusters have been shown to exist with positive probability in some percolation model, the next natural question is: *How many infinite clusters can exist simultaneously?* For Bernoulli site or bond percolation on $\mathbf{Z}^d$, Aizenman, Kesten and Newman [12] obtained the following, now classical uniqueness result: with probability 1, there exists at most one infinite cluster. Simpler proofs were found later by Gandolfi, Grimmett and Russo [89] and by Burton and Keane [42]. The argument of Burton and Keane is not only the shortest (and, arguably, the most elegant) so far. Also, it requires much weaker assumptions on the percolation model, namely: translation invariance and the finite-energy condition below, which is a strong way of stating that all local configurations are really possible; its significance for percolation theory had been discovered before by Newman and Schulman [183].

Definition 5.16 *A probability measure μ on $\{0,1\}^{\mathcal{L}}$, with $\mathcal{L}$ a countable set, is said to have* **finite energy** *if, for every finite region $\Lambda \subset \mathcal{L}$,*

$$\mu(X \equiv \eta \text{ on } \Lambda \,|\, X \equiv \xi \text{ off } \Lambda) > 0$$

for all $\eta \in \{0,1\}^\Lambda$ and μ-a.e. $\xi \in \{0,1\}^{\Lambda^c}$.

Theorem 5.17 (The Burton–Keane uniqueness theorem) *Let μ be a probability measure on $\{0,1\}^{\mathbf{Z}^d}$ which is translation invariant and has finite energy. Then, μ-a.s., there exists at most one infinite open cluster.*

Sketch proof: Without loss of generality we can assume that μ is ergodic with respect to translations. For, one can easily show that the measures in the ergodic decomposition of μ admit the same conditional probabilities, and thus inherit the finite-energy property. Since the number N of infinite clusters is obviously invariant under translations, it then follows that N is almost surely equal to some constant $k \in \{0,1,\ldots,\infty\}$. In fact, $k \in \{0,1,\infty\}$. Otherwise, with positive probability each of the k clusters would meet a sufficiently large cube Λ; by the finite-energy property, this would imply that with positive probability all these clusters are connected within Λ, so that in fact $k=1$, in contradiction to the hypothesis. (This part of the argument goes back to [183].)

We thus only need to exclude the case $k=\infty$. In this case, $\mu(N \geq 3) = 1$, and the finite-energy property implies again that $\mu(A_x) = \delta > 0$, where A_x is the event that x is a triple point, in that there exist three disjoint infinite open paths with starting point x. By the (norm-) ergodic theorem, for any sufficiently large cubic box Λ we have

$$\mu\Big(|\Lambda|^{-1} \sum_{x\in\Lambda} I_{A_x} \geq \delta/2\Big) \geq 1/2\,. \tag{22}$$

On the other hand, for geometrical reasons (which are intuitively obvious but need some work when made precise), there cannot be more triple points in Λ than points in the boundary $\partial\Lambda$ of Λ. Indeed, each of the three paths leaving a triple point meets $\partial\Lambda$, which gives three boundary points associated to each triple point in Λ. If one identifies these boundary points successively for one triple point after the other one sees that, at each step, at least one of the boundary points must be different from those obtained before. Hence,

$$|\Lambda|^{-1} \sum_{x\in\Lambda} I_{A_x} \leq |\Lambda|^{-1}|\partial\Lambda| < \delta/2$$

when Λ is large enough. Inserting this into (22) we arrive at the contradiction $\mu(\emptyset) \geq 1/2$, and the theorem is proved. For more details we refer to the original paper [42]. □

We stress that the last argument relies essentially on the amenability property of $\mathbf{Z}^d$ discussed in Section 5.2. The finite-energy condition is also indispensable: In another paper [43], Burton and Keane construct, for any $k \in \{2, 3, \ldots, \infty\}$, translation invariant percolation models on $\mathbf{Z}^2$ for which finite energy fails and which have exactly k infinite open clusters. For example, we have $k = \infty$ in Example 6.1. Fortunately, the finite-energy condition holds in most of the dependent percolation models which show up in stochastic-geometric studies of Gibbs measures.

The situation becomes radically different when $\mathbf{Z}^d$ is replaced by the non-amenable tree $\mathbf{T}_d$. Instead of having a unique infinite cluster, supercritical percolation models on $\mathbf{T}_d$ tend to produce infinitely many infinite clusters. It is not hard to verify that this is indeed the case for supercritical Bernoulli site or bond percolation (except in the trivial case when the retention probability p is 1), and a corresponding result for automorphism invariant percolation on $\mathbf{T}_d$ can be found in [118]. On more general nonamenable graph structures, the uniqueness of the infinite cluster property can fail in more interesting ways than on trees; see e.g. Grimmett and Newman [111] and Häggström and Peres [122].

Let us next consider the particular case of (possibly dependent) site percolation on $\mathbf{Z}^2$. We know from Theorem 5.17 that under fairly general assumptions there is almost surely at most one infinite open cluster. Under the same asumptions there is almost surely at most one infinite *closed* cluster (i.e., at most one infinite connected component of closed vertices). In fact, the proof of Theorem 5.17 even shows that almost surely there is at most one infinite closed $*$-cluster. (Here, a closed $*$-cluster is a maximal set C of closed sites which is $*$-connected, in that any two $x, y \in C$ are connected by a $*$-path in C; $*$-paths were introduced in the proof of Theorem 5.3. Any closed cluster is part of some closed $*$-cluster.) But perhaps an infinite open cluster and an infinite closed $*$-cluster can coexist? Theorem 5.18 below asserts that under reasonably general circumstances this cannot happen. Under slightly different conditions (replacing the finite-energy assumption by separate ergodicity under translations in the two coordinate directions), it was proved by Gandolfi, Keane and Russo [90].

Theorem 5.18 *Let μ be an automorphism invariant and ergodic probability measure on $\{0,1\}^{\mathbf{Z}^2}$ with finite energy and positive correlations. Then*

$$\mu(\exists \text{ infinite open cluster}, \exists \text{ infinite closed } *\text{-cluster}) = 0\,.$$

Note that automorphism invariance in the $\mathbf{Z}^2$-case means that, in addition to translation invariance, μ is also invariant under reflection in and exchange of coordinate axes. Under the conditions of the theorem, we have in fact some information on the geometric shape of infinite clusters: If an infinite open cluster exists and thus all closed $*$-clusters are finite, each finite box of $\mathbf{Z}^2$ is surrounded by an open circuit, and all these circuits are part of the (necessarily unique) infinite open cluster. Hence the infinite open cluster is a sea, in the sense that all "islands" (i.e., the $*$-clusters of its complement) are finite. Similarly, if a closed $*$-cluster in $\mathbf{Z}^2$ exists, it is necessarily a sea (and in particular unique).

The corresponding result is false in higher dimensions. To see this, consider Bernoulli site percolation on $\mathbf{Z}^3$. The critical value p_c for this model is strictly less than $1/2$ (see Campanino and Russo [47]), whence for $p = 1/2$ there exist almost surely both an infinite open cluster and an infinite closed cluster.

The proof of Theorem 5.18 below is based on a geometric argument of Yu Zhang who gave a new proof of Harris' [128] classical result that the critical value p_c for bond percolation on $\mathbf{Z}^2$ is at least $1/2$. (Recall that this bound is actually sharp.) Zhang's proof appeared first in [108] and was exploited later in other contexts in [117, 120].

Proof of Theorem 5.18: Let A be the event that there exists an infinite open cluster, let B be the event that there exists an infinite closed $*$-cluster, and assume by contradiction that $\mu(A \cap B) > 0$. Then, by ergodicity, $\mu(A \cap B) = 1$. (This is the only use of ergodicity we make, and ergodicity could clearly be replaced by tail triviality or some other mixing condition.) Next we pick n so large that

$$\mu(A_n) > 1 - 10^{-3} \text{ and } \mu(B_n) > 1 - 10^{-3} ,$$

where A_n (resp. B_n) is the event that some infinite open cluster (resp. some infinite closed $*$-cluster) intersects $\Lambda_n = [-n, n]^2 \cap \mathbf{Z}^2$. Let A_n^L (resp. A_n^R, A_n^T and A_n^B) be the event that some vertex in the left (resp. right, top and bottom) side of the square-shaped vertex set $\Lambda_n \setminus \Lambda_{n-1}$ belongs to some infinite open path which contains no other vertex of Λ_n, and define B_n^L, B_n^R, B_n^T and B_n^B analogously. Then

$$A_n = A_n^L \cup A_n^R \cup A_n^T \cup A_n^B .$$

Since all four events in the right hand side are increasing and μ has positive correlations,

$$\begin{aligned} \mu(A_n) &= \mu(A_n^L \cup A_n^R \cup A_n^T \cup A_n^B) \\ &= 1 - \mu(\neg A_n^L \cap \neg A_n^R \cap \neg A_n^T \cap \neg A_n^B) \\ &\leq 1 - \mu(\neg A_n^L)\mu(\neg A_n^R)\mu(\neg A_n^T)\mu(\neg A_n^B) , \end{aligned}$$

where $\neg$ indicates the complement of a set (for typographical reasons). By the automorphism invariance of μ, A_n^L, A_n^R, A_n^T and A_n^B all have the same μ-probability, so that

$$\mu(\neg A_n^L) \leq (1 - \mu(A_n))^{1/4}$$

and therefore

$$\mu(A_n^L) = \mu(A_n^R) \geq 1 - (1 - \mu(A_n))^{1/4} = 1 - 10^{-3/4} > 0.82 . \tag{23}$$

In the same way, we get

$$\mu(B_n^T) = \mu(B_n^B) > 0.82 . \tag{24}$$

Now define the event $D = A_n^L \cap A_n^R \cap B_n^T \cap B_n^B$. From (23) and (24) we obtain

$$\mu(D) \geq 1 - 4(1 - 0.82) = 0.28 > 0 .$$

When D occurs, both the left side and the right side of Λ_n are intersected by some infinite open cluster. By Theorem 5.17, these infinite open clusters are identical and separate their (common) complement into (at least) two pieces, preventing the infinite closed $*$-clusters intersecting the top and bottom sides of Λ_n from reaching each other (see the picture on p. 196 of [108]). Consequently, there exist two infinite closed $*$-clusters, in contradiction to Theorem 5.17. □

Theorem 5.18 admits some variants. First, the assumption of ergodicity can be avoided if the assumption of positive correlations is strenghtened to the condition that μ is monotone in the sense of Definition 4.9. (This is because monotonicity is preserved under ergodic decomposition, so that Theorem 4.11 implies positive correlations for each ergodic component.) Moreover, as the preceding proof shows, ergodicity is only needed to show that infinite clusters exist with probability either 0 or 1, and translation invariance and finite energy are only used for the uniqueness of infinite clusters. We also need only the invariance under lattice rotations rather than all reflections, and closed $*$-clusters can be replaced by closed clusters. We may thus state the following result.

Proposition 5.19 *There exists no probability measure μ on $\{0,1\}^{\mathbf{Z}^2}$ which has positive correlations, is invariant under lattice rotations and the interchange of the states 1 ("open") and 0 ("closed"), and satisfies*

$$\mu(\,\exists \text{ a unique infinite open cluster}) = 1\ .$$

This proposition will be applied to the ferromagnetic Ising model in Section 8.2.

6 Random-cluster representations

In the previous section we saw a number of dependent percolation models. Here we shall focus on a particular class of such models, namely the (Fortuin–Kasteleyn) random-cluster model (and some of its relatives), which has turned out to be of great value in analyzing the phase transition behavior of Ising and Potts models. An alternative source for much of the material in the present section is Häggström [119]. In Sections 6.1 and 6.2, we introduce the random-cluster model and discuss its relation to Ising and Potts models. This relation is then applied in Section 6.3 to prove Theorems 3.1 and 3.2. Despite the fact that Theorems 3.1 and 3.2 concern infinite systems, these applications only require defining finite volume random-cluster measures. However, it may be interesting in its own right to study infinite volume random-cluster measures on graphs such as $\mathbf{Z}^d$; this is done in Section 6.4. In Section 6.6, we describe how the random-cluster representation of Ising and Potts models can be used to construct highly efficient Monte Carlo simulation algorithms. Finally, in Section 6.7, we discuss a variant of the random-cluster model which is applicable to the Widom–Rowlinson model rather than to Ising and Potts models.

6.1 Random-cluster and Potts models

The random-cluster model, also known as the Fortuin–Kasteleyn (FK) model after its inventors [82, 80, 81], is a two-parameter family of dependent bond percolation models living on a finite graph. Let $G = (\mathcal{L}, \sim)$ be a finite graph with vertex set $\mathcal{L}$ and edge set $\mathcal{B}$. For a bond configuration $\eta \in \{0,1\}^{\mathcal{B}}$, we write $k(\eta)$ for the number of connected components (including isolated vertices) in the subgraph of G containing all vertices but only the open edges (i.e. those $e \in \mathcal{B}$ for which $\eta(e) = 1$).

Definition 6.1 *The* **random-cluster measure** $\phi^G_{p,q}$ *for* G *with parameters* $p \in [0,1]$ *and* $q > 0$ *is the probability measure on* $\{0,1\}^{\mathcal{B}}$ *which to each* $\eta \in \{0,1\}^{\mathcal{B}}$ *assigns probability*

$$\phi^G_{p,q}(\eta) = \frac{1}{Z^G_{p,q}} \left\{ \prod_{e \in \mathcal{B}} p^{\eta(e)}(1-p)^{1-\eta(e)} \right\} q^{k(\eta)},$$

where $Z^G_{p,q}$ *is a normalizing constant.*

Note that taking $q = 1$ yields the Bernoulli bond percolation measure ϕ_p defined at the end of Section 5.1. All other choices of q give rise to dependencies between edges (as long as p is not 0 or 1, and G is not a tree).

Taking $q \in \{2, 3, \ldots\}$ yields a model which is intimately related to the q-state Potts model, in a way which we will explain now. Let $\mu^G_{\beta,q}$ be the Gibbs measure for the q-state Potts model on G at inverse temperature β, i.e. $\mu^G_{\beta,q}$ is the measure on $\{1, \ldots, q\}^{\mathcal{L}}$ which to each $\sigma \in \{1, \ldots, q\}^{\mathcal{L}}$ assigns probability

$$\mu^G_{q,\beta}(\sigma) = \frac{1}{Z^G_{\beta,q}} \exp\left(-2\beta \sum_{x \sim y} I_{\{\sigma(x) \neq \sigma(y)\}} \right),$$

where again $Z^G_{\beta,q}$ is a normalizing constant.

For $q \in \{2, 3, \ldots\}$ and $p \in [0,1]$, let $P^G_{p,q}$ be the probability measure on $\{1, \ldots, q\}^{\mathcal{L}} \times \{0,1\}^{\mathcal{B}}$ corresponding to picking a random element of $\{1, \ldots, q\}^{\mathcal{L}} \times \{0,1\}^{\mathcal{B}}$ according to 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.

In other words, $P^G_{p,q}$ is the measure which to each element (σ, η) of $\{1, \ldots, q\}^{\mathcal{L}} \times \{0,1\}^{\mathcal{B}}$ assigns probability proportional to

$$\prod_{e=\langle xy\rangle \in \mathcal{B}} \left(p^{\eta(e)}(1-p)^{1-\eta(e)} \; I_{\{(\sigma(x)-\sigma(y))\eta(e)=0\}} \right) \, .$$

Here, $\langle xy \rangle$ denotes the edge linking x and y. The measure $P^G_{p,q}$ was introduced by Swendsen and Wang [219] and made more explicit by Edwards and Sokal [73], and is therefore called the *Edwards–Sokal measure*. The following theorem states that the edge marginal of $P^G_{p,q}$ is a random-cluster measure, and the vertex marginal is a Gibbs measure for the Potts model, meaning that $P^G_{p,q}$ is a coupling of $\mu^G_{\beta,q}$ and $\phi^G_{p,q}$.

Theorem 6.2 *Let $P^{G,vertex}_{p,q}$ and $P^{G,edge}_{p,q}$ be the probability measures obtained by projecting $P^G_{p,q}$ on $\{1, \ldots, q\}^{\mathcal{L}}$ and $\{0,1\}^{\mathcal{B}}$, respectively. Then*

$$P^{G,vertex}_{p,q} = \mu^G_{\beta,q} \tag{25}$$

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

$$P^{G,edge}_{p,q} = \phi^G_{p,q} \, . \tag{26}$$

Proof: The proof is just a matter of summing out the marginals. Letting Z be the normalizing constant in $P^G_{p,q}$, fixing $\sigma \in \{1, \ldots, q\}^{\mathcal{L}}$, and summing over all $\eta \in \{0,1\}^{\mathcal{B}}$ we find

$$\begin{aligned}
P^{G,vertex}_{p,q}(\sigma) &= \sum_{\eta \in \{0,1\}^{\mathcal{B}}} P^G_{p,q}(\sigma, \eta) \\
&= \frac{1}{Z} \sum_{\eta \in \{0,1\}^{\mathcal{B}}} \prod_{e=\langle xy\rangle \in \mathcal{B}} p^{\eta(e)}(1-p)^{1-\eta(e)} I_{\{(\sigma(x)-\sigma(y))\eta(e)=0\}} \\
&= \frac{1}{Z} \prod_{e=\langle xy\rangle \in \mathcal{B}} (1-p)^{I_{\{\sigma(x)\neq\sigma(y)\}}} \\
&= \frac{1}{Z} \exp\Big(-2\beta \sum_{x \sim y} I_{\{\sigma(x)\neq\sigma(y)\}} \Big) \\
&= \mu^G_{\beta,q}(\sigma) \, ,
\end{aligned}$$

since Z must be equal to $Z^{\beta,q}_G$ by normalization. This proves (25). To verify (26) we proceed similarly, fixing $\eta \in \{0,1\}^{\mathcal{B}}$ and summing over $\sigma \in \{1, \ldots, q\}^{\mathcal{L}}$. Note that, given η, there are exactly $q^{k(\eta)}$ spin configurations σ that are allowed, in that any two neighboring vertices $x \sim y$ with $\eta(\langle xy \rangle) = 1$ have the same spin. We get

$$\begin{aligned}
P^{G,edge}_{p,q}(\eta) &= \sum_{\sigma \in \{1,\ldots,q\}^{\mathcal{L}}} P^G_{p,q}(\sigma, \eta) \\
&= q^{k(\eta)} \frac{1}{Z} \prod_{e \in \mathcal{B}} p^{\eta(e)}(1-p)^{1-\eta(e)} \\
&= \mu^G_{p,q}(\eta) \, ,
\end{aligned}$$

again by normalization. □

The Edwards–Sokal coupling $P^G_{p,q}$ of $\mu^G_{\beta,q}$ and $\phi^G_{p,q}$ is the key to using the random-cluster model in analyzing the Potts model. The following two results are each other's dual, and are immediate consequences of Theorem 6.2 and the definition of $P^G_{p,q}$.

Corollary 6.3 *Let $p = 1 - e^{-2\beta}$, and suppose we pick a random spin configuration $X \in \{1, \ldots, q\}^{\mathcal{L}}$ as follows:*

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

Then X is distributed according to the Gibbs measure $\mu^G_{\beta,q}$.

Corollary 6.4 *Let $p = 1 - e^{-2\beta}$, and suppose we pick a random edge configuration $Y \in \{0,1\}^{\mathcal{B}}$ as follows:*

1. *Pick a random spin configuration $X \in \{1, \ldots, q\}^{\mathcal{L}}$ according to the Gibbs measure $\mu^G_{\beta,q}$.*
2. *Given X, assign each edge $e = \langle xy \rangle$ 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 $\phi^G_{p,q}$.

As a warm-up for the phase transition considerations in Section 6.3, we give the following result as a typical application of the random-cluster representation.

Corollary 6.5 *If we pick a random spin configuration $X \in \{1, \ldots, q\}^{\mathcal{L}}$ according to the Gibbs measure $\mu^G_{\beta,q}$, then for $i \in \{1, \ldots, q\}$ and two vertices $x, y \in \mathcal{L}$, the two events $\{X(x) = i\}$ and $\{X(y) = i\}$ are positively correlated, i.e.*

$$\mu^G_{\beta,q}(X(x) = i, X(y) = i) \geq \mu^G_{\beta,q}(X(x) = i)\, \mu^G_{\beta,q}(X(y) = i)\,.$$

Proof: The measure $\mu^G_{\beta,q}$ is invariant under permutation of the spin set $\{1, \ldots, q\}$, so that

$$\mu^G_{\beta,q}(X(x) = i) = \mu^G_{\beta,q}(X(y) = i) = \frac{1}{q}\,.$$

We therefore need to show that

$$\mu^G_{\beta,q}(X(x) = i, X(y) = i) \geq \frac{1}{q^2}\,.$$

We may now think of X as being obtained as in Corollary 6.3 by first picking an edge configuration $Y \in \{0,1\}^{\mathcal{B}}$ according to the random-cluster measure $\phi^G_{p,q}$ and then assigning i.i.d. uniform spins to the connected components. Given Y, the conditional probability that $X(x) = X(y) = i$ is $1/q$ if u and v are in the same connected component of Y, and $1/q^2$ if they are in different connected components. Hence, for some $\alpha \in [0,1]$,

$$\mu^G_{\beta,q}(X(x) = i, X(y) = i) = \alpha\frac{1}{q} + (1-\alpha)\frac{1}{q^2} \geq \frac{1}{q^2}\,.$$

□

An easy modification of the above proof shows that if G is connected and $\beta > 0$, then the correlation between $I_{\{X(x)=i\}}$ and $I_{\{X(y)=i\}}$ is in fact *strictly* positive.

Note that the relation between the random-cluster model and the Potts model depends crucially on the fact that all spins in $\{1,\ldots,q\}$ are *a priori* equivalent. This is no longer the case when a nonzero external field is present in the Ising model. Several attempts to find useful random-cluster representations of the Ising model with external field have been made, but progress has been limited. Perhaps the recent duplication idea of Chayes, Machta and Redner [60] represents a breakthrough on this problem.

6.2 Infinite-volume limits

In this subsection we will exploit some stochastic monotonicity properties of random-cluster distributions on finite subgraphs of $\mathbf{Z}^d$. This will give us the existence of certain limiting random-cluster distributions, and also the existence of certain limiting Gibbs measures for the Potts model.

The basic observation is stated in the lemma below which follows directly from definitions.

Lemma 6.6 *Consider the random-cluster model with parameters p and q on a finite graph G with edge set $\mathcal{B}$. For any edge $e = \langle xy\rangle \in \mathcal{B}$, and any configuration $\eta \in \{0,1\}^{\mathcal{B}\setminus\{e\}}$, we have that*

$$\phi^G_{p,q}(e \text{ is open} \,|\, \eta) = \begin{cases} p & \text{if } x \text{ and } y \text{ are connected via open edges in } \eta \\ \frac{p}{p+(1-p)q} & \text{otherwise.} \end{cases} \tag{27}$$

For $q \geq 1$, Lemma 6.6 means in particular that the conditional probability in (27) is increasing in η (and also in p). This allows us to use Holley's Theorem and the FKG inequality to prove the following very useful result. We write ϕ^G_p for Bernoulli bond percolation on G with parameter p.

Corollary 6.7 *For a finite graph G and the random-cluster measure $\phi^G_{p,q}$ with $p \in [0,1]$ and $q \geq 1$, we have*

(a) $\phi^G_{p,q}$ *is monotone, and therefore it has positive correlations,*

(b) $\phi^G_{p,q} \preceq_{\mathcal{D}} \phi^G_p$,

(c) $\phi^G_{p,q} \succeq_{\mathcal{D}} \phi^G_{\frac{p}{p+(1-p)q}}$.

Furthermore, for $0 \leq p_1 \leq p_2 \leq 1$ and $q \geq 1$, we have

(d) $\phi^G_{p_1,q} \preceq_{\mathcal{D}} \phi^G_{p_2,q}$.

Proof: The monotonicity in (a) is just the observation that the conditional probability in (27) is increasing in p and in η. Positive correlations then follows from Theorem 4.11. Next, note that (27) implies that

$$\frac{p}{p+(1-p)q} \leq \phi^G_{p,q}(e \text{ is open} \,|\, \eta) \leq p \tag{28}$$

for all η as in Lemma 6.6. Theorem 4.8 in conjunction with the second (resp. first) inequality in (28) implies (b) (resp. (a)). Finally, (d) is just another application of (d) and Theorem 4.8. □

Consider now the integer lattice $\mathbf{Z}^d$ (for definiteness and simplicity) with its usual graph structure. We associate with any finite region $\Lambda \subset \mathbf{Z}^d$ two specific random-cluster distributions which correspond to two different choices of boundary condition. The latter will we distinguished by a parameter $b \in \{0,1\}$. Let $\mathcal{B}$ be the set of all nearest-neighbor bonds in $\mathbf{Z}^d$, $\mathcal{B}^0_\Lambda$ the set of all edges of $\mathcal{B}$ that are contained in Λ, and $\mathcal{B}^1_\Lambda$ the set of edges with at least one endpoint in Λ. (The difference $\mathcal{B}^1_\Lambda \setminus \mathcal{B}^0_\Lambda$ thus consists of all edges leading from a point of Λ to a point of Λ^c.) We then let $\phi^b_{p,q,\Lambda}$ be the probability measure on $\{0,1\}^{\mathcal{B}}$ in which each $\eta \in \{0,1\}^{\mathcal{B}}$ is assigned probability proportional to

$$I_{\{\eta \equiv b \text{ off } \mathcal{B}^b_\Lambda\}} \left\{ \prod_{e \in \mathcal{B}^b_\Lambda} p^{\eta(e)} (1-p)^{1-\eta(e)} \right\} q^{k(\eta,\Lambda)} ,$$

where $k(\eta,\Lambda)$ is the number of all η-open clusters meeting Λ. We call $\phi^b_{p,q,\Lambda}$ the random-cluster distribution in Λ with parameters p and q and boundary condition b. In the case $b = 0$, $k(\cdot,\Lambda)$ is simply the number of all clusters that are contained in Λ; this corresponds to forgetting all sites in Λ^c and is therefore referred to as the *free boundary condition*. On the other hand, suppose that Λ has no holes, in the sense that Λ^c has no finite connected components; since we can always assume without loss of generality that Λ is connected, we call such a Λ *simply connected*. Then, in the case $b = 1$, all sites of Λ^c may be thought of as being firmly wired together, whence this is called the *wired boundary condition*.

Suppose now that $\Lambda \subset \Delta$ are two finite regions in $\mathbf{Z}^d$. Then $\phi^b_{p,q,\Lambda}$ is obtained from $\phi^b_{p,q,\Delta}$ by conditioning on the event $\{\eta \equiv b \text{ on } \mathcal{B}^b_\Delta \setminus \mathcal{B}^b_\Lambda\}$ which is increasing for $b = 1$ and decreasing for $b = 0$. Hence, if $q \geq 1$ then Corollary 6.7 (a) implies that

$$\phi^0_{p,q,\Lambda} \preceq_{\mathcal{D}} \phi^0_{p,q,\Delta} \text{ and } \phi^1_{p,q,\Lambda} \succeq_{\mathcal{D}} \phi^1_{p,q,\Delta} \text{ when } \Lambda \subset \Delta \,, \tag{29}$$

in complete analogy to (14). Moreover, we obtain the following counterpart of Proposition 4.14 on the existence of infinite-volume limits. We write $\Lambda \uparrow \mathbf{Z}^d$ for the limit along some (any) increasing sequence of finite simply connected subsets of $\mathbf{Z}^d$, converging to $\mathbf{Z}^d$ in the usual way.

Lemma 6.8 *For $p \in [0,1]$ and $q \geq 1$, the limiting measures*

$$\phi^b_{p,q} = \lim_{\Lambda \uparrow \mathbf{Z}^d} \phi^b_{p,q,\Lambda} \,, \quad b \in \{0,1\} \,,$$

exist and are translation invariant.

This convergence result has consequences for the convergence of Gibbs distributions for the Potts model, as we will show next. Let $q \in \{2,3,\ldots\}$, and for $i \in \{1,\ldots,q\}$ and any finite region Λ in $\mathbf{Z}^d$ let $\mu^i_{\beta,q,\Lambda}$ denote the Gibbs distribution in Λ for the Potts model at inverse temperature β with boundary condition $\eta \equiv i$ on Λ^c. For $i = 0$, let $\mu^0_{\beta,q,\Lambda}$ be the corresponding Gibbs distribution with free boundary condition, which is defined by letting $\mathcal{L} = \Lambda$ in (4), i.e., by ignoring all sites outside Λ; we think of $\mu^0_{\beta,q,\Lambda}$ as a probability measure on the full configuration space $\{1,\dots,q\}^{\mathbf{Z}^d}$ by using an arbitrary extension.

Still for $i = 0$, Theorem 6.2 shows that $\mu^0_{\beta,q,\Lambda}$ and $\phi^0_{p,q,\Lambda}$ admit an Edwards–Sokal coupling (on Λ) when $p = 1 - e^{-2\beta}$. A similar Edwards–Sokal coupling is possible for $i \in \{1,\dots,q\}$ when Λ is simply connected. Indeed, let $P^i_{p,q,\Lambda}$ be the probability measure on $\{1,\dots,q\}^{\mathcal{L}} \times \{0,1\}^{\mathcal{B}}$ corresponding to picking a random site-and-bond configuration according to the following procedure.

1. Assign to each vertex of Λ^c value i, and to all edges of $\mathcal{B} \setminus \mathcal{B}^1_\Lambda$ value 1.
2. Assign to each vertex in Λ a spin value chosen from $\{1,\dots,q\}$ according to uniform distribution, assign to each edge in $\mathcal{B}^1_\Lambda$ value 1 or 0 with respective probabilities p and $1-p$, and do this independently for all vertices and edges.
3. Condition on the event that no two vertices with different spins have an edge with value 1 connecting them.

It is now a simple modification of the proof of Theorem 6.2 to check that the vertex and edge marginals of $P^i_{p,q,\Lambda}$ are $\mu^i_{\beta,q,\Lambda}$ and $\phi^1_{p,q,\Lambda}$, respectively. (Note that by the simple connectedness of Λ there is always a unique component containing Λ^c.) Analogues of Corollaries 6.3 and 6.4 follow easily. This leads us to the following result extending Proposition 4.14 to the Potts model.

Proposition 6.9 *For any $i \in \{0,1,\dots,q\}$, the limiting probability measure*

$$\mu^i_{\beta,q} = \lim_{\Lambda \uparrow \mathbf{Z}^d} \mu^i_{\beta,q,\Lambda}$$

on $\{1,\dots,q\}^{\mathbf{Z}^d}$ exists and is a translation invariant Gibbs measure for the q-state Potts model on $\mathbf{Z}^d$ at inverse temperature β.

Proof: In view of the general facts reported in Section 2.6, the limits are Gibbs measures whenever they exist. We thus need to show that $\mu^i_{\beta,q,\Lambda}(f)$ converges as $\Lambda \uparrow \mathbf{Z}^d$, for any local observable f. For definiteness, we do this for $i \in \{1,\dots,q\}$; the case $i = 0$ is completely similar.

Fix an f as above, and let $\Delta \subset \mathbf{Z}^d$ be the finite region on which f depends. As shown above, for a simply connected Λ we may think of a $\{1,\dots,q\}^{\mathbf{Z}^d}$-valued random element X with distribution $\mu^i_{\beta,q,\Lambda}$ as arising by first picking an edge configuration $Y \in \{0,1\}^\Lambda$ according to $\phi^1_{p,q,\Lambda}$ (with $p = 1 - e^{-2\beta}$) and then assigning random spins to the connected components, forcing spin i to the (unique) infinite cluster. For $x, y \in \Delta$, we write $\{x \leftrightarrow y\}$ for the event that x and y are in the same connected component in Y, and $\{x \leftrightarrow \infty\}$ for the event that x is in an infinite cluster. Clearly, the conditional distribution of f given Y depends only on the indicator functions $(I_{\{x\leftrightarrow y\}})_{x,y\in\Delta}$ and $(I_{\{x\leftrightarrow\infty\}})_{x\in\Delta}$, since the conditional distribution of X on Δ is uniform over all elements of $\{1,\dots,q\}^\Delta$ such that firstly $X(x) = X(y)$ whenever $x \leftrightarrow y$, and secondly $X(x) = i$ whenever $x \leftrightarrow \infty$. Hence, the desired convergence of $\mu^i_{\beta,q,\Lambda}(f)$ follows if we can show that the joint distribution of $(I_{\{x\leftrightarrow y\}})_{x,y\in\Delta}$ and $(I_{\{x\leftrightarrow\infty\}})_{x\in\Delta}$ converges as $n\to\infty$. This, however, follows from Lemma 6.8 upon noting that $(I_{\{x\leftrightarrow y\}})_{x,y\in\Delta}$ and $(I_{\{x\leftrightarrow\infty\}})_{x\in\Delta}$ are increasing functions. □

6.3 Phase transition in the Potts model

As promised, this subsection is devoted to proving Theorems 3.1 and 3.2, using random-cluster arguments. The original source for the material in this subsection is Aizenman, Chayes, Chayes and Newman [9]; see also [119] for a slightly different presentation.

We consider the Potts model on $\mathbf{Z}^d$, $d\geq 2$. All the arguments to be used here, except those showing that the critical inverse temperature β_c is strictly between 0 and ∞, go through on arbitrary infinite graphs; we stick to the $\mathbf{Z}^d$ case for definiteness and simplicity of notation. We consider the limiting Gibbs measures $\mu^i_{\beta,q}$ obtained in Proposition 6.9. For $i\in\{1,\dots,q\}$, these play a role similar to that of the "plus" and "minus" measures μ^+_β and μ^-_β for the Ising model. In fact, we have the following result which extends Theorem 4.15 to the Potts model and also gives a characterization of phase transition in terms of percolation in the random-cluster model.

Theorem 6.10 *Let $\beta>0$ and $p=1-e^{-2\beta}$. For any $x\in\mathbf{Z}^d$ and any $i\in\{1,\dots,q\}$, the following statements are equivalent.*

(i) *There is a unique Gibbs measure for the q-state Potts model on $\mathbf{Z}^d$ at inverse temperature β.*

(ii) $\mu^i_{\beta,q}(X(x)=i)=1/q$.

(iii) $\phi^1_{p,q}(x\leftrightarrow\infty)=0$.

As we will see in a moment, it is the percolation criterion (iii) which is most convenient to apply. In this context we note that

$$\phi^1_{p,q}(x\leftrightarrow\infty)=\inf_{\Lambda,\Delta}\phi^1_{p,q,\Lambda}(x\leftrightarrow\Delta^c)=\lim_{\Lambda\uparrow\mathbf{Z}^d}\phi^1_{p,q,\Lambda}(x\leftrightarrow\Lambda^c)\ , \qquad (30)$$

where $\{x\leftrightarrow\Delta^c\}$ stands for the event that there exists an open path from x to some site in Δ^c. This follows from (29) and the fact that $\{x\leftrightarrow\Delta^c\}$ decreases to $\{x\leftrightarrow\infty\}$ as $\Delta\uparrow\mathbf{Z}^d$.

The usefulness of the percolation criterion is demonstrated by the next result which extends the scenario for Bernoulli percolation to the random-cluster model. Together with Theorem 6.10, this gives Theorem 3.2 with $\beta_c=\frac{1}{2}\log(1-p_c)$.

Proposition 6.11 *For the random-cluster model on $\mathbf{Z}^d$, $d\geq 2$, and any fixed $q\geq 1$, there exists a percolation threshold $p_c\in(0,1)$ (depending on d and q) such that*

$$\phi^1_{p,q}(x\leftrightarrow\infty)\begin{cases}=0 & \text{for } p<p_c\ ,\\ >0 & \text{for } p>p_c\ .\end{cases}$$

Proof: The statement of the proposition consists of the following three parts:

(i) $\phi^1_{p,q}(x\leftrightarrow\infty)=0$ for p sufficiently small,

(ii) $\phi^1_{p,q}(x\leftrightarrow\infty)>0$ for p sufficiently close to 1, and

(iii) $\phi^1_{p,q}(x \leftrightarrow \infty)$ is increasing in p.

We first prove (i). Suppose $p < p_c(\mathbf{Z}^d, \text{bond})$, the critical value for Bernoulli bond percolation on $\mathbf{Z}^d$. For $\varepsilon > 0$, we can then pick Δ large enough so that

$$\phi_p(0 \leftrightarrow \Delta^c) \leq \varepsilon\,.$$

By Corollary 6.7 (b), we have that the projection of ϕ_p on $\{0,1\}^{\mathcal{B}^1_\Delta}$ stochastically dominates the projection of $\phi^1_{p,q,\Lambda}$ on $\{0,1\}^{\mathcal{B}^1_\Delta}$ for any $\Lambda \supset \Delta$, so that

$$\phi^1_{p,q,\Lambda}(0 \leftrightarrow \infty) \leq \phi^1_{p,q,\Lambda}(0 \leftrightarrow \Delta^c) \leq \varepsilon$$

for any $\Lambda \supset \Delta$. Since ε was arbitrary, we find

$$\lim_{\Lambda \uparrow \mathbf{Z}^d} \phi^1_{p,q,\Lambda}(0 \leftrightarrow \infty) = 0$$

which in conjunction with (30) implies (i).

Next, (ii) can be established by a similar argument: Let p be such that $p^* = p/[p + (1-p)q] > p_c(\mathbf{Z}^d, \text{bond})$. Corollary 6.7 (c) then shows that $\phi^1_{p,q,\Lambda} \succeq_{\mathcal{D}} \phi_{p^*}$ for every Λ, so that

$$\lim_{\Lambda \uparrow \mathbf{Z}^d} \phi^1_{p,q,\Lambda}(0 \leftrightarrow \infty) > 0,$$

proving (ii).

To check (iii) we note that Corollary 6.7 (d) implies that, for any Λ,

$$\phi^1_{p_1,q,\Lambda} \preceq_{\mathcal{D}} \phi^1_{p_2,q,\Lambda} \text{ whenever } p_1 \leq p_2. \tag{31}$$

This proves (iii) and thereby the proposition. □

Before the proof of Theorem 6.10 we need another definition and a couple of lemmas. For a finite box Λ in $\mathbf{Z}^d$ and a spin configuration $\xi \in \{1,\ldots,q\}^{\partial\Lambda}$, let

$$A^i_\xi = \{x \in \partial\Lambda : \xi(x) = i\}$$

for $i = 1,\ldots,q$. We now define the *random-cluster distribution* $\phi^\xi_{p,q,\Lambda}$ *for* Λ *with boundary condition* ξ, as the probability measure on $\{0,1\}^{\mathcal{B}^1_\Lambda}$ which to each $\eta \in \{0,1\}^{\mathcal{B}^1_\Lambda}$ assigns probability proportional to

$$I_{D(\xi,\eta)} \left\{ \prod_{e \in \mathcal{B}^1_\Lambda} p^{\eta(e)}(1-p)^{1-\eta(e)} \right\} q^{k_\xi(\eta)}$$

where $k_\xi(\eta)$ is the number of connected components in η that do not intersect $\partial\Lambda$, and $D(\xi,\eta)$ is the event that there is no open path in η connecting any two vertices in A^i_ξ and A^j_ξ for any $i \neq j$.

Lemma 6.12 *Let $p = 1 - e^{-2\beta}$, let Λ be a finite region in $\mathbf{Z}^d$, and fix some boundary condition $\xi \in \{1,\ldots,q\}^{\partial\Lambda}$. Suppose that we pick a random spin configuration $X \in \{1,\ldots,q\}^\Lambda$ as follows.*

1. *Pick $Y \in \{0,1\}^{\mathcal{B}^1_\Lambda}$ according to $\phi^\xi_{p,q,\Lambda}$.*

2. *For each $i \in \{1,\ldots,q\}$ and each connected component C in η intersecting A^i_ξ, assign spin i to every vertex in C.*

3. *For all other connected components C in Y, pick a spin at random (uniformly) from $\{1,\ldots,q\}$, assign this spin to every vertex of C, and do this independently for different connected components.*

Then X is distributed according to the Gibbs distribution $\mu^\xi_{\beta,q,\Lambda}$ for the Potts model on Λ with boundary condition ξ.

Proof: This is a straightforward generalization of the proofs of Theorem 6.2 and Corollary 6.3. □

Lemma 6.13 *With notation as above, we have, for any $\xi \in \{1,\ldots,q\}^{\partial\Lambda}$, that the projection of $\phi^1_{p,q,\Lambda}$ on $\{0,1\}^{\mathcal{B}^1_\Lambda}$ stochastically dominates $\phi^\xi_{p,q,\Lambda}$.*

Proof: Just write down single-edge conditional distributions for $\phi^\xi_{p,q,\Lambda}$ and $\phi^1_{p,q,\Lambda}$ (as in Lemma 6.6) and invoke Theorem 4.8. □

We are finally ready for the proof of Theorem 6.10.

Proof of Theorem 6.10: We begin with the implication (i) ⇒ (ii). If there is a unique Gibbs measure for the q-state Potts model on $\mathbf{Z}^d$ at inverse temperature β, then we have in particular that

$$\mu^1_{\beta,q} = \ldots = \mu^q_{\beta,q}\,.$$

But since by symmetry $\mu^i_{\beta,q}(X(x)=i) = \mu^j_{\beta,q}(X(x)=j)$ for any $i,j \in \{1,\ldots,q\}$, we must have $\mu^i_{\beta,q}(X(x)=i) = 1/q$, and (ii) is established.

Next we turn to the implication (ii) ⇒ (iii). By the Edwards–Sokal coupling of edge and site processes introduced before Proposition 6.9, we have

$$\begin{aligned}\mu^i_{\beta,q}(X(0)=i) &= \lim_{\Lambda\uparrow\mathbf{Z}^d}\mu^i_{\beta,q,\Lambda}(X(0)=i)\\ &= \frac{1}{q} + \frac{q-1}{q}\lim_{\Lambda\uparrow\mathbf{Z}^d}\phi^1_{p,q,\Lambda}(0\leftrightarrow\Lambda^c)\,. \qquad (32)\end{aligned}$$

Together with (30), the result follows.

Most of the work is needed for the implication (iii) ⇒ (i). Roughly speaking, the absence of percolation in the random-cluster model implies that every finite region is cut off from infinity by a set of closed edges. Thus, independently of what happens macroscopically, the local spins feel as if they are in a system with free boundary condition. This makes a phase transition impossible.

To make this intuition precise we let μ be an arbitrary Gibbs measure for the Potts model at inverse temperature β. We will show that $\mu = \mu^0_{\beta,q}$, the limiting measure with free boundary condition obtained in Proposition 6.9. We fix any local observable f and some $\varepsilon > 0$. We then can find a finite box $\Delta \subset \mathbf{Z}^d$ such that

$$\left|\,\mu^0_{\beta,q,\Gamma}(f) - \mu^0_{\beta,q}(f)\,\right| < \varepsilon \ \text{ for all finite } \Gamma \supset \Delta\,. \qquad (33)$$

In view of (30) we can also choose a finite box $\Lambda \supset \Delta$ satisfying $\phi^1_{p,q,\Lambda}(x \leftrightarrow \Lambda^c) < \varepsilon/|\Delta|$ for all $x \in \Delta$, and thus

$$\phi^1_{p,q,\Lambda}(\Delta \leftrightarrow \Lambda^c) < \varepsilon\,.$$

Here, $\{\Delta \leftrightarrow \Lambda^c\}$ is the event that there exists an open path from Δ to Λ^c.

Consider the complementary event $C = \{\Delta \leftrightarrow \Lambda^c\}^c$. For any edge configuration $\eta \in C$, Δ is cut off from Λ^c by a set of closed edges. Indeed, let Γ be the union of all η-open clusters meeting Δ. Then

(a) $\Delta \subset \Gamma \subset \Lambda$, and

(b) $\eta(e) = 0$ for all edges from Γ to Γ^c, i.e., $\eta \equiv 0$ on $\mathcal{B}^1_\Gamma \setminus \mathcal{B}^0_\Gamma$.

Since these properties are stable under finite unions, there exists a largest set $\Gamma(\eta)$ satisfying (a) and (b). The maximality implies that, for each fixed region Γ, the event $\{\eta : \Gamma(\eta) = \Gamma\}$ only depends on the status of the edges off $\mathcal{B}^0_\Gamma$.

Coming to the core of the argument, we fix any boundary spin configuration $\xi \in \{1, \ldots, q\}^{\partial\Lambda}$ and consider the Edwards–Sokal coupling $P = P^\xi_{p,q,\Lambda}$ of $\mu^\xi_{\beta,q,\Lambda}$ and $\phi^\xi_{p,q,\Lambda}$ introduced in Lemma 6.12. We write X for the random spin configuration in $\{1, \ldots, q\}^{\mathbf{Z}^d}$ and Y for the random edge configuration in $\{0,1\}^{\mathcal{B}}$. Then $P(f \circ X) = \mu^\xi_{\beta,q,\Lambda}(f)$, and

$$P(Y \in C) = \phi^\xi_{p,q,\Lambda}(C) \geq \phi^1_{p,q,\Lambda}(C) > 1 - \varepsilon$$

by Lemma 6.13 and the choice of Λ. Assuming without loss of generality that $\|f\| \leq 1$, we can therefore conclude that

$$\left| P(f \circ X \,|\, Y \in C) - \mu^\xi_{\beta,q,\Lambda}(f) \right| < 2\varepsilon \,. \tag{34}$$

However, the conditional expectation on the left is an average of the conditional expectations $P(f \circ X \,|\, \Gamma(Y) = \Gamma)$ with $\Delta \subset \Gamma \subset \Lambda$, and these in turn are averages of conditional expectations of the form

$$P(f \circ X \mid Y = \eta \text{ off } \mathcal{B}^1_\Gamma,\ Y \equiv 0 \text{ on } \mathcal{B}^1_\Gamma \setminus \mathcal{B}^0_\Gamma)$$

which, by construction of P, are equal to $\mu^0_{\beta,q,\Gamma}(f)$. Together with (33) and (34), we conclude that

$$\left| \, \mu^\xi_{\beta,q,\Lambda}(f) - \mu^0_{\beta,q}(f) \, \right| < 3\varepsilon \,.$$

Taking the μ-average over ξ and letting $\varepsilon \to 0$ we finally get $\mu(f) = \mu^0_{\beta,q}(f)$, and the proof is complete. □

6.4 Infinite volume random-cluster measures

The random-cluster arguments used in the previous subsections for studying infinite volume Ising and Potts models only required defining finite volume random-cluster distributions, although we have seen already the limiting random-cluster measures $\phi^b_{p,q}$. (Their existence was convenient in the formulation of Theorem 6.10, but not really needed for the arguments.) Recent years have nevertheless witnessed a rapid development of a theory for infinite volume random-cluster measures, defined in the DLR spirit. Here we shall discuss the basics of such a theory. A similar theory of infinite volume Edwards–Sokal measures for joint spin and edge distributions was recently developed in [31].

Let G be an infinite (locally finite) graph with vertex set $\mathcal{L}$ and edge set $\mathcal{B}$. Fix $p \in [0,1]$ and $q > 0$, and let $B \subset \mathcal{B}$ be a finite simply connected region; since the random-cluster model lives on edges rather than on vertices, we let "region" refer to edge sets rather than vertex sets in this subsection. Let $V(B) = \{x \in \mathcal{L} : \exists\, e \in B \text{ incident to } x\}$. For an edge configuration $\xi \in \{0,1\}^{B^c}$, define the *random-cluster distribution* $\phi^{B,\xi}_{p,q}$ in B as the probability measure on $\{0,1\}^{\mathcal{B}}$ in which each $\eta \in \{0,1\}^{\mathcal{B}}$ is assigned probability proportional to

$$I_{\{\eta=\xi \text{ off } B\}} \left\{ \prod_{e\in B} p^{\eta(e)}(1-p)^{1-\eta(e)} \right\} q^{k(\eta,B)} \ , \tag{35}$$

where $k(\eta, B)$ is the number of connected components of η which intersect $V(B)$. This is a generalization of the random-cluster distributions $\phi^b_{p,q,\Lambda}$ defined in Section 6.2, which are recovered by taking $B = \mathcal{B}^b_\Lambda$ and $\xi \equiv b$, $b \in \{0,1\}$. It is easy to see that the random-cluster distributions are consistent in the sense that conditioning on a configuration in some $B' \subset B$ yields the corresponding random-cluster distribution in $B \setminus B'$.

Definition 6.14 *A probability measure ϕ on $\{0,1\}^{\mathcal{B}}$ is said to be a random-cluster measure with parameters p and q if its conditional probabilities satisfy*

$$\phi(\eta\,|\,\xi) \equiv \phi(\eta \text{ in } B\,|\,\xi \text{ off } B) = \phi^{B,\xi}_{p,q}(\eta)$$

for all finite simply connected $B \subset \mathcal{B}$, ϕ-almost all ξ and all η such that $\eta = \xi$ off B.

This is the direct analogue of the definition of a Gibbs measure in the random-cluster setting. There is, however, also another possibility which differs from the preceding one for graphs in which the complements of finite regions are not connected. The idea is to connect all infinite clusters at infinity. This corresponds to a one-point compactification of G. Accordingly, we shall use a prefix 'C' which stands for "compactified". Thus, we define a *C-random-cluster distribution* $\hat\phi^{B,\xi}_{p,q}$ as in (35), except that $k(\eta, B)$ is replaced by $\hat k(\eta, B)$, defined as the number of all *finite* connected components of η intersecting $V(B)$.

Definition 6.15 *A probability measure ϕ on $\{0,1\}^{\mathcal{B}}$ is said to be a C-random-cluster measure for p and q if its conditional probabilities satisfy*

$$\phi(\eta\,|\,\xi) = \hat\phi^{B,\xi}_{p,q}(\eta)$$

for all finite simply connected $B \subset \mathcal{B}$, ϕ-almost all ξ and all η such that $\eta = \xi$ off B.

The study of random-cluster measures in the case $G = \mathbf{Z}^d$ was initiated by Grimmett [109], and about simultaneously by Pfister and Vande Velde [191] and Borgs and Chayes [32]; see also Seppäläinen [211] for some even more recent developments. The C-variant (Definition 6.15) was introduced in the regular tree case by Häggström [114, 115], and further studied in a general graph context by Jonasson [139]. In the following, we will try to convince the reader that random-cluster measures of both types are of interest, and also discuss their relation to each other.

We shall concentrate mainly on the case $q \geq 1$. The reason for this is that it is only for $q \geq 1$ that the conditional probability in (27) is increasing in η, which allows the use of the stochastic domination and correlation results in Section 4 (Theorems 4.8 and 4.11). For $q < 1$, these tools are not available, and for this reason the random-cluster model with $q < 1$ is much less understood than the $q \geq 1$ case, although in Grimmett [109], Häggström [113] and Seppäläinen [211] one can find at least some results in the $q < 1$ regime of the parameter space.

We write $\phi^{B,0}_{p,q}$ for $\phi^{B,\xi}_{p,q}$ with $\xi \equiv 0$, and $\hat\phi^{B,1}_{p,q}$ for $\hat\phi^{B,\xi}_{p,q}$ with $\xi \equiv 1$. We can omit the hat when $G = \mathbf{Z}^d$, because there is always exactly one infinite cluster regardless of the configuration on B (this is related to Proposition 6.19 below). On the other hand, the two different ways of counting clusters with wired boundary condition are not equivalent for all graph structures; a simple counterexample is $G = \mathbf{T}_d$ in which the wired boundary condition gives rise to several infinite clusters.

Still in the context of general infinite graphs, we write $B \uparrow \mathcal{B}$ for the limit along some (any) sequence of finite simply connected regions increasing to $\mathcal{B}$ in the usual way. In complete analogy to (29) and Lemma 6.8 we then obtain the following monotonicity and convergence result.

Lemma 6.16 *For $p \in [0,1]$, $q \geq 1$, and any two finite bond sets $B_1 \subset B_2$, we have*

(i) $\phi^{B_1,0}_{p,q} \preceq_{\mathcal{D}} \phi^{B_2,0}_{p,q}$, *so that the limit* $\phi^0_{p,q} = \lim_{B\uparrow\mathcal{B}} \phi^{B,0}_{p,q}$ *exists; and*

(ii) $\hat\phi^{B_1,1}_{p,q} \succeq_{\mathcal{D}} \hat\phi^{B_2,1}_{p,q}$, *so that the limit* $\hat\phi^1_{p,q} = \lim_{B\uparrow\mathcal{B}} \hat\phi^{B,1}_{p,q}$ *exists.*

Now let ϕ be any random-cluster measure of either type, compactified or not, with the given parameters p and q. Further application of Theorem 4.8 (or Corollary 6.7) implies that

$$\phi^0_{p,q} \preceq_{\mathcal{D}} \phi \preceq_{\mathcal{D}} \hat\phi^1_{p,q}\,.$$

This is analogous to the sandwiching relation (15) for the Ising model.

Furthermore, the arguments of Section 6.3 go through to show that the q-state Potts model on G at inverse temperature β has a unique Gibbs measure if and only if the $\hat\phi^1_{p,q}$-probability of having an infinite cluster is 0, where as usual $p = 1 - e^{-2\beta}$.

For Definitions 6.14 and 6.15 to be of interest, we have to establish at least the existence of random-cluster measures of the two types. The following theorem tells us that at least for $q \geq 1$, such measures do exist. (The existence problem for $q < 1$ remains open in the setting of general graphs, although existence has been established for $\mathbf{Z}^d$ and $\mathbf{T}_d$; see [109] and [114], respectively.)

Theorem 6.17 *For $p \in [0,1]$ and $q \geq 1$, we have that (i) $\phi^0_{p,q}$ is a random-cluster measure, and (ii) $\hat\phi^1_{p,q}$ is a C-random-cluster measure.*

For the proof we use the following lemma which characterizes random-cluster measures in terms of single-edge conditional probabilities. For $e = \langle xy \rangle \in \mathcal{B}$ and $\xi \in \{0,1\}^{\mathcal{B}\setminus\{e\}}$, we write as usual $\{x \leftrightarrow y\}$ for the event that there exists an open path in ξ from x to y. We also write $\{x \overset{C}{\longleftrightarrow} y\}$ for the event that there either exists an open path in ξ from x to y, or x and y are both in infinite clusters of ξ. We think of this C-connectivity notion $x \overset{C}{\longleftrightarrow} y$ as allowing paths between x and y to go "via infinity".

Lemma 6.18 *Fix $p \in [0,1]$ and $q > 0$, and let ϕ be a probability measure on $\{0,1\}^{\mathcal{B}}$. Then ϕ is a random-cluster measure for p and q if and only if for each $e = \langle xy \rangle \in \mathcal{B}$ and ϕ-a.e. $\xi \in \{0,1\}^{\mathcal{B}\setminus\{e\}}$ we have*

$$\phi(e \text{ is open} \mid \xi) = \begin{cases} p & \text{if } x \leftrightarrow y \text{ in } \xi \\ \frac{p}{p+(1-p)q} & \text{otherwise,} \end{cases} \tag{36}$$

Similarly, ϕ is a C-random-cluster measure for p and q if and only if (36) holds with $x \overset{C}{\longleftrightarrow} y$ instead of $x \leftrightarrow y$.

Proof: We consider only the first statement, as the C-case is completely similar. For the "only if" part we only need to note that the right-hand side of (36) is equal to $\phi^{\{e\},\xi}_{p,q}(e \text{ is open })$. Passing to the "if" part, we may restrict ourselves to the case of $p \in (0,1)$. Assume that ϕ satisfies (36) for each $e = \langle xy\rangle \in \mathcal{B}$ and ϕ-a.e. $\xi \in \{0,1\}^{\mathcal{B}\setminus\{e\}}$. Let $B \subset \mathcal{B}$ be some finite edge set. We need to show that the conditional distribution $\phi(\cdot|\xi)$ of ϕ given the configuration ξ on B^c equals $\phi^{B,\xi}_{p,q}$ for ϕ-a.e. ξ. For this, it suffices to check that for any two configurations $\eta, \eta' \in \{0,1\}^{\mathcal{B}}$ which agree with ξ on B^c we have

$$\frac{\phi(\eta\,|\,\xi)}{\phi(\eta'\,|\,\xi)} = \frac{\{\prod_{e\in B} p^{\eta(e)}(1-p)^{1-\eta(e)}\}q^{k(\eta,B)}}{\{\prod_{e\in B} p^{\eta'(e)}(1-p)^{1-\eta'(e)}\}q^{k(\eta',B)}} \tag{37}$$

with k defined as in (35). If η and η' differ only at a single edge e, then $k(\eta, B) - k(\eta', B) = k(\eta, e) - k(\eta', e)$, whence (37) is immediate from (36). In the general case, we interpolate η and η' by a sequence of configurations which successively differ in at most one edge, and use a telescoping argument. $\square$

Proof of Theorem 6.17: We prove (ii) only, as (i) follows from a similar argument and is also better known, see e.g. Borgs and Chayes [32]. Fix $e = \langle xy\rangle \in \mathcal{B}$. By Lemma 6.18, it is sufficient to establish (36) with the C-connectivity relation $\overset{C}{\longleftrightarrow}$ in place of the standard connectivity relation $\leftrightarrow$.

Let $B_1, B_2, \ldots$ be an increasing sequence of finite edge sets containing e and converging to $\mathcal{B}$ in the usual sense. We write, with slight abuse of notation, $\xi(B_i)$ for the restriction of ξ to $B_i \setminus \{e\}$. We recall from the martingale convergence theorem that

$$\hat\phi^1_{p,q}(e \text{ is open}\,|\,\xi) = \lim_{j\to\infty} \hat\phi^1_{p,q}(e \text{ is open}\,|\,\xi(B_j)) \tag{38}$$

for $\hat\phi^1_{p,q}$-a.e. $\xi \in \{0,1\}^{\mathcal{B}\setminus\{e\}}$.

We suppose first that $x \overset{C}{\longleftrightarrow} y$ fails in ξ. Then at least one of the vertices x and y is in a finite cluster of ξ, and consequently there is some m (depending on ξ) such that $\neg(x \overset{C}{\longleftrightarrow} y)$ can be verified by just looking at $\xi(B_m)$. (This is a consequence of the special concept of C-connectivity.) For any $n \ge j \ge m$ we then have

$$\hat\phi^{B_n,1}_{p,q}(e \text{ is open}\,|\,\xi(B_j)) = \frac{p}{p+(1-p)q}$$

so that by the definition of $\hat\phi^1_{p,q}$ we get

$$\hat\phi^1_{p,q}(e \text{ is open}\,|\,\xi(B_j)) = \frac{p}{p+(1-p)q}$$

by letting $n \to \infty$. Then we let $j \to \infty$ and use (38) to deduce the C-version of (36) in the case $\xi \notin \{x \overset{C}{\longleftrightarrow} y\}$.

We go on to the case $\xi \in \{x \overset{C}{\longleftrightarrow} y\}$. In analogy to (38), we have

$$\lim_{j\to\infty} \hat\phi^1_{p,q}(x \overset{C}{\longleftrightarrow} y\,|\,\xi(B_j)) = 1$$

for $\hat\phi^1_{p,q}$-a.e. $\xi \in \{x \overset{C}{\longleftrightarrow} y\}$. For such ξ and any $\varepsilon > 0$, we can thus find an m (depending on ξ) such that

$$\hat\phi^1_{p,q}(x \overset{C}{\longleftrightarrow} y\,|\,\xi(B_j)) \ge 1-\varepsilon \tag{39}$$

for any $j \geq m$. Next we use the definition of $\hat{\phi}^1_{p,q}$. For any $n = 1, 2, \ldots$, let Y and Y_n be $\{0,1\}^{\mathcal{B}}$-valued random edge configurations with distributions $\hat{\phi}^1_{p,q}$ and $\hat{\phi}^{B_n,1}_{p,q}$ satisfying $Y_n \succeq Y$; this is possible by Lemma 6.16. We write P_n for the probability measure underlying this coupling. By the same lemma and the order relation $Y_n \succeq Y$, we have

$$\lim_{n\to\infty} P_n(Y_n(B_j) \neq Y(B_j)) = 0 . \tag{40}$$

Since $Y_n \in \{x \overset{C}{\longleftrightarrow} y\}$ whenever $Y \in \{x \overset{C}{\longleftrightarrow} y\}$, we can write

$$\begin{aligned}
& \left| \, \hat{\phi}^{B_n,1}_{p,q}(x \overset{C}{\longleftrightarrow} y \,,\, \xi(B_j)) - \hat{\phi}^1_{p,q}(x \overset{C}{\longleftrightarrow} y \,,\, \xi(B_j)) \, \right| \\
\leq \;\; & P_n\Big(\{x \overset{C}{\longleftrightarrow} y \text{ in } Y_n,\ Y_n(B_j) = \xi(B_j)\} \bigtriangleup \{x \overset{C}{\longleftrightarrow} y \text{ in } Y,\ Y(B_j) = \xi(B_j)\}\Big) \\
\leq \;\; & \Big(\hat{\phi}^{B_n,1}_{p,q}(x \overset{C}{\longleftrightarrow} y) - \hat{\phi}^1_{p,q}(x \overset{C}{\longleftrightarrow} y)\Big) + P_n(Y_n(B_j) \neq Y(B_j)) \,,
\end{aligned}$$

where $\bigtriangleup$ denotes symmetric difference. Since $\{x \overset{C}{\longleftrightarrow} y\}$ is the decreasing limit of the local events $\{x \leftrightarrow y \text{ in } \Delta\} \cup \{x \leftrightarrow \Delta^c,\, y \leftrightarrow \Delta^c\}$ as $\Delta \uparrow \mathcal{L}$, an analogue of (30) together with (40) shows that the last expression tends to zero as $n \to \infty$. It follows that

$$\lim_{n\to\infty} \hat{\phi}^{B_n,1}_{p,q}(x \overset{C}{\longleftrightarrow} y \,|\, \xi(B_j)) = \hat{\phi}^1_{p,q}(x \overset{C}{\longleftrightarrow} y \,|\, \xi(B_j))$$

which is at least $1 - \varepsilon$ by (39). But since $\hat{\phi}^{B_n,1}_{p,q}(e \text{ is open} \,|\, \xi') = p$ for each n and all $\xi' \in \{x \overset{C}{\longleftrightarrow} y\}$, we get

$$p - \varepsilon \leq \lim_{n\to\infty} \hat{\phi}^{B_n,1}_{p,q}(e \text{ is open} \,|\, \xi(B_j)) \leq p \,.$$

Hence,

$$p - \varepsilon \leq \hat{\phi}^1_{p,q}(e \text{ is open} \,|\, \xi(B_j)) \leq p \,,$$

and since ε was arbitrary we can use (38) to deduce the C-version of (36) in the case $\xi \in \{x \overset{C}{\longleftrightarrow} y\}$. $\square$

Let us now briefly address the issue of whether the two types of random-cluster measures are any different. The following result says that very often they are the same.

Proposition 6.19 *Let ϕ be a probability measure on $\{0,1\}^{\mathcal{B}}$ with*

$$\phi(\exists \text{ at most one infinite open cluster}) = 1 \,.$$

Then, for any $p \in [0,1]$ and $q > 0$, ϕ is a random-cluster measure for p and q if and only if it is a C-random-cluster measure for p and q.

This means that whenever "uniqueness of the infinite cluster" can be verified, the two types of random-cluster measures coincide. An example is obtained if we consider translation invariant random-cluster measures for $\mathbf{Z}^d$, since the Burton–Keane uniqueness theorem (Theorem 5.17) applies in this situation. For $\mathbf{Z}^d$, the measures $\phi^0_{p,q}$ and $\hat{\phi}^1_{p,q}$ are translation invariant, by Lemma 6.8. On the other hand, uniqueness of the infinite cluster typically fails on trees, leading to very different behavior for the two types of random-cluster measures; see [114, 115] for a discussion.

Proof of Proposition 6.19: For $p = 0$ or 1 the result is trivial, so we may assume that $p \in (0, 1)$. The conditional probabilities in (36) and its C-counterpart differ only on the event

$$A_{xy} = \{x \stackrel{C}{\longleftrightarrow} y\} \setminus \{x \leftrightarrow y\}\,.$$

Hence if ϕ is a random-cluster measure but not a C-random-cluster measure (or vice versa), then A_{xy} has to have positive ϕ-probability for some edge $e = \langle xy \rangle \in \mathcal{B}$. But then the event $A_{xy} \cap \{e \text{ is closed}\}$ has positive ϕ-probability, and since this event implies the existence of at least two infinite clusters, we are done. $\square$

Much of the study of infinite volume random-cluster measures that has been done so far concerns the issue of uniqueness (or non-uniqueness) of random-cluster measures. A discussion of this issue would, however, lead us too far, so instead we advise the reader to consult Grimmett [109], Häggström [114] and Jonasson [139] to find out what is known and what is conjectured in this field.

6.5 An application to percolation in the Ising model

In Theorem 5.10 we have seen that the probability of percolation of plus spins in the Ising model is an increasing function of the external field. A much harder question is to determine monotonicity properties of percolation probabilities as β (rather than h) is varied. An interesting open problem is to decide whether for $G = \mathbf{Z}^d$, $d \geq 2$, the probability

$$\mu^+_\beta(x \stackrel{+}{\longleftrightarrow} \infty)$$

is increasing in β. Here we write μ^+_β for the plus phase in the Ising model at inverse temperature β with external field $h = 0$, and $\{x \stackrel{+}{\longleftrightarrow} \infty\}$ is the event that there exists an infinite path of plus spins starting at x. At first sight, one might be seduced into thinking that this would be a consequence of the connection between Ising and random-cluster models, and the stochastic monotonicity of random-cluster measures as p varies; see (31). However, such a conclusion is unwarranted. For example, in the coupling of Theorem 6.2 the existence of an open path between x and y in the random-cluster representation is a sufficient *but not necessary* condition for x and y to be in the same spin cluster. In fact, Häggström [116] showed, by means of a simple counterexample and in response to a question of Cammarota [45], that the probability that x and y are in the same spin cluster need not be increasing in β, and similarly for the expected size of the spin cluster containing x.

However, when the underlying graph is a tree, monotonicity in β of the probability of plus percolation can be established:

Theorem 6.20 *For the Ising model on the regular tree $\mathbf{T}_d$, $d \geq 2$, with a distinguished vertex x, the percolation probability $\mu^+_\beta(x \stackrel{+}{\longleftrightarrow} \infty)$ is increasing in β.*

An interesting aspect of this result is that its proof, unlike those of the monotonicity results mentioned earlier in this section, is *not* based on stochastic domination between the probability measures in question. In fact, stochastic domination fails, i.e. it is not always the case (in the setting of Theorem 6.20) that

$$\mu^+_{\beta_1} \preceq_{\mathcal{D}} \mu^+_{\beta_2} \tag{41}$$

when $\beta_1 \leq \beta_2$. An easy way to see this is as follows. Just as in Theorem 3.1, let β_c be the critical inverse temperature for non-uniqueness of the Gibbs measure. (It is straightforward to show, using either the random-cluster approach or the methods in Section 7, that $\beta_c > 0$ for $\mathcal{L} = \mathbf{T}_d$.) Pick $\beta_1 < \beta_2$ in $(0, \beta_c)$. By Theorem 4.15, we then have $\mu^+_{\beta_1}(\xi : \xi(y) = +1) = \mu^+_{\beta_2}(\xi : \xi(y) = +1) = 1/2$ for every vertex y. If now (41) was true we would have $\mu^+_{\beta_1} = \mu^+_{\beta_2}$ by Proposition 4.12. This, however, is impossible because the two measures have different conditional distributions on finite regions.

Theorem 6.20 can be proved using the exact calculations for the Ising model on $\mathbf{T}_d$, which can be found in e.g. Spitzer [215] and Georgii [96]. Here we present a simpler proof which does not require any exact calculation, but which exploits random-cluster methods.

Proof of Theorem 6.20: As usual we write $\mathcal{L}$ and $\mathcal{B}$ for the vertex and edge sets of $\mathbf{T}_d$. Since

$$\lim_{\Lambda \uparrow \mathcal{L}} \mu^+_{\beta,\Lambda}(x \stackrel{+}{\longleftrightarrow} \infty) = \mu^+_{\beta}(x \stackrel{+}{\longleftrightarrow} \infty)$$

for any β in analogy to (30), it suffices to show that for $\beta_1 \leq \beta_2$ and any Λ, we have

$$\mu^+_{\beta_1,\Lambda}(x \stackrel{+}{\longleftrightarrow} \infty) \leq \mu^+_{\beta_2,\Lambda}(x \stackrel{+}{\longleftrightarrow} \infty)\,. \tag{42}$$

This we will do by constructing a coupling P of two $\{-1,+1\}^{\mathcal{L}}$-valued random objects X_1 and X_2 with respective distributions $\mu^+_{\beta_1,\Lambda}$ and $\mu^+_{\beta_2,\Lambda}$ and the property that if $x \stackrel{+}{\longleftrightarrow} \infty$ in X_1, then the same thing happens in X_2.

Recall the Edwards–Sokal coupling of spin and edge configurations described ahead of Proposition 6.9. In the present case of the tree $\mathbf{T}_d$, this construction requires the C-version of counting clusters, which corresponds to making the complement of Λ connected. Therefore we will work with the C-random-cluster distributions. Let $p_1 = 1 - e^{-2\beta_1}$ and $p_2 = 1 - e^{-2\beta_2}$, and let $B = \mathcal{B}^1_\Lambda \subset \mathcal{B}$ be the set of edges incident to at least one vertex in Λ. We first let Y_1 and Y_2 be two $\{0,1\}^{\mathcal{B}}$-valued random edge configurations distributed according to the random-cluster measures $\hat{\phi}^{B,1}_{p_1,2}$ and $\hat{\phi}^{B,1}_{p_2,2}$, and such that $P(Y_1 \preceq Y_2) = 1$; this is possible by the $\hat{\phi}^{B,1}_{p,2}$-analogue of (31). X_1 and X_2 can now be obtained by assigning spins to the connected components of Y_1 and Y_2 in the usual way; these spin assignments are coupled as follows. First we must assign spin $+1$ to all infinite clusters in Y_1 and Y_2. Then we let $(Z(y))_{y \in \Lambda}$ be i.i.d. random variables taking values $+1$ and -1 with probability 1/2 each, and assign to each finite cluster C of Y_1 and Y_2 the value $Z(y)$, where y is the (unique) vertex of C that minimizes the distance to x. This defined X_1 and X_2. A moment's thought reveals that the set of vertices that can be reached from x via spins in X_1 is almost surely contained in the corresponding set for X_2. Hence (42) is established, and we are done. $\Box$

Note that this proof did not use any property of $\mathbf{T}_d$ except for the tree structure, so Theorem 6.20 can immediately be extended to the setting of arbitrary trees.

6.6 Cluster algorithms for computer simulation

An issue of great importance in statistical mechanics which we have not touched upon so far is the ability to perform computer simulations of large Gibbs systems. Many (most?) questions about phase transition behavior etc. can with current knowledge only be answered partially (or not at all) using rigorous mathematical arguments. Computer simulations are then important for supporting (or rejecting) heuristic arguments, or (in case not even a good heuristic can be found) to provide ideas for what a good conjecture might be. This topic is somewhat beside the main issue of our survey, but since random-cluster representations have played a key role in simulation algorithms for more than a decade we feel that it is appropriate to describe some of these algorithms. In fact, it was the need of efficient simulation which, in the late 1980's, sparked the revival of the random-cluster model (Swendsen and Wang [219]) which up to then had raised only little interest since its introduction by Fortuin and Kasteleyn in the early 1970's.

Consider for instance the Ising model with free boundary condition on a large cubic region $\Lambda \subset \mathbf{Z}^d$. Direct sampling from the Gibbs distribution $\mu_{h,\beta,\Lambda}$ with free boundary condition is not feasible, due to the huge cardinality of the state space Ω, and the (related) intractability of computing the normalizing constant for the Gibbs measure. The most widely used way to handle this problem is the *Markov chain Monte Carlo* method, which dates back to the 1953 paper by Metropolis et al. [178]. The idea is to define an ergodic Markov chain having as unique stationary distribution the target distribution $\mu_{h,\beta,\Lambda}$. Starting the chain in an arbitrary state and running the chain for long enough will then produce an output with a distribution close to the target distribution. An example of such a chain is the single-site heat bath algorithm, whose evolution is as follows. At each integer time, a vertex $x \in \Lambda$ is chosen at random, and the spin at x is replaced by a new value according to the conditional distribution (under $\mu_{h,\beta,\Lambda}$) of the spin at x given the spins at its neighbors. It is immediate that $\mu_{h,\beta,\Lambda}$ is stationary for this chain, and ergodicity of the chain follows from elementary Markov chain theory upon checking that it is aperiodic and irreducible. The problem with this approach is that the time taken to come close to equilibrium may be very long. For example, let $h = 0$. Then, for $\beta < \beta_c$ (with β_c defined as in Theorem 3.1), the time taken to come within a fixed small variational distance from the target distribution grows only like $n \log n$ in the size of the system (here n is the number of vertices in Λ) whereas in contrast the time grows (stretched) exponentially in the size of the system for $\beta > \beta_c$; see e.g. [169, 168]. This means that simulation using this heat bath algorithm is computationally feasible even for fairly large systems provided that $\beta < \beta_c$, but not for $\beta > \beta_c$. What happens for $\beta > \beta_c$ is that if the chain starts in a configuration dominated by plus spins, then the plus spins continue to dominate for an astronomical amount of time, and similarly for starting configurations dominated by minus spins. The set of configurations where the fraction of plus spins is around $1/2$ (rather than around the fractions predicted by the magnetization in the infinite-volume Gibbs measures μ_β^+ and μ_β^-) has small probability and thus can be seen as a "bottleneck" in the state space, slowing down the convergence rate.

A way to tackle the exceedingly slow convergence rate in the phase coexistence regime is to use the heat bath algorithm for the corresponding random-cluster model rather than for the Ising model itself, and only in the end go over to the Ising model by the random mapping described in Corollary 6.3. This has the disadvantage that the calculation of single-site (or, rather, single-edge) conditional probabilities become computationally more complicated due to the possible dependence on edges arbitrarily far away (see Lemma 6.6). This disadvantage, however, seems to be by far outweighed by the fact that the convergence rate of the Markov chain (for $\beta > \beta_c$) appears to be very much faster than for the heat bath applied directly to the spin variables. The reason for this phenomenon is that the random-cluster representation "doesn't see any difference" between the plus state and the minus state. This approach can of course be used also for the $q \geq 3$ Potts model, and is due to Sweeny [218]. Later, Propp and Wilson [200] built on this approach by coupling several such Markov chains (i.e. running them in parallel) in an ingenious way, producing an algorithm which runs for a random amount of time (determined by the algorithm itself) and then outputs a state which has *exactly* the target distribution. The running time of this algorithm turns out (from experiments) to be moderate except for the case of large q and β close to the critical value. The Propp–Wilson approach, known as exact or perfect simulation, has received a vast amount of attention among statisticians during the last few years (see e.g. the annotated bibliography [225]) and we believe that it has interesting potential also in physics.

There is, however, another Markov chain which appears to converge even faster than those of Sweeny, Propp and Wilson. We are talking about the Swendsen–Wang [219] algorithm, which runs as follows for Ising and Potts models on a graph with vertex set $\mathcal{L}$ and edge set $\mathcal{B}$: Starting with a spin configuration $X_0 \in \{1, \ldots, q\}^{\mathcal{L}}$, a bond configuration $Y_0 \in \{0, 1\}^{\mathcal{B}}$ is chosen according to the random mapping defined in Corollary 6.4. Then another spin configuration X_1 is produced from Y_0 by assigning random spins to the connected components, i.e. by the random mapping of Corollary 6.3. This procedure is then iterated, producing a new edge configuration Y_1 and a new spin configuration X_2, etc. By combining the two corollaries, we see that if X_0 is chosen according to the target distribution, then the same holds for X_1, and consequently for $X_2, X_3, \ldots$. In other words, the target distribution is stationary for the chain $\{X_k\}_{k=0}^{\infty}$, and by the (easily verified) ergodicity of the chain we have a valid Markov chain Monte Carlo algorithm. Although it is not exact in the sense of the Propp–Wilson algorithm, it appears to converge much faster, thus in practice allowing simulation of systems that are orders of magnitude larger. Heuristically, the reason for this faster convergence is that large chunks of spins may flip simultaneously, allowing the chain to tunnel through any bottlenecks in the target distribution. However, rigorous upper and lower bounds on the time taken to come close to equilibrium are to a large extent lacking, although Li and Sokal [159] have provided a lower bound demonstrating the phenomenon of "critical slowing down" as β approaches β_c.

The Swendsen–Wang algorithm has, since its introduction in 1987, become the standard approach to simulating Ising and Potts models. Interesting variants and modifications of this algorithm have been developed by Wolff [226] and Machta et al. [165]; the last paper is an interesting attempt at combining the original approach of Swendsen and Wang with ideas from so called *invasion percolation* (see [55]) to get an algorithm specifically aimed at sampling from a Gibbs distribution at the critical inverse temperature β_c, i.e. where the use of other algorithms have proved to be most difficult. Generalizations of the Swendsen–Wang algorithm for various models other than Ising and Potts models have also been obtained, see e.g. Campbell and Chayes [48], Chayes and Machta [58, 59], and Häggström et al. [121].

6.7 Random-cluster representation of the Widom–Rowlinson model

The random-cluster model can be seen as a perturbation of Bernoulli bond percolation, where the probability measure is changed in favour of configurations with many (for $q > 1$) or few (for $q < 1$) connected components. A fairly natural question is what happens if we perturb Bernoulli site percolation in the same way. For lack of an established name, we call the resulting model the site-random-cluster model. Let G be a finite graph with vertex set $\mathcal{L}$ and edge set $\mathcal{B}$. For a site configuration $\eta \in \{0,1\}^{\mathcal{L}}$, we write $k(\eta)$ for the number of connected components in the subgraph of G obtained by deleting all vertices x with $\eta(x) = 0$ and their incident edges.

Definition 6.21 *The* **site-random-cluster measure** $\psi^G_{p,q}$ *for G with parameters $p \in [0,1]$ and $q > 0$ is the probability measure on $\{0,1\}^{\mathcal{L}}$ which to each $\eta \in \{0,1\}^{\mathcal{L}}$ assigns probability*

$$\psi^G_{p,q}(\eta) = \frac{1}{Z^G_{p,q}} \left\{ \prod_{x \in \mathcal{L}} p^{\eta(x)}(1-p)^{1-\eta(x)} \right\} q^{k(\eta)},$$

where $Z^G_{p,q}$ is a normalizing constant.

Analogously to the usual random-cluster model living on bonds, taking $q = 1$ gives the ordinary Bernoulli site percolation ψ_p, while other choices of q lead to dependence between vertices.

Taking $q = 2$ is of particular interest because it leads to a representation of the Widom–Rowlinson model which is similar to (and in fact slightly simpler than) the usual random-cluster representation of the Ising model. Let μ^G_λ be the Gibbs measure for the Widom–Rowlinson model with activity λ on G, i.e. μ^G_λ is the probability measure on $\{-1,0,+1\}^{\mathcal{L}}$ which to each $\xi \in \{-1,0,+1\}^{\mathcal{L}}$ assigns probability proportional to

$$\prod_{\langle xy \rangle \in \mathcal{B}} I_{\{\xi(x)\xi(y) \neq -1\}} \prod_{x \in \mathcal{L}} \lambda^{|\xi(x)|}.$$

The following analogues of Corollaries 6.3 and 6.4 are trivial to check.

Proposition 6.22 *Let $p = \frac{\lambda}{1+\lambda}$, and suppose we pick a random spin configuration $X \in \{-1,0,+1\}^{\mathcal{L}}$ as follows.*

1. *Pick $Y \in \{0,1\}^{\mathcal{L}}$ according to $\psi^G_{p,2}$.*
2. *Set $X(x) = 0$ for each $x \in \mathcal{L}$ such that $Y(x) = 0$.*
3. *For each open cluster C of Y, flip a fair coin to decide whether to give spin $+1$ or -1 in X to all vertices of C.*

Then X is distributed according to the Widom–Rowlinson Gibbs measure μ^G_λ.

Proposition 6.23 *Let $p = \frac{\lambda}{1+\lambda}$, and suppose we pick a random spin configuration $Y \in \{0,1\}^{\mathcal{L}}$ as follows.*

1. *Pick $X \in \{-1,0,+1\}^{\mathcal{L}}$ according to μ^G_λ.*
2. *Set $Y(x) = |X(x)|$ for each $x \in \mathcal{L}$.*

Then Y is distributed according to the site-random-cluster measure $\psi^G_{p,2}$.

We remark that for $q \in \{3,4,\ldots\}$, these results extend in the obvious way to a connection between $\psi^G_{p,q}$ and the generalized Widom–Rowlinson model whith q types of particles rather than just 2 (and strict repulsion between all particles of different type).

Many of the arguments applied to Ising and Potts models in Section 6.3 can now be applied to the Widom–Rowlinson model in a similar manner. To apply Theorem 4.8, we need to calculate the conditional probability in the site-random-cluster model that a given vertex is open given the status of all other vertices. For $x \in \mathcal{L}$ and $\eta \in \{0,1\}^{\mathcal{L}\setminus\{x\}}$, we get

$$\psi^G_{p,q}(x \text{ is open} \mid \eta) = \frac{p\, q^{1-\kappa(x,\eta)}}{p\, q^{1-\kappa(x,\eta)} + 1 - p} \tag{43}$$

where $\kappa(x,\eta)$ is the number of open clusters of η which intersect x's neighborhood $\{y \in \mathcal{L} : y \sim x\}$. If the degree of the vertices in G is bounded by N, say, then $0 \le \kappa(x,\eta) \le N$ for any $x \in \mathcal{L}$ and $\eta \in \{0,1\}^{\mathcal{L}\setminus\{x\}}$. For fixed q and any $p^* \in (0,1)$, we can thus apply Theorem 4.8 to show that $\psi^G_{p,q}$ stochastically dominates ψ_{p^*} for p sufficiently close to 1, and is dominated by ψ_{p^*} for p small enough. The arguments of Section 6.3 leading to a proof of Theorem 3.1, with the random-cluster model replaced by the site-random-cluster model, therefore go through to show Theorem 3.4.

One thing that does *not* go through in this context, however, is the analogue of (31). The reason for this is that, in contrast to (27), the conditional probability in (43) fails to be increasing in η, so that Theorem 4.8 is not applicable for comparison between site-random-cluster measures with different values of p. In fact, the analogue of (31) for site-random-cluster measures sometimes fails, and moreover the occurrence of phase transition for the Widom–Rowlinson model on certain graphs fails to be increasing in λ, as demonstrated by Brightwell, Häggström and Winkler [38].

Another consequence of the failure of the conditional probability in (43) to be increasing is that the FKG inequality (Theorem 4.11) cannot be applied to $\psi^G_{p,q}$. As a consequence, the proof of Theorem 6.10 cannot be adapted to the case of the multitype ($q \ge 3$) Widom–Rowlinson model. In fact, such a Widom–Rowlinson analogue of Theorem 6.10 is known to be *false*, as shown by Runnels and Lebowitz [207]; see also [57] and [184].

7 Uniqueness and exponential mixing from non-percolation

In the previous section we saw examples where phase transition in one system was equivalent to the existence of infinite clusters in another, suitably defined, system. In this section we shall discuss various approaches where conclusions about the phase transition behavior can only be drawn from nonexistence (and not from existence) of infinite clusters. On the other hand, these approaches typically apply to a much wider range of models. We address two problems: the uniqueness of the Gibbs measure, and the decay of correlations for a given Gibbs measure. In fact, the general theme of this section can be stated as follows: To which extent can a given spin be influenced by a configuration far away? If such an influence disappears in the limit of infinite distance, it follows (depending on the setting) that either there is no long-range influence of boundary conditions at all (implying uniqueness of the Gibbs measure), or that a specific low temperature phase exhibits some mixing properties. In both cases, the decreasing influence comes from the absence of infinite clusters of suitable type which could transport a dependence between spins. So, both uniqueness and mixing will appear here as a consequence of non-percolation.

In a first part, we will address the problem of uniqueness. In fact, we will encounter conditions which not only imply the uniqueness of the Gibbs measure, but also lead us into a regime where 'all good things' happen, i.e., where the unique Gibbs measure exhibits nice exponential mixing properties and the free energy depends analytically on all relevant parameters. (In general, the uniqueness of the Gibbs measure does not imply the absence of other critical phenomena, which might manifest themselves as singularities of the free energy or other thermodynamic quantities. For example, in Section 9 we will see that in the so-called Griffiths' regime of a disordered system there is a unique Gibbs measure, but the free energy is not analytic.)

The 'nice regime' above is usually referred to as the high temperature, or weak coupling, low density, or also analytic regime, and is usually studied by high temperature cluster expansions. Dobrushin and Shlosman [70, 71] developed a beautiful and general theory describing a regime of 'complete analyticity' by various equivalent properties. One of these ranks at the top of a hierarchy of mixing properties. While complete analyticity makes precise what actually the 'nice regime' is, and applies mainly to high temperatures or large external fields, it is not limited to this case only [75]. The relationships between this and related notions and also with dynamical properties have been studied in many papers. Although some of these have an explicit geometric flavor, we do not discuss them here because of limitations of space. We rather refer to the sources [70, 71, 217, 168, 170] and also to the references following condition (68).

In Section 7.3 we shall discuss an application of the percolation method to the low temperature regime, and see how percolation estimates for the covariance between two distant observables, combined with contour estimates, give rise to exponential mixing properties.

7.1 Disagreement paths

Let $(\mathcal{L}, \sim)$ be an arbitrary locally finite graph, and suppose we are given a neighbor interaction $U : S \times S \to \mathbf{R}$ and a self-potential $V : S \to \mathbf{R}$. Consider the associated Gibbs distributions $\mu^{\eta}_{\beta,\Lambda}$ introduced in (4). More generally, we could consider an arbitrary

Markov specification $(G_\Lambda)_{\Lambda\in\mathcal{E}}$ in the sense of Section 2.6. Such specifications appear, in particular, if we have an interaction of finite range R, say on $\mathbf{Z}^d$, and draw edges between all sites of distance at most R. However, for definiteness and simplicity we stick to the setting described by the Hamiltonian (1). We will often consider the inverse temperature β as fixed and then simply write μ_Λ^η instead of $\mu_{\beta,\Lambda}^\eta$. If Λ is a singleton, we use the shorthand x for $\{x\}$.

We look for a condition implying that there is only one Gibbs measure μ for the Hamiltonian (1), i.e., a unique probability measure on $\Omega = S^{\mathcal{L}}$ satisfying

$$\mu(\,\cdot\,|\,X \equiv \eta \text{ off } \Lambda) = \mu_\Lambda^\eta \quad \text{ for } \mu\text{-almost all } \eta \in \Omega\,.$$

Since this property needs only to be checked for singletons $\Lambda = \{x\}$ (cf. Theorem 1.33 of [96]), it is sufficient to look for conditions on the single-spin Gibbs distributions μ_x^η with $x \in \mathcal{L}$. Intuitively, we want to express that $\mu_x^\eta(X(x) = a)$ depends only weakly on η (which can be expected to hold for small β). This dependence can be measured by the maximal variation

$$p_x = \max_{\eta,\eta'\in\Omega} \|\mu_x^\eta - \mu_x^{\eta'}\|_x\,, \tag{44}$$

where

$$\|\nu\|_\Delta = \sup_{A\in\mathcal{F}_\Delta} |\nu(A)| \tag{45}$$

is the total variation norm on the sub-σ-algebra $\mathcal{F}_\Delta$ of events which depend only on the spins in Δ. We write $\mathbf{p}$ as a shorthand for the family $(p_x)_{x\in\mathcal{L}}$.

Given two configurations $\xi,\ \xi' \in \Omega$, a path in $\mathcal{L}$ will be called a *path of disagreement* (for ξ and ξ') if $\xi(x) \neq \xi'(x)$ for all its vertices x. For each finite region $\Lambda \subset \mathcal{L}$ and any two configurations η, η' on Λ^c we will construct a coupling P of μ_Λ^η and $\mu_\Lambda^{\eta'}$ describing the difference of these measures in terms of paths of disagreement running from the boundary $\partial\Lambda$ into the interior of Λ. Intuitively, these paths of disagreement then show how deep inside the influence of the boundary conditions can still be felt. We write $\{\Delta \stackrel{\neq}{\longleftrightarrow} \partial\Lambda\}$ for the event in $S^\Lambda \times S^\Lambda$ that there exists a path of disagreement from some point of a set $\Delta \subset \Lambda$ to some point of $\partial\Lambda$.

Although the coupling P to be constructed is not best suited for direct use, it has a useful special feature: its disagreement distribution is stochastically dominated by a Bernoulli measure. This will allow us to conclude that absence of percolation for the latter implies uniqueness of the Gibbs measure for the Hamiltonian (1).

We write $\psi_{\mathbf{p}}$ for the Bernoulli measure on $\{0,1\}^{\mathcal{L}}$ with $\psi_{\mathbf{p}}(X(x) = 1) = p_x$ for all $x \in \mathcal{L}$, and $\psi_{\mathbf{p},\Lambda}$ for the analogous product measure on $\{0,1\}^\Lambda$. As in Section 4, we use the notation $X(x)$ and $X'(x)$ for the projections from $\Omega \times \Omega$ to S. The following theorem is due to van den Berg and Maes [27].

Theorem 7.1 *For each finite $\Lambda \subset \mathcal{L}$ and each pair $\eta,\ \eta' \in \Omega$ there exists a coupling $P = P_{\Lambda,\eta,\eta'}$ of μ_Λ^η and $\mu_\Lambda^{\eta'}$ having the following properties:*

(i) *For each $x \in \Lambda$, $\{X(x) \neq X'(x)\} = \{x \stackrel{\neq}{\longleftrightarrow} \partial\Lambda\}$ P-a.s.*

(ii) *For the distribution $P_\Lambda^{\neq}$ of $(I_{\{(X(x)\neq X'(x)\}})_{x\in\Lambda}$ under P, we have $P_\Lambda^{\neq} \preceq_{\mathcal{D}} \psi_{\mathbf{p},\Lambda}$.*

(iii) *For each $\Delta \subset \Lambda$,*

$$\|\mu_\Lambda^\eta - \mu_\Lambda^{\eta'}\|_\Delta \le P(\Delta \stackrel{\neq}{\longleftrightarrow} \partial\Lambda) \le \psi_{\mathbf{p}}(\Delta \leftrightarrow \partial\Lambda)\,. \tag{46}$$

Proof: We construct a coupling (X, X') of μ_Λ^η and $\mu_\Lambda^{\eta'}$ by the following algorithm. In a preparatory step we introduce an arbitrary linear ordering on Λ , set $\Delta = \Lambda$, and define $X(x) = \eta(x), X'(x) = \eta'(x)$ for $x \in \Delta^c$.

For fixing the main iteration step, suppose that (X, X') is already defined on the complement of a non-empty set $\Delta \subset \Lambda$ and is realized as a pair (ξ, ξ') off Δ, where $(\xi, \xi') \equiv (\eta, \eta')$ off Λ. Conditional on the event that $(X, X') \equiv (\xi, \xi')$ off Δ, we consider the Gibbs distributions μ_Δ^ξ and $\mu_\Delta^{\xi'}$ obtained by conditioning μ_Λ^η and $\mu_\Lambda^{\eta'}$ on $X \equiv \xi$ resp. ξ' off Δ, and we pick the smallest vertex $x = x(\xi, \xi') \in \Delta$ for which there exists some vertex $y \in \Delta^c$ with $y \sim x$ and $\xi(y) \neq \xi'(y)$. If such an x does not exist, we have $\mu_\Delta^\xi = \mu_\Delta^{\xi'}$ on $\mathcal{F}_\Delta$ by the Markov property, so that we can take the obvious optimal coupling for which $X \equiv X'$ on Δ, and we are done. If such an x does exist, we consider the single vertex distributions $\mu_{\Delta,x}^\xi = \mu_\Delta^\xi(X(x) = \cdot\,)$ and $\mu_{\Delta,x}^{\xi'} = \mu_\Delta^{\xi'}(X(x) = \cdot\,)$ on S. Conditionally on $(X, X') \equiv (\xi, \xi')$ off Δ, we then let $(X(x), X'(x))$ be distributed according to an optimal coupling (as in Definition 4.3) of $\mu_{\Delta,x}^\xi$ and $\mu_{\Delta,x}^{\xi'}$. The coupling (X, X') is then defined on the set $x \cup \Delta^c$, so that we can replace Δ by $\Delta \setminus x$ and repeat the preceding iteration step.

It is clear that the algorithm above stops after finitely many iterations and gives us a coupling of μ_Λ^η and $\mu_\Lambda^{\eta'}$. Property (i) is evident from the construction, since disagreement at a vertex is only possible if a path of disagreement leads from this vertex to the boundary. For (ii), we note that the measures $\mu_{\Delta,x}^\xi$ and $\mu_{\Delta,x}^{\xi'}$ are mixtures of the Gibbs distributions μ_x^σ with suitable boundary conditions σ, by the consistency of Gibbs distributions. Hence

$$\|\mu_{\Delta,x}^\xi - \mu_{\Delta,x}^{\xi'}\|_x \leq p_x \ .$$

By construction, this means that in each iteration of the main step we have

$$P(X(x) \neq X'(x) \mid (X, X') \equiv (\xi, \xi') \text{ off } \Delta) \leq p_x$$

for $x = x(\xi, \xi')$, so that (ii) follows by induction. Finally, (iii) follows directly from (i) and (ii) because for each $\Delta \subset \Lambda$

$$\|\mu_\Lambda^\eta - \mu_\Lambda^{\eta'}\|_\Delta \leq P(X(x) \neq X'(x) \text{ for some } x \in \Delta)$$

by the coupling inequality (10). The proof is therefore complete. □

Although the algorithm in the proof above is quite explicit, it is not easy to deal with directly. In particular, it is not clear in which way the coupling depends on the chosen ordering, because the site x to be selected in each step depends on (ξ, ξ') and is therefore random. Nevertheless, if the Gibbs distributions are monotone (in the sense of Definition 4.9), we get some extra properties.

Remark: Suppose S is linearly ordered and the conditional distributions μ_x^ξ are stochastically increasing in ξ. Then, if $\eta \preceq \eta'$, the coupling P of Theorem 7.1 can be chosen in such a way that, in addition to properties (i) to (iii), $X \preceq X'$ P–a.s. and, for each $x \in \Lambda$,

$$\|\mu_\Lambda^\eta - \mu_\Lambda^{\eta'}\|_x \leq P(x \overset{\neq}{\longleftrightarrow} \partial\Lambda) \leq (|S| - 1)\, \|\mu_\Lambda^\eta - \mu_\Lambda^{\eta'}\|_x \ . \tag{47}$$

This is because in each step of the algorithm proving Theorem 7.1 we can achieve that $X(x) \leq X'(x)$, and for the second inequality in (47) it is sufficient to note that

$$P(x \overset{\neq}{\longleftrightarrow} \partial\Lambda) = P(X(x) < X'(x)) \leq \sum_{a \in S \setminus \{m\}} [P(X(x) \leq a) - P(X'(x) \leq a)] \ ,$$

where m is the maximal element of S. For details we refer to [27]. In particular, for $|S| = 2$ we have equality in (47).

Let us apply this remark to the ferromagnetic Ising model with external field $h = 0$ and any inverse temperature β, with boundary conditions $\eta \equiv +1$ and $\eta' \equiv -1$ outside of some finite region $\Lambda \in \mathcal{E}$. Then, by the spin flip symmetry and stochastic monotonicity,

$$\mu^{+}_{\beta,\Lambda}(X(x)) = \mu^{+}_{\beta,\Lambda}(X(x) = 1) - \mu^{-}_{\beta,\Lambda}(X(x) = 1) = \|\mu^{+}_{\beta,\Lambda} - \mu^{-}_{\beta,\Lambda}\|_x$$

and therefore, by (47),

$$\mu^{+}_{\beta,\Lambda}(X(x)) = P(x \stackrel{\neq}{\longleftrightarrow} \partial\Lambda) \ .$$

We emphasize that this relation is completely similar to what we obtained for the random-cluster representation, viz.

$$\mu^{+}_{\beta,\Lambda}(X(x)) = \phi^{1}_{p,2,\Lambda}(x \leftrightarrow \partial\Lambda)$$

for $p = 1 - e^{-2\beta}$; cf. equation (32). The coupling P, however, is less explicit, and the geometric event involves site percolation rather than bond percolation as for the random–cluster measure, but the exact correspondence between the magnetization for the spin system and the percolation probability of the geometric system is the same.

Let us now turn to the main result of this subsection, the uniqueness theorem. Let μ, μ' be any two Gibbs measures for the Hamiltonian (1) at some inverse temperature β. Inequality (46) then shows that

$$\|\mu - \mu'\|_\Delta \le \sup_{\eta,\eta' \in \Omega} \|\mu^{\eta}_{\Lambda} - \mu^{\eta'}_{\Lambda}\|_\Delta \le \psi_{\mathbf{p}}(\Delta \leftrightarrow \partial\Lambda)$$

whenever $\Delta \subset \Lambda \in \mathcal{E}$. Letting $\Lambda \uparrow \mathcal{L}$ we find

$$\|\mu - \mu'\|_\Delta \le \psi_{\mathbf{p}}(\Delta \leftrightarrow \infty)$$

which gives the following uniqueness result.

Theorem 7.2 *If* $\psi_{\mathbf{p}}(\,\exists$ an infinite open cluster $) = 0$ *then the set* $\mathcal{G}(\beta H)$ *of Gibbs measures for the Hamiltonian (1) at inverse temperature* β *is a singleton. In particular, this holds if* $\sup_x p_x < p_c$*, the critical density for Bernoulli site percolation on* $(\mathcal{L}, \sim)$*.*

A weaker version of Theorem 7.2 was obtained first in [24] using a product coupling instead of Theorem 7.1; see Proposition 7.10 below and also [28]. In some cases, the simple product coupling nevertheless gives equivalent results; cf. the discussion in [27].

For a large class of regular graphs such as $\mathbf{Z}^d$, the assumption of Theorem 7.2 not only implies the uniqueness of the Gibbs measure but even yields certain exponential mixing properties. This can be seen almost immediately by combining inequality (46) with Theorem 5.6 on the exponential tail of the distribution of the cluster diameter in sub-critical Bernoulli percolation. We will use similar arguments in Section 9.2 in the context of random interactions.

Let us discuss now some special cases. Clearly, the conditions of Theorem 7.2 hold when $\mathcal{L} = \mathbf{Z}$ with the usual graph structure, since then $p_c = 1$. This gives uniqueness of the Gibbs measure for one-dimensional nearest-neighbor systems. Next we consider the case $\mathcal{L} = \mathbf{Z}^d$, $d \ge 2$. Recall the bound (19) for the percolation threshold p_c when $d = 2$, and the large–dimensions asymptotics of p_c in (20).

Example 7.3 *The Ising ferromagnet.* Let $\beta > 0$ be any inverse temperature and h an external field. Then, for any x, we obtain from (13) by a short computation

$$p_x = \|\mu^+_{h,\beta,x} - \mu^-_{h,\beta,x}\|_x = [\tanh(\beta(h+2d)) - \tanh(\beta(h-2d))]/2\,.$$

Hence, the Gibbs measure is unique when $h = 0$ and $\tanh(2d\beta) < p_c$, or if $|h| > 2d$ is so large that $2d < p_c \cosh^2(\beta(|h| - 2d))$, for example.

Example 7.4 *The hard-core lattice gas.* Setting $\beta = 1$, we see that $p_x = \lambda/(1+\lambda)$ for any x, so that uniqueness of the Gibbs measure follows for $\lambda < p_c/(1-p_c)$. (This can also be obtained by using the product coupling mentioned above, cf. [28].)

Example 7.5 *The Widom–Rowlinson lattice gas.* We take again $\beta = 1$ and set $\lambda_+ = \lambda_- = \lambda$. It turns out that the maximum in equation (44) is attained for the boundary conditions $\eta \equiv 0$ and η' equal to $+1$ and -1 on (at least) two different neighbors of x, whence $p_x = 2\lambda/(1+2\lambda)$ for any x. It follows that the Gibbs measure is unique when $\lambda < p_c/(2(1-p_c))$.

It is interesting to compare the uniqueness condition of Theorem 7.2 with the celebrated Dobrushin uniqueness condition, cf. [96] and the original papers [65, 66]. This condition reads

$$\sup_x \sum_y \max_{\eta \equiv \eta' \text{ off } y} \|\mu^\eta_x - \mu^{\eta'}_x\|_x < 1. \tag{48}$$

The constraint "$\eta \equiv \eta'$ off y" means that the configurations η, η' differ only at the vertex y. For systems with hard-core exclusion or in certain antiferromagnetic models it often happens that, for every $y \in \partial x$, the maximum in (48) is actually the same as that in (44), see [27]. Dobrushin's uniqueness condition then takes the form $\sup_x |\partial x|\, p_x < 1$. For $\mathcal{L} = \mathbf{Z}^d$ and $p_x = p_0$ independently of x, this means that $p_0 < 1/(2d)$, while Theorem 7.2 only requires $p_0 < p_c$, and it is known that $p_c > 1/(2d-1)$ for $d > 1$. However, if the constrained maximum in Dobrushin's condition is much smaller than the unconstrained maximum in (44), then Dobrushin's condition will be weaker than that of Theorem 7.2. For example, for the Ising ferromagnet on $\mathbf{Z}^d$ with external field $h = 0$, Dobrushin's condition requires that $2d \tanh\beta < 1$ which, in view of (20), is less restrictive than the condition obtained in Example 7.3. Thus, roughly speaking, Theorem 7.2 works best for "constrained" systems with strong repulsive interactions and low-dimensional lattices (or graphs with small $|\partial x|$'s) for which reasonable lower bounds of the critical probability p_c are available. Examples are the hard-core lattice gas and the Widom–Rowlinson lattice gas on $\mathbf{Z}^2$ considered above.

There is also another reason why Theorem 7.2 is useful. Namely, its condition of non-percolation is a global condition: the absence of percolation does not depend on the value of p_x at any single site x. In particular, p_x could be large or even be equal to 1 for all x's in an infinite subset (say, a periodic sublattice) of $\mathcal{L}$; once the p_x's are sufficiently small on the complementary set, there is still no infinite open cluster. This can be applied to non-translation invariant interactions where, in general, it is impossible to obtain uniform small bounds on the p_x's (or on the strength of the interaction, as would be required by the Dobrushin condition or for some standard cluster-expansion argument). We will come back to this point in Section 7.3.

7.2 Stochastic domination by random-cluster measures

Recently, Alexander and Chayes [14] introduced a variant of the random-cluster technique that applies to a substantially greater class of systems than those considered in Section 6. This approach involves a so-called *graphical representation* of the original system. The graphical representation is stochastically dominated by a random-cluster model, and absence of infinite clusters in this random-cluster model implies uniqueness of the Gibbs measure for the original spin system. The price to pay for the greater generality is that the implication goes only one way: percolation in the random-cluster model does not, in general, imply non-uniqueness of Gibbs measures.

We assume that the state space S is a *finite group* with unit element 1; the inverse element of $a \in S$ is denoted by a^{-1}, so that $a^{-1}a = aa^{-1} = 1$. For simplicity we assume that the underlying graph is $\mathcal{L} = \mathbf{Z}^d$ (although this will not really matter). We consider the Hamiltonian (1) for a pair potential U and with no self-energy, $V = 0$. By adding some constant to U (which does not change the relative Hamiltonian) we can arrange that $U \le 0$. The basic assumption is that U is *left-invariant*, so that

$$U(a,b) = u(a^{-1}b) \tag{49}$$

for all $a, b \in S$ and the even function $u = U(1,\cdot) \le 0$. Note that this setting includes the q-state Potts model for which $S = \mathbf{Z}_q$ and $u = -2\, I_{\{0\}}$. For any finite $\Lambda \subset \mathbf{Z}^d$ we consider the Gibbs distribution

$$\mu^\eta_{\beta,\Lambda}(\sigma) = \frac{I_{\{\sigma\equiv\eta \text{ off } \Lambda\}}}{Z_\Lambda(\beta,\eta)} \exp\Big[-\beta \sum_{\langle xy\rangle \in \mathcal{B}_\Lambda} u(\sigma(x)^{-1}\sigma(y)) \Big]$$

at inverse temperature β with boundary condition $\eta \in \Omega$. Here we write $\mathcal{B}_\Lambda$ for the set of all bonds $b \in \mathcal{B}$ with at least one endpoint in Λ. The graphical representation of $\mu^\eta_{\beta,\Lambda}$ will be based on bond configurations $\omega \in \{0,1\}^{\mathcal{B}_\Lambda}$. Each such ω will also be viewed as a subset of $\mathcal{B}_\Lambda$, and the bonds in ω will be called open. The key idea of this representation is taken from the classical high temperature expansion. For fixed $\beta > 0$ and any $a \in S$ we introduce the difference

$$R_a = e^{-\beta\, u(a)} - 1 \ge 0\,. \tag{50}$$

With this notation we can write

$$\begin{aligned} \mu^\eta_{\beta,\Lambda}(\sigma) &= \frac{I_{\{\sigma\equiv\eta \text{ off } \Lambda\}}}{Z_\Lambda(\beta,\eta)} \prod_{\langle xy\rangle \in \mathcal{B}_\Lambda} (1 + R_{\sigma(x)^{-1}\sigma(y)}) \\ &= \frac{I_{\{\sigma\equiv\eta \text{ off } \Lambda\}}}{Z_\Lambda(\beta,\eta)} \sum_{\omega \in \{0,1\}^{\mathcal{B}_\Lambda}} \prod_{\langle xy\rangle \in \omega} R_{\sigma(x)^{-1}\sigma(y)}\ . \end{aligned}$$

This shows that $\mu^\eta_{\beta,\Lambda}$ is the first marginal distribution of a probability measure $P^\eta_{\beta,\Lambda}$ on $\Omega \times \{0,1\}^{\mathcal{B}_\Lambda}$, namely

$$P^\eta_{\beta,\Lambda}(\sigma,\omega) = \frac{I_{\{\sigma\equiv\eta \text{ off } \Lambda\}}}{Z_\Lambda(\beta,\eta)} \prod_{\langle xy\rangle \in \omega} R_{\sigma(x)^{-1}\sigma(y)}\ ,$$

$\sigma \in \Omega$, $\omega \in \{0,1\}^{\mathcal{B}_\Lambda}$. The second marginal distribution of $P^\eta_{\beta,\Lambda}$ is equal to

$$\gamma^\eta_{\beta,\Lambda}(\omega) = W^\eta_{\beta,\Lambda}(\omega) \Big/ Z_\Lambda(\beta,\eta)\ ,$$

where

$$W^\eta_{\beta,\Lambda}(\omega) = \sum_{\sigma\equiv\eta \text{ off } \Lambda} \prod_{\langle xy\rangle\in\omega} R_{\sigma(x)^{-1}\sigma(y)} \tag{51}$$

is the "graphical weight" of any $\omega \in \{0,1\}^{\mathcal{B}_\Lambda}$. The probability measure $\gamma^\eta_{\beta,\Lambda}$ on $\{0,1\}^{\mathcal{B}_\Lambda}$ is called the *graphical distribution* or the *grey measure* (since it ignores the spins which are considered as colors). The graphical representation of $\mu^\eta_{\beta,\Lambda}$ thus obtained is analoguous to the random-cluster representation of the Potts model and can be summarized as follows.

Lemma 7.6 *In the set-up described above, the Gibbs distribution $\mu^\eta_{\beta,\Lambda}$ can be derived from the graphical distribution $\gamma^\eta_{\beta,\Lambda}$ by means of the conditional probabilities*

$$P^\eta_{\beta,\Lambda}(\sigma|\omega) = W^\eta_{\beta,\Lambda}(\omega)^{-1} \prod_{b=\langle xy\rangle\in\omega} R_{\sigma(x)^{-1}\sigma(y)}\ .$$

That is,

$$\mu^\eta_{\beta,\Lambda}(\sigma) = \sum_{\omega\in\{0,1\}^{\mathcal{B}_\Lambda}} \gamma^\eta_{\beta,\Lambda}(\omega)\, P^\eta_{\beta,\Lambda}(\sigma|\omega)\ .$$

For the Potts interaction $u = -2\, I_{\{0\}}$ with state space $S = \mathbf{Z}_q$, the graphical representation above is easily seen to coincide with the random-cluster representation studied in Section 6. One important feature is that the graphical weights factorize into cluster terms. Indeed, each bond configuration ω divides Λ into connected components called open clusters (which may possibly consist of isolated sites). The set of bonds belonging to an open cluster C is denoted by ω_C. Writing $\mathcal{C}(\omega)$ for the set of all open clusters we then obtain that

$$W^\eta_{\beta,\Lambda}(\omega) = \prod_{C\in\mathcal{C}(\omega)} \bar{W}^\eta_{\beta,\Lambda}(C,\omega_C) \tag{52}$$

with

$$\bar{W}^\eta_{\beta,\Lambda}(C,\omega_C) = \sum_{\sigma\in S^C:\,\sigma\equiv\eta \text{ on } C\cap\partial\Lambda} \ \prod_{\langle xy\rangle\in\omega_C} R_{\sigma(x)^{-1}\sigma(y)}\ .$$

(We make the usual convention that the empty product is equal to 1; hence $\bar{W}^\eta_{\beta,\Lambda}(C,\omega_C) = |S|$ if C is an isolated site.) Together with Lemma 7.6, equation (52) shows that the spins belonging to disjoint open clusters are conditionally independent. In particular, we can simulate the spin system by first drawing a bond configuration ω with weights (51) and then obtain in each open cluster a spin configuration according to $P^\eta_{\beta,\Lambda}(\sigma|\omega)$.

Suppose we knew that there is no percolation in the graphical representation, in the sense that $\max_\eta \gamma^\eta_{\beta,\Lambda}(0 \leftrightarrow \partial\Lambda) \to 0$ as $\Lambda \uparrow \mathcal{L}$. The conditional independence of spins in different open clusters would then suggest that there is only one Gibbs measure for the spin system. Unfortunately, this is not known (though weaker statements are established in [58]). However, one can make a stochastic comparison of the graphical distributions with wired random-cluster distributions (Lemma 7.7 below), and the absence of percolation in the dominating random-cluster distribution will then guarantee that the original system has a unique Gibbs measure. This will be achieved in Theorem 7.8 allowing to bound the dependence on boundary conditions in terms of the connectivity probability in a random-cluster model.

To this end we also need to consider Gibbs distributions $\mu^f_{\beta,\Lambda}$ with *free boundary condition*. These admit similar graphical representations $\gamma^f_{\beta,\Lambda}$ based on bond configurations inside Λ; that is, the bonds leading from Λ to Λ^c are removed. In the following, the superscript f will refer to this case.

The stochastic comparison with random-cluster distributions will be formulated using

$$R^* = \max_{a\in S} R_a\ , \quad \bar{R} = \frac{1}{|S|}\sum_{a\in S} R_a\ , \quad p = R^*/(1+R^*)\ , \quad q = R^*/\bar{R}\ . \tag{53}$$

Note that these quantities depend on β since the R_a in (50) do. In the case of the r-state Potts model when $u = -2\, I_{\{0\}}$, we have $R^* = 1 - e^{-2\beta}$ and $q = r$; that is, in this case the parameters p and q are nothing but the standard parameters of the random-cluster representation. For p and q as above we consider now the wired (resp. free) random-cluster distribution $\phi^1_{p,q,\Lambda}$ (resp. $\phi^0_{p,q,\Lambda}$) in Λ as introduced in Section 6.2.

Lemma 7.7 *For any $\Lambda \in \mathcal{E}$, $\beta > 0$ and p, q as above, $\gamma^\eta_{\beta,\Lambda} \preceq_{\mathcal{D}} \phi^1_{p,q,\Lambda}$ and $\gamma^f_{\beta,\Lambda} \preceq_{\mathcal{D}} \phi^0_{p,q,\Lambda}$.*

Proof: We only prove the first statement since the second is similar and simpler. According to Section 6.1, the weights of the random-cluster distribution $\phi^1_{p,q,\Lambda}$ are proportional to

$$\left(\frac{p}{1-p}\right)^{|\omega|} q^{k(\omega,\Lambda)}$$

with $|\omega|$ the number of open bonds and $k(\omega,\Lambda)$ the number of open clusters meeting Λ (where all clusters touching $\partial\Lambda$ are wired together into a single cluster). Up to a constant factor, the Radon–Nikodym density of $\gamma^\eta_{\beta,\Lambda}$ relative to $\phi^1_{p,q,\Lambda}$ is thus given by

$$F(\omega) = W^\eta_{\beta,\Lambda}(\omega) \Big/ (R^*)^{|\omega|} (R^*/\bar{R})^{k(\omega,\Lambda)}\ .$$

Since $\phi^1_{p,q,\Lambda}$ has positive correlations, the lemma will therefore be proved once we have shown that F is a decreasing function of ω. To this end we let $\omega \preceq \omega'$ be such that $\omega' = \omega \cup \{b\}$ for a bond $b \in \mathcal{B}_\Lambda \setminus \omega$.

We first consider the case when $b = \langle xy\rangle$ is not connected to $\partial\Lambda$ and joins two open clusters C_x, $C_y \in \mathcal{C}(\omega)$. For each open cluster C let $\bar{W}(C) = \bar{W}^\eta_{\beta,\Lambda}(C,\omega_C)$ be as in (52). Suppose we stipulate that the spin $\sigma(z)$ at any site $z \in C$ is equal to some $a \in S$. It is then easy to see that the remaining sum in the definition of $\bar{W}(C)$ does not depend on a and thus has the value $\bar{W}(C)/|S|$. Prescribing the values of $\sigma(x)$ and $\sigma(y)$ in this way we thus find that

$$\bar{W}(C_x \cup C_y \cup b) = \bar{W}(C_x)\, \bar{W}(C_y)\ |S|^{-2} \sum_{\sigma(x),\, \sigma(y)} R_{\sigma(x)^{-1}\sigma(y)},$$

and therefore $W^\eta_{\beta,\Lambda}(\omega') = \bar{R}\, W^\eta_{\beta,\Lambda}(\omega)$. Since $k(\omega',\Lambda) = k(\omega,\Lambda) - 1$ and $|\omega'| = |\omega| + 1$, it follows that $F(\omega') = F(\omega)$, proving the claim in the first case. If b links some cluster to the boundary which otherwise was separated from the boundary, then the argument above shows again that $F(\omega') = F(\omega)$.

Next we consider the case when $b = \langle xy\rangle$ closes a loop in ω but is still not connected to the boundary. Since clearly

$$R_{\sigma(x)^{-1}\sigma(y)} \le \max_{a\in S} R_a = R^*\ ,$$

we find that $W^\eta_{\beta,\Lambda}(\omega') \le W^\eta_{\beta,\Lambda}(\omega)\, R^*$. On the other hand, in this case we have $|\omega'| = |\omega| + 1$ and $k(\omega', \Lambda) = k(\omega, \Lambda)$, so that $F(\omega') \le F(\omega)$. As we are considering the wired random-cluster measure, this argument remains valid if b joins two clusters already attached to the boundary. $\square$

We are now in a position to state the main result of Alexander and Chayes [14], an estimate on the dependence of Gibbs distributions on their boundary condition in terms of percolation in the wired random-cluster distribution. Recall the notation (45) for the total variation norm on the sub-σ-algebra $\mathcal{F}_\Delta$ of events in some Δ.

Theorem 7.8 *Consider the spin system with pair interaction (49) at some inverse temperature $\beta > 0$, and let p, q be given by (53). Then, for any $\Delta \subset \Lambda \in \mathcal{E}$ and any pair of boundary conditions $\eta, \eta' \in \Omega$,*

$$\|\mu^\eta_{\beta,\Lambda} - \mu^{\eta'}_{\beta,\Lambda}\|_\Delta \le \phi^1_{p,q,\Lambda}(\Delta \leftrightarrow \partial\Lambda)\,.$$

Proof: (This proof is different from the one that appeared in [14].) Let A be any event in $\mathcal{F}_\Delta$. From Lemma 7.6 we know that

$$\mu^\eta_{\beta,\Lambda}(A) = \sum_\omega \gamma^\eta_{\beta,\Lambda}(\omega)\, P^\eta_{\beta,\Lambda}(A|\omega)\,.$$

To control the η-dependence of this probability we will proceed in analogy to the argument for the implication (ii) $\Rightarrow$ (iii) of Theorem 6.10. If $\omega \in \{\Delta \not\leftrightarrow \partial\Lambda\}$ then equation (52) shows that the conditional distribution $P^\eta_{\beta,\Lambda}(A|\omega)$ does not depend on η. So we need to control the η-dependence of $\gamma^\eta_{\beta,\Lambda}(\Delta \not\leftrightarrow \partial\Lambda)$. This, however, does not seem possible directly. So we will replace $\gamma^\eta_{\beta,\Lambda}$ by the η-independent $\phi^1_{p,q,\Lambda}$ by using a suitable coupling trick.

By Lemma 7.7 and Strassen's theorem (Theorem 4.6) there exists a coupling $(\tilde Y, \tilde Y')$ of $\gamma^\eta_{\beta,\Lambda}$ and $\phi^1_{p,q,\Lambda}$ such that $\tilde Y \preceq \tilde Y'$ almost surely. If $\tilde Y' \in \{\Delta \not\leftrightarrow \partial\Lambda\}$, there exists a largest (random) set $\Gamma = \Gamma(\tilde Y')$ such that

(a) $\Delta \subset \Gamma \subset \Lambda$, and

(b) $\tilde Y'(b) = 0$ for all bonds connecting Γ with Γ^c.

For $\tilde Y' \in \{\Delta \leftrightarrow \partial\Lambda\}$ we set $\Gamma = \emptyset$. Conditional on Γ, Lemma 7.7 and Strassen's theorem provide us further with a coupling $(\tilde Y_\Gamma, \tilde Y'_\Gamma)$ of $\gamma^f_{\beta,\Gamma}$ and $\phi^0_{p,q,\Gamma}$ such that $\tilde Y_\Gamma \preceq \tilde Y'_\Gamma$. It is then easy to see that the pair of random variables (Y, Y') defined by

$$(Y, Y')(b) = \begin{cases} (\tilde Y_\Gamma, \tilde Y'_\Gamma)(b) & \text{if } b \text{ is contained in } \Gamma, \\ (\tilde Y, \tilde Y')(b) & \text{otherwise} \end{cases}$$

is still a coupling of $\gamma^\eta_{\beta,\Lambda}$ and $\phi^1_{p,q,\Lambda}$ such that $Y \preceq Y'$ almost surely. (Notice that also $\tilde Y(b) = 0$ for all bonds from Γ to Γ^c.) We denote the underlying probability measure by Q^η. Now we can write

$$\begin{aligned} \mu^\eta_{\beta,\Lambda}(A) &= Q^\eta\Big(P^\eta_{\beta,\Lambda}(A|Y)\Big) \\ &= Q^\eta\Big(P^\eta_{\beta,\Lambda}(A|Y)\, I_{\{\Gamma=\emptyset\}}\Big) + Q^\eta\Big(P^\eta_{\beta,\Lambda}(A|Y)\, I_{\{\Gamma\neq\emptyset\}}\Big)\,. \end{aligned}$$

The first term in the last sum is at most $Q^\eta(\Gamma = \emptyset) = \phi^1_{p,q,\Lambda}(\Delta \leftrightarrow \partial\Lambda)$. We claim that the second term does not depend on η. Indeed, if $\Gamma \neq \emptyset$ then, by (52), $P^\eta_{\beta,\Lambda}(A|Y) = P^f_{\beta,\Gamma}(A|Y_\Gamma)$ only depends on the restriction Y_Γ of Y to the set of bonds inside Γ. The second term can thus be written explicitly as

$$\sum_{G \neq \emptyset} \phi^1_{p,q,\Lambda}(\Gamma = G) \sum_{\omega \text{ in } G} \gamma^f_{\beta,G}(\omega)\ P^f_{\beta,G}(A|\omega)\ ,$$

which is obviously independent of η. The theorem now follows immediately. □

To apply the theorem we consider the limiting random-cluster measure $\phi^1_{p,q}$ with arbitrary parameters $p \in]0,1[$ and $q \geq 1$ and wired boundary condition; recall from Section 6.2 that this limiting measure exists. By Corollary 6.7(d), it makes sense to define the percolation threshold

$$p_c(q) = \inf\{p : \phi^1_{p,q}(0 \leftrightarrow \infty) > 0\}.$$

We also consider the threshold $w_c(q)$ for exponential decay of connectivities, which is defined as the supremum of all p's for which

$$\phi^1_{p,q}(0 \leftrightarrow \partial\Lambda) \leq C e^{-c\, d(0,\partial\Lambda)}$$

uniformly in Λ (or, at least, for Λ in a prescribed sequence increasing to $\mathcal{L}$) with suitable constants $c > 0$ and $C < \infty$. It is evident that $w_c(q) \leq p_c(q)$; for large q it is known that $w_c(q) = p_c(q)$ [75]. Theorem 7.8 then gives us the following conditions for high-temperature behavior; compare with Theorem 7.2.

Corollary 7.9 *Whenever β is so small that $p < p_c(q)$, there is a unique Gibbs measure for the Hamiltonian H with pair interaction (49). Furthermore, if in fact $p < w_c(q)$ then the spin system is exponentially weak-mixing in the sense that there are positive constants $C < \infty, c > 0$ such that for all $\Delta \subset \Lambda \in \mathcal{E}$ and all boundary conditions $\eta, \eta' \in \Omega$*

$$\|\mu^\eta_{\beta,\Lambda} - \mu^{\eta'}_{\beta,\Lambda}\|_\Delta \leq C\,|\partial\Delta|\, e^{-c\, d(\Delta,\Lambda^c)}\,.$$

In fact, Alexander and Chayes [14] go a bit further in their exploration of 'nice' high temperature behavior, showing that for $p < p_c(1)$ the unique Gibbs measure satisfies the condition of 'complete analyticity' (investigated in [71], for example).

7.3 Exponential mixing at low temperatures

In the previous subsections we have seen how stochastic-geometric methods can be used to analyze the high temperature behavior of a spin system and, in particular, for establishing exponential decay of correlations. Here we want to demonstrate that similar percolation techniques can also be used in the low temperature regime in the presence of phase transition. We will present a method to show that, for a given phase μ, the covariance $\mu(f;g)$ of any two local observables f and g decays exponentially fast with the distance between their dependence sets. (The problem of phase transition at low temperatures will be addressed in Section 8.)

As a matter of fact, the problem of exponential decay of covariances (or truncated correlation functions) arises in many physical situations. Correlation functions are related to interesting response functions or to fluctuations of specific order parameters. Exponential decay of covariances also provides estimates on higher order correlation functions, eventually providing infinite differentiability of the free energy with respect to an external field [72]. Motivated by these interests, a variety of techniques have been developed. The most familiar approach are cluster expansions which apply equally well to both the high temperature (or low density) regime and the low temperature (or high density) regime. Although they often employ geometric concepts, it seems useful to combine them with ideas of percolation theory to make geometry more visible. An example of this is the method to be described below which is taken from a paper by Burton and Steif [37], where it was used to show that certain Gibbs measures exhibit a powerful mixing property called 'quite weak Bernoulli with exponential rate'.

We consider a spin system on an arbitrary graph $(\mathcal{L}, \sim)$ with Hamiltonian (1). As before, the essential feature of this Hamiltonian is that only adjacent spins interact, so that the Gibbs distributions $\mu^{\eta}_{\beta,\Lambda}$ in finite regions Λ have the Markov property. The inverse temperature $\beta > 0$ does not play any role for the moment, so we set it equal to 1 and drop it from our notation.

Our starting point is the following estimate on the η-dependence in terms of disagreement paths for two *independent* copies of μ^{η}_{Λ}. This result is a weak version (and, in fact, a forerunner [24]) of Theorem 7.1. It is a pleasant surprise that although developed with high temperature situations in mind, it also provides a useful alternative to some aspects of the standard low temperature expansions.

Proposition 7.10 *For any $\Delta \subset \Lambda \in \mathcal{E}$ and $\eta, \eta' \in \Omega$,*

$$\|\mu^{\eta}_{\Lambda} - \mu^{\eta'}_{\Lambda}\|_{\Delta} \le \mu^{\eta}_{\Lambda} \times \mu^{\eta'}_{\Lambda}(\Delta \overset{\neq}{\longleftrightarrow} \partial\Lambda) \ .$$

Proof: For brevity let $P = \mu^{\eta}_{\Lambda} \times \mu^{\eta'}_{\Lambda}$, and write X, X' for the two projections from $\Omega \times \Omega$ to Ω. Then for any $A \in \mathcal{F}_{\Delta}$ we have

$$\mu^{\eta}_{\Lambda}(A) - \mu^{\eta'}_{\Lambda}(A) = P(X \in A) - P(X' \in A) \ .$$

We decompose the probabilities on the right-hand side into the two contributions according to whether the event $\{\Delta \overset{\neq}{\longleftrightarrow} \partial\Lambda\}$ occurs or not. In the latter case, there exists a random set $\Gamma \subset \Lambda$ containing Δ such that $X \equiv X'$ on $\partial\Gamma$. (The union of all disagreement clusters in Λ meeting Δ is such a set.) Let Γ be the maximal random subset of Λ with this property. Then for each G the event $\{\Gamma = G\}$ only depends on the configuration outside G, and $X \equiv X'$ on ∂G. The Markov property therefore implies that, conditionally on $\{\Gamma = G\}$ and (X_{G^c}, X'_{G^c}), X_G and X'_G are independent and identically distributed, and this shows that $P(X \in A, \, \Delta \overset{\neq}{\not\leftrightarrow} \partial\Lambda) = P(X' \in A, \, \Delta \overset{\neq}{\not\leftrightarrow} \partial\Lambda)$. The proposition now follows immediately. □

What is gained with the disagreement estimate above? First, let us observe that this estimate provides bounds on covariances of local functions in terms of disagreement percolation.

Corollary 7.11 *Fix any $\Lambda \in \mathcal{E}$ and $\eta \in \Omega$. Let f and g be any two local functions depending on the spins in two disjoint subsets Δ resp. Δ' of Λ. Then*

$$|\mu^{\eta}_{\Lambda}(f; g)| \le \delta(f)\, \delta(g)\; \mu^{\eta}_{\Lambda} \times \mu^{\eta}_{\Lambda}(\Delta \overset{\neq}{\longleftrightarrow} \Delta' \textit{ in } \Lambda)$$

where $\delta(f) = \max_{\xi} f(\xi) - \min_{\xi} f(\xi)$ is the total oscillation of f.

Proof: By rescaling and addition of suitable constants we can assume that $0 \le f \le 1$ and $0 \le g \le 1$. Proposition 7.10 then shows that

$$\begin{aligned}|\mu^\eta_\Lambda(f;g)| &\le \int \mu^\eta_\Lambda(d\xi) \int \mu^\eta_\Lambda(d\xi')\, g(\xi') \left|\mu^\xi_{\Lambda\setminus\Delta'}(f) - \mu^{\xi'}_{\Lambda\setminus\Delta'}(f)\right| \\ &\le \delta(f)\,\delta(g) \int \mu^\eta_\Lambda(d\xi) \int \mu^\eta_\Lambda(d\xi')\, \mu^\xi_{\Lambda\setminus\Delta'} \times \mu^{\xi'}_{\Lambda\setminus\Delta'}(\Delta \overset{\neq}{\longleftrightarrow} \Delta' \text{ in } \Lambda)\end{aligned}$$

because $\xi \equiv \xi' \equiv \eta$ on $\partial\Lambda$. By carrying out the last integration we obtain the result. □

The bounds above leave us with the task of estimating the probability of disagreement paths in a duplicated system. In contrast to the situation in Section 7.1, we are looking now for estimates valid at low temperatures. If a cluster expansion works, there is no need to look any further. For instance, a low temperature analysis and estimates of semi-invariants for the Ising model can be obtained using standard contour representations; see [67]. It needs to be emphasized, however, that the main step of cluster expansions consists in expanding the logarithm of the partition function. Only afterwards, by taking ratios of partition functions, does one obtain expressions for covariances and higher order correlation functions. Therefore, a point to appreciate is that the bound of Corollary 7.11 provides a direct geometric bound on covariances which avoids the machinery of cluster expansions and, in particular, the problems coming from taking logarithms. As a consequence, this estimate also applies to some cases where standard cluster expansions are doomed to fail.

Let us illustrate this for the case of the Ising ferromagnet. This might not be the best example because other methods can also be applied to it; nevertheless, it is useful to demonstrate the technique in this simple case. Afterwards we will discuss a case where cluster expansion techniques cannot be used equally easily.

Consider the low temperature plus phase μ^+_β of the ferromagnetic Ising model on the square lattice $\mathcal{L} = \mathbf{Z}^2$ with zero magnetic field. By taking the infinite volume limit in Corollary 7.11 with $\eta \equiv +1$ we obtain the estimate

$$|\mu^+_\beta(f;g)| \le \delta(f)\,\delta(g)\; \mu^+_\beta \times \mu^+_\beta(\Delta \overset{\neq}{\longleftrightarrow} \Delta') \tag{54}$$

for the covariance of any two local functions f, g with disjoint dependence sets Δ, Δ'. Next we observe that the event $\{\Delta \overset{\neq}{\longleftrightarrow} \Delta'\}$ is clearly contained in the event that there exists a path from Δ to Δ' along which $(X, X') \neq (+,+)$. The latter event will be denoted by $\{\Delta \overset{\neq(+,+)}{\longleftrightarrow} \Delta'\}$. Now, Burton and Steif [37] have shown how to estimate the probability of this event to occur in the duplicated plus phase. The result is the following.

Theorem 7.12 *For the Ising ferromagnet on $\mathbf{Z}^2$ at sufficiently large β, there exist constants $c > 0$ and $C < \infty$ (depending on β) such that for any two disjoint sets $\Delta, \Delta' \in \mathcal{E}$,*

$$\mu^+_\beta \times \mu^+_\beta(\Delta \overset{\neq(+,+)}{\longleftrightarrow} \Delta') \le C \min\{|\partial_i\Delta|, |\partial_i\Delta'|\}\; e^{-c\, d(\Delta,\Delta')}\,,$$

where $\partial_i\Delta = \partial(\Delta^c)$ is the inner boundary of Δ.

Combining this theorem with (54) we obtain an exponential bound for the covariance of any two local observables in the low temperature Ising plus phase. While this result is well-known, its proof below shows how one can proceed in more general cases.

Sketch proof of Theorem 7.12: It is sufficient to prove the statement with μ_β^+ replaced by $\mu_{\beta,\Lambda}^+$, where Λ is a sufficiently large square box containing Δ, Δ'. For brevity, let $P_\Lambda = \mu_{\beta,\Lambda}^+ \times \mu_{\beta,\Lambda}^+$. Suppose that $|\partial\Delta'| \leq |\partial\Delta|$, fix an arbitrary $x \in \partial_i\Delta'$, and suppose that the event $\{x \stackrel{\neq(+,+)}{\longleftrightarrow} \Delta\}$ occurs. Let $\Gamma_0 = \Gamma_0(X, X')$ be the maximal connected set containing x on which $(X, X') \neq (+,+)$. Also, let Γ be the union of Γ_0 and all finite components of ${\Gamma_0}^c$. Γ is enclosing in the sense that both Γ and Γ^c are connected. In fact, Γ can be identified with a contour, the broken line which separates Γ from its complement.

Consider the set $\mathcal{C}$ of all enclosing sets containing x and contained in Λ. For any integer $\ell \geq 2$, let $\mathcal{C}_\ell$ be the set of all $C \in \mathcal{C}$ such that $|\partial_i C| = \ell$. Since the number of contours with length ℓ surrounding a given site on the square lattice is bounded by $\ell 3^\ell$ and since each enclosing set with $|\partial_i C| = \ell$ uniquely defines a contour with length between ℓ and 4ℓ, we have that $|\mathcal{C}_\ell| \leq 3(4\ell+1)/2(81)^\ell$ growing exponentially with ℓ. Now we can write

$$P_\Lambda(x \stackrel{\neq(+,+)}{\longleftrightarrow} \Delta) = P_\Lambda(\Gamma \cap \Delta \neq \emptyset) \leq \sum_{\ell \geq d(\Delta,\Delta')} \sum_{C \in \mathcal{C}_\ell} P_\Lambda(\Gamma = C) \ .$$

Furthermore, for $C \in \mathcal{C}_\ell$ we have

$$\begin{aligned} P_\Lambda(\Gamma = C) \quad \leq \quad & \sum_{D \subset \partial_i C} \mu_{\beta,\Lambda}^+(X \equiv -1 \text{ on } D,\, X \equiv +1 \text{ on } \partial C) \\ & \qquad \times \mu_{\beta,\Lambda}^+(X \equiv -1 \text{ on } \partial_i C \setminus D,\, X \equiv +1 \text{ on } \partial C) \ . \end{aligned} \tag{55}$$

Now, a standard Peierls estimate (see e.g. [213]) shows that

$$\mu_{\beta,\Lambda}^+(X \equiv -1 \text{ on } D, X \equiv +1 \text{ on } \partial C) \leq C(\beta)\, e^{-c(\beta)\,|D|} \tag{56}$$

with constants $c(\beta), C(\beta)$ independent of Λ satisfying $c(\beta) \to \infty$ as $\beta \to \infty$. Substituting (56) into (55) we obtain the theorem by simple combinatorics and summations. □

We emphasize that the specific properties of the plus phase μ_β^+ are used only in the last step, the Peierls estimate (56). Before, we needed only the Markov property. Therefore it is useful to note that the Peierls estimate is not limited to the Ising model; it remains valid under the conditions of the standard Pirogov–Sinai theory [213]. In particular, it follows that the results of Burton and Steif [37] on the ergodic properties of the Ising model carry over to more general Markovian models of Pirogov–Sinai type.

Let us finally discuss a case in which a Peierls estimate of the form (56) is not available. Namely, we ask for covariance estimates of local functions, still for the ferromagnetic Ising model in a large square Λ, but now for some boundary condition η not identically equal to $+1$. This question arises, for example, in the context of correlations atop of a disordered surface, or in the problem of establishing a Gibbsian description of non-Gibbsian measures. In fact, in [167] the method to be described below is used to prove that the projection to a line of the low-temperature plus phase of the two-dimensional Ising model is weakly Gibbsian.

To be specific, suppose that the boundary condition η on $\partial\Lambda$ is not identically plus but contains a large proportion of plus spins; we stipulate that $\eta \equiv +1$ on three sides of Λ while on the remaining side only a large fraction of the spins is plus. In this case, (56) cannot be true because D can be small and close to the boundary of Λ. For example, if

$E = \partial D \cap \partial\Lambda \neq \emptyset$ and η happens to be minus on E then it is not very unlikely that all spins in D are minus, and (56) will not hold. However, one can take advantage of the fact that for small D its complement in $\partial_i C$ is large, and vice versa. In other words, to estimate the right-hand side of (55) one should not apply (56) separately to each factor, but rather one can hope to estimate their product.

To make these general remarks precise we consider the Ising model on the half-plane $\mathbf{Z}\times\mathbf{Z}_+$. For any n we consider the square $\Lambda_n = \{(x_1, x_2) \in \mathbf{Z}^2 : -n \le x_1 \le n,\ 0 < x_2 \le n\}$ touching the boundary line $\mathbf{Z}\times\{0\}$. On this line we fix a configuration $\xi \in \{+1, -1\}^{\mathbf{Z}}$, thereby defining a boundary condition on one part of $\partial\Lambda$. On the remaining part of $\partial\Lambda$ we impose plus boundary condition. That is, we choose the boundary condition $\eta \equiv +1$ on $\partial\Lambda_n \setminus (\mathbf{Z}\times\{0\})$ and $\eta \equiv \xi$ on $\mathbf{Z}\times\{0\}$. We ask for the correlation of the spins at the sites $x = (0, 1)$ and $y = (k, 1)$ with $0 < |k| < n$.

Theorem 7.13 *In the situation just described, suppose ξ is such that*

$$\sum_{j=0}^{m} \xi(j,0) \ge 8m/9 \quad \textit{and} \quad \sum_{j=-m}^{-1} \xi(j,0) \ge 8m/9$$

for sufficiently large m, and let η be defined as above. Then there are constants $c > 0$ and $C < \infty$ (not depending on n) such that

$$|\mu^{\eta}_{\beta,\Lambda_n}(X(0,1); X(k,1))| \le C\, e^{-c|k|}$$

whenever β and $|k|$ are sufficiently large and $n > |k|$.

Sketch proof: We proceed as in the proof of Theorem 7.12. In dealing with the right-hand side of (55) we must take into account that possibly $\partial C \cap \partial\Lambda \neq \emptyset$. We therefore replace ∂C by $\partial C \setminus \partial\Lambda$ in the product term and also estimate the probabilities of intersections by conditional probabilities, yielding the upper bound

$$\mu^{+,\eta}_{\beta,C}(X \equiv -1 \text{ on } D)\, \mu^{+,\eta}_{\beta,C}(X \equiv -1 \text{ on } \partial_i C \setminus D)$$

for the summands on the right-hand side of (55). Here, $\mu^{+,\eta}_{\beta,C}$ stands for the Gibbs distribution in C with boundary condition equal to $+1$ on $\partial C \cap \Lambda$ and equal to η on $\partial C \cap \partial\Lambda$. To derive the theorem we need to replace the Peierls estimate (56) by a similar bound on the last product. The exponential decay of correlations then again follows by simple combinatorics and summations.

To make the influence of the boundary condition η explicit we exploit a contour representation leading to the estimate

$$\begin{aligned} &\mu^{+,\eta}_{\beta,C}(X \equiv -1 \text{ on } D)\, \mu^{+,\eta}_{\beta,C}(X \equiv -1 \text{ on } \partial_i C \setminus D) \\ &\qquad \le \sum_{\substack{\Gamma,\Gamma' \text{ inside } C \\ \Gamma \text{ compatible with } D \\ \Gamma' \text{ compatible with } \partial_i C\setminus D}} \prod_{\gamma\in\Gamma} w_\eta(\gamma) \prod_{\gamma'\in\Gamma'} w_\eta(\gamma')\,. \end{aligned} \tag{57}$$

The right-hand side is defined as follows. For any configuration $\sigma \in \{+1, -1\}^{\Lambda}$ we draw horizontal resp. vertical lines of unit length between neighboring sites of opposite spins, doing as if the boundary spins were all plus; we then obtain a disjoint union of closed non-self-intersecting polygonal curves. Each of these curves is called a contour γ, and a set Γ of contours arising in this way is called compatible. We thus have a one-to-one correspondence between spin configurations σ and compatible sets Γ of contours. If $\sigma \equiv -1$ on D, then each component of D is surrounded by some contour γ (i.e., belongs to the interior $\mathrm{Int}\,\gamma$ of γ); the smallest contours surrounding the components of D are collected into a set Γ of contours. Each set Γ arising in this way is called *compatible with* D. The probability that a given set Γ of contours occurs is not larger than $\prod_{\gamma\in\Gamma} w_\eta(\gamma)$, where

$$w_\eta(\gamma) = \exp\Big[-2\beta|\gamma| + 2\beta \sum_{x\in \mathrm{Int}\,\gamma} \sum_{y\in\partial\Lambda: y\sim x} (1-\eta(y))\Big]$$

and $|\gamma|$ is the length of γ; this can be seen by comparing the probability of a configuration containing Γ with the probability of the configuration obtained by flipping the spins in $\bigcup_{\gamma\in\Gamma} \mathrm{Int}\,\gamma$. These observations establish the inequality (57).

Note that the weight $w_\eta(\gamma)$ of a contour γ depends on the boundary configuration η; this is because we have chosen to draw the contours for plus boundary conditions rather than η. It follows that $w_\eta(\gamma)$ does not necessarily tend to zero when $|\gamma|$ grows to infinity; this is in contrast with the case $\eta \equiv +1$. The standard low temperature expansion would therefore become much more complicated. However, if the density of plus spins in η is sufficiently large, or if $\partial\mathrm{Int}\,\gamma \cap \partial\Lambda$ is rather small, the standard weight $\exp[-2\beta|\gamma|]$ of the Ising contours will dominate, and the right-hand side of (57) can be estimated, as we will show now.

We unite the sets Γ, Γ' in (57) into a single set of contours $\tilde\Gamma = \Gamma \cup \Gamma'$. The contours in $\tilde\Gamma$ can overlap, but a site of $\partial_i C$ can only belong to the interior of at most two contours. On the other hand, every site of $\partial_i C$ is in the interior of at least one contour of $\tilde\Gamma$, and $|\partial_i C| \geq k$, the distance of the two spins considered. These ingredients allow us to control the sum on the right-hand side of (57). If k is so large that the density of plus spins in ξ between 0 and $(k,0)$ exceeds $8/9$ then we find for any collection of contours $\tilde\Gamma = \Gamma \cup \Gamma'$ as above

$$\sum_{\tilde\gamma\in\tilde\Gamma} \sum_{x\in\mathrm{Int}\,\tilde\gamma} \sum_{y\in\partial\Lambda: y\sim x} (1-\eta(y)) \leq 5/9 \sum_{\tilde\gamma\in\tilde\Gamma} |\tilde\gamma| \,.$$

This yields

$$\prod_{\gamma\in\Gamma} w_\eta(\gamma) \prod_{\gamma'\in\Gamma'} w_\eta(\gamma') \leq \prod_{\gamma\in\Gamma} \exp[-8/9\,\beta\,|\gamma|] \prod_{\gamma'\in\Gamma'} \exp[-8/9\,\beta\,|\gamma'|] \,.$$

At this point the standard arguments take over (with β replaced by $4\beta/9$), leading to an exponential estimate of (57). For example, one can conclude the proof along the lines of Lemma 2.5 of [37]. □

8 Phase transition and percolation

Typically, two ends of the phase diagram are amenable to mathematical analysis. One is the high temperature, or low density, regime which was discussed in the previous section and in which the system can be viewed as a small perturbation of an independent spin system. The other end is the low temperature regime which we will now consider. At low temperatures, the energy dominates over the entropy which comes from the thermal fluctuations of the spins. One therefore expects that the spin configuration is typically similar to some frozen zero temperature state, which is a configuration of minimal energy and thus called a *ground state*. The similarity of a low temperature state with a ground state is conveniently described in geometric terms: one imagines that the spins which agree with the given ground state form an infinitely extended sea, whereas those spins which have chosen to deviate from the ground state are confined to interspersed finite islands. This is, of course, a picture of percolation theory: spins that agree with the ground state form a unique infinite cluster. We are thus led to the concept of *agreement percolation*, which will be discussed in the first part of this section.

In fact, agreement percolation is intimately related to the existence of a phase transition. If several distinct ground states exist, we may hope to find at low temperatures also several equilibrium phases which can be distinguished by agreement percolation with respect to the different ground states. One may ask further whether the geometric picture that applies to low temperatures remains valid throughout the whole non-uniqueness region. Physically, this is a matter of stability of the ground states. Mathematically, it means to look for conditions under which distinct Gibbs measures allow distinct stochastic-geometric characterizations.

We will approach this question from two different sides. In Sections 8.1 to 8.4 we investigate whether "phase transition implies percolation". We study a fixed equilibrium phase μ in the non-uniqueness region which, by its very construction, can be viewed as a random perturbation of some ground state η. We then will see that, in many cases, spins that agree with η do percolate. After a general discussion of agreement percolation in Section 8.1, we investigate this concept in the subsequent subsections for some specific models including the Ising ferromagnet and the Potts model. In the case of the planar lattice $\mathbf{Z}^2$ with its limited geometric possibilities we will also see that conversely, the absence of phase transition sometimes implies an absence of percolation, and that in the case of phase transition one has restrictions on the number of phases. (Methodologically, these results still run under the heading "phase transition implies percolation".) In the last Section 8.5 the converse will be treated more systematically and under a different aspect: we will show that at low temperatures one has percolation of bonds along which the interaction energy is minimal, and we will see that such a *ground-energy bond percolation* often implies a phase transition. Taken together, these results will show that in various models phase transition comes along with the existence of a ground-state sea with finite islands (deviation islands) on which the spins deviate from the ground state, and vice versa.

A theory developing this picture in much more detail is the Pirogov–Sinai theory of phase transition which deals with the low temperature phase diagram in the presence of several stable ground states. One basic idea of this theory is to treat the finite deviation islands of a low temperature system as the constituents of a low density gas of hard-core particles. While the Pirogov–Sinai theory is intimately related to the subject of this section, it is much too involved to be developed here. There is, however, a number of expositions which may serve as general introductions and in which many additional references can be found. We mention only [37, 69, 76, 213, 214] and [228] to [230]. Here we will concentrate on more specific results which are partly beyond the Pirogov–Sinai theory, in that they are not limited to low temperatures but rather apply to the full non-uniqueness region, and comment occasionally on some relationships. In particular, the results of Section 8.5 are similar in spirit to this theory.

8.1 Agreement percolation from phase coexistence

We consider again the general setting of Section 2. $(\mathcal{L}, \sim)$ is an arbitrary locally finite graph, S is a finite set, and $\Omega = S^{\mathcal{L}}$. Suppose μ is a random field and $\eta \in \Omega$ a fixed configuration. We consider the event $\{x \overset{\eta}{\longleftrightarrow} \infty\}$ that $x \in \mathcal{L}$ belongs to an infinite cluster of the random set $R(\eta) = \{y \in \mathcal{L} : X(y) = \eta(y)\}$, and we say that μ *exhibits agreement percolation for* η if $\mu(x \overset{\eta}{\longleftrightarrow} \infty) > 0$ for some $x \in \mathcal{L}$. In short, we will then simply speak of η-percolation. To visualize such an agreement, it may be convenient to think of a reduced description of μ in terms of its image under the map $s_\eta : \Omega \to \{0,1\}^{\mathcal{L}}$, which describes local agreement and disagreement with η, and is defined by

$$(s_\eta(\sigma))(x) = \begin{cases} 1 & \text{if } \sigma(x) = \eta(x), \\ 0 & \text{otherwise.} \end{cases} \tag{58}$$

With this mapping, we can write $\{x \overset{\eta}{\longleftrightarrow} \infty\} = s_\eta^{-1}\{x \leftrightarrow \infty\}$.

We are interested here in the case when μ is a Gibbs measure for the Hamiltonian (1), and η is an associated ground state. We say that a configuration $\eta \in \Omega$ is a *ground state* or, more explicitly, a ground state configuration for the relative Hamiltonian (2), if $H(\sigma|\eta) \geq 0$ for any local modification (or "excitation") σ of η. In other words, η is a ground state if, for any region $\Lambda \in \mathcal{E}$, the configuration η minimizes the energy in Λ when η_{Λ^c} is fixed. One should note in this context that, in the low temperature limit $\beta \uparrow \infty$, the finite-volume Gibbs distribution $\mu^\eta_{\beta,\Lambda}$ from (4) tends to the equidistribution on the set of all configurations σ of minimal energy $H(\sigma|\eta)$. This fact suggests that, at least in some cases, the low temperature phase diagram is only a slight deformation of the zero-temperature phase diagram describing the structure of ground states. This is precisely the subject of Pirogov–Sinai theory which provides sufficient conditions for this to hold, proposes a construction of low temperature phases as perturbations of ground states, and also shows that the size distribution of the deviation islands has exponential decay.

Suppose next that the Gibbs measure μ is related to the ground state η in some way. For example, μ might be obtained as the infinite volume limit of the finite volume Gibbs distributions $\mu^\eta_{\beta,\Lambda}$ with boundary condition η, possibly along some subsequence. (Under the conditions of the Pirogov–Sinai theory such a limit always exists.) In the case of a phase transition, when other phases than μ exist and one is interested in characteristic properties of μ, one expects that the relationship between μ and η becomes manifest in a macroscopic pattern of the typical configurations, in that μ shows η-percolation. In short, we ask for the validity of the hypothesis

$$\begin{aligned} &|\mathcal{G}(\beta H)| > 1,\ \mu \text{ is extremal in } \mathcal{G}(\beta H) \text{ and related to a ground state } \eta \in \Omega \\ &\Longrightarrow \ \mu(x \overset{\eta}{\longleftrightarrow} \infty) > 0 \quad \forall\, x \in \mathcal{L}\,. \end{aligned} \tag{59}$$

In the specific cases considered below it will always be clear in what sense μ and η are related; typically, μ will be a limiting Gibbs measure with boundary condition η. We emphasize that (59) does not hold in general; a counter-example can be constructed by combining many independent copies of the Ising ferromagnet to a layered system, see the discussion after Proposition 8.3 below. Also, even when (59) holds, it does not necessarily imply that the phase μ is uniquely characterized by the property of η-percolation.

How can one establish (59)? In the context of the Ising model, Coniglio et al. [62] and Russo [208] developed a convenient citerion which is based on a multidimensional analog of the strong Markov property and thus can be used for general Markov random fields [36, 100]. One version is as follows.

Theorem 8.1 *Let $(\mathcal{L},\sim)$ be a locally finite graph, μ a Markov field on $\Omega = S^{\mathcal{L}}$, and $\eta \in \Omega$ any configuration. Suppose there exist a constant $c \in \mathbf{R}$ and a local function $f : \Omega \to \mathbf{R}$ depending only on the configuration in a connected set Δ, such that $\mu(f) > c$ but*

$$\mu(f \,|\, X \equiv \xi \text{ on } \partial\Gamma) \leq c \tag{60}$$

for all finite connected sets $\Gamma \supset \Delta$ and all $\xi \in \Omega$ with $s_\eta(\xi) \equiv 0$ on $\partial\Gamma$. Then $\mu(\Delta \overset{\eta}{\longleftrightarrow} \infty) > 0$, i.e., μ exhibits agreement percolation for η.

Proof: Suppose by contraposition that $\mu(\Delta \overset{\eta}{\longleftrightarrow} \infty) = 0$. For any $\varepsilon > 0$ we can then choose some finite $\Lambda \supset \Delta$ such that $\mu(\Delta \overset{\eta}{\longleftrightarrow} \Lambda^c) < \varepsilon$. For $\xi \notin \{\Delta \overset{\eta}{\longleftrightarrow} \Lambda^c\}$, there exists a connected set Γ such that $\Delta \subset \Gamma \subset \Lambda$ and $s_\eta(\xi) \equiv 0$ on $\partial\Gamma$; we simply let Γ be the union of Δ and all η-clusters meeting $\partial\Delta$. As in the proof of Theorem 6.10, we let $\Gamma(\xi)$ be the largest such set. For $\xi \in \{\Delta \overset{\eta}{\longleftrightarrow} \Lambda^c\}$ we put $\Gamma(\xi) = \emptyset$. Then, for each finite connected set $\Gamma \neq \emptyset$, the event $\{\xi : \Gamma(\xi) = \Gamma\}$ depends only on the configuration in $\Lambda \setminus \Gamma$, whence by the Markov property $\mu(f \,|\, \Gamma(\cdot) = \Gamma)$ is an average of the conditional probabilities that appear in assumption (60). ¿From this we obtain

$$\mu(f) \leq c\,\mu(\Gamma(\cdot) \neq \emptyset) + \mu(|f|\, I_{\{\Gamma(\cdot)=\emptyset\}}) < c + \varepsilon \, \|f\| \, .$$

Letting $\varepsilon \to 0$ we find $\mu(f) \leq c$, contradicting our assumption. □

In most applications we will have a natural candidate for the function f. Whenever distinct phases do exist, they can be distinguished by some order parameter, viz. a local function f having different expectations for the two phases. If, in addition, some stochastic monotonicity is available then we can hope to establish (60). In fact, the percolation phenomena stated in Examples 5.11 to 5.15 can be deduced from a slight modification of Theorem 8.1. We will not go into the details of these examples which are treated in [36], but rather apply Theorem 8.1 to our standard examples.

8.2 Plus-clusters for the Ising ferromagnet

The idea of agreement percolation was first developed in the context of the ferromagnetic Ising model [62, 208]. Let us apply Theorem 8.1 to this standard case. We only consider the case of no external field, i.e., we set $h = 0$, so the only parameter is the inverse temperature $\beta > 0$. We are interested in agreement percolation for the constant configurations $\eta \equiv +1$ resp. $\eta \equiv -1$, which are the only periodic ground states of the model. We write $\overset{+}{\longleftrightarrow}$ resp. $\overset{-}{\longleftrightarrow}$ for the corresponding connectedness relation. Our first result shows that if there is a phase transition then there is plus-percolation for each Gibbs measure except the minus-phase μ_β^-; that is, assertion (59) holds for $\eta \equiv +1$. This result (due to [208]) is valid for an arbitrary locally finite graph $(\mathcal{L},\sim)$ with finite critical inverse temperature β_c.

Theorem 8.2 *Let μ be an arbitrary Gibbs measure for the ferromagnetic Ising model with parameters $\beta > 0$, $h = 0$. If $\mu \neq \mu_\beta^-$ then $\mu(x \stackrel{+}{\longleftrightarrow} \infty) > 0$ for all $x \in \mathcal{L}$. In particular, if $\beta > \beta_c$ then $\mu_\beta^+(x \stackrel{+}{\longleftrightarrow} \infty) > 0$ for all $x \in \mathcal{L}$.*

Proof: By the sandwiching inequality (15) and Proposition 4.12, there exists a site $x \in \mathcal{L}$ such that $\mu(X(x) = 1) > c \equiv \mu_\beta^-(X(x) = 1)$. On the other hand, the analogue of inequality (14) for the minus boundary condition shows that

$$\mu(X(x) = 1 \,|\, X \equiv -1 \text{ on } \partial\Gamma) = \mu_{\beta,\Gamma}^-(X(x) = 1) \leq c$$

for every finite $\Gamma \ni x$. Theorem 8.1 thus gives the result for the x at hand. In view of the finite energy property of μ, this extends easily to all other $x \in \mathcal{L}$. □

Let us rephrase the last statement of Theorem 8.2 as follows: below the critical temperature the plus spins percolate in the plus phase and, by symmetry, the minus spins percolate in the minus phase μ_β^-. In the case of graphs with symmetry axes, this statement allows an interesting refinement.

Proposition 8.3 *Suppose $(\mathcal{L}, \sim)$ admits an involutive graph automorphism r which maps a subset $\mathcal{H} \subset \mathcal{L}$ onto its complement $\mathcal{H}^c$, and that for $x \in \mathcal{H}$, $y \in \mathcal{H}^c$ either $x \not\sim y$ or $y = rx$. For $x \in \mathcal{H}$ let $\{x \stackrel{+r+}{\longleftrightarrow} \infty \text{ in } \mathcal{H}\}$ be the event that there exists an infinite path γ in $\mathcal{H}$ starting from x such that all spins along both γ and its reflection image $r\gamma$ are positive. If $\mu_\beta^+ \neq \mu_\beta^-$ then $\mu_\beta^+(x \stackrel{+r+}{\longleftrightarrow} \infty \text{ in } \mathcal{H}) > 0$ for all $x \in \mathcal{H}$.*

One natural case to think of is when $\mathcal{L} = \mathbf{Z}^d$ for $d \geq 2$, $\mathcal{H}$ a halfspace with boundary orthogonal to an axis, and r the associated reflection. The proposition then asserts that μ_β^+-almost surely there exists an infinite connected mirror-symmetric set of plus spins. Another interesting case is when $\mathcal{L}$ consists of two disjoint copies of a graph $\mathcal{H}$ which are not connected to each other by any bond. In this case, $\mu_\beta^+ = \mu_{\beta,\mathcal{H}}^+ \times \mu_{\beta,\mathcal{H}}^+$, and the statement is that two independent realizations of $\mu_{\beta,\mathcal{H}}^+$ exhibit simultaneous plus-percolation; in this case, the preceding proposition was observed by Giacomin et al. [100]. It is, however, not possible to take an arbitrarily large number k of independent realizations $X_1, \ldots, X_k$ of $\mu_{\beta,\mathcal{H}}^+$, at least when $\mathcal{H}$ has bounded degree N. For, if p_c is the Bernoulli site percolation threshold of $\mathcal{H}$ and k is so large that

$$\sup_{x \in \mathcal{H}} \mu_{\beta,\{x\}}^+(X(x) = 1)^k < p_c$$

then the set $\{x \in \mathcal{H} : X_1(x) = \cdots = X_k(x) = 1\}$ does not percolate. This follows from a standard domination argument. Since the layered system consisting of k independent copies of the Ising model with $\beta > \beta_c$ certainly exhibits a phase transition, we see that hypothesis (59) does not hold in general.

Proof of Proposition 8.3: We identify each $\xi \in \Omega$ with $(\xi(x), \xi(rx))_{x \in \mathcal{H}} \in S^{\mathcal{H}}$, where $S = \{-1, 1\}^2$. The event under consideration then corresponds to η-percolation for the configuration $\eta \in S^{\mathcal{H}}$ with $\eta(x) = (1,1)$ for all $x \in \mathcal{H}$. Let $f = X(x) + X(rx)$. Then $\mu_\beta^+(f) = 2\,\mu_\beta^+(X(x)) > 0$ by the r-invariance of μ_β^+. On the other hand, let $\Gamma \subset \mathcal{H}$ be a finite set containing x, and $\tilde{\Gamma} = \Gamma \cup r\Gamma$. If $(\xi, \xi') \in S^{\mathcal{H}} = \Omega$ with $s_\eta(\xi, \xi') \equiv 0$ on $\partial_{\mathcal{H}}\Gamma = \partial\Gamma \cap \mathcal{H}$, then $\xi' \preceq -\xi$ on $\partial_{\mathcal{H}}\Gamma$, and therefore $(\xi, \xi') \preceq (\xi, -\xi)$ (as elements of Ω) on $\partial\tilde{\Gamma} = \partial_{\mathcal{H}}\Gamma \cup r\,\partial_{\mathcal{H}}\Gamma$. We can thus write

$$\begin{aligned}
\mu_\beta^+(f \,|\, (X, X') \equiv (\xi, \xi') \text{ on } \partial\Gamma) &= \mu_{\beta,\tilde{\Gamma}}^{(\xi,\xi')}(X(x)) + \mu_{\beta,\tilde{\Gamma}}^{(\xi,\xi')}(X(rx)) \\
&\leq \mu_{\beta,\tilde{\Gamma}}^{(\xi,-\xi)}(X(x)) + \mu_{\beta,\tilde{\Gamma}}^{(\xi,-\xi)}(X(rx)) \;=\; 0
\end{aligned}$$

by Lemma 4.13 and the symmetry under r and simultaneous spin flip. The proposition thus follows from Theorem 8.1. $\square$

In the remaining part of this subsection we consider the case of the square lattice $\mathcal{L} = \mathbf{Z}^2$, in which we can obtain much stronger conclusions. The following result gives a complete characterization of the non-uniqueness regime of the parameter space in terms of percolation of plus spins in the Gibbs measure μ_β^+. It is due to Coniglio et al. [62]; see also [131].

Corollary 8.4 *For the Ising ferromagnet on the square lattice $\mathbf{Z}^2$ with no external field and inverse temperature β, the μ_β^+-probability of having an infinite plus-cluster is* 0 *in the uniqueness regime $\beta \le \beta_c$, and* 1 *in the non-uniqueness regime $\beta > \beta_c$.*

Proof: The existence of an infinite cluster of plus spins is a tail event and thus, by the extremality of μ_β^+, has probability 0 or 1. The case $\beta > \beta_c$ is thus covered by Theorem 8.2. For $\beta \le \beta_c$, μ_β^+ coincides with μ_β^-. Thus, if an infinite plus-cluster existed with probability 1 then, by symmetry, an infinite minus-cluster would also exist, in contradiction to Theorem 5.18; the assumptions of this theorem are satisfied by Proposition 4.14. $\square$

Combining the corollary above with Proposition 4.16, we can also obtain some bounds for the percolative region of the Ising model for $h \neq 0$; see [8] for a detailed discussion.

The equivalence of non-uniqueness and percolation just observed for the Ising model on $\mathbf{Z}^2$ cannot be expected to hold for non-planar graphs. Consider, for example, the Ising model on the cubic lattice $\mathbf{Z}^3$. For $\beta = 0$ uniqueness certainly holds, and plus-percolation is equivalent to Bernoulli site percolation on $\mathbf{Z}^3$ with parameter 1/2. But a result of [47] states that $p_c(\mathbf{Z}^3) < 1/2$. The plus spins thus percolate at $\beta = 0$. In view of Proposition 4.16, this is still the case for sufficiently small β, so that plus-percolation does occur in a non-trivial part of the uniqueness region.

For the planar graph $\mathbf{Z}^2$, however, Theorem 5.18 does not only imply the equivalence of phase transition and percolation, but also gives some information on the number of phases in the non-uniqueness region. As a warm-up let us show that, for the Ising ferromagnet on $\mathbf{Z}^2$ at inverse temperature $\beta > \beta_c$, there are no other translation and rotation invariant extremal Gibbs measures than μ_β^+ and μ_β^-. For, suppose another such phase μ existed. By Theorem 8.2 and the Burton-Keane uniqueness theorem 5.17, there exist unique infinite plus- and minus-clusters with μ-probability 1. As an extremal Gibbs measure, μ has positive correlations; recall the paragraph below Proposition 4.14. Proposition 5.19 thus shows that μ cannot exist.

The statement just shown is a weak version of the following result which characterizes all translation invariant Gibbs measures. In fact, it is sufficient to assume periodicity, which means invariance under the translation subgroup $(\theta_x)_{x\in p\mathbf{Z}^2}$ for some $p > 1$.

Proposition 8.5 *Any periodic Gibbs measure μ for the Ising ferromagnet on $\mathcal{L} = \mathbf{Z}^2$ with no external field and inverse temperature $\beta > \beta_c$ is a mixture of the two phases μ_β^+ and μ_β^-.*

Under the condition of translation invariance, this proposition was first derived for large β by Gallavotti and Miracle-Sole [88], and later for all $\beta > \beta_c$ by Messager and Miracle-Sole [177] using some specific correlation inequalities; it follows also from the Onsager-formula for the free energy density and a result of Lebowitz [152]. We will give a geometric proof below.

Remarkably enough, one can go one step further: each (not necessarily periodic) Gibbs measure for the Ising model on the square lattice is a mixture of the plus-phase and the minus-phase, and thus automatically automorphism invariant. This beautiful result was obtained independently by Aizenman [3] and Higuchi [130] based on the work of Russo [208]. For more general two-dimensional systems the absence of non-translation-invariant Gibbs measures at sufficiently low temperatures was proved in [69]. In three or more dimensions, however, non-translation invariant phases of the Ising model do exist; this is a famous result of Dobrushin [66], see also [22] for a short proof.

Theorem 8.6 (Aizenman–Higuchi) *For the Ising ferromagnet on $\mathcal{L} = \mathbf{Z}^2$ with no external field and inverse temperature $\beta > \beta_c$, μ_β^+ and μ_β^- are the only phases, and any other Gibbs measure is a mixture of these two.*

The proof is a masterpiece of random-geometric analysis of equilibrium phases and contains various ingenious ideas, but unfortunately it is too long to be sketched here. For the full result we thus need to refer to the original papers cited above, as well as to the survey [4]. However, to provide an idea of some of the geometric ideas involved we will now give a (new) geometric proof of Proposition 8.5. This proof resulted from discussions of the first author with Y. Higuchi.

In this proof we need to consider infinite clusters in halfplanes. Here, we say that a set $\mathcal{H} \subset \mathbf{Z}^2$ is a halfplane if $\mathcal{H}$ is a translate of either the upper halfplane $\{x = (x_1, x_2) \in \mathbf{Z}^2 : x_2 \geq 0\}$ or its complement, the lower halfplane, or a translate of the right and left halfplanes which are similarly defined. The next lemma provides a first step in the proof of Proposition 8.5.

Lemma 8.7 *Consider the Ising ferromagnet on $\mathbf{Z}^2$, and let D be the event that for at least one halfplane $\mathcal{H}$ in $\mathbf{Z}^2$, both $\mathcal{H}$ and $\mathcal{H}^c$ contain an infinite cluster of the same sign. Then $\mu(D) = 1$ for all $\mu \in \mathcal{G}(\beta H)$ and $\beta > \beta_c$.*

Proof: Since each Gibbs measure is a mixture of extremal Gibbs measures, we only need to show that $\mu(D) = 1$ for any extremal μ. Suppose the contrary. Since D is tail measurable, it then follows that $\mu(D) = 0$ for some extremal μ. We will show that this is impossible.

Step 1: Let $\mathcal{H}$ be any halfplane, r the reflection of $\mathbf{Z}^2$ which maps $\mathcal{H}$ onto $\mathcal{H}^c$, and $\tau : \sigma \to -\sigma$ the spin flip on Ω. We show that $\mu = \mu \circ r \circ \tau$. Since $\mu(D) = 0$, at least one of the halfplanes $\mathcal{H}$ and $\mathcal{H}^c$ contains no infinite minus-cluster, and this or the other halfplane contains no infinite plus-cluster. In view of the tail triviality of μ, we can assume that $\mathcal{H}$ contains no infinite minus-cluster μ-almost surely. Hence, for any given $\Delta \in \mathcal{E}$ and μ-almost every $\xi \in \Omega$, there exists an r-symmetric region $\Gamma(\xi) \in \mathcal{E}$ such that $\Gamma(\xi) \supset \Delta$ and $\xi \equiv 1$ on $\partial\Gamma(\xi) \cap \mathcal{H}$. The last property implies that $\xi \succeq r \circ \tau(\xi)$ on $\partial\Gamma(\xi)$, and using Lemma 4.13 and the flip-reflection symmetry of H we find that

$$\mu^{\xi}_{\beta,\Gamma(\xi)} \succeq_{\mathcal{D}} \mu^{r\circ\tau(\xi)}_{\beta,\Gamma(\xi)} = \mu^{\xi}_{\beta,\Gamma(\xi)} \circ r \circ \tau \text{ on } \mathcal{F}_\Delta \, .$$

Assuming that $\Gamma(\xi)$ is maximal in a large box $\Lambda \supset \Delta$, we can apply the Markov property of μ in the same way as in the proof of Theorem 8.1. This yields that $\mu \succeq_{\mathcal{D}} \mu \circ r \circ \tau$ on $\mathcal{F}_\Delta$ for any Δ, and thus $\mu \succeq_{\mathcal{D}} \mu \circ r \circ \tau$. (The preceding argument is a variant of an idea of Russo [208].) Using the absence of infinite plus-clusters in $\mathcal{H}$ or $\mathcal{H}^c$ we find analogously that $\mu \preceq_{\mathcal{D}} \mu \circ r \circ \tau$. Hence $\mu = \mu \circ r \circ \tau$ as claimed.

Step 2: Here we use a variant of Zhang's argument which was explained in the proof of Theorem 5.18. To begin, we observe that the composition of two reflections in parallel axes is a translation. Step 1 therefore implies that μ is periodic with period 2. The flip-reflection symmetry of μ implies further that μ is different from μ_β^+ and μ_β^-, so that (by Theorem 8.2) there exist both an infinite plus- and an infinite minus-cluster μ-almost surely. By the Burton–Keane uniqueness theorem 5.17, these infinite clusters are almost surely unique. We now choose a square $\Lambda = [-n, n-1]^2 \cap \mathbf{Z}^2$ so large that $\mu(\Lambda \stackrel{+}{\longleftrightarrow} \infty) > 1 - 10^{-3}$. Since μ is extremal, μ has positive correlations. By the argument leading to (23) we thus obtain that $\mu(\partial_k\Lambda \stackrel{+}{\longleftrightarrow} \infty) > 1 - 10^{-3/4}$ for some $k \in \{1, \ldots, 4\}$, where $\partial_k\Lambda$ is the intersection of $\partial\Lambda$ with the k'th quadrant (relative to the axes $\{x_2 = -1/2\}$ and $\{x_1 = -1/2\}$). For definiteness, we assume that $k = 1$. By the flip-reflection symmetry, it follows that the intersection

$$\{\partial_1\Lambda \stackrel{+}{\longleftrightarrow} \infty,\ \partial_2\Lambda \stackrel{-}{\longleftrightarrow} \infty,\ \partial_3\Lambda \stackrel{+}{\longleftrightarrow} \infty,\ \partial_4\Lambda \stackrel{-}{\longleftrightarrow} \infty\}$$

has probability at least $1 - 4 \cdot 10^{-3/4} > 0$, which is impossible because of the uniqueness of the infinite clusters. This contradiction concludes the proof of the lemma. □

Proof of Proposition 8.5: Let μ be any Gibbs measure invariant under $(\theta_x)_{x \in p\mathbf{Z}^2}$ for some $p > 1$. Using the ergodic decomposition, we can assume that μ is in fact ergodic with respect to this group of translations. By Lemma 8.7, there exists a pair $(\mathcal{H}, \mathcal{H}^c)$ of halfplanes such that, with positive probability, both $\mathcal{H}$ and $\mathcal{H}^c$ contain infinite clusters of spins of the same constant sign. For definiteness, suppose $\mathcal{H}$ is the upper halfplane, and the sign is plus. In view of the finite energy property, it then follows that also $\mu(A_0) > 0$, where for $k \in \mathbf{Z}$

$$A_k = \{(k, 0) \stackrel{+}{\longleftrightarrow} \infty \text{ both in } \mathcal{H} \text{ and } \mathcal{H}^c\}\ .$$

Let A be the event that A_k occurs for infinitely many $k < 0$ and infinitely many $k > 0$. The horizontal periodicity and Poincaré's recurrence theorem (or the ergodic theorem) then show that $\mu(A_0 \setminus A) = 0$, and therefore $\mu(A) > 0$.

Next, let B be the event that there exists an infinite minus-cluster. We claim that $\mu(A \cap B) = 0$. Indeed, suppose $\mu(A \cap B) > 0$. Since A is tail measurable and horizontally periodic, we can use the finite energy property and horizontal periodicity of μ as above to show that the event

$$C = A \cap \{(k, 0) \stackrel{-}{\longleftrightarrow} \infty \text{ for infinitely many } k < 0 \text{ and infinitely many } k > 0\}$$

has positive probability. But on C there exist infinitely many minus-clusters, which is impossible by the Burton–Keane theorem.

To complete the proof, we note that $\mu(B) \le \mu(A^c) < 1$, and thus $\mu(B) = 0$ by ergodicity. In view of Theorem 8.2, this means that $\mu = \mu_\beta^+$. In the case considered, the proposition is thus proved. The other cases are similar; in particular, in the case of negative sign we find that $\mu = \mu_\beta^-$. □

8.3 Constant-spin clusters in the Potts model

Consider the q-state Potts model on the lattice $\mathcal{L} = \mathbf{Z}^d$ introduced in Section 3.3, $q, d \ge 2$, and recall the results of Section 6.3 on the phase transition in this model. The periodic ground states are the constant configurations $\eta_i \equiv i$, $1 \le i \le q$. We write $\stackrel{i}{\longleftrightarrow}$ for the agreement connectivity relation relative to η_i, and we consider the limiting Gibbs measure $\mu^i_{\beta,q}$ at inverse temperature β associated to η_i, which exists by Proposition 6.9. As a further illustration of assertion (59), we show that $\mu^i_{\beta,q}$ exhibits i-percolation whenever there is a phase transition. This is a Potts-counterpart of Theorem 8.2. For its proof, we use the random-cluster representation rather than Theorem 8.1 because for $q > 2$ there is no stochastic monotonicity available in the spin configuration.

Theorem 8.8 *For the Potts model on $\mathbf{Z}^d$ at any inverse temperature β with $|\mathcal{G}(\beta H)| > 1$, we have $\mu^i_{\beta,q}(x \stackrel{i}{\longleftrightarrow} \infty) > 0$ for all $x \in \mathbf{Z}^d$ and $i \in \{1,\dots,q\}$.*

Proof: By translation invariance we can choose $x = 0$. In Theorem 6.10 we have seen that $\phi^1_{p,q}(0 \leftrightarrow \infty) = c > 0$ for $\beta > \beta_c$, where $p = 1 - e^{-2\beta}$ as usual. In view of (30), this means that $\phi^1_{p,q,\Lambda}(0 \leftrightarrow \Lambda^c) \geq c$ for all $\Lambda \ni 0$. But for the Edwards-Sokal coupling $P^i_{p,q,\Lambda}$ of $\mu^i_{\beta,q,\Lambda}$ and $\phi^1_{p,q,\Lambda}$ (defined before Proposition 6.9) we have $\{0 \leftrightarrow \Lambda^c\} \subset \{0 \stackrel{i}{\longleftrightarrow} \Lambda^c\}$ almost surely, so that $\mu^i_{\beta,q,\Lambda}(0 \stackrel{i}{\longleftrightarrow} \Lambda^c) \geq c$. In particular, $\mu^i_{\beta,q,\Lambda}(0 \stackrel{i}{\longleftrightarrow} \Delta^c) \geq c$ whenever $0 \in \Delta \subset \Lambda$. Letting first $\Lambda \uparrow \mathbf{Z}^d$ and then $\Delta \uparrow \mathbf{Z}^d$ we find that $\mu^i_{\beta,q}(0 \stackrel{i}{\longleftrightarrow} \infty) \geq c$, and the theorem follows. □

Next we ask for a converse stating that "agreement percolation implies phase transition". As we already noticed in the case of the Ising model, this can be expected to hold only in the case of a planar lattice. But then a counterpart of Corollary 8.4 does indeed hold, as was shown by L. Chayes [56].

Theorem 8.9 *For the unique Gibbs measure $\mu_{\beta,q}$ of the q-state Potts model on the square lattice $\mathbf{Z}^2$ at inverse temperature $\beta < \beta_c$, we have $\mu_{\beta,q}(\exists \text{ an infinite } i\text{-cluster}) = 0$ for all $i \in \{1,\dots,q\}$.*

The strategy of proving this theorem is the same as that in the proof of Corollary 8.4. Suppose the i-spins percolate in $\mu_{\beta,q}$ for some i. Then, by symmetry, this holds for all i, so that in particular the 1-spins *and* the other spins percolate. Hence, Theorem 5.18 leads to a contradiction, provided we can show that the set of 1's has positive correlations. Theorem 8.9 thus follows from the following lemma.

Lemma 8.10 *Consider the phase $\mu^i_{\beta,q}$ of the q-state Potts model at any inverse temperature $\beta > 0$, and let the mapping s_i be defined by (58) with $\eta = \eta_i$, $i \in \{1,\dots,q\}$. Then the measure $\nu_{\beta,q} = \mu^i_{\beta,q} \circ s_i^{-1}$ has positive correlations.*

Sketch of proof: By symmetry, $\nu_{\beta,q}$ does not depend on i. For definiteness we set $i = 1$ in the following. Since the property of positive correlations is preserved under weak limits, it is sufficient to consider the finite volume Gibbs distribution $\mu^1_{\beta,q,\Lambda}$ and its image $\nu_{\beta,q,\Lambda} = \mu^1_{\beta,q,\Lambda} \circ s_1^{-1}$. By the FKG inequality, Theorem 4.11, it is further sufficient to show that $\nu_{\beta,q,\Lambda}$ is monotone. In terms of $\mu^1_{\beta,q,\Lambda}$ and the random field $Y = s_1(X)$, this means that the conditional probability

$$q_x(\xi) = \mu^1_{\beta,q,\Lambda}(Y(x) = 1 \,|\, Y \equiv \xi \text{ off } x)$$

is increasing in $\xi \in \{0,1\}^{\mathbf{Z}^d}$ for any $x \in \Lambda$. Since the boundary condition is fixed to be equal to 1 off Λ, we can assume that ξ is equal to 1 off Λ, and it is sufficient to prove the inequality $q_x(\xi) \leq q_x(\xi')$ for any two such ξ, ξ' that differ only at a single site $y \in \Lambda$ and are such that $\xi(y) = 0$ and $\xi'(y) = 1$. For such ξ, ξ', the inequality $q_x(\xi) \leq q_x(\xi')$ simply means that $Y(x)$ and $Y(y)$ are positively correlated under the conditional distribution $\mu_{x,y|\xi}$ of $\mu^1_{\beta,q,\Lambda}$ given that $Y \equiv \xi$ off $\{x,y\}$.

To show this we fix x, y, ξ. For $\mu_{x,y|\xi}$, we have $X \equiv 1$ on the complement of $\Delta = \{x,y\} \cup \{v \in \Lambda \setminus \{x,y\} : \xi(v) = 0\}$. We thus consider the graph G with vertex set Δ and edge set $\mathcal{B}(\Delta)$ consisting of all edges of $\mathcal{B}$ with both endpoints lying in Δ. If we knew that $Y(x) = Y(y) = 0$, then $\mu_{x,y|\xi}$ would be the distribution of a $(q-1)$-state Potts model on G with state space $\{2,3,\ldots,q\}$. Now that we don't know $Y(x)$ and $Y(y)$, $\mu_{x,y|\xi}$ is still a modification of this $(q-1)$-state Potts model, in which x and y are allowed to have the q'th spin value 1.

To describe this modification we suppose first that x and y are not adjacent. Let n_x be the number of neighbors v of x with $\xi(v) = 1$, and define n_y accordingly. The probability weight of $\mu_{x,y|\xi}$ then contains the additional biasing factor

$$\exp[2\beta(n_x I_{\{X(x)=1\}} + n_y I_{\{X(y)=1\}})]$$

which acts like an external field at x and y. For this modified Potts model, we can still define a modified random-cluster representation which gives any edge configuration $\zeta \in \{0,1\}^{\mathcal{B}(\Delta)}$ a probability proportional to

$$(q-1)^{k(\zeta)}(q-1+e^{2\beta n_x})^{k_x(\zeta)}(q-1+e^{2\beta n_y})^{k_y(\zeta)} \prod_{e \in \mathcal{B}(\Delta)} p^{\zeta(e)}(1-p)^{1-\zeta(e)} .$$

Here $p = 1 - e^{-2\beta}$, $k(\zeta)$ is the number of connected components *excluding singletons at x or y*, and $k_x(\zeta)$ and $k_y(\zeta)$ are the indicator functions of having a singleton connected component at x resp. y. A spin configuration with distribution $\mu_{x,y|\xi}$ is then obtained from the edge configuration by assigning spins at random uniformly from $\{2,\ldots,q\}$ to connected components, except for a singleton at x, where the spin is taken from $\{1,\ldots,q\}$ with probabilities proportional to $(e^{2\beta n_x}, 1, \ldots, 1)$, and similarly for a singleton at y. Just as in Corollary 6.5, this representation gives the desired positive correlation of $Y(x)$ and $Y(y)$ under $\mu_{x,y|\xi}$, provided we can show that k_x and k_y are positively correlated in the modified random-cluster model. Since these indicator variables are decreasing, it suffices to check that the modified random-cluster model has positive correlations, which follows from Theorem 4.11 by verifying that it is monotone; this, however, is similar to Lemma 6.6.

The case when x and y are neighbors is handled analogously; in fact, the positive correlation can only become stronger when x and y have an edge in common. $\square$

8.4 Further examples of agreement percolation

Here we treat the Ising antiferromagnet, the hard-core lattice gas, and the Widom–Rowlinson lattice model, and shortly mention the Ashkin–Teller model.

The Ising antiferromagnet. Consider the setting of Section 3.2. We need to assume that the underlying lattice $\mathcal{L}$ is bipartite, and thus splits off into two parts, $\mathcal{L}_{even}$ and $\mathcal{L}_{odd}$. If $|h| < 2d$, there exist two periodic ground states, η_{even} and $\eta_{odd} = -\eta_{even}$, where $\eta_{even} \equiv 1$ on $\mathcal{L}_{even}$ and $\eta_{even} \equiv -1$ on $\mathcal{L}_{odd}$. (There are no other periodic ground states, see for example [68].) The phase transition in this model has been studied in [65] and [129]. Because of the bipartite structure, we can flip all spins on a sublattice as in (6), which turns the model into an Ising ferromagnet in a staggered magnetic field of alternating sign on $\mathcal{L}_{even}$ and $\mathcal{L}_{odd}$. The latter model still satisfies the FKG inequality.

As pointed out in Section 3.2, for $h = 0$ there is a one-to-one correspondence between all Gibbs measures for the Ising ferromagnet and the Ising antiferromagnet. In particular, both models then have the same critical inverse temperature β_c. For general $|h| < 2d$, we still have two limiting Gibbs measures $\mu_\beta^{\eta_{even}}$ and $\mu_\beta^{\eta_{odd}}$, and these measures have positive correlations relative to the "staggered" ordering $\sigma \preceq \sigma'$ iff $\sigma(x)\eta_{even}(x) \leq \sigma'(x)\eta_{even}(x)$ for all $x \in \mathcal{L}$. Relative to this ordering, an analogue of the sandwiching inequality (15) holds; for more details see Section 9 of [198]. Here is a version of statement (59) for this model.

Theorem 8.11 *Consider the Ising antiferromagnet on a bipartite graph $(\mathcal{L}, \sim)$ in an external field h at any inverse temperature $\beta > 0$. If $|\mathcal{G}(\beta H)| > 1$, we have*

$$\mu_\beta^{\eta_{even}}(x \overset{\eta_{even}}{\longleftrightarrow} \infty) > 0$$

for all $x \in \mathcal{L}$.

This follows from Theorem 8.1 in the same way as Theorem 8.2. For $\mathcal{L} = \mathbf{Z}^2$, the obvious counterparts of Corollary 8.4 and Proposition 8.5 are also valid since the proofs of these results carry over to the case of a staggered external field.

The hard-core lattice gas. As we have seen in Section 3.4, this model has state space $S = \{0, 1\}$ and corresponds to setting $U(a, b) = \infty I_{\{a=b=1\}}$ and $V(a) = -a \log \lambda$ in (1), $a, b \in S$; $\lambda > 0$ is an activity parameter. The hard-core model is the limit of the Ising antiferromagnet for $\beta \to \infty$ and $h \to 2d$ along $\beta(2d - h) = \frac{1}{2} \log \lambda$, provided a configuration $\sigma \in \{-1, +1\}^{\mathcal{L}}$ is mapped to $(1-\sigma)/2 \in \{0, 1\}^{\mathcal{L}}$; see [68] for details. (The phase diagram point $h = 2d, \beta = +\infty$ of the Ising antiferromagnet is highly degenerate since there are infinitely many, in general nonperiodic, ground states.) For $\mathcal{L} = \mathbf{Z}^d$, the hard-core lattice gas can be seen as a gas of hard (i.e., non-overlapping) diamonds. In general, we still assume that $\mathcal{L}$ is bipartite. For $\lambda > 1$, the hard-core model then has two periodic ground states of chessboard type, namely η_{even} which is equal to 1 on $\mathcal{L}_{even}$ and 0 otherwise, and $\eta_{odd} = 1 - \eta_{even}$. As noticed in Section 4.4, the associated limiting Gibbs states μ_λ^{even} and μ_λ^{odd} exist. So, following the program stated in (59), we may ask whether these Gibbs measures exibit agreement percolation in the case of phase transition. The answer is again positive:

Theorem 8.12 *For the hard-core model on a bipartite graph $\mathcal{L}$ we have for any activity $\lambda > 0$: If $\mu_\lambda^{even} \neq \mu_\lambda^{odd}$ then $\mu_\lambda^{even}(x \overset{\eta_{even}}{\longleftrightarrow} \infty) > 0$ for all $x \in \mathcal{L}$, and similarly with 'odd' in place of 'even'.*

This result is completely analogous to Theorem 8.11, and was conjectured by Hu and Mak [135, 136] from computer simulations. In these papers, the authors also discuss the case of hard-core particles on a triangular lattice, the hard hexagon model. While Theorem 8.12 does apply to the hard triangle model on the hexagonal lattice (which is bipartite), the non-bipartite triangular lattice with nearest-neighbor bonds is excluded. The results of [135, 136] suggest that Theorem 8.12 still holds for the triangular lattice. A geometric proof of this conjecture would be of particular interest.

The hard-core model on the square lattice $\mathbf{Z}^2$ admits an analogue to Corollary 8.4, in that nonuniqueness of the Gibbs measure is equivalent to η_{even}-percolation for the Gibbs measure μ_λ^{even}; see [100] or [117] for more details.

The Widom–Rowlinson lattice model. Consider the set-up of Section 3.5, with equal activities $\lambda_+ = \lambda_- = \lambda > 0$ for the plus and minus particles. For $\lambda > 1$ we have two distinct periodic ground states $\eta_+ \equiv +1$ and $\eta_- \equiv -1$. From Section 4.4 we know that the associated limiting Gibbs measures $\mu_\lambda^+ = \lim_{\Lambda\uparrow\mathcal{L}} \mu_{\lambda,\Lambda}^{\eta+}$ and μ_λ^- exist. Moreover, Theorem 4.17 asserts that a phase transition occurs for some activity λ if and only if $\mu_\lambda^+(X(x) = 1) > \mu_\lambda^+(X(x) = -1)$ for some $x \in \mathcal{L}$. Now, it turns out that in this model not only hypothesis (59) holds, but that the nonuniqueness of the Gibbs measure is in fact equivalent to agreement percolation, not only for the square lattice but *for any graph.* This comes from the nature of the random-cluster representation of Section 6.7, which is related to the sites rather than the bonds of the lattice, and is a curious exception from the fact that, on the whole, the Widom–Rowlinson model is less amenable to sharp results than the Ising model. However, by the reasons discussed in Section 6.7, this result does *not* carry over to the multitype Widom–Rowlinson lattice model with $q \geq 3$ types of particles.

Theorem 8.13 *Consider the Widom–Rowlinson lattice model on an arbitrary graph* $(\mathcal{L}, \sim)$ *for any activity* $\lambda > 0$. *Then the following statements are equivalent.*

(i) *The Gibbs measure for the parameter* λ *is non-unique.*

(ii) $\mu_\lambda^+(x \overset{\eta+}{\longleftrightarrow} \infty) > 0$ *for some, and thus all* $x \in \mathcal{L}$.

Sketch of Proof: Consider $\mu_{\lambda,\Lambda}^{\eta+}$ for some finite Λ. In the same way as the random-cluster representation of Section 6.1 was modified in Section 6.2 to deal with boundary conditions, we can modify the site-random-cluster representation of Section 6.7 to obtain a coupling of $\mu_{\lambda,\Lambda}^{\eta+}$ and a wired site-random cluster distribution $\psi_{p,2,\Lambda}^1$, so that analogues of Propositions 6.22 and 6.23 hold. As a counterpart to equation (32) and by the specific nature of the site-random-cluster representation, we then find that

$$\mu_{\lambda,\Lambda}^{\eta+}(X(x) = 1) - \mu_{\lambda,\Lambda}^{\eta+}(X(x) = -1) = \psi_{p,2,\Lambda}^1(x \leftrightarrow \partial\Lambda) = \mu_{\lambda,\Lambda}^{\eta+}(x \overset{\eta+}{\longleftrightarrow} \partial\Lambda)$$

for all $x \in \mathcal{L}$. In the limit $\Lambda \uparrow \mathcal{L}$ we obtain by an analogue to (30)

$$\mu_\lambda^+(X(x) = 1) - \mu_\lambda^+(X(x) = -1) = \mu_\lambda^+(x \overset{\eta+}{\longleftrightarrow} \infty) ,$$

and the theorem follows immediately. □

To conclude this subsection, we note that hypothesis (59) also holds in other models. We mention here only the *Ashkin–Teller model* [16], a 4-state model which interpolates in an interesting way between the 4-state Potts and the so called 4-state clock model, which is also accessible to random-cluster methods; we refer to [58, 61, 192, 210].

8.5 Percolation of ground-energy bonds

So far in this section we considered a number of models which are known to show a phase transition, and asked whether this phase transition goes hand in hand with agreement percolation. These results run under the heading "phase transition implies percolation", even though for the square lattice we established results of converse type coming from the impossibility of simultaneous occupied and vacant percolation on $\mathbf{Z}^2$.

We now take the opposite point of view and ask whether "percolation implies phase transition". More precisely, we want to deduce the existence of a phase transition (at low temperatures or high densities) from a percolation result. In fact, such an idea is already implicit in Peierls' [187] and Dobrushin's [63] proof of phase transition in the Ising model, and is an integral part of the Pirogov–Sinai theory. For models with neighbor interaction as in the Hamiltonian (1), the underlying principle can be sketched as follows. At low temperatures (or high densities), each pair of adjacent spins (or particles) tries to minimize its pair interaction energy. Note that this minimization involves the bonds rather than the sites of the lattice. So, one expects that bonds of minimal energy – the *ground-energy bonds* – prevail, forming regions separated by boundaries that consist of bonds of higher energy. Such boundaries, which are known as *contours*, cost an energy proportional to their size, and are therefore typically small when β is large. This implies that the ground-energy bonds should percolate. Now, the point is that if the spins along a bond can choose between different states of minimal energy then this ambiguity can be transmitted to the macroscopic level by an infinite ground-energy cluster, and this gives rise to phase transition. In other words, the classical contour argument for the existence of phase transition can be summarized in the phrase: ground-energy bond percolation together with a (clear-cut) non-uniqueness of the local ground state implies non-uniqueness of Gibbs measures. We will now describe this picture in detail.

We consider the cubic lattice $\mathcal{L} = \mathbf{Z}^d$ of dimension $d \geq 2$ with its usual graph structure. For definiteness we consider the Hamiltonian (1) for some pair potential $U : S \times S \to \mathbf{R}$. We can and will assume that the self-potential V vanishes; this is because otherwise we can replace U by

$$U'(a,b) = U(a,b) + \frac{1}{2d}\left[V(a) + V(b)\right], \quad a,b \in S, \tag{61}$$

which, together with the self-potential $V' \equiv 0$, leads to the same Hamiltonian. Let

$$m = \min_{a,b \in S} U(a,b) \tag{62}$$

be the minimal value of U.

Given an arbitrary configuration $\sigma \in \Omega$, we will say that an edge $e = \{x,y\} \in \mathcal{B}$ is a *ground-energy bond* for σ if $U(\sigma(x),\sigma(y)) = m$. The subgraph of $\mathbf{Z}^d$ consisting of all vertices of $\mathbf{Z}^d$ and only the ground-energy bonds for σ splits then off into connected components which will be called *ground-energy clusters* for σ. We are interested in the existence of infinite ground-energy clusters, and we also need to identify specific such clusters. Unfortunately, the Burton–Keane uniqueness theorem 5.17 does not apply here because, for any Gibbs measure, the distribution of the set of ground-energy bonds fails to have the finite-energy property. We therefore resort to considering ground-energy clusters in any fixed two-dimensional layer of $\mathbf{Z}^d$; the uniqueness of planar infinite clusters can be shown in our case. (An alternative argument avoiding the use of planar layers but requiring stronger conditions on the temperature has been suggested in [87].) In fact, we have the following result.

Theorem 8.14 *Consider the Hamiltonian (1) on the lattice $\mathcal{L} = \mathbf{Z}^d$, $d \geq 2$, with neighbor interaction U and no self-potential, and let $\mathcal{P}$ be any planar layer in $\mathcal{L}$. (So $\mathcal{P} = \mathcal{L}$ for $d = 2$.) If β is large enough, there exists a Gibbs measure $\mu \in \mathcal{G}(\beta H)$ which is invariant under all automorphisms of $\mathcal{L}$ and all symmetries of U such that*

$$\mu(\exists \text{ a unique infinite ground-energy cluster in } \mathcal{P}) = 1\ .$$

In the above, a symmetry of U is a transformation τ of S such that $U(\tau a, \tau b) = U(a,b)$ for all $a, b \in S$; such a τ acts coordinatewise on configurations.

Theorem 8.14 is a particular case of a result first derived in [92] and presented in detail in Chapter 18 of [96]. We will sketch its proof below. The remarkable fact is that this type of percolation often implies that μ has a non-trivial extremal decomposition, so that there must be a phase transition. In fact, this happens whenever the set

$$G_U = \{(a,b) \in S \times S : U(a,b) = m\} \tag{63}$$

of bond ground states splits into sufficiently disjoint parts. To explain the underlying mechanism (which may be viewed as the core of the classical Peierls argument, and a rudimentary version of Pirogov–Sinai theory) we consider first the standard Ising model.

Example 8.15 *The Ising ferromagnet at zero external field.* In this model, we have as usual $S = \{-1, 1\}$, $U(a,b) = -ab$ for $a, b \in S$, $m = -1$, and $G_U = \{(-1,-1), (1,1)\}$. Hence, either all spins of a ground-energy cluster are negative, or else all these spins are positive. In other words, each ground-energy cluster is either a minus-cluster or a plus-cluster. This implies that

$$\{\exists \text{ a unique infinite ground-energy cluster in } \mathcal{P}\} \subset A_- \cup A_+ \,,$$

where A_- and A_+ are the events that there exists an infinite cluster of negative, resp. positive, spins in $\mathcal{P}$. For the Gibbs measure μ of Theorem 8.14 we thus have $\mu(A_- \cup A_+) = 1$ and, by the spin-flip symmetry of U and thus μ, $\mu(A_-) = \mu(A_+)$. Hence $\mu(A_-) > 0$ and $\mu(A_+) > 0$, so that the measures $\mu^- = \mu(\cdot\,|A_-)$ and $\mu^+ = \mu(\cdot\,|A_+)$ are well-defined. Since A_-, A_+ are tail events, it follows immediately that μ^-, μ^+ are Gibbs measures for βH. Also, $A_- \cap A_+$ is contained in the event that there are two distinct ground-energy clusters in $\mathcal{P}$, and therefore has μ-measure 0. Hence μ^- and μ^+ are mutually singular, whence $|\mathcal{G}(\beta H)| > 1$.

The same argument as in the preceding example yields the following theorem on phase transition by symmetry breaking. A detailed proof (in a slightly different setting) can be found in Section 18.2 of [96].

Theorem 8.16 *Under the conditions of Theorem 8.14 suppose that the set G_U defined by (63) admits a decomposition $G_U = G_1 \cup \ldots \cup G_N$ such that*

1. *the sets G_n, $1 \le n \le N$, have disjoint projections, i.e., if $(a,b) \in G_n$, $(a',b') \in G_{n'}$, and $n \neq n'$, then $a \neq a', b \neq b'$, and*

2. *for any two indices $n, n' \in \{1, \ldots, N\}$ we have $\bar\tau(G_n) = G_{n'}$ for some transformation $\bar\tau$ of $S \times S$ which is either the reflection, or the coordinatewise application of some symmetry of U, or a composition of both.*

Then, if β is sufficiently large, there exist N mutually singular Gibbs measures $\mu^1, \ldots \mu^N \in \mathcal{G}(\beta H)$, invariant under all even automorphisms of $\mathbf{Z}^d$ and such that

$$\mu^n(\exists \text{ an infinite } n\text{-cluster in } \mathcal{P}) = 1$$

for all $1 \le n \le N$. In particular, there exist N distinct phases for βH.

In the statement above, an infinite n-cluster for a configuration σ is an infinite cluster of the subgraph of $\mathbf{Z}^d$ obtained by keeping only those edges $e \in \mathcal{B}$ with $(\sigma(x), \sigma(y)) \in G_n$, where x is the endpoint of e in the even sublattice $\mathcal{L}_{even}$ and $y \in \mathcal{L}_{odd}$ is the other endpoint of e. Also, an even automorphism of $\mathbf{Z}^d$ is an automorphism leaving $\mathcal{L}_{even}$ invariant.

We illustrate this theorem by applying it to our other standard examples.

Example 8.17 *The Ising antiferromagnet in an external field.* We have again $S = \{-1, 1\}$, but the interaction is now $U(a, b) = ab - \frac{h}{2d}(a + b)$ for some constant $h \in \mathbf{R}$. (Here we applied the transformation (61).) If $|h| < 2d$ then $m = -1$ and $G_U = \{(-1, 1), (1, -1)\}$. G_U splits up into the singletons $G_1 = \{(1, -1)\}$ and $G_2 = \{(-1, 1)\}$. This decomposition meets the conditions of the theorem; in particular, G_1 and G_2 are related to each other by the reflection of $S \times S$. Consequently, there exist two mutually singular Gibbs measures μ^1 and μ^2 which are invariant under even automorphisms and have an infinite cluster of chessboard type, either with plus spins on the even cluster sites and minus spins at the odd cluster sites, or vice versa.

Example 8.18 *The Potts model.* In this case, $S = \{1, \ldots, q\}$ for some integer $q \geq 2$ and $U(a, b) = 1 - 2I_{\{a=b\}}$. Again $m = -1$, and $G_U = \{(n, n) : 1 \leq n \leq q\}$. Theorem 8.16 is obviously applicable, and we recover the result that for sufficiently large β there exist q mutually singular, automorphism invariant Gibbs measures, the nth of which has an infinite cluster of spins with value n.

Example 8.19 *The hard-core lattice gas.* This model has state space $S = \{0, 1\}$ and neighbor interaction U of the form $U(a, b) = \infty$ if $a = b = 1$, and $U(a, b) = -\frac{\log \lambda}{2d}(a + b)$ for all other $(a, b) \in S^2$. Here $\lambda > 0$ is an activity parameter, and we have again used the transformation (61). For $\lambda > 1$ we have $G_U = \{(0, 1), (1, 0)\}$, so that Theorem 8.16 applies. Since multiplying U with a factor β amounts to changing λ, we obtain that for sufficiently large λ there exist two distinct Gibbs measures with infinite clusters of chessboard type, just as for the Ising antiferromagnet at low temperatures.

Example 8.20 *The Widom–Rowlinson lattice model.* Here we have $S = \{-1, 0, 1\}$ and $U(a, b) = \infty$ if $ab = -1$, $U(a, b) = -\frac{\log \lambda}{2d}(|a| + |b|)$ otherwise, $a, b \in S$. If $\lambda > 1$ then $G_U = \{(-1, -1), (1, 1)\}$. Theorem 8.16 thus shows that for sufficiently large λ there exist two translation invariant Gibbs measures having infinite clusters of plus- resp. minus-particles.

Although the results in the examples above are weaker than those obtained by the random-cluster methods of Section 6 (when these apply), the ideas presented here have the advantage of providing a general picture of the geometric mechanisms that imply a phase transition, and Theorem 8.16 can quite easily be applied. Moreover, the ideas can be extended immediately to systems with arbitrary state space and suitable interactions. In this way one obtains phase transitions in anisotropic plane rotor models, classical Heisenberg ferromagnets or antiferromagnets, and related N-vector models; see Chapter 16 of [96]. One can also consider next-nearest neighbor interactions, and thus obtain various other interesting examples; for this one has to consider percolation of ground-energy plaquettes rather than ground-energy bonds, which is in fact the set-up of [96]. Last but not least, the symmetry assumption of Theorem 8.16 can often be replaced by either some direct argument, or a Peierls condition in the spirit of the Pirogov–Sinai theory; see Chapter 19 of [96]. One such extension will be used in our next example.

Example 8.21 *First-order phase transition in the Potts model.* Consider again the Potts model of Example 8.18, and suppose for simplicity that $d = 2$. Any translate in $\mathbf{Z}^2$ of the quadratic cell $\{0, 1\}^2$ is called a plaquette. For a given configuration $\sigma \in \Omega$, a plaquette P is called *ordered* if all spins in P agree, *disordered* if no two adjacent spins in P agree, and *pure* if one of these two cases occurs. If q (the number of distinct spin values) is large enough then, for arbitrary β, there exists an automorphism invariant Gibbs measure μ supported on configurations with a unique infinite cluster of pure plaquettes. This variant of Theorem 8.14 is due to Kotecký and Shlosman [148], see also Section 19.3.2 of [96]. Clearly, each cluster of pure plaquettes only contains plaquettes of the same type, either ordered or disordered. For some specific critical value $\beta_c(q)$ both possibilities must occur with positive probability; this follows from thermodynamic considerations, namely by convexity of the free energy as a function of β [148, 96]. Conditioning on each of these two cases yields two mutually singular Gibbs measures with an infinite cluster of ordered resp. disordered plaquettes. Furthermore, all spins of a cluster of ordered plaquettes must have the same value, so that by symmetry the "ordered" Gibbs measure can be decomposed further into q Gibbs measures with infinite clusters of constant spin value. As a result, for large q and $\beta = \beta_c(q)$ there exist $q + 1$ mutually singular Gibbs measures which behave qualitatively similar to the disordered phase for $\beta < \beta_c(q)$ resp. the q ordered phases for $\beta > \beta_c(q)$. This is the first-order phase transition in the Potts model for large q. For further discussions we refer to [227, 148, 150, 149] and the references therein.

We now give an outline of the proof of Theorem 8.14.

Sketch proof of Theorem 8.14: For simplicity we stick to the case $d = 2$. For any inverse temperature $\beta > 0$ and any square box $\Lambda_n = [-n, n-1]^2 \cap \mathbf{Z}^2$ we write $\mu^{per}_{\beta,n}$ for the Gibbs distribution relative to βH in the box Λ_n with periodic boundary condition. The latter means that Λ_n is viewed as a torus, so that $(i, n-1) \sim (i, -n)$ and $(n-1, i) \sim (-n, i)$ for $i \in [-n, n-1] \cap \mathbf{Z}$; the Hamiltonian H^{per}_n in Λ_n with periodic boundary condition is then defined in the natural way. Let μ^{per} be an arbitrary limit point of the sequence $(\mu^{per}_{\beta,n})_{n\geq 1}$. Evidently, μ^{per} has the symmetry properties required of μ in Theorem 8.14, and $\mu^{per} \in \mathcal{G}(\beta H)$.

To establish percolation of ground-energy bonds we fix some $\alpha < 1$ and consider the wedge $\mathcal{W} = \{x = (x_1, x_2) \in \mathcal{L} : x_1 \geq 0, |x_2| \leq \alpha x_1\}$. Let $A_{\mathcal{W}}$ be the event that there exists an infinite path of ground-energy bonds in $\mathcal{W}$ starting from the origin. We want to show that $\mu^{per}(A_{\mathcal{W}}) > 3/4$ when β is large enough. Suppose $\xi \notin A_{\mathcal{W}}$. Then there exists a contour crossing $\mathcal{W}$, i.e., a path γ in the dual lattice $\mathcal{L}^* = \mathbf{Z}^2 + (\frac{1}{2}, \frac{1}{2})$ which crosses no ground-energy bond for ξ and connects the two half-lines bordering $\mathcal{W}$. For each such path γ we will establish the contour estimate

$$\mu^{per}(\gamma \text{ is a contour}) \leq (|S|e^{-\beta\delta})^{|\gamma|}, \tag{64}$$

where $|\gamma|$ is the length (the number of vertices) of γ, and $\delta > 0$ is such that $m + 2\delta$ is the second lowest value of U.

Assuming (64) we obtain the theorem as follows. The number of paths of length k crossing $\mathcal{W}$ is at most $ck\,3^k$ for some $c < \infty$ depending on α. Hence

$$\mu^{per}(A^c_{\mathcal{W}}) \leq c \sum_{k\geq 1} k\,(3|S|e^{-\beta\delta})^k < \frac{1}{4}$$

for sufficiently large β. By the rotation invariance of μ^{per}, it follows that $\mu^{per}(A_0) > 0$, where A_0 is the intersection of $A_{\mathcal{W}}$ with its three counterparts obtained by lattice rotations. Roughly speaking, A_0 is the event that the origin belongs to two doubly infinite ground-energy paths, one being quasi-horizontal and the other quasi-vertical.

Since μ^{per} is invariant under horizontal and vertical translations, the Poincaré recurrence theorem (or the ergodic theorem) implies that the event

$$A_\infty = \{\xi \in \Omega : \theta_x \xi \in A_0 \text{ for infinitely many } x \text{ in each of the four half-axes}\}$$

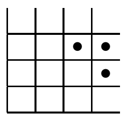
has also positive μ^{per}-probability. Each configuration in A_∞ has infinitely many quasi-horizontal and quasi-vertical ground-energy paths in each of the four directions of the compass, and by planarity all paths of different orientation must intersect. Therefore all these paths belong to the same infinite ground-energy cluster which has only finite holes, and is therefore unique. Hence A_∞ is contained in the event B that there exists a ground-energy cluster surrounding each finite set of $\mathcal{L}$, and $\mu^{per}(B) > 0$. As B is a tail-event and invariant under all automorphisms of $\mathcal{L}$ and all symmetries of U, the theorem follows by setting $\mu = \mu^{per}(\,\cdot\,|B)$.

It remains to establish the contour estimate (64). For this it is sufficient to show that

$$\mu^{per}_{\beta,n}(\gamma \text{ is a contour}) \leq (|S|e^{-\beta\delta})^{|\gamma|} \tag{65}$$

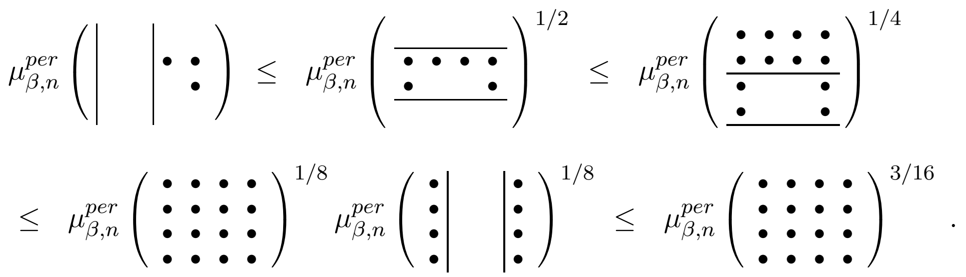
when n is so large that γ is contained in Λ_n. This bound is based on reflection positivity and the chessboard estimate, which are treated at length in Chapter 17 of [96]. Here we give only the principal ideas. The basic observation is the following consequence of the toroidal symmetry of $\mu^{per}_{\beta,n}$: for any $i \in \{0, \ldots, n-1\}$, the configurations on the two parts $\Lambda^+_{n,i} = \{x \in \Lambda_n : x_1 \geq i \text{ or } x_1 \leq i-n\}$ and $\Lambda^-_{n,i} = \{x \in \Lambda_n : i-n \leq x_1 \leq i\}$ of Λ_n are conditionally independent and, up to reflection, identically distributed given the spin values on the two separating lines $\{x_1 = i\}$ and $\{x_1 = i-n\}$. Hence, if f, g are real functions on S^{Λ_n} depending only on the configuration in $\Lambda^+_{n,i}$, and $g^{(i)}$ is the function obtained from g by reflection in these two separating lines (and thus depending on the configuration in $\Lambda^-_{n,i}$), then the bilinear form $(f,g) \to \mu^{per}_{\beta,n}(fg^{(i)})$ is positive definite and thus satisfies the Cauchy–Schwarz inequality. Similar Cauchy–Schwarz inequalities hold for vertical reflections. The chessboard inequality is obtained by suitable combinations of all these, as we will illustrate next.

Let us mark the plaquettes around the vertices of γ with a $\bullet$; this gives $|\gamma|$ marked plaquettes. Marking a plaquette indicates that at least one of its bonds has non-minimal energy; leaving it unmarked does not say that it consists of ground-energy bonds, but that we don't need any information on its spins . In the case $n = 2$, this might lead to the picture



.

To estimate its probability we use repeatedly the Cauchy–Schwarz inequality relative to suitable pairs of reflection lines. Indicating each time only the pair of lines used next, and omitting the event that no plaquette is marked (which has probability 1), we obtain



In general, we obtain in this way

$$\mu_{\beta,n}^{per}(\gamma \text{ is a contour}) \le \mu_{\beta,n}^{per}(C_n)^{|\gamma|/|\Lambda_n|} ,$$

where C_n is the event that all plaquettes in Λ_n contain at least one bond of non-minimal energy. But if C_n occurs then at least $|\Lambda_n|/2$ of the $2|\Lambda_n|$ edges in Λ_n are no ground-energy bonds. The Hamiltonian H_n^{per} with periodic boundary condition is therefore at least $(2m+\delta)|\Lambda_n|$ on C_n. Since there is at least one $\sigma \in S^{\Lambda_n}$ with $H_n^{per}(\sigma) = 2|\Lambda_n|m$, it follows that

$$\mu_{\beta,n}^{per}(C_n) \le \sum_{\xi \in C_n} e^{-\beta(2m+\delta)|\Lambda_n|} \Big/ e^{-\beta 2m|\Lambda_n|} \le (|S|\, e^{-\beta\delta})^{|\Lambda_n|} .$$

This gives estimate (65) and completes the proof of Theorem 8.14. □

9 Random interactions

So far in this review, the spin systems considered had an interaction which was invariant under all automorphisms of the underlying graph $\mathcal{L}$. Here we will assume for convenience that $\mathcal{L}$ is the cubic lattice $(\mathbf{Z}^d, \mathcal{B})$, $d \geq 2$, but the interaction between adjacent spins will no longer be translation invariant. That is, instead of the Hamiltonian (1) we now consider a modified Hamiltonian of the form

$$H(\sigma) = \sum_{b=\langle xy\rangle \in \mathcal{B}} J_b \, U(\sigma(x), \sigma(y)) + \sum_{x \in \mathbf{Z}^d} h_x \, V(\sigma(x)) \tag{66}$$

where the J_b and the h_x may vary from bond to bond, resp. from site to site. In fact, we are interested in the case where these coupling coefficients show no regular structure, and thus assume that they are *random*. Such systems of spins interacting differently depending on their position and in a way governed by chance are known as *disordered systems*. We will not elaborate on the physical origins of such random interactions. We merely mention that they can be related to the presence of impurities or defects in an originally homogeneous system, and are used to model quenched alloys of magnetic and nonmagnetic materials like $FeAu$. For details we refer to [84, 30, 78].

We assume that the family $\mathbf{J} = (J_b)_{b \in \mathcal{B}}$ of coupling coefficients and the external fields $\mathbf{h} = (h_x)_{x \in \mathbf{Z}^d}$ are independent, and each collection constitutes a family of mutually independent and identically distributed real random variables. Hence, while no realization of the coupling coefficients is translation invariant, we still have translation invariance in a statistical sense. We will not specify the underlying probability space, except that the letter P will be used to denote the probability measure and the associated expectation. The random families $\mathbf{J}$ and $\mathbf{h}$ are often referred to as the *disorder*. The disorder is called bounded if $P(|J_b| > c) = 0$ for some finite c. Physically, this is the most relevant case.

(In some physical applications it is natural to assume that the J_b are not independent but rather have some finite-range dependence structure, but we will not include this case here. We also assumed for simplicity that the disorder is real valued, although some of the following also applies to the case when J_b or h_x are allowed to take the value $+\infty$ with positive probability.)

In Section 9.1 we will discuss *diluted ferromagnets*. A bond-diluted Ising or Potts ferromagnet on $\mathbf{Z}^d$ can be viewed as an Ising or Potts model on the open clusters for Bernoulli bond percolation on $\mathbf{Z}^d$. As observed in [9], these models can quite easily be understood using their random-cluster representation. They form just about the only class of disordered systems where the phase transition can be investigated in such detail.

Then, in Section 9.2, we study the so-called *Griffiths regime* which is the only nontrivial regime for disordered systems where by now quite general results are available. It occurs at intermediate temperatures if the disorder is bounded, or at arbitrary high temperatures if the disorder is unbounded, and is characterized by the fact that the Gibbs measure is still unique but fails to have a nice high temperature behavior uniformly in the disorder. The study of random Gibbs measures in the Griffiths regime has started in the early 1980's and has reached a satisfactory stage only recently. The simplest and also most powerful methods use stochastic-geometric representations and will be presented here.

As representative for the large literature on the subject we refer to [13, 18, 29, 46, 72, 84, 85, 86, 102, 106, 189, 190]. Dynamical problems (which are not touched upon here) are treated e.g. in [49, 50, 51, 15, 103, 112, 101].

9.1 Diluted and random Ising and Potts ferromagnets

The random Potts model is defined as follows. Spins take values in the state space $S = \{1, \ldots, q\}$, and the interaction is given by the Hamiltonian (66) with

$$U(\sigma(x), \sigma(y)) = 2\, I_{\{\sigma(x) \neq \sigma(y)\}}$$

and $V \equiv 0$. Note that this choice of U and V coincides with that used in Section 6.1 for the standard Potts model and differs from that in Section 3.3 only by constants which cancel in the definition of Gibbs distributions. The Ising model corresponds to the choice $q = 2$. As for the disorder, we make the essential assumption that the random coupling coefficients J_b are nonnegative, so that the interaction is still ferromagnetic. Of course, we also make the general assumption of this section that the J_b are independent with the same distribution, say π. A particular case of special interest is that of *dilution*, in which the J_b take the values 1 and 0 with probabilities p and $1 - p$, respectively, which means that $\pi = p\delta_1 + (1 - p)\delta_0$. For $p = 1$ we then recover the homogeneous Potts model of Section 3.3.

In the following, the distribution π of the J_b will enter only through the quantities

$$\bar{p}(\beta, \pi) = P\Big(1 - e^{-2\beta J_b}\Big)\,, \quad \underline{p}(\beta, \pi) = P\Big(\frac{1 - e^{-2\beta J_b}}{1 + (q-1)e^{-2\beta J_b}}\Big)$$

for $\beta > 0$. Note that they do not depend on $b \in \mathcal{B}$, and that $0 \leq \underline{p}(\beta, \pi) \leq \bar{p}(\beta, \pi) \leq p$ with $p = P(J_b > 0)$.

For a given realization $\mathbf{J} = (J_b)_{b \in \mathcal{B}}$ of the disorder and inverse temperature β, we can introduce the Gibbs measure $\mu^i_{\beta\mathbf{J},q}$ obtained from the Gibbs distributions with constant boundary condition $i \in \{1, \ldots, q\}$ in the infinite volume limit. This limit exists by the arguments of Proposition 6.9, since these use only the stochastic monotonicity coming from Corollary 6.7 (a) as well as the random-cluster representation, which both remain valid in the non-homogeneous case.

The key quantity for phase transition, the order parameter, is the "quenched magnetization"

$$m(\beta, \pi) = \frac{q}{q-1}\, P\Big(\mu^i_{\beta\mathbf{J},q}(X(0) = i) - \frac{1}{q}\Big)\,.$$

Indeed, an inhomogeneous version of equation (32) shows that $m(\beta, \pi) = P(\theta_q(\beta\mathbf{J}))$, where $\theta_q(\beta\mathbf{J}) = \phi^1_{\mathbf{p},q}(0 \leftrightarrow \infty)$ is the percolation probability for the wired infinite-volume random-cluster measure with bond probabilities $p_b = 1 - e^{-2\beta J_b}$. Hence, we have $m(\beta, \pi) > 0$ if and only if $\theta_q(\beta\mathbf{J}) > 0$ with positive P-probability. But whether or not $\theta_q(\beta\mathbf{J}) > 0$ does not depend on the value of J_b for a single bond b. So, Kolmogorov's zero–one law implies that $m(\beta, \pi) > 0$ if and only if $\theta_q(\beta\mathbf{J}) > 0$ P-almost surely, and by an inhomogeneous version of Theorem 6.10 the latter means that multiple Gibbs measures for $\beta\mathbf{J}$ exist with P-probability 1. Moreover, an inhomogeneous version of relation (31) shows that $\theta_q(\beta\mathbf{J})$ is an increasing function of $\beta\mathbf{J}$. It follows that $m(\beta, \pi)$ is increasing in β and also in π (relative to $\preceq_{\mathcal{D}}$). In particular, for each π there exists a critical inverse temperature $\beta_c(\pi)$, possibly $= +\infty$, such that $m(\beta, \pi) > 0$ for $\beta > \beta_c(\pi)$ and $m(\beta, \pi) = 0$ for $\beta < \beta_c(\pi)$.

It remains to investigate the quenched magnetization $m(\beta, \pi)$. The following lemma shows how $m(\beta, \pi)$ can be estimated in terms of Bernoulli bond percolation; recall the end of Section 5.1.

Lemma 9.1 *Let $\theta(p) = \phi_p(0 \leftrightarrow \infty)$ be the Bernoulli bond percolation probability on $\mathbf{Z}^d$ with parameter p. Then*

$$\theta(\underline{p}(\beta,\pi)) \le m(\beta,\pi) \le \theta(\bar{p}(\beta,\pi)) \, .$$

Proof: We will use an inhomogeneous limiting version of the domination bounds (b) and (c) of Corollary 6.7. Although they were stated only in the case of a homogeneous interaction, they do extend also to the inhomogeneous case. Define two families $\mathbf{p} = (p_b)_{b\in\mathcal{B}}$ and $\mathbf{p}' = (p'_b)_{b\in\mathcal{B}}$ in terms of a realization $\mathbf{J}$ by $p_b = 1 - e^{-2\beta J_b}$, $p'_b = (1 - e^{-2\beta J_b})/(1 + (q-1)e^{-2\beta J_b}) = p_b/(p_b + q(1-p_b))$. Let $\phi^1_{\mathbf{p},q}$ be the associated wired random-cluster measure, and $\phi_{\mathbf{p}}$, $\phi_{\mathbf{p}'}$ the corresponding product measures on $\{0,1\}^{\mathcal{B}}$. An inhomogeneous version of Corollary 6.7 then shows that

$$\phi_{\mathbf{p}'}(0 \leftrightarrow \infty) \le \phi^1_{\mathbf{p},q}(0 \leftrightarrow \infty) \le \phi_{\mathbf{p}}(0 \leftrightarrow \infty) \, .$$

We now take the expectation with respect to P. In view of the preceding remarks, the middle term has P-expectation $m(\beta,\pi)$, while the P-integration of the Bernoulli measures $\phi_{\mathbf{p}'}$ and $\phi_{\mathbf{p}}$ again leads to Bernoulli measures, namely the homogeneous Bernoulli measures $\phi_{\underline{p}(\beta,\pi)}$ and $\phi_{\bar{p}(\beta,\pi)}$. The lemma follows immediately. □

Combining the lemma with the discussion before we arrive at the following result on phase transition in the random Potts model.

Theorem 9.2 *Consider the random Potts model on $\mathbf{Z}^d$ at inverse temperature $\beta > 0$ with coupling distribution π. Set $p(\pi) = \pi(]0,\infty[) = P(J_b > 0)$, and let p_c be the Bernoulli bond percolation threshold of $\mathbf{Z}^d$; so $p_c = 1/2$ when $d = 2$.*

(i) *If $\bar{p}(\beta,\pi) < p_c$ then with P-probability 1 there exists only one Gibbs measure with interaction $\beta\mathbf{J}$. In particular, this holds when $p(\pi) < p_c$ or β is small enough.*

(ii) *If $\underline{p}(\beta,\pi) > p_c$ then $m(\beta,\pi) > 0$, and with P-probability 1 there exist q distinct phases for the interaction $\beta\mathbf{J}$. In particular, this holds when $p(\pi) > p_c$ and β is large enough.*

Another way of stating this result is the following. Suppose $\pi = (1-p)\delta_0 + p\,\pi_+$ with $\pi_+ = \pi(\cdot\,|]0,\infty[)$, and let $L_+(\beta) = \int_0^\infty e^{-\beta t}\pi_+(dt)$ be the Laplace transform of π_+. (Note that then $\bar{p}(\beta,\pi) = p(1 - L_+(2\beta))$ and $\underline{p}(\beta,\pi) \ge p(1 - q\,L_+(2\beta))$.) Then there is no phase transition for $p < p_c$, whereas for $p > p_c$ the critical inverse temperature $\beta_c(p,\pi_+) \equiv \beta_c(\pi)$ is finite (and decreasing in p) and satifies the bounds

$$\frac{p - p_c}{pq} \le L_+(2\,\beta_c(p,\pi_+)) \le \frac{p - p_c}{p} \, .$$

If $\theta(p_c) = 0$ (which is known to hold for $d = 2$, and is expected to hold for all dimensions) then uniqueness holds when $p = p_c$ or $\beta = \beta_c(p,\pi_+)$. In physical terminology, the preceding bounds on $\beta_c(p,\pi_+)$ imply that the so-called crossover exponent is 1.

Example 9.3 *The case of dilution.* If J_b is 1 or 0 with probability p resp. $1-p$ then $L_+(\beta) = e^{-\beta}$. Hence, for $p > p_c$ the critical inverse temperature satisfies the logarithmic bounds

$$-\ln\frac{p - p_c}{p} \le 2\,\beta_c(p,\delta_1) \le -\ln\frac{p - p_c}{pq} \, .$$

For $q = 2$, the diluted Ising model, assertion (ii) of Theorem 9.2 gives the slightly sharper upper bound $\beta_c(p,\delta_1) \le \tanh^{-1}(p_c/p)$.

Example 9.4 *The case of power law singularities.* Suppose $\pi_+(dt) = \Gamma(a)^{-1} t^{a-1} e^{-t} dt$ is the Gamma distribution with parameter $a > 0$. Then $L_+(\beta) = (1+\beta)^{-a}$, so that for $p > p_c$ the critical inverse temperature satisfies a power law with exponent $-1/a$:

$$\left(\frac{p_c}{p - p_c}\right)^{1/a} - 1 \;\leq\; 2\,\beta_c(p, \pi_+) \;\leq\; \left(\frac{p_c q}{p - p_c}\right)^{1/a}.$$

Examples with other kinds of singularities can easily be produced [94].

Theorem 9.2 is due to [9]. Earlier, a generalized Peierls argument was used in [93] to show that for the diluted Ising model ($q = 2$) in $d = 2$ dimensions a phase transition occurs almost surely when $p > p_c = 1/2$ and β is large enough. In fact, this paper dealt mainly with the case of site dilution, in which sites rather than bonds are randomly removed from the lattice, and which in the present framework can be described by setting $J_{\langle xy \rangle} = \xi(x)\xi(y)$ for a family $(\xi(x))_{x \in \mathbf{Z}^d}$ of Bernoulli variables; the J_b are thus 1-dependent. This was continued in [94, 95] for a class of random interaction models including the random-bond Ising model as considered here, obtaining improved bounds on $\beta_c(p, \pi_+)$ for $d = 2$ as $p \downarrow 1/2$. Extensions, in particular to $d \geq 3$, were obtained in [53]. The diluted Ising model with a non-random external field $h \neq 0$ does not exhibit a phase transition; this was shown in [93] for $\mathcal{L} = \mathbf{Z}^d$ and recently extended to quite general graphs in [123].

For the diluted Ising model there is also a dynamical phase transition at the point $p = p_c$. For $p > p_c$ and $\beta_c(1, \delta_1) < \beta < \beta_c(p, \delta_1)$ the relaxation to equilibrium is no longer exponentially fast [15]. This illustrates that uniqueness of the Gibbs measure does not in itself imply the absence of a critical phenomenon. Beside such dynamical phenomena, there are also some static effects of the disorder in the uniqueness regime, albeit these are perhaps less remarkable. These are the subject of the next subsection.

9.2 Mixing properties in the Griffiths regime

As we have seen above, the diluted Ising ferromagnet shows spontaneous magnetization when $p > p_c$ and $\beta > \beta_c(p) \equiv \beta_c(p, \delta_1)$, and multiple Gibbs measures for $\beta\mathbf{J}$ exist almost surely. In the uniqueness region when still $p > p_c$ but $\beta < \beta_c(p)$ we need to distinguish between two different regimes. At high temperatures when actually $\beta < \beta_c \equiv \beta_c(1)$, the critical inverse temperature of the undiluted system, we are in the so-called paramagnetic case. This is comparable to the usual uniqueness regime for translation invariant Ising models. At intermediate temperatures, namely when $\beta_c < \beta < \beta_c(p)$, we encounter different behavior arising from the fact that the system starts to feel the disorder. This regime is called the *Griffiths regime*, since it was he [105] who discovered in this parameter region the phenomenon now called Griffiths' singularities. He studied site-diluted ferromagnets , but the arguments remain valid also in the bond-diluted case. The basic fact is the following: adding a complex magnetic field h to the Hamiltonian of the diluted Ising model we find that the partition function in a box with plus boundary conditions, as a function of h, can take values arbitrarily close to zero. The reason is that typically a large part of the box is filled by a huge cluster of interacting bonds, giving a contribution corresponding to an Ising partition function in the phase transition region. The radius of analyticity of the free energy around $h = 0$ is thus zero. In other words, the magnetization $m(\beta, p, h)$ cannot be continued analytically from $h > 0$ to $h < 0$ through $h = 0$ when $p > p_c$ and $\beta > \beta_c$. So, the presence of macrosopic clusters of strongly interacting spins (on which the spins show the low temperature behavior of the corresponding translation invariant system) gives rise to singular behavior. Related phenomena show up in a large variety of other random models, though not necessarily in the form of non-analyticity in the uniqueness regime; in general it may be difficult to pinpoint their precise nature. Nevertheless, we will speak of Griffiths' phase or Griffiths' regime whenever such singularities are expected to occur, even when a proof is still lacking. These terms then simply indicate that the usual high temperature techniques cannot be applied as such.

As another illustration we consider a random Ising model with unbounded, say Gaussian coupling variables J_b. Then βJ_b is also unbounded, even for arbitrarily small β, and with high probability a large box contains a positive fraction of strongly interacting spins. In particular, there is no paramagnetic regime, and the whole uniqueness region belongs to the Griffiths phase. For this reason, it is a non-trivial problem to show the uniqueness of the Gibbs measure. For example, the standard Dobrushin uniqueness condition encountered in (48) (cf. [65, 70]) is useless in this case; similarly, a naive use of standard cluster expansion techniques fails. In fact, these methods are bound to fail since they also imply analyticity which is probably too much to hope for (even though we cannot disprove it).

In the following we will not deal with the singular behavior in the Griffiths phase. Instead, we address the problem of showing nice behavior, which we specify here as good mixing properties of the system. We shall present two techniques: the use of random-cluster representations, and the use of disagreement percolation.

Application of random-cluster representations. Consider a random Ising model. Spins take values $\sigma(x) = \pm 1$, and the formal Hamiltonian is

$$H(\sigma) = - \sum_{b=\langle xy \rangle \in \mathcal{B}} J_b \, \sigma(x)\sigma(y) \; . \tag{67}$$

We set $\beta = 1$. Let $\mu^{\eta}_{\mathbf{J},\Lambda}$ be the associated Gibbs distribution in $\Lambda \in \mathcal{E}$ with boundary condition $\eta \in \Omega$. For many applications it is important to have good estimates on the variational distance $\|\cdot\|_\Delta$ on $\Delta \subset \Lambda$ (see (45)) of these measures with different boundary conditions.

Definition 9.5 *The random spin system above is called* **exponentially weak-mixing** *with rate* $m > 0$ *if for some* $C < \infty$ *and all* $\Lambda \in \mathcal{E}$ *and* $\Delta \subset \Lambda$

$$P\Big(\max_{\eta,\eta' \in \Omega} \|\mu^{\eta}_{\mathbf{J},\Lambda} - \mu^{\eta'}_{\mathbf{J},\Lambda}\|_\Delta \Big) \leq C|\Delta| \, e^{-m \, d(\Delta,\Lambda^c)} \; , \tag{68}$$

where $d(\Delta, \Lambda^c)$ *is the Euclidean distance of* Δ *and* Λ^c*.*

Various other mixing conditions can also be considered. A stronger condition requires that the variational distance in (68) is exponentially small in the distance between the set Δ and the region where the boundary conditions η and η' really differ. One could also restrict Λ and/or Δ to regular boxes. See [170, 217, 70, 71, 25, 49, 50, 15].

Let us comment on the significance of the exponential weak-mixing condition above.

Remarks: (1) Suppose condition (68) holds. A straightforward application of the Borel-Cantelli lemma then shows that for any $m' < m$

$$\max_{\eta,\eta' \in \Omega} \|\mu^{\eta}_{\mathbf{J},\Lambda} - \mu^{\eta'}_{\mathbf{J},\Lambda}\|_\Delta \leq C_{\mathbf{J}} \, |\Delta| \, e^{-m' \, d(\Delta,\Lambda^c)}$$

with some realization-dependent $C_{\mathbf{J}} < \infty$ P-almost surely. Integrating over η' for any Gibbs measure $\mu_{\mathbf{J}}$ we find in particular that $\mu_{\mathbf{J}} = \lim_{\Lambda \uparrow \mathbf{Z}^d} \mu^{\eta}_{\mathbf{J},\Lambda}$ for all η, implying that $\mu_{\mathbf{J}}$ is the only Gibbs measure (and depends measurably on $\mathbf{J}$). Moreover, noting that $\mu_{\mathbf{J}}(A|B) = \int \mu^{\eta}_{\mathbf{J},\Lambda}(A)\, \mu_{\mathbf{J}}(d\eta|B)$ for $A \in \mathcal{F}_{\Delta}$ and $B \in \mathcal{F}_{\Lambda^c}$, we see that this realization-dependent Gibbs measure $\mu_{\mathbf{J}}$ satisfies the exponential weak-mixing condition

$$\sup_{A \in \mathcal{F}_{\Delta},\, B \in \mathcal{F}_{\Lambda^c},\, \mu_{\mathbf{J}}(B) > 0} |\mu_{\mathbf{J}}(A|B) - \mu_{\mathbf{J}}(A)| \leq C_{\mathbf{J}}\, |\Delta|\, e^{-m'\, d(\Delta, \Lambda^c)} \, . \tag{69}$$

(2) Condition (69) above also implies an exponential decay of covariances. Let f be any local observable with dependence set $\Delta \in \mathcal{E}$ and g be any bounded observable depending only on the spins off Λ, where $\Delta \subset \Lambda$. Also, let $\delta(f) = \max_{\sigma,\sigma'} |f(\sigma) - f(\sigma')|$ be the total oscillation of f and $\delta(g)$ that of g. The covariance $\mu_{\mathbf{J}}(f;g)$ of f and g then satisfies P-almost surely the inequality

$$|\mu_{\mathbf{J}}(f;g)| \leq C_{\mathbf{J}}\, |\Delta|\, \delta(f)\, \delta(g)\, e^{-m'\, d(\Delta, \Lambda^c)}/2 \, .$$

Indeed, a short computation shows that $|\mu_{\mathbf{J},\mathbf{h}}(f;g)|$ is not larger than the left-hand side of inequality (69) times $\delta(f)\, \delta(g)/2$, cf. inequality (8.33) in [96]. If g is local, a similar inequality holds for covariances relative to finite volume Gibbs distributions in sufficiently large regions with arbitrary boundary conditions.

We will now investigate under which conditions the random Ising system with Hamiltonian (67) is exponentially weak-mixing. We start from the estimate

$$\|\mu^{\eta}_{\mathbf{J},\Lambda} - \mu^{\eta'}_{\mathbf{J},\Lambda}\|_{\Delta} \leq \phi^1_{\mathbf{p},2,\Lambda}(\Delta \leftrightarrow \partial\Lambda)$$

obtained in Theorem 7.8. As before, this bound is also valid in the inhomogeneous case considered here, and $\phi^1_{\mathbf{p},2,\Lambda}$ stands for the wired random-cluster distribution in Λ with bond-probabilities $\mathbf{p} \equiv \mathbf{p}(|\mathbf{J}|) = (p_b)_{b \in \mathcal{B}}$ given by $p_b = 1 - \exp[-2|J_b|]$. Next we can use a recent concavity result of [15]:

Lemma 9.6 *Let* $\mathbf{K} = (K_b)_{b \in \mathcal{B}}$ *be a collection of positive real numbers and* $\mathbf{p} = \mathbf{p}(\mathbf{K})$ *denote the collection of densities* $p_b = 1 - e^{-2K_b}$*. For any increasing function* f*, the expectation* $\phi^1_{\mathbf{p},2,\Lambda}(f)$ *is then a concave function of each* K_b*.*

Proof: For brevity we set $\phi^1_{\mathbf{p},2,\Lambda} = \phi$. For any fixed bond b we consider the functions $F(p_b) = \phi(f)$ and $G(K_b) = F(1 - e^{-2K_b})$. Using equation (27) of Lemma 6.6 we then find that

$$F'(p_b) = \phi(f\, g_{p_b}) - \phi(f)\, \phi(g_{p_b})$$

and

$$F''(p_b) = -2F'(p_b)\phi(g_{p_b}) \, ,$$

where for $\eta_b \in \{0,1\}$

$$g_{p_b}(\eta_b) = \frac{2\eta_b - 1}{p_b^{\eta_b}(1-p_b)^{1-\eta_b}} \, .$$

This implies

$$G''(K_b) = -4e^{-2K_b} F'(p_b)\, [1 + 2e^{-2K_b}\phi(g_{p_b})].$$

Now, F' is nonnegative because f and g_{p_b} are increasing and ϕ has positive correlations by Corollary 6.7(a). Another explicit computation shows that

$$1 + 2e^{-2K_b}\phi(g_{p_b}) = 2\phi(\eta_b = 1)/p_b - 1 \geq 2/(2 - p_b) - 1 \geq 0,$$

where the first inequality uses the fact that the random-cluster distribution for $q = 2$ dominates an independent percolation model with densities $p'_b = p_b/(2-p_b)$, see Corollary 6.7(c). It follows that $G'' \le 0$, proving the claimed concavity. □

The preceding lemma implies that

$$P\Big(\phi^1_{\mathbf{p}(|\mathbf{J}|),2,\Lambda}(\Delta \leftrightarrow \partial\Lambda)\Big) \le \phi^1_{p,2,\Lambda}(\Delta \leftrightarrow \partial\Lambda) \le \sum_{x\in\partial\Delta} \phi^1_{p,2,\Lambda}(x \leftrightarrow \partial\Lambda),$$

where $p = 1 - \exp[-2\,P(|J_b|]$. The weak mixing property (68) thus follows provided we can show an exponential decay of connectivity in the wired random-cluster distribution $\phi^1_{p,2,\Lambda}$. The latter certainly holds when $p < p_c$ because $\phi^1_{p,2,\Lambda}$ is dominated (on Λ) by the Bernoulli measure ϕ_p (cf. Corollary 6.7(b)), and the connectivity of a subcritical Bernoulli model decays exponentially fast (recall Theorem 5.6). So we arrive in particular at the following result.

Theorem 9.7 *Consider a random Ising model on $\mathbf{Z}^d$ with Hamiltonian (67). If $2\,P(|J_b|) < -\ln(1-p_c)$ for the Bernoulli bond percolation threshold p_c of $\mathbf{Z}^d$ then the system is exponentially weak-mixing.*

As the proof above shows, the factor $|\Delta|$ on the right-hand side of (68) can actually be replaced by $|\partial\Delta|$.

In the diluted Ising model at inverse temperature β (see Example 9.3), the condition of the preceding theorem reads $2\beta p < -\ln(1-p_c)$. Therefore, if d is so large that $2\beta_c\,p_c < -\ln(1-p_c)$ then the theorem covers part of the Griffiths regime. This fact is evident in the case of an unbounded, say Gaussian, disorder because then (as explained above) there is no paramagnetic phase.

Exponential weak-mixing for random Ising models can also be shown by other applications of random-cluster domination. Let us sketch such an alternative route. Using the random-cluster representation, Newman [182] showed that (pointwise in the disorder $\mathbf{J}$)

$$\max_{\eta,\eta'} \|\mu^{\eta}_{\mathbf{J},\Lambda} - \mu^{\eta'}_{\mathbf{J},\Lambda}\|_\Delta \le 2\sum_{x\in\Delta} \mu^+_{|\mathbf{J}|,\Lambda}(X(x)), \tag{70}$$

where $\mu^+_{|\mathbf{J}|,\Lambda}$ is the Gibbs distribution in Λ with plus boundary conditions for the Hamiltonian (67) with J_b replaced by $|J_b|$. On the other hand, Higuchi [132] obtained the estimate

$$\mu^+_{|\mathbf{J}|,\Lambda}(X(x)) \le \sum_{y\notin\Lambda}\ \sum_{z\in\Lambda:\, z\sim y} \mu^f_{|\mathbf{J}|,\Lambda}(X(x)X(z))$$

(the superscript 'f' referring to the free boundary condition), while Olivieri, Perez and Goulart Rosa [185] proved that

$$P\Big(\mu^f_{|\mathbf{J}|,\Lambda}(X(x)X(z))\Big) \le \mu^f_{J,\Lambda}(X(x)X(z)) \tag{71}$$

with $J = P(|J_b|)$. We are thus back to the standard Ising Gibbs distribution in Λ with zero external field, free boundary condition and coupling constant J. Now we can take advantage of the second Griffiths inequality (which we did not discuss so far in this text) stating that correlation functions such as on the right-hand side above are monotone in the coupling coefficients, see e.g. [160]. This implies that the right-hand side of (71) is an increasing function of Λ and thus bounded above by its infinite volume limit. But for

J less than J_c, the critical coupling, the Gibbs measure is unique and has an exponential decay of correlations, see [7]. We thus find that for $J < J_c$ the right-hand side of (71) has an exponential upper bound $C \exp[-m|x-z|]$ for suitable constants $C < \infty$ and $m > 0$. Together with the previous estimates, we conclude that *under the condition $P(|J_b|) < J_c$, the random Ising model on $\mathbf{Z}^d$ with Hamiltonian (67) is exponentially weak-mixing.* Again, this condition covers part of the Griffiths regime for the diluted Ising ferromagnet.

As an alternative to the use of Griffiths' inequalities above we can also apply Theorem 6.2 and Corollary 6.7(b), giving

$$P\Big(\mu^f_{|\mathbf{J}|,\Lambda}(X(x)X(z))\Big) \le \phi_p(x \leftrightarrow z)$$

where ϕ_p is the bond Bernoulli measure with density $p = P(1 - \exp[-2|J_b|])$. As mentioned above, its connectivity function decays exponentially fast when $p < p_c$. We therefore conclude that *the random Ising model on $\mathbf{Z}^d$ is also exponentially weak-mixing when $P(1 - \exp[-2|J_b|]) < p_c$.*

The above estimate (70) does not hold when we add a random magnetic field to the Hamiltonian (67),

$$H(\sigma) = -\sum_{b=\langle xy\rangle \in \mathcal{B}} J_b\,\sigma(x)\sigma(y) - \sum_{x \in \mathbf{Z}^d} h_x\,\sigma(x), \tag{72}$$

with i.i.d. real random variables h_x independent from the J_b. However, if $J_b, h_x \ge 0$ then we can take advantage of Section 2 of [132] to replace (70) with

$$\max_{\eta,\eta'} \|\mu^\eta_{\mathbf{J},\mathbf{h},\Lambda} - \mu^{\eta'}_{\mathbf{J},\mathbf{h},\Lambda}\|_\Delta \le 2 \sum_{x\in\Delta} \mu^+_{\mathbf{J},0,\Lambda}(X(x)),$$

and we can continue as above.

Application of disagreement percolation. As we have indicated at the end of Section 7.1, the idea of disagreement percolation can be applied to study the Griffiths regime for rather general random-interaction systems. To be specific we consider Ising spins with the Hamiltonian (72). We consider the finite volume Gibbs distribution $\mu^\eta_{\mathbf{J},\mathbf{h},\Lambda}$ in a box Λ with boundary condition η; as before, the subscripts $\mathbf{J}, \mathbf{h}$ describe the random interaction. We are going to apply Theorems 7.1 and 7.2 pointwise in the disorder. As in (44) we thus have to consider the variational single-spin oscillations

$$p_x^{\mathbf{J},\mathbf{h}} = \max_{\eta,\eta'\in\Omega} \|\mu^\eta_{\mathbf{J},\mathbf{h},x} - \mu^{\eta'}_{\mathbf{J},\mathbf{h},x}\|_x \tag{73}$$

which depend on the disorder $\mathbf{J}, \mathbf{h}$. Under the present assumptions, $(p_x^{\mathbf{J},\mathbf{h}})_{x\in\mathbf{Z}^d}$ is a 1-dependent random field; this is the only property of the disorder needed below. Now, from Theorems 7.1 and 7.2 we can conclude that if with P-probability 1 there is no Bernoulli site-percolation with densities (73) then, again with P-probability 1, there is a unique Gibbs measure for $\mathbf{J}, \mathbf{h}$. (In fact, similarly to the results of Section 7.2 the random system can be dominated by a random percolation system which more or less coincides with a stochastic-geometric representation of the diluted ferromagnet [101].) The following theorem is an immediate consequence of the results of Section 7.1.

Theorem 9.8 *An Ising system with random Hamiltonian (72) satisfying*

$$P(p_x^{\mathbf{J},\mathbf{h}}) < \frac{1}{(2d-1)^2}$$

is exponentially weak-mixing.

Of course, we now have to add in (68) the subscript $\mathbf{h}$ referring to the random external field.

Proof: Consider Theorem 7.1. Evidently, the right hand side of (46) can only increase if the density for an open site is set to be 1 on the odd sublattice $\mathcal{L}_{odd}$ of $\mathbf{Z}^d$. Due to the 1-dependence of the random field $\mathbf{p} = (p_x^{\mathbf{J},\mathbf{h}})_{x\in\mathbf{Z}^d}$ noticed above, the remaining $p_x^{\mathbf{J},\mathbf{h}}$ with $x \in \mathcal{L}_{even}$ are mutually independent. Taking the P-expectation in (46) we thus obtain on the right-hand side the Bernoulli percolation probability $\psi_{p,\,even}(\Delta \leftrightarrow \partial\Lambda)$, where $\psi_{p,\,even}$ is the Bernoulli measure with density $p = P(p_x^{\mathbf{J},\mathbf{h}})$ on the even sublattice $\mathcal{L}_{even}$ and density 1 on $\mathcal{L}_{odd}$. The exponential weak-mixing property now follows by an argument similar to that following (18); note that, relative to $\psi_{p,\,even}$, a path of length k is open with probability at most $p^{\lfloor k/2 \rfloor}$. □

Remarks: (1) Suppose we add a uniform magnetic field h to the random Hamiltonian (72). Under the conditions of Theorem 9.8 it is then not too difficult to show that the disorder-averaged expectation $P(\mu_{\mathbf{J},\mathbf{h}+h}(f))$ of any local function f is an infinitely differentiable function of h, see e.g. [72].

(2) Following Theorem 7.2, Theorem 9.8 requires that the single-point densities (73) are globally small enough to prevent percolation in the associated Bernoulli model. This condition can be extended to so-called constructive conditions involving finite boxes rather than single sites [25].

(3) Another very powerful and transparent treatment of the Griffiths regime (based on very similar percolation ideas) has been developed in [72]. This paper proposes a technique similar to [18] and [85] of a resummation in the high temperature cluster expansion. The bounds then allow a probabilistic interpretation of the expansion linking it with a bond percolation process.

10 Continuum models

All models considered so far lived on a lattice. Physical systems like real gases, however, are more realistically modelled by particles living in continuous space. This section is an outline of how some of the stochastic-geometric ideas developed in previous sections can be applied to a continuum setting. First, in Section 10.1, we consider the natural continuum analogues, based on Poisson processes, of the Bernoulli percolation models introduced in Section 5. Then, in Section 10.2, we consider a continuum variant of the Widom–Rowlinson model introduced in Section 3.5 and discuss its phase transition behavior. (As mentioned in Section 10.2, this continuum variant is the one originally considered by Widom and Rowlinson [223], so it predates the lattice model.)

10.1 Continuum percolation

Here we consider the basic models of continuum percolation. For a thorough treatment of the mathematical theory of continuum percolation we refer to Meester and Roy [173].

We first need to introduce the Poisson process on $\mathbf{R}^d$ and its subsets. Heuristically, a Poisson process with intensity $\lambda > 0$ on $\mathbf{R}^d$ is a random set X of points in $\mathbf{R}^d$ with the properties that

(i) for any bounded Borel set Λ of $\mathbf{R}^d$ with volume $|\Lambda|$, the number of points of X in Λ is Poisson distributed with mean $\lambda|\Lambda|$, i.e., for $n = 0, 1, 2, \ldots$ the probability of seeing exactly n points in Λ equals $\exp(-\lambda|\Lambda|)(\lambda|\Lambda|)^n/n!$;

(ii) for any two disjoint such subsets Λ_1 and Λ_2, the numbers of points observed in Λ_1 and in Λ_2 are independent.

For a construction of such a process, we first consider a bounded Borel set Λ of $\mathbf{R}^d$. Let Ω_Λ be the set of all finite subsets of Λ. A Poisson process on Λ with intensity $\lambda > 0$ is then given by a random element of Ω_Λ having distribution $\pi_{\lambda,\Lambda}$, where

$$\pi_{\lambda,\Lambda}(F) = e^{-\lambda|\Lambda|} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \int \cdots \int I_{\{\{x_1,\ldots,x_n\}\in F\}} \, dx_1 \cdots dx_n$$

for all $F \in \mathcal{F}_\Lambda$, the smallest σ-field which allows us to count the number of points in each Borel subset of Λ.

Next, let Ω be the set of all locally finite point configurations on $\mathbf{R}^d$; locally finite means that any bounded set contains only finitely many points. The Poisson process on $\mathbf{R}^d$ with intensity λ is a random point configuration X distributed according to the unique probability measure π_λ on Ω which, when projected on any bounded Borel set $\Lambda \subset \mathbf{R}^d$, yields $\pi_{\lambda,\Lambda}$. Properties (i) and (ii) above are easily verified, and make Poisson processes the natural analogues of Bernoulli measures for lattice models.

To study percolation properties of the Poisson process X, we need to introduce some notion of connectivity. A natural way is to imagine a closed Euclidean ball $B(x, R)$ of fixed radius R around each point x of the Poisson process, which leads us to considering the random subset $\bar{X} = \bigcup_{x\in X} B(x, R)$ of $\mathbf{R}^d$. Such random subsets are widely known as *Boolean models.* (More generally, $B(x, R)$ could be replaced by a closed compact random set centered at x.) This particular Boolean model is often referred to as the *Poisson blob model* or *lily pond model.* Two points $x, y \in X$ are then considered as connected to each other if they are connected in $\bar{X}$, meaning that there exists a continuous path from x to y in $\bar{X}$. By scaling, there is no loss of generality in setting $R = 1/2$, so that two balls centered at x and y intersect if and only if $|x - y| \le 1$, where $|\cdot|$ denotes Euclidean distance. The basic result on Boolean continuum percolation, analogous to Theorem 5.3 for ordinary site percolation, is the following.

Theorem 10.1 *Pick a Poisson process X on $\mathbf{R}^d$, $d \ge 2$, with intensity λ. Let $\bar{X} = \bigcup_{x \in X} B(x, 1/2)$ be the associated Boolean model, and let $\theta(\lambda)$ denote the probability that the origin belongs to an unbounded connected component of $\bar{X}$. Then there exists a critical value $\lambda_c = \lambda_c(d) \in (0, \infty)$ such that $\theta(\lambda) = 0$ if $\lambda < \lambda_c$ and $\theta(\lambda) > 0$ if $\lambda > \lambda_c$.*

The standard proof of this result is based on a partitioning of $\mathbf{R}^d$ into small cubes, reducing the problem to its lattice analogue, Theorem 5.3; see [173] or [108].

Another continuum percolation model is the so called *random connection model*, or *Poisson random edge model*, which was introduced by M. Penrose [188]. Let $g : [0, \infty) \to [0, 1]$ be a decreasing function with bounded support (that is $g(x) = 0$ when x exceeds some $R < \infty$). The random connection model with intensity λ and connectivity function g arises by taking a Poisson process X in $\mathbf{R}^d$ with intensity λ and independently drawing an edge between each pair of points x and y of X with probability $g(|x-y|)$. This setting includes the Boolean model, which corresponds to the choice $g = I_{[0,1]}$. Theorem 10.1 extends to this model: For g as above with $\int_0^\infty g(x)dx > 0$ and dimensions $d \ge 2$, there is a critical value $\lambda_c = \lambda_c(d, g)$ such that infinite connected components a.s. occur (resp. do not occur) whenever $\lambda > \lambda_c$ (resp. $\lambda < \lambda_c$) in the random connection model with intensity λ and connectivity function g.

Much of the theory of standard (lattice) percolation has analogues for these continuum models. An example is the uniqueness of the infinite cluster (Theorem 5.17), which goes through for the Boolean and random connection models. See [173] for this and much more.

10.2 The continuum Widom–Rowlinson model

The continuum Widom–Rowlinson model is a marked point process where the points are of two types: we call them plus-points and minus-points. (From now on and throughout this section, we drop the term "continuum" when refering to this model, and instead add the word "lattice" when talking about the model of Section 3.5.) For the model defined on a region $\Lambda \subseteq \mathbf{R}^d$ realizations take values in $\Omega_\Lambda \times \Omega_\Lambda$; the first coordinate describes the locations of the plus-points, and the second coordinate the minus-points. There is a hard sphere interaction preventing two points from coming within Euclidean distance R from each other; again, we set $R = 1$ without loss of generality. This interaction corresponds to the Hamiltonian

$$H(\mathbf{x}, \mathbf{y}) = \sum_{x \in \mathbf{x}, y \in \mathbf{y}} \infty \, I_{\{|x-y| \le 1\}} \, ,$$

$\mathbf{x}, \mathbf{y} \in \Omega_\Lambda$.

When Λ is bounded, the Widom–Rowlinson model on Λ with intensity λ is obtained by conditioning the Poisson product measure $\pi_{\lambda,\Lambda} \times \pi_{\lambda,\Lambda}$ on the event that there is no pair of points of opposite type within unit distance from each other. The extension to $\mathbf{R}^d$ is done in the usual DLR fashion: a probability measure μ on $\Omega \times \Omega$ is a Gibbs measure for the Widom–Rowlinson model at intensity λ if, for any bounded Borel set Λ, the conditional distribution of the point configuration on Λ given the point configuration on $\mathbf{R}^d \setminus \Lambda$ is that of two independent Poisson processes conditioned on the event that no point in Λ is placed within unit distance from a point of the opposite type, either inside or outside Λ. The resemblence with the lattice Widom–Rowlinson model of Section 3.5 is evident. We have the following analogue of Theorem 3.4.

Theorem 10.2 *For the Widom–Rowlinson model on $\mathbf{R}^d$, $d \geq 2$, with activity λ, there exist constants $0 < \lambda'_c \leq \lambda''_c < \infty$ (depending on d) such that for $\lambda < \lambda'_c$ the model has a unique Gibbs measure, while for $\lambda > \lambda''_c$ there are multiple Gibbs measures.*

The proof of this result splits naturally into two parts: first, we need to demonstrate uniqueness of Gibbs measures for λ sufficiently small, and secondly we need to show non-uniqueness for λ sufficiently large. The first half can be done by a variety of techniques. For instance, one can partition $\mathbf{R}^d$ into cubes of unit sidelength and apply disagreement percolation (Theorem 7.1). Two observations are crucial in order to make this work: that the conditional distribution of the configuration in such a cube given everything else only depends on the configurations in its neighboring cubes, and that the conditional probability of seeing no point at all in a cube tends to 1 as $\lambda \to 0$, uniformly in the neighbors' configurations.

The more difficult part, the nonuniqueness for large λ, was first obtained by Ruelle [204] using a Peierls-type argument. Here we shall sketch a modern stochastic-geometric approach using a random-cluster representation. This approach is due mainly to Chayes, Chayes and Kotecký [54] (but see also [100]), and works in showing both parts of Theorem 3.4. The so called *continuum random-cluster model* is defined as follows.

Definition 10.3 *The* **continuum random-cluster distribution** *$\phi_{\lambda,\Lambda}$ with intensity λ for the compact region $\Lambda \subset \mathbf{R}^d$ is the probability measure on Ω_Λ with density*

$$f(\mathbf{x}) = \frac{1}{Z_{\lambda,\Lambda}}\, 2^{k(\mathbf{x})}\,, \quad \mathbf{x} \in \Omega_\Lambda \tag{74}$$

with respect to the Poisson process $\pi_{\lambda,\Lambda}$ of intensity λ; here $Z_{\lambda,\Lambda}$ is a normalizing constant and $k(\mathbf{x})$ is the number of connected components of the set $\bar{\mathbf{x}} = \bigcup_{x\in\mathbf{x}} B(x, 1/2)$.

In analogy to the correspondence between the lattice Widom–Rowlinson model and its random-cluster representation in Propositions 6.22 and 6.23, we obtain the continuum random-cluster model by simply disregarding the types of the points in the Widom–Rowlinson model, with the same choice of the parameter λ. Conversely, the Widom–Rowlinson model is obtained when the connected components in the continuum random-cluster model are assigned independent types, plus or minus with probability 1/2 each. To see why this is true, note that for any $\mathbf{x} \in \Omega_\Lambda$ there are exactly $2^{k(\mathbf{x})}$ elements of $\Omega_\Lambda \times \Omega_\Lambda$ which do not contradict the hard sphere condition of the Widom–Rowlinson model and which map into $\mathbf{x}$ when we disregard the types of the points.

Besides the random-cluster representation, we can also take advantage of stochastic monotonicity properties. Let us define a partial order $\preceq$ on $\Omega \times \Omega$ by setting

$$(\mathbf{x}, \mathbf{y}) \preceq (\mathbf{x}', \mathbf{y}') \quad \text{if} \quad \mathbf{x} \subseteq \mathbf{x}' \text{ and } \mathbf{y} \supseteq \mathbf{y}'\,, \tag{75}$$

so that in other words a configuration increases with respect to $\preceq$ if plus-points are added and minus-points are deleted. A straightforward extension of Theorem 10.4 below then implies that the Gibbs distributions for the Widom–Rowlinson model have positive correlations relative to this order. The methods of Sections 4.1 and 4.3 can therefore be adapted to show that the Widom–Rowlinson model on $\mathbf{R}^d$ at intensity λ admits two particular phases μ_λ^+ and μ_λ^-, where μ_λ^+ is obtained as a weak limit of Gibbs measures on compact sets (tending to $\mathbf{R}^d$) with boundary condition consisting of a dense crowd of plus-points, and μ_λ^- is obtained similarly. We also have the sandwiching relation

$$\mu_\lambda^- \preceq_{\mathcal{D}} \mu \preceq_{\mathcal{D}} \mu_\lambda^+ \tag{76}$$

for any Gibbs measure μ for the intensity λ Widom–Rowlinson model on $\mathbf{R}^d$, so that uniqueness of Gibbs measures is equivalent to having $\mu_\lambda^- = \mu_\lambda^+$.

The Gibbs measure for the Widom–Rowlinson model on a box Λ with "plus" (or "minus") boundary condition corresponds to the wired continuum random-cluster model $\phi^1_{\lambda,\Lambda}$ on Λ where all connected components within distance 1/2 from the boundary count as a single component. Arguing as in Sections 6.3 and 6.7 we find that uniqueness of Gibbs measures for the Widom–Rowlinson model is equivalent to not having any infinite connected components in the continuum random-cluster model. Let λ_c be as in Theorem 10.1. Theorem 10.2 follows if we can show that the continuum random-cluster model $\phi^1_{\lambda,\Lambda}$ with sufficiently large intensity λ stochastically dominates $\pi_{\lambda_1,\Lambda}$ for some $\lambda_1 > \lambda_c$, whereas $\phi^1_{\lambda,\Lambda} \preceq_{\mathcal{D}} \pi_{\lambda_2,\Lambda}$ for some $\lambda_2 < \lambda_c$ when λ is sufficiently small.

To this end we need a point process analogue of Theorem 4.8, which is based on the concept of *Papangelou (conditional) intensities* for point processes. Suppose μ is a probability measure on Ω_Λ which is absolutely continuous with density $f(\mathbf{x})$ relative to the unit intensity Poisson process $\pi_{1,\Lambda}$. For $x \in \Lambda$ and a point configuration $\mathbf{x} \in \Omega_\Lambda$ not containing x, the Papangelou intensity of μ at x given $\mathbf{x}$ is, if it exists,

$$\lambda(x|\mathbf{x}) = \frac{f(\mathbf{x} \cup \{x\})}{f(\mathbf{x})} \; . \tag{77}$$

Heuristically, $\lambda(x|\mathbf{x})dx$ can be interpreted as the probability of finding a point inside an infinitesimal region dx around x, given that the point configuration outside this region is $\mathbf{x}$. Alternatively, $\lambda(\cdot\,|\,\cdot)$ can be characterized as the Radon–Nikodym density of the measure $\int \mu(d\mathbf{x}) \sum_{x\in\mathbf{x}} \delta_{(x,\mathbf{x}\setminus\{x\})}$ on $\Lambda \times \Omega_\Lambda$, the so-called reduced Campbell measure of μ, relative to the Lebesgue measure times μ [98]. It is easily checked that the Poisson process $\pi_{\lambda,\Lambda}$ has Papangelou intensity $\lambda(x|\mathbf{x}) = \lambda$.

The following point process analogue of Theorem 4.8 was proved by Preston [199] under an additional technical assumption, using a coupling of so called spatial birth-and-death processes similar to the coupling used in the proof of Theorem 4.8. Later, the full result was proved in Georgii and Küneth [98] by a discretization argument.

Theorem 10.4 *Suppose μ and $\tilde{\mu}$ are probability measures on Ω_Λ with Papangelou intensities $\lambda(\cdot\,|\,\cdot)$ and $\tilde{\lambda}(\cdot\,|\,\cdot)$ satisfying*

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

whenever $x \in \Lambda$ and $\mathbf{x}, \tilde{\mathbf{x}} \in \Omega_\Lambda$ are such that $\mathbf{x} \subseteq \tilde{\mathbf{x}}$. Then $\mu \preceq_{\mathcal{D}} \tilde{\mu}$, in that there exists a coupling $(X, \tilde{X})$ of μ and $\tilde{\mu}$ such that $X \subseteq \tilde{X}$ a.s.

Plugging (74) into (77) we find that the continuum random-cluster measure $\phi_{\lambda,\Lambda}$ has Papangelou intensity

$$\lambda(x|\mathbf{x}) = \lambda\, 2^{1-\kappa(x,\mathbf{x})} \; , \tag{78}$$

where $\kappa(x,\mathbf{x})$ is the number of connected components of $\bigcup_{y\in\mathbf{x}} B(y,1/2)$ intersecting $B(x,1/2)$. It is a simple geometric fact that there exists a constant $\kappa_{max} = \kappa_{max}(d) < \infty$ such that $\kappa(x,\mathbf{x}) \le \kappa_{max}$ for all x and $\mathbf{x}$; for $d = 2$ we may take $\kappa_{max} = 5$. It follows that

$$\lambda\, 2^{1-\kappa_{max}} \le \lambda(x|\mathbf{x}) \le 2\lambda \tag{79}$$

for all x and $\mathbf{x}$. Hence, applying Theorems 10.4 and 10.1 we find that $\phi^1_{\lambda,\Lambda} \preceq_{\mathcal{D}} \pi_{2\lambda,\Lambda}$, so that for $\lambda < \lambda_c/2$ we obtain the absence of unbounded connected components of $\bigcup_{x\in\mathbf{x}} B(x,1/2)$ in the limit $\Lambda \uparrow \mathbf{R}^d$ of continuum random-cluster measures. On the other hand, taking $\lambda > \lambda_c\, 2^{\kappa_{max}-1}$ yields the presence of unbounded connected components in the same limit. Theorem 10.2 follows immediately.

It is important to note that this approach does not allow us to show that nonuniqueness of Gibbs measures depends monotonically on λ. The reason is similar to what we saw for the lattice Widom–Rowlinson model in Section 6.7: the right hand side of (78) fails to be increasing in $\mathbf{x}$. It thus remains an open problem whether one can actually take $\lambda'_c = \lambda''_c$ in Theorem 10.2.

There are several interesting generalizations of the Widom–Rowlinson model. Let us mention one of them, in which neighboring pairs of particles of the opposite type are not forbidden, but merely discouraged. Let $h : [0,\infty) \to [0,\infty]$ be an "interspecies repulsion function" which is decreasing and has bounded support. For $\Lambda \subset \mathbf{R}^d$ compact and $\lambda > 0$, the associated Gibbs distribution $\mu_{h,\lambda,\Lambda}$ on $\Omega_\Lambda \times \Omega_\Lambda$ is given by its density

$$f(\mathbf{x},\mathbf{y}) = \frac{1}{Z_{h,\lambda,\Lambda}} \exp\Big(-\sum_{x\in\mathbf{x},y\in\mathbf{y}} h(|x-y|)\Big).$$

relative to $\pi_{\lambda,\Lambda} \times \pi_{\lambda,\Lambda}$. Infinite volume Gibbs measures on $\Omega \times \Omega$ are then defined in the usual way. Lebowitz and Lieb [154] proved nonuniqueness of Gibbs measures for large λ when $h(x)$ is large enough in a neighborhood of the origin. Georgii and Häggström [97] later established the same behavior without this condition, and for a larger class of systems, using the random-cluster approach. This involves a generalization of the continuum random-cluster model, which arises by taking the random connection model of Section 10.1 with connectivity function $g(x) = 1 - e^{-h(x)}$ and biasing it with a factor $2^{k(\mathbf{z})}$, where $k(\mathbf{z})$ is the number of connected components of a configuration $\mathbf{z}$ of points and edges. To establish the phase transition behavior of this "soft-core Widom–Rowlinson model" (i.e., uniqueness of Gibbs measures for small λ and nonuniqueness for large λ) one can basically use the same arguments as the ones sketched above for the standard Widom–Rowlinson model. However, due to the extra randomness of the edges some parts of the argument become more involved. In particular, there is no longer a deterministic bound (corresponding to κ_{max}) on how much the number of connected components can decrease when a point is added to the random-cluster configuration; thus more work is needed to obtain an analogue of the first inequality in (79).

To conclude, we note that the Widom–Rowlinson model on $\mathbf{R}^d$ has an obvious multitype analogue with $q \ge 3$ different types of particles. This multitype model still admits a random-cluster representation from which the existence of a phase transition can be derived [97]. There is, however, no partial ordering like (75) giving rise to stochastic monotonicity or an analogue of (76).

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] Adler, R.J. (1998) On excursion sets, tube formulae, and maxima of random fields, preprint, PAS 1060.

[3] Aizenman, M. (1980) Translation invariance and instability of phase coexistence in the two-dimensional Ising system, *Commun. Math. Phys.* **73**, 83–94.

[4] Aizenman, M. (1980) Instability of phase coexistence and translation invariance in two dimensions, in K. Osterwalder (ed.), *Mathematical problems in theoretical physics*, Lecture Notes in Physics 116, 143–147, Springer, Berlin, Heidelberg, New York.

[5] Aizenman, M. (1982) Geometric analysis of Φ^4 fields and Ising models. Parts I and II, *Commun. Math. Phys.* **86**, 1–48.

[6] Aizenman, M. and Barsky, D.J. (1987) Sharpness of the phase transition in percolation models, *Commun. Math. Phys.* **108**, 489–526.

[7] Aizenman, M., Barsky, D.J. and Fernandez, R. (1987) The phase transition in a general class of Ising-type models is sharp, *J. Stat. Phys.* **47**, 343–374.

[8] Aizenman, M., Bricmont, J. and Lebowitz, J.L. (1987) Percolation of the minority spins in high-dimensional Ising models, *J. Stat. Phys.* **49**, 859–865.

[9] 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.

[10] 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. Stat. Phys.* **50**, 1–40.

[11] Aizenman, M. and Fernandez, R. (1986) On the critical behavior of the magnetization in high-dimensional Ising models, *J. Stat. Phys.* **44**, 393–454.

[12] Aizenman, M., Kesten H. and Newman, C.M. (1987) Uniqueness of the infinite cluster and continuity of connectivity functions for short- and long-range percolation, *Commun. Math. Phys.* **111**, 505–532.

[13] Aizenman, N., Klein, A. and Newman, C. M. (1992) Percolation methods for disordered quantum Ising models. In: Kotecký, R. (ed.), *Phase Transitions: Mathematics, Physics, Biology, ...* 129–137, World Scientific, New Jersey, London.

[14] Alexander, K.S. and Chayes L. (1997) Non-perturbative criteria for Gibbsian uniqueness, *Commun. Math. Phys.* **189**, 447–464.

[15] Alexander, K.S., Cesi, F., Chayes, L., Maes, C. and Martinelli, F. (1998) Convergence to equilibrium of random Ising models in the Griffiths' phase, *J. Stat. Phys* **92**, 337–351.

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

[17] Averintsev, M.B. (1975) Gibbs description of random fields whose conditional probabilities may vanish *Probl. Inform. Transmission* **11**, 326–334.

[18] Bassalygo, L.A. and Dobrushin, R.L. (1986) Uniqueness of a Gibbs field with random potential – an elementary approach, *Theory Prob. Appl.* **31**, 572–589.

[19] Basuev, A.G. (1984) Complete phase diagrams with respect to external fields at low temperatures for models with nearest-neighbor interaction in the case of a finite or countable number of ground states, *Theor. Math. Phys.* **58**, 171–182.

[20] Basuev, A.G. (1987) Hamiltonian of the phase separation border and phase transitions of the first kind.II. The simplest disordered phases, *Theor. Math. Phys.* **72**, 861–871.

[21] Batty, C.J.K. and Bollmann, H.W. (1980) Generalized Holley–Preston inequalities on measure spaces and their products. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **53**, 157–174.

[22] van Beijeren, H. (1975) Interface sharpness in Ising systems, *Commun. Math. Phys.* **40**, 1–6.

[23] Benjamini, I., Lyons, R., Peres, Y. and Schramm, O. (1999) Group-invariant percolation on graphs, *Geom. Funct. Analysis* **9**, 29–66.

[24] 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.

[25] van den Berg, J. (1997) A constructive mixing condition for 2–D Gibbs measures with random interactions, *Ann. Probab.* **25**, 1316–1333.

[26] van den Berg, J. and Ermakov, A. (1996) A new lower bound for the critical probability of site percolation on the square lattice, *Random Structures Algorithms* **8**, 199–212.

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

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

[29] Berretti, A. (1985) Some properties of random Ising models, *J. Stat. Phys.* **38**, 483–496.

[30] Binder, K. and Young, A.P. (1988) Spin glasses: experimental facts, theoretical concepts, and open questions, *Rev. Mod. Phys.* **58**, 801–976.

[31] Biskup, M., Borgs, C., Chayes, J.T. and Kotecký, R. (1999) The Potts model with external fields: Gibbs states of Edwards–Sokal, random cluster and spin representations, preprint.

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

[33] Brascamp, H.J. and Lieb, E.H. (1976) On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation, *J. Funct. Anal.* **22**, 366–389.

[34] Bricmont, J and Kupiainen, A (1997) Infinite dimensional SRB measures, *Physica D* **103**, 18–33.

[35] Bricmont, J and Kupiainen, A (1996) High temperature expansions and dynamical systems, *Commun. Math. Phys.* **178**, 703–732.

[36] Bricmont, J., Lebowitz, J.L. and Maes, C. (1987) Percolation in strongly correlated systems: the massless Gaussian field, *J. Stat. Phys.* **46**, 1249–1268.

[37] Bricmont, J. and Slawny J. (1989) Phase transitions in systems with a finite number of dominant ground states, *J. Stat. Phys.* **54**, 89–161.

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

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

[40] Brydges, D., Fröhlich J. and Sokal, A.D. (1983) The random walk representation of classical spin-systems and correlation inequalities. II. The skeleton inequalities, *Commun. Math. Phys.* **91**. 117–139.

[41] Brydges, D., Fröhlich J. and Spencer, T. (1982) The random walk representation of classical spin-systems and correlation inequalities, *Commun. Math. Phys.* **83**. 123–150.

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

[43] Burton, R. and Keane, M. (1991) Topological and metric properties of infinite clusters in stationary two-dimensional site percolation, *Israel J. Math.* **76**, 299–316.

[44] Burton, R. and Steif, J. (1995) Quite weak Bernoulli with exponential rate and percolation for random fields, *Stoch. Proc. Appl.* **58**, 35–55.

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

[46] Campanino, M. and Klein, A. (1991) Decay of two–point functions for $(d+1)$–dimensional percolation, Ising and Potts model with d–dimensional disorder, *Commun. Math. Phys.* **135**, 499–515.

[47] Campanino, M. and Russo, L. (1985) An upper bound for the critical percolation probability for the three-dimensional cubic lattice, *Ann. Probab.* **13**, 478–491.

[48] Campbell, M. and Chayes, L. (1998) The isotropic O(3) model and the Wolff representation, *J. Phys. A*, **31**, L255–L259.

[49] Cesi, F., Maes, C. and Martinelli, F. (1997) Relaxation of disordered magnets in the Griffiths' regime, *Commun. Math. Phys.* **188**, 135–173.

[50] Cesi, F., Maes, C. and Martinelli, F. (1997) Relaxation to equilibrium for two dimensional disordered Ising systems in the Griffiths' phase, *Commun. Math. Phys. 189*, 323–335.

[51] Cesi, F., Maes, C. and Martinelli, F. (1997) Relaxation of random magnets : exact results for the Griffiths' regime, *Markov Proc. Rel. Fields* **3**, 465–474.

[52] Chayes, J.T. and Chayes, L. (1988) Les Houches Lectures.

[53] Chayes, J.T., Chayes, L. and Fröhlich, J. (1985) The Low-Temperature Behavior of Disordered Magnets, *Commun. Math. Phys.* **100**, 399–437.

[54] 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.

[55] Chayes, J.T., Chayes, L. and Newman, C.M. (1985) The stochastic geometry of invasion percolation, *Commun. Math. Phys.* **101**, 383–407.

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

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

[58] Chayes, L. and Machta, J. (1997) Graphical representations and cluster algorithms I. Discrete spin systems, *Physica A* **239**, 542–601.

[59] Chayes, L. and Machta, J. (1998) Graphical representations and cluster algorithms. Part II, *Physica A* **254**, 477–516.

[60] Chayes, L., Machta, J. and Redner, O. (1998) Graphical representations for Ising systems in external fields, *J. Stat. Phys.* **93**, 17–32.

[61] Chayes, L., McKellar, D. and Winn, B. (1998) Percolation and Gibbs states multiplicity for ferromagnetic Ashkin–Teller models on $\mathbf{Z}^2$, *J. Phys. A* **31**, 9055–9063.

[62] 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.

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

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

[65] 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.

[66] Dobrushin, R.L. (1972) Gibbs state describing coexistence of phases for a three-dimensional Ising model, *Th. Probab. Appl.* **17**, 582–601.

[67] Dobrushin, R.L. (1996) Estimates of semi-invariants for the Ising model at low temperatures, *Amer. Math. Soc. Transl.* **177**, 59–81.

[68] Dobrushin, R.L., Kolafa, J. and Shlosman, S.B. (1985) Phase diagram of the two–dimensional Ising antiferromagnet (computer–assisted proof), *Commun. Math. Phys.* **102**, 89–103.

[69] Dobrushin, R.L. and Shlosman S.B. (1985) The problem of translation invariance of Gibbs states at low temperatures, in *Mathematical Physics Reviews, Soviet Science Reviews Section C*, Vol. 5, S.P. Novikov, ed. (Gordon and Breach, New York).

[70] Dobrushin, R.L. and Shlosman, S.B. (1985) Constructive criterion for the uniqueness of a Gibbs field. In: Fritz, J., Jaffe, A., Szasz, D. (eds.), *Statistical mechanics and dynamical systems.* 347–370, Birkhauser, Boston.

[71] Dobrushin, R.L. and Shlosman, S.B. (1987) Completely analytical interaction: constructive description, *J. Stat. Phys.* **46**, 983–1014.

[72] von Dreifus, H., Klein, A. and Perez, J.F. (1995) Taming Griffiths' singularities: infinite differentiability of quenched correlation functions, *Comm. Math. Phys.* **170**, 21–39.

[73] 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.

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

[75] van Enter, A.C.D., Fernandez, R., Schonmann, R.H and Shlosman, S.B. (1997) On complete analyticity of the 2D Potts model, *Commun. Math. Phys.* **189**,

[76] van Enter, A.C.D., Fernandez, R. and Sokal A.D. (1993) Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory, *J. Stat. Phys.* **72**, 879–1167.

[77] Fernandez, R., Fröhlich J. and Sokal, A.D. (1992) *Random Walks, Critical phenomena, and Triviality in Quantum Field Theory*, Springer-Verlag, Berlin.

[78] Fisher, K.H. and Hertz, J.A. (1991) *Spin glasses*, Cambridge University Press, Cambridge, U.K.

[79] Flory, P.J. (1941) Molecular size distribution in three-dimensional gelation I–III, *J. Am. Chem. Soc.***63**, 3083–3100.

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

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

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

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

[84] Fröhlich, J. (1984) Mathematical aspects of the physics of disordered systems. In: Osterwalder, K., Stora, R. (eds.), *Critical phenomena, random systems and gauge theories.* 725–894, North–Holland, Amsterdam.

[85] Fröhlich, J. and Imbrie, J. (1984) Improved perturbation expansion for disordered systems: beating Griffiths' singularities, *Commun. Math. Phys.* **96**, 145–180.

[86] Fröhlich, J. and Zegarlinski, B. (1987) The high temperature phase of a long range spin glass, *Commun. Math. Phys.* **110**, 121–155.

[87] Fukuyama, M. (1998) Discrete symmetry breaking for certain short-range interactions, preprint, Graduate School of Mathematical Science, University of Tokyo.

[88] Gallavotti, G. and Miracle-Sole, S. (1972) Equilibrium states of the Ising model in the two-phase region. *Phys. Rev. B* **5**, 2555–2559.

[89] Gandolfi, A., Grimmett, G. R., and Russo, L. (1988) On the uniqueness of the infinite open cluster in the percolation model, *Commun. Math. Phys.* **114**, 549–552.

[90] Gandolfi, A., Keane, M. and Russo, L. (1988) On the uniqueness of the infinite occupied cluster in dependent two-dimensional site percolation, *Ann. Probab.* **16**, 1147–1157.

[91] Georgii, H.-O. (1973) Two remarks on extremal equilibrium states, *Comm. Math Phys.* **32**, 107–118.

[92] Georgii, H.-O. (1981) Percolation for low energy clusters and discrete symmetry breaking in classical spin systems, *Commun. Math. Phys.* **81**, 455–473.

[93] Georgii, H.-O. (1981) Spontaneous magnetization of randomly dilute ferromagnets, *J. Stat. Phys.* **25**, 369–396.

[94] Georgii, H.-O. (1984) On the ferromagnetic and the percolative region of random spin systems, *Adv. Appl. Prob.* **16**, 732–765.

[95] Georgii, H.-O. (1985) On the ferromagnetic and percolative region of random spin systems, *Contemp. Math.* **41**, 135–147.

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

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

[98] Georgii, H.-O. and Küneth, T. (1997) Stochastic comparison of point random fields, *J. Appl. Probab.* **34**, 868–881.

[99] Gertsik, V.M. (1976) Conditions for the non-uniqueness of the Gibbs state for lattice models having finite interaction potentials, *Math. USSR Izvestija* **10**, 429–443.

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

[101] Gielis, G. (1998) Spin systems with random interactions: the Griffiths' regime, Ph.D. thesis, K.U. Leuven.

[102] Gielis, G. and Maes, C. (1995) The uniqueness regime of Gibbs fields with unbounded disorder, *J. Stat. Phys.* **81**, 829–835.

[103] Gielis, G., Maes, C. and Vande Velde, K. (1999) Randomly interacting particle systems: the uniqueness regime, *Ann. Inst. H. Poincaré, Phys. Theor.*, to appear (May 1999).

[104] Gordon, D.M. (1991) Percolation in high dimensions, *J. London Math. Soc. (2)* **44**, 373–384.

[105] Griffiths, R. (1969) Non–analytic behaviour above the critical point in a random Ising ferromagnet, *Phys. Rev. Lett.* **23**, 17–19.

[106] Griffiths, R. and Lebowitz, J.L. (1972) Random spin systems: rigorous results, *J. Math. Phys.* **9**, 1284–1292.

[107] Grimmett, G.R. (1973) A theorem about random fields, *Bull London Math. Soc.* **5**, 81–84.

[108] Grimmett, G.R. (1999) *Percolation*, 2. edition, Springer, New York.

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

[110] Grimmett, G.R. (1997) Percolation and disordered systems, *Lectures on Probability Theory and Statistics (Saint-Flour 1996)*, 153–300, *Lecture Notes in Math.* **1665**, Springer, Berlin.

[111] Grimmett, G.R. and Newman, C.M. (1990) Percolation in $\infty + 1$ dimensions, *Disorder in Physical Systems* (G. R. Grimmett and D. J. A. Welsh, editors), 219–240, Clarendon Press, Oxford.

[112] Guionnet, A. and Zegarlinski, B. (1997) Decay to equilibrium in random spin systems on a lattice II, *J. Stat. Phys.* **86**, 899–904.

[113] Häggström, O. (1995) Random-cluster measures and uniform spanning trees, *Stoch. Proc. Appl.* **59**, 267–275.

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

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

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

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

[118] Häggström, O. (1997) Infinite clusters in dependent automorphism invariant percolation on trees, *Ann. Probab.* **25**, 1423–1436.

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

[120] Häggström, O. and Jonasson, J. (1998) Phase transition in the random triangle model, *J. Appl. Probab.*, to appear.

[121] Häggström, O., van Lieshout, M.N.M. and Møller, J. (1997) Characterisation results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes, *Bernoulli*, to appear.

[122] Häggström, O. and Peres, Y. (1999) Monotonicity of uniqueness for percolation on Cayley graphs: all infinite clusters are born simultaneously, *Probab. Th. Rel. Fields* **113**, 273–285.

[123] Häggström, O., Schonmann, R.H. and Steif, J.E. (1999) The Ising model on diluted graphs and strong amenability, preprint.

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

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

[126] Hara, T. and Slade, G. (1994) Mean field behavior and the lace expansion, in *Probability and Phase Transition* (ed. G. Grimmett), 87–122, Proceedings of the NATO ASI meeting in Cambridge 1993, Kluwer.

[127] Hara, T. and Slade, G. (1995) The self-avoiding walk and percolation critical points in high dimensions, *Combin. Probab. Comput.* **4**, 197–215.

[128] Harris, T.E. (1960) A lower bound on the critical probability in a certain percolation process, *Proc. Cambridge Phil. Soc.***56**, 13–20.

[129] Heilmann, O.J. (1974) The use of reflection as symmetry operation in connection with Peierls' argument, *Comm. Math. Phys.* **36**, 91–114.

[130] Higuchi, Y. (1981) On the absence of non-translation invariant Gibbs states for the two-dimensional Ising model, in : J. Fritz, J.L. Lebowitz and D. Szász (eds.), *Random fields*, Esztergom (Hungary) 1979. Amsterdam: North-Holland, Vol. I, 517–534.

[131] Higuchi, Y. (1985) A weak version of RSW theorem for the two-dimensional Ising model, in: R. Durrett (ed.), Particle systems, random media, and large deviations. *Contemporary Math.* **41**, Amer. Math. Soc., 207–214.

[132] Higuchi, Y. (1993) Coexistence of infinite (*)–clusters II:–Ising percolation in two dimension, *Prob. Th. Rel. Fields* **97**, 1–33.

[133] Higuchi, Y. (1993) A sharp transition for the two-dimensional Ising percolation, *Prob. Th. Rel. Fields* **97**, 489–514.

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

[135] Hu, C.-K. and Mak, K.-S. (1989) Percolation and phase transitions of hard–core particles on lattices : Monte Carlo approach, *Phys. Rev. B* **39**, 2948–2951.

[136] Hu, C.-K. and Mak, K.-S. (1990) Percolation and phase transitions of hard–core particles on lattices with pair interactions, *Phys. Rev. B* **42**, 965–968.

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

[138] Israel, R.B. (1979) Convexity in the theory of lattice gases, Princeton University Press, Princeton.

[139] Jonasson, J. (1999) The random-cluster model on a general graph and a phase transition characterization of nonamenability, *Stoch. Proc. Appl.* **79**, 335–354.

[140] Jonasson, J. and Steif, J. (1997) Amenability and phase transition in the Ising model, *J. Theor. Probab.*, to appear.

[141] Kamae, T., Krengel, U. and O'Brien, G.L. (1977) Stochastic inequalities on partially ordered spaces, *Ann. Probab.* **5**, 899–912.

[142] Kesten, H. (1980) The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$, *Commun. Math. Phys.* **74**, 41–59.

[143] Kesten, H. (1982) *Percolation theory for mathematicians.* Boston MA, Birkhauser.

[144] Kesten, H. (1990) Asymptotics in high dimensions for percolation, *Disorder in physical systems* (eds G. Grimmett and D. Welsh), 219–240, Oxford University Press.

[145] Klain, D.A. and Rota, G.-C. (1997) *Introduction to geometric probability*, Cambridge University Press, Cambridge.

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

[147] Kotecký, R. (1995) Phase transitions of lattice models. Université de Rennes I lectures.

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

[149] 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.

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

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

[152] Lebowitz, J.L. (1972) Bounds on the correlations and analyticity properties of Ising spin systems, *Commun. Math. Phys.* **28**, 313–321.

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

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

[155] Lebowitz, J.L. and Martin-Löf, A. (1972) On the uniqueness of the equilibrium state for Ising spin systems, *Commun. Math. Phys.* **25**, 276–282.

[156] Lebowitz, J.L. and Penrose, O. (1977) Cluster and percolation inequalities for lattice systems with interactions, *J. Stat. Phys.* **16**, 321–337.

[157] Lebowitz, J.L. and Saleur, H. (1986) Percolation in strongly correlated systems, *Physica* **138**A, 194–205.

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

[159] Li, X.-J. and Sokal, A. (1989) Rigorous lower bounds on the dynamical critical exponents of the Swendsen–Wang algorithm, *Phys. Rev. Lett.* **63**, 827–830.

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

[161] Liggett, T.M., Schonmann, R.H. and Stacey, A.M. (1997) Domination by product measures, *Ann. Probab.* **25**, 71–96.

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

[163] Lörinczi, J., Maes, C. and Vande Velde K. (1998) Transformations of Gibbs measures, *Prob. Th. Rel. Fields* **112**, 121–147.

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

[165] 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.

[166] Maes, C. (1993) Coupling interacting particle systems, *Rev. Math. Phys.* **5**, 457–475.

[167] Maes, C. and K. Vande Velde (1997) Relative energies for non-Gibbsian states, *Commun. Math. Phys.* **189**, 277–286.

[168] Martinelli, F. (1997) Lectures on Glauber dynamics for discrete spin models. Saint-Flour Lecture Notes.

[169] Martinelli, F. and Olivieri, E. (1994) Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case. *Commun. Math. Phys.* **161**, 447–486.

[170] Martinelli, F. and Olivieri, E. (1994) Approach to equilibrium of Glauber dynamics in the one phase region II: The general case, *Commun. Math. Phys.* **161**, 487.

[171] McCoy, B.M. and Wu, T.T. (1973) *The two-dimensional Ising model*, Harvard Univ. Press, Cambridge, Mass.

[172] Mecke, K.R. (1998) Integral geometry and statistical physics, preprint.

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

[174] Menshikov, M.V. (1986) Coincidence of critical points in percolation models, *Soviet Mathematics Doklady* **33**, 856–859.

[175] Menshikov, M.V., Molchanov, S.A. and Sidovenko, S.A. (1986) Percolation theory and some applications, *Itogi Nauki i Techniki* (Series of Probability theory, Mathematical statistics and Theoretical cybernetics) **24**, 53–110.

[176] Menshikov, M.V. and Pelikh, K.D. (1989) Percolation with several defect types, an estimate of critical probability for a square lattice, *Math. Notes Acad. Sc. USSR* **46**, 778–785.

[177] Messager, A. and Miracle-Sole, S. (1975) Equilibrium states of the two-dimensional Ising model in the two-phase region, *Commun. Math. Phys.* **40**, 187–196.

[178] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) Equations of state calculations by fast computing machine, *J. Chem. Phys.* **21**, 1087–1091.

[179] Minlos R.A. (1967) Gibbs' limit distribution, *Funct. Anal. Appl.* **1**, 60–73.

[180] S. Miracle Sole (1994) On the microscopic theory of phase transitions, in *Proc. XIII Sitges Conf.* "25 years of statistical mechanics."

[181] Molchanov, S.A. and Stepanov, A.K. (1983) Percolation in random fields. I, *Teoret. Mat. Fiz.* **55**, 246–256 & 419–430.

[182] Newman, C.M. (1994) Disordered Ising systems and random cluster representations, *Probability and Phase Transition* (ed. G. Grimmett), NATO ASI series vol. 120, Kluwer, Dordrecht, 247–260.

[183] Newman, C.M. and Schulman, L.S. (1981) Infinite clusters in percolation models, *J. Stat. Phys.* **26**, 613–628.

[184] Nielaba, P. and Lebowitz, J.L. (1997) Phase transitions in the multicomponent Widom–Rowlinson model and in hard cubes on the BCC-lattice, preprint.

[185] Olivieri, E., Perez, F. and Goulart Rosa Jr., S. (1983) Some rigorous results on the phase diagram of the dilute Ising model, *Phys. Lett.* **94A**, 309.

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

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

[188] Penrose, M.D. (1991): On a continuum percolation model, *Adv. Appl. Probab.* **23**, 536–556.

[189] Perez, J.F. (1993) Controlling the effect of Griffiths' singularities in random ferromagnets, *Braz. J. Phys.* **23**, 356-362.

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

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

[192] Pfister, C.-E. and Velenik, Y. (1997) Random-cluster representation of the Ashkin–Teller model, *J. Stat. Phys.* **88**, 1295–1331.

[193] Pirogov, S.A. and Sinai, Y.G. (1976) Phase diagrams for classical lattice systems, *Theor. Math. Phys.* **25**, 1185–1192

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

[195] Preston, C.J. (1974) *Gibbs states on countable sets*, Cambridge University Press, Cambridge.

[196] Preston, C.J. (1974) An application of the GHS inequality to show the absence of phase transitions for Ising spin systems, *Commun. Math. Phys.* **35**, 253–255.

[197] Preston, C.J. (1974) A generalization of the FKG inequalities, *Commun. Math. Phys.* **36**, 233–242.

[198] Preston, C.J. (1976) *Random Fields*, Lect. Notes in Math. **534**, Springer–Verlag, Berlin.

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

[200] 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.

[201] Radulescu, D.C. (1997) *A computer-assisted proof of uniqueness of phase for the hard-square lattice gas model in two dimensions*, Rutgers University Ph.D. thesis (New Brunswick, New Jersey).

[202] Ruelle, D. (1969) *Statistical Mechanics. Rigorous results*, Benjamin, New York.

[203] Ruelle, D. (1971) An extension of the Lee–Yang circle theorem, *Phys. Rev. Lett.* **26**, 303–304.

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

[205] Ruelle, D. (1972) On the use of "small external fields" in the problem of symmetry breakdown in statistical mechanics, *Ann. Phys.* **69**, 364–374.

[206] Ruelle, D. (1978) *Thermodynamic Formalism*, Addison-Wesley, Reading.

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

[208] Russo, L. (1979) The infinite cluster method in the two-dimensional Ising model, *Commun. Math. Phys.* **67**, 251–266.

[209] Russo, L. (1981) On the critical percolation probabilities, *Z. Wahrsch. Verw. Gebiete* **56**, 229–237.

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

[211] Seppäläinen, T. (1998) Entropy for translation-invariant random-cluster measures, *Ann. Probab.* **26**, 1139–1178.

[212] Simon, B. (1993) *The Statistical Mechanics of Lattice Gases, Volume 1*, Princeton University Press, Princeton.

[213] Sinai, Y.G. (1982) *Theory of Phase Transitions: Rigorous Results*, Pergamon Press, Oxford.

[214] Slawny, J. (1986) Low temperature properties of classical lattice systems: phase transitions and phase diagrams, *Phase Transitions and Critical Phemomena* **5** (C. Domb and J.L. Lebowitz, eds.), Academic Press, London and New York.

[215] Spitzer, F. (1975) Markov random fields on an infinite tree, *Ann. Probab.* **3**, 387–398.

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

[217] Stroock, D.W. and Zegarlinski, B. (1992) The logarithmic Sobolev inequality for discrete systems on a lattice, *Commun. Math. Phys.* **149**, 175–193.

[218] Sweeny, M. (1983) Monte Carlo study of weighted percolation clusters relevant to the Potts model, *Phys. Rev.* **27**, 4445–4455.

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

[220] Thompson, C.J. (1972) *Mathematical Statistical Mechanics*, McMillan, New York, reprinted by Princeton Univ. Press.

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

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

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

[224] Wierman, J.C. (1995) Substitution method critical probability bounds for the square lattice site percolation model, *Combin. Probab. Comput.* **4**, 181–188.

[225] Wilson, D. (1998) Annotated bibliography of perfectly random sampling with Markov chains, *Microsurveys in Discrete Probability* (D. Aldous and J. Propp, eds), pp 209–220, American Mathematical Society. See also
`http://dimacs.rutgers.edu/~dbwilson/exact/` .

[226] Wolff, U. (1989) Collective Monte Carlo updating for spin systems, *Phys. Rev. Lett.* **62**, 361–364.

[227] Wu, F.Y. (1982) The Potts model, *Rev. Mod. Phys.* **54**, 235– 268.

[228] Zahradník, M. (1984) An alternate version of Pirogov–Sinai theory, *Commun. Math. Phys.* **93**, 559–581.

[229] Zahradník, M. (1987) Analyticity of low-temperature phase diagrams of lattice spin models, *J. Stat. Phys.* **47**, 725–755.

[230] Zahradník, M. (1996) A short course on the Pirogov–Sinai theory, Rome II lectures.