Percolation and number of phases in the 2D Ising model

Hans-Otto Georgii
Mathematisches Institut der Universität München
Theresienstr. 39, D-80333 München, Germany.

Yasunari Higuchi
Department of Mathematics, Faculty of Science
Kobe University, Rokko, Kobe 657-8501, Japan.

We reconsider the percolation approach of Russo, Aizenman and Higuchi for showing that there exist only two phases in the Ising model on the square lattice. We give a fairly short alternative proof which is only based on FKG monotonicity and avoids the use of GKS-type inequalities originally needed for some background results. Our proof extends to the Ising model on other planar lattices such as the triangular and honeycomb lattice. We can also treat the Ising antiferromagnet in an external field and the hard-core lattice gas model on $\mathbf{Z}^2$.

1 Introduction

One of the fundamental results on the two-dimensional ferromagnetic Ising model is the following theorem obtained independently in the late 1970s by Aizenman [1] and Higuchi [8] on the basis of the seminal work of Russo [11].

Theorem. *For the ferromagnetic Ising model on $\mathbf{Z}^2$ with no external field and inverse temperature $\beta > \beta_c$, there exist precisely two extremal Gibbs measures μ^+ and μ^-.*

The basic technique initiated by Russo consists of an interplay of three features of the Ising model:

- the strong Markov property for random sets defined by geometric conditions involving clusters of constant spin,
- the symmetry of the interaction under spin-flip and lattice automorphisms, and
- the ferromagnetic character of the interaction which manifests itself in FKG order and positive correlations.

These ingredients led to a detailed understanding of the geometric features of typical configurations as described by the concepts of percolation theory. In addition to these tools, the authors of [1, 8, 11] also needed the result that the limiting Gibbs measure with $\pm$ boundary condition is a mixture of the two pure phases. This result was obtained by quite different means, namely some symmetry correlation inequalities of GKS and Lebowitz type [10]. While these symmetry inequalities are a beautiful and powerful tool, they are quite different in character from the FKG inequality and have their own restrictions. It is therefore natural to ask whether Russo's random cluster method is flexible enough to prove the theorem without recourse to symmetry inequalities. On the one hand, this would allow to extend the theorem to models with less symmetries, while on the other hand one might gain a deeper understanding of possible geometric features of typical configurations.

In this paper we propose such a purely geometric reasoning which is only based on the three features above and avoids the use of the symmetry inequalities of Messager and Miracle-Sole [10]. Despite this reduction of tools we could simplify the proof by an efficient combination of known geometric arguments. These include

- the Burton-Keane uniqueness theorem for infinite clusters [2],
- a version of Zhang's argument for the impossibility of simultaneous plus- and minus-percolation in $\mathbf{Z}^2$ (cf. Theorem 5.18 of [6]),
- Russo's symmetry trick for simultaneous flipping of spins and reflection of the lattice [11], and
- Aizenman's idea of looking at contour intersections in a duplicated system [1].

We have tried to keep the paper reasonably self-contained, so that the reader will find a complete proof of the theorem. As a payoff of the method we also obtain some generalizations. On the one hand, the arguments carry over to the Ising model on other planar lattices such as the triangular or the hexagonal lattice. On the other hand, in the case of the square lattice they cover also the antiferromagnetic Ising model in an external field as well as the hard-core lattice gas model.

2 Set-up and basic facts

Although we assume that the reader is familiar with the definition of the Ising model, let us start recalling a number of fundamental facts and introducing some notations. We assume throughout that the inverse temperature β exceeds the Onsager threshold β_c, and that there is no external field, $h = 0$. The main ingredients we need are the following:

• the *configuration space* $\Omega = \{-1, 1\}^{\mathbf{Z}^2}$, which is equipped with the Borel σ-algebra $\mathcal{F}$ and the local σ-algebras $\mathcal{F}_\Lambda$ of events depending only on the spins in $\Lambda \subset \mathbf{Z}^2$.

• the *Gibbs distributions* μ_Λ^ω in finite regions $\Lambda \subset \mathbf{Z}^2$ with boundary condition $\omega \in \Omega$; these enjoy the *Markov property* which says that $\mu_\Lambda^\omega(A)$ for $A \in \mathcal{F}_\Lambda$ depends only on the restriction of ω to the boundary $\partial\Lambda = \{x \notin \Lambda : |x - y| = 1 \text{ for some } y \in \Lambda\}$ of Λ, and the *finite-energy property*, which states that $\mu_\Lambda^\omega(A) > 0$.

• the *Gibbs measures* μ on $(\Omega,\mathcal{F})$ which, by definition, satisfy $\mu(\cdot\,|\mathcal{F}_{\Lambda^c})(\omega)=\mu_\Lambda^\omega$ for μ-almost all ω and any finite Λ; we write $\mathcal{G}$ for the set of all Gibbs measures and $\operatorname{ex}\mathcal{G}$ for the set of all extremal Gibbs measures.

• the *strong Markov property* of Gibbs measures, stating that $\mu(\cdot\,|\mathcal{F}_{\Gamma^c})(\omega)=\mu_{\Gamma(\omega)}^\omega$ for μ-almost all ω whenever Γ is a *random* finite subset of $\mathbf{Z}^2$ satisfying $\{\Gamma=\Lambda\}\in\mathcal{F}_{\Lambda^c}$ for all finite Λ; $\mathcal{F}_{\Gamma^c}$ is the set of all events A satisfying $A\cap\{\Gamma=\Lambda\}\in\mathcal{F}_{\Lambda^c}$ for all finite Λ.

• the *stochastic monotonicity* (or FKG order) of Gibbs distributions; writing $\mu\preceq\nu$ when $\mu(f)\le\nu(f)$ for all increasing local (or, equivalently, all increasing bounded measurable) real functions f on Ω, we have $\mu_\Lambda^\omega\preceq\mu_\Lambda^{\omega'}$ when $\omega\le\omega'$, and $\mu_\Lambda^\omega\preceq\mu_\Delta^\omega$ when $\Delta\subset\Lambda$ and $\omega\equiv+1$ on $\Lambda\setminus\Delta$ (the opposite relation holds when $\omega\equiv-1$ on $\Lambda\setminus\Delta$).

• the *pure phases* μ^+, $\mu^-\in\mathcal{G}$ obtained as limits for $\Lambda\uparrow\mathbf{Z}^2$ of μ_Λ^ω with $\omega\equiv+1$ resp. -1, their invariance under all graph automorphisms of $\mathbf{Z}^2$, the sandwich relation $\mu^-\preceq\mu\preceq\mu^+$ for any other $\mu\in\mathcal{G}$, and the resulting extremality of μ^+ and μ^-.

• the characterization of extremal Gibbs measures by their *triviality on the tail σ-algebra* $\mathcal{T}=\bigcap\{\mathcal{F}_{\Lambda^c}:\Lambda\subset\mathbf{Z}^2\text{ finite}\}$; the fact that extremal Gibbs measures have *positive correlations*; and the *extremal decomposition* representing any Gibbs measure as the barycenter of a mass distribution on $\operatorname{ex}\mathcal{G}$.

A general account of Gibbs measures can be found in [5], and [6] contains an exposition of the Ising model and its properties related to stochastic monotonicity.

We will also use a class of transformations of Ω which preserve the Ising Hamiltonian, and thereby the class $\mathcal{G}$ of Gibbs measures. These transformations are

• the *spin-flip transformation* $T:\omega=(\omega(x))_{x\in\mathbf{Z}^2}\to(-\omega(x))_{x\in\mathbf{Z}^2}$;

• the *translations* ϑ_x, $x\in\mathbf{Z}^2$, which are defined by $\vartheta_x\omega(y)=\omega(y-x)$ for $y\in\mathbf{Z}^2$, and in particular the horizontal and vertical shifts $\vartheta_{\text{hor}}=\vartheta_{(1,0)}$ resp. $\vartheta_{\text{vert}}=\vartheta_{(0,1)}$; and

• the *reflections* in lines ℓ through lattice sites: for any $k\in\mathbf{Z}$ we write

$$R_{k,\text{hor}}:\mathbf{Z}^2\ni x=(x_1,x_2)\to(x_1,2k-x_2)\in\mathbf{Z}^2$$

for the reflection in the horizontal line $\{x_2=k\}$, and similarly $R_{k,\text{vert}}$ for the reflection in the vertical line $\{x_1=k\}$. For $k=0$ we simply write $R_{\text{hor}}=R_{0,\text{hor}}$ and $R_{\text{vert}}=R_{0,\text{vert}}$. All these reflections act canonically on Ω.

We will investigate the geometric behavior of typical configurations in *half-planes* of $\mathbf{Z}^2$. These are sets of the form

$$\pi=\{x=(x_1,x_2)\in\mathbf{Z}^2:\ x_i\ge k\}$$

with $k\in\mathbf{Z}$, $i\in\{1,2\}$, or with '$\ge$' replaced by '$\le$'. The line $\ell=\{x\in\mathbf{Z}^2:\ x_i=k\}$ is called the associated *boundary line.* In particular, we will consider

• the upper half-plane $\pi_{\text{up}}=\{x=(x_1,x_2)\in\mathbf{Z}^2:\ x_2\ge 0\}$,

• the downwards half-plane $\pi_{\text{down}}=\{x=(x_1,x_2)\in\mathbf{Z}^2:\ x_2\le 0\}$,

and the analogously defined right half-plane π_{right} and left half-plane π_{left}. We will also work with

- the left horizontal semiaxis $\ell_{\text{left}} = \{x = (x_1, x_2) \in \mathbf{Z}^2 : \; x_1 \le 0, \, x_2 = 0\}$, and
- the right semiaxis $\ell_{\text{right}} = \{x = (x_1, x_2) \in \mathbf{Z}^2 : \; x_1 \ge 0, \, x_2 = 0\}$.

In the rest of this section we state three fundamental results on percolation in the Ising model. By the symmetry between the spin values $+1$ and -1, these results also hold when '$-$' and '$+$' are interchanged. Similarly, all notations introduced with one sign will be used accordingly for the opposite sign.

We assume that the reader is familiar with the basic concepts of percolation theory such as paths and $*$paths, $+$paths and $+*$paths (consisting of plus spins), circuits and $*$circuits, semicircuits in half-planes, clusters, $+$ and $+*$clusters, and so on. These can be found in the paper of Russo [11].

The starting point is the following result of [3, 11]. Let E^+ denote the event that there exists an infinite $+$cluster in $\mathbf{Z}^2$, and define E^-, E^{+*}, E^{-*} analogously. (Throughout this paper we will use the letter E to denote events concerning existence of infinite clusters.)

Lemma 2.1 (Existence of infinite clusters) *If $\mu \in \mathcal{G}$ is different from μ^-, there exists with positive probability an infinite $+$cluster. That is, $\mu(E^+) > 0$ when $\mu \neq \mu^-$.*

Sketch Proof. Suppose that $\mu(E^+) = 0$. Then any given square Δ is almost surely surrounded by a $-*$circuit, and with probability close to 1 such a circuit can already be found within a square $\Lambda \supset \Delta$ provided Λ is large enough. Define a random set Γ as the largest subset of Λ which is the interior of such a $-*$circuit. (A largest such set exists because the union of such sets is again the interior of a $-*$circuit.) By maximality, Γ satisfies the conditions of the strong Markov property. This together with the stochastic monotonicity $\mu_\Gamma^- \preceq \mu^-$ implies (in the limit $\Lambda \uparrow \mathbf{Z}^2$) that $\mu \preceq \mu^-$ on $\mathcal{F}_\Delta$. Since Δ was arbitrary and μ^- is minimal we find that $\mu = \mu^-$, and the lemma is proved. □

The next lemma is a variant of another result of Russo [11].

Lemma 2.2 (Flip-reflection domination) *Let $\mu \in \mathcal{G}$ and R any reflection, and suppose that for μ-almost all ω each finite $\Delta \subset \mathbf{Z}^2$ is surrounded by an R-invariant $*$circuit c such that $\omega \ge R \circ T(\omega)$ on c. Then $\mu \succeq \mu \circ R \circ T$.*

Proof. Another way of stating the assumption is that for any finite R-invariant Δ and μ-almost all ω there exists a finite R-invariant random set $\Gamma(\omega) \supset \Delta$ such that $\omega \ge R \circ T(\omega)$ on $\partial\Gamma(\omega)$. Given any $\varepsilon > 0$, we can thus find an R-invariant Λ so large that with probability at least $1 - \varepsilon$ such an R-invariant $\Gamma(\omega)$ exists within Λ. Since the union of any two such $\Gamma(\omega)$'s enjoys the same properties, we can assume that $\Gamma(\omega)$ is chosen maximal in Λ; in the case when no such $\Gamma(\omega)$ exists we set $\Gamma(\omega) = \emptyset$. The maximality of Γ implies that the events $\{\Gamma = G\}$ are measurable with respect to $\mathcal{F}_{\Lambda \setminus G}$. For any increasing $\mathcal{F}_\Delta$-measurable function $f \ge 0$ we thus get from the strong Markov property

$$\mu(f) \ge \mu\Big(\mu_\Gamma^{\cdot}(f) \; ; \; \Gamma \neq \emptyset\Big) \; .$$

However, if $\Gamma(\omega) \neq \emptyset$ then $\omega \geq R \circ T(\omega)$ on $\partial\Gamma(\omega)$. By stochastic monotonicity, for such ω we have

$$\mu^{\omega}_{\Gamma(\omega)}(f) \geq \mu^{R\circ T(\omega)}_{\Gamma(\omega)}(f) = \mu^{\omega}_{\Gamma(\omega)}(f \circ R \circ T) \, ,$$

where the identity follows from the R-invariance of Γ and the $R \circ T$ -invariance of the interaction. Hence

$$\mu(f) \geq \mu\Big(f \circ R \circ T \, ; \, \Gamma \neq \emptyset\Big) \geq \mu(f \circ R \circ T) - \varepsilon \, \|f\|_{\infty} \, .$$

The lemma thus follows by letting $\varepsilon \to 0$ and $\Delta \uparrow \mathbf{Z}^2$. $\square$

A third useful result of Russo [11] is the following. To state it we need to introduce two notations. First, let

$$\theta = \mu^{+}(0 \in I^{+*})$$

be the μ^{+}-probability that the origin belongs to an infinite $+*$cluster. Lemma 2.1 implies that $\theta > 0$. Secondly, for a half-plane π with boundary line ℓ and a $*$semicircuit σ in π we write $\mathrm{Int}\,\sigma$ for the unique subset of $\mathbf{Z}^2$ which is invariant under the reflection R in ℓ and satisfies $\pi \cap \partial(\mathrm{Int}\,\sigma) = \sigma$; we call $\mathrm{Int}\,\sigma$ the interior of σ.

Lemma 2.3 (Point-to-semicircuit lemma) *Let π be some half-plane with boundary line ℓ, $x \in \ell$, and σ a $*$semicircuit in π with interior $\Lambda = \mathrm{Int}\,\sigma \ni x$. Let $\omega \in \Omega$ be such that $\omega \equiv +1$ on σ. Then*

$$\mu^{\omega}_{\Lambda}\Big(x \text{ is in } \Lambda \ +*\text{connected to } \sigma\Big) \geq \theta/2 \, .$$

Proof. By hypothesis we have $\omega \geq R \circ T(\omega)$ on $\partial\Lambda$, and therefore $\mu^{\omega}_{\Lambda} \succeq \mu^{\omega}_{\Lambda} \circ R \circ T$. To exploit this relation we let $B_{x,\sigma}$ be the event that there exists a $+*$paths in Λ from x to σ, $C_{x,\sigma}$ the event that x is surrounded by a $+*$circuit in Λ which is $+*$connected to σ, and $D_{x,\sigma} = B_{x,\sigma} \cup C_{x,\sigma}$. Then $\mu^{\omega}_{\Lambda}(D_{x,\sigma} \cup R\circ T(D_{x,\sigma})) = 1$, and therefore $\mu^{\omega}_{\Lambda}(D_{x,\sigma}) \geq 1/2$. Hence

$$\mu^{\omega}_{\Lambda}(B_{x,\sigma}) \geq \mu^{\omega}_{\Lambda}(B_{x,\sigma}|C_{x,\sigma}) \, \mu(D_{x,\sigma}) \geq \mu^{\omega}_{\Lambda}(B_{x,\sigma}|C_{x,\sigma})/2 \, .$$

But if $C_{x,\sigma}$ occurs then there exists a largest random set $\Gamma \subset \Lambda$ containing x such that $\partial\Gamma$ forms a $+*$circuit and is $+*$connected to σ. Writing $B_{x,\partial\Gamma}$ for the event that x is $+*$connected to $\partial\Gamma$ and using the strong Markov property we thus find that

$$\mu^{\omega}_{\Lambda}(B_{x,\sigma}|C_{x,\sigma}) = \mu^{\omega}_{\Lambda}\,(\,\mu^{+}_{\Gamma}(B_{x,\partial\Gamma})\,|\,C_{x,\sigma}\,) \geq \theta$$

because $\mu^{+}_{\Gamma}(B_{x,\partial\Gamma}) \geq \theta$ by stochastic monotonicity. Together with the previous inequality this gives the result. $\square$

3 Percolation in half-planes

In this section we will prove that there exist plenty of infinite clusters of constant spin in the half-planes of $\mathbf{Z}^2$. In particular, this will show that all translation invariant $\mu \in \mathcal{G}$ are mixtures of μ^{+} and μ^{-}. We will use two pearls of percolation theory, the

Burton-Keane uniqueness theorem [2], and Zhang's argument for the non-existence of two infinite clusters of opposite sign in $\mathbf{Z}^2$.

For a given half-plane π we let E_π^+ denote the event that there exists an infinite +cluster in π. When this occurs, we will write I_π^+ for such an infinite +cluster in π. (As we will see, such clusters are unique, so that this notation does not lead into conflicts.) In case of the standard half-planes, we will only keep the directional index and omit the π; for example, we write E_{up}^+ for $E_{\pi_{\text{up}}}^+$. Similar notations will be used for +∗clusters and for the sign $-$ instead of $+$.

Let us say that (π, π') is a pair of *conjugate half-planes* if π and π' share only a common boundary line. An associated pair $(I_\pi^+, I_{\pi'}^+)$ or $(I_\pi^-, I_{\pi'}^-)$ of infinite clusters of the same sign in π and π' will be called an *infinite butterfly*. (This name alludes to the assumption that the two infinite 'wings' have the same 'color', but is not meant to suggest that they are symmetric and connected to each other, although the latter will turn out to be true.)

Lemma 3.1 (Butterfly lemma) *$\mathcal{G}$-almost surely there exists at least one infinite butterfly.*

Proof: Suppose the contrary. By the extremal decomposition theorem and the fact that the existence of infinite butterflies is a tail measurable event, there is then some $\mu \in \operatorname{ex}\mathcal{G}$ for which there exists no infinite butterfly μ-almost surely. We will show that this is impossible.

Step 1. First we observe that μ is $R \circ T$-invariant for all reflections $R = R_{k,\text{hor}}$ or $R_{k,\text{vert}}$, and in particular is periodic under translations. Indeed, let (π, π') be conjugate half-planes with common boundary line ℓ and R the reflection in ℓ mapping π onto π'. By the absence of infinite butterflies, at least one of the half-planes π and π' contains no infinite $-$cluster, and this or the other half-plane contains no infinite +cluster. In view of the tail triviality of μ, we can assume that $\mu(E_\pi^-) = 0$. This means that for μ-almost all ω every finite $\Delta \subset \pi$ is surrounded by some +∗semicircuit γ in π. For such a γ, $c = \gamma \cup R(\gamma)$ is an R-invariant ∗circuit that surrounds $\Delta \cup R(\Delta)$ and satisfies $\omega \geq R \circ T(\omega)$ on c. By Lemma 2.2, this gives the flip-reflection domination $\mu \succeq \mu \circ R \circ T$. Since also $\mu(E_\pi^+) = 0$ or $\mu(E_{\pi'}^+) = 0$, we conclude in the same way that $\mu \preceq \mu \circ R \circ T$, so that $\mu = \mu \circ R \circ T$. Since both ϑ_{hor}^2 and $\vartheta_{\text{vert}}^2$ are compositions of two reflections, the invariance under the translation group $(\vartheta_x)_{x \in 2\,\mathbf{Z}^2}$ follows.

Step 2. We now take advantage of the Burton–Keane uniqueness theorem [2], stating that for every periodic μ with finite energy there exists at most one infinite + (resp. $-$) cluster, and Zhang's symmetry argument (cf. [6], Theorem 5.18) deducing from this uniqueness the absence of simultaneous + and $-$percolation.

We start noting that, by the flip-reflection symmetry of μ, μ is different from μ^+ and μ^-, so that by Lemma 2.1 and the Burton–Keane uniqueness theorem there exist both a unique infinite +cluster I^+ and a unique infinite $-$cluster I^- in the whole plane $\mathbf{Z}^2$ μ-almost surely. We now choose a square $\Lambda = [-n, n]^2 \cap \mathbf{Z}^2$ so large that $\mu(\Lambda \cap I^+ \neq \emptyset) > 1 - 2^{-12}$. Let $\partial_k \Lambda$ be the intersection of $\partial\Lambda$ with the k'th quadrant, and let A_k^+ be the increasing event that there exists an infinite +path in Λ^c starting from some site in $\partial_k\Lambda$. Define A_k^- analogously. Since

$$\{\Lambda \cap I^+ \neq \emptyset\} \subset \bigcup_{k=1}^{4} A_k^+$$

and μ (as an extremal Gibbs measure) has positive correlations, it follows that

$$\prod_{k=1}^{4} \mu(\Omega \setminus A_k^+) \leq \mu(\bigcap_{k=1}^{4} \Omega \setminus A_k^+) \leq \mu(\Lambda \cap I^+ = \emptyset) < 2^{-12} ,$$

whence there exists some $k \in \{1, \ldots, 4\}$ such that $\mu(\Omega \setminus A_k^+) < 2^{-3}$. For notational convenience we assume that $k = 1$. By the flip-reflection symmetry shown above, we find that

$$\mu(A_1^+ \cap A_2^- \cap A_3^+ \cap A_4^-) > 1 - 4 \cdot 2^{-3} = 1/2 ,$$

which is impossible because on this intersection the infinite clusters I^+ and I^- cannot be both unique. This contradiction concludes the proof of the lemma. □

The butterfly lemma leads immediately to the following result first obtained by Messager and Miracle-Sole [10] by means of correlation inequalities of symmetry type; the following proof appeared first in [6].

Corollary 3.2 (Periodic Gibbs measures) *Any periodic $\mu \in \mathcal{G}$ is a mixture of μ^+ and μ^-.*

Proof. Suppose $\mu \in \mathcal{G}$ is invariant under $(\vartheta_x)_{x \in p\mathbf{Z}^2}$ for some period $p \geq 1$. Conditioning μ on any periodic tail event E we obtain again a periodic Gibbs measure. It is therefore sufficient to show that $\mu(E^+ \cap E^-) = 0$. Indeed, the butterfly lemma then shows that $\mu(E^+) + \mu(E^-) = 1$, and Lemma 2.1 implies that $\mu(\,\cdot\,|E^+) = \mu^+$ and $\mu(\,\cdot\,|E^-) = \mu^-$ whenever these conditional probabilities are defined. Hence $\mu = \mu(E^+)\,\mu^+ + \mu(E^-)\,\mu^-$.

Suppose by contraposition that $\mu(E^+ \cap E^-) > 0$. Since $E^+ \cap E^-$ is invariant and tail measurable, we can in fact assume that $\mu(E^+ \cap E^-) = 1$; otherwise we replace μ by $\mu(\,\cdot\,|E^+ \cap E^-)$. By the butterfly lemma, there exists a pair (π, π') of conjugate halfplanes, say $\pi_{\rm up}$ and $\pi_{\rm down}$, and a sign, say $+$, such that both half-planes contain an infinite clusters of this sign with positive probability. Since $\mu(E^-) = 1$ by assumption, we can find a large square Δ such that with positive probability Δ meets infinite $+$clusters in $\pi_{\rm up}$ and $\pi_{\rm down}$ and also an infinite $-$cluster. This $-$cluster leaves Δ either on the left or on the right between the two infinite $+$clusters. We can assume that the latter occurs with positive probability. By the finite energy property, it then follows that also $\mu(A_0) > 0$, where for $k \in p\mathbf{Z}$ we write A_k for the event that the point $(k, 0)$ belongs to a two-sided infinite $+$path with its two halves staying in $\pi_{\rm up}$ resp. $\pi_{\rm down}$, and $(k+1, 0)$ belongs to an infinite $-$cluster.

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 (cf. Lemma (18.15) of [5]) then show that $\mu(A_0 \setminus A) = 0$, and therefore $\mu(A) > 0$. But on A there exist infinitely many $-$clusters which are separated from each other by the infinitely many 'vertical' $+$paths. This contradicts the Burton–Keane theorem. □

The preceding argument actually shows that $\mu(E^{-*} \cap E^{+*}) = 0$ whenever $\mu \in \mathcal{G}$ is periodic. Since $\mu^+(E^+) = 1$ by Lemma 2.1 and tail triviality, this shows that in the +phase the +spins form an infinite sea with only finite islands.

Corollary 3.3 (Plus-sea in the plus-phase) $\mu^+(E^{-*}) = 0$. *Hence, μ^+-almost surely there exists a unique infinite +cluster I^+ in $\mathbf{Z}^2$ which surrounds each finite set.*

We note that in contrast to Zhang's argument (cf. Theorem 5.18 of [6]) our proof of the preceding corollary does not rely on the reflection invariance of μ^+ but only on its periodicity, and thus can be extended to the setting of Section 6 below.

We conclude this section with the observation that percolation in half-planes is not affected by spatial shifts.

Lemma 3.4 (Shift lemma) *Let π and $\tilde{\pi}$ be two half-planes such that $\pi \supset \tilde{\pi}$, i.e., π and $\tilde{\pi}$ are translates of each other. Then $E^+_\pi = E^+_{\tilde{\pi}}$ $\mathcal{G}$-almost surely, and similarly with $-$ instead of $+$.*

Proof. Since trivially $E^+_\pi \supset E^+_{\tilde{\pi}}$, we only need to show that $E^+_\pi \subset E^+_{\tilde{\pi}}$ $\mathcal{G}$-almost surely. For definiteness we consider the case when $\pi = \pi_{\mathrm{up}} = \{x_2 \geq 0\}$ and $\tilde{\pi} = \{x_2 \geq 1\}$. Take any $\mu \in \operatorname{ex}\mathcal{G}$, and suppose that $\mu(E^+_{\tilde{\pi}}) = 0$. Then for almost all ω and any $n \geq 1$ there exists a smallest $-*$semicircuit $\sigma_n(\omega)$ in $\tilde{\pi}$ containing $\Delta_n \cup \sigma_{n-1}(\omega)$ in its interior; here $\Delta_n = [-n, n] \times [1, n]$ and $\sigma_0 = \emptyset$. Let $x_n(\omega) \in \ell_{\mathrm{left}}$ and $y_n(\omega) \in \ell_{\mathrm{right}}$ be the two points facing the two endpoints of $\sigma_n(\omega)$; these are $\mathcal{F}_{\tilde{\pi}}$-measurable functions of ω, and the random sets $\{x_n, y_n\}$ are pairwise disjoint. Let A_n be the event that the spins at x_n and y_n take value -1.

We claim that A_n occurs for infinitely many n with probability 1. Indeed, for each $N \geq 1$ and any $x \in \ell_{\mathrm{left}}$, $y \in \ell_{\mathrm{right}}$ we have

$$\begin{aligned} &\mu\left(A_N \,\middle|\, x_N = x, y_N = y, \bigcap_{n>N} A_n^c\right) \\ &= \mu\left(\mu^{\cdot}_{\{x,y\}}(\omega(x) = \omega(y) = -1) \,\middle|\, x_N = x, y_N = y, \bigcap_{n>N} A_n^c\right) \\ &\geq \delta^2 > 0 \end{aligned}$$

because the event in the condition is measurable with respect to $\{x, y\}^c$ and the one-point conditional probabilities of μ are bounded from below by $\delta = [1 + e^{8\beta}]^{-1}$. Hence $\mu(A_N^c | \bigcap_{n>N} A_n^c) \leq 1 - \delta^2$, and by iteration we get $\mu(\bigcap_{n \geq N} A_n^c) = 0$, proving the claim.

We thus conclude that with probability 1 each box $[-n, n] \times [0, n]$ is surrounded by a $-*$semicircuit in π_{up}, which means that $\mu(E^+_{\mathrm{up}}) = 0$. As $\mu(E^+_{\tilde{\pi}})$ is either 0 or 1, the lemma follows. □

4 Uniqueness of semi-infinite clusters

Our next subject is the uniqueness of infinite clusters in half-planes, together with the stronger property that such clusters touch the boundary line infinitely often. This result has already been obtained by Russo [11] on the basis of the subsequent Lemma 4.2 which we derive here differently from the preceding Corollary 3.2.

Lemma 4.1 (Line touching lemma) *For any half-plane π, there exists $\mathcal{G}$-almost surely at most one infinite $+$ (resp. $+*$) cluster I_π^+ (resp. I_π^{+*}) in π. When it exists, this infinite cluster $\mathcal{G}$-almost surely intersects the boundary line ℓ of π infinitely often, in the sense that outside any finite Δ one can find an infinite path in this cluster starting from ℓ.*

Proof. For definiteness we assume that $\pi = \pi_{\mathrm{up}}$; other half-planes merely correspond to a change of coordinates. We consider only infinite $+$clusters in π_{up}; the case of $+*$clusters is similar. It is also clear that any result proved for the $+$sign is also valid with the $-$sign.

Uniqueness: The uniqueness of infinite $+$clusters in π_{up} is a consequence of the second statement, the line-touching property for infinite $-*$clusters. Indeed, suppose there exists no infinite $-*$cluster in π_{up}; then each finite set in π_{up} is surrounded by a $+$semicircuit, so that any two infinite $+$paths are necessarily $+$connected to each other. In the alternative case when an infinite $-*$cluster I_{up}^{-*} in π_{up} exists, this I_{up}^{-*} meets ℓ_{left} or ℓ_{right} infinitely often, so that each infinite $+$cluster must meet the other half-line infinitely often. Hence, two such $+$clusters must cross each other, and are thus identical.

Line touching: Let $\mu \in \mathrm{ex}\,\mathcal{G}$ and $x \in \pi_{\mathrm{up}}$ and consider the event A_x^+ that x belongs to an infinite $+$cluster in π_{up} which does not touch the horizontal axis ℓ_{hor}. We will show that $\mu(A_x^+) = 0$. Once this is established, we can take the union over all x and use the finite energy property to see that for each finite Δ the event "an infinite $+$cluster in π_{up} is not connected to ℓ_{hor} outside Δ" also has probability zero, which means that almost surely any infinite $+$cluster in π_{up} must meet ℓ_{hor} infinitely often.

Intuitively, if A_x^+ occurs then the infinite $+$cluster containing x is separated from ℓ_{hor} by an infinite $-*$path; but the spins 'above' this path feel only the $-$boundary condition and thus believe to be in the $-$phase μ^-, so that they will not form an infinite $+$cluster.

To make this intuition precise we fix some integer $k \geq 1$ and consider the event $A_{x,k}^+$ that x belongs to a $+$cluster of size at least k which does not meet ℓ_{hor}. Take a large box $\Delta \subset \pi_{\mathrm{up}}$ containing x. For $\omega \in A_{x,k}^+$ we consider the largest set $\Gamma(\omega) \subset \Delta$ containing x such that $\omega = -1$ on $\partial\Gamma(\omega) \setminus \partial_{\mathrm{up}}\Delta$, where $\partial_{\mathrm{up}}\Delta = \partial\Delta \cap \pi_{\mathrm{up}}$. By the strong Markov property and the stochastic monotonicity of Gibbs distributions, we find that

$$\mu(A_{x,k}^+) \leq \mu\left(\mu_\Gamma^{\cdot}(E_{x,k}^+)\right) \leq \mu_\Delta^{\pm}(E_{x,k}^+)\ ,$$

where $E_{x,k}^+$ is the event that x belongs to a $+$cluster in π_{up} of size at least k and $\pm$ stands for the configuration which is $+1$ on π_{up} and -1 on π_{up}^c. Now, again by stochastic monotonicity the semi-infinite limit $\mu_{\mathrm{up}}^{\pm} = \lim_{\Delta\uparrow\pi_{\mathrm{up}}} \mu_\Delta^{\pm}$ exists. Letting first $\Delta \uparrow \mathbf{Z}^2$ and then $k \to \infty$ we thus find that $\mu(A_x^+) \leq \mu_{\mathrm{up}}^{\pm}(E_{\mathrm{up}}^+)$. The lemma thus follows from the subsequent lemma. □

By the argument of Lemma 2.1, the following result implies the uniqueness of the Gibbs measure on π_{up} with $-$boundary condition in π_{down}.

Lemma 4.2 (No percolation on a bordered half-plane) $\mu_{\mathrm{up}}^{\pm}(E_{\mathrm{up}}^{+*}) = 0.$

Proof. To begin we note that $\mu_{\mathrm{up}}^{\pm}$ is invariant under horizontal translations and stochastically maximal in the set of all Gibbs measures on π_{up} with $-$boundary condition in

π_{down}. This follows just as in the case of the plus-phase μ^+ on the whole lattice. In particular, $\mu^{\pm}_{\text{up}}$ is trivial on the π_{up}-tail $\mathcal{T}_{\text{up}} = \bigcap\{\mathcal{F}_{\pi_{\text{up}}\setminus\Lambda} : \Lambda \subset \pi_{\text{up}} \text{ finite}\}$. We think of $\mu^{\pm}_{\text{up}}$ as a probability measure on Ω for which almost all configurations are identically equal to -1 on π^c_{up}.

Next we consider the downwards translates $\mu^+_{n,-} = \mu^{\pm}_{\text{up}} \circ \vartheta^{-n}_{\text{vert}}$, $n \geq 0$. Evidently, $\mu^+_{n,-}$ is obtained by an analogous infinite-volume limit in the half-plane $\{x_2 \geq -n\}$. This shows that $\mu^+_{n,-} \preceq \mu^+_{n+1,-}$ by stochastic monotonicity, so that the stochastically increasing limit $\mu^+_- = \lim_{n\to\infty} \mu^+_{n,-}$ exists. Clearly $\mu^+_- \in \mathcal{G}$. Also, μ^+_- inherits the horizontal invariance of the $\mu^+_{n,-}$ and is in addition vertically invariant. Corollary 3.2 therefore implies that $\mu^+_- = a\,\mu^- + (1-a)\mu^+$ for some coefficient $a \in [0,1]$.

We claim that $a > 0$. For $n \geq 1$ let B_n denote the event that the origin is $-*$connected to the horizontal line $\{x_2 = -n\}$. By the horizontal ergodicity of $\mu^+_{n,-}$, there exist for $\mu^+_{n,-}$-almost all ω some random integers $m_{\text{left}}(\omega) < 0 < m_{\text{right}}(\omega)$ such that $\omega \equiv -1$ on

$$\sigma(\omega) = \left\{x \in \mathbf{Z}^2 : x_1 \in \{m_{\text{left}}(\omega), m_{\text{right}}(\omega)\},\ -n \leq x_2 \leq 0\right\} .$$

Together with a segment of the line $\{x_2 = -n-1\}$ on which $\omega = -1$ $\mu^+_{n,-}$-almost surely, $\sigma(\omega)$ forms a $-$semicircuit in π_{down} surrounding the origin. An immediate application of the strong Markov property and the point-to-semicircuit lemma thus implies that $\mu^+_{n,-}(B_n) \geq \theta/2$. Therefore, writing $E^{-*}_{0,m}$ for the event that the origin belongs to some $-*$ cluster of size at least m we find $\mu^+_{n,-}(E^{-*}_{0,m}) \geq \theta/2$ when $n \geq m$. Letting first $n \to \infty$ and then $m \to \infty$ we see that $\mu^+_-(E^{-*}) \geq \theta/2$. Since $\mu^+(E^{-*}) = 0$ by Corollary 3.3, it follows that $a \geq \theta/2$, and the claim is proved.

To conclude the proof we observe that

$$\mu^{\pm}_{\text{up}}(E^{+*}_{\text{up}}) \leq \mu^+_-(E^{+*}) = 1 - a < 1\ ,$$

again by Corollary 3.3. Since $\mu^{\pm}_{\text{up}}$ is trivial on $\mathcal{T}_{\text{up}}$, the lemma follows. □

The butterfly lemma and shift lemma together still leave the possibility that all infinite butterflies have the same orientation, either horizontal or vertical. As a consequence of the line touching lemma, we can now show that both orientations must occur.

Lemma 4.3 (Orthogonal butterflies) *$\mathcal{G}$-almost surely there exist both a horizontal infinite butterfly in π_{up} and π_{down} as well as a vertical infinite butterfly in π_{left} and π_{right}.*

Proof. Suppose there exists some $\mu \in \text{ex}\,\mathcal{G}$ having almost surely no vertical infinite butterfly. By the first step in the proof of the butterfly lemma, it then follows that $\mu = \mu \circ R_{k,\text{vert}} \circ T$ for all $k \in \mathbf{Z}$, and thus $\mu = \mu \circ \vartheta^{-2}_{\text{hor}}$. By the tail triviality, μ is in fact ergodic under ϑ^2_{hor}. By the butterfly lemma, horizontal infinite butterflies do exist, say of color $+$.

We now use an argument similar to that in Corollary 3.2, with the line touching lemma in place of the Burton-Keane theorem. Fix any $n \geq 1$. For $k \in \mathbf{Z}$ let A_k denote the event that the point $(k,0)$ is connected by straight $+$paths to infinite $+$clusters in both $\pi_{n,\text{up}} = \{x \in \mathbf{Z}^2 : x_2 \geq n\}$ and $\pi_{n,\text{down}} = \{x_2 \leq -n\}$. Let A be the event that A_k

occurs for infinitely many $k < 0$ and infinitely many $k > 0$. The finite energy property then shows that $\mu(A_0) > 0$, and the horizontal ergodicity and Poincaré's recurrence theorem (or the ergodic theorem) imply that $\mu(A) = 1$. But the line touching lemma guarantees that the infinitely many doubly-infinite 'vertical' +paths passing through the horizontal axis are connected to each other in $\pi_{n,\text{up}}$ and $\pi_{n,\text{down}}$. As n was arbitrary, it follows that almost surely each finite set is surrounded by a +circuit, and an infinite −cluster cannot exist. In view of Lemma 2.1, this implies that $\mu = \mu^+$. But μ^+ is not invariant under $R_{\text{vert}} \circ T$, in contradiction to what we derived for μ. □

The preceding argument can be used to derive the result of Russo [11] that μ^+ and μ^- are the only phases which are periodic in one direction. We will not need this intermediate result.

5 Non-coexistence of phases

In this section we will prove the following proposition.

Proposition 5.1 (Absence of non-periodic phases) *Any Gibbs measure $\mu \in \mathcal{G}$ is invariant under translations, i.e., $\mu = \mu \circ \vartheta_{\text{hor}}^{-1}$ and $\mu = \mu \circ \vartheta_{\text{vert}}^{-1}$.*

Together with Corollary 3.2 this will immediately imply the main theorem that each Gibbs measure is a mixture of the two phases μ^+ and μ^-. Our starting point is the following lemma estimating the probability that a semi-infinite cluster can be pinned at a prescribed point.

Lemma 5.2 (Pinning lemma) *Let $\mu \in \mathcal{G}$, and suppose that there exists an infinite $+*$cluster I_{up}^{+*} in π_{up} which meets the right semiaxis ℓ_{right} infinitely often. Then for each finite square $\Delta = [-n, n]^2$ and $x \in \ell_{\text{right}}$ we have*

$$\mu\Big(x \text{ is } +*\text{connected in } (\Delta \cup \ell_{\text{left}})^c \text{ to } I_{\text{up}}^{+*}\Big) \geq \theta/4$$

provided x lies sufficiently far to the right. The same holds when 'left' and 'right' or 'up' and 'down' are interchanged.

Proof. By hypothesis, the infinite component of $I_{\text{up}}^{+*} \setminus \Delta$ almost surely contains infinitely many points of ℓ_{right}. Thus, if $x \in \ell_{\text{right}}$ is located far enough to the right then, with probability exceeding $1/2$, at least one such point can be found left from x, and another such point can be found right from x. This means that x is surrounded by a $+*$semicircuit σ in π_{up} which belongs to I_{up}^{+*} and satisfies $\Delta \cap \operatorname{Int} \sigma = \emptyset$.

Let Λ be a large square box containing x. If Λ is large enough, a semicircuit σ as above can be found within Λ with probability still larger than $1/2$. We then can assume that σ has the largest interior among all such semicircuits in Λ. Using the strong Markov property and the point-to-semicircuit lemma we get the result. □

Our main task in the following is to analyze the situation when a half-plane contains both an infinite +cluster and an infinite −cluster. (The line-touching lemma still allows this possibility.) In this situation it is useful to consider contours.

Remark 5.3 (Contours in half-planes) As is usually done in the Ising model, we draw lines of unit length between adjacent spins of opposite sign. We then obtain a system of polygonal curves running through the sites of the dual lattice $\mathbf{Z}^2 + (\frac{1}{2}, \frac{1}{2})$. A *contour* in a half-plane π is a part of these polygonal curves which separates a $-$cluster in π from a $+*$cluster in π. This corresponds to the convention that at crossing points the contours are supposed to bend around the $-$spins. (The artificial asymmetry between $+$ and $-$ does not matter, and we could clearly make the opposite convention.)

Suppose now that π contains both an infinite $+*$cluster I_π^{+*} and an infinite $-$cluster I_π^-. Since these are unique by the line touching lemma, it follows that π contains a unique semi-infinite contour γ_π which starts between two points of the boundary line ℓ. On its two sides, γ_π is accompanied by an infinite $+*$path f_π^+ and an infinite $-$path f_π^-. We call f_π^+ and f_π^- the $+$ resp. $-$face of γ_π.

By the shift lemma, γ_π intersects each line $\ell' \subset \pi$ parallel to ℓ only finitely often. Indeed, the shifted half-plane π' with boundary line ℓ' also contains a unique semi-infinite contour $\gamma_{\pi'}$. Since the faces of $\gamma_{\pi'}$ are connected to ℓ by the line touching lemma, $\gamma_{\pi'}$ necessarily coincides with γ_π up to finitely many steps.

From now on we consider a fixed extremal Gibbs measure $\mu \in \operatorname{ex}\mathcal{G}$. We want to prove that μ is horizontally invariant. (The proof of vertical invariance is similar.) To this end we consider its horizontal translate $\hat{\mu} = \mu \circ \vartheta_{\text{hor}}^{-1}$, as well as the product measure $\hat{\nu} = \mu \times \hat{\mu}$ on $\Omega \times \Omega$. It is convenient two think of the latter as a duplicated system consisting of two independent layers. The following lemma is a slight variation of a result of Aizenman [1] in his proof of the main theorem.

Lemma 5.4 (Fluctuations of the infinite contour) *Suppose π_{up} contains a semi-infinite contour γ_{up} μ-almost surely. Then for $\hat{\nu}$-almost all $(\omega, \hat{\omega}) \in \Omega^2$, $\gamma_{\text{up}}(\omega)$ and $\gamma_{\text{up}}(\hat{\omega})$ intersect each other infinitely often.*

Proof. By the line touching lemma and tail triviality, we can assume that I_{up}^- intersects ℓ_{right} infinitely often and I_{up}^{+*} intersects ℓ_{left} infinitely often. (The alternative case is analogous.) For $n \geq 1$ let $\pi_{n,\text{up}} = \vartheta_{\text{vert}}^n \pi_{\text{up}}$ be the half-plane above the horizontal line through $(0, n)$, $I_{n,\text{up}}^-$ the infinite $-$cluster in this half-plane and

$$a_n = \min\{k \in \mathbf{Z} : (k, n) \in I_{n,\text{up}}^-\}$$

the abscissa of the point at which γ_{up} enters definitely into $\pi_{n,\text{up}}$. Since γ_{up} is a continuous curve, it is sufficient to show that for $\hat{\nu}$-almost all $(\omega, \hat{\omega})$ we have $a_n(\omega) \geq a_n(\hat{\omega})$ and $a_{n+1}(\omega) < a_{n+1}(\hat{\omega})$ infinitely often. Using the definition $\hat{\mu} = \mu \circ \vartheta_{\text{hor}}^{-1}$ and setting $\nu = \mu \times \mu$ and $d_n(\omega, \omega') = a_n(\omega) - a_n(\omega')$ for $\omega, \omega' \in \Omega$, we need to show that $d_n \geq 1 > d_{n+1}$ infinitely often ν-almost surely.

We observe first that $d_n = 0$ infinitely often ν-almost surely. Indeed, the complementary event consists of the two parts $A = \{d_n \geq 1 \text{ eventually}\}$ and $B = \{d_n \leq -1 \text{ eventually}\}$. Symmetry implies that $\nu(A) = \nu(B)$. On the other hand, the tail-triviality of μ implies that $\nu(A) = 0$ or 1. This follows from Fubini's theorem because A is measurable with respect to the 'product-tail' $\mathcal{T}^{(2)} = \bigcap\{\mathcal{F}_{\Lambda^c}^2 : \Lambda \subset \mathbf{Z}^2 \text{ finite}\}$ in Ω^2. (One should not be mistaken to believe that A was measurable with respect to the smaller 'tail-product' $\mathcal{T}^2$. It is only the case that the ω-section A_ω of A belongs to $\mathcal{T}$ for any ω, and the function $\omega \to \mu(A_\omega)$ is $\mathcal{T}$-measurable.) The disjointness of A and B thus implies that $\nu(A) = \nu(B) = 0$.

Next we claim that there exists some constant $\delta > 0$ such that

$$\nu(d_{n-1} \geq 1 | d_n, d_{n+1}, \dots) \geq \delta \quad \text{on } \{d_n = 0\} .$$

To see this let $A_{n,k} = \{(\omega, \omega') : a_n(\omega) = a_n(\omega') = k\}$, $\Delta_{n,k}$ the two-point set consisting of the points $(k, n-1)$ and $(k+1, n-1)$, and $B_{n,k}$ the event that $\omega = (-1,-1)$ on $\Delta_{n,k}$ and $\omega' = (-1,+1)$ on $\Delta_{n,k}$. Then

$$\nu(B_{n,k} | \mathcal{F}^2_{\Delta^c_{n,k}})(\omega, \omega') = \mu^{\omega}_{\Delta_{n,k}} \times \mu^{\omega'}_{\Delta_{n,k}}(B_{n,k}) \geq [1 + e^{8\beta}]^{-4} \equiv \delta$$

and thus

$$\nu(d_{n-1} \geq 1 | \mathcal{F}^2_{n,\mathrm{up}}) \geq \nu(B_{n,k} | \mathcal{F}^2_{n,\mathrm{up}}) \geq \delta \quad \text{on } A_{n,k} \ \nu\text{-almost surely}$$

because $A_{n,k} \cap B_{n,k} \subset \{d_{n-1} \geq 1\}$ and $\pi_{n,\mathrm{up}} \subset \Delta^c_{n,k}$. Summing over k and conditioning on $\sigma(d_n, d_{n+1}, \dots)$ we get the claim.

Now, given any integers $1 \leq N < L$ let $\tau = \tau_{N,L}$ be the largest $n \in \{N, \dots, L\}$ such that $d_n = 0$, provided such an n exists, and $\tau = 0$ otherwise. Then $\{\tau = n\} \in \sigma(d_n, d_{n+1}, \dots)$ for all n, and therefore

$$\begin{aligned} \nu\Big(\exists\, n \geq N : d_{n-1} \geq 1 \ > \ d_n\Big) &\geq \sum_{n=N}^{L} \nu(d_{n-1} \geq 1,\, \tau = n) \\ = \sum_{n=N}^{L} \nu\Big(\nu(d_{n-1} \geq 1 | d_n, d_{n+1}, \dots)\, 1_{\{\tau = n\}}\Big) &\ \geq \ \delta\, \nu\Big(\exists\, N \leq n \leq L : d_n = 0\Big) . \end{aligned}$$

Letting first $L \to \infty$ and then $N \to \infty$ we see that $d_{n-1} \geq 1 > d_n$ infinitely often with probability at least δ. This gives the result because ν is trivial on the product-tail $\mathcal{T}^{(2)}$. $\square$

Our key observation is the following percolation result for the duplicated system with distribution $\hat{\nu}$. We will say that a path in $\mathbf{Z}^2$ is a $\leq$path for a pair $(\omega, \hat{\omega}) \in \Omega^2$ if $\omega(x) \leq \hat{\omega}(x)$ for all its sites x. In the same way we define $\leq *$paths, and we can speak of $\leq *$circuits and $\leq *$clusters.

Lemma 5.5 (No $(+,-)$percolation in the duplicated system) *$\hat{\nu}$-almost surely each finite square $\Delta = [-n, n]^2$ is surrounded by a $\leq *$circuit in $\mathbf{Z}^2$.*

Proof. Consider any two points $x \in \ell_{\mathrm{left}}$ and $y \in \ell_{\mathrm{right}}$. We claim that with $\hat{\nu}$-probability at least $(\theta/4)^2$ there exists a $\leq *$path from x to y 'above' Δ, provided x and y are located sufficiently far to the left resp. to the right. We distinguish three cases.

Case 1: $\mu(E^+_{\mathrm{up}}) = 0$. By Lemma 4.3, π_{up} then almost surely contains an infinite $-$cluster I^-_{up}, and each finite subset of π_{up} is surrounded by a $-*$semicircuit in π_{up}. In other words, an infinite $-*$cluster I^{-*}_{up} in π_{up} exists and touches both ℓ_{left} and ℓ_{right} infinitely often. By the pinning lemma and the positive correlations of μ, with μ-probability at least $(\theta/4)^2$ both x and y are $-*$connected to I^{-*}_{up} outside Δ, and therefore also $-*$connected to each other by a $-*$path p above Δ. However, this $-*$path p on the first layer is certainly also a $\leq*$path for the duplicated system, and the claim follows.

Case 2: $\mu(E^-_{\rm up}) = 0$. In this case we also have $\hat{\mu}(E^-_{\rm up}) = 0$. Interchanging $+$ and $-$ and replacing μ by $\hat{\mu}$ in Case 1, we find that with $\hat{\mu}$-probability at least $(\theta/4)^2$, there exists a $+*$path $\hat{p}$ in the second layer above Δ from x to y. Since $\hat{p}$ is again a $\leq*$path for the duplicated system, the claim follows as in the first case.

Case 3: $\mu(E^+_{\rm up}) = \mu(E^-_{\rm up}) = 1$. Then μ-almost surely there exists a unique semi-infinite contour $\gamma_{\rm up}$, and by tail triviality we can assume (for definiteness) that $\gamma_{\rm up}$ has its $+$face on the left-hand side μ-almost surely, and thus also $\hat{\mu}$-almost surely. By the pinning lemma and the independence of the two layers, the following event has $\hat{\nu}$-probability at least $(\theta/4)^2$:

- in the first layer, y is $-*$connected off Δ to $I^-_{\rm up}(\omega)$, and thus to the $-$face $f^-_{\rm up}(\omega)$ of $\gamma_{\rm up}(\omega)$; that is, there exists an infinite $-*$path $p^-_y(\omega)$ from y outside Δ eventually running along $\gamma_{\rm up}(\omega)$;
- in the second layer, x is $+*$connected off Δ to $I^{+*}_{\rm up}(\hat{\omega})$, and thus to the $+$face $f^+_{\rm up}(\hat{\omega})$ of $\gamma_{\rm up}(\hat{\omega})$; that is, there exists an infinite $+*$path $p^+_x(\hat{\omega})$ from x outside Δ eventually running along $\gamma_{\rm up}(\hat{\omega})$.

Since $\gamma_{\rm up}(\omega)$ and $\gamma_{\rm up}(\hat{\omega})$ intersect each other infinitely often by Lemma 5.4, the union of $p^-_y(\omega)$ and $p^+_x(\hat{\omega})$ contains a $*$path from x to y which by construction is a $\leq*$path for the duplicated system. This proves the claim in the final case.

To conclude the proof of the lemma, we let $A_{x,y}$ denote the event that there exist a $\leq*$path from x to y above Δ, and $B_{x,y}$ the event that such a path exists below Δ. The indicator functions of these events can be written as increasing functions f resp. g of the difference configuration $\hat{\omega} - \omega$. Using the positive correlations of μ and $\hat{\mu}$ we thus obtain

$$\begin{aligned}\hat{\nu}(A_{x,y} \cap B_{x,y}) &= \int \mu(d\omega) \int \hat{\mu}(d\hat{\omega})\ f(\hat{\omega} - \omega)\, g(\hat{\omega} - \omega) \\ &\geq \int \mu(d\omega)\ \hat{\mu}(f(\cdot - \omega))\, \hat{\mu}(g(\cdot - \omega)) \\ &\geq \hat{\nu}(A_{x,y})\, \hat{\nu}(B_{x,y}) \ \geq\ (\theta/4)^4 \ .\end{aligned}$$

The last inequality follows from the claim and its analogue for the lower half-plane. However, if $A_{x,y} \cap B_{x,y}$ occurs then Δ is surrounded by a $\leq*$circuit for the duplicated system. Letting $\Delta \uparrow \mathbf{Z}^2$ we see that with probability at least $(\theta/4)^4$ each finite set is surrounded by a $\leq*$circuit. Since this event is measurable with respect to the product-tail $\mathcal{T}^{(2)}$ on which $\hat{\nu}$ is trivial, the lemma follows. □

It is now easy to complete the proof of Proposition 5.1.

Proof of Proposition 5.1. Consider any square $\Delta = [-n, n]^2$, and let $\varepsilon > 0$. By Lemma 5.5, Δ is $\hat{\nu}$-almost surely surrounded by a $\leq*$circuit, and with probability at least $1 - \varepsilon$ such a $\leq*$circuit can be found in a sufficiently large square Λ. Let Γ be the interior of the largest such $\leq*$circuit; if no such $\leq*$circuit exists let $\Gamma = \emptyset$. Then we find for any increasing $\mathcal{F}_\Delta$-measurable function $0 \le f \le 1$, using the strong Markov property of $\hat{\nu}$ and the fact that $\mu_\Gamma^\omega \preceq \mu_\Gamma^{\hat\omega}$ when $\Gamma(\omega,\omega') \neq \emptyset$,

$$\begin{aligned}\mu(f) &= \hat{\nu}(f\otimes 1) \le \int_{\{\Gamma\neq\emptyset\}} d\hat{\nu}(\omega,\omega')\, \mu^{\omega}_{\Gamma(\omega,\omega')}(f) + \varepsilon \\ &\le \int d\hat{\nu}(\omega,\omega')\, \mu^{\omega'}_{\Gamma(\omega,\omega')}(f) + \varepsilon = \hat{\nu}(1\otimes f) + \varepsilon = \hat{\mu}(f) + \varepsilon \,.\end{aligned}$$

Letting $\varepsilon \to 0$ and $\Delta \uparrow \mathbf{Z}^2$ we find that $\mu \preceq \hat{\mu}$. Interchanging μ and $\hat{\mu}$ (i.e., the roles of the layers) we get the reverse relation. Hence $\mu = \hat{\mu}$, so that μ is horizontally invariant. The vertical invariance follows similarly by an interchange of coordinates. □

6 Extensions

Which properties of the square lattice $\mathbf{Z}^2$ entered into the preceding arguments? The only essential feature was its invariance under the reflections in all horizontal and vertical lines with integer coordinates. We claim that the theorem remains true for the Ising model on any graph $\mathcal{L}$ with these properties. (The Ising model on triangular and honeycomb lattices has already been treated in [4].)

To be more precise, let $\mathcal{R} = \{R_{k,\mathrm{hor}}, R_{k,\mathrm{vert}} : k \in \mathbf{Z}\}$ denote the set of all reflections of the Euclidean plane $\mathbf{R}^2$ in horizontal or vertical lines with integer coordinates, and suppose $\mathcal{L}$ is a countable subset of $\mathbf{R}^2$ which (after suitable scaling and rotation) is R-invariant for all $R \in \mathcal{R}$. Such an $\mathcal{L}$ is uniquely determined by its finite intersection with the unit cube $[0,1]^2$, and it is periodic with period 2. Suppose further that $\mathcal{L}$ is equipped with a symmetric neighbor relation '$\sim$' satisfying

(L1) each $x \in \mathcal{L}$ has only finitely many 'neighbors' $y \in \mathcal{L}$ satisfying $x \sim y$;

(L2) $x \sim y$ if and only if $Rx \sim Ry$ for all $R \in \mathcal{R}$.

If $x \sim y$ we say that x and y are connected by an edge, which is visualized by a straight line segment between x and y. The preceding assumptions simply mean that $(\mathcal{L}, \sim)$ is a locally finite graph admitting the reflections $R \in \mathcal{R}$ (and thereby the translations ϑ_x, $x \in 2\mathbf{Z}^2$) as graph automorphisms. The fundamental further assumption is

(L3) $(\mathcal{L}, \sim)$ is planar, i.e., the edges between different pairs of neighboring points do not cross each other except possibly at some lattice points.

The complement (in $\mathbf{R}^2$) of the union of all edges then splits into connected components called the faces of $(\mathcal{L}, \sim)$.

One basic consequence of planarity is that $(\mathcal{L}, \sim)$ admits a conjugate matching graph $(\mathcal{L}, \stackrel{*}{\sim})$. As indicated by the notation, this conjugate graph has the same set of vertices, but the relation $x \stackrel{*}{\sim} y$ holds if either $x \sim y$ or x and y are distinct points (on the border) of the same face of $(\mathcal{L}, \sim)$. (Note that this matching dual is in general not planar.) The edges of $(\mathcal{L}, \stackrel{*}{\sim})$ are then used to define the concept of $*$connectedness. The construction implies that every path in $(\mathcal{L}, \sim)$ is also a $*$path (i.e., a path in $(\mathcal{L}, \stackrel{*}{\sim})$), and that the outer boundary of any cluster is a $*$path, and vice versa. (The latter property holds for arbitrary matching pairs of graphs as defined in Kesten [9], e.g. However, we also used repeatedly the former property which does not extend to general matching pairs. In particular, this means that our results do not apply to the Ising model on the matching conjugate of $\mathbf{Z}^2$ having nearest-neighbor interactions *and* diagonal interactions.)

Another consequence of planarity is that we can draw contours separating clusters from $*$clusters. Such contours can either be visualized by broken lines passing through the edges of $(\mathcal{L}, \sim)$, or simply as a pair consisting of a path and an adjacent $*$path, namely the two faces of the contour.

In order to see how to work with a graph $(\mathcal{L}, \sim)$ as above we will discuss the proper definition of half-planes and their boundary lines. A half-plane π is still the intersection of $\mathcal{L}$ with a set of the form $\{x \in \mathbf{R}^2 : x_i \geq k\}$, $k \in \mathbf{Z}$, $i \in \{1, 2\}$, or with $\leq$ instead of $\geq$. However, the 'boundary line' ℓ is now in general not a straight line but rather the set $\ell = \{x \in \pi : x \sim y \text{ for some } y \notin \pi\}$. In particular, ℓ is not necessarily a line of fixed points for the reflection $R \in \mathcal{R}$ mapping π onto its conjugate halfplane π'. However, we can simply replace a point $x \in \ell$ by a pair (x, x') consisting of x and its R-image x' (which either coincides with x or is a neighbor of x). Similarly, the semiaxes ℓ_{up} etc. should be considered as sets of such pairs. With these modifications, all geometric arguments still work in the obvious way.

So, as a consequence of the preceding discussion we see that the square lattice $\mathbf{Z}^2$ can be replaced, for example, by

- the *triangular lattice* $\mathbf{T}$. This is the $\mathcal{R}$-invariant lattice satisfying $\mathbf{T} \cap [0, 1]^2 = \{(1, 0), (0, 1)\}$ and $(-1, 0) \sim (1, 0) \sim (0, 1) \sim (2, 1)$; the remaining edges result from (L2). Since all faces are triangles, $\mathbf{T}$ is self-matching. While π_{up} and π_{down} have a common straight boundary line, the boundaries of π_{right} and π_{left} are not straight; besides a common part on the vertical axis they also contain the adjacent points $(1, k)$ resp. $(-1, k)$, $k \in 2\mathbf{Z}$.
- the hexagonal or *honeycomb lattice* $\mathbf{H}$. Here, for example, $\mathbf{H} \cap [0, 1]^2 = \{(\frac{1}{3}, 1), (\frac{2}{3}, 0)\}$ and $(-\frac{1}{3}, 1) \sim (\frac{1}{3}, 1) \sim (\frac{2}{3}, 0) \sim (\frac{4}{3}, 0)$; all other edges are again determined by (L2). As in the triangular lattice, π_{up} and π_{down} have a common straight boundary line, but π_{right} and π_{left} have no common points.
- the *diced lattice*. This is obtained from the honeycomb lattice by placing points in the centers of the hexagonal faces and connecting them to the three points in the west, north-east and south-east of these faces; to obtain reflection symmetry an additional shift by $(-\frac{1}{3}, 0)$ is necessary. See p. 16 of [9] for more details.
- the covering lattice of the honeycomb lattice, the *Kagomé lattice*, cf. p. 37 of [9].

As for the interaction, it is neither necessary that the interaction along all bonds is the same, nor that it is invariant under the spin flip. Except for attractivity, we need only the invariance under simultaneous flip-reflections (which in particular implies periodicity with period 2). For example, we did not need that μ^+ and μ^- are invariant under all $R \in \mathcal{R}$ and related to each other by the spin flip T (cf. the comments after Corollary 3.3). We rather needed that $\mu^+ = \mu^- \circ R \circ T$ for all $R \in \mathcal{R}$ (implying that μ^+ and μ^- are periodic, and that any flip-reflection invariant μ is different from these phases; the latter was used in Lemmas 3.1 and 4.3). This, however, holds whenever the interaction is invariant under flip-reflections. The same invariance is sufficient for spin-reflection domination and the point-to-semicircuit lemma. So, the only thing to observe is that whenever we work with translations (as in the proof of Lemma 4.2 and in Section 5 5.4 and below) we have to confine ourselves to *even* translations, which does not raise any problems. As a result, we can consider any system of spins $\omega(x) = \pm 1$ with formal Hamiltonian of the form

$$H(\omega) = \sum_{x \sim y} U_{x,y}(\omega(x), \omega(y)) + \sum_{x \in \mathcal{L}} V_x(\omega(x)) \ , \tag{1}$$

where for all $a, b \in \{-1, 1\}$ we have $U_{x,y}(a, b) = U_{y,x}(b, a)$ and

(H1) $U_{x,y}(1, \cdot) - U_{x,y}(-1, \cdot)$ is decreasing on $\{-1, 1\}$;

(H2) $U_{x,y}(a, b) = U_{Rx,Ry}(-a, -b)$ and $V_x(a) = V_{Rx}(-a)$ for all $R \in \mathcal{R}$.

Assumption (H1) implies that the FKG inequality is applicable. We thus obtain the following general result.

Theorem 6.1 *Consider a planar graph $(\mathcal{L}, \sim)$ as above and an interaction of the form (1) satisfying (H1) and (H2). Then there exist no more than two extremal Gibbs measures.*

The standard case, of course, is the ferromagnetic Ising model without external field; this corresponds to the choice $U_{x,y}(a, b) = -\beta ab$ and $V_x \equiv 0$. But there is also another case of particular interest. Consider $\mathcal{L} = \mathbf{Z}^2 + (\frac{1}{2}, \frac{1}{2})$, the shifted square lattice with its usual graph structure. $\mathcal{L}$ is bipartite, in the sense that $\mathcal{L}$ splits into to disjoint sublattices, $\mathcal{L}_{even}$ and $\mathcal{L}_{odd}$, such that all edges run from one sublattice to the other. If we set $U_{x,y}(a, b) = -\beta ab$ and define a staggered external field

$$V_x(a) = \begin{cases} -ha & \text{if } x \in \mathcal{L}_{even} \\ ha & \text{if } x \in \mathcal{L}_{odd} \end{cases}$$

with $h \in \mathbf{R}$ then the conditions (H1) and (H2) hold; here we take advantage of the fact that the reflections $R \in \mathcal{R}$ map $\mathcal{L}_{even}$ into $\mathcal{L}_{odd}$ and vice versa. But it is well-known that this model is isomorphic to the *anti*ferromagnetic Ising model on $\mathbf{Z}^2$ with external field h; the isomorphism consists in flipping all spins in $\mathcal{L}_{odd}$. This gives us the following result.

Corollary 6.2 *For the Ising antiferromagnet on $\mathbf{Z}^2$ for any inverse temperature and arbitrary external field there exist at most two extremal Gibbs measures.*

This corollary does not extend to non-bipartite lattices such as the triangular lattice. In fact, for the Ising antiferromagnet on $\mathbf{T}$ one expects the existence of three different phases for suitable h.

However, there is a similar repulsive model to which our theorem applies, namely the hard-core lattice gas on $\mathbf{Z}^2$. In this model, the values -1 and 1 are interpreted as the absence resp. presence of a particle, and no particles are allowed to sit on adjacent places. Interchanging the values ± 1 on $\mathcal{L}_{odd}$ we obtain an isomorphic model which is defined by setting $U_{x,y}(a,b) = \infty$ if $x \in \mathcal{L}_{even}$ and $a = -b = 1$, or $x \in \mathcal{L}_{odd}$ and $a = -b = -1$, and $U_{x,y}(a,b) = 0$ otherwise; and $V_x(a) = -\log\lambda$ if $x \in \mathcal{L}_{even}$ and $a = 1$, or $x \in \mathcal{L}_{odd}$ and $a = -1$, and $V_x(a) = 0$ otherwise. The parameter $\lambda > 0$ is called the activity. This model satisfies all conditions of Theorem 6.1, except that the interaction takes the value $+\infty$. This implies that the finite energy condition does not hold as it stands. However, it is easily seen that there are still enough admissible configurations to satisfy all needs of the Burton-Keane theorem and our other applications of the finite energy property. This leads us to the following corollary.

Corollary 6.3 *For the hard-core lattice gas on* $\mathbf{Z}^2$ *at any activity* $\lambda > 0$ *there exist at most two extremal Gibbs measures.*

References

[1] Aizenman, M. (1980) Translation invariance and instability of phase coexistence in the two-dimensional Ising system, *Commun. Math. Phys.* **73**, 83–94.

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

[3] 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.

[4] Fukuda, H. (1995) The two dimensional Ising model and Gibbs states, an approach with percolation method, Master thesis, Osaka University (in Japanese).

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

[6] Georgii, H.-O., Häggström, O. and Maes, C. (1999) The random geometry of equilibrium phases, in: Domb and J.L. Lebowitz (eds.), *Critical phenomena*, Academic Press.

[7] Grimmett, G.R. (1999) *Percolation*, 2. edition, Springer, New York.

[8] 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.

[9] Kesten, H. (1982) *Percolation theory for mathematicians.* Boston MA, Birkhäuser.

[10] 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.

[11] Russo, L. (1979) The infinite cluster method in the two-dimensional Ising model, *Commun. Math. Phys.* **67**, 251–266.