Representation and poly-time approximation for pressure of $\mathbb{Z}^2$ lattice models in the non-uniqueness region

Stefan Adams · **Raimundo Briceño** ·
Brian Marcus · **Ronnie Pavlov**

Abstract We develop a new pressure representation theorem for nearest-neighbour Gibbs interactions and apply this to obtain the existence of efficient algorithms for approximating the pressure in the 2-dimensional ferromagnetic Potts, multi-type Widom-Rowlinson and hard-core models. For Potts, our results apply to every inverse temperature but the critical. For Widom-Rowlinson and hard-core, they apply to certain subsets of both the subcritical and supercritical regions. The main novelty of our work is in the latter.

Contents

Stefan Adams
Mathematics Institute
University of Warwick
Zeeman Building
Coventry CV4 7AL
United Kingdom
E-mail: s.adams@warwick.ac.uk

Brian Marcus
Department of Mathematics
The University of British Columbia
1984 Mathematics Road
Vancouver, B.C., V6T 1Z2
Canada
E-mail: marcus@math.ubc.ca

Raimundo Briceño
Department of Mathematics
The University of British Columbia
1984 Mathematics Road
Vancouver, B.C., V6T 1Z2
Canada
E-mail: raimundo@math.ubc.ca

Ronnie Pavlov
Department of Mathematics
University of Denver
2280 S. Vine St.
Denver, CO 80208
USA
E-mail: rpavlov@du.edu

1 Introduction

The *pressure* of an interaction is a crucial quantity studied in statistical mechanics and dynamical systems. In the former, it coincides with the *specific Gibbs free energy* of a statistical mechanical system (e.g. [16, Part III] and [35, Chapter 3-4]). In the latter, it is a generalization of *topological entropy* and has many applications in a wide variety of classes of dynamical systems, ranging from symbolic to smooth systems (e.g. [8, 23, 40]).

In this paper, we continue the development in [15, 29, 9] of representing pressure with a simplified expression and using this to prove the existence of efficient algorithms for approximating pressure.

We consider *nearest-neighbour (n.n.) real-valued interactions* Φ on $\mathbb{Z}^d$, i.e. interactions defined only on configurations on single sites and pairs of adjacent sites. Since pressure is normally defined for stationary interactions, we assume that our interactions are stationary here. Also, we allow the possibility of forbidden configurations $\mathscr{E}$ on pairs of adjacent sites, and so the space of feasible configurations on $\mathbb{Z}^d$ may be constrained. In the dynamical systems literature, the space of such feasible configurations is known as a *nearest-neighbour shift of finite type (n.n. SFT)*, that here we denote $\Omega(\mathscr{E})$ (see Section 3.1).

A *specification* π for a n.n. interaction Φ is a uniquely determined collection of Borel probability measures π_Λ^ξ given in an explicit form in terms of Φ, for configurations on finite subsets Λ of $\mathbb{Z}^d$ and feasible configurations ξ on the boundary of Λ.

A *Gibbs measure* μ for a n.n. interaction Φ is a Borel probability measure on $\Omega(\mathscr{E})$, whose conditional probability distributions on any such Λ agree with the specification for Φ for all boundary conditions ξ of positive μ-measure.

Gibbs measures exist for all n.n. interactions (and, indeed, for much more general interactions), but a given n.n. interaction may have more than one Gibbs measure. In many cases, including the ones of most interest to us here, there is a n.n. interaction Φ which gives rise to a parameterized family of interactions $\{\zeta\Phi\}_{\zeta>0}$, and uniqueness of Gibbs measures holds for sufficiently small ζ (the so-called *subcritical region*) and uniqueness fails for sufficiently large ζ (the so-called *supercritical region*).

Given a n.n. interaction Φ on a n.n. SFT $\Omega(\mathscr{E})$, we can associate an *energy* to any feasible configuration on a finite subset Λ of $\mathbb{Z}^d$. The *partition function* Z_Λ^Φ of Φ on Λ corresponds to the sum over all feasible configurations on Λ of a function (namely, e^{-x}) of their corresponding energy, and the pressure $\mathrm{P}(\Phi)$ is defined as the asymptotic exponential growth rate of the partition function $\mathrm{Z}_{\mathrm{B}_n}^\Phi$ on an increasing sequence of boxes B_n which exhausts $\mathbb{Z}^d$, as $n \to \infty$. Note that $\mathrm{P}(\Phi)$ implicitly depends on $\Omega(\mathscr{E})$.

When $d = 1$, there is a closed-form expression for $\mathrm{P}(\Phi)$ in terms of the largest eigenvalue of an adjacency matrix formed from Φ (see [28, p. 99]). In contrast, when $d \geq 2$, there are very few n.n. interactions Φ for which $\mathrm{P}(\Phi)$ is known exactly.

There is much work in the literature on numerical approximations of $\mathrm{P}(\Phi)$, both for somewhat general Φ and somewhat specific Φ (see [4, 14]). In our paper, we take a theoretical computer science point of view (see [26]): an algorithm for computing a real number r is said to be *poly-time* if for every $N \in \mathbb{N}$, the algorithm outputs an approximation r_N to r, which is guaranteed to be accurate within $\frac{1}{N}$ and takes time at most polynomial in N to compute. In that case, we say that r is *poly-time computable*.

One of our goals is to prove the existence of poly-time algorithms for $\mathrm{P}(\Phi)$ under certain assumptions on Φ and $\Omega(\mathscr{E})$. While one might expect such algorithms to exist for most Φ and $\Omega(\mathscr{E})$ of practical interest, there exist $\Omega(\mathscr{E})$ for which even $\mathrm{P}(0)$ (which corresponds to the topological entropy of $\Omega(\mathscr{E})$, when the n.n. interaction is $\Phi \equiv 0$) is not poly-time computable and some for which $\mathrm{P}(0)$ is not computable at any rate (see [22]). However, the closed-form expression when $d = 1$ mentioned above, always gives a poly-time algorithm in that case.

We follow an approach initiated by Gamarnik and Katz [15], and further developed by two of the authors [29] of the present paper. The basic idea is motivated by the *variational principle* [23, Section 4.4], which asserts that $\mathrm{P}(\Phi)$ is the supremum over all stationary Borel probability measures μ on $\Omega(\mathscr{E})$ of the sum of two quantities: one quantity is the *measure-theoretic entropy* $h(\mu)$ of μ and the other quantity is the integral, with respect to μ, of a simple explicit function $A_\Phi : \Omega(\mathscr{E}) \to \mathbb{R}$, determined by Φ. The entropy $h(\mu)$ can be expressed as the integral, also with respect to μ, of a function known as the *information function* I_μ, i.e $h(\mu) = \int I_\mu d\mu$. The supremum is always achieved by a Gibbs measure μ for Φ, and so for such μ, we can write $\mathrm{P}(\Phi) = \int (I_\mu + A_\Phi) d\mu$.

The idea of [15] was to represent $\mathrm{P}(\Phi)$ as the integral of the same integrand, but with respect to a simpler measure ν, i.e. $\mathrm{P}(\Phi) = \int (I_\mu + A_\Phi) d\nu$. This is what we call a *pressure representation* and requires some assumptions on μ, ν and $\Omega(\mathscr{E})$.

A pressure representation becomes especially useful for approximating $P(\Phi)$ in the case that ν is a periodic point measure, i.e. a measure which assigns equal weight to each distinct translation of a given periodic configuration (this was the only case considered in [15]). Then $\int (I_\mu + A_\Phi) d\nu$ becomes a finite sum. The terms in this sum corresponding to A_Φ are easy to compute. In this way, the problem of approximating $P(\Phi)$ (and therefore proving that $P(\Phi)$ is poly-time computable) reduces to approximating I_μ on a single periodic configuration and its translates.

The pressure representation theorems in [15] and [29], as well as in our paper (see Theorem 6.3), work in all dimensions d. Among other conditions, these results require conditions on $\Omega(\mathscr{E})$ and a convergence condition for certain sequences of finite volume half-plane measures (different convergence conditions in the different results). In the case $d = 2$, if the convergence holds at exponential rate, then one obtains a poly-time algorithm for approximating $P(\Phi)$ (see Theorem 9.1). For $d > 2$, one can deduce an algorithm for approximating $P(\Phi)$ with sub-exponential but not polynomial rate.

In [15] and [29], the convergence condition is given in terms of the information function I_μ of a stationary Gibbs measure μ for the interaction. In our paper, the condition is given in terms of a closely related function $\hat{I}_\pi$, which depends only on the specification π of the interaction (see Section 6.2), in contrast with [15] and [29]. This is natural, since the pressure depends only on the interaction and not on any particular Gibbs measure μ.

In [15], the convergence condition is *strong spatial mixing* of a Gibbs measure μ for the n.n. interaction Φ. This condition is known to imply that there is a unique Gibbs measure for Φ and thus can be applied only in the uniqueness (subcritical) region of a given model. The convergence conditions in [29] are weaker but also apply primarily to this region. However, in our paper, since our convergence condition depends only on the interaction, one might expect that the pressure representation and approximation results can apply in the non-uniqueness region as well. Indeed, they do. As illustrations, we apply these results to explicit subcritical and supercritical sub-regions of the 2-dimensional *(ferromagnetic) Potts*, *(multi-type) Widom-Rowlinson* and *hard-core* models. In particular, for the pressure approximation results for these models, we establish the required exponential convergence conditions. However, we believe that our results are applicable to a much broader class of models, in particular satisfying weaker conditions on $\Omega(\mathscr{E})$ (e.g. the *topological strong spatial mixing property*, introduced in [9]). We remark that the strong spatial mixing condition of [15] is a much stronger version of our condition, and so in this sense our results generalize some results of that paper (in particular, for the hard-core model on $\mathbb{Z}^2$).

In the case of the 2-dimensional ferromagnetic Potts model, we obtain a pressure representation and efficient pressure approximation for all $\beta \neq \beta_c(q)$, where q is the number of colours, β is the inverse temperature and $\beta_c(q) = \log(1 + \sqrt{q})$ is the *critical value* which separates the uniqueness and non-uniqueness regions. Our proof in the non-uniqueness region generalizes a result from [11] for $q = 2$ (i.e. the *Ising model*) and we closely follow their proof, which relies heavily on a coupling with the *bond random-cluster model* and planar duality. For the uniqueness region, our result follows from [3]. (See Corollary 2, part 1.)

For the Widom-Rowlinson and hard-core models, our results are not as complete as in the Potts case, since the subcritical and supercritical regions for these two models haven't been completely determined, in contrast with the Potts model. We also expect our results can be improved, because they only apply to proper subsets of the currently known uniqueness/non-uniqueness regions.

For the Widom-Rowlinson model, in the supercritical region, we use a variation of the disagreement percolation technique introduced in [7], combined with the connection between the Widom-Rowlinson model and the *site random-cluster model*. In the subcritical region, we apply directly the results in [7]. (See Corollary 2, part 2.)

For the hard-core model, in the supercritical region, we combine the coupling in [7] and a Peierls argument used by Dobrushin (see [13]). In the subcritical region, we use a recent result on strong spatial mixing for the hard-core model in $\mathbb{Z}^2$. (See Corollary 2, part 3.)

For the Potts model, we also extend the pressure representation, by a continuity argument, to give an expression for the pressure at criticality. It is of interest that there is an exact, explicit, but non-rigorous, formula for the pressure at criticality due to Baxter [5]. So, our rigorously obtained expression should agree with that formula, though we do not know how to prove this statement. It seems that Baxter's explicit expression gives a poly-time approximation algorithm, but we cannot justify that our expression is poly-time computable.

We remark that the finite volume half-plane measures mentioned above typically are constant on their bottom boundaries and thus are related to *wetting models* (see [34, 38]). Our proofs are related with such models where the interaction with the hard-wall is the same as the bulk interaction.

The remainder of the paper is organized as follows. Since we have drawn heavily on many concepts from many different sources, for the convenience of the reader we have collected a good deal of relevant background material early in the paper. This can be found in Section 2, Section 3, Section 4, Section 5 and Section 7, with the notable exception of Lemma 5.5 in Section 5, there is very little new material in those sections. In Section 2 and Section 3, we review the fundamentals on configuration spaces on $\mathbb{Z}^d$, Gibbs measures and pressure. In Section 4, we review the specific lattice spin systems models to which we apply our main results, and in Section 5 we review the bond and site random-cluster models which are intimately connected with two of our models. Our pressure representation theorem is contained in Section 6. We review spatial mixing and stochastic dominance in Section 7 and use these concepts in Section 8 to help establish exponential convergence results for our models. Finally, in Section 9, we combine our pressure representation theorem and our exponential convergence results in Section 8 to obtain pressure representations and poly-time algorithms for our models.

2 Definitions and preliminaries

2.1 Hypercubic lattice $\mathbb{Z}^d$

Given $d \in \mathbb{N}$, we consider the *d-dimensional hypercubic lattice* $\mathbb{Z}^d$, which can be regarded as a countable graph with regular degree $2d$, where $V(\mathbb{Z}^d) = \mathbb{Z}^d$ is the set of sites and $E(\mathbb{Z}^d) = \left\{ \{x,y\} : x,y \in \mathbb{Z}^d, \|x-y\| = 1 \right\}$ is the set of bonds, with $\|x\| = \sum_{i=1}^{d} |x_i|$ the 1-norm. We will mainly focus our attention on the case $d = 2$.

Two sites $x,y \in \mathbb{Z}^d$ are *adjacent* if $\{x,y\} \in E(\mathbb{Z}^d)$ and we will denote this by $x \sim y$. All subsets of sites in $\mathbb{Z}^d$ will be denoted with uppercase Greek letters (e.g. Λ, Δ, Θ, etc.). Whenever a finite set Δ is contained in an infinite set Λ, we denote this by $\Delta \Subset \Lambda$. The *(outer) boundary* of $\Lambda \subseteq \mathbb{Z}^d$ is the set $\partial \Lambda$ of $x \in \mathbb{Z}^d \setminus \Lambda$ which are adjacent to some element of Λ, i.e. $\partial \Lambda := \{x \in \Lambda^c : \mathrm{dist}(\{x\}, \Lambda) = 1\}$, where $\mathrm{dist}(\Lambda_1, \Lambda_2) = \min_{x \in \Lambda_1, y \in \Lambda_2} \|x - y\|$, for $\Lambda_1, \Lambda_2 \subseteq \mathbb{Z}^d$. We also write the *closure of* Λ as $\overline{\Lambda} := \Lambda \cup \partial \Lambda$. On the other hand, the *inner boundary* of $\Lambda \subseteq \mathbb{Z}^d$ is the set $\underline{\partial}\Lambda := \partial \Lambda^c$ of $x \in \Lambda$ which are adjacent to some element of Λ^c. When denoting subsets of $\mathbb{Z}^d$ that are singletons, brackets will be usually omitted, e.g. $\mathrm{dist}(x, \Lambda)$ will be regarded to be the same as $\mathrm{dist}(\{x\}, \Lambda)$.

A *path* $\mathrm{T} \Subset \mathbb{Z}^d$ will be any sequence of distinct sites $x_1, \ldots, x_n$ such that $x_i \sim x_{i+1}$, for all $1 \leq i < n$. Similarly, a *circuit* $\mathrm{C} \Subset \mathbb{Z}^d$ will be any path $x_1, \ldots, x_n$ with $n \geq 4$ such that, in addition, $x_n \sim x_1$. We will say that the circuit is *simple* if $x_i \sim x_j$ iff $|i - j| = 1$ or $\{i, j\} = \{1, n\}$ (in particular, $x_1, \ldots, x_n$ are all distinct). For $\Delta, \Theta \subseteq \mathbb{Z}^d$, a *path from* Δ *to* Θ is a path T whose first site is in Δ and whose last site is in Θ. A set $\Lambda \subseteq \mathbb{Z}^d$ is said to be *connected* if for every $x, y \in \Lambda$, there is a path T from x to y contained in Λ (i.e. $\mathrm{T} \subseteq \Lambda$). A set $\Lambda \Subset \mathbb{Z}^2$ is said to be *simply lattice-connected* if Λ and Λ^c are both connected.

In $\mathbb{Z}^d$ we can also define an alternative notion of adjacency and therefore, an alternative notion of boundary, inner boundary, closure, path, connectedness, etc., by replacing the 1-norm $\| \cdot \|$ with the ∞-norm $\| \cdot \|_\infty$, defined as $\|x\|_\infty = \max_{i=1,\ldots,d} |x_i|$, for $x \in \mathbb{Z}^d$. When referring to these notions with respect to the ∞-norm, we will always add a $\star$ superscript and talk about $\star$-adjacency $x \overset{\star}{\sim} y$, $\star$-boundary $\partial^\star \Lambda$, inner $\star$-boundary $\underline{\partial}^\star \Lambda$, $\star$-closure $\overline{\Lambda}^\star$, $\star$-path, $\star$-connectedness, etc. Notice that two sites x and y are $\star$-adjacent if they are adjacent in a version of the d-dimensional hypercubic lattice $\mathbb{Z}^d$ including in addition diagonal bonds. We will denote this version of the lattice by $\mathbb{Z}^{d,\star}$.

A natural order on $\mathbb{Z}^d$ is the so-called *lexicographic order*, where $y \prec x$ (or $x \succ y$) if and only if $y \neq x$ and, for the smallest i for which $y_i \neq x_i$, y_i is strictly smaller than x_i. We also denote $y \preceq x$ (or $x \succeq y$) if $y \prec x$ or $y = x$. Considering this order, we define the family of sets $S_{y,z} \Subset \mathbb{Z}^d$ as:

$$S_{y,z} := \{x \succeq \mathbf{0} : -y \leq x \leq z\}, \tag{2.1}$$

where $y, z \in \mathbb{Z}^d$ are such that $y, z \geq \mathbf{0}$ (here $\mathbf{0}$ denotes the vector $(0, \ldots, 0) \in \mathbb{Z}^d$ and $\geq$, the coordinate-wise comparison of vectors). In addition, given $n \in \mathbb{N}$, we define the *n-block* as the set $\mathrm{B}_n := [-n, n]^d \cap \mathbb{Z}^d$ and we abbreviate by S_n the set $S_{\mathbf{1}n, \mathbf{1}n} = \mathrm{B}_n \setminus \mathscr{P}$, where $\mathscr{P} := \{x \in \mathbb{Z}^d : x \prec \mathbf{0}\}$ denotes the *(lexicographic) past* of $\mathbb{Z}^d$ and $\mathbf{1}$, the vector $(1, \ldots, 1) \in \mathbb{Z}^d$.

2.2 Configuration spaces

Consider a finite set of *symbols* $\mathscr{A}$ called the *alphabet*. A *configuration* is a map $\theta : \Lambda \to \mathscr{A}$, for some $\emptyset \neq \Lambda \subseteq \mathbb{Z}^d$ (i.e. $\theta \in \mathscr{A}^\Lambda$), which will be usually denoted with lowercase Greek letters θ, τ, υ. The set Λ is called the *shape* of θ, and a configuration will be said to be finite if its shape is finite. For any configuration θ with shape Λ and $\Delta \subseteq \Lambda$, $\theta(\Delta)$ denotes the restriction of θ to Δ, i.e. the *sub-configuration* of θ occupying Δ. We will usually save the Greek letters ξ and η to denote configurations whose shape is the boundary $\partial \Lambda$ of some given set Λ. For Λ_1 and Λ_2 disjoint sets, $\theta \in \mathscr{A}^{\Lambda_1}$ and $\tau \in \mathscr{A}^{\Lambda_2}$, $\theta\tau$ will be the configuration on $\Lambda_1 \sqcup \Lambda_2$ defined by $(\theta\tau)(\Lambda_1) = \theta$ and $(\theta\tau)(\Lambda_2) = \tau$. For $a \in \mathscr{A}$ and $\Lambda \subseteq \mathbb{Z}^d$, a^Λ denotes the configuration of all a's on Λ. A *point* is a configuration with shape $\mathbb{Z}^d$, i.e. an element of $\mathscr{A}^{\mathbb{Z}^d}$, usually denoted with the Greek letter ω.

Given sets $\Lambda_1, \Lambda_2 \subseteq \mathbb{Z}^d$, $\Delta \subseteq \Lambda_1 \cap \Lambda_2$ and a pair of configurations $\theta \in \mathscr{A}^{\Lambda_1}$, $\tau \in \mathscr{A}^{\Lambda_2}$, we define the *set of Δ-disagreement* as:

$$\Sigma_\Delta(\theta, \tau) := \{x \in \Delta : \theta(x) \neq \tau(x)\}, \tag{2.2}$$

i.e. the set of sites in Δ where θ and η differ.

The map $\sigma : \mathbb{Z}^d \times \mathscr{A}^{\mathbb{Z}^d} \to \mathscr{A}^{\mathbb{Z}^d}$ will be the *shift action* on $\mathscr{A}^{\mathbb{Z}^d}$ defined by $(x, \omega) \mapsto \sigma_x(\omega)$, where $x \in \mathbb{Z}^d$ and $\omega \in \mathscr{A}^{\mathbb{Z}^d}$, with $(\sigma_x(\omega))(y) = \omega(x + y)$, for $y \in \mathbb{Z}^d$. We also extend the shift action σ_x to configurations with arbitrary shapes, i.e. given $\theta \in \mathscr{A}^\Lambda$, we define $\sigma_x(\theta) \in \mathscr{A}^{\Lambda - x}$ as the configuration such that $(\sigma_x(\theta))(y) = \theta(x + y)$, for $y \in \Lambda - x$.

Given a point $\omega \in \mathscr{A}^{\mathbb{Z}^d}$, we define its *orbit* as the set $\mathrm{O}(\omega) := \{\sigma_x(\omega)\}_{x \in \mathbb{Z}^d}$. We will say that a point ω is *periodic* if $|\mathrm{O}(\omega)| < \infty$.

2.3 Borel probability measures

Given a configuration $\theta \in \mathscr{A}^\Lambda$, we define the *cylinder set* $[\theta]_\Lambda := \{\omega \in \mathscr{A}^{\mathbb{Z}^d} : \omega(\Lambda) = \theta\}$ (or just $[\theta]$, if Λ is understood). We denote by $\mathscr{F}_\Lambda$ the σ-algebra generated by all the cylinder sets with shape Λ and set $\mathscr{F} := \mathscr{F}_{\mathbb{Z}^d}$.

A *Borel probability measure* μ on $\mathscr{F}$ is a measure determined by its values on cylinder sets of finite configurations such that $\mu(\mathscr{A}^{\mathbb{Z}^d}) = 1$. Given a cylinder set $[\theta]$, we will just write $\mu(\theta)$ for the value of $\mu([\theta])$. The *support* of such a measure μ is defined as:

$$\mathrm{supp}(\mu) := \left\{\omega \in \mathscr{A}^{\mathbb{Z}^d} : \mu(\omega(\Lambda)) > 0, \text{ for all } \Lambda \Subset \mathbb{Z}^d\right\}. \tag{2.3}$$

Given $\Delta \subseteq \Lambda \subseteq \mathbb{Z}^d$ and a measure μ on $\mathscr{F}_\Lambda$, we denote by $\mu|_\Delta$ the restriction (or projection or marginalization) of μ to $\mathscr{F}_\Delta$.

A measure μ is *shift-invariant* (or *stationary*) if $\mu(\sigma_x(A)) = \mu(A)$, for all measurable sets $A \in \mathscr{F}$ and $x \in \mathbb{Z}^d$. Given any point $\omega \in \mathscr{A}^{\mathbb{Z}^d}$ and $A \in \mathscr{F}$, we define the

delta-measure supported on ω as the measure:

$$\delta_\omega(A) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

If ω is a periodic point with orbit $O(\omega) = \{\omega_1, \ldots, \omega_k\}$, we define ν^ω to be the shift-invariant Borel probability measure supported on $O(\omega)$ given by:

$$\nu^\omega := \frac{1}{k} \left(\delta_{\omega_1} + \cdots + \delta_{\omega_k} \right). \tag{2.5}$$

2.4 Markov random fields

Definition 2.1. *Given $\Lambda \subseteq \mathbb{Z}^d$, a probability measure ρ on $\mathscr{A}^\Lambda$ is a* Markov random field (Λ-MRF) *if, for any subset $\Theta \Subset \Lambda$, any $\theta \in \mathscr{A}^\Theta$, any $\Delta \Subset \Lambda$ s.t. $\partial\Theta \cap \Lambda \subseteq \Delta \subseteq \Lambda \setminus \Theta$, and any $\tau \in \mathscr{A}^\Delta$ with $\rho(\tau) > 0$, it is the case that:*

$$\rho(\theta|\tau) = \rho(\theta|\tau(\partial\Theta \cap \Lambda)). \tag{2.6}$$

In other words, an MRF is a measure where every finite configuration conditioned to its boundary is independent of the configuration on the complement.

3 Specifications, Gibbs measures and pressure

3.1 Gibbs specifications

Fix a dimension $d \in \mathbb{N}$ and let $\mathscr{E} = (\mathscr{E}_1, \ldots, \mathscr{E}_d)$ be a *set of constraints* such that $\mathscr{E}_i \subseteq \mathscr{A}^2$, for $i = 1, \ldots, d$. Given any set $\Lambda \subseteq \mathbb{Z}^d$ and a configuration $\theta \in \mathscr{A}^\Lambda$, we say that θ is *feasible* for $\mathscr{E}$ if for every $x \in \Lambda$ such that $\{x, x + e_i\} \subseteq \Lambda$, we have that $(\theta(x), \theta(x + e_i)) \notin \mathscr{E}_i$, where $e_1, \ldots, e_d$ is the canonical basis. The *nearest-neighbour shift of finite type (n.n. SFT)* $\Omega(\mathscr{E})$ induced by $\mathscr{E}$, is the set of points:

$$\Omega(\mathscr{E}) := \left\{ \omega \in \mathscr{A}^{\mathbb{Z}^d} : \omega \text{ is feasible} \right\}. \tag{3.1}$$

We will always assume that $\Omega(\mathscr{E}) \neq \emptyset$.

In the symbolic dynamics literature, a feasible configuration on a set Λ is called *locally admissible*, and is called *globally admissible* if it also extends to a point of $\Omega(\mathscr{E})$.

Notice that $\Omega(\mathscr{E})$ is always a shift-invariant set, i.e. $\sigma_x(\Omega(\mathscr{E})) = \Omega(\mathscr{E})$, for all $x \in \mathbb{Z}^d$. Given a n.n. SFT $\Omega(\mathscr{E})$, $\mathscr{M}_1(\Omega(\mathscr{E}))$ denotes the set of Borel probability measures whose support $\text{supp}(\mu)$ is contained in $\Omega(\mathscr{E})$ and $\mathscr{M}_{1,\sigma}(\Omega(\mathscr{E})) \subseteq \mathscr{M}_1(\Omega(\mathscr{E}))$, the corresponding subset of shift-invariant Borel probability measures. Given a configuration $\theta \in \mathscr{A}^\Lambda$, $[\theta]_\Lambda^{\Omega(\mathscr{E})}$ will denote the set $[\theta]_\Lambda \cap \Omega(\mathscr{E})$ (or just $[\theta]^{\Omega(\mathscr{E})}$ if Λ is understood).

Definition 3.1. *A nearest-neighbour (n.n.) interaction for a set of constraints $\mathscr{E}$ is a real-valued shift-invariant function Φ from the set of configurations on sites x and feasible configurations on bonds $\{x, x + e_i\}$ to $\mathbb{R}$, for $x \in \mathbb{Z}^d$ and $i = 1, \ldots, d$. Here, shift-invariance means that $\Phi(\sigma_x(\theta)) = \Phi(\theta)$ for configurations θ on sites and bonds, and for all $x \in \mathbb{Z}^d$.*

Often in the literature a n.n. interaction is not required to be shift-invariant. Our assumption of shift-invariance on a n.n. interaction fits naturally with the shift-invariance of a n.n. SFT. Clearly, a n.n. interaction is defined by only finitely many numbers, namely the values of the interaction on configurations on $\{\mathbf{0}\}$ and bonds $\{\mathbf{0}, e_i\}$, for $i = 1, \ldots, d$.

We can view an interaction Φ as implicitly determining the constraints $\mathscr{E}$, and hence $\Omega(\mathscr{E})$, by the absence of $\mathscr{E}$ from the domain of Φ. Some authors incorporate the constraints by allowing the interaction to take the value $+\infty$.

Definition 3.2. *Given a n.n. interaction Φ for a set of constraints $\mathscr{E}$ and a set $\Lambda \Subset \mathbb{Z}^d$, we define the* energy function $\mathrm{E}_\Lambda^\Phi : \mathscr{A}^\Lambda \to \mathbb{R}$ *as:*

$$\mathrm{E}_\Lambda^\Phi(\theta) := \sum_{x \in \Lambda} \Phi(\theta(x)) + \sum_{i=1}^d \sum_{\{x, x+e_i\} \subseteq \Lambda} \Phi(\theta(\{x, x + e_i\})), \tag{3.2}$$

where θ is any feasible configuration in $\mathscr{A}^\Lambda$. We define the partition function *of Λ as:*

$$Z_\Lambda^\Phi := \sum_{\theta \text{ feasible}} \exp\left(-\mathrm{E}_\Lambda^\Phi(\theta)\right), \tag{3.3}$$

and the following boundary-free *probability measure on $\mathscr{A}^\Lambda$:*

$$\pi_\Lambda^{(f)}(\theta) := \begin{cases} \frac{1}{Z_\Lambda^\Phi} \exp\left(-\mathrm{E}_\Lambda^\Phi(\theta)\right) & \text{if } \theta \text{ is feasible}, \\ 0 & \text{otherwise}. \end{cases} \tag{3.4}$$

Analogously, for an arbitrary $\omega \in \Omega(\mathscr{E})$, we can take $\xi = \omega(\partial \Lambda)$ and consider:

$$Z_\Lambda^{\Phi, \xi} := \sum_{\theta: \theta \xi \text{ feasible}} \exp\left(-\mathrm{E}_\Lambda^\Phi(\theta \xi)\right), \tag{3.5}$$

and then define the ξ-boundary probability measure on $\mathscr{A}^\Lambda$:

$$\pi_\Lambda^\xi(\theta) := \begin{cases} \frac{1}{Z_\Lambda^{\Phi, \xi}} \exp\left(-\mathrm{E}_\Lambda^\Phi(\theta \xi)\right) & \text{if } \theta \xi \text{ is feasible}, \\ 0 & \text{otherwise}. \end{cases} \tag{3.6}$$

The collection $\pi = \{\pi_\Lambda^\xi\}_{\Lambda, \xi}$ is called a $\mathbb{Z}^d$ Gibbs specification for the n.n. interaction Φ. For $\Delta \subseteq \Lambda$ and $\tau \in \mathscr{A}^\Delta$, we marginalize as follows:

$$\pi_\Lambda^\xi(\tau) = \sum_{\theta \in \mathscr{A}^\Lambda : \theta(\Delta) = \tau} \pi_\Lambda^\xi(\theta). \tag{3.7}$$

Notice that each π_Λ^ξ is an MRF on $\mathscr{A}^\Lambda$. In addition, a Gibbs specification π as defined above is always stationary, in the sense that $\pi_{\Lambda-x}^{\sigma_x(\xi)}(\sigma_x(A)) = \pi_\Lambda^\xi(A)$, for every $A \subseteq \mathscr{A}^\Lambda$. We will usually think of the set of restrictions $\mathscr{E}$ implicit when considering a n.n. interaction Φ. Given a point $\omega \in \Omega(\mathscr{E})$, we will abbreviate:

$$\pi_\Lambda^\omega(\cdot) := \pi_\Lambda^{\omega(\partial\Lambda)}(\cdot). \tag{3.8}$$

3.2 Gibbs measures

Definition 3.3. *A* nearest-neighbour (n.n.) Gibbs measure *for a n.n. interaction Φ is a measure $\mu \in \mathscr{M}_1(\Omega(\mathscr{E}))$ such that for any $\Lambda \Subset \mathbb{Z}^d$ and $\omega \in \mathscr{A}^{\mathbb{Z}^d}$ with $\mu(\omega(\partial\Lambda)) > 0$, we have that $Z_\Lambda^{\Phi,\omega(\partial\Lambda)} > 0$ and:*

$$\mu(\theta|\mathscr{F}_{\Lambda^c})(\omega) = \pi_\Lambda^\omega(\theta) \ \mu\text{-a.s.}, \tag{3.9}$$

for $\theta \in \mathscr{A}^\Lambda$, where $\{\pi_\Lambda^\xi\}_{\Lambda,\xi}$ is the stationary $\mathbb{Z}^d$ Gibbs specification for Φ.

While our interactions and specifications are assumed to be shift-invariant, a Gibbs measure for such an interaction may or may not be stationary. The definition of n.n. Gibbs measure, shows that such a measure is an MRF. The definition is stated only for cylinder events $[\theta]$ in Λ, but this is equivalent to the usual definition with general events $A \in \mathscr{F}$ instead.

Every n.n. interaction Φ has at least one (stationary) n.n. Gibbs measure (special case of a general result in [35, Theorem 3.7 and Theorem 4.2]). For a single Φ, multiple Gibbs measures can exist. This phenomenon is usually called a *phase transition*.

3.3 Pressure

Now we proceed to define the pressure of a n.n. interaction Φ.

Definition 3.4. *Given a n.n. interaction Φ for a set of restrictions $\mathscr{E}$, the* pressure of Φ *is defined as:*

$$\mathrm{P}(\Phi) := \lim_{n\to\infty} \frac{1}{|\mathrm{B}_n|} \log Z_{\mathrm{B}_n}^\Phi. \tag{3.10}$$

Given $n \in \mathbb{N}$, we can also define an analogous version $\hat{Z}_{\mathrm{B}_n}^\Phi$ of the partition function $Z_{\mathrm{B}_n}^\Phi$, but over globally admissible configurations:

$$\hat{Z}_{\mathrm{B}_n}^\Phi := \sum_{\theta \in \mathscr{A}^{\mathrm{B}_n}:[\theta]^{\Omega(\mathscr{E})}\neq\emptyset} \exp\left(-\mathrm{E}_{\mathrm{B}_n}^\Phi(\theta)\right). \tag{3.11}$$

Notice that $\hat{Z}_{\mathrm{B}_n}^\Phi \leq Z_{\mathrm{B}_n}^\Phi$. The following result states that in the normalized limit, both quantities coincide.

Theorem 3.1 ([35, Theorem 3.4], see also [14, Theorem 2.5])**.** *Given a n.n. interaction Φ for a set of restrictions $\mathscr{E}$:*

$$\mathrm{P}(\Phi) = \lim_{n \to \infty} \frac{1}{|\mathrm{B}_n|} \log \hat{Z}_{\mathrm{B}_n}^{\Phi}. \qquad (3.12)$$

The pressure is the main quantity of interest in this paper. Our goals are to find simple representations of pressure in terms of very special configurations and use this to develop efficient (in principle) algorithms to approximate the pressure.

4 Main models: Potts, Widom-Rowlinson and hard-core

In this section we introduce the three main families of lattice models studied in this paper. The first one will be the Potts model, which can be regarded as a generalization of the Ising model by considering more than two types of particles. The second one, the Widom-Rowlinson model, is also a multi-type particle system but with hard-core exclusion between particles of different type. The third one is the classical hard-core model.

4.1 The (ferromagnetic) Potts model

Given $d, q \in \mathbb{N}$ and $\beta > 0$, the $\mathbb{Z}^d$ *(ferromagnetic) Potts model with q types and inverse temperature* β is defined over the alphabet $\mathscr{A}_q = \{1, \ldots, q\}$ and given by the n.n. interaction:

$$\Phi_\beta(\theta) = \begin{cases} -\beta & \text{if } \theta(x) = \theta(x + e_i), \\ 0 & \text{if } \theta(x) \neq \theta(x + e_i), \end{cases} \qquad (4.1)$$

for $\theta \in \mathscr{A}_q^{\{x, x+e_i\}}$, $x \in \mathbb{Z}^d$, $i = 1, \ldots, d$, where the constraints $\mathscr{E}_i$ are empty. The specification $\pi_\beta^{\mathrm{FP}} = \{\pi_{\beta,\Lambda}^{\xi}\}_{\Lambda,\xi}$ induced by Φ_β defines the (ferromagnetic) Potts model, where neighbouring sites preferably align to each other with the same type or "colour" from the alphabet $\mathscr{A}_q$.

A measure $\mu \in \mathscr{M}_1(\mathscr{A}_q^{\mathbb{Z}^d})$ is called a *Potts Gibbs measure* for q types and inverse temperature $\beta > 0$ if it is a n.n. Gibbs measure for the specification π_β^{FP} above.

Theorem 4.1 ([6])**.** *For the $\mathbb{Z}^2$ (ferromagnetic) Potts model with q types and inverse temperature β, there exists a critical inverse temperature $\beta_{\mathrm{c}}(q) := \log(1 + \sqrt{q})$ such that uniqueness of Gibbs measures holds for $\beta < \beta_{\mathrm{c}}(q)$ and for $\beta > \beta_{\mathrm{c}}(q)$ there is a phase transition.*

4.2 The (multi-type) Widom-Rowlinson model

Given $d, q \in \mathbb{N}$ and $\lambda > 0$, the $\mathbb{Z}^d$ *Widom-Rowlinson model with q types and activity* λ is defined over the alphabet $\mathscr{B}_q = \{0, 1, \ldots, q\}$, and given by the set of constraints

$\mathscr{E} = (\mathscr{E}_1, \ldots, \mathscr{E}_d)$, where $\mathscr{E}_i = \{\theta \in (\mathscr{B}_q \setminus \{0\})^2 : \theta(1) \neq \theta(2)\}$, for all $i = 1, \ldots, d$, and by the n.n. interaction for $\mathscr{E}$ over configurations on sites:

$$\Phi_\lambda(\theta) = \begin{cases} -\log(\lambda) & \text{if } \theta \in \{1, \ldots, q\}, \\ 0 & \text{if } \theta = 0, \end{cases} \tag{4.2}$$

where $\theta \in \mathscr{B}_q^{\{x\}}$ and $x \in \mathbb{Z}^d$. The specification $\pi_\lambda^{\mathrm{WR}} = \{\pi_{\lambda,\Lambda}^\xi\}_{\Lambda,\xi}$ induced by Φ_λ defines the (multi-type) Widom-Rowlinson model, where neighbouring sites are forced to align to each other with the same type or "colour" from the alphabet $\mathscr{B}_q$ or with 0.

A measure $\mu \in \mathscr{M}_1(\mathscr{B}_q^{\mathbb{Z}^d})$ is called a *Widom-Rowlinson Gibbs measure* for q types and activity $\lambda > 0$ if it is a n.n. Gibbs measure for the specification $\pi_\lambda^{\mathrm{WR}}$ above.

Theorem 4.2 ([36], see also [19])**.** *For the $\mathbb{Z}^2$ Widom-Rowlinson model with q types and activity λ, uniqueness of Gibbs measures holds for sufficiently small λ and there is a phase transition for sufficiently large λ.*

4.3 The hard-core lattice gas model

Given $\gamma > 0$, the $\mathbb{Z}^d$ *hard-core model with activity* γ is defined over the alphabet $\{0, 1\}$, and given by the set of constraints $\mathscr{E}$, where $\mathscr{E}_i = \{(1, 1)\}$, for all $i = 1, \ldots, d$, and the the n.n. interaction for $\mathscr{E}$ over configurations on sites:

$$\Phi_\gamma(\theta) = \begin{cases} -\log(\gamma) & \text{if } \theta = 1, \\ 0 & \text{if } \theta = 0, \end{cases} \tag{4.3}$$

for $\theta \in \{0, 1\}^{\{x\}}$, $x \in \mathbb{Z}^d$. The specification $\pi_\gamma^{\mathrm{HC}} = \{\pi_{\gamma,\Lambda}^\xi\}_{\Lambda,\xi}$ induced by Φ_γ defines the hard-core model, where neighbouring sites cannot be both 1.

A measure $\mu \in \mathscr{M}_1(\{0, 1\}^{\mathbb{Z}^d})$ is called a *hard-core Gibbs measure* for activity $\gamma > 0$ if it is a n.n. Gibbs measure for the specification π_γ^{HC} above.

Theorem 4.3 ([17, Theorem 3.3])**.** *For the $\mathbb{Z}^2$ hard-core model with activity γ, uniqueness of Gibbs measures holds for sufficiently small γ and there is a phase transition for sufficiently large γ.*

For both the Potts and Widom-Rowlinson models we will also distinguish a particular type of particle or colour in the alphabet. W.l.o.g., we can take the type q in $\mathscr{A}_q$ or $\mathscr{B}_q \setminus \{0\}$, respectively. Given this colour, we will denote by ω_q the fixed point $q^{\mathbb{Z}^d}$. For the hard-core model, we will consider the two special points $\omega^{(e)}$ and $\omega^{(o)}$, given by:

$$\omega^{(e)}(x) := \begin{cases} 0 & \text{if } \sum_i x_i \text{ is even}, \\ 1 & \text{if } \sum_i x_i \text{ is odd}, \end{cases} \tag{4.4}$$

and $\omega^{(o)} = \sigma_{e_1}(\omega^{(e)})$.

5 Random-cluster models

The Potts and Widom-Rowlinson models have interpretations in terms of a random-cluster representation. The Potts model is related to a random-cluster model on bonds (via the so-called Edwards-Sokal coupling), while the Widom-Rowlinson is naturally related to a random-cluster model on sites.

Definition 5.1. *A coupling of two probability measures ρ_1 on a finite set X and ρ_2 on a finite set Y, is a probability measure $\mathbb{P}$ on the set $X \times Y$ such that, for any $A \subseteq X$ and $B \subseteq Y$, we have that:*

$$\mathbb{P}(A \times Y) = \rho_1(A) \text{ and } \mathbb{P}(X \times B) = \rho_2(B). \tag{5.1}$$

5.1 The bond random-cluster model and the Potts model

We will make use of the *bond random-cluster model*. One of our main results, Part I of Theorem 8.1, is proven using arguments based on this model. This model is a two parameter family of dependent bond percolation models on a finite graph. We are mainly interested in finite subgraphs of $\mathbb{Z}^2$ and we describe the model with boundary conditions indexed by $i = 0, 1$.

Fix a finite simply lattice-connected set of sites Λ. Let $E^0(\Lambda)$ denote the set of bonds with both endpoints in Λ and $E^1(\Lambda)$ the set of bonds with at least one endpoint in Λ. We speak of a bond e as being *open* if $w(e) = 1$, and as being *closed* if $w(e) = 0$.

Definition 5.2. *Given a finite simply lattice-connected set Λ, and parameters $p \in [0,1]$ and $q > 0$, we define the* free *($i = 0$) and* wired *($i = 1$) bond random-cluster distributions on $E^i(\Lambda)$ ($i = 0, 1$) as the measures $\phi^i_{p,q,\Lambda}$ that to each $w \in \{0,1\}^{E^i(\Lambda)}$ assigns probability proportional to:*

$$\phi^{(i)}_{p,q,\Lambda}(w) \propto \left\{ \prod_{e \in E^i(\Lambda)} p^{w(e)}(1-p)^{1-w(e)} \right\} q^{k^i_\Lambda(w)} = \left(\frac{p}{1-p} \right)^{\#_1(w)} q^{k^i_\Lambda(w)}, \tag{5.2}$$

where $\#_1(w)$ is the number of open bonds in w and $k^0_\Lambda(w)$ and $k^1_\Lambda(w)$ are the number of connected components (including isolated sites) in the graphs $(\Lambda, \{e \in E^0(\Lambda) : w(e) = 1\})$ and $(\mathbb{Z}^2, E^0(\mathbb{Z}^2 \setminus \Lambda) \cup \{e \in E^1(\Lambda) : w(e) = 1\})$, respectively.

Notice that when $q = 1$, we recover the ordinary Bernoulli bond percolation measure $\phi_{p,\Lambda}$, while other choices of q lead to dependence between bonds. For given p and q, one can also define bond random-cluster measures $\phi^{(i)}_{p,q}$ on $\mathbb{Z}^2$ as a limit of finite volume measures $\phi^{(i)}_{p,q,\Lambda}$ ($i = 0, 1$).

Theorem 5.1 ([17, Lemma 6.8]). *For $p \in [0,1]$ and $q \in \mathbb{N}$, the limiting measures:*

$$\phi^{(i)}_{p,q} = \lim_{n \to \infty} \phi^{(i)}_{p,q,\Lambda_n}, \qquad i \in \{0,1\}, \tag{5.3}$$

exist and are translation invariant, where $\{\Lambda_n\}_n$ is any increasing sequence of finite simply lattice-connected sets that exhausts $\mathbb{Z}^2$.

General bond random-cluster measures on $\mathbb{Z}^2$ can be defined using an analogue of the DLR condition [20, Definition 4.29]. For $q \geq 1$, there is a value $p_c(q)$ that delimits exactly the transition for existence of an infinite open cluster for these measures. It is known [20, p. 107] that for $q \geq 1$ and $p < p_c(q)$, there is a unique such measure which we denote by $\phi_{p,q}$ (characterized by the nonexistence of infinite open clusters), and that coincides with $\phi_{p,q}^{(0)}$ and $\phi_{p,q}^{(1)}$ in this region. It was recently proven (see [6]) that $p_c(q) = \frac{\sqrt{q}}{1+\sqrt{q}}$, for every $q \geq 1$.

Let $p = 1 - e^{-\beta}$. The *free Edwards-Sokal coupling* $\mathbb{P}_{p,q,\Lambda}^{(0)}$ (see [20]) is a coupling between the boundary-free Potts measure $\pi_{\beta,\Lambda}^{(f)}$ and $\phi_{p,q,\Lambda}^{(0)}$. The *wired Edwards-Sokal coupling* $\mathbb{P}_{p,q,\Lambda}^{(1)}$ is a coupling between $\pi_{\beta,\Lambda}^{\omega_q}$ and $\phi_{p,q,\Lambda}^{(1)}$. Notice that $p_c(q) = 1 - e^{-\beta_c(q)}$.

These couplings are measures on pairs of site configurations and corresponding bond configurations. The projection to site configurations is the boundary-free/ω_q-boundary Potts measure, and the projection to bond configurations is the free/wired bond random-cluster measure, respectively.

Theorem 5.2 ([20, Theorem 1.13]). *Let Λ be a finite simply lattice-connected set, $q \in \mathbb{N}$, and let $p \in [0,1]$ and $\beta > 0$ be such that $p = 1 - e^{-\beta}$. Then:*

1. *For $w \in \{0,1\}^{E^1(\Lambda)}$, the conditional measure $\mathbb{P}_{p,q,\Lambda}^{(1)}\left(\cdot \big| \mathscr{A}_q^\Lambda \times \{w\}\right)$ on $\mathscr{A}_q^\Lambda$ is obtained by putting random colours on entire clusters of w not connected with $\mathbb{Z}^2 \setminus \Lambda$ (of which there are $k_\Lambda^1(w) - 1$) and colour q on the clusters connected with $\mathbb{Z}^2 \setminus \Lambda$. These colours are constant on given clusters, are independent between clusters, and the random ones are uniformly distributed on the set $\mathscr{A}_q$.*

2. *For $\theta \in \mathscr{A}^\Lambda$, the conditional measure $\mathbb{P}_{p,q,\Lambda}^{(1)}\left(\cdot \big| \{\theta\} \times \{0,1\}^{E^1(\Lambda)}\right)$ on $\{0,1\}^{E^1(\Lambda)}$ is obtained as follows. Consider the extended configuration $\hat{\theta} = \theta q^{\partial \Lambda}$ and an arbitrary bond $e = \{x,y\} \in E^1(\Lambda)$. If $\hat{\theta}(x) \neq \hat{\theta}(y)$, we set $w(e) = 0$. If $\hat{\theta}(x) = \hat{\theta}(y)$, we set:*

$$w(e) = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{otherwise,} \end{cases} \tag{5.4}$$

the values of different $w(e)$ being (conditionally) independent random variables.

The couplings can be used to relate probabilities and expectations for the Potts model to corresponding events and expectations in the associated bond random-cluster model. A main example is a relation between the two-point correlation function in the Potts model and the connectivity function in the bond random-cluster model [20, Theorem 1.16].

By considering a displaced version of $\mathbb{Z}^2$, namely $\frac{1}{2}\mathbf{1} + \mathbb{Z}^2$ (the *dual lattice*), we can define a notion of duality for bond configurations w. Notice that every bond $e \in E(\mathbb{Z}^2)$ (if we think of bonds as unitary vertical and horizontal straight segments) is intersected perpendicularly by one and only one *dual bond* $e^* \in E(\frac{1}{2}\mathbf{1} + \mathbb{Z}^2)$, so there is a clear correspondence between $E(\mathbb{Z}^2)$ and $E(\frac{1}{2}\mathbf{1} + \mathbb{Z}^2)$. We are mainly interested in wired bond random-cluster distributions on the set of sites $\tilde{B}_n := [-n+1,n]^2 \cap \mathbb{Z}^2$.

Given $n \in \mathbb{N}$, if we consider the set of bonds $E^1(\tilde{B}_n)$, it is easy to check that there is a correspondence $e \mapsto e^*$ between this set and the set of bonds from $\frac{1}{2}\mathbf{1} + \mathbb{Z}^2$ with both endpoints in $[-n, n]^2 \cap \left(\frac{1}{2}\mathbf{1} + \mathbb{Z}^2\right)$, which can be identified with the set $E^0(B_n)$. Then, given a bond configuration $w \in E^1(\tilde{B}_n)$ we can associate a dual bond configuration $w^* \in E^0(B_n)$ such that $w^*(e^*) = 0$ if and only if $w(e) = 1$.

Considering this, we have the corresponding equality:

Proposition 5.3 ([20], Equation 6.12 and Theorem 6.13]). *Given $n \in \mathbb{N}$, $p \in [0, 1]$ and $q \in \mathbb{N}$:*

$$\phi^{(1)}_{p,q,\tilde{B}_n}(w) = \phi^{(0)}_{p^*,q,B_n}(w^*), \qquad (5.5)$$

for any bond configuration $w \in \{0, 1\}^{E^1(\tilde{B}_n)}$, where $\tilde{B}_n = [-n+1, n]^2 \cap \mathbb{Z}^2$ and $p^* \in [0, 1]$ is the dual value of p, which is given by:

$$\frac{p^*}{1 - p^*} = \frac{q(1 - p)}{p}. \qquad (5.6)$$

The previous duality result can be generalized to more arbitrary shapes and it has also a counterpart from free-to-wired boundary conditions, instead of from wired-to-free.

The unique fixed point of the map $p \mapsto p^*$ defined by (5.6) is $\frac{\sqrt{q}}{1+\sqrt{q}}$ and, as mentioned above, is known to coincide with the critical point $p_c(q)$ for the existence of an infinite open cluster for the bond random-cluster model (see [6], Theorem 1]). It is easy to see that $p > p_c(q)$ iff $p^* < p_c(q)$.

5.2 The site random-cluster model and the Widom-Rowlinson model

In a similar fashion to the bond random-cluster model, we can perturb Bernoulli site percolation, where the probability measure is changed in favour of configurations with many (for $q > 1$) or few (for $q < 1$) connected components. The resulting model is called the *site random-cluster model*.

Definition 5.3. *Given $\Lambda \Subset \mathbb{Z}^2$, and parameters $p \in [0, 1]$ and $q > 0$, the* wired site random-cluster measure $\psi^{(1)}_{p,q,\Lambda}$ *is the probability measure on $\{0, 1\}^{\Lambda}$ which to each $\theta \in \{0, 1\}^{\Lambda}$ assigns probability proportional to:*

$$\psi^{(1)}_{p,q,\Lambda}(\theta) \propto \left\{ \prod_{x \in \Lambda} p^{\theta(x)}(1 - p)^{1 - \theta(x)} \right\} q^{\kappa_{\Lambda}(\theta)} = \lambda^{\#_1(\theta)} q^{\kappa_{\Lambda}(\theta)}, \qquad (5.7)$$

where $\lambda = \frac{p}{1-p}$, $\#_1(\theta)$ is the number of 1's in θ and $\kappa_{\Lambda}(\theta)$ is the number of connected components in $\{x \in \Lambda : \theta(x) = 1\}$ that do not intersect $\underline{\partial}\Lambda$.

The *free site random-cluster measure* $\psi^{(0)}_{p,q,\Lambda}$ is defined as in (5.7) by replacing $\kappa_{\Lambda}(\theta)$ by the total number of connected components in Λ. However, we will not require that measure in this work. In any case, taking $q = 1$ gives the ordinary Bernoulli site percolation $\psi_{p,\Lambda}$, while other choices of q lead to dependence between sites, similarly to the bond random-cluster model.

Proposition 5.4. *Given a set $\Lambda \Subset \mathbb{Z}^2$ and parameters $\lambda > 0$ and $q \in \mathbb{N}$, consider the Widom-Rowlinson with q types distribution and monochromatic boundary condition $\pi_{\lambda,\Lambda}^{\omega_q}$. Now, let $f : \mathscr{B}_q^{\Lambda} \to \{0,1\}^{\Lambda}$ be defined site-wise as:*

$$(f(\theta))(x) = \begin{cases} 0 & \text{if } \theta(x) = 0, \\ 1 & \text{if } \theta(x) \neq 0, \end{cases} \tag{5.8}$$

for $\theta \in \mathscr{B}_q^{\Lambda}$ and $x \in \Lambda$, and let $p = \frac{\lambda}{1+\lambda}$. Then, $f_* \pi_{\lambda,\Lambda}^{\omega_q} = \psi_{p,q,\Lambda}^{(1)}$, where $f_* \pi_{\lambda,\Lambda}^{\omega_q}(\cdot) := \pi_{\lambda,\Lambda}^{\omega_q}(f^{-1}(\cdot))$ denotes the push-forward measure on $\{0,1\}^{\Lambda}$.

The requirement that $\kappa_{\Lambda}(\cdot)$ does not count connected components that intersect the inner boundary of Λ in the site random-cluster model, corresponds to the fact that non 0 sites adjacent to the monochromatic boundary $\omega_q(\partial\Lambda)$ in the Widom-Rowlinson model must have the same colour q.

For $q = 2$, Proposition 5.4 is proven in [21, Lemma 5.1 (ii)], and the proof extends easily for general q. Proposition 5.4 can be regarded as a coupling between $\pi_{\lambda,\Lambda}^{\omega_q}$ and $\psi_{p,q,\Lambda}^{(1)}$, because a push-forward measure can be naturally coupled with the original measure.

It is important to notice that $\psi_{p,q,\Lambda}^{(1)}$ is itself not an MRF: given sites on a simple circuit C, the inside and outside of C are generally not conditionally independent, because knowledge of sites outside C could cause connected components of 1's in C to "amalgamate" into a single component, which would affect the conditional distribution of configurations inside C. The following lemma shows that in certain situations, when conditioning on a circuit C labeled entirely by 1's, this kind of amalgamation does not occur.

Lemma 5.5. *Let $\emptyset \neq \Theta \subseteq \Lambda \Subset \mathbb{Z}^2$ be such that $\Lambda^c \cup \overline{\Theta}^{\star}$ is connected. Take $\Delta := \partial^{\star}\Theta \cap \Lambda$. Consider an event $A \in \mathscr{F}_{\Theta}$ and a configuration $\tau \in \{0,1\}^{\Sigma}$, where $\Sigma \subseteq \Lambda \setminus \overline{\Theta}^{\star}$. Then:*

$$\psi_{p,q,\Lambda}^{(1)}(A|1^{\Delta}\tau) = \psi_{p,q,\Lambda}^{(1)}(A|1^{\Delta}0^{\Lambda \setminus \overline{\Theta}^{\star}}). \tag{5.9}$$

Proof. W.l.o.g., we may assume that A is a cylinder event $[\theta]$ with $\theta \in \{0,1\}^{\Theta}$ (by linearity) and $\Sigma = \Lambda \setminus \overline{\Theta}^{\star}$ (by taking weighted averages).

Now, $\Sigma = \Lambda \setminus \overline{\Theta}^{\star}$ can be written as a disjoint union of $\star$-connected components $\Sigma = K_1 \sqcup \cdots \sqcup K_n$. For every i, $\partial^{\star}K_i \subseteq \Lambda^c \cup \overline{\Theta}^{\star}$ (in fact, $\partial^{\star}K_i \subseteq \Lambda^c \cup \Delta$). Since $\Lambda^c \cup \overline{\Theta}^{\star}$ is connected and Λ is finite, for every site in $\partial^{\star}K_i$ there is a path to infinity that does not intersect K_i.

Then, by application of a result of Kesten (see [24, Lemma 2.23]), $\partial^{\star}K_i$ is connected, for every i. In addition, we have that $\Lambda = \Theta \sqcup \Delta \sqcup \Sigma$ and $\partial^{\star}K_i \subseteq \Lambda^c \cup \Delta$.

We claim that:

$$\kappa_{\Lambda}(\upsilon) = \kappa_{\Lambda}(\upsilon(\Theta)1^{\Delta}0^{\Sigma}) + \sum_{i=1}^{n} \kappa_{K_i}(\upsilon(K_i)) = \kappa_{\Lambda}(\upsilon(\Theta)1^{\Delta}0^{\Sigma}) + \kappa_{\Sigma}(\tau), \tag{5.10}$$

for any $\upsilon \in \{0,1\}^{\Lambda}$ such that $\upsilon(\Delta) = 1^{\Delta}$ and $\upsilon(\Sigma) = \tau$.

To see this, given such υ, we exhibit a bijection r between the connected components of υ that do not intersect $\underline{\partial}\Lambda$ and the union of: (a) the connected components of $\upsilon(\Theta)1^\Delta 0^\Sigma$ that do not intersect $\underline{\partial}\Lambda$, and (b) the connected components of $\upsilon(K_i)$ that do not intersect $\underline{\partial}K_i$, for all i; namely, if $C \subseteq \Lambda$ is a connected component of υ, then r is defined as follows:

$$r(C) = \begin{cases} C \cap \overline{\Theta}^\star & \text{if } C \cap \overline{\Theta}^\star \neq \emptyset, \\ C & \text{if } C \subseteq \Sigma. \end{cases} \tag{5.11}$$

In order to see that r is well-defined, note that if C intersects $\overline{\Theta}^\star$ and Σ, the set $C \cap \overline{\Theta}^\star$ is still connected thanks to the fact that $\partial^\star K_i$ is connected and $\upsilon(\Delta) = 1^\Delta$. To see that r is onto, observe that if C' is a connected component of $\upsilon(\Theta)1^\Delta 0^\Sigma$, then there is a unique component C of υ such that $C \cap \overline{\Theta}^\star = C'$, due again to the fact that $\partial^\star K_i$ is connected. And r is clearly injective because two distinct connected components cannot intersect.
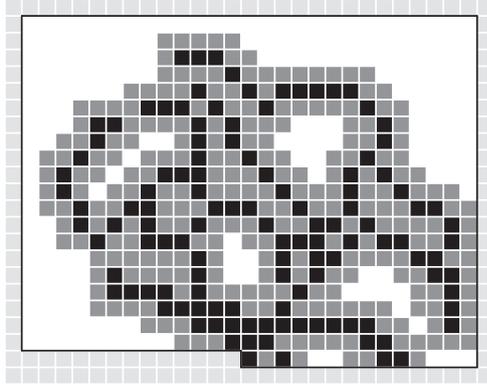
Fig. 1 A $\star$-connected Θ (in black), the set $\Delta = \partial^\star\Theta \cap \Lambda$ (in dark grey) and Λ^c (in light grey) for $\Lambda = S_{y,z}$.

Finally, we conclude from (5.10) that:

$$\psi^{(1)}_{p,q,\Lambda}(\theta \mid 1^\Delta\tau) = \frac{\lambda^{\#_1(\theta 1^\Delta\tau)} q^{\kappa_\Lambda(\theta 1^\Delta\tau)}}{\sum_{\upsilon \in \{0,1\}^\Lambda : \upsilon(\Delta) = 1^\Delta, \upsilon(\Sigma) = \tau} \lambda^{\#_1(\upsilon)} q^{\kappa_\Lambda(\upsilon)}} \tag{5.12}$$

$$= \frac{\lambda^{\#_1(\theta 1^\Delta) + \#_1(\tau)} q^{\kappa_\Lambda(\theta 1^\Delta 0^\Sigma) + \kappa_\Sigma(\tau)}}{\sum_{\upsilon \in \{0,1\}^\Lambda : \upsilon(\Sigma) = \tau} \lambda^{\#_1(\upsilon(\Theta)1^\Delta) + \#_1(\tau)} q^{\kappa_\Lambda(\upsilon(\Theta)1^\Delta 0^\Sigma) + \kappa_\Sigma(\tau)}} \tag{5.13}$$

$$= \frac{\lambda^{\#_1(\theta 1^\Delta)} q^{\kappa_\Lambda(\theta 1^\Delta 0^{\Lambda\backslash\overline{\Theta}^\star})}}{\sum_{\tilde\theta \in \{0,1\}^\Theta} \lambda^{\#_1(\tilde\theta 1^\Delta)} q^{\kappa_\Lambda(\tilde\theta 1^\Delta 0^{\Lambda\backslash\overline{\Theta}^\star})}} = \psi^{(1)}_{p,q,\Lambda}(\theta \mid 1^\Delta 0^{\Lambda\backslash\overline{\Theta}^\star}), \quad (5.14)$$

as we wanted. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Remark 1. *We claim that if $\emptyset \neq \Theta \subseteq \Lambda \Subset \mathbb{Z}^2$ are such that Λ^c is connected, Θ is $\star$-connected and $\overline{\Theta}^\star \cap \underline{\partial}\Lambda \neq \emptyset$, then $\Lambda^c \cup \overline{\Theta}^\star$ is connected, which is the main hypothesis of Lemma 5.5. This follows from the easy fact that the $\star$-closure of a $\star$-connected set is connected.*

6 Pressure representation

6.1 Variational principle

The variational principle states that the pressure of an interaction has a variational characterization in terms of shift-invariant measures. We state the variational principle below for the case of an n.n. interaction Φ for a set of restrictions $\mathscr{E}$.

Theorem 6.1 (Variational principle [23, 33, 35]). *Given a n.n. interaction Φ for a set of restrictions $\mathscr{E}$, we have that:*

$$\mathrm{P}(\Phi) = \sup_{\mu \in \mathscr{M}_{1,\sigma}(\Omega(\mathscr{E}))} \left(h(\mu) + \int A_\Phi d\mu \right), \tag{6.1}$$

where:

- $A_\Phi(\omega) := -\Phi(\omega(\mathbf{0})) - \sum_{i=1}^d \Phi(\omega(\{\mathbf{0}, e_i\}))$, *for* $\omega \in \Omega(\mathscr{E})$, *and*
- $h(\mu) := \lim_{n \to \infty} \frac{-1}{|\mathrm{B}_n|} \sum_{\theta \in \mathscr{A}^{\mathrm{B}_n}} \mu(\theta) \log(\mu(\theta))$ *is the* measure-theoretic entropy *of* μ, *where* $0 \log 0 = 0$.

In this case, the supremum is also always achieved (see [23, Section 4.2]) and any measure which achieves the supremum is called an *equilibrium state* for A_Φ. So, if μ is an equilibrium state, then:

$$\mathrm{P}(\Phi) = h(\mu) + \int A_\Phi d\mu. \tag{6.2}$$

For a shift-invariant measure μ and $\Lambda \Subset \mathbb{Z}^d \setminus \{\mathbf{0}\}$, define:

$$p_{\mu,\Lambda}(\omega) := \mu(\omega(\mathbf{0})|\omega(\Lambda)), \tag{6.3}$$

and let $p_\mu(\omega) := \lim_{n \to \infty} p_{\mu, \mathrm{B}_n \cap \mathscr{P}}(\omega)$, which exists μ-a.s. [23, Theorem 3.1.10] by Lévy's zero-one law. In addition, let:

$$I_\mu(\omega) := -\log p_\mu(\omega), \tag{6.4}$$

which is also defined μ-a.s. and is usually called the *information function*. It is well-known (see [16, p. 318, Equation 15.18] or [27, Theorem 2.4, p. 283]) that for any shift-invariant measure μ, $h(\mu) = \int I_\mu d\mu$. Therefore, if μ is an equilibrium state for Φ, we can rewrite the preceding formula for $\mathrm{P}(\Phi)$ as:

$$\mathrm{P}(\Phi) = \int \left(I_\mu + A_\Phi \right) d\mu. \tag{6.5}$$

So, the pressure can be represented as the integral of a function, determined by an equilibrium state μ and Φ, with respect to μ.

In this section, we show that the pressure can be represented as the integral of a function similar to $I_\mu + A_\Phi$, with respect to any invariant measure ν, assuming some conditions. This is useful for approximation of pressure when ν is an atomic measure supported on a periodic configuration (see Section 9).

One of the conditions involves the SFT $\Omega(\mathscr{E})$.

Definition 6.1. *A n.n. SFT $\Omega(\mathscr{E})$ for a set of constraints $\mathscr{E}$ satisfies the* square block D-condition *if there exists a sequence of integers $\{r_n\}_{n \geq 1}$ such that $\frac{r_n}{n} \to 0$ as $n \to \infty$ and, for any finite set $\Lambda \Subset \mathrm{B}_{n+r_n}^c$, $\theta \in \mathscr{A}^{\mathrm{B}_n}$ and $\tau \in \mathscr{A}^\Lambda$:*

$$[\theta]^{\Omega(\mathscr{E})}, [\tau]^{\Omega(\mathscr{E})} \neq \emptyset \implies [\theta\tau]^{\Omega(\mathscr{E})} \neq \emptyset. \tag{6.6}$$

This condition is a strengthened version of the classical D-condition (see [35, Section 4.1]) which guarantees that the set of Gibbs measures for Φ coincides with the set of equilibrium states for A_Φ.

Definition 6.2. *Given a set of restrictions $\mathscr{E}$, the corresponding n.n. SFT $\Omega(\mathscr{E}) \subseteq \mathscr{A}^{\mathbb{Z}^d}$ and $a \in \mathscr{A}$, we say that $\Omega(\mathscr{E})$ has a* safe symbol *a if $(a,b),(b,a) \notin \mathscr{E}_i$, for every $b \in \mathscr{A}$, for all $i = 1, \ldots, d$.*

It is easy to see that if $\Omega(\mathscr{E})$ has a safe symbol, then it satisfies the square block D-condition. For the sets of restrictions $\mathscr{E}$ in the Potts, Widom-Rowlinson and hard-core models, the corresponding n.n. SFT $\Omega(\mathscr{E})$ has a safe symbol in each case (any $a \in \mathscr{A}_q$, $0 \in \mathscr{B}_q$, and $0 \in \{0,1\}$, respectively), so $\Omega(\mathscr{E})$ satisfies the square block D-condition for the three models.

6.2 The function $\hat{\pi}$ and additional notation

Given a n.n. interaction Φ for a set of constraints $\mathscr{E}$, we will define some useful functions from $\Omega(\mathscr{E})$ to $\mathbb{R}$. First, given $\mathbf{0} \in \Lambda \Subset \mathbb{Z}^d$ and $\omega \in \Omega(\mathscr{E})$, we define:

$$\pi_\Lambda(\omega) := \pi_\Lambda^\omega(\theta(\mathbf{0}) = \omega(\mathbf{0})) = \pi_\Lambda^{\omega(\partial\Lambda)}(\theta(\mathbf{0}) = \omega(\mathbf{0})). \tag{6.7}$$

Recall that, for $y,z \in \mathbb{Z}^d$ such that $y,z \geq \mathbf{0}$, we have defined the set $S_{y,z}$ as $\{x \succeq \mathbf{0} : -y \leq x \leq z\}$. Now, given $y,z \geq \mathbf{0}$ and $\omega \in \Omega(\mathscr{E})$, define $\pi_{y,z}(\omega) := \pi_{S_{y,z}}(\omega)$ and, given $n \in \mathbb{N}$, abbreviate $\pi_n(\omega) := \pi_{\mathbf{1}n,\mathbf{1}n}(\omega)$. Considering this, we also define the limit $\hat{\pi}(\omega) := \lim_{n \to \infty} \pi_n(\omega)$, whenever it exists. If such limit exists, we will also denote $\hat{I}_\pi(\omega) := -\log \hat{\pi}(\omega)$.

It is not difficult to prove that under some mixing assumptions over an MRF μ, namely the SSM property introduced in Definition 7.1 (see Section 7), and assuming that $\mathrm{supp}(\mu) = \Omega(\mathscr{E})$, one has that the original information function I_μ coincides with $\hat{I}_\pi$ in $\Omega(\mathscr{E})$. In this sense, our definition provides a generalization of previous results (see [15]), where I_μ may not be even well-defined.

Now, suppose we have a shift-invariant measure ν such that $\mathrm{supp}(\nu) \subseteq \Omega(\mathscr{E})$. We say that:

$$\lim_{y,z \to \infty} \pi_{y,z}(\omega) = \hat{\pi}(\omega) \text{ uniformly over } \omega \in \mathrm{supp}(\nu), \tag{6.8}$$

if for all $\varepsilon > 0$, exists $k \in \mathbb{N}$ such that for all $\forall y, z \geq \mathbf{1}k$:

$$\left| \pi_{y,z}(\omega) - \hat{\pi}(\omega) \right| < \varepsilon, \text{ for all } \omega \in \text{supp}(\nu). \tag{6.9}$$

In addition, we introduce the following bound:

$$c_\pi(\nu) := \inf\{\pi_\Lambda(\omega) : \mathbf{0} \in \Lambda \Subset \mathbb{Z}^d, \omega \in \text{supp}(\nu)\}. \tag{6.10}$$

Lemma 6.2. *Let π be a n.n. interaction Φ for a set of restrictions $\mathcal{E}$, with π and $\Omega(\mathcal{E})$ the corresponding specification and n.n. SFT. Then, if $\Omega(\mathcal{E})$ has a a safe symbol, we have that $c_\pi(\nu) > 0$, for any shift-invariant measure ν such that $\text{supp}(\nu) \subseteq \Omega(\mathcal{E})$.*

Proof. The proof is analogous and a particular case of [29, Proposition 2.17]. In that reference, under these assumptions, it is shown that $c_\mu(\nu) := \inf\{\mu(\omega(\mathbf{0})|\omega(\Lambda)) : \Lambda \in \mathbb{Z}^d \setminus \{\mathbf{0}\}, \omega \in \text{supp}(\nu)\} > 0$, for a given n.n. Gibbs measure μ for Φ. We leave it to the reader to verify that $c_\pi(\nu) \geq c_\mu(\nu)$, for any such μ. $\qquad\square$

In fact, much weaker conditions than the existence of a safe symbol are sufficient for the result of Lemma 6.2 and also for having the square block D-condition. See, for example, the *single-site fillability* property [29] and the *topological strong spatial mixing property* [9]. Notice that, since the Potts, Widom-Rowlinson and hard-core models have a safe symbol, we have that $c_\pi(\nu) > 0$, for any shift-invariant ν with $\text{supp}(\nu) \subseteq \Omega(\mathcal{E})$.

6.3 Pressure representation theorem

Pressure representation results can be found in [29, Theorems 3.1 and 3.6]. Those results are not adequate for the application to the specific models we are considering in this paper. Instead we will use the following result, whose proof is adapted from the proof of [29, Theorem 3.1], as well as an idea of [29, Theorem 3.6]. In contrast to the results of [29], our result makes assumptions on the specification rather than a Gibbs measure.

Theorem 6.3. *Let Φ be a n.n. interaction for a set of restrictions $\mathcal{E}$ and suppose that $\Omega(\mathcal{E})$ satisfies the square block D-condition. Let ν be a shift-invariant measure such that $\text{supp}(\nu) \subseteq \Omega(\mathcal{E})$ and $c_\pi(\nu) > 0$. In addition, suppose that:*

$$\lim_{y,z \to \infty} \pi_{y,z}(\omega) = \hat{\pi}(\omega) \text{ uniformly over } \omega \in \text{supp}(\nu). \tag{6.11}$$

Then:

$$P(\Phi) = \int \left(\hat{I}_\pi + A_\Phi \right) d\nu. \tag{6.12}$$

Proof. Choose $\ell < 0$ and $L > 0$ to be lower and upper bounds respectively on values of Φ. Given $n \in \mathbb{N}$, let r_n be as in the definition of the square block D-condition and consider the sets B_n and $\Lambda_n := B_{n+r_n}$. We begin by proving that:

$$\frac{1}{|B_n|}(\log Z_{B_n}^\Phi + \log \pi_{\Lambda_n}^\omega(\omega(B_n)) + E_{B_n}^\Phi(\omega(B_n))) \to 0, \tag{6.13}$$

uniformly in $\omega \in \Omega$. For this, we will only use the square block D-condition. We fix $n \in \mathbb{N}$, $\omega \in \operatorname{supp}(\nu)$ and let $m_n := |\Lambda_n| - |\mathrm{B}_n|$. Let $C_d \geq 1$ be a constant such that for any $\Delta \Subset \mathbb{Z}^d$, the total number of sites and bonds contained in $\overline{\Delta}$ is bounded from above by $C_d|\Delta|$.

$$\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)) \geq \pi_{\Lambda_n}^{\omega}(\omega(\Lambda_n)) \tag{6.14}$$

$$= \frac{\exp(-\mathrm{E}_{\Lambda_n}^{\Phi}(\omega(\overline{\Lambda_n})))}{\sum_{\theta:\theta\omega(\partial\Lambda_n) \text{ feasible}} \exp(-E_{\Lambda_n}^{\Phi}(\theta\omega(\partial\Lambda_n)))} \tag{6.15}$$

$$\geq \frac{\exp(-\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\omega(\mathrm{B}_n)) - C_d m_n L)}{\sum_{\tau\in\mathscr{A}^{\mathrm{B}_n}:\tau \text{ feasible}} \exp(-\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\tau))|\mathscr{A}|^{C_d m_n}\exp(-C_d m_n \ell)} \tag{6.16}$$

$$= \frac{\exp(-\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\omega(\mathrm{B}_n)))}{\mathrm{Z}_{\mathrm{B}_n}^{\Phi}}\exp(m_n(C_d\ell - C_d L - C_d\log|\mathscr{A}|)). \tag{6.17}$$

Now, if $\tau_{\max}$ achieves the maximum of $\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)\tau)$ over $\tau \in \mathscr{A}^{\Lambda_n \setminus \mathrm{B}_n}$, then:

$$\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)) = \sum_{\tau\in\mathscr{A}^{\Lambda_n\setminus\mathrm{B}_n}:\omega(\mathrm{B}_n)\tau \text{ feasible}} \pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)\tau) \tag{6.18}$$

$$\leq |\mathscr{A}|^{m_n}\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)\tau_{\max}) \tag{6.19}$$

$$= |\mathscr{A}|^{m_n}\frac{\exp(-\mathrm{E}_{\Lambda_n}^{\Phi}(\omega(\mathrm{B}_n)\tau_{\max}\omega(\partial\Lambda_n)))}{\sum_{\theta:\theta\omega(\partial\Lambda_n) \text{ feasible}} \exp(-\mathrm{E}_{\Lambda_n}^{\Phi}(\theta\omega(\partial\Lambda_n)))} \tag{6.20}$$

$$\leq |\mathscr{A}|^{m_n}\frac{\exp(-\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\omega(\mathrm{B}_n)) - C_d m_n \ell)}{\sum_{\tau\in\mathscr{A}^{\mathrm{B}_n}:[\tau]^{\Omega}\neq\emptyset} e^{-\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\tau)}\exp(-C_d m_n L)} \tag{6.21}$$

$$\leq \frac{\exp(-\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\omega(\mathrm{B}_n)))}{\hat{Z}_{\mathrm{B}_n}^{\Phi}}\exp(-m_n(C_d\ell - C_d L - C_d\log|\mathscr{A}|)), \tag{6.22}$$

where the square block D-condition has been used in (6.21). Therefore,

$$\alpha^{-m_n} \leq \pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n))\mathrm{Z}_{\mathrm{B}_n}^{\Phi}\exp(\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\omega(\mathrm{B}_n))) \leq \frac{\mathrm{Z}_{\mathrm{B}_n}^{\Phi}}{\hat{Z}_{\mathrm{B}_n}^{\Phi}}\alpha^{m_n}, \tag{6.23}$$

where $\alpha := e^{-(C_d\ell - C_d L - C_d\log|\mathscr{A}|)}$. Since $\frac{m_n}{|\mathrm{B}_n|} \to 0$ and $\frac{1}{|\mathrm{B}_n|}\left(\log\mathrm{Z}_{\mathrm{B}_n}^{\Phi} - \log\hat{Z}_{\mathrm{B}_n}^{\Phi}\right) \to 0$ (thanks to Theorem 3.1), we have obtained (6.13).

We use (6.13) to represent pressure:

$$\mathrm{P}(\Phi) = \lim_{n\to\infty}\frac{\log\mathrm{Z}_{\mathrm{B}_n}^{\Phi}}{|\mathrm{B}_n|} = \lim_{n\to\infty}\int\frac{\log\mathrm{Z}_{\mathrm{B}_n}^{\Phi}}{|\mathrm{B}_n|}d\nu \tag{6.24}$$

$$= \lim_{n\to\infty}\int\frac{-\log\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)) - \mathrm{E}_{\mathrm{B}_n}^{\Phi}(\omega(\mathrm{B}_n))}{|\mathrm{B}_n|}d\nu. \tag{6.25}$$

(Here the second equality comes from the fact that $\frac{\log Z_{\mathrm{B}_n}^{\Phi}}{|\mathrm{B}_n|}$ is independent of ω, and the third from (6.13).) Since ν is shift-invariant, it can be checked that:

$$\lim_{n\to\infty} \int \frac{-\mathrm{E}_{\mathrm{B}_n}^{\Phi}(\omega(\mathrm{B}_n))}{|\mathrm{B}_n|} d\nu = \int A_{\Phi} d\nu, \qquad (6.26)$$

and so we can write:

$$\mathrm{P}(\Phi) = \int A_{\Phi} d\nu - \lim_{n\to\infty} \int \frac{\log \pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n))}{|\mathrm{B}_n|} d\nu. \qquad (6.27)$$

It remains to show that:

$$\lim_{n\to\infty} \int \frac{-\log \pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n))}{|\mathrm{B}_n|} d\nu = \int \hat{I}_{\pi} d\nu. \qquad (6.28)$$

Fix $\omega \in \mathrm{supp}(\nu)$ and denote $c := c_{\pi}(\nu)$. We will decompose $\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n))$ as a product of conditional probabilities. By (6.11), for any $\varepsilon > 0$, there exists $k := k_{\varepsilon}$ so that for $y, z \geq \mathbf{1}k$, $|\pi_{y,z}(\omega) - \hat{\pi}(\omega)| < \varepsilon$ for all $\omega \in \mathrm{supp}(\nu)$. For $x \in \mathrm{B}_{n-1}$, we denote $\mathrm{B}_n^-(x) := \{y \in \mathrm{B}_{n-1} : y \prec x\}$. Then, we can decompose $\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n))$ as:

$$\pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)) = \pi_{\Lambda_n}^{\omega}(\omega(\partial \mathrm{B}_n)) \prod_{x\in\mathrm{B}_{n-1}} \pi_{\Lambda_n}^{\omega}\left(\omega(x)\big|\omega\left(\mathrm{B}_n^-(x)\cup\partial\mathrm{B}_n\right)\right) \qquad (6.29)$$

$$= \pi_{\Lambda_n}^{\omega}(\omega(\partial \mathrm{B}_n)) \prod_{x\in\mathrm{B}_{n-1}} \pi_{y(x),z(x)}(\sigma_x(\omega)), \qquad (6.30)$$

where $y(x) := \mathbf{1}n + x$ and $z(x) := \mathbf{1}n - x$, thanks to the MRF property and stationarity of the specification.

Let's denote $R_{n,k} := \mathrm{B}_n \setminus \mathrm{B}_{n-k}$. Then, $\mathrm{B}_n = \partial\mathrm{B}_n \sqcup \mathrm{B}_{n-k-1} \sqcup R_{n-1,k}$ and we have:

$$c^{|\partial\mathrm{B}_n|+|R_{n-1,k}|} \prod_{x\in\mathrm{B}_{n-k-1}} \pi_{y(x),z(x)}(\sigma_x(\omega)) \leq \pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)) \qquad (6.31)$$

$$\leq \prod_{x\in\mathrm{B}_{n-k-1}} \pi_{y(x),z(x)}(\sigma_x(\omega)). \qquad (6.32)$$

Taking $-\log(\cdot)$, we have that:

$$0 \leq -\log \pi_{\Lambda_n}^{\omega}(\omega(\mathrm{B}_n)) - \sum_{x\in\mathrm{B}_{n-k-1}} -\log \pi_{y(x),z(x)}(\sigma_x(\omega)) \qquad (6.33)$$

$$\leq (|\partial\mathrm{B}_n| + |R_{n-1,k}|) \log\left(c^{-1}\right). \qquad (6.34)$$

So, by the choice of k, for $x \in \mathrm{B}_{n-k-1}$,

$$\left|\pi_{y(x),z(x)}(\sigma_x(\omega)) - \hat{\pi}(\sigma_x(\omega))\right| < \varepsilon, \qquad (6.35)$$

and since $\pi_{y(x),z(x)}(\sigma_x(\omega)), \hat{\pi}(\sigma_x(\omega)) \geq c > 0$, by the Mean Value Theorem:

$$\left|-\log \pi_{y(x),z(x)}(\sigma_x(\omega)) - \hat{I}_{\pi}(\sigma_x(\omega))\right| < \varepsilon c^{-1}, \qquad (6.36)$$

It follows from (6.11) that $\hat{\pi}$ is the uniform limit of continuous functions on supp(ν). In addition, $\hat{\pi}(\omega) \geq c > 0$, for all $\omega \in$ supp(ν). Therefore, we can integrate with respect to ν to see that:

$$\left| \int -\log \pi_{y(x),z(x)}(\sigma_x(\omega)) d\nu - \int \hat{I}_\pi(\omega) d\nu \right| < \varepsilon c^{-1}. \tag{6.37}$$

We now combine the previous equations to see that:

$$\left| \int -\log \pi_{\Lambda_n}^\omega(\omega(B_{n+1})) d\nu - \int \hat{I}_\pi(\omega) d\nu | B_{n-k-1} | \right| \tag{6.38}$$

$$\leq |B_{n-k-1}| \varepsilon c^{-1} + (|\partial B_n| + |R_{n-1,k}|) \log\left(c^{-1}\right). \tag{6.39}$$

Notice that, for a fixed k, $\lim_{n\to\infty} \frac{|\partial B_n| + |R_{n-1,k}|}{|B_n|} = 0$ and $\lim_{n\to\infty} \frac{|B_{n-k-1}|}{|B_n|} = 1$. Therefore,

$$-\varepsilon c^{-1} + \int \hat{I}_\pi(\omega) d\nu \leq \liminf_{n\to\infty} \int \frac{-\log \pi_{\Lambda_n}^\omega(\omega(B_n))}{|B_n|} d\nu \tag{6.40}$$

$$\leq \limsup_{n\to\infty} \int \frac{-\log \pi_{\Lambda_n}^\omega(\omega(B_n))}{|B_n|} d\nu \tag{6.41}$$

$$\leq \int \hat{I}_\pi(\omega) d\nu + \varepsilon c^{-1}. \tag{6.42}$$

By letting $\varepsilon \to 0$, we see that:

$$\lim_{n\to\infty} \int \frac{-\log \pi_{\Lambda_n}^\omega(\omega(B_n))}{|B_n|} d\nu = \int \hat{I}_\pi(\omega) d\nu, \tag{6.43}$$

completing the proof. $\qquad\square$

7 Spatial mixing and stochastic dominance

From now on, when talking about specifications for the Potts, Widom-Rowlinson and hard-core lattice models, we will distinguish them by the subindex corresponding to the parameter β, λ or γ of the model, i.e. $\pi_{\beta,\Lambda}^\xi$ should be understood as a probability measure in the Potts model, $\pi_{\lambda,\Lambda}^\xi$ in the Widom-Rowlinson and $\pi_{\gamma,\Lambda}^\xi$ in the hard-core lattice model, and π_β, π_λ and π_γ will denote the corresponding specifications. Also, we will write π_Λ^β, $\hat{\pi}^\beta$ and $\hat{I}_\pi^\beta$ for the functions π_Λ, $\hat{\pi}$ and $\hat{I}_\pi$ in the Potts model, and short-hand notations when $\Lambda = S_n$ or $S_{y,z}$. For example,

$$\pi_n^\beta(\omega) := \pi_{S_n}^\beta(\omega) := \pi_{\beta,S_n}^{\omega(\partial S_n)}(\theta(0) = \omega(0))$$

The analogous notation will be used for the Widom-Rowlinson and hard-core cases, but using the parameters λ and γ, respectively.

7.1 Spatial mixing properties

We now introduce concepts of spatial mixing that we will need in this paper. Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a function such that $f(n) \searrow 0$ as $n \to \infty$.

Definition 7.1. *Given $\Lambda \subseteq \mathbb{Z}^d$, we say that a Λ-MRF μ satisfies* strong spatial mixing *(SSM) with rate $f(n)$ for a class of finite sets $\mathscr{C}$ if for any $\Delta \in \mathscr{C}$ such that $\overline{\Delta} \subseteq \Lambda$, any $\Theta \subseteq \Delta$, $\theta \in \mathscr{A}^{\Theta}$ and $\xi, \eta \in \mathscr{A}^{\partial \Delta}$ with $\mu(\xi)\mu(\eta) > 0$,*

$$|\mu(\theta|\xi) - \mu(\theta|\eta)| \leq |\Theta| f\left(\mathrm{dist}(\Theta, \Sigma_{\partial \Delta}(\xi, \eta))\right). \tag{7.1}$$

We say that a Gibbs specification $\pi = \{\pi_{\Lambda}^{\xi}\}_{\Lambda,\xi}$ satisfies SSM with rate $f(n)$ for a class of finite sets $\mathscr{C}$ if each element π_{Λ}^{ξ} satisfies SSM with rate $f(n)$ for the class $\mathscr{C}$ restricted to subsets of Λ.

If there exists $C, \alpha > 0$ such that f can be chosen to be $f(n) = Ce^{-\alpha n}$, we say that exponential SSM *holds.*

Definition 7.2. *([3, p. 445]) A $\mathbb{Z}^d$-MRF μ satisfies the* ratio strong mixing property *for a class of finite sets $\mathscr{C}$ if there exists $C, \alpha > 0$ such that for any $\Delta \in \mathscr{C}$, any $\Theta, \Sigma \subseteq \Delta$ and $\xi \in \mathscr{A}^{\partial \Delta}$ with $\mu(\xi) > 0$,*

$$\sup\left\{\left|\frac{\mu(A \cap B|\xi)}{\mu(A|\xi)\mu(B|\xi)} - 1\right| : A \in \mathscr{F}_{\Theta}, B \in \mathscr{F}_{\Sigma}, \mu(A|\xi)\mu(B|\xi) > 0\right\}$$
$$\leq C \sum_{x \in \Theta, y \in \Sigma} e^{-\alpha \mathrm{dist}(x,y)}. \tag{7.2}$$

Proposition 7.1. *Let μ be a $\mathbb{Z}^2$-MRF with $\mathrm{supp}(\mu) = \mathscr{A}^{\mathbb{Z}^2}$ that satisfies the ratio strong mixing property for the class of finite simply lattice-connected sets. Then, μ satisfies exponential SSM for the family of sets $\{S_{y,z}\}_{y,z \geq 0}$.*

Proof. Fix $y, z \geq 0$ and the corresponding set $S_{y,z} \Subset \mathbb{Z}^2$. Let $\Theta \subseteq S_{y,z}$, $\theta \in \mathscr{A}^{\Theta}$ and $\xi_1, \xi_2 \in \mathscr{A}^{\partial S_{y,z}}$ with $\mu(\xi_1)\mu(\xi_2) > 0$, consider:

1. the sets $\Sigma := \Sigma_{\partial S_{y,z}}(\xi_1, \xi_2)$ and $\Delta := S_{y,z} \cup \Sigma$,
2. an arbitrary configuration $\tilde{\xi} \in \mathscr{A}^{\partial \Delta}$ such that $\tilde{\xi}(\partial S_{y,z} \setminus \Sigma) = \xi_i(\partial S_{y,z} \setminus \Sigma)$ $(i = 1, 2)$, and
3. the events $A := [\theta] \in \mathscr{F}_{\Theta}$ and $B_i := [\xi_i(\Sigma)] \in \mathscr{F}_{\Sigma}$, for $i = 1, 2$.

Notice that Δ is a finite simply lattice-connected set and, since $\mathrm{supp}(\mu) = \mathscr{A}^{\mathbb{Z}^d}$, we can be sure that $\mu(\tilde{\xi}) > 0$. Then:

$$|\mu(\theta|\xi_1) - \mu(\theta|\xi_2)| = \left|\mu(A|[\tilde{\xi}] \cap B_1) - \mu(A|[\tilde{\xi}] \cap B_2)\right| \tag{7.3}$$

$$= \left|\frac{\mu(A \cap B_1|\tilde{\xi})}{\mu(B_1|\tilde{\xi})} - \mu(A|\tilde{\xi}) + \mu(A|\tilde{\xi}) - \frac{\mu(A \cap B_2|\tilde{\xi})}{\mu(B_2|\tilde{\xi})}\right| \tag{7.4}$$

$$\leq \left|\frac{\mu(A \cap B_1|\tilde{\xi})}{\mu(B_1|\tilde{\xi})\mu(A|\tilde{\xi})} - 1\right| + \left|1 - \frac{\mu(A \cap B_2|\tilde{\xi})}{\mu(B_2|\tilde{\xi})\mu(A|\tilde{\xi})}\right| \tag{7.5}$$

$$\leq 2C \sum_{x \in \Theta, y \in \Sigma} e^{-\alpha \operatorname{dist}(x,y)} \tag{7.6}$$

$$\leq |\Theta| 2C \sum_{y \in \Sigma} e^{-\alpha \operatorname{dist}(\Delta, y)}. \tag{7.7}$$

W.l.o.g., we can assume that $|\Sigma| = 1$ (see [9, Corollary 2]). Therefore, by taking $C' = 2C$, we have:

$$|\mu(\theta|\xi_1) - \mu(\theta|\xi_2)| \leq |\Theta| 2K \sum_{y \in \Sigma} e^{-\alpha \operatorname{dist}(\Delta, y)} = |\Theta| C' e^{-\alpha \operatorname{dist}(\Theta, \Sigma)}. \tag{7.8}$$

$\square$

Remark 2. *The proof of Proposition 7.1 seems to require some assumption on the support of μ (for the existence of $\tilde{\xi}$ in the enumerated item list above). Fully supported (i.e. $\operatorname{supp}(\mu) = \mathscr{A}^{\mathbb{Z}^2}$) suffices, and is the only case in which we will apply this result (see Corollary 1), but the conclusion probably holds under weaker assumptions.*

Given $y, z \geq 0$, we define the *bottom boundary of $S_{y,z}$* as $\partial_\downarrow S_{y,z} := \partial S_{y,z} \cap \mathscr{P}$, i.e. the portion of the boundary of $S_{y,z}$ included in the past, and the *top boundary of $S_{y,z}$* as the complement $\partial_\uparrow S_{y,z} := \partial S_{y,z} \setminus \mathscr{P}$. Clearly, $\partial S_{y,z} = \partial_\downarrow S_{y,z} \sqcup \partial_\uparrow S_{y,z}$.

Proposition 7.2. *Let π be a specification satisfying exponential SSM with parameters $C, \alpha > 0$. Then, for all $n \in \mathbb{N}$, $y, z \geq \mathbf{1}n$ and $a \in \mathscr{A}$:*

$$\left| \pi_{S_n}^{\omega_1}(\theta(\mathbf{0}) = a) - \pi_{S_{y,z}}^{\omega_2}(\theta(\mathbf{0}) = a) \right| \leq C e^{-\alpha n}, \tag{7.9}$$

uniformly over $\omega_1, \omega_2 \in \Omega(\mathscr{E})$ such that $\omega_1(\mathscr{P}) = \omega_2(\mathscr{P})$.

Proof. Fix $n \in \mathbb{N}$, $y, z \geq \mathbf{1}n$, $a \in \mathscr{A}$ and $\omega_1, \omega_2 \in \Omega(\mathscr{E})$ such $\omega_1(\mathscr{P}) = \omega_2(\mathscr{P})$. Denote $\xi := \omega_1(\partial S_n)$. Then:

$$\left| \pi_{S_n}^{\omega_1}(\theta(\mathbf{0}) = a) - \pi_{S_{y,z}}^{\omega_2}(\theta(\mathbf{0}) = a) \right| \tag{7.10}$$

$$= \left| \pi_{S_n}^{\xi}(\theta(\mathbf{0}) = a) - \sum_{\eta} \pi_{S_{y,z}}^{\omega_2}(\theta(\mathbf{0}) = a | \eta) \pi_{S_{y,z}}^{\omega_2}(\eta) \right| \tag{7.11}$$

$$\leq \sum_{\eta} \left| \pi_{S_n}^{\xi}(\theta(\mathbf{0}) = a) - \pi_{S_n}^{\eta}(\theta(\mathbf{0}) = a) \right| \pi_{S_{y,z}}^{\omega_2}(\eta) \tag{7.12}$$

$$\leq \sum_{\eta} C e^{-\alpha \operatorname{dist}(\mathbf{0}, \Sigma_{\partial S_n}(\xi, \eta))} \pi_{S_{y,z}}^{\omega_2}(\eta) \leq C e^{-\alpha n}, \tag{7.13}$$

where the summation $\sum_\eta$ is taken over all $\eta \in \mathscr{A}^{\partial S_n}$ such that $\pi_{S_{y,z}}^{\omega_2}(\eta) > 0$ and $\eta(\partial_\downarrow S_n) = \omega_2(\partial_\downarrow S_n)$. The last inequality above follows from the fact that for any such η, $\Sigma_{\partial S_n}(\xi, \eta) \subseteq \partial_\uparrow S_n$, so:

$$\operatorname{dist}(\mathbf{0}, \Sigma_{\partial S_n}(\xi, \eta)) \geq \operatorname{dist}(\mathbf{0}, \partial_\uparrow S_n) = n. \tag{7.14}$$

$\square$

Definition 7.3 (Variational distance). *Let S be a finite set and let X_1 and X_2 be two S-valued random variables with distribution ρ_1 and ρ_2, respectively. The* variational distance *d_{TV} of X_1 and X_2 (or equivalently, of ρ_1 and ρ_2) is defined by:*

$$d_{TV}(\rho_1, \rho_2) := \frac{1}{2} \sum_{x \in S} |\rho_1(x) - \rho_2(x)|. \tag{7.15}$$

It is well-known that $d_{TV}(\rho_1, \rho_2)$ is a lower bound on $\mathbb{P}(X_1 \neq X_2)$ over all couplings $\mathbb{P}$ of ρ_1 and ρ_2 and that there is a coupling, called the *optimal coupling*, that achieves this lower bound.

Given a Gibbs specification π, we define:

$$Q(\pi) := \max_{\omega_1, \omega_2 \in \Omega(\mathscr{E})} d_{TV}\left(\pi_{\{\mathbf{0}\}}^{\omega_1}(\cdot), \pi_{\{\mathbf{0}\}}^{\omega_2}(\cdot)\right). \tag{7.16}$$

The following result is essentially in [7].

Theorem 7.3. *Let π be a Gibbs specification for a n.n. interaction Φ and a set of constraints $\mathscr{E}$, such that $\Omega(\mathscr{E})$ has a safe symbol. Then, if p_c denotes the critical value of site percolation on $\mathbb{Z}^2$ and $Q(\pi) < p_c$, we have that π satisfies exponential SSM.*

Proof. Take μ any n.n. Gibbs measure for Φ. Since $\Omega(\mathscr{E})$ has a safe symbol, μ is fully supported, i.e. $\operatorname{supp}(\mu) = \Omega(\mathscr{E})$ (very special case of [35, Remark 1.14]). Given a $\mathbb{Z}^d$-MRF μ, define:

$$Q(\mu) := \max_{\eta_1, \eta_2} d_{TV}(\mu(\theta(\mathbf{0}) = \cdot | \eta_1), \mu(\theta(\mathbf{0}) = \cdot | \eta_2)), \tag{7.17}$$

where η_1 and η_2 range over all configurations on $\partial\{\mathbf{0}\}$ such that $\mu(\eta_1), \mu(\eta_2) > 0$. Then, $Q(\mu) \leq Q(\pi) < p_c$, so by [7, Theorem 1] and shift-invariance of Φ, μ satisfies exponential SSM (see [30, Theorem 3.10]). Finally, since μ is fully supported, we can conclude that π satisfies exponential SSM. $\qquad\square$

7.2 Stochastic dominance

Suppose that $\mathscr{A}$ is a finite linearly ordered set. Then for any set L (in our context, usually a set of sites or bonds), $\mathscr{A}^L$ is equipped with a natural partial order $\preceq$ which is defined coordinate-wise: for $\theta_1, \theta_2 \in \mathscr{A}^L$, we write $\theta_1 \preceq \theta_2$ if $\theta_1(x) \leq \theta_2(x)$ for every $x \in L$. A function $f : \mathscr{A}^L \to \mathbb{R}$ is said to be *increasing* if $f(\theta_1) \leq f(\theta_2)$ whenever $\theta_1 \preceq \theta_2$. An event A is said to be increasing if its characteristic function χ_A is increasing.

Definition 7.4. *Let ρ_1 and ρ_2 be two probability measures on $\mathscr{A}^L$. We say that ρ_1 is* stochastically dominated by *ρ_2, writing $\rho_1 \leq_D \rho_2$, if for every bounded increasing function $f : \mathscr{A}^L \to \mathbb{R}$ we have $\rho_1(f) \leq \rho_2(f)$, where $\rho(f)$ denotes the expected value $\mathbb{E}_\rho(f)$ of f according to the measure ρ.*

7.2.1 Stochastic dominance and connectivity decay for the bond random-cluster model

Recall from Section 5.1 the bond random-cluster model on finite subsets of $\mathbb{Z}^2$ with boundary conditions $i = 0, 1$, and the bond random-cluster model $\phi_{p,q}$ on $\mathbb{Z}^2$ (see page 14).

Theorem 7.4 ([17, Equation (29)]). *For any $p \in [0,1]$ and $q \in \mathbb{N}$, and any $\Delta \subseteq \Lambda \Subset \mathbb{Z}^2$:*

$$\phi^{(0)}_{p,q,\Delta} \leq_D \phi^{(0)}_{p,q,\Lambda} \text{ and } \phi^{(1)}_{p,q,\Lambda} \leq_D \phi^{(1)}_{p,q,\Delta}. \tag{7.18}$$

In particular, if $p < p_c(q)$, we have that, for any $\Lambda \Subset \mathbb{Z}^2$:

$$\phi^{(0)}_{p,q,\Lambda} \leq_D \phi_{p,q} \leq_D \phi^{(1)}_{p,q,\Lambda}, \tag{7.19}$$

where $\leq_D$ is with respect to the restriction of each measure to events on $E^0(\Lambda)$.

The following result was a key element of the proof that $\beta_c(q) = \log(1+\sqrt{q})$ is the critical inverse temperature for the Potts model. We will use this result in a crucial way.

Recall that for $p < p_c(q)$, $\phi_{p,q}$ is the unique bond random cluster measure with parameters p and q.

Theorem 7.5 ([6, Theorem 2]). *Let $q \geq 1$. For any $p < p_c(q) = \frac{\sqrt{q}}{1+\sqrt{q}}$, the two-point connectivity function decays exponentially, i.e. there exist $0 < C(p,q), c(p,q) < \infty$ such that for any $x, y \in \mathbb{Z}^2$:*

$$\phi_{p,q}(x \leftrightarrow y) \leq C(p,q)e^{-c(p,q)\|x-y\|_2}, \tag{7.20}$$

where $\{x \leftrightarrow y\}$ is the event that the sites x and y are connected by an open path and $\|\cdot\|_2$ is the Euclidean norm.

7.2.2 Stochastic dominance for the site random-cluster model

Lemma 7.6. *Given a set $\Lambda \Subset \mathbb{Z}^d$ and parameters $p \in [0,1]$ and $q > 0$, we have that for any $x \in \Lambda$ and any $\tau \in \{0,1\}^{\Lambda \setminus \{x\}}$:*

$$p_1(q) \leq \psi^{(1)}_{p,q,\Lambda}(\theta(x) = 1|\tau) \leq p_2(q), \tag{7.21}$$

where $p_1(q) = \frac{pq}{pq+(1-p)q^{2d}}$ and $p_2(q) = \frac{pq}{pq+(1-p)}$. In consequence,

$$\Psi_{p_1(q),\Lambda} \leq_D \psi^{(1)}_{p,q,\Lambda} \leq_D \Psi_{p_2(q),\Lambda}. \tag{7.22}$$

(Recall that $\Psi_{p,\Lambda}$ denotes Bernoulli site percolation).

Proof. This result is obtained by adapting the discussion on [20, p. 339] to the wired site random-cluster model. See also [21, Lemma 5.4] for the case $q = 2$. $\square$

7.2.3 Stochastic dominance for the Potts model

As before, let $q \in \mathscr{A}_q$ denote a fixed, but arbitrary, choice of a colour. Let $\Lambda \Subset \mathbb{Z}^d$ and consider $g : \mathscr{A}_q^\Lambda \to \{+,-\}^\Lambda$ be defined by:

$$(g(\theta))(x) = \begin{cases} + & \text{if } \theta(x) = q, \\ - & \text{if } \theta(x) \neq q. \end{cases} \tag{7.23}$$

The function g makes the non-q colours indistinguishable and gives a *reduced model*. We say $\theta \simeq \theta'$ if $g(\theta) = g(\theta')$. This relation defines a partition of $\mathscr{A}_q^\Lambda$ and unions of elements of this partition form a sub-algebra of $\mathscr{A}_q^\Lambda$, which can be identified with the collection of all subsets of $\{+,-\}^\Lambda$. Let $\pi_{\beta,\Lambda}^+ := g_* \pi_{\beta,\Lambda}^{\omega_q}$ be the push-forward measure, which is nothing more than the restriction (projection) of $\pi_{\beta,\Lambda}^{\omega_q}$ to $\{+,-\}^\Lambda$. Chayes showed that the FKG property holds on events in this reduced model. In particular:

Proposition 7.7 ([11, Lemma on p. 211]). *For all $\beta > 0$ and $\Lambda \Subset \mathbb{Z}^2$, $\pi_{\beta,\Lambda}^+$ satisfies the following properties:*

1. *For increasing subsets $A, B \subseteq \{+,-\}^\Lambda$: $\pi_{\beta,\Lambda}^+(A \mid B) \geq \pi_{\beta,\Lambda}^+(A)$.*
2. *If A is decreasing and B is increasing, then: $\pi_{\beta,\Lambda}^+(A \mid B) \leq \pi_{\beta,\Lambda}^+(A)$.*
3. *If $\Delta \subseteq \Lambda$ and A is an increasing subset of $\{+,-\}^\Delta$, then: $\pi_{\beta,\Delta}^+(A) \geq \pi_{\beta,\Lambda}^+(A)$.*

Proof.

1. This is contained in [11, Lemma on p. 211].
2. This is an immediate consequence of (1).
3. This is a standard consequence of (1): Let $B = +^{\partial\Delta}$. Since $g^{-1}(B)$ is a single configuration, namely $q^{\partial\Delta}$, we obtain from the Markov property of $\pi_{\beta,\Lambda}^{\omega_q}$ that $\pi_{\beta,\Delta}^+(A) = \pi_{\beta,\Lambda}^+(A \mid B)$. From (1), we have $\pi_{\beta,\Lambda}^+(A \mid B) \geq \pi_{\beta,\Lambda}^+(A)$. Now, combine the previous two statements.

$\square$

Remark 3. *The preceding result immediately applies to $\pi_{\beta,\Lambda}^{\omega_q}$ for events in $\mathscr{A}_q^\Lambda$ that are measurable with respect to $\{+,-\}^\Lambda$, viewed as a sub-algebra of $\mathscr{A}_q^\Lambda$.*

7.2.4 Volume monotonicity for the Widom-Rowlinson model with 2 types

For the classical Widom-Rowlinson model ($q = 2$), Higuchi and Takei showed that the FKG property holds. In particular,

Proposition 7.8 ([21, Lemma 2.3]). *Fix $q = 2$ and let $\Delta \subseteq \Lambda \Subset \mathbb{Z}^d$ and $\lambda > 0$. Then:*

$$\pi_\Lambda^\lambda(\omega_q) \leq \pi_\Delta^\lambda(\omega_q). \tag{7.24}$$

However, this kind of stochastic monotonicity can fail for general q (see [17, p. 60]).

8 Exponential convergence of π_n in $\mathbb{Z}^2$ lattice models

In this section, we consider the Potts, Widom-Rowlinson and hard-core models and establish exponential convergence results that will lead to pressure representation and approximation algorithms for these lattice models.

Recall that for the Potts model, $\pi_{y,z}^{\beta}(\omega) = \pi_{\beta,S_n}^{\omega(\partial S_{y,z})}(\theta(\mathbf{0}) = \omega(\mathbf{0}))$ and, in particular, $\pi_n^{\beta}(\omega) = \pi_{\beta,S_n}^{\omega(\partial S_n)}(\theta(\mathbf{0}) = \omega(\mathbf{0}))$, with similar notation for the Widom-Rowlinson and hard core models.

8.1 Exponential convergence in the Potts model

Theorem 8.1. *For the Potts model with q types and inverse temperature β, there exists a critical parameter $\beta_c(q) > 0$ such that for $0 < \beta \neq \beta_c(q)$, there exists $C, \alpha > 0$ such that, for every $y, z \geq \mathbf{1}n$:*

$$\left| \pi_n^{\beta}(\omega_q) - \pi_{y,z}^{\beta}(\omega_q) \right| \leq Ce^{-\alpha n}. \tag{8.1}$$

Proof. In the supercritical region $\beta > \beta_c(q)$, our proof very closely follows [10, Theorem 3], which treated the Ising case. We fill in some details of their proof, adapting that proof in two ways: to a half-plane version of their result (the quantities in (8.1) are effectively half-plane quantities) and to the general Potts case. For the subcritical region $\beta < \beta_c(q)$, the proposition will follow easily from [3, Theorem 1.8 (ii)].

Part I: $\beta > \beta_c(q)$. Let $T_{\partial S_n}^{-\star}$ denote the event that there is a $\star$-path of $-$ from $\mathbf{0}$ to ∂S_n, i.e. a path that runs along ordinary $\mathbb{Z}^2$ bonds and diagonal bonds where the colour at each site is *not* q (in our context below, the configuration on the bottom piece $\partial_{\downarrow} S_{y,z}$ of ∂S_n will be all q and thus a $\star$-path of $-$ from $\mathbf{0}$ to ∂S_n cannot terminate on $\partial_{\downarrow} S_n$). Note that $T_{\partial S_n}^{-\star}$ is an event that is measurable with respect to the sub-algebra $\{+,-\}^{\Lambda}$, for any finite set Λ containing S_n, introduced in Section 7.2.3 (recall that this sub-algebra corresponds to the reduced Potts model).

By decomposing $\pi_{y,z}^{\beta}(\omega_q)$ into probabilities conditional on $T_{\partial S_n}^{-\star}$ and $(T_{\partial S_n}^{-\star})^c$, we obtain:

$$\pi_n^{\beta}(\omega_q) - \pi_{y,z}^{\beta}(\omega_q) \tag{8.2}$$

$$= \pi_{\beta,S_n}^{\omega_q}(\theta(\mathbf{0}) = q) - \pi_{\beta,S_{y,z}}^{\omega_q}(\theta(\mathbf{0}) = q) \tag{8.3}$$

$$= \pi_{\beta,S_{y,z}}^{\omega_q}(T_{\partial S_n}^{-\star}) \left(\pi_{\beta,S_n}^{\omega_q}(\theta(\mathbf{0}) = q) - \pi_{\beta,S_{y,z}}^{\omega_q}(\theta(\mathbf{0}) = q | T_{\partial S_n}^{-\star}) \right)$$

$$+ (1 - \pi_{\beta,S_{y,z}}^{\omega_q}(T_{\partial S_n}^{-\star})) \left(\pi_{\beta,S_n}^{\omega_q}(\theta(\mathbf{0}) = q) - \pi_{\beta,S_{y,z}}^{\omega_q}(\theta(\mathbf{0}) = q | (T_{\partial S_n}^{-\star})^c) \right). \tag{8.4}$$

We claim that the expression in (8.2) is nonnegative. To see this, observe that the events $\{\omega(\mathbf{0}) = q\}$, $\{\omega(\partial S_n) = q^{\partial S_n}\}$ and $\{\omega(\partial S_{y,z}) = q^{\partial S_{y,z}}\}$ may be viewed as the events $\{\omega(\mathbf{0}) = +\}$, $\{\omega(\partial S_n) = +^{\partial S_n}\}$ and $\{\omega(\partial S_n) = +^{\partial S_{y,z}}\}$ in the sub-algebra $\{+,-\}^{S_{y,z}}$ of the reduced model, as discussed in Section 7.2.3. Now, apply Proposition 7.7 (part 3) and Remark 3.

We next claim that:

$$\pi_{\beta,S_{y,z}}^{\omega_q}(\theta(\mathbf{0}) = q|(\mathrm{T}_{\partial S_n}^{-\star})^c) \geq \pi_{\beta,S_n}^{\omega_q}(\theta(\mathbf{0}) = q). \tag{8.5}$$

To be precise, first observe that $\omega \in (\mathrm{T}_{\partial S_n}^{-\star})^c$ iff ω contains an all-q path in S_n from $\partial \mathscr{P} \cap \{x_1 < 0\}$ to $\partial \mathscr{P} \cap \{x_1 > 0\}$. So, $(\mathrm{T}_{\partial S_n}^{-\star})^c$ can be decomposed into a disjoint collection of events determined by the unique furthest such path from $\mathbf{0}$. Using the MRF property of Gibbs measures, it follows that we can regard each of these events as an increasing event in $\{+,-\}^{S_m}$. Now, apply Proposition 7.7 and Remark 3. (The reader may notice that here we have essentially used the strong Markov property (see [18, p. 1154]).)

Thus, (8.4) is nonpositive. This, together with the fact that $\pi_{\beta,S_n}^{\omega_q}(\theta(\mathbf{0}) = q|\mathrm{T}_{\partial S_n}^{-\star}) = 0$, yields:

$$0 \leq \pi_n^{\beta}(\omega_q) - \pi_{y,z}^{\beta}(\omega_q) \leq \pi_{\beta,S_{y,z}}^{\omega_q}(\mathrm{T}_{\partial S_n}^{-\star})\pi_{\beta,S_n}^{\omega_q}(\theta(\mathbf{0}) = q) \leq \pi_{\beta,S_{y,z}}^{\omega_q}(\mathrm{T}_{\partial S_n}^{-\star}). \tag{8.6}$$

So, it suffices to show that $\sup_{y,z \geq \mathbf{1}n} \pi_{\beta,S_{y,z}}^{\omega_q}(\mathrm{T}_{\partial S_n}^{-\star})$ decays exponentially in n. Fix $y,z \geq \mathbf{1}n$ and let $m > n$ such that $\mathbf{1}m \geq y,z$. By Proposition 7.7 (parts 2 and 3) and Remark 3,

$$\pi_{\beta,S_{y,z}}^{\omega_q}(\mathrm{T}_{\partial S_n}^{-\star}) \leq \pi_{\beta,S_m}^{\omega_q}(\mathrm{T}_{\partial S_n}^{-\star}) \tag{8.7}$$

$$= \pi_{\beta,B_m}^{\omega_q}(\mathrm{T}_{\partial S_n}^{-\star}|q^{\mathscr{P}}) \leq \pi_{\beta,B_m}^{\omega_q}(\mathrm{T}_{\partial S_n}^{-\star}) \leq \pi_{\beta,B_m}^{\omega_q}(\mathrm{T}_{\partial B_n}^{-\star}). \tag{8.8}$$

So, it suffices to show that $\sup_{m>n} \pi_{\beta,B_m}^{\omega_q}(\mathrm{T}_{\partial B_n}^{-\star})$ decays exponentially in n. Recall the Edwards-Sokal coupling $\mathbb{P}_{p,q,B_m}^{(1)}$ for the Gibbs distribution and the corresponding bond random-cluster measure with wired boundary condition $\phi_{p,q,B_m}^{(1)}$ (see Section 5.1).

W.l.o.g., let's suppose that n is even, i.e. $n = 2k < m$, for some $k \in \mathbb{N}$. We consider the following two events in the bond random-cluster model, as in [11, Theorem 3]. Let $\mathcal{R}_n$ be the event of an open circuit in $B_{2k} \setminus B_k$ that surrounds B_k. Let $\mathcal{M}_{n,m}$ be the event in which there is an open path from some site in B_k to ∂B_m. The joint occurrence of these two events forces the Potts event $(\mathrm{T}_{\partial B_n}^{-\star})^c$ in the coupling: $\mathcal{R}_n \cap \mathcal{M}_{n,m} \subseteq (\mathrm{T}_{\partial B_n}^{-\star})^c$ (here, technically, we are identifying these events with their inverse images of the projections in the coupling).

Then, by the coupling property:

$$\pi_{\beta,B_m}^{\omega_q}\left((\mathrm{T}_{\partial B_n}^{-\star})^c\right) = \mathbb{P}_{p,q,B_m}^{(1)}\left((\mathrm{T}_{\partial B_n}^{-\star})^c\right) \tag{8.9}$$

$$\geq \mathbb{P}_{p,q,B_m}^{(1)}\left((\mathrm{T}_{\partial B_n}^{-\star})^c\Big|\mathcal{R}_n \cap \mathcal{M}_{n,m}\right)\mathbb{P}_{p,q,B_m}^{(1)}\left(\mathcal{R}_n \cap \mathcal{M}_{n,m}\right) \tag{8.10}$$

$$= \phi_{p,q,B_m}^{(1)}\left(\mathcal{R}_n \cap \mathcal{M}_{n,m}\right), \tag{8.11}$$

so:

$$\pi_{\beta,B_m}^{\omega_q}(\mathrm{T}_{\partial B_n}^{-\star}) \leq 1 - \phi_{p,q,B_m}^{(1)}(\mathcal{R}_n \cap \mathcal{M}_{n,m}) \leq \phi_{p,q,B_m}^{(1)}(\mathcal{R}_n^c) + \phi_{p,q,B_m}^{(1)}(\mathcal{M}_{n,m}^c). \tag{8.12}$$

Therefore,

$$\sup_{m>n} \pi^{\omega_q}_{\beta,\mathrm{B}_m}(\mathrm{T}^{-\star}_{\partial \mathrm{B}_n}) \leq \sup_{m>n} \phi^{(1)}_{p,q,\mathrm{B}_m}(\mathcal{R}^{\mathrm{c}}_n) + \sup_{m>n} \phi^{(1)}_{p,q,\mathrm{B}_m}(\mathcal{M}^{\mathrm{c}}_{n,m}). \qquad (8.13)$$

The first term on the right hand side of (8.13) is bounded from above as follows:

$$\phi^{(1)}_{p,q,\mathrm{B}_m}(\mathcal{R}^{\mathrm{c}}_n) \leq \phi^{(1)}_{p,q,\tilde{\mathrm{B}}_{m+1}}(\mathcal{R}^{\mathrm{c}}_n) \qquad (8.14)$$

$$\leq \sum_{x \in \partial \mathrm{B}_k, y \in \underline{\partial}\mathrm{B}_{2k}} \phi^{(0)}_{p^*,q,\mathrm{B}_{m+1}}(x \leftrightarrow y) \qquad (8.15)$$

$$\leq \sum_{x \in \partial \mathrm{B}_k, y \in \underline{\partial}\mathrm{B}_{2k}} \phi_{p^*,q}(x \leftrightarrow y), \qquad (8.16)$$

where $\tilde{\mathrm{B}}_m = [-m+1,m]^2 \cap \mathbb{Z}^2$ and p^* denotes the dual of p and the inequalities follow from Proposition 5.3 and Theorem 7.4.

If $p > p_{\mathrm{c}}(q)$, then $p^* < p_{\mathrm{c}}(q)$, and by Theorem 7.5, the first term on the right side of (8.13) is upper bounded by $64C(p^*,q)n^2 \exp(-c(p^*,q)n/4)$, since $|\partial \mathrm{B}_k||\underline{\partial}\mathrm{B}_{2k}| \leq 64n^2$ and $\|x-y\|_2 \geq k-1 \geq \frac{n}{4}$, for all $x \in \partial \mathrm{B}_k$ and $y \in \underline{\partial}\mathrm{B}_{2k}$. So, the first term on the right side of (8.13) decays exponentially.

As for the second term, in order for $\mathcal{M}_{n,m}$ to fail to occur, there must be a closed circuit in $\mathrm{B}_m \setminus \mathrm{B}_k$ and in particular a closed path from $L_{m,n} := \mathrm{B}_m \setminus \mathrm{B}_k \cap \{x_1 < 0, x_2 = 0\}$ to $R_{m,n} := \mathrm{B}_m \setminus \mathrm{B}_k \cap \{x_1 > 0, x_2 = 0\}$ in B_m. Thus,

$$\phi^{(1)}_{p,q,\mathrm{B}_m}(\mathcal{M}^{\mathrm{c}}_{n,m}) \leq \phi^{(1)}_{p,q,\tilde{\mathrm{B}}_{m+1}}(\mathcal{M}^{\mathrm{c}}_{n,m}) \qquad (8.17)$$

$$\leq \sum_{x \in L_{m,n}, y \in R_{m,n}} \phi^{(0)}_{p^*,q,\mathrm{B}_{m+1}}(x \leftrightarrow y) \qquad (8.18)$$

$$\leq \sum_{x \in L_{m,n}, y \in R_{m,n}} \phi_{p^*,q}(x \leftrightarrow y), \qquad (8.19)$$

where the last inequality follows by Proposition 5.3 and Proposition 7.4. By Theorem 7.5, this is less than:

$$\sum_{i=n, j=n} C(p^*,q)e^{-c(p^*,q)(i+j)} \leq C(p^*,q)\left(e^{-c(p^*,q)n}\frac{1}{1-e^{-c(p^*,q)}}\right)^2 \qquad (8.20)$$

$$= \frac{C(p^*,q)}{(1-e^{-c(p^*,q)})^2}e^{-2c(p^*,q)n}. \qquad (8.21)$$

Thus, the 2nd term on the right side of (8.13) also decays exponentially. Thus, $\sup_{m>n} \pi^{\omega_q}_{\beta,m}(\mathrm{T}^{-\star}_{\partial B_n})$ decays exponentially in n. Thus, by (8.7) $\sup_{m>n} \pi^{\omega_q}_{\beta,m}(\mathrm{T}^{-\star}_{\partial S_n})$ also decays exponentially in n, as desired.

Part II: $\beta < \beta_{\mathrm{c}}(q)$. Recall from Section 7 the notions of strong spatial mixing and ratio strong mixing property.

Theorem 8.2 ([3, Theorem 1.8 (ii)]). *For the $\mathbb{Z}^2$ Potts model with q types and inverse temperature β, if $0 < \beta < \beta_c(q)$ and exponential decay of the two-point connectivity function holds for the corresponding random-cluster model, then the (unique) Potts Gibbs measure satisfies the ratio strong mixing property for the class of finite simply lattice-connected sets.*

Corollary 1. *For the $\mathbb{Z}^2$ Potts model with q types and inverse temperature $0 < \beta < \beta_c(q)$, the specification π_β^{FP} satisfies exponential SSM for the family of sets $\{S_{y,z}\}_{y,z \geq 0}$.*

Proof. This follows immediately from Theorem 7.5, Theorem 8.2 and Proposition 7.1. ☐

Then, since exponential SSM holds for the class of finite simply lattice-connected sets when $\beta < \beta_c(q)$, the desired result follows directly from Proposition 7.2.

This completes the proof of Theorem 8.1. ☐

8.2 Exponential convergence in the Widom-Rowlinson model

Recall that for Bernoulli site percolation in $\mathbb{Z}^2$ there exists a probability parameter p_c, known as the *percolation threshold*, such that for $p < p_c$, there is no infinite cluster of 1's $\psi_{p,\mathbb{Z}^2}$-almost surely and for $p > p_c$, there is such a cluster $\psi_{p,\mathbb{Z}^2}$-almost surely. Similarly, one can define an analogous parameter $p_c^\star$ for the lattice $\mathbb{Z}^{2,\star}$, which satisfies $p_c + p_c^\star = 1$ (see [37]).

Theorem 8.3. *For the Widom-Rowlinson model with q types and activity λ, there exist two critical parameters $0 < \lambda_1(q) < \lambda_2(q)$ such that for $\lambda < \lambda_1(q)$ or $\lambda > \lambda_2(q)$, there exists $C, \alpha > 0$ such that, for every $y, z \geq \mathbf{1}n$:*

$$\left| \pi_n^\lambda(\omega_q) - \pi_{y,z}^\lambda(\omega_q) \right| \leq C e^{-\alpha n}. \tag{8.22}$$

Proof. As in the proof of Theorem 8.1, we split the proof in two parts.

Part I: $\lambda > \lambda_2(q) := q^3 \left(\frac{p_c}{1-p_c} \right)$. Fix $n \in \mathbb{N}$ and $y, z \geq \mathbf{1}n$. Notice that, due to the constraints of the Widom-Rowlinson model, and recalling Proposition 5.4:

$$\pi_{y,z}^\lambda(\omega_q) = \pi_{\lambda,S_{y,z}}^{\omega_q}(\theta(\mathbf{0}) = q) = \psi_{p,q,S_{y,z}}^{(1)}(\theta(\mathbf{0}) = 1), \tag{8.23}$$

where $p = \frac{\lambda}{1+\lambda}$, and the same holds for $\pi_n^\lambda(\omega_q)$. Then, it suffices to prove that:

$$\left| \psi_{p,q,S_n}^{(1)}(\theta(\mathbf{0}) = 1) - \psi_{p,q,S_{y,z}}^{(1)}(\theta(\mathbf{0}) = 1) \right| \leq C e^{-\alpha n}, \tag{8.24}$$

for some $C, \alpha > 0$.

Notice that $\mathbf{0} \in S_n \subseteq S_{y,z} =: \Lambda$. Fix any ordering on the set $\overline{\Lambda}$. From now on, when we talk about comparing sites in $\overline{\Lambda}$, it is assumed we are speaking of this ordering. For convenience, we will extend configurations on S_n and Λ to configurations on $\overline{\Lambda}$ by appending $1^{\overline{\Lambda} \setminus S_n}$ and $1^{\partial \Lambda}$, respectively.

Now, we will proceed to define a coupling $\mathbb{P}_{n,y,z}$ of $\psi_{p,q,S_n}^{(1)}$ and $\psi_{p,q,\Lambda}^{(1)}$, defined on pairs of configurations $(\theta_1, \theta_2) \in \{0,1\}^{\overline{\Lambda}} \times \{0,1\}^{\overline{\Lambda}}$. The coupling is defined one site at a time, using values from previously defined sites.

We use (τ_1^t, τ_2^t) to denote the (incomplete) configurations on $\overline{\Lambda} \times \overline{\Lambda}$ at step $t = 0, 1, \ldots, |S_n|$. We therefore begin with $\tau_1^0 = 1^{\overline{\Lambda} \setminus S_n}$ and $\tau_2^0 = 1^{\partial \Lambda}$. Next, we set $\tau_1^1 = \tau_1^0$ and form τ_2^1 by extending τ_2^0 to $\overline{\Lambda} \setminus S_n$, choosing randomly according to the distribution $\psi_{p,q,\Lambda}^{(1)}\left(\cdot \,\middle|\, 1^{\partial \Lambda}\right)$. At this point of the construction, both τ_1^1 and τ_2^1 have shape $\overline{\Lambda} \setminus S_n$. In the end, $(\tau_1^{|S_n|}, \tau_2^{|S_n|})$ will give as a result a pair (θ_1, θ_2).

At any step t, we use W^t to denote the set of sites in $\overline{\Lambda}$ on which τ_1^t and τ_2^t have already received values in previous steps. In particular, $W^1 = \overline{\Lambda} \setminus S_n$. At an arbitrary step t of the construction, we choose the next site x^{t+1} on which to assign values in τ_1^{t+1} and τ_2^{t+1} as follows:

(i) If possible, take x^{t+1} to be the smallest site in $\partial^\star W^t$ that is $\star$-adjacent to a site $y \in W^t$ for which $(\tau_1^t(y), \tau_2^t(y)) \neq (1,1)$.
(ii) Otherwise, just take x^{t+1} to be the smallest site in $\partial^\star W^t$.

Notice that at any step t, W^t is a $\star$-connected set, and that it it always possible to find the next site x^{t+1} for any $t < |S_n|$ (i.e. the two rules above give a well defined procedure).

Now we are ready to augment the coupling from W^t to $W^t \cup \{x^{t+1}\}$ by assigning $\tau_1^{t+1}(x^{t+1})$ and $\tau_2^{t+1}(x^{t+1})$ according to an optimal coupling of $\psi_{p,q,S_n}^{(1)}\left(\cdot \,\middle|\, \tau_1^t\right)\Big|_{\{x^{t+1}\}}$ and $\psi_{p,q,S_{y,z}}^{(1)}\left(\cdot \mid \tau_2^t\right)\Big|_{\{x^{t+1}\}}$, i.e. a coupling which minimizes the probability that, given (τ_1^t, τ_2^t), $\theta_1(x^{t+1}) \neq \theta_2(x^{t+1})$. Since $\mathbb{P}_{n,y,z}$ is defined site-wise, and at each step is assigned according to $\psi_{p,q,S_n}^{(1)}\left(\cdot | \tau_1^t\right)$ in the first coordinate and $\psi_{p,q,S_{y,z}}^{(1)}\left(\cdot \mid \tau_2^t\right)$ in the second, the reader may check that it is indeed a coupling of $\psi_{p,q,S_n}^{(1)}$ and $\psi_{p,q,S_{y,z}}^{(1)}$. The key property of $\mathbb{P}_{n,y,z}$ is the following.

Lemma 8.4. $\theta_1(\mathbf{0}) \neq \theta_2(\mathbf{0})$ $\mathbb{P}_{n,y,z}$-a.s. if and only if there exists a path T of $\star$-adjacent sites from $\mathbf{0}$ to ∂S_n, such that for each site $y \in \mathrm{T}$, $(\theta_1(y), \theta_2(y)) \neq (1,1)$.

Proof. Suppose, for a contradiction, that $\theta_1(\mathbf{0}) \neq \theta_2(\mathbf{0})$ and there exists no such path. This implies that there exists a circuit C surrounding $\mathbf{0}$ (when we include the bottom boundary as part of C) and contained in $\overline{S}_n$ such that for all $y \in \mathrm{C}$, $(\theta_1(y), \theta_2(y)) = (1,1)$. Define by I the simply lattice-$\star$-connected set of sites in the interior of C and, let's say that at time t_0, x^{t_0} was the first site within I defined according to the site-by-site evolution of $\mathbb{P}_{n,y,z}$. Then, $(\tau_1^{t_0}(x^{t_0}), \tau_2^{t_0}(x^{t_0}))$ cannot have been defined according to rule (i) since all sites $\star$-adjacent to x^{t_0} are either in I (and therefore not yet defined by definition of x^{t_0}), or on C (and therefore either not yet defined or sites at which θ_1 and θ_2 are both 1).

Therefore, $(\theta_1(x^{t_0}), \theta_2(x^{t_0}))$ was defined according to rule (ii). We therefore define the set $D := \overline{\Lambda} \setminus W^{t_0 - 1} \supseteq I$, and note that $\mathbf{0}$ and x^{t_0} belong to the same $\star$-connected component Θ of D. We also know that $\tau_1^{t_0 - 1}(\partial^\star D) = \tau_2^{t_0 - 1}(\partial^\star D) = 1^{\partial^\star D}$, otherwise some unassigned site in D would be $\star$-adjacent to a 0 in either $\tau_1^{t_0 - 1}(\partial^\star D)$

or $\tau_2^{t_0-1}(\partial^\star D)$, and so rule (i) would be applied instead. We may now apply Lemma 5.5 (combined with Remark 1) to Θ and Λ in order to see that $\psi_{p,q,S_n}^{(1)}(\theta_1(\Theta)|\tau_1^{t_0-1})$ and $\psi_{p,q,S_{y,z}}^{(1)}(\theta_2(\Theta)|\tau_2^{t_0-1})$ are identical. This means that the optimal coupling according to which $\tau_1^{t_0}(x^{t_0})$ and $\tau_2^{t_0}(x^{t_0})$ are assigned is supported on the diagonal, and so $\tau_1^{t_0}(x^{t_0}) = \tau_2^{t_0}(x^{t_0})$, $\mathbb{P}_{n,y,z}$-almost surely. This will not change the conditions under which we applied Lemma 5.5, and so inductively, the same will be true for each site in I as it is assigned, including $\mathbf{0}$. We have shown that $\theta_1(\mathbf{0}) = \theta_2(\mathbf{0})$, $\mathbb{P}_{n,y,z}$-almost surely, regardless of when $\mathbf{0}$ is assigned in the site-by-site evolution of $\mathbb{P}_{n,y,z}$. This is a contradiction, and so our original assumption was incorrect, implying that the desired path T exists. $\qquad\square$

Given an arbitrary time t, let:

$$\rho_1^t(\cdot) := \psi_{p,q,S_n}^{(1)}(\cdot|\tau_i^{t-1})\Big|_{\{x^t\}} \quad \text{and} \quad \rho_2^t(\cdot) := \psi_{p,q,\Lambda}^{(1)}(\cdot|\tau_i^{t-1})\Big|_{\{x^t\}} \tag{8.25}$$

be the two corresponding probability measures defined on the set $\{0,1\}^{\{x^t\}}$. Note that at any step within the site-by-site definition of $\mathbb{P}_{n,y,z}$, Lemma 7.6 implies that $\frac{\lambda}{\lambda+q^3} \leq \rho_i^t(1)$, where $\lambda = \frac{p}{1-p}$ and $i = 1,2$. Now, w.l.o.g., suppose that $\rho_2^t(0) \geq \rho_1^t(0)$. Then, an optimal coupling $\mathbb{Q}^t$ of ρ_1^t and ρ_2^t will assign $\mathbb{Q}^t(\{(0,0)\}) = \rho_1^t(0)$, $\mathbb{Q}^t(\{(0,1)\}) = 0$, $\mathbb{Q}^t(\{(1,0)\}) = \rho_2^t(0) - \rho_1^t(0)$, and $\mathbb{Q}^t(\{(1,1)\}) = 1 - \rho_2^t(0)$. Therefore,

$$\mathbb{Q}^t(\{(1,1)\}^c) = \rho_2^t(0) \leq \frac{q^3}{\lambda+q^3}. \tag{8.26}$$

Next, define the map $h : \{0,1\}^{S_n} \times \{0,1\}^{S_n} \to \{0,1\}^{S_n}$ given by:

$$(h(\theta_1,\theta_2))(x) = \begin{cases} 1 & \text{if } (\theta_1(x),\theta_2(x)) \neq (1,1), \\ 0 & \text{if } (\theta_1(x),\theta_2(x)) = (1,1). \end{cases} \tag{8.27}$$

By (8.26), $h_*\mathbb{P}_{n,y,z}$ (the push-forward measure) can be coupled against an i.i.d. measure on $\{0,1\}^{S_n}$ which assigns 1 with probability $\frac{q^3}{\lambda+q^3}$ and 0 with probability $\frac{\lambda}{\lambda+q^3}$, and that the former is stochastically dominated by the latter. This, together with Lemma 8.4, yields

$$\left| \psi_{p,q,S_n}^{(1)}(\theta(\mathbf{0}) = 1) - \psi_{p,q,S_{y,z}}^{(1)}(\theta(\mathbf{0}) = 1) \right| \leq \mathbb{P}_{n,y,z}(\theta_1(\mathbf{0}) \neq \theta_2(\mathbf{0})) \tag{8.28}$$

$$\leq \psi_{\frac{q^3}{\lambda+q^3},S_n}(\mathbf{0} \overset{\star}{\leftrightarrow} \partial S_n), \tag{8.29}$$

Since we have assumed $\lambda > q^3\left(\frac{p_c}{1-p_c}\right)$ and $p_c + p_c^\star = 1$, we have $\frac{q^3}{\lambda+q^3} < p_c^\star$. It follows by [1, 32] that the expression in (8.29) decays exponentially in n. This completes the proof.

Part II: $\lambda < \lambda_1(q) := \frac{1}{q}\left(\frac{p_c}{1-p_c}\right)$**.** Observe that, by virtue of Proposition 7.2, it suffices to prove that $\pi_\lambda^{\mathrm{WR}}$ satisfies exponential SSM. For this, we use Theorem 7.3.

By considering all cases of nearest-neighbour configurations at the origin, one can compute:

$$Q(\pi_\lambda^{\mathrm{WR}}) = \max_{\omega_1,\omega_2\in\Omega(\mathscr{E})} d_{TV}(\pi_{\lambda,\{\mathbf{0}\}}^{\omega_1},\pi_{\lambda,\{\mathbf{0}\}}^{\omega_2}) = \frac{q\lambda}{1+q\lambda}. \tag{8.30}$$

By Theorem 7.3, we obtain exponential SSM when:

$$\lambda < \frac{1}{q}\left(\frac{p_c}{1-p_c}\right) = \lambda_1(q). \tag{8.31}$$

Uniqueness of Gibbs states in this same region was mentioned in [19, p. 40], by appealing to [7, Theorem 1] (which is the crux of Theorem 7.3).   □

Remark 4. *In the case $q = 2$, it is possible to give an alternative proof of Theorem 8.3, Part I, using the framework of the proof of Theorem 8.1, Part I. The arguments through (8.7) go through, with an appropriate re-definition of events and use of Proposition 7.8 for stochastic dominance. One can then apply Lemma 7.6 to give estimates based on the site random-cluster model. (In contrast to Theorem 8.1, Part I, this does not require the use of planar duality). So far, this approach is limited to $q = 2$ because we do not know appropriate versions of Proposition 7.8 for $q > 2$.*

8.3 Exponential convergence in the hard-core model

Our argument again relies on proving exponential convergence for conditional measures with respect to certain "extremal" boundaries on S_n, but these now will consist of alternating 0 and 1 symbols rather than a single symbol (recall from Section 4.3 that $\omega^{(o)}$ is defined as the configuration of 1's on all even sites and 0 on all odd sites).

Theorem 8.5. *For the $\mathbb{Z}^2$ hard-core model with activity γ, there exist two critical parameters $0 < \gamma_1 < \gamma_2$ such that for any $0 < \gamma < \gamma_1$ or $\gamma > \gamma_2$, there exist $C,\alpha > 0$ such that for every $y,z \geq \mathbf{1}n$,*

$$\left|\pi_n^\gamma(\omega^{(o)}) - \pi_{y,z}^\gamma(\omega^{(o)})\right| \leq Ce^{-\alpha n}. \tag{8.32}$$

Proof. As in the previous two theorems, we consider two cases.

Part I: $\gamma > \gamma_2 := 468.$ Our proof essentially combines the disagreement percolation techniques of [7] and the proof of non-uniqueness of equilibrium state for the hard-core model due to Dobrushin (see [13]). We need enough details not technically contained in either proof that we present a mostly self-contained argument here. From [7, Theorem 1] and an averaging argument (as in the proof of Proposition 7.2) on $\partial_\uparrow S_n$ induced by a boundary condition on $S_{y,z}$, we know that for any $y,z \geq \mathbf{1}n$,

$$\left|\pi_n^\gamma(\omega^{(o)}) - \pi_{y,z}^\gamma(\omega^{(o)})\right| \leq \mathbb{P}_{n,y,z}\left(\exists \text{ a path of disagreement from } \mathbf{0} \leftrightarrow \partial_\uparrow S_n\right) \tag{8.33}$$

for a certain coupling $\mathbb{P}_{n,y,z}$ of $\pi_{\gamma,S_n}^{\omega^{(o)}}$ and $\pi_{\gamma,S_{y,z}}^{\omega^{(o)}}\big|_{S_n}$. We do not need the structure of $\mathbb{P}_{n,y,z}$ here, but instead note the following: a path of disagreement for the boundaries

$\omega^{(o)}(\partial S_n)$ and $\omega^{(o)}(\partial S_{y,z})$ implies that in one of the configurations, all entries on the path will be "out of phase" with respect to $\omega^{(o)}$, i.e. that all entries along the path will have 1 at every odd site and 0 at every even site rather than the opposite alternating pattern of $\omega^{(o)}$. Then, if we denote by $\mathscr{T}_n$ the event that there is a path T from $\mathbf{0} \leftrightarrow \partial_\uparrow S_n$ with 1 at every odd site and 0 at every even site, it is clear that:

$$\mathbb{P}_{n,y,z}\big(\exists \text{ a path of disagreement from } \mathbf{0} \leftrightarrow \partial_\uparrow S_n\big) \leq \pi_{\gamma,S_n}^{\omega^{(o)}}(\mathscr{T}_n) + \pi_{\gamma,S_{y,z}}^{\omega^{(o)}}(\mathscr{T}_n). \quad (8.34)$$

Since $y,z \geq \mathbf{1}n$ are arbitrary (in particular, y and z can be chosen to be $\mathbf{1}n$), it suffices to prove that $\sup_{y,z \geq \mathbf{1}n} \pi_{\gamma,S_{y,z}}^{\omega^{(o)}}(\mathscr{T}_n)$ decays exponentially with n. Define the set:

$$\Theta_{y,z} = \{\theta \in \{0,1\}^{\overline{S}_{y,z}^\star} : \theta \text{ is feasible and } \theta(\partial^\star S_{y,z}) = \omega^{(o)}(\partial^\star S_{y,z})\}. \quad (8.35)$$

For any $\theta \in \Theta_{y,z}$, we define $\Sigma_\mathbf{0}(\theta)$ to be the connected component of $\Sigma_{S_{y,z}}(\theta,\omega^{(o)})$ $(= \{x \in S_{y,z} : \theta(x) \neq \omega^{(o)}(x)\})$ containing the origin $\mathbf{0}$. Since $\mathscr{T}_n \subseteq \{\Sigma_\mathbf{0}(\theta) \cap \partial_\uparrow S_n \neq \emptyset\}$, our proof will then be complete if we can show that there exist $C, \alpha > 0$ so that for any n and $y,z \geq \mathbf{1}n$, the following holds:

$$\pi_{\gamma,S_{y,z}}^{\omega^{(o)}}(\Sigma_\mathbf{0}(\theta) \cap \partial_\uparrow S_n \neq \emptyset) \leq C e^{-\alpha n}. \quad (8.36)$$

To prove this, we use a Peierls argument, similar to [13].

Fix any $y,z \geq \mathbf{1}n$ and for any $\theta \in \Theta_{y,z}$, define $\Sigma_\mathbf{0}(\theta)$ as above, and let $K(\theta)$ to be the connected component of $\{x \in \overline{S}_{y,z}^\star : \theta(x) = \omega^{(o)}(x)\}$ containing $\partial^\star S_{y,z}$. Clearly, $\Sigma_\mathbf{0}(\theta)$ and $K(\theta)$ are disjoint, $K(\theta) \neq \emptyset$ and, provided $\theta(\mathbf{0}) = 0$, $\Sigma_\mathbf{0}(\theta) \neq \emptyset$. Then, define $\Gamma(\theta) := \Sigma_\mathbf{0}(\theta) \cap \partial K(\theta) \subseteq S_{y,z}$. We note that for any $\theta \in \Theta_{y,z}$ with $\theta(\mathbf{0}) = 0$, we have that $\theta(\Gamma(\theta)) = 0^{\Gamma(\theta)}$, since adjacent sites in $\Sigma_\mathbf{0}(\theta)$ and $K(\theta)$ must have the same letter by definition of $\Sigma_\mathbf{0}(\theta)$, and adjacent 1 symbols are forbidden in the hard-core model. Therefore, every $x \in \Gamma(\theta)$ is even.

We need the concept of *inner external boundary* for a connected set $\Sigma \Subset \mathbb{Z}^2$. The inner external boundary of Σ is defined to be the inner boundary of the simply lattice-connected set consisting of the union of Σ and the union of all the finite components of $\mathbb{Z}^2 \setminus \Sigma$. Intuitively, the inner external boundary of Σ is the inner boundary of the set Σ obtained after "filling in the holes" of Σ. Notice that the set $\Gamma(\theta)$ corresponds exactly to the inner external boundary of $\Sigma_\mathbf{0}(\theta)$. In addition, by [12, Lemma 2.1 (i)], we know that the inner external boundary of a finite connected set (more generally a finite $\star$-connected set) is $\star$-connected. Thus, $\Gamma(\theta) \subseteq S_{y,z}$ is a $\star$-connected set $C^\star$ that consists only of even sites and contains the origin $\mathbf{0}$, for any $\theta \in \Theta_{y,z}$ with $\theta(\mathbf{0}) = 0$.

Then, for $C^\star \subseteq S_{y,z}$, we define the event $E_{C^\star} := \{\theta \in \Theta_{y,z} : \Gamma(\theta) = C^\star\}$, and will bound from above $\pi_{\gamma,S_{y,z}}^{\omega^{(o)}}(E_{C^\star})$, for every $C^\star$ such that $E_{C^\star}$ is nonempty. We make some more notation: for every such a set $C^\star$, define $O(C^\star)$ (for 'outside') as the connected component of $(C^\star)^c$ containing $\partial^\star S_{y,z}$, and define $I(C^\star)$ (for 'inside') as $S_{y,z} \setminus (C^\star \cup O(C^\star))$. Then $C^\star$, $I(C^\star)$, and $O(C^\star)$ form a partition of $\overline{S}_{y,z}^\star$. We note that there cannot be a pair of adjacent sites from $I(C^\star)$ and $O(C^\star)$ respectively, since they would then be in the same connected component of $(C^\star)^c$. We also note that for every $\theta \in E_{C^\star}$, $C^\star \subseteq \Sigma_\mathbf{0}(\theta) \subseteq C^\star \cup I(C^\star)$ and $K(\theta) \subseteq O(C^\star)$ though the sets need not be

equal, since $\Sigma_{\mathbf{0}}(\theta)$ or $K(\theta)$ could contain "holes" which are "filled in" in $I(\mathrm{C}^\star)$ and $O(\mathrm{C}^\star)$, respectively.

Choose any set $\mathrm{C}^\star$ such that $E_{\mathrm{C}^\star} \neq \emptyset$. For each $\theta \in E_{\mathrm{C}^\star}$ and $x \in \mathrm{C}^\star$, using the definition of $\mathrm{C}^\star$ and the fact that $K(\theta) \subseteq O(\mathrm{C}^\star)$, there exists $x_0 \in \{e_1, -e_1, e_2, -e_2\}$ for which $x - x_0 \in O(\mathrm{C}^\star)$. Fix an x_0 which is associated to at least $|\mathrm{C}^\star|/4$ of the sites in $\mathrm{C}^\star$ in this way. Then, we define a function $s : E_{\mathrm{C}^\star} \to \{0,1\}^{\overline{S}^\star_{y,z}}$ that, given $\theta \in E_{\mathrm{C}^\star}$, defines a new configuration $s(\theta)$ as follows:

$$(s(\theta))(x) = \begin{cases} \theta(x - x_0) & \text{if } x \in I(\mathrm{C}^\star), \\ \theta(x) & \text{if } x \in O(\mathrm{C}^\star), \\ 1 & \text{if } x \in \mathrm{C}^\star \text{ and } x - x_0 \in O(\mathrm{C}^\star), \\ 0 & \text{if } x \in \mathrm{C}^\star \text{ and } x - x_0 \in I(\mathrm{C}^\star). \end{cases} \tag{8.37}$$

Informally, we move all 1 symbols inside $I(\mathrm{C}^\star)$ in the x_0-direction by 1 unit (even if those symbols were not part of $\Sigma_{\mathbf{0}}(\theta)$), add new 1 symbols at some sites in $\mathrm{C}^\star$, and leave everything in $O(\mathrm{C}^\star)$ unchanged.
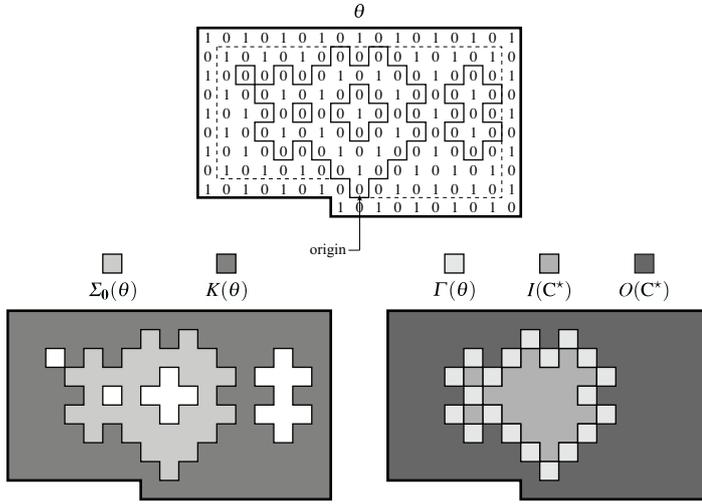
Fig. 2 A configuration $\theta \in E_{\mathrm{C}^\star}$. On the left, the associated sets $\Sigma_{\mathbf{0}}(\theta)$ and $K(\theta)$. On the right, the sets $I(\mathrm{C}^\star)$ and $O(\mathrm{C}^\star)$ for $\Gamma(\theta) = \mathrm{C}^\star$.

It should be clear that $s(\theta)$ has at least $|\mathrm{C}^\star|/4$ more 1 symbols than θ did. We make the following two claims: s is injective on $E_{\mathrm{C}^\star}$, and for every $\theta \in E_{\mathrm{C}^\star}$, $s(\theta) \in \Theta_{y,z}$. If these claims are true, then clearly $\pi^{\omega^{(o)}}_{\gamma, S_{y,z}}(s(E_{\mathrm{C}^\star})) \geq \gamma^{|\mathrm{C}^\star|/4} \pi^{\omega^{(o)}}_{\gamma, S_{y,z}}(E_{\mathrm{C}^\star})$, implying that:

$$\pi^{\omega^{(o)}}_{\gamma, S_{y,z}}(E_{\mathrm{C}^\star}) \leq \gamma^{-|\mathrm{C}^\star|/4}. \tag{8.38}$$

Firstly, we show that s is injective. Suppose that $\theta_1 \neq \theta_2$, for $\theta_1, \theta_2 \in E_{\mathrm{C}^\star}$. Then there is a site x at which $\theta_1(x) \neq \theta_2(x)$. If $x \in O(\mathrm{C}^\star)$, then $(s(\theta_1))(x) = \theta_1(x) \neq$

$\theta_2(x) = (s(\theta_2))(x)$ and so $s(\theta_1) \neq s(\theta_2)$. If $x \in I(C^\star)$, then $(s(\theta_1))(x+x_0) = \theta_1(x) \neq \theta_2(x) = (s(\theta_2))(x+x_0)$, and again $s(\theta_1) \neq s(\theta_2)$. Finally, we note that x cannot be in $C^\star$, since at all sites in $C^\star$, both θ_1 and θ_2 must have 0 symbols.

Secondly, we show that for any $\theta \in E_{C^\star}$, $s(\theta)$ is feasible. All that must be shown is that $s(\theta)$ does not contain adjacent 1 symbols. We break 1 symbols in $s(\theta)$ into three categories:

1. *shifted*, meaning that the 1 symbol came from shifting a 1 symbol at a site in $I(C^\star)$ in the x_0-direction,
2. *new*, meaning that the 1 symbol was placed at a site $x \in C^\star$ such that $x - x_0 \in O(C^\star)$, or
3. *untouched*, meaning that the 1 symbol was at a site in $O(C^\star)$ ($\supseteq \partial^\star S_{y,z}$).

Note that untouched 1 symbols cannot be adjacent to $C^\star$: θ contains all 0 symbols on $C^\star$, and so since $C^\star \subseteq \Sigma_0(\theta)$, a 1 symbol adjacent to a symbol in $C^\star$ would be in $\Sigma_0(\theta)$ as well, a contradiction since $\Sigma_0(\theta) \subseteq C^\star \cup I(C^\star)$, and so $\Sigma_0(\theta)$ and $O(C^\star)$ are disjoint.

Clearly shifted 1 symbols cannot be adjacent to each other, since there were no adjacent 1 symbols in θ. All new 1's were placed at sites in $C^\star$, and all sites in $C^\star$ are even, so new 1 symbols can't be adjacent to each other. Untouched 1's can't be adjacent for the same reason as shifted 1's. We now address the possibility of adjacent 1 symbols in $s(\theta)$ from different categories. A shifted or new 1 in $s(\theta)$ is at a site in $C^\star \cup I(C^\star)$, and an untouched 1 can't be adjacent to a site in $C^\star$ as explained above, and also cannot be adjacent to a site in $I(C^\star)$ since $I(C^\star)$ and $O(C^\star)$ do not contain adjacent sites. Therefore, shifted or new 1's can't be adjacent to untouched 1's. The only remaining case which we need to rule out is a new 1 adjacent to a shifted 1. Suppose that $(s(\theta))(x)$ is a new 1 and $(s(\theta))(x')$ is a shifted 1. Then by definition, $x' - x_0 \in I(C^\star)$ and $x - x_0 \in O(C^\star)$. We know that $I(C^\star)$ and $O(C^\star)$ do not contain adjacent sites, so $x - x_0$ and $x' - x_0$ are not adjacent, implying that x and x' are not adjacent. We've then shown that $s(\theta)$ is feasible and then, since $\partial^\star S_{y,z} \subseteq O(C^\star)$, $s(\theta) \in \Theta_{y,z}$, completing the proof of (8.38).

Recall that every set $C^\star$ which we are considering is $\star$-connected, occupies only even sites, and contains the origin $\mathbf{0}$. Then, given $k \in \mathbb{N}$, it is direct to see that the number of such $C^\star$ with $|C^\star| = k$ is less than or equal to $k \cdot t(k)$, where $t(k)$ denotes the number of *site animals* (see [25] for the definition) of size k (the first k factor comes from the fact that site animals are defined up to translation, and here given a site animal of size k, exactly k translations of it will contain the origin $\mathbf{0}$). We know that for every $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that $t(k) \leq C_\varepsilon(\delta + \varepsilon)^k$ for every k, where $\delta := \lim_{k \to \infty}(t(k))^{1/k} \leq 4.649551$ (see [25]).

If $\Sigma_0(\theta) \cap \partial_\uparrow S_n \neq \emptyset$, then $\Sigma_0(\theta)$ has to intersect the left, top or right boundary of S_n. W.l.o.g., we may assume that $\Sigma_0(\theta)$ intersects the right boundary of S_n. Then, every vertical segment in the right half of S_n must intersect $\Sigma_0(\theta)$ and, therefore, at least one element of its inner external boundary, namely $\Gamma(\theta)$. Then:

$$\Sigma_0(\theta) \cap \partial_\uparrow S_n \neq \emptyset \implies |\Gamma(\theta)| \geq n. \tag{8.39}$$

Therefore, taking an arbitrary $\varepsilon > 0$, we may bound $\pi_{\gamma,S_{y,z}}^{\omega^{(o)}}(\Sigma_{\mathbf{0}}(\theta) \cap \partial_{\uparrow}S_n \neq \emptyset)$ from above:

$$\pi_{\gamma,S_{y,z}}^{\omega^{(o)}}(\Sigma_{\mathbf{0}}(\theta) \cap \partial_{\uparrow}S_n \neq \emptyset) \leq \sum_{C^{\star}:|C^{\star}|\geq n} \gamma^{-|C^{\star}|/4} \leq \sum_{k=n}^{\infty} kC_{\varepsilon}(\delta + \varepsilon)^k \cdot \gamma^{-k/4}, \qquad (8.40)$$

which decays exponentially in n as long as $\gamma > (\delta + \varepsilon)^4$, independently of y and z. Since ε was arbitrary, $\gamma > 468 > \delta^4$ suffices for justifying (8.36), completing the proof.

Part II: $\gamma < \gamma_1 := 2.48$. It is known (see [39]) that when $d = 2$ and $\gamma < 2.48$, $\pi_{\gamma}^{\mathrm{HC}}$ satisfies exponential SSM. Then, by applying Proposition 7.2, we conclude. $\qquad\square$

9 Poly-time approximation for pressure of $\mathbb{Z}^2$ lattice models

By a *poly-time approximation algorithm* to compute a number r, we mean an algorithm that, given $N \in \mathbb{N}$, produces an estimate r_N such that $|r - r_N| < \frac{1}{N}$ and the time to compute r_N is polynomial in N.

Theorem 9.1. *Let Φ be a n.n. interaction for a set of restrictions $\mathscr{E}$ and suppose that $\Omega(\mathscr{E})$ satisfies the square block D-condition. Let $\overline{\omega} \in \Omega(\mathscr{E})$ be a periodic point such that $c_{\pi}(v^{\overline{\omega}}) > 0$. In addition, suppose that there exists $C, \alpha > 0$ such that, for every $y, z \geq \mathbf{1}n$:*

$$\left|\pi_n(\omega) - \pi_{y,z}(\omega)\right| \leq Ce^{-\alpha n} \text{ over } \omega \in O(\overline{\omega}). \qquad (9.1)$$

Then:

$$P(\Phi) = \frac{1}{|O(\overline{\omega})|} \sum_{\omega \in O(\overline{\omega})} \hat{I}_{\pi}(\omega) + A_{\Phi}(\omega), \qquad (9.2)$$

and there is a poly-time approximation algorithm to compute $P(\Phi)$, when $d = 2$.

Proof. Notice that $\mathrm{supp}(v^{\overline{\omega}}) = O(\overline{\omega}) \subseteq \Omega(\mathscr{E})$, since $\Omega(\mathscr{E})$ is shift-invariant and $\overline{\omega} \in \Omega(\mathscr{E})$. Now, since $\left|\pi_n(\omega) - \pi_{y,z}(\omega)\right| \leq Ce^{-\alpha n}$ over $\omega \in \mathrm{supp}(v^{\overline{\omega}})$, we can easily conclude that $\lim_{y,z \to \infty} \pi_{y,z}(\omega) = \hat{\pi}(\omega)$ uniformly over $\omega \in \mathrm{supp}(v^{\overline{\omega}})$. This, combined with $\Omega(\mathscr{E})$ satisfying the square block D-condition and $c_{\pi}(v^{\overline{\omega}}) > 0$, gives us

$$P(\Phi) = \int \left(\hat{I}_{\pi} + A_{\Phi}\right) dv^{\overline{\omega}} = \frac{1}{|\mathrm{supp}(v^{\overline{\omega}})|} \sum_{\omega \in \mathrm{supp}(v^{\overline{\omega}})} \hat{I}_{\pi}(\omega) + A_{\Phi}(\omega), \qquad (9.3)$$

thanks to Theorem 6.3.

For the algorithm, it suffices to show that there is a poly-time algorithm to compute $\hat{\pi}(\omega)$, for any $\omega \in O(\overline{\omega})$.

By Equation 9.1, there exist $C, \alpha > 0$ such that $|\pi_n(\omega) - \hat{\pi}(\omega)| < Ce^{-\alpha n}$. Since $|\partial S_n|$ is linear in n when $d = 2$, by a modified transfer matrix approach (see [31, Lemma 4.8]), we can compute $\pi_n(\omega)$ in exponential time $Ke^{\rho n}$ for some $K, \rho > 0$. Combining the exponential time to compute $\pi_n(\omega)$ for the exponential decay of $|\pi_n(\omega) - \hat{\pi}(\omega)|$, we get a poly-time algorithm to compute $P(\Phi)$: namely, given $N \in$

$\mathbb{N}$, let n be the smallest integer such that $Ce^{-\alpha(n+1)} < \frac{1}{N}$. Then $\pi_{n+1}(\omega)$ is within $\frac{1}{N}$ of $\hat{\pi}(\omega)$ and since $\frac{1}{N} \leq Ce^{-\alpha n}$, the time to compute $\pi_{n+1}(\omega)$ is at most:

$$Ke^{\rho(n+1)} = (Ke^\rho C^{\rho/\alpha}) \frac{1}{(Ce^{-\alpha n})^{\rho/\alpha}} \leq (Ke^\rho C^{\rho/\alpha}) N^{\rho/\alpha}, \qquad (9.4)$$

which is a polynomial in N. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Corollary 2. *The following holds:*

1. *For the $\mathbb{Z}^2$ Potts model with q types and inverse temperature $\beta > 0$:*

$$\mathrm{P}(\Phi_\beta) = \hat{I}_\pi^\beta(\omega_q) + 2\beta. \qquad (9.5)$$

2. *For the $\mathbb{Z}^2$ Widom-Rowlinson model with q types and activity $\lambda \in (0, \lambda_1(q)) \cup (\lambda_2(q), \infty)$:*

$$\mathrm{P}(\Phi_\lambda) = \hat{I}_\pi^\lambda(\omega_q) + \log\lambda, \qquad (9.6)$$

   where $\lambda_1(q) := \frac{1}{q}\left(\frac{p_c}{1-p_c}\right)$ and $\lambda_2(q) := q^3\left(\frac{p_c}{1-p_c}\right)$.
3. *For the $\mathbb{Z}^2$ hard-core model with activity $\gamma \in (0, \gamma_1) \cup (\gamma_2, \infty)$:*

$$\mathrm{P}(\Phi_\gamma) = \frac{1}{2}\hat{I}_\pi^\gamma(\omega^{(o)}) + \frac{1}{2}\log\gamma, \qquad (9.7)$$

   where $\gamma_1 = 2.48$ and $\gamma_2 = 468$.

   Moreover, for the three models in the corresponding regions (except in the case when $\beta = \beta_c(q)$ in the Potts model), the pressure can be approximated in poly-time, where the polynomial involved depends on the parameters of the models.

Proof. The representation of the pressure given in the previous statement for the $\mathbb{Z}^2$ Potts model with q types and inverse temperature $\beta \neq \beta_c(q)$, the $\mathbb{Z}^2$ Widom-Rowlinson model with q types and activity $\lambda \in (0, \lambda_1(q)) \cup (\lambda_2(q), \infty)$ and the $\mathbb{Z}^2$ hard-core model with activity $\gamma \in (0, \gamma_1) \cup (\gamma_2, \infty)$, is a direct consequence of Theorem 9.1, by virtue of the following facts:

- Recall that the corresponding n.n. SFT $\Omega(\mathscr{E})$ for the Potts, Widom-Rowlinson and hard-core model has a safe symbol, respectively, so $\Omega(\mathscr{E})$ satisfies the square block D-condition and $c_\pi(\nu) > 0$, for any shift-invariant ν with $\mathrm{supp}(\nu) \subseteq \Omega(\mathscr{E})$, in each case.
- If we consider the delta-measure $\nu^{\omega_q} = \delta_{\omega_q}$, both in the Potts and Widom-Rowlinson cases (in a slight abuse of notation, since the Potts and Widom-Rowlinson σ-algebras are defined in different alphabets), and the measure $\nu = \nu^{\omega^{(o)}} = \frac{1}{2}\delta_{\omega^{(e)}} + \frac{1}{2}\delta_{\omega^{(o)}}$ in the hard-core case, we have that in all three models, for the range of parameters specified, except for when $\beta = \beta_c(q)$ in the Potts model, there exists $C, \alpha > 0$ such that, for every $y, z \geq \mathbf{1}n$:

$$\left|\pi_n(\omega) - \pi_{y,z}(\omega)\right| \leq Ce^{-\alpha n}, \text{ over } \omega \in \mathrm{supp}(\nu), \qquad (9.8)$$

thanks to Theorem 8.1, Theorem 8.3 and Theorem 8.5, respectively. (Notice that $\hat{I}_\pi^\gamma(\omega^{(e)}) = A_\Phi(\omega^{(e)}) = 0$.)

This proves (9.5), (9.6) and (9.7), except in the Potts case when $\beta = \beta_c$. To establish this case, first note that it is easy to prove that $P(\Phi_\beta)$ is continuous with respect to β. Second, if $\beta_1 \leq \beta_2$, then $\pi_n^{\beta_1}(\omega_q) \leq \pi_n^{\beta_2}(\omega_q)$. This follows by the Edwards-Sokal coupling (see Theorem 5.2) and the comparison inequalities for the bond random-cluster model [2, Theorem 4.1].

As an exercise in analysis, it is not difficult to prove that if $a_{m,n} \geq 0$, and each $a_{m+1,n} \leq a_{m,n}$ and $a_{m,n+1} \leq a_{m,n}$, then $\lim_m \lim_n a_{m,n} = \lim_n \lim_m a_{m,n} = a$, for some $a \geq 0$.

Now, consider the sequence $a_{m,n} := \pi_n^{\beta_c(q)+\frac{1}{m}}(\omega_q)$. By stochastic dominance (see Proposition 7.7), $a_{m,n}$ is decreasing in n. By the previous discussion (Edwards-Sokal coupling), it is also decreasing in m. Therefore, and since $a_{m,n} \geq 0$, we conclude that $\lim_m \lim_n a_{m,n} = \lim_n \lim_m a_{m,n} = a$, for some a.

Then, we have that:

$$P(\Phi_{\beta_c(q)}) = \lim_m P(\Phi_{\beta_c(q)+\frac{1}{m}}) \tag{9.9}$$

$$= \lim_m -\log \lim_n \pi_n^{\beta_c(q)+\frac{1}{m}}(\omega_q) + 2\left(\beta_c(q) + \frac{1}{m}\right) \tag{9.10}$$

$$= -\log \lim_m \lim_n \pi_n^{\beta_c(q)+\frac{1}{m}}(\omega_q) + 2\beta_c(q) \tag{9.11}$$

$$= -\log \lim_n \lim_m \pi_n^{\beta_c(q)+\frac{1}{m}}(\omega_q) + 2\beta_c(q) \tag{9.12}$$

$$= -\log \lim_n \pi_n^{\beta_c(q)}(\omega_q) + 2\beta_c(q) \tag{9.13}$$

$$= \hat{I}_\pi^{\beta_c(q)}(\omega_q) + 2\beta_c(q). \tag{9.14}$$

(To prove that $\lim_m \pi_n^{\beta_c(q)+\frac{1}{m}}(\omega_q) = \pi_n^{\beta_c(q)}(\omega_q)$ is straightforward.)

Finally, the algorithmic implications are also a direct application of Theorem 9.1. $\qquad\square$

Remark 5. *The algorithm given in Theorem 9.1 seems to require explicit bounds on the constants C and α, so that given $N \in \mathbb{N}$, we can find an explicit n such that $Ce^{-\alpha(n+1)} < \frac{1}{N}$. Without such bounds, while there exists a poly-time algorithm, we do not always know how to exhibit an explicit algorithm. However, for all three models, for regions sufficiently deep within the supercritical region (i.e. β, λ or γ sufficiently large), one can find crude, but adequate, estimates on C and α and thus can exhibit a poly-time algorithm. This is the case for the hard-core model, where our proof does allow an explicit estimate of the constants for any $\gamma > 468$. On the other hand, in the regions specified in Corollary 2 within the subcritical region, all three models satisfy exponential SSM and then using [31, Corollary 4.7], one can, in principle, exhibit a poly-time algorithm (even without estimates on C and α).*

References

[1] M. Aizenman and D. J. Barsky. "Sharpness of the phase transition in perco-
     lation models". In: *Comm. Math. Phys.* 108.3 (1987), pp. 489–526 (cit. on
     p. 34).

[2] M. Aizenman et al. "Discontinuity of the magnetisation in one-dimensional
     $1/|x-y|^2$ Ising and Potts models". In: *J. Stat. Phys.* 50 (1988), pp. 1–40 (cit.
     on p. 41).

[3] K. Alexander. "Mixing properties and exponential decay for lattice systems in
     finite volumes". In: *Ann. Probab.* 32 (2004), pp. 441–487 (cit. on pp. 4, 24, 29,
     32).

[4] R. Baxter. "Variational approximations for square lattice models in statistical
     mechanics". In: *J. Stat. Phys.* 19.5 (1978), pp. 461–478 (cit. on p. 3).

[5] R. J. Baxter. "Hard hexagons: exact solution". In: *J. Phys. A* 13.3 (1980),
     p. L61 (cit. on p. 5).

[6] V. Beffara and H. Duminil-Copin. "The self-dual point of the two-dimensional
     random-cluster model is critical for $q \geqslant 1$". In: *Probab. Theory Related Fields*
     153.3-4 (2012), pp. 511–542 (cit. on pp. 11, 14, 15, 27).

[7] J. van den Berg and C. Maes. "Disagreement percolation in the study of Markov
     fields". In: *Ann. Probab.* 22.2 (1994), pp. 749–763 (cit. on pp. 5, 26, 35).

[8] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomor-
     phisms*. Ed. by J.-R. Chazottes. 2nd ed. Springer-Verlag, 2008 (cit. on p. 2).

[9] R. Briceño. *The topological strong spatial mixing property and new conditions
     for pressure approximation*. Version 1. Nov. 9, 2014. arXiv: 1411.2289 (cit.
     on pp. 2, 4, 20, 25).

[10] J. T. Chayes, L. Chayes, and R. H. Schonmann. "Exponential decay of connec-
     tivities in the two-dimensional Ising model". In: *J. Stat. Phys.* 49.3-4 (1987),
     pp. 433–445 (cit. on p. 29).

[11] L. Chayes. "Percolation and ferromagnetism on $\mathbb{Z}^2$: the q-state Potts cases".
     In: *Stochastic Process. Appl.* 65.2 (1996), pp. 209–216 (cit. on pp. 4, 28, 30).

[12] J.-D. Deuschel and A. Pisztora. "Surface order large deviations for high-density
     percolation". In: *Probab. Theory Related Fields* 104 (1996), pp. 467–482 (cit.
     on p. 36).

[13] R. L. Dobrushin. "The problem of uniqueness of a Gibssian random field and
     the problem of phase transitions". In: *Funct. Anal. Appl.* 2 (1968), pp. 302–312
     (cit. on pp. 5, 35, 36).

[14] S. Friedland. "On the entropy of $\mathbb{Z}^d$ subshifts of finite type". In: *Linear Algebra
     Appl.* 252 (1997), pp. 199–220 (cit. on pp. 3, 11).

[15] D. Gamarnik and D. Katz. "Sequential cavity method for computing free en-
     ergy and surface pressure". In: *J. Stat. Phys.* 137.2 (2009), pp. 205–232 (cit. on
     pp. 2–4, 19).

[16] H.-O. Georgii. *Gibbs Measures and Phase Transitions*. 2nd ed. Vol. 9. De
     Gruyter Studies in Mathematics. Berlin, 2011 (cit. on pp. 2, 18).

[17] H.-O. Georgii, O. Häggström, and C. Maes. *The random geometry of equilib-
     rium phases*. Ed. by C. Domb and J. Lebowitz. Vol. 18. Phase Transitions and

Critical Phenomena. London: Academic Press, 2001, pp. 1–142 (cit. on pp. 12, 13, 27, 28).

[18] H.-O. Georgii and Y. Higuchi. "Percolation and number of phases in the two-dimensional Ising model". In: *J. Math. Phys.* 41.3 (2000), pp. 1153–1169 (cit. on p. 30).

[19] H.-O. Georgii and V. Zagrebnov. "Entropy-driven phase transitions in multi-type lattice gas models". In: *J. Stat. Phys.* 102 (2001), pp. 35–67 (cit. on pp. 12, 35).

[20] G. R. Grimmett. *The Random-Cluster Model*. Springer, 2006 (cit. on pp. 14, 15, 27).

[21] Y. Higuchi and M. Takei. "Some results on the phase structure of the two-dimensional Widom-Rowlinson model". In: *Osaka J. Math.* 41 (2004), pp. 237–255 (cit. on pp. 16, 27, 28).

[22] M. Hochman and T. Meyerovitch. "A characterization of the entropies of multidimensional shifts of finite type". In: *Ann. of Math. (2)* 171.3 (2010), pp. 2011–2038 (cit. on p. 3).

[23] G. Keller. *Equilibrium States in Ergodic Theory*. Cambridge: Cambridge University Press, 1998 (cit. on pp. 2, 3, 18).

[24] H. Kesten. "Aspects of first passage percolation". In: *Éc. Été Probab. St.-Flour XIV*. Ed. by P. Hennequin. Vol. 1180. Lecture Notes in Mathematics. Springer-Verlag, 1986, pp. 125–264 (cit. on p. 16).

[25] D. A. Klarner and R. L. Rivest. "A procedure for improving the upper bound for the number of *n*-ominoes". In: *Canad. J. Math.* 25.3 (1973), pp. 585–602 (cit. on p. 38).

[26] K.-I. Ko. *Complexity Theory of Real Functions*. Boston: Birkhäuser, 1991 (cit. on p. 3).

[27] U. Krengel and A. Brunel. *Ergodic Theorems*. De Gruyter Studies in Mathematics. W. de Gruyter, 1985 (cit. on p. 18).

[28] W. Krieger. "On the uniqueness of the equilibrium state". In: *Math. Syst. Theory* 8.2 (1974), pp. 97–104 (cit. on p. 3).

[29] B. Marcus and R. Pavlov. "An integral representation for topological pressure in terms of conditional probabilities". In: *Israel J. Math.* 207.1 (2015), pp. 395–433 (cit. on pp. 2–4, 20).

[30] B. Marcus and R. Pavlov. "Approximating entropy for a class of $\mathbb{Z}^2$ Markov random fields and pressure for a class of functions on $\mathbb{Z}^2$ shifts of finite type". In: *Ergodic Theory Dynam. Systems* 33 (2013), pp. 186–220 (cit. on p. 26).

[31] B. Marcus and R. Pavlov. "Computing bounds for entropy of stationary $\mathbb{Z}^d$ Markov random fields". In: *SIAM J. Discrete Math.* 27.3 (2013), pp. 1544–1558 (cit. on pp. 39, 41).

[32] M. Menshikov. "Coincidence of critical points in percolation problems". In: *Soviet Math. Dokl.* 33 (1986), pp. 856–859 (cit. on p. 34).

[33] M. Misiurewicz. "A short proof of the variational principle for a $\mathbb{Z}_+^N$ action on a compact space". In: *International Conference on Dynamical Systems in Mathematical Physics*. Vol. 40. Astérisque. Paris: Soc. Math. France, 1976, pp. 147–157 (cit. on p. 18).

[34] C.-E. Pfister and Y. Velenik. "Mathematical theory of the wetting phenomenon in the 2D Ising model". In: *Helv. Phys. Acta* 69.5-6 (1996), pp. 949–973 (cit. on p. 5).

[35] D. Ruelle. *Thermodynamic Formalism. The Mathematical Structure of Equilibrium Statistical Mechanics*. 2nd ed. Cambridge Math. Lib. Cambridge: Cambridge Univ. Press, 2004 (cit. on pp. 2, 10, 11, 18, 19, 26).

[36] L. K. Runnels and J. L. Lebowitz. "Phase transitions of a multicomponent Widom-Rowlinson model". In: *J. Math. Phys.* 15.10 (1974), pp. 1712–1717 (cit. on p. 12).

[37] L. Russo. "On the critical percolation probabilities". In: *Z. Wahrscheinlichkeit.* 56.2 (1981), pp. 229–237 (cit. on p. 32).

[38] L. Russo. "The infinite cluster method in the two-dimensional Ising model". In: *Comm. Math. Phys.* 67.3 (1979), pp. 251–266 (cit. on p. 5).

[39] J. C. Vera, E. Vigoda, and L. Yang. "Improved bounds on the phase transition for the hard-core model in 2-dimensions". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Vol. 8096. Lecture Notes in Computer Science. Springer-Verlag, 2013, pp. 699–713 (cit. on p. 39).

[40] P. Walters. *An Introduction to Ergodic Theory*. Graduate Texts in Mathematics 79. Springer-Verlag, 1982 (cit. on p. 2).