

# Potts model with invisible colours

## Random-cluster representation and Pirogov-Sinai analysis

Aernout C. D. van Enter<sup>1,\*</sup>      Giulio Iacobelli<sup>1,†</sup>  
Siamak Taati<sup>1,2,‡</sup>

<sup>1</sup>Johann Bernoulli Institute, University of Groningen, The Netherlands

<sup>2</sup>Mathematical Institute, Utrecht University, The Netherlands

### Abstract

We study a variant of the ferromagnetic Potts model, recently introduced by Tamura, Tanaka and Kawashima, consisting of a ferromagnetic interaction among  $q$  “visible” colours along with the presence of  $r$  non-interacting “invisible” colours. We introduce a random-cluster representation for the model, for which we prove the existence of a first-order transition for any  $q > 0$ , as long as  $r$  is large enough. When  $q > 1$ , the low-temperature regime displays a  $q$ -fold symmetry breaking. The proof involves a Pirogov-Sinai analysis applied to this random-cluster representation of the model.

**Keywords:** Potts model with invisible colours, biased random-cluster model, phase transition, symmetry breaking

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Potts Model with Invisible Colours</b>	<b>3</b>
2.1	The Model . . . . .	3
2.2	First-Order Phase Transition . . . . .	5
2.3	The Main Result . . . . .	5
<b>3</b>	<b>Biased Random-Cluster Representation</b>	<b>6</b>
<b>4</b>	<b>Reduction to Contour Model</b>	<b>10</b>
4.1	Contour Representation . . . . .	10
4.2	Contour Models . . . . .	14

---

Last update: January 20, 2012  
\*Email: [A.C.D.van.Enter@rug.nl](mailto:A.C.D.van.Enter@rug.nl)  
†Email: [dajegiu@gmail.com](mailto:dajegiu@gmail.com)  
‡Email: [siamak.taati@gmail.com](mailto:siamak.taati@gmail.com)

<b>5</b>	<b>Damping of Contour Weights</b>	<b>20</b>
<b>6</b>	<b>Proof of the Main Result</b>	<b>28</b>
<b>7</b>	<b>Conclusion</b>	<b>31</b>
<b>A</b>	<b>Appendix</b>	<b>32</b>
A.1	Derivation of the Biased Random-Cluster Representation . . . . .	32
A.2	Derivation of the Contour Representation . . . . .	33
A.3	A Few Properties of the Biased Random-Cluster Model . . . . .	35

# 1 Introduction

Recently, in a series of papers [21, 22, 23], Tamura, Tanaka and Kawashima introduced a variant of the ferromagnetic Potts model to study the relation between symmetry breaking and the order of the phase transition. The model consists of a ferromagnetic Potts interaction taking place between  $q$  “visible” colours along with the presence of  $r$  “invisible” colours without any interaction. They observed, through numerical simulations, that in two dimensions with  $q = 2, 3, 4$  and  $r$  large, the model undergoes a first-order phase transition with  $q$ -fold symmetry breaking. This is in contrast with the ordinary two-dimensional  $q$ -colour Potts model with  $q = 2, 3, 4$ , in which the transition accompanied by the  $q$ -fold symmetry breaking is known to be of second order [1]. This provides a simple example that a  $q$ -fold symmetry breaking in two dimensions does not universally identify the order of the transition.

The transition in this model (as well as in the standard  $q$ -colour Potts model) occurs from an *ordered* state (which has a favoured direction among  $q$  possibilities) to a *disordered* state (which has no favoured direction) when the temperature is increased. In the standard  $q$ -colour Potts model with  $q$  small, this transition is of *second* order (no latent heat at the transition point) whereas in the Tamura-Tanaka-Kawashima version of the Potts model, for the same values of  $q$  but  $r$  chosen sufficiently large, the transition is of *first* order (the system absorbs heat during the transition, without changing its temperature).

For the standard  $q$ -colour Potts model, when  $q$  is large enough, there is a variety of different rigorous proofs that the transition is of first order [15, 4, 18, 17] (see also Section 6.4 and Chapter 7 of [11], and [16]). As announced in an earlier communication [6], in the present paper, we prove, by minor adaptations of the proofs in [17, 16], that when  $q + r$  is large enough, the Potts model with  $q$  visible colours and  $r$  invisible colours undergoes a first-order phase transition. The proof is based on an application of the Pirogov-Sinai method to a random-cluster representation of the model.

The phase transition could be better understood if one thinks of a state of the system as a possible resolution of the conflict between order and disorder. The conflict should be resolved locally and in every region. In an ordered region, the neighbouring sites tend to take the same colour so as to minimize the energy, while in a disordered region, the neighbouring sites take their colours independently to maximize the entropy. To establish the resolution of the order-disorder conflict, one needs to take into account the disturbance present at the interface between ordered and disordered regions (the *contours*).

For the standard  $q$ -colour Potts model, the order-disorder conflict is niftily depicted in the Fortuin-Kasteleyn (a.k.a. random-cluster) representation of the model [7, 11, 9, 13]. In this repre-

sentation, order is associated with the presence of bonds between neighbouring sites and disorder with the absence of bonds. In the same spirit, we introduce a variant of the Fortuin-Kasteleyn representation for the Potts model with invisible colours. The advantage of this new formulation is that it admits a neat definition of the interface between ordered and disordered regions. Now, having two reference configurations describing complete order and complete disorder — the one with every bond present, and the one with every bond absent — as in [17], we can apply the Pirogov-Sinai method [16, 20, 24, 3].

In Section 2, we describe the model and recall the formulation of first-order phase transition in the Gibbsian setup. Section 3 is dedicated to the introduction of a variant of the random-cluster model and its connection with the Potts model with invisible colours. In Section 4.1, formal definitions for contours are provided, and it is shown how to rewrite the partition functions of the model in terms of contours. These contour representations are then reduced, in Section 4.2, to two abstract contour models, on which standard techniques can be applied. In Section 5, starting from the two contour models, we obtain two approximations for the free energy of the Potts model with invisible colours. If  $q+r$  is large, each of these approximations turns out to be accurate in an interval of temperatures, one whenever order prevails and the other when disorder is dominant. The two intervals exhaust all the temperatures and have a unique common point, which is the transition point of the system. Finally, the above two approximations are used in Section 6 to prove a first-order transition at the transition point. The occurrence of the symmetry breaking at the same transition point then follows, using standard properties of the random-cluster representation, which are reviewed in Appendix A.3.

**Acknowledgments.** G.I. and S.T. thank NWO for support. We thank Roberto Fernández for helpful conversations. We thank A. N. Berker for bringing reference [2] to our attention.

## 2 Potts Model with Invisible Colours

### 2.1 The Model

Let  $\mathbb{L}$  denote the two-dimensional square *lattice*, which we think of as a graph  $(\mathbb{S}, \mathbb{B})$ , where  $\mathbb{S}$  denotes the set of sites (identified by  $\mathbb{Z}^2$ ) and  $\mathbb{B}$  the set of nearest neighbour bonds. In the  $(q, r)$ -Potts model, each site  $i \in \mathbb{S}$  is in one of  $(q+r)$  colours  $1, 2, \dots, q, q+1, \dots, q+r$ . Therefore, a *configuration*  $\sigma$  of the model is an assignment of values from the set  $\{1, 2, \dots, q, q+1, \dots, q+r\}$  to the sites in  $\mathbb{S}$ . The  $(q, r)$ -Potts model [21, 22, 23] is expressed by the formal Hamiltonian

$$H(\sigma) = - \sum_{\{i,j\} \in \mathbb{B}} \delta(\sigma_i = \sigma_j \leq q), \quad (1)$$

where  $\delta(\sigma_i = \sigma_j \leq q)$  is 1 if  $\sigma_i = \sigma_j \leq q$  and 0 otherwise. Each pair of neighbouring sites that have the same colour  $\alpha \leq q$  contributes with energy  $-1$ , while sites with colours  $\alpha > q$  do not contribute to the energy. The first  $q$  colours are hence called the *visible* colours, and the rest the *invisible* colours. If there are no invisible colours (i.e. if  $r = 0$ ), the model reduces to the ordinary Potts model with  $q$  colours. As in the ordinary  $q$ -colour Potts model, the  $(q, r)$ -Potts model has precisely  $q$  periodic ground-state configurations, in which every site has the same visible colour.

Following the usual approach, we describe the system in thermal equilibrium via probability distributions on the space of all possible configurations of the model. The *Boltzmann distribution*

on a finite volume  $\Lambda \subseteq \mathbb{L}$  with boundary condition  $\omega$  at inverse temperature  $\beta$  is defined by

$$\mu_{\beta, \Lambda}^{\omega}(\sigma_{\Lambda}) = \frac{1}{Z_{\beta}^{\omega}(\Lambda)} e^{-\beta H_{\Lambda}(\sigma_{\Lambda} \omega_{\Lambda^c})}, \quad (2)$$

where  $H_{\Lambda}(\sigma_{\Lambda} \omega_{\Lambda^c})$  consists of a finite number of terms in the formal Hamiltonian (1) corresponding to the energy of  $\sigma_{\Lambda}$  and its interaction with the boundary condition  $\omega$ . Namely,

$$H_{\Lambda}(\sigma) = - \sum_{\substack{\{i, j\} \in \mathbb{B} \\ i \in S(\Lambda) \text{ or } j \in S(\Lambda)}} \delta(\sigma_i = \sigma_j \leq q), \quad (3)$$

which can be decomposed as a sum

$$H_{\Lambda}(\sigma_{\Lambda} \omega_{\Lambda^c}) = H_{\Lambda}^{\text{int}}(\sigma_{\Lambda}) + H_{\Lambda}^{\text{bound}}(\sigma_{\Lambda} \omega_{\Lambda^c}), \quad (4)$$

where  $H_{\Lambda}^{\text{int}}(\sigma_{\Lambda})$  involves the interaction terms within  $\Lambda$ , and  $H_{\Lambda}^{\text{bound}}(\sigma_{\Lambda} \omega_{\Lambda^c})$  represents the terms corresponding to the interaction of  $\Lambda$  with its boundary. The factor  $Z_{\beta}^{\omega}(\Lambda)$  is a normalizing constant — the *partition function* — making  $\mu_{\beta, \Lambda}^{\omega}$  a probability distribution. More specifically, the partition function of volume  $\Lambda$  with boundary condition  $\omega$  is given by

$$Z_{\beta}^{\omega}(\Lambda) = \sum_{\sigma_{\Lambda}} e^{-\beta H_{\Lambda}^{\text{int}}(\sigma_{\Lambda}) - \beta H_{\Lambda}^{\text{bound}}(\sigma_{\Lambda} \omega_{\Lambda^c})}. \quad (5)$$

If we ignore the boundary term, then we obtain the *free-boundary* partition function of volume  $\Lambda$ :

$$Z_{\beta}^{\text{free}}(\Lambda) = \sum_{\sigma_{\Lambda}} e^{-\beta H_{\Lambda}^{\text{int}}(\sigma_{\Lambda})}. \quad (6)$$

This is the normalizing factor for the free-boundary Boltzmann distribution on  $\Lambda$ . Note that if  $\omega$  is a configuration in which every site has an invisible colour, the two partition functions  $Z_{\beta}^{\omega}(\Lambda)$  and  $Z_{\beta}^{\text{free}}(\Lambda)$  coincide. A *Gibbs measure* on the space of all configurations of the infinite lattice system, at inverse temperature  $\beta$ , is a probability measure  $\mu$  whose conditional probabilities for every finite volume  $\Lambda$ , given the configuration  $\omega$  outside  $\Lambda$ , are given by the Boltzmann distribution  $\mu_{\beta, \Lambda}^{\omega}$ . More specifically,

$$\mu(A \text{ and } B) = \int_B \mu_{\beta, \Lambda}^{\omega}(A) \mu(d\omega), \quad (7)$$

for every event  $A$  not depending on the colours of the sites outside  $\Lambda$  and every event  $B$  not depending on the colours of the sites in  $\Lambda$ . It follows from a compactness argument that such measures exist at every temperature. However, when the temperature is sufficiently low, it is possible to have several distinct Gibbs measures. The multiplicity of Gibbs measures is then interpreted as the possibility of co-existence of distinguishable phases of the physical system (in this case, the possibility of spontaneous magnetization in  $q$  different directions). We refer to [8] for an exhaustive treatment.

One way to obtain Gibbs measures consists of taking the thermodynamic limit of the Boltzmann distribution with or without a fixed boundary condition. For a visible colour  $k$ , let  $\omega^k$  denote the configuration of the lattice in which every site has colour  $k$ . Let  $\mu_{\beta}^k$  denote a Gibbs measure obtained

by taking a weak limit of finite-volume Boltzmann distributions with boundary condition  $\omega^k$  at inverse temperature  $\beta$ , when the finite-volume grows to the whole lattice. Similarly, we obtain a Gibbs measure  $\mu_\beta^{\text{free}}$  by taking a weak limit of free-boundary Boltzmann distributions.

For every  $n > 0$ , let  $\Lambda_n$  denote the  $(2n+1) \times (2n+1)$  central square in the lattice, which we see as the subgraph of the lattice induced by the sites in  $[-n, n]^2$ . The *pressure* of the model is defined by

$$f(\beta) = \lim_{n \rightarrow \infty} \frac{1}{|S(\Lambda_n)|} \log Z_\beta^\omega(\Lambda_n), \quad (8)$$

in which  $S(\Lambda_n)$  denotes the set of sites in  $\Lambda_n$ . The function  $-\frac{1}{\beta}f(\beta)$  is the *free energy* per site. The limit exists and is independent of the boundary condition  $\omega$  (see e.g. Theorem I.2.3 in [14] or Theorem 3.4 in [19]). We would also get the same limit as in (8) if we used the free boundary partition function. The particular choice of volumes used above is not crucial, and can be replaced by any sequence satisfying the van Hove property (see [14], Theorem I.2.4). The function  $f(\beta)$  is convex and Lipschitz continuous (Theorem I.2.3 in [14] or Theorem 3.4 in [19]). In particular, its left and right derivatives exist and are finite at any point  $\beta$ .

## 2.2 First-Order Phase Transition

A *first-order phase transition* in temperature is characterized by the presence of *latent heat* at the transition point [5]. This means that at the transition point, the system absorbs or gives out heat without a change in temperature. The presence of latent heat, therefore, corresponds to a jump in the internal energy.

In the Gibbsian setup (see [8]), the state of a system in thermal equilibrium is represented by a Gibbs measure. If the Gibbs measure is translation-invariant, the *internal energy density* of the system is described by the expected value of energy per site. The presence of latent heat at a temperature means that the limits of the internal energy density from above and below the transition temperature are different. This implies, by continuity, the existence of two translation-invariant Gibbs measures at that temperature having different internal energy density.

If the pressure function  $f(\beta)$  is differentiable at a point  $\beta$ , its derivative at  $\beta$  coincides with the internal energy per site with respect to every translation-invariant Gibbs measure at  $\beta$  (see [14], Chapters II and III, or [19], Chapters 3 and 4). (This, however, does not rule out the possibility of the existence of several translation-invariant Gibbs measures at  $\beta$ .) If, on the other hand, the pressure function  $f(\beta)$  is non-differentiable at a point  $\beta$ , its left and right derivatives at  $\beta$  (which exist due to convexity) are different and coincide with the internal energy per site with respect to two different translation-invariant Gibbs measures at  $\beta$ . The difference between these two derivatives corresponds to a latent heat at  $\beta$ , implying that the system undergoes a first-order phase transition at  $\beta$ .

## 2.3 The Main Result

In this paper, we show that for  $r$  large, the  $(q, r)$ -Potts model undergoes a first-order phase transition in temperature accompanied by a  $q$ -fold symmetry breaking. This is obtained by proving that the pressure function  $f(\beta)$  has a unique non-differentiable point  $\beta_c$  at which the permutation symmetry between the  $q$  visible colours is broken. Below  $\beta_c$ , the system has a “disordered” translation-invariant Gibbs measure, which can be seen as a perturbation of the uniform Bernoulli measure:

there is a unique infinite sea of independent colours with finite islands of disturbance. Above  $\beta_c$ , the system admits  $q$  “ordered” translation-invariant Gibbs measures. Each “ordered” measure can be thought of as a perturbation of one of the  $q$  ground-state configurations, in the sense that with probability 1, the configuration of the model consists of a unique infinite sea of one of the visible colours with finite islands of disturbance. At  $\beta_c$ , the  $q$  “ordered” measures co-exist with the “disordered” one.

**Theorem.** *For  $\varepsilon > 0$ , there exists  $Q_\varepsilon > 0$  such that for every  $q > 1$  and  $r \geq 0$  satisfying  $q + r \geq Q_\varepsilon$ , the two-dimensional  $(q, r)$ -Potts model undergoes a first-order transition in temperature with breaking of permutation symmetry. Namely, there exists a critical temperature at which the pressure function is not differentiable, and for which the following statements hold:*

- i) *Above the transition temperature, the model has a unique Gibbs state  $\mu^{\text{free}}$ , which is “disordered”.*
- ii) *Below the transition temperature, there exist at least  $q$  different “ordered” Gibbs states  $\mu^1, \mu^2, \dots, \mu^q$ .*
- iii) *At the transition temperature,  $q$  “ordered” Gibbs states  $\mu^1, \mu^2, \dots, \mu^q$  coexist with a “disordered” Gibbs state  $\mu^{\text{free}}$ .*

The “ordered” and “disordered” states can be distinguished by

$$\mu^k(\{\sigma : \sigma_i = k\}) > 1 - \varepsilon, \quad \text{for every visible } k, \quad (9)$$

$$\mu^{\text{free}}(\{\sigma : \sigma_i = k\}) < \varepsilon, \quad \text{for every } k, \quad (10)$$

for every site  $i$  in the lattice.

Let us remark that the above theorem remains valid even if  $q = 1$ , although in that case there is no breaking of permutation symmetry.

### 3 Biased Random-Cluster Representation

In analogy with the standard Potts model, it is possible to rewrite the partition function for the  $(q, r)$ -Potts model in terms of the partition function for a variant of the random-cluster model (see [7, 11] or Section 6 of [9]). While the former is a model defined on sites, the latter will be a model defined on bonds. The random-cluster representation of the Potts model allows for an elegant formulation of the intuitive concepts of “order” and “disorder”: the presence of a bond in the random-cluster representation is interpreted as “order”, while the absence of a bond as “disorder” [17].

Although for the purpose of our problem, it suffices to present the connection for squares  $\Lambda_n$  in the lattice, we elucidate the connection for an arbitrary finite graph, where there is no boundary condition. Later, we explain how the boundary conditions affect this connection.

Let  $\mathbb{G} = (S, B)$  be a finite graph. The  $r$ -biased random-cluster model on  $\mathbb{G}$  is given by a probability distribution on the sets  $X \subseteq B$ . The distribution has three parameters  $0 \leq p \leq 1$ ,  $q > 0$  and  $r \geq 0$ , and is defined by

$$\phi_{p,q,r}(X) = \frac{1}{Z_{p,q,r}^{\text{RC}}(\mathbb{G})} \left[ \prod_{b \in B} p^{\delta(b \in X)} (1-p)^{\delta(b \notin X)} \right] (q+r)^{\kappa_0(S,X)} q^{\kappa_1(S,X)}, \quad (11)$$

in which  $\kappa_0(S, X)$  denotes the number of isolated sites of the graph  $(S, X)$  and  $\kappa_1(S, X)$  the number of non-singleton connected components of  $(S, X)$  and  $Z_{p,q,r}^{\text{RC}}(\mathbb{G})$  the partition function. Notice that for  $r = 0$ , the model reduces to the standard random-cluster model, in which both singleton and non-singleton connected components have weight  $q$ . For  $r > 0$ , the above model induces a bias towards singleton connected components. Namely, the singleton connected components have weight  $(q + r)$  whereas the non-singleton connected components have weight  $q$ .

Let us now see how the  $(q, r)$ -Potts model is related to the  $r$ -biased random-cluster model. This is a mere generalization of the standard relation between the Potts and random-cluster models (see e.g., Section 1.4 of [11]). Let  $\Omega$  be the set of  $(q, r)$ -Potts configurations on  $\mathbb{G}$ . The partition function of this model can be rewritten as

$$\begin{aligned}
Z_\beta(\mathbb{G}) &= \sum_{\sigma \in \Omega} e^{\beta \sum_{\{i,j\} \in B} \delta(\sigma_i = \sigma_j \leq q)} \\
&= \sum_{\sigma \in \Omega} \prod_{\{i,j\} \in B} e^{\beta \delta(\sigma_i = \sigma_j \leq q)} \\
&= \sum_{\sigma \in \Omega} \prod_{\{i,j\} \in B} [1 + \delta(\sigma_i = \sigma_j \leq q)(e^\beta - 1)] \\
&= \sum_{\sigma \in \Omega} \sum_{X \subseteq B} (e^\beta - 1)^{|X|} \prod_{\{i,j\} \in X} \delta(\sigma_i = \sigma_j \leq q) \\
&= \sum_{\sigma \in \Omega} \sum_{X \subseteq B} \pi(\sigma, X),
\end{aligned} \tag{12}$$

where

$$\pi(\sigma, X) = e^{\beta|B|} \prod_{\{i,j\} \in B} [\delta(\{i,j\} \in X) \delta(\sigma_i = \sigma_j \leq q)(1 - e^{-\beta}) + \delta(\{i,j\} \notin X) e^{-\beta}]. \tag{13}$$

The latter expression can be seen as a coupling of the  $(q, r)$ -Potts distribution on  $\Omega$  and a probability distribution on the space  $\{0, 1\}^B$ . The marginal of this coupling on  $\{0, 1\}^B$  is simply the  $r$ -biased random-cluster distribution  $\phi_{p_\beta, q, r}$  with  $p_\beta = 1 - e^{-\beta}$ . In particular, the weight  $\pi(\sigma, X)$  can also be expressed as

$$\pi(\sigma, X) = e^{\beta|B|} \cdot 1_{F_r}(\sigma, X) \cdot \prod_{\{i,j\} \in B} [p_\beta \delta(\{i,j\} \in X) + (1 - p_\beta) \delta(\{i,j\} \notin X)], \tag{14}$$

where

$$F_r \triangleq \{(\sigma, X) : \sigma_i = \sigma_j \leq q \text{ for all } \{i, j\} \in X\}. \tag{15}$$

The effect of the bias in the  $r$ -biased random-cluster model reduces to an increase in the number of compatible configurations with a given  $X$ , which is driven by a larger number of choices for the colour of those sites constituting the singleton connected components. In short, for each  $X \subseteq B$ , we have

$$\sum_{\sigma \in \Omega} 1_{F_r}(\sigma, X) = q^{\kappa_1(S, X)} (q + r)^{\kappa_0(S, X)}. \tag{16}$$

The above coupling could be interpreted in either of the following ways (Section 1.4 of [11] or Section 6 of [9]):

- I. We first sample  $\sigma$  according to the  $(q, r)$ -Potts distribution. Then, we choose the elements of  $X$  from  $B$ , randomly and independently, as follows: for each bond  $\{i, j\} \in B$  with  $\sigma_i = \sigma_j$ , we put  $\{i, j\}$  in  $X$  with probability  $p_\beta$ ; for each bond  $\{i, j\} \in B$  with  $\sigma_i \neq \sigma_j$ , we do not put  $\{i, j\}$  in  $X$ .
- II. We first sample  $X$  according to the  $r$ -biased random-cluster distribution  $\phi_{p_\beta, q, r}$ . Then, for each non-singleton connected component of  $(S, X)$ , we pick a random colour uniformly among the visible colours, and colour every site in the component with that colour. Last, for every isolated site in  $(S, X)$ , we choose a random colour uniformly among all the possible colours. (The choices of colours ought to be independent of each other.)

We can now use (13) to obtain

$$Z_\beta(\mathbb{G}) = e^{\beta|B|} Z_{p_\beta, q, r}^{\text{RC}}(\mathbb{G}) \quad (17)$$

with  $p_\beta = 1 - e^{-\beta}$ .

For finite subgraphs of the infinite lattice, we will be using only two types of partition functions for the  $(q, r)$ -Potts model, namely the one with free boundary and the ones with homogenous boundary conditions. In the following, we see how the above two types of boundary conditions translate into the so-called *disordered* and *ordered* boundary conditions for the  $r$ -biased random-cluster model. Although, setting  $\mathbb{G} = \Lambda_n$ , equation (17) already provides a relation between the free-boundary partition functions of the two models, we will work with a slightly different relation, connecting the free-boundary partition function of the  $(q, r)$ -Potts model to a partition function for the  $r$ -biased random-cluster model that involves a boundary condition. This new relation will turn out to be more convenient in the sequel.

The free-boundary partition function for the  $(q, r)$ -Potts model can be written as

$$Z_\beta(\Lambda_n) = (q + r)^{-|S(\Lambda_{n+1}) \setminus S(\Lambda_n)|} \cdot e^{\beta|B(\Lambda_{n+1})|} \cdot Z_{p_\beta, q, r}^{\text{RC, disord}}(\Lambda_{n+1}), \quad (18)$$

where  $Z_{p_\beta, q, r}^{\text{RC, disord}}(\Lambda_{n+1})$  is the partition function with *disordered boundary condition* for the  $r$ -biased random-cluster model. The latter is defined by

$$Z_{p_\beta, q, r}^{\text{RC, disord}}(\Lambda_{n+1}) = \sum_{X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}} p_\beta^{|X|} (1 - p_\beta)^{|B(\Lambda_{n+1}) \setminus X|} (q + r)^{\kappa_0(S(\Lambda_{n+1}), X)} q^{\kappa_1(S(\Lambda_{n+1}), X)}, \quad (19)$$

where

$$\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}} = \{X \subseteq B(\Lambda_{n+1}) : X \cap (B(\Lambda_{n+1}) \setminus B(\Lambda_n)) = \emptyset\}. \quad (20)$$

Similarly, for the boundary condition  $\omega^k$  we get

$$Z_\beta^{\omega^k}(\Lambda_n) = q^{-1} \cdot (e^\beta - 1)^{-|B(\Lambda_{n+1}) \setminus \Lambda_n|} \cdot e^{\beta|B(\Lambda_{n+1})|} \cdot Z_{p_\beta, q, r}^{\text{RC, ord}}(\Lambda_{n+1}), \quad (21)$$

where  $Z_{p_\beta, q, r}^{\text{RC, ord}}(\Lambda_{n+1})$  is the partition function with *ordered boundary condition* for the  $r$ -biased random-cluster model, which is defined by

$$Z_{p_\beta, q, r}^{\text{RC, ord}}(\Lambda_{n+1}) = \sum_{X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}} p_\beta^{|X|} (1 - p_\beta)^{|B(\Lambda_{n+1}) \setminus X|} (q + r)^{\kappa_0(S(\Lambda_{n+1}), X)} q^{\kappa_1(S(\Lambda_{n+1}), X)}, \quad (22)$$



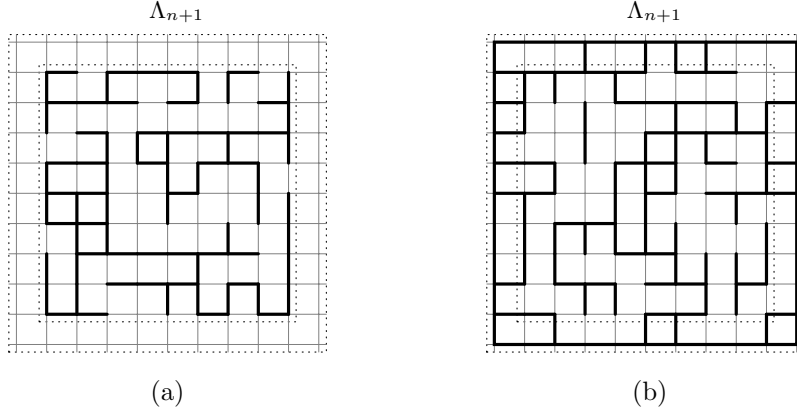


Figure 1: (a) A configuration in  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$ . (b) A configuration in  $\mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$ .

where

$$\mathcal{X}_{\Lambda_{n+1}}^{\text{ord}} = \{X \subseteq B(\Lambda_{n+1}) : X \supseteq B(\Lambda_{n+1} \setminus \Lambda_n)\} . \quad (23)$$

By  $\Lambda_{n+1} \setminus \Lambda_n$  we mean the graph obtained from  $\Lambda_{n+1}$  by removing all the sites in  $\Lambda_n$  and the bonds attached to them. Let us remark that although mathematically  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  is simply the collection of all subsets  $X \subseteq B(\Lambda_n)$ , we wrote it as above to emphasize that the elements of  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  are configurations of  $B(\Lambda_{n+1})$ . See Figure 1 for typical examples of elements in  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  and  $\mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$ .

In the following section, we will extend the definition of  $Z_{p\beta,q,r}^{\text{RC,disord}}$  and  $Z_{p\beta,q,r}^{\text{RC,ord}}$  to arbitrary subgraphs of the lattice.

Using the above relationships, we obtain that the pressure of the  $(q,r)$ -Potts model can be written as

$$f(\beta) = 2 \left( \beta + \lim_{n \rightarrow \infty} \frac{\log Z_{p\beta,q,r}^{\text{RC,disord}}(\Lambda_n)}{|B(\Lambda_n)|} \right) = 2 \left( \beta + \lim_{n \rightarrow \infty} \frac{\log Z_{p\beta,q,r}^{\text{RC,ord}}(\Lambda_n)}{|B(\Lambda_n)|} \right) . \quad (24)$$

We call the limit

$$f^{\text{RC}}(\beta) = \lim_{n \rightarrow \infty} \frac{\log Z_{p\beta,q,r}^{\text{RC,disord}}(\Lambda_n)}{|B(\Lambda_n)|} = \lim_{n \rightarrow \infty} \frac{\log Z_{p\beta,q,r}^{\text{RC,ord}}(\Lambda_n)}{|B(\Lambda_n)|} \quad (25)$$

the *pressure* (per bond) of the  $r$ -biased random-cluster representation. Note that the singularities of the  $(q,r)$ -Potts pressure function  $f(\beta)$  can be detected by studying the pressure function  $f^{\text{RC}}(\beta)$ . One advantage of this random-cluster representation is that it has a more transparent expression in terms of “contours”, which helps us study the function  $f^{\text{RC}}(\beta)$ .

## 4 Reduction to Contour Model

### 4.1 Contour Representation

Any configuration of the  $r$ -biased random-cluster model in a volume is a subset  $X$  of bonds in the volume. We interpret each bond in  $X$  as “ordered” and each bond outside  $X$  as “disordered”. Any configuration  $X$  can then be seen as clusters of ordered and disordered bonds. Whether an equilibrium state is ordered or disordered can be seen as the result of a competition between ordered and disordered regions. The selection criterion for this competition is “energy”. The term “energy” refers to an abstract notion of energy for the  $r$ -biased random-cluster model, which in analogy with the Boltzmann distribution, corresponds to minus logarithm of probability. We remark that this abstract notion of “energy” should be interpreted as free energy in the original spin model.

Let us define the “energy” of an ordered bond as the “energy” per bond of the fully ordered configuration; that is,

$$e(\mathbb{B}) \triangleq \lim_{n \rightarrow \infty} \frac{-\log [q(1 - e^{-\beta})^{|B(\Lambda_n)|}]}{|B(\Lambda_n)|} = -\log(1 - e^{-\beta}) . \quad (26)$$

Similarly, define the “energy” of a disordered bond as the “energy” per bond of the fully disordered configuration:

$$e(\emptyset) \triangleq \lim_{n \rightarrow \infty} \frac{-\log [e^{-\beta|B(\Lambda_n)|}(q+r)^{|S(\Lambda_n)|}]}{|B(\Lambda_n)|} = -\log [e^{-\beta}\sqrt{q+r}] . \quad (27)$$

The “energy” of the ordered and the disordered regions can now be expressed as  $|R^o| \cdot e(\mathbb{B})$  and  $|R^d| \cdot e(\emptyset)$ , respectively, where  $|R^o|$  and  $|R^d|$  denote the size of the ordered and disordered regions. The “energy” of  $X$ , in turn, can be written in terms of the “energy” of ordered and disordered regions plus a correction term due to the effect of the boundaries separating them. If the effect of these boundaries is negligible (which will turn out to be the case whenever  $q+r$  is large), the selection criterion for the competition between order and disorder boils down to determining which of  $e(\mathbb{B})$  and  $e(\emptyset)$  is minimal. This is the starting point of the Pirogov-Sinai approach to study phase transitions (see e.g. [16]).

The presence of the correction term at the boundaries can be explained as follows: In the probability weight of a configuration  $X$ , each isolated site contributes with a factor  $(q+r)$ . To express the “energy” of the disordered regions purely in terms of bonds, we evenly distribute the contribution of the isolated sites among the 4 incident bonds. Doing so, every disorder bond acquires zero, one or two “energy”-shares, depending on the number of isolated sites it is incident to. Since in the fully disordered configuration  $\emptyset$  there is no ordered region, every bond is incident to precisely two isolated sites and receives two “energy”-shares, leading to the factor  $(q+r)^{\frac{2}{4}}$  in the expression of  $e(\emptyset)$ . In an arbitrary configuration, however, the disordered bonds on the borderline between the ordered and disordered regions, receive one or no “energy”-share, hence the need for a correction term.

It is possible to define a suitable notion of boundary between ordered and disordered regions, so that each configuration  $X$  is uniquely identified by its boundary (see below). We could then rewrite the partition functions as sums running over “admissible” boundaries, that is, those corresponding to configurations of bonds. Each admissible boundary is split into “primary” objects termed contours whose “energy” add up to the corresponding boundary effect.

In the following, we specify rigorously the above heuristic notions of “boundary” and “contours”. We define the *boundary* of a configuration  $X \subseteq \mathbb{B}$  as the set

$$\bar{\partial}X \triangleq \{(i, b) \in \mathbb{S} \times \mathbb{B} : i \sim b \text{ and } i \in S(X) \text{ and } b \notin X\}, \quad (28)$$

where  $i \sim b$  means site  $i$  and bond  $b$  are incident, and  $S(X)$  is the set of sites incident to bonds in  $X$ . The set  $\bar{\partial}X$  uniquely determines  $X$ . We say that two bonds  $b$  and  $b'$  in the lattice are *co-adjacent* if they belong to the same unit square. More intuitively, co-adjacency is equivalent to adjacency in the dual lattice. A set of bonds  $X$  is *co-connected* if for every two bonds  $b, b' \in X$ , there is a sequence  $b = b_1, b_2, \dots, b_n = b'$  of bonds in  $X$  such that  $b_i$  and  $b_{i+1}$  are co-adjacent. A *contour* is a set  $\gamma \subseteq \mathbb{S} \times \mathbb{B}$  such that

- i) the set of bonds appearing in  $\gamma$ , denoted by  $B(\gamma)$ , is co-connected, and
- ii) there exists a configuration  $X$  such that  $(S(X), X)$  is connected and  $\gamma = \bar{\partial}X$ .

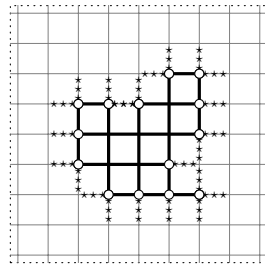
We shall denote by  $\Gamma$  the set of all finite contours in  $\mathbb{L}$ . If  $\gamma$  is a contour, then removing the bonds  $B(\gamma)$  breaks the lattice  $\mathbb{L}$  into connected subgraphs. If  $\gamma$  is finite, the graph  $\mathbb{L} \setminus B(\gamma)$  has a unique infinite connected component, which we call the *exterior* of  $\gamma$  and denote by  $\text{ext } \gamma$ . The subgraph  $\mathbb{L} \setminus B(\gamma) \setminus \text{ext } \gamma$  (which could be empty or disconnected) is called the *interior* of  $\gamma$  and is denoted by  $\text{int } \gamma$ . By  $V(\gamma)$  we will mean the union of  $\text{int } \gamma$  and the subgraph induced by  $B(\gamma)$ .<sup>1</sup> Let  $\gamma$  be a finite contour. The configuration  $X$  such that  $(S(X), X)$  is connected and  $\gamma = \bar{\partial}X$  (which exists by definition) is either finite or co-finite. If  $X$  is finite, we call  $\gamma$  a *disorder* contour, and if  $X$  is co-finite, we call  $\gamma$  an *order* contour. Note that, if  $\gamma$  is a disorder contour, all the sites appearing in  $\gamma$  are in the interior of  $\gamma$ , whereas if  $\gamma$  is an order contour, all the sites appearing in  $\gamma$  are in the exterior of  $\gamma$ . As a result, we can safely represent a finite contour  $\gamma$  by the pair  $(B(\gamma), x)$  where  $x$  is a label specifying the type of the contour (disorder or order). This also means that the set of all finite contours  $\Gamma$  can be partitioned into two subsets: the set of disorder contours, which we denote by  $\Gamma^d$ , and the set of order contours, which we denote by  $\Gamma^o$ .

Two contours are said to be *mutually compatible* if they are disjoint (as subsets of  $\mathbb{S} \times \mathbb{B}$ ). Let us emphasize that two mutually compatible contours are allowed to share either sites or bonds, but not pairs.

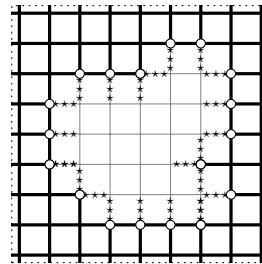
If  $X \subseteq \mathbb{B}$  is an arbitrary configuration, there could be several ways to partition its boundary  $\bar{\partial}X$  into mutually compatible contours. One way to construct such decomposition in an unambiguous way is as follows: first, we partition  $(S(X), X)$  into its maximal connected components  $(S(C_i), C_i)$ . Then  $\bar{\partial}C_i$  form a partitioning of  $\bar{\partial}X$ . Now, the maximal co-connected components of every  $C_i$  are contours that we identify as the *contours of  $X$* .

The above decomposition allows us to think of  $\bar{\partial}X$  as a family of mutually compatible contours, which we call the *contour family* of  $X$ . Let us recall that the contour family of a configuration  $X$  uniquely determines  $X$ . However, note that not every family of mutually compatible contours corresponds to a configuration. In particular, in a contour family of a configuration  $X$ , between every two nested finite contours of the same type, there necessarily lies a contour of the other type. This requirement induces a long-range constraint among contours, which raises some difficulties in dealing with the contours. We will see later how to get rid of such a constraint. Let us call a family  $\partial$  of contours *admissible* if it is the contour family of a configuration  $X \subseteq \mathbb{B}$ . We shall denote by  $\Delta$  the set of all admissible contour families. A contour  $\gamma$  in a mutually compatible family  $\partial$  of

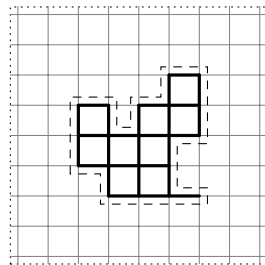
<sup>1</sup>By the subgraph induced by a set of bonds we mean the graph obtained by those bonds and their endpoints.



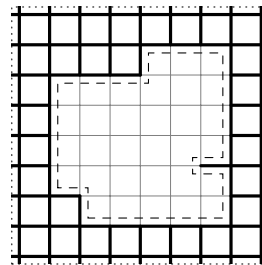
(a)



(b)



(a')



(b')

Figure 2: (a) A disorder contour and its defining configuration. (b) An order contour and its defining configuration. (a') and (b') Geometric illustrations of (a) and (b).

contours is said to be *external* if it is not in the interior of any other contour in  $\partial$ . Note that if  $\partial$  is an admissible contour family with no infinite contours, all the external contours in  $\partial$  are necessarily of the same type.

Having formalized the notions of boundary and contours, we can now express the weight of a configuration of the  $r$ -biased random-cluster model in terms of the “energy” of its ordered and disordered regions and the correction term due to the contours separating them. The one-to-one correspondence between the configurations and the admissible families of contours allows us to write the partition functions as a sum over contour families. The ordered/disordered boundary conditions on the configurations translate into the constraints for the corresponding contour family that the outermost contours in the volume be of the order/disorder type.

Let  $\Lambda$  be a *volume* in the lattice, by which, from now on, we shall mean a finite subgraph of  $\mathbb{L}$  without “holes”. More precisely, we assume that if we remove the subgraph  $\Lambda$  from  $\mathbb{L}$ , the remaining subgraph is connected. Let us denote by  $\Delta_\Lambda^{\text{disord}}$  the set of all admissible contour families whose contours are in  $\Lambda$  (i.e., their bonds are chosen from the bonds of  $\Lambda$ ) and whose external contours are all of the disorder type. Similarly, let  $\Delta_\Lambda^{\text{ord}}$  denote the set of admissible contour families in  $\Lambda$  whose external contours are all of the order type. The partition function for the  $r$ -biased random-cluster model in a volume  $\Lambda$  with *disordered* (resp., *ordered*) *boundary conditions* can be defined as

$$Z_{p_\beta, q, r}^{\text{RC, disord}}(\Lambda) = (q+r)^{\frac{|\partial B(\Lambda)|}{4}} \sum_{\partial \in \Delta_\Lambda^{\text{disord}}} e^{-|R_\Lambda^{\text{o}}(\partial) \cdot e(\mathbb{B}) - |R_\Lambda^{\text{d}}(\partial) \cdot e(\emptyset)} \prod_{\gamma \in \partial} \rho(\gamma), \quad (29)$$

$$Z_{p_\beta, q, r}^{\text{RC, ord}}(\Lambda) = q \sum_{\partial \in \Delta_\Lambda^{\text{ord}}} e^{-|R_\Lambda^{\text{o}}(\partial) \cdot e(\mathbb{B}) - |R_\Lambda^{\text{d}}(\partial) \cdot e(\emptyset)} \prod_{\gamma \in \partial} \rho(\gamma), \quad (30)$$

where  $\rho(\gamma)$  is the *weight* of a contour  $\gamma$  and is given by

$$\rho(\gamma) \triangleq \begin{cases} (q+r)^{-\frac{1}{4}|\gamma|}, & \text{if } \gamma \text{ order,} \\ q(q+r)^{-\frac{1}{4}|\gamma|}, & \text{if } \gamma \text{ disorder,} \end{cases} \quad (31)$$

and  $R_\Lambda^{\text{o}}(\partial)$  and  $R_\Lambda^{\text{d}}(\partial)$  denote, respectively, the sets of ordered and disordered bonds in  $\Lambda$  of the configuration corresponding to  $\partial$ .

The above definitions are consistent with the definitions given in (19) and (22) when  $\Lambda = \Lambda_{n+1}$  is a square. Namely, for  $\Lambda = \Lambda_{n+1}$ , if  $X$  is the corresponding configuration of a family  $\partial \in \Delta_{\Lambda_{n+1}}^{\text{disord}}$ , the restriction of  $X$  to  $B(\Lambda_{n+1})$  is an element of  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$ . Conversely, every element of  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  has a unique infinite-volume extension whose corresponding contour family is in  $\Delta_{\Lambda_{n+1}}^{\text{disord}}$ . A similar correspondence holds between  $\mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$  and  $\Delta_{\Lambda_{n+1}}^{\text{ord}}$ . For the proof of the equivalence of the two definitions see Appendix A.2.

We emphasize that the factors  $(q+r)^{\frac{|\partial B(\Lambda)|}{4}}$  and  $q$  in front of the partition functions (29) and (30) do not contribute to the pressure function  $f^{\text{RC}}(\beta)$ : in the thermodynamic limit, they are swallowed by the size of the volume. Hence, to avoid heavy notation — with all due apologies to the reader — we *re-define* the partition functions of the  $r$ -biased random-cluster model with disordered/ordered boundary conditions as

$$Z^{\text{RC, disord}}(\Lambda) = \sum_{\partial \in \Delta_\Lambda^{\text{disord}}} e^{-|R_\Lambda^{\text{o}}(\partial) \cdot e(\mathbb{B}) - |R_\Lambda^{\text{d}}(\partial) \cdot e(\emptyset)} \prod_{\gamma \in \partial} \rho(\gamma), \quad (32)$$

$$Z^{\text{RC,ord}}(\Lambda) = \sum_{\partial \in \Delta_{\Lambda}^{\text{ord}}} e^{-|R_{\Lambda}^{\circ}(\partial)| \cdot e(\mathbb{B}) - |R_{\Lambda}^{\square}(\partial)| \cdot e(\emptyset)} \prod_{\gamma \in \partial} \rho(\gamma). \quad (33)$$

From now on, every time we talk about the partition function of the  $r$ -biased random-cluster model, we will be referring to the latter definitions.

As was mentioned in the introduction, we would like to express the two partition functions in terms of two (standard) contour models. The purpose of this is to make use of the machinery available for contour models; namely, a result providing an estimate on the convergence of the corresponding free energy functions (Proposition 5), and the Peierls estimate for the probability of the appearance of a contour. The main features of the contour models that are required in the above tools are (see [16])

- i) independence, and
- ii) damping.

Unfortunately, the contours of the contour representation of the  $r$ -biased random-cluster partition functions are not independent (due to the long-range constraint). In the following section, we will see how to achieve the independence among contours, by rewriting the partition functions in terms of abstract contour models. As in the standard random-cluster model (see [17]), we need two different such contour models, one for each of the two boundary conditions.

## 4.2 Contour Models

In this section, we want to resolve the issue of long-range constraints between contours. Recall that the admissibility condition requires the contours of a family to be alternating between disorder and order contours, and this imposes a long-range constraint between contours. For example, two nested contours of the disorder type (no matter how far from each other) are “aware” of the presence of an order contour separating them. As a result, if we remove a contour from an admissible family, the admissibility could be lost.

In order to get rid of this constraint, we use two abstract contour models in which the contours are all of the same type and the admissibility condition is replaced by mere mutual compatibility. The weights of the contours in each of the abstract models will be chosen in such a way to guarantee that the ensuing partition functions are equal (up to a factor) to each of the partition functions for the  $r$ -biased random-cluster model.

A *contour model* is specified by a function  $\chi : \Gamma \rightarrow \mathbb{R}$ , assigning a weight  $\chi(\gamma)$  to each contour  $\gamma \in \Gamma$ . The configurations of the model are families of mutually compatible (i.e., disjoint) contours in  $\mathbb{L}$ . Let us denote the set of all such families by  $\mathcal{M}$ , and the set of all elements of  $\mathcal{M}$  whose contours are in a volume  $\Lambda$  by  $\mathcal{M}_{\Lambda}$ . The partition function of the model in  $\Lambda$  is given by

$$\mathcal{Z}(\Lambda | \chi) = \sum_{\partial \in \mathcal{M}_{\Lambda}} \prod_{\gamma \in \partial} \chi(\gamma). \quad (34)$$

In the following lemma, we will see how to represent the partition functions of the  $r$ -biased random-cluster model with disordered and ordered boundary conditions, each in terms of the partition function of a contour model, with a particular choice of the weight function. In fact, the contour model associated to the disordered boundary condition will not involve order contours. This is

reflected by the fact that in this model each order contour has weight zero. Similarly, the contour model for the ordered boundary condition involves only order contours.

To set the stage for the lemma, we rewrite the partition functions  $Z^{\text{RC.disord}}(\Lambda)$  and  $Z^{\text{RC.ord}}(\Lambda)$  in a form resembling more the contour model partition function  $\mathcal{Z}(\Lambda | \chi)$ . That is,

$$Z^{\text{RC.disord}}(\Lambda) = e^{-|B(\Lambda)| \cdot e(\emptyset)} \sum_{\partial \in \Delta_{\Lambda}^{\text{disord}}} \prod_{\gamma \in \partial} \tilde{\rho}(\gamma), \quad (35)$$

$$Z^{\text{RC.ord}}(\Lambda) = e^{-|B(\Lambda)| \cdot e(\mathbb{B})} \sum_{\partial \in \Delta_{\Lambda}^{\text{ord}}} \prod_{\gamma \in \partial} \tilde{\rho}(\gamma), \quad (36)$$

where

$$\tilde{\rho}(\gamma) = \begin{cases} \rho(\gamma) \cdot e^{-|B(\text{int } \gamma)| \cdot (e(\mathbb{B}) - e(\emptyset))}, & \text{if } \gamma \text{ is disorder,} \\ \rho(\gamma) \cdot e^{-|B(V(\gamma))| \cdot (e(\emptyset) - e(\mathbb{B}))}, & \text{if } \gamma \text{ is order.} \end{cases} \quad (37)$$

Let us recall that the set  $\Delta_{\Lambda}^{\text{disord}}$  (resp.,  $\Delta_{\Lambda}^{\text{ord}}$ ) does not contain only families of disorder (resp., order) contours, but all families compatible with the disordered (resp., ordered) boundary condition. To make the proof more transparent, let us define

$$Y^{\text{d}}(\Lambda) = \sum_{\partial \in \Delta_{\Lambda}^{\text{disord}}} \prod_{\gamma \in \partial} \tilde{\rho}(\gamma), \quad (38)$$

$$Y^{\text{o}}(\Lambda) = \sum_{\partial \in \Delta_{\Lambda}^{\text{ord}}} \prod_{\gamma \in \partial} \tilde{\rho}(\gamma), \quad (39)$$

so that

$$Z^{\text{RC.disord}}(\Lambda) = e^{-|B(\Lambda)| e(\emptyset)} \cdot Y^{\text{d}}(\Lambda), \quad (40)$$

$$Z^{\text{RC.ord}}(\Lambda) = e^{-|B(\Lambda)| e(\mathbb{B})} \cdot Y^{\text{o}}(\Lambda). \quad (41)$$

Notice that the above contour representation for the partition functions  $Z^{\text{RC.disord}}(\Lambda)$  and  $Z^{\text{RC.ord}}(\Lambda)$  lacks the condition of independence between compatible contours.

The following lemma is similar to Lemma 1 in [16].

**Lemma 1.** *The partition functions for the  $r$ -biased random-cluster model on volume  $\Lambda$  with the disordered and ordered boundary conditions can be written as*

$$Z^{\text{RC.disord}}(\Lambda) = e^{-|B(\Lambda)| e(\emptyset)} \mathcal{Z}(\Lambda | \xi^{\text{d}}), \quad (42)$$

$$Z^{\text{RC.ord}}(\Lambda) = e^{-|B(\Lambda)| e(\mathbb{B})} \mathcal{Z}(\Lambda | \xi^{\text{o}}), \quad (43)$$

where the weights  $\xi^{\text{d}}$  and  $\xi^{\text{o}}$  are defined by

$$\xi^{\text{d}}(\gamma) = \begin{cases} \rho(\gamma) \frac{Z^{\text{RC.ord}}(\text{int } \gamma)}{Z^{\text{RC.disord}}(\text{int } \gamma)}, & \text{if } \gamma \text{ disorder,} \\ 0, & \text{otherwise,} \end{cases} \quad (44)$$

and

$$\xi^{\text{o}}(\gamma) = \begin{cases} \rho(\gamma) e^{|B(\gamma)| e(\mathbb{B})} \frac{Z^{\text{RC.disord}}(V(\gamma))}{Z^{\text{RC.ord}}(\text{int } \gamma)}, & \text{if } \gamma \text{ is order,} \\ 0, & \text{otherwise.} \end{cases} \quad (45)$$

*Proof.* The key step to prove the lemma is to write a recursion for the above partition functions by factoring the contribution of the interior of each external contour. Let us denote by  $\mathcal{E}_\Lambda^{\text{disord}}$  the set of mutually compatible families of disorder contours whose elements are all external. (We include the empty family in  $\mathcal{E}_\Lambda^{\text{disord}}$ .) Note that the elements of  $\mathcal{E}_\Lambda^{\text{disord}}$  are all admissible and in  $\Delta_\Lambda^{\text{disord}}$ . Moreover, for each admissible family  $\partial \in \Delta_\Lambda^{\text{disord}}$ , the sub-family of  $\partial$  consisting of its external contours is in  $\mathcal{E}_\Lambda^{\text{disord}}$ . Similarly, we denote by  $\mathcal{E}_\Lambda^{\text{ord}}$  the set of mutually compatible families of order contours whose elements are all external. The partition functions  $Y^{\text{d}}$  and  $Y^{\text{o}}$  satisfy the following recursions:

$$Y^{\text{d}}(\Lambda) = \sum_{\theta \in \mathcal{E}_\Lambda^{\text{disord}}} \prod_{\gamma \in \theta} \tilde{\rho}(\gamma) \cdot Y^{\text{o}}(\text{int } \gamma), \quad (46)$$

$$Y^{\text{o}}(\Lambda) = \sum_{\theta \in \mathcal{E}_\Lambda^{\text{ord}}} \prod_{\gamma \in \theta} \tilde{\rho}(\gamma) \cdot Y^{\text{d}}(V(\gamma)). \quad (47)$$

Similar recursions hold for the contour model partition functions  $\mathcal{Z}(\cdot | \xi^{\text{d}})$  and  $\mathcal{Z}(\cdot | \xi^{\text{o}})$ :

$$\mathcal{Z}(\Lambda | \xi^{\text{d}}) = \sum_{\theta \in \mathcal{E}_\Lambda^{\text{disord}}} \prod_{\gamma \in \theta} \xi^{\text{d}}(\gamma) \cdot \mathcal{Z}(\text{int } \gamma | \xi^{\text{d}}), \quad (48)$$

$$\mathcal{Z}(\Lambda | \xi^{\text{o}}) = \sum_{\theta \in \mathcal{E}_\Lambda^{\text{ord}}} \prod_{\gamma \in \theta} \xi^{\text{o}}(\gamma) \cdot \mathcal{Z}(\text{int } \gamma | \xi^{\text{o}}). \quad (49)$$

Note that, since every order contour is weighted 0 by  $\xi^{\text{d}}$ , we can ignore in  $\mathcal{Z}(\cdot | \xi^{\text{d}})$  the families containing order contours, and similarly the disorder contours can be ignored in  $\mathcal{Z}(\cdot | \xi^{\text{o}})$ .

We use induction on volume  $\Lambda$  to prove that  $Y^{\text{d}}(\Lambda) = \mathcal{Z}(\Lambda | \xi^{\text{d}})$ . Suppose that for every sub-volume  $\Lambda' \subsetneq \Lambda$  we have  $Y^{\text{d}}(\Lambda') = \mathcal{Z}(\Lambda' | \xi^{\text{d}})$ . Let  $\theta \in \mathcal{E}_\Lambda^{\text{disord}}$ . We want to show that the terms corresponding to  $\theta$  in the recursion formulas (46) and (48) for  $Y^{\text{d}}(\Lambda)$  and  $\mathcal{Z}(\Lambda | \xi^{\text{d}})$  are equal. If  $\theta$  is empty, the equality is trivial (we consider the product over an empty set to be 1). Otherwise, for every  $\gamma \in \theta$ , we have  $\text{int } \gamma \subsetneq \Lambda$ , which implies  $\mathcal{Z}(\text{int } \gamma | \xi^{\text{d}}) = Y^{\text{d}}(\text{int } \gamma)$ . Using the definitions of  $\tilde{\rho}$  and  $\xi^{\text{d}}$  we obtain that

$$\prod_{\gamma \in \theta} \tilde{\rho}(\gamma) \cdot Y^{\text{o}}(\text{int } \gamma) = \prod_{\gamma \in \theta} \xi^{\text{d}}(\gamma) \cdot Y^{\text{d}}(\text{int } \gamma). \quad (50)$$

Therefore,  $Y^{\text{d}}(\Lambda) = \mathcal{Z}(\Lambda | \xi^{\text{d}})$ . The starting point of the induction is when the only element of  $\mathcal{E}_\Lambda^{\text{disord}}$  is the empty family.

The argument for  $Y^{\text{o}}(\Lambda) = \mathcal{Z}(\Lambda | \xi^{\text{o}})$  is similar.  $\square$

Note that there is no complete correspondence between the configurations of the  $r$ -biased random-cluster model and the contour families of the corresponding abstract contour model. Nevertheless, the probability of appearance of a contour as an external contour is the same in both models. Let  $\phi_\Lambda^{\text{ord}}$  denote the probability distribution associated to  $Z^{\text{RC,ord}}(\Lambda)$ . We consider  $\phi_\Lambda^{\text{ord}}$  as a measure on the infinite-volume bond configurations  $X \subseteq \mathbb{B}$ , which is concentrated on the set  $\{X : \bar{\partial}X \in \Delta_\Lambda^{\text{ord}}\}$ , and is defined by

$$\phi_\Lambda^{\text{ord}}(X) \triangleq \frac{\prod_{\gamma \in \bar{\partial}X} \tilde{\rho}(\gamma)}{\sum_{\partial \in \Delta_\Lambda^{\text{ord}}} \prod_{\gamma \in \bar{\partial}X} \tilde{\rho}(\gamma)} = \frac{\prod_{\gamma \in \bar{\partial}X} \tilde{\rho}(\gamma)}{Y^{\text{o}}(\Lambda)} \quad (51)$$



for every  $X \subseteq \mathbb{B}$  such that  $\bar{\partial}X \in \Delta_\Lambda^{\text{ord}}$ . Likewise,  $\phi_\Lambda^{\text{disord}}$  will denote the measure corresponding to  $Z^{\text{RC,disord}}(\Lambda)$ , which is concentrated on the set  $\{X : \bar{\partial}X \in \Delta_\Lambda^{\text{disord}}\}$ .

**Corollary 2.** *Let  $\Lambda$  be a finite volume and  $\theta \in \mathcal{E}_\Lambda^{\text{ord}}$  a family of external mutually compatible order contours. Then,*

$$\phi_\Lambda^{\text{ord}}\{X : \bar{\partial}_{\text{ext}}X = \theta\} = \frac{\prod_{\gamma \in \theta} \xi^\circ(\gamma) \mathcal{Z}(\text{int } \gamma | \xi^\circ)}{\mathcal{Z}(\Lambda | \xi^\circ)}, \quad (52)$$

where  $\bar{\partial}_{\text{ext}}X$  is the family of external contours of  $X$ . A similar statement holds for the probability of families of external mutually compatible disorder contours under  $\phi_\Lambda^{\text{disord}}$ .

*Proof.*

$$\begin{aligned} \phi_\Lambda^{\text{ord}}\{X : \bar{\partial}_{\text{ext}}X = \theta\} &= \frac{\prod_{\gamma \in \theta} \tilde{\rho}(\gamma) Y^{\text{d}}(V(\gamma))}{Y^\circ(\Lambda)} \\ &= \frac{\prod_{\gamma \in \theta} \left[ \rho(\gamma) \cdot e^{-|B(V(\gamma))| \cdot (e(\emptyset - e(\mathbb{B})))} \frac{Y^{\text{d}}(V(\gamma))}{Y^\circ(\text{int } \gamma)} \right] Y^\circ(\text{int } \gamma)}{Y^\circ(\Lambda)} \\ &= \frac{\prod_{\gamma \in \theta} \xi^\circ(\gamma) \mathcal{Z}(\text{int } \gamma | \xi^\circ)}{\mathcal{Z}(\Lambda | \xi^\circ)} \end{aligned} \quad (53)$$

□

The next corollary provides an estimate for the probability that a finite region of the lattice is surrounded by a contour (see e.g. Theorem 7.32 in [11]). For a finite set of sites  $A$  in the lattice, let  $\Gamma_A$  denote the set of all finite contours that have  $A$  in their interiors.

**Corollary 3.** *For every finite volume  $\Lambda$  and every finite set  $A \subseteq S(\Lambda)$*

$$\phi_\Lambda^{\text{ord}}\{X : \bar{\partial}X \cap \Gamma_A \neq \emptyset\} \leq \sum_{\gamma \in \Gamma_A} \xi^\circ(\gamma) = \sum_{\substack{\gamma \in \Gamma_A \\ \gamma \text{ order}}} \xi^\circ(\gamma). \quad (54)$$

*A similar bound holds in the disordered case.*

*Proof.* Taking into account the ordered boundary condition, we have that if  $A$  is surrounded by a contour in  $\Lambda$ , it is also surrounded by an external order contour in  $\Lambda$ , that is,

$$\{X : \bar{\partial}X \cap \Gamma_A \neq \emptyset\} = \{X : \bar{\partial}_{\text{ext}}X \cap \Gamma_A \neq \emptyset\}. \quad (55)$$

By the previous corollary, we can bound the probability of a contour  $\gamma$  appearing as an external

contour by

$$\begin{aligned}
\phi_\Lambda^{\text{ord}}\{X : \bar{\partial}_{\text{ext}}X \ni \gamma\} &= \sum_{\substack{\theta \in \xi_\Lambda^{\text{ord}} \\ \theta \ni \gamma}} \frac{\prod_{\hat{\gamma} \in \theta} \xi^\circ(\hat{\gamma}) \mathcal{Z}(\text{int } \hat{\gamma} | \xi^\circ)}{\mathcal{Z}(\Lambda | \xi^\circ)} \\
&= \frac{\xi^\circ(\gamma) \mathcal{Z}(\text{int } \gamma | \xi^\circ) \sum_{\substack{\theta \in \xi_\Lambda^{\text{ord}} \\ \theta \ni \gamma}} \prod_{\substack{\hat{\gamma} \in \theta \\ \hat{\gamma} \neq \gamma}} \xi^\circ(\hat{\gamma}) \mathcal{Z}(\text{int } \hat{\gamma} | \xi^\circ)}{\mathcal{Z}(\Lambda | \xi^\circ)} \\
&\leq \frac{\xi^\circ(\gamma) \mathcal{Z}(\text{int } \gamma | \xi^\circ) \mathcal{Z}(\Lambda \setminus \text{int } \gamma | \xi^\circ)}{\mathcal{Z}(\Lambda | \xi^\circ)} \\
&\leq \xi^\circ(\gamma).
\end{aligned} \tag{56}$$

The last step follows from the fact that all the terms in the partition function  $\mathcal{Z}(\Lambda | \xi^\circ)$  are non-negative, hence

$$\mathcal{Z}(\Lambda | \xi^\circ) \geq \mathcal{Z}(\text{int } \gamma | \xi^\circ) \mathcal{Z}(\Lambda \setminus \text{int } \gamma | \xi^\circ). \tag{57}$$

We obtain that

$$\begin{aligned}
\phi_\Lambda^{\text{ord}}\{X : \bar{\partial}X \cap \Gamma_A \neq \emptyset\} &= \phi_\Lambda^{\text{ord}}\{X : \bar{\partial}_{\text{ext}}X \cap \Gamma_A \neq \emptyset\} \\
&= \sum_{\gamma \in \Gamma_A} \phi_\Lambda^{\text{ord}}\{X : \bar{\partial}_{\text{ext}}X \ni \gamma\} \\
&\leq \sum_{\gamma \in \Gamma_A} \xi^\circ(\gamma).
\end{aligned} \tag{58}$$

□

A standard argument using the positive correlation property of  $\phi_\Lambda^{\text{ord}}$  (resp.,  $\phi_\Lambda^{\text{disord}}$ ) can be used to show that the thermodynamic limit of  $\phi_\Lambda^{\text{ord}}$  (resp.,  $\phi_\Lambda^{\text{disord}}$ ) exists and is unique (see Appendix A.3). The limit measure  $\phi^{\text{ord}}$  (resp.,  $\phi^{\text{disord}}$ ) satisfies the same bound as in the above corollary. If the weights  $\xi^\circ$  (resp.,  $\xi^{\text{d}}$ ) decay sufficiently fast, the latter bound implies that under  $\phi^{\text{ord}}$  (resp.,  $\phi^{\text{disord}}$ ), the configuration of the model almost surely consists of a unique infinite sea of order (resp., disorder) with finite islands of disorder (resp., order). By a “sea” of order (resp., disorder) in a random-cluster configuration we mean a connected component of present (resp., absent) bonds.

**Corollary 4.** *For every finite set  $A \subseteq \mathbb{S}$*

$$\phi^{\text{ord}}\{X : \bar{\partial}X \cap \Gamma_A \neq \emptyset\} \leq \sum_{\gamma \in \Gamma_A} \xi^\circ(\gamma) = \sum_{\substack{\gamma \in \Gamma_A \\ \gamma \text{ order}}} \xi^\circ(\gamma). \tag{59}$$

Furthermore, if the sum  $\sum_{\gamma \in \Gamma_A} \xi^\circ(\gamma)$  converges,

$$\phi^{\text{ord}} \left( \begin{array}{l} \exists \text{ unique infinite sea of order} \\ \text{with finite islands of disorder} \end{array} \right) = 1, \tag{60}$$

A similar statement holds in the disordered case.

*Proof.* As before, let  $\Lambda_n$  denote the  $(2n+1) \times (2n+1)$  central square in the lattice. For every  $n$  let us define  $\Gamma_{A,\Lambda_n}$  as the set of all contours in  $\Lambda_n$  having  $A$  in their interiors. From the previous corollary, we know that for every  $m > n$ , the following bound holds:

$$\phi_{\Lambda_m}^{\text{ord}} \{X : \bar{\partial}X \cap \Gamma_{A,\Lambda_n} \neq \emptyset\} \leq \sum_{\gamma \in \Gamma_A} \xi^o(\gamma). \quad (61)$$

Since the event  $\{X : \bar{\partial}X \cap \Gamma_{A,\Lambda_n} \neq \emptyset\}$  is local, we obtain

$$\phi^{\text{ord}} \{X : \bar{\partial}X \cap \Gamma_{A,\Lambda_n} \neq \emptyset\} \leq \sum_{\gamma \in \Gamma_A} \xi^o(\gamma). \quad (62)$$

due to weak convergence of  $\phi_{\Lambda_m}^{\text{ord}}$  to  $\phi^{\text{ord}}$ . Letting  $n \rightarrow \infty$  the first claim follows.

If  $\sum_{\gamma \in \Gamma_A} \xi^o(\gamma)$  converges, using a Borel-Cantelli argument, with probability 1, no infinite cascade of contours appears on the lattice. In particular, if we define

$$\mathcal{S}^o \triangleq \left\{ X : \begin{array}{l} X \text{ has a unique infinite sea of order} \\ \text{with finite islands of disorder} \end{array} \right\}, \quad (63)$$

$$\mathcal{S}^d \triangleq \left\{ X : \begin{array}{l} X \text{ has a unique infinite sea of disorder} \\ \text{with finite islands of order} \end{array} \right\}, \quad (64)$$

the latter implies that  $\phi^{\text{ord}}(\mathcal{S}^o \cup \mathcal{S}^d) = 1$ . We show that, in fact,  $\phi^{\text{ord}}(\mathcal{S}^d) = 0$ .

Let  $A$  be a finite set of sites in the lattice. For every  $X \in \mathcal{S}^d$  one can find a volume  $\Lambda$  containing  $A$  such that  $\bar{\partial}(X \cap B(\Lambda)) \in \Delta_{\Lambda}^{\text{disord}}$  (i.e., the restriction of  $X$  to  $\Lambda$  is compatible with the disordered boundary condition). In particular, if we define

$$\mathcal{C}_{A,\Lambda_n} \triangleq \{X : \exists \text{ a finite volume } \Lambda \subseteq \Lambda_n \text{ with } S(\Lambda) \supseteq A \text{ and } \bar{\partial}(X \cap B(\Lambda)) \in \Delta_{\Lambda}^{\text{disord}}\}, \quad (65)$$

we have  $\mathcal{C}_{A,\Lambda_1} \subseteq \mathcal{C}_{A,\Lambda_2} \subseteq \dots$  and  $\mathcal{S}^d \subseteq \bigcup_n \mathcal{C}_{A,\Lambda_n}$ .

If  $m, n$  are integers with  $m > n$ , every configuration  $X$  that is compatible with the ordered boundary condition on  $\Lambda_m$  (i.e.,  $\bar{\partial}X \in \Delta_{\Lambda_m}^{\text{ord}}$ ) and is in  $\mathcal{C}_{A,\Lambda_n}$  necessarily has an order contour surrounding  $A$ . Therefore, by the previous corollary, we have

$$\phi_{\Lambda_m}^{\text{ord}}(\mathcal{C}_{A,\Lambda_n}) \leq \phi_{\Lambda_m}^{\text{ord}} \{X : \bar{\partial}X \cap \Gamma_A \neq \emptyset\} \leq \sum_{\gamma \in \Gamma_A} \xi^o(\gamma). \quad (66)$$

Since  $\mathcal{C}_{A,\Lambda_n}$  is a local event, by weak convergence of  $\phi_{\Lambda_m}^{\text{ord}}$  to  $\phi^{\text{ord}}$  we have

$$\phi^{\text{ord}}(\mathcal{C}_{A,\Lambda_n}) \leq \sum_{\gamma \in \Gamma_A} \xi^o(\gamma). \quad (67)$$

Letting  $n \rightarrow \infty$ , we obtain

$$\phi^{\text{ord}}(\mathcal{S}^d) \leq \lim_{n \rightarrow \infty} \phi^{\text{ord}}(\mathcal{C}_{A,\Lambda_n}) \leq \sum_{\gamma \in \Gamma_A} \xi^o(\gamma). \quad (68)$$

The latter holds for every finite set  $A \subseteq \mathbb{S}$ , which by the convergence of the series, implies that  $\phi^{\text{ord}}(\mathcal{S}^d) = 0$ .  $\square$

## 5 Damping of Contour Weights

One advantage of working with contour models is that when the contour weights are sufficiently “damped” (i.e., decay exponentially in the length with a sufficiently fast rate) the free energy exists and is bounded, and moreover, the error in the finite-volume approximations of the free energy is of the order of the size of the boundary of the finite volume. This is the message of the following well-known proposition (see e.g. Section 2 of [24], or Section 9, Proposition 2.3 in [20]).

**Proposition 5.** *Let  $\tau > 0$  be sufficiently large, and suppose that the weight function  $\chi : \Gamma \rightarrow \mathbb{R}$  of a contour model satisfies  $0 \leq \chi(\gamma) \leq e^{-\tau|\gamma|}$  for every contour  $\gamma$ . Then, the limit*

$$g(\chi) = \lim_{n \rightarrow \infty} \frac{1}{|B(\Lambda_n)|} \log \mathcal{Z}(\Lambda_n | \chi) \quad (69)$$

*exists and satisfies  $g(\chi) \leq \sum_{\gamma: S(\gamma) \ni 0} \chi(\gamma) \leq e^{-\tau/2}$ . In particular  $g(\chi) \rightarrow 0$  as  $\tau$  tends to infinity.*

*Furthermore, there is a constant  $C = C(\tau)$ , such that  $C \rightarrow 0$  as  $\tau$  goes to infinity, and for each finite volume  $\Lambda \in \mathbb{Z}^d$*

$$e^{g(\chi)|B(\Lambda)| - C(\tau)|\partial\Lambda|} \leq \mathcal{Z}(\Lambda | \chi) \leq e^{g(\chi)|B(\Lambda)| + C(\tau)|\partial\Lambda|}, \quad (70)$$

*where  $\partial\Lambda$  denotes the boundary of the volume  $\Lambda$  and can be defined as the set of bonds that are not in  $B(\Lambda)$  but are incident to  $\Lambda$ .*

The main purpose of this section is to identify conditions on the parameters  $(q+r)$  and  $\beta$  under which the weights  $\xi^d$  and  $\xi^o$  are damped (i.e., satisfy the condition of the above proposition). We will see that when  $(q+r)$  is large, for any value of  $\beta > 0$  at least one of  $\xi^d$  and  $\xi^o$  is damped. Moreover, we shall show the existence of a unique  $\beta$  below which  $\xi^d$  is damped and above which  $\xi^o$  is damped. Let us remark that for sufficiently damped weights, the sum appearing in Corollary 4 converges, implying that the corresponding phase is stable.

For  $\tau > 0$  large enough, let us introduce the *truncated* (i.e., artificially damped) weights

$$\bar{\xi}^o(\gamma) = \begin{cases} \xi^o(\gamma), & \text{if } \xi^o(\gamma) \leq e^{-\tau|\gamma|}, \\ 0, & \text{otherwise,} \end{cases} \quad (71)$$

and similarly for  $\bar{\xi}^d(\gamma)$  (see e.g. [16]). The term truncated refers to the suppression of all contours whose weights are not damped. If we replace the original weight  $\xi^o$  by the artificially damped one  $\bar{\xi}^o$ , we obtain the following *truncated* partition function, which can be thought of as an approximation of the partition function of the  $r$ -biased random cluster model with ordered boundary condition:

$$\bar{Z}^{\text{RC.ord}}(\Lambda) = e^{-|B(\Lambda)| \cdot e(\mathbb{B})} \mathcal{Z}(\Lambda | \bar{\xi}^o). \quad (72)$$

Similarly, replacing  $\xi^d$  by  $\bar{\xi}^d$  leads to the truncated partition function for the  $r$ -biased random cluster model with disordered boundary condition:

$$\bar{Z}^{\text{RC.disord}}(\Lambda) = e^{-|B(\Lambda)| \cdot e(\emptyset)} \mathcal{Z}(\Lambda | \bar{\xi}^d). \quad (73)$$

The advantage of introducing these truncated partition functions is that we can apply Proposition 5. Note that if the original weights are “damped” (that is,  $\xi^o(\gamma) \leq e^{-\tau|\gamma|}$  or  $\xi^d(\gamma) \leq e^{-\tau|\gamma|}$ ), the corresponding truncated partition functions coincide with the original ones.

From Proposition 5 we have the following bounds for the truncated partition functions:

$$e^{(g(\bar{\xi}^o) - e(\mathbb{B}))|B(\Lambda) - C(\tau)|\partial\Lambda|} \leq \bar{Z}^{\text{RC.ord}}(\Lambda) \leq e^{(g(\bar{\xi}^o) - e(\mathbb{B}))|B(\Lambda) + C(\tau)|\partial\Lambda|}, \quad (74)$$

$$e^{(g(\bar{\xi}^d) - e(\emptyset))|B(\Lambda) - C(\tau)|\partial\Lambda|} \leq \bar{Z}^{\text{RC.disord}}(\Lambda) \leq e^{(g(\bar{\xi}^d) - e(\emptyset))|B(\Lambda) + C(\tau)|\partial\Lambda|}. \quad (75)$$

The pressure functions associated to the truncated partition functions are

$$f^o(\beta) = \lim_{n \rightarrow \infty} \frac{1}{|B(\Lambda_n)|} \log \bar{Z}^{\text{RC.ord}}(\Lambda_n) = -e(\mathbb{B}) + g(\bar{\xi}^o), \quad (76)$$

$$f^d(\beta) = \lim_{n \rightarrow \infty} \frac{1}{|B(\Lambda_n)|} \log \bar{Z}^{\text{RC.disord}}(\Lambda_n) = -e(\emptyset) + g(\bar{\xi}^d). \quad (77)$$

The functions  $f^o(\beta)$  and  $f^d(\beta)$  are lower approximations of the pressure  $f^{\text{RC}}(\beta)$  of the  $r$ -biased random-cluster representation. The next lemma states that in fact when  $(q+r)$  is large enough, the maximum of  $f^o$  and  $f^d$  coincides with  $f^{\text{RC}}$ . As we will see in the next section, for  $(q+r)$  large enough, the functions  $g(\bar{\xi}^o)$  and  $g(\bar{\xi}^d)$  and their  $\beta$ -derivatives are small, and therefore, the dominant terms of  $f^o$  and  $f^d$  are  $-e(\mathbb{B})$  and  $-e(\emptyset)$ . This means that  $f^{\text{RC}}$  is approximated by the maximum of the curves  $-e(\mathbb{B})$  and  $-e(\emptyset)$ , which intersect at a unique value  $\beta$ , with significantly different slopes.

By the diameter of a contour  $\gamma$ , denoted by  $\text{diam } \gamma$ , we shall mean the maximum lattice distance between two bonds in  $B(\gamma)$ . The next lemma is parallel to Lemma 2 in [16] or Theorem 3.1 in [3].

**Lemma 6.** *Let  $(q+r)$  be sufficiently large. If  $f^d \leq f^o$ , then*

*i) for every disorder contour  $\gamma$  with  $\text{diam } \gamma \leq \frac{1}{f^o - f^d}$*

$$\xi^d(\gamma) \leq e^{-\tau|\gamma|}, \quad (78)$$

*ii) for every order contour  $\gamma$*

$$\xi^o(\gamma) \leq e^{-\tau|\gamma|}. \quad (79)$$

*A similar statement holds if  $f^o \leq f^d$ .*

*Proof.* We prove the two claims simultaneously by induction on  $\text{diam } \gamma$ . Let  $K > 0$  and suppose that the claims hold for all (disorder/order) contours with diameter less than  $K$ .

Let  $\gamma$  be a disorder contour with diameter  $K$  that satisfies  $\text{diam } \gamma \leq \frac{1}{f^o - f^d}$ . Then,

$$\begin{aligned} \xi^d(\gamma) &= \rho(\gamma) \frac{Z^{\text{RC.ord}}(\text{int } \gamma)}{Z^{\text{RC.disord}}(\text{int } \gamma)} \\ &= \rho(\gamma) \frac{\bar{Z}^{\text{RC.ord}}(\text{int } \gamma)}{\bar{Z}^{\text{RC.disord}}(\text{int } \gamma)} \\ &\leq \rho(\gamma) \frac{e^{f^o \cdot |B(\text{int } \gamma) + C(\tau)|\partial \text{int } \gamma|}}{e^{f^d \cdot |B(\text{int } \gamma) - C(\tau)|\partial \text{int } \gamma|}} \\ &= \rho(\gamma) e^{(f^o - f^d)|B(\text{int } \gamma) + 2C(\tau)|\partial \text{int } \gamma|}, \end{aligned} \quad (80)$$

where in the second equality we have used the induction hypothesis. Namely, every contour in  $\text{int } \gamma$  has diameter less than  $K$ , allowing us to replace the original partition functions with the truncated ones. Notice that

- $\rho(\gamma) = q(q+r)^{-\frac{1}{4}|\gamma|}$ ,
- $|B(\text{int } \gamma)| \leq \frac{1}{2} |\gamma| \cdot \text{diam } \gamma$ ,
- $(f^\circ - f^d) \cdot \text{diam } \gamma \leq 1$ , and
- $|\partial \text{int } \gamma| \leq |\gamma|$ .

Hence, we obtain that

$$\xi^d(\gamma) \leq q e^{-\left(\frac{1}{4} \log(q+r) - 1 - 2C(\tau)\right) \cdot |\gamma|} . \quad (81)$$

For  $(q+r)$  large enough (uniformly in  $\gamma$ ) the righthand side is bounded by  $e^{-\tau|\gamma|}$ , hence the claim.

Next, suppose that  $\gamma$  is an order contour with diameter  $K$ . We need to show

$$\xi^\circ(\gamma) = \rho(\gamma) e^{|B(\gamma)|e(\mathbb{B})} \frac{Z^{\text{RC.disord}}(V(\gamma))}{Z^{\text{RC.ord}}(\text{int } \gamma)} \leq e^{-\tau|\gamma|} . \quad (82)$$

By the induction hypothesis, the partition function  $Z^{\text{RC.ord}}(\text{int } \gamma)$  is equal to the corresponding truncated partition function, which can be bounded using Proposition 5. As for  $Z^{\text{RC.disord}}(V(\gamma))$ , if we suppress all the contours that are “big”, we can get a similar bound using the induction hypothesis.

To render the argument more transparent, we work with the partition functions  $Y^d(\Lambda)$  and  $Y^\circ(\Lambda)$  (see (38) and (39)) for which we have

$$Z^{\text{RC.disord}}(\Lambda) = e^{-|B(\Lambda)|e(\emptyset)} Y^d(\Lambda) , \quad (83)$$

$$Z^{\text{RC.ord}}(\Lambda) = e^{-|B(\Lambda)|e(\mathbb{B})} Y^\circ(\Lambda) . \quad (84)$$

Let us call a disorder contour *small* if its diameter is less than or equal to  $\frac{1}{f^d - f^\circ}$ . Otherwise, we call the contour *big*.

As before, let us denote by  $\mathcal{E}_\Lambda^{\text{disord}}$  the set of all mutually compatible families of disorder contours in  $\Lambda$  whose elements are external. Factoring the contribution of the interior of big external contours we have the recursion<sup>2</sup>

$$Y^d(\Lambda) = \sum_{\substack{\theta \in \mathcal{E}_\Lambda^{\text{disord}} \\ \theta \text{ big}}} Y_{\text{small}}^d(\Lambda \setminus \text{int } \theta) \prod_{\gamma' \in \theta} \tilde{\rho}(\gamma') Y^\circ(\text{int } \gamma') \quad (85)$$

where  $\text{int } \theta \triangleq \bigcup_{\gamma' \in \theta} \text{int } \gamma'$ , and  $Y_{\text{small}}^d(\Lambda) \triangleq \mathcal{Z}(\Lambda | \xi_{\text{small}}^d)$ , in which the weight function  $\xi_{\text{small}}^d$  is obtained from  $\xi^d$  by replacing the weights of all big contours with 0, that is,

$$\xi_{\text{small}}^d(\gamma') = \begin{cases} \xi^d(\gamma') & \text{if } \gamma' \text{ small,} \\ 0 & \text{if } \gamma' \text{ big.} \end{cases} \quad (86)$$

---

<sup>2</sup>Although  $\Lambda \setminus \text{int } \theta$  does not match our requirement for being a volume (i.e., not having holes), it does not cause any problem. In fact, since the contours in  $\mathcal{E}_{\text{small}}^d$  are small, they cannot surround the holes in  $\Lambda \setminus \text{int } \theta$ , hence they do not distinguish the holes from the outside of  $\Lambda$ .

The expression for the weight  $\xi^\circ(\gamma)$  then reads

$$\xi^\circ(\gamma) = \rho(\gamma) \cdot e^{|B(V(\gamma))| \cdot (e(\mathbb{B}) - e(\emptyset))} \sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} \frac{Y_{\text{small}}^{\text{d}}(V(\gamma) \setminus \text{int } \theta) \cdot Y^\circ(\text{int } \theta)}{Y^\circ(\text{int } \gamma)} \prod_{\gamma' \in \theta} \tilde{\rho}(\gamma'). \quad (87)$$

The induction hypothesis and Proposition 5 tell us:

- $Y_{\text{small}}^{\text{d}}(V(\gamma) \setminus \text{int } \theta) \leq e^{g(\xi_{\text{small}}^{\text{d}}) \cdot |B(V(\gamma) \setminus \text{int } \theta)| + C(\tau) \cdot |\partial(V(\gamma) \setminus \text{int } \theta)|}$ ,
- $Y^\circ(\text{int } \theta) \leq e^{g(\bar{\xi}^\circ) \cdot |B(\text{int } \theta)| + C(\tau) \cdot |\partial \text{int } \theta|}$ ,
- $Y^\circ(\text{int } \gamma) \geq e^{g(\bar{\xi}^\circ) \cdot |B(\text{int } \gamma)| - C(\tau) \cdot |\partial \text{int } \gamma|}$ , and
- $g(\bar{\xi}^\circ) \leq e^{-\tau/2} \leq 1$ .

Moreover

- $|\partial(V(\gamma) \setminus \text{int } \theta)| \leq |\partial V(\gamma)| + |\partial \text{int } \theta| \leq 3|\gamma| + \sum_{\gamma' \in \theta} |\gamma'|$ , and
- $|B(\gamma)| \leq |\gamma|$ .

Hence we have

$$\begin{aligned} \xi^\circ(\gamma) &\leq \rho(\gamma) \cdot e^{(1+4C(\tau)) \cdot |\gamma|} \sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} e^{|B(V(\gamma) \setminus \text{int } \theta)| \cdot [e(\mathbb{B}) - e(\emptyset) + g(\xi_{\text{small}}^{\text{d}}) - g(\bar{\xi}^\circ)]} \prod_{\gamma' \in \theta} \rho(\gamma') \cdot e^{2C(\tau) \cdot |\gamma'|} \\ &= \rho(\gamma) \cdot e^{(1+4C(\tau)) \cdot |\gamma|} \sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} e^{|B(V(\gamma) \setminus \text{int } \theta)| \cdot [(g(\xi_{\text{small}}^{\text{d}}) - g(\bar{\xi}^{\text{d}})) - (f^\circ - f^{\text{d}})]} \prod_{\gamma' \in \theta} \rho(\gamma') \cdot e^{2C(\tau) \cdot |\gamma'|}. \end{aligned} \quad (88)$$

As we shall see shortly, if  $(q+r)$  is large enough, the sum appearing in the above expression can be bounded by  $e^{3C(\tau) \cdot |\gamma|}$ , so that

$$\xi^\circ(\gamma) \leq \rho(\gamma) \cdot e^{(1+4C(\tau)) \cdot |\gamma|} e^{3C(\tau) \cdot |\gamma|} \quad (89)$$

$$= e^{-\left(\frac{1}{4} \log(q+r) - 1 - 7C(\tau)\right) \cdot |\gamma|} \quad (90)$$

For large  $(q+r)$ , the righthand side is bounded by  $e^{-\tau|\gamma|}$ , proving the claim.

It remains to show that for  $(q+r)$  sufficiently large,

$$\sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} e^{|B(V(\gamma) \setminus \text{int } \theta)| \cdot [(g(\xi_{\text{small}}^{\text{d}}) - g(\bar{\xi}^{\text{d}})) - (f^\circ - f^{\text{d}})]} \prod_{\gamma' \in \theta} \rho(\gamma') \cdot e^{2C(\tau) \cdot |\gamma'|} \leq e^{3C(\tau) \cdot |\gamma|}. \quad (91)$$

To show this, let us consider a contour model with weight function

$$\tilde{\rho}(\gamma') = \begin{cases} \rho(\gamma') \cdot e^{3C(\tau) \cdot |\gamma'|}, & \text{if } \gamma' \text{ big and disorder,} \\ 0, & \text{otherwise.} \end{cases} \quad (92)$$

Assuming that  $g(\widehat{\rho}) \leq f^o - f^d + g(\bar{\xi}^d) - g(\xi_{\text{small}}^d)$ , and  $(q+r)$  in such a way that  $\widehat{\rho}(\gamma') \leq e^{-\tau|\gamma'|}$ , we can use Proposition 5 to obtain

$$\begin{aligned}
& \sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} e^{|B(V(\gamma) \setminus \text{int } \theta)| \cdot [(g(\xi_{\text{small}}^d) - g(\bar{\xi}^d)) - (f^o - f^d)]} \prod_{\gamma' \in \theta} \rho(\gamma') \cdot e^{2C(\tau) \cdot |\gamma'|} \\
& \leq \sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} e^{-|B(V(\gamma) \setminus \text{int } \theta)| \cdot g(\widehat{\rho})} \prod_{\gamma' \in \theta} \rho(\gamma') \cdot e^{2C(\tau) \cdot |\gamma'|} \\
& = e^{-|B(V(\gamma))| \cdot g(\widehat{\rho})} \sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} \prod_{\gamma' \in \theta} \rho(\gamma') \cdot e^{2C(\tau) \cdot |\gamma'|} \cdot e^{|B(\text{int } \gamma')| \cdot g(\widehat{\rho})} \\
& \leq e^{-|B(V(\gamma))| \cdot g(\widehat{\rho})} \sum_{\substack{\theta \in \mathcal{E}_{V(\gamma)}^{\text{disord}} \\ \theta \text{ big}}} \prod_{\gamma' \in \theta} \rho(\gamma') \cdot e^{2C(\tau) \cdot |\gamma'|} \cdot \mathcal{Z}(\text{int } \gamma' | \widehat{\rho}) \cdot e^{C(\tau) \cdot |\gamma'|} \\
& = e^{-|B(V(\gamma))| \cdot g(\widehat{\rho})} \cdot \mathcal{Z}(V(\gamma) | \widehat{\rho}) \\
& \leq e^{3C(\tau) \cdot |\gamma|} .
\end{aligned} \tag{93}$$

Finally, to see that  $g(\widehat{\rho}) \leq f^o - f^d + g(\bar{\xi}^d) - g(\xi_{\text{small}}^d)$ , note that  $g(\bar{\xi}^d) - g(\xi_{\text{small}}^d) \geq 0$  (due to the fact that  $\xi_{\text{small}}^d \leq \bar{\xi}^d$ ) and by Proposition 5

$$g(\widehat{\rho}) \leq \sum_{\substack{\gamma': S(\gamma') \ni 0 \\ \gamma' \text{ big}}} \widehat{\rho}(\gamma') \leq e^{-\tau/(f^o - f^d)} , \tag{94}$$

using the fact that  $\gamma'$  is big only if  $|\gamma'| \geq \frac{2}{f^o - f^d}$ . For  $\tau$  not too small we have  $e^{-\tau/(f^o - f^d)} \leq f^o - f^d$ .  $\square$

From the above lemma, we know that the pressure  $f^{\text{RC}}(\beta)$  of the  $r$ -biased random-cluster representation coincides with the maximum of the functions  $f^o(\beta)$  and  $f^d(\beta)$ , that is,

$$f^{\text{RC}}(\beta) = \max \{ f^o(\beta), f^d(\beta) \} . \tag{95}$$

Recall that

$$f^o(\beta) = -e(\mathbb{B}) + g(\bar{\xi}^o) , \tag{96}$$

$$f^d(\beta) = -e(\emptyset) + g(\bar{\xi}^d) . \tag{97}$$

If  $\tau$  is large, Proposition 5 says that  $g(\bar{\xi}^o)$  and  $g(\bar{\xi}^d)$  are small, so that  $f^{\text{RC}}(\beta)$  can be nearly expressed in terms of the ‘‘energy’’ per bond of the fully ordered and fully disordered configurations. More precisely, if we define

$$F(\beta) \triangleq \max \{ -e(\mathbb{B}), -e(\emptyset) \} . \tag{98}$$

we have

$$0 \leq f^{\text{RC}}(\beta) - F(\beta) \leq e^{-\tau/2} . \tag{99}$$



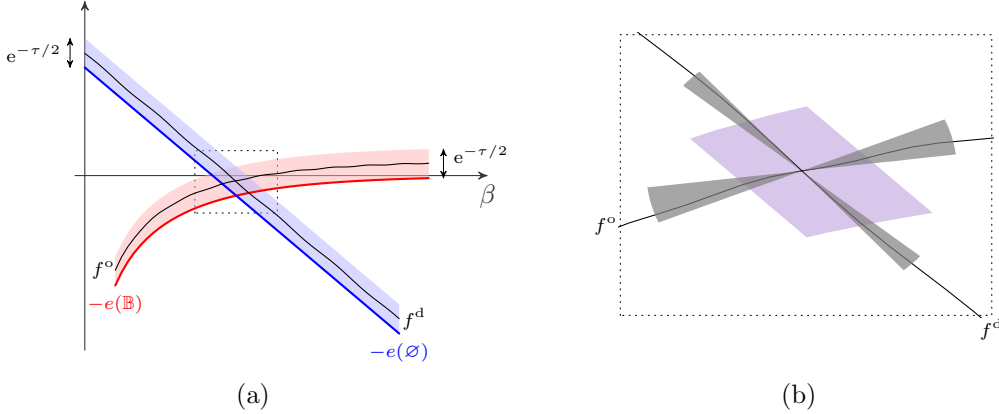


Figure 3: (a) The curves of  $f^o(\beta)$  and  $f^d(\beta)$  are within a narrow margin above  $-e(\mathbb{B})$  and  $-e(\emptyset)$ . (b) The slopes of  $f^o(\beta)$  and  $f^d(\beta)$  are close to the slopes of  $-e(\mathbb{B})$  and  $-e(\emptyset)$ .

The two curves  $-e(\mathbb{B})$  and  $-e(\emptyset)$  (as functions of  $\beta$ ) intersect at a single point

$$\bar{\beta}_c = \log(1 + \sqrt{q+r}) , \quad (100)$$

above which  $F(\beta) = -e(\mathbb{B})$  and below which  $F(\beta) = -e(\emptyset)$ . Furthermore, these two curves have significantly different slopes, implying that  $F(\beta)$  is not differentiable at  $\bar{\beta}_c$ . What we are after is to infer that  $f^{\text{RC}}(\beta)$  has a similar behavior. In other words, we would like to show that there exists a unique solution  $\beta_c$  for the equation  $f^o(\beta) = f^d(\beta)$ , at which  $f^{\text{RC}}(\beta)$  is not differentiable, above which  $f^{\text{RC}}(\beta) = f^o(\beta)$  and below which  $f^{\text{RC}}(\beta) = f^d(\beta)$ . Note that condition (99) guarantees that  $f^{\text{RC}}(\beta) = f^o(\beta) > f^d(\beta)$  for  $\beta \gg \bar{\beta}_c$  and  $f^{\text{RC}}(\beta) = f^d(\beta) > f^o(\beta)$  for  $\beta \ll \bar{\beta}_c$ . In fact, it states that  $f^{\text{RC}}(\beta)$  lives in a margin of width  $e^{-\tau/2}$  above  $F(\beta)$  (see Figure 3(a)). To infer a sharp transition, we further need to look at the angle with which the two functions  $f^o$  and  $f^d$  cross each other. This is addressed in the following lemma, which is analogous to Theorem 3.3 in [3].

In the sequel, we shall write  $\xi_\beta^o$  and  $\xi_\beta^d$  to emphasize the dependence of the weight functions on  $\beta$ .

**Lemma 7.** *Let  $\beta_2 > \beta_1 > 0$  be such that at each inverse temperatures  $\beta \in [\beta_1, \beta_2]$  the weight function  $\xi_\beta^o$  is damped (i.e.,  $\xi_\beta^o(\gamma) \leq e^{-\tau|\gamma|}$  for every order contour  $\gamma$ ). Then, the right and left derivatives of  $g(\xi_\beta^o)$  on  $[\beta_1, \beta_2]$  satisfy the following bound.*

$$\left| \frac{\partial}{\partial \beta^\pm} g(\xi_\beta^o) \right| \leq \frac{2}{e^\beta - 1} e^{-\tau/2} . \quad (101)$$

A similar bound holds for the left and right derivatives of  $g(\xi_\beta^d)$  on any interval in which  $\xi_\beta^d$  is damped.

*Proof.* From Lemma 1, together with equations (24), (25), and (26), we know that

$$g(\xi_\beta^o) = f^{\text{RC}}(\beta) + e(\mathbb{B}) = \frac{1}{2} f(\beta) - \beta - \log(1 - e^{-\beta}) . \quad (102)$$

It therefore follows from the convexity of the pressure function  $f(\beta)$  that the right and left derivatives of  $g(\xi_\beta^\circ)$  exist.

We now prove the claimed bound. For a finite volume  $\Lambda$ , if we denote by  $\Gamma(\Lambda)$  the set of all contours in  $\Lambda$ , we have

$$\begin{aligned}
\left| \frac{\partial}{\partial \beta} \log \mathcal{Z}(\Lambda | \xi_\beta^\circ) \right| &= \frac{1}{\mathcal{Z}(\Lambda | \xi_\beta^\circ)} \left| \sum_{\partial \in \mathcal{M}_\Lambda} \frac{\partial}{\partial \beta} \prod_{\gamma \in \partial} \xi_\beta^\circ(\gamma) \right| \\
&\leq \frac{1}{\mathcal{Z}(\Lambda | \xi_\beta^\circ)} \sum_{\partial \in \mathcal{M}_\Lambda} \sum_{\gamma \in \partial} \left| \frac{\partial \xi_\beta^\circ(\gamma)}{\partial \beta} \right| \prod_{\substack{\hat{\gamma} \in \partial \\ \hat{\gamma} \neq \gamma}} \xi_\beta^\circ(\hat{\gamma}) \\
&\leq \sum_{\gamma \in \Gamma(\Lambda)} \left| \frac{\partial \xi_\beta^\circ(\gamma)}{\partial \beta} \right| \frac{\mathcal{Z}(\text{int } \gamma | \xi_\beta^\circ) \cdot \mathcal{Z}(\Lambda \cap \text{ext } \gamma | \xi_\beta^\circ)}{\mathcal{Z}(\Lambda | \xi_\beta^\circ)} \\
&\leq \sum_{\gamma \in \Gamma(\Lambda)} \left| \frac{\partial \xi_\beta^\circ(\gamma)}{\partial \beta} \right|.
\end{aligned} \tag{103}$$

Using the definition of  $\xi_\beta^\circ(\gamma)$  (Eq. (45)) we have

$$\left| \frac{\partial \xi_\beta^\circ(\gamma)}{\partial \beta} \right| \leq \left| |B(\gamma)| \cdot \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \xi_\beta^\circ(\gamma) \right| + \rho(\gamma) \cdot e^{|B(\gamma)| \cdot e(\mathbb{B})} \cdot \left| \frac{\partial}{\partial \beta} \frac{Z^{\text{RC.disord}}(V(\gamma))}{Z^{\text{RC.ord}}(\text{int } \gamma)} \right|. \tag{104}$$

The derivative of the partition functions appearing on the righthand side can be bounded directly from the definitions (Eq. (32) and (33)) by

$$0 \leq \left| \frac{\partial}{\partial \beta} Z^{\text{RC.disord}}(V(\gamma)) \right| \leq -|B(V(\gamma))| \cdot \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot Z^{\text{RC.disord}}(V(\gamma)), \tag{105}$$

$$0 \leq \left| \frac{\partial}{\partial \beta} Z^{\text{RC.ord}}(\text{int } \gamma) \right| \leq -|B(\text{int } \gamma)| \cdot \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot Z^{\text{RC.ord}}(\text{int } \gamma), \tag{106}$$

leading to

$$\left| \frac{\partial}{\partial \beta} \frac{Z^{\text{RC.disord}}(V(\gamma))}{Z^{\text{RC.ord}}(\text{int } \gamma)} \right| \leq -(|B(V(\gamma))| + |B(\text{int } \gamma)|) \cdot \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \frac{Z^{\text{RC.disord}}(V(\gamma))}{Z^{\text{RC.ord}}(\text{int } \gamma)}. \tag{107}$$

Therefore, recalling the definition of  $\xi_\beta^\circ(\gamma)$  (Eq. (45)) and that  $|B(V(\gamma))| = |B(\gamma)| + |B(\text{int } \gamma)|$  we obtain

$$\left| \frac{\partial \xi_\beta^\circ(\gamma)}{\partial \beta} \right| \leq -2 |B(V(\gamma))| \cdot \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \xi_\beta^\circ(\gamma). \tag{108}$$

We can now write

$$\begin{aligned}
\left| \frac{\partial}{\partial \beta} \frac{1}{|B(\Lambda)|} \log \mathcal{Z}(\Lambda | \xi_\beta^\circ) \right| &\leq -2 \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \frac{1}{|B(\Lambda)|} \sum_{\gamma \in \Gamma(\Lambda)} |B(V(\gamma))| \cdot \xi_\beta^\circ(\gamma) \\
&= -2 \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \frac{1}{|B(\Lambda)|} \sum_{\gamma \in \Gamma(\Lambda)} \sum_{b \in B(V(\gamma))} \xi_\beta^\circ(\gamma)
\end{aligned} \tag{109}$$

$$\begin{aligned}
&= -2 \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \frac{1}{|B(\Lambda)|} \sum_{b \in B(\Lambda)} \sum_{\substack{\gamma \in \Gamma(\Lambda) \\ B(V(\gamma)) \ni b}} \xi_\beta^\circ(\gamma) \\
&\leq -2 \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \frac{1}{|B(\Lambda)|} \sum_{b \in B(\Lambda)} \sum_{\substack{\gamma \in \Gamma \\ B(V(\gamma)) \ni b}} \xi_\beta^\circ(\gamma) \\
&\leq -2 \frac{\partial e(\mathbb{B})}{\partial \beta} \cdot \frac{|B(\Lambda)|}{|B(\Lambda)|} \sum_{\substack{\gamma \in \Gamma \\ S(V(\gamma)) \ni 0}} \xi_\beta^\circ(\gamma) \\
&\leq \frac{2}{e^\beta - 1} e^{-\tau/2}.
\end{aligned}$$

For  $n \geq 0$ , let  $\Lambda_n$  denotes the central  $(2n+1) \times (2n+1)$  square in the lattice, and define

$$g_n(\xi_\beta^\circ) \triangleq \frac{1}{|B(\Lambda_n)|} \log \mathcal{Z}(\Lambda_n | \xi_\beta^\circ). \quad (110)$$

From Lemma 1, and equations (24) and (25), we know that  $g_n(\xi_\beta^\circ) \rightarrow g(\xi_\beta^\circ)$  as  $n \rightarrow \infty$ .

Let  $\beta \in [\beta_1, \beta_2)$ . For every  $x > 0$  such that  $\beta + x < \beta_2$ ,

$$|g(\xi_{\beta+x}^\circ) - g(\xi_\beta^\circ)| \leq \lim_{n \rightarrow \infty} \int_\beta^{\beta+x} \left| \frac{\partial}{\partial s} g_n(\xi_s^\circ) \right| ds \leq \frac{2}{e^\beta - 1} e^{-\tau/2} x. \quad (111)$$

Dividing by  $x$  and letting  $x \downarrow 0$ , we obtain

$$\left| \frac{\partial}{\partial \beta^+} g(\xi_\beta^\circ) \right| \leq \frac{2}{e^\beta - 1} e^{-\tau/2}. \quad (112)$$

The bound for the left derivatives at  $\beta \in (\beta_1, \beta_2]$  follows similarly.  $\square$

Using the above lemma, we show that the functions  $f^d$  and  $f^\circ$  cross each other at a single point  $\beta_c$  with different slopes. Although, intuitively this statement may sound clear, a subtle argument is needed to rule out unexpected possibilities (e.g., the possibility of having a dense set of crossings on an interval).

Let  $T^d$  be the set of inverse temperatures  $\beta \in [0, \infty)$  at which the weight function  $\xi_\beta^d$  is damped, that is,

$$T^d \triangleq \bigcap_{\gamma \in \Gamma} \left\{ \beta \in [0, \infty) : \xi_\beta^d(\gamma) \leq e^{-\tau|\gamma|} \right\}. \quad (113)$$

The set  $T^\circ$  is defined similarly. Note that  $T^d$  and  $T^\circ$  are closed in  $[0, \infty)$ , because the weights  $\xi_\beta^d(\gamma)$  and  $\xi_\beta^\circ(\gamma)$  are continuous functions of  $\beta$ . Let  $T \triangleq T^d \cap T^\circ$  be the set of inverse temperatures where both weight functions are damped.

The set  $(0, \infty) \setminus T$  is open and dense in  $(0, \infty)$ . To see the denseness, suppose to the contrary that there is an interval  $(\beta_1, \beta_2) \subset (0, \infty)$  such that  $T \supseteq (\beta_1, \beta_2)$ . Then, for every  $\beta \in [\beta_1, \beta_2]$ , both weight functions  $\xi_\beta^d$  and  $\xi_\beta^\circ$  are damped, which by Lemma 7 implies that

$$\left| \frac{\partial}{\partial \beta^\pm} g(\xi_\beta^d) \right|, \left| \frac{\partial}{\partial \beta^\pm} g(\xi_\beta^\circ) \right| \leq \frac{2}{e^\beta - 1} e^{-\tau/2}. \quad (114)$$

However, for  $\tau$  sufficiently large, this leads to a contradiction, because on the interval  $(\beta_1, \beta_2)$ ,  $f^d(\beta)$  and  $f^o(\beta)$  coincide with the function  $f^{\text{RC}}(\beta)$ , and hence cannot have different slopes.

For every  $\beta \in (0, \infty) \setminus T$ , either  $\xi_\beta^d$  is damped and  $\xi_\beta^o$  is not damped, or vice versa. Let  $D^d$  (resp.,  $D^o$ ) denote the set of  $\beta \in (0, \infty) \setminus T$  where  $\xi_\beta^d$  (resp.,  $\xi_\beta^o$ ) is damped. Then,  $D^d$  and  $D^o$  form a partitioning of  $(0, \infty) \setminus T$ . Furthermore,  $D^d$  and  $D^o$  are open. For, suppose that  $D^d$  is not open. Then, there exists at least a point  $\beta_0 \in D^d$  such that for every neighbourhood  $N \ni \beta_0$ , there is a point from  $T^o$  (Lemma 6). Since  $T^o$  is closed, it must contain  $\beta_0$ , hence a contradiction. A similar argument shows that  $D^o$  is open.

We now claim that the pair  $(D^d, D^o)$  is a Dedekind cut of  $(0, \infty) \setminus T$ ; that is, for every  $\beta_1 \in D^d$  and every  $\beta_2 \in D^o$  we have  $\beta_1 < \beta_2$ . For, let  $\beta_1 \in D^d$  and  $\beta_2 \in D^o$ . Then, there exist open neighbourhoods  $N_{\beta_1} \ni \beta_1$  and  $N_{\beta_2} \ni \beta_2$  such that  $N_{\beta_1} \subseteq D^d$  and  $N_{\beta_2} \subseteq D^o$ . Using Lemma 7, the above implies that

$$\frac{\partial}{\partial \beta^-} f^d(\beta_1) = \frac{\partial}{\partial \beta^-} f^{\text{RC}}(\beta_1) < \frac{\partial}{\partial \beta^+} f^{\text{RC}}(\beta_2) = \frac{\partial}{\partial \beta^+} f^o(\beta_2) \quad (115)$$

which by the convexity of  $f^{\text{RC}}$  implies that  $\beta_1 < \beta_2$ .

Let us define

$$\beta_c \triangleq \sup D^d = \inf D^o . \quad (116)$$

The equality of the supremum and infimum follows from the denseness of  $D^d \cup D^o$  in  $(0, \infty)$ . The denseness of  $D^d \cup D^o$  in  $(0, \infty)$  and the fact that  $(D^d, D^o)$  is a Dedekind cut also imply that  $D^d$  and  $D^o$  are dense, respectively in  $[0, \beta_c]$  and  $[\beta_c, \infty)$ . Therefore,  $T^d \supseteq \overline{D^d} = [0, \beta_c]$  and  $T^o \supseteq \overline{D^o} = [\beta_c, \infty)$ . This, together with Lemma 7, concludes that

$$\frac{\partial}{\partial \beta^-} f^d(\beta_c) = \frac{\partial}{\partial \beta^-} f^{\text{RC}}(\beta_c) < \frac{\partial}{\partial \beta^+} f^{\text{RC}}(\beta_c) = \frac{\partial}{\partial \beta^+} f^o(\beta_c) . \quad (117)$$

## 6 Proof of the Main Result

In the previous sections, we have introduced the main ingredients to prove the occurrence of a first-order transition. Below, we first put these ingredients together so as to obtain a recipe for the proof. Afterwards, we shall see how these ingredients, along with basic properties of the biased random-cluster model (see Appendix A.3), can be used to prove the symmetry breaking at the transition temperature.

*Proof of the Theorem.* The first step was to reduce the partition function of the  $(q, r)$ -Potts model to the partition function of the  $r$ -biased random-cluster model. This was done for the free and homogeneous visible boundary conditions, which led to the disordered and ordered boundary conditions for the  $r$ -biased random-cluster model (see Eq. (18) and (21)). By means of this we could rewrite the pressure  $f(\beta)$  for the  $(q, r)$ -Potts model as

$$f(\beta) = 2\beta + 2f^{\text{RC}}(\beta) , \quad (118)$$

where

$$f^{\text{RC}}(\beta) = \lim_{n \rightarrow \infty} \frac{\log Z_{p\beta, q, r}^{\text{RC, ord}}(\Lambda_n)}{|B(\Lambda_n)|} = \lim_{n \rightarrow \infty} \frac{\log Z_{p\beta, q, r}^{\text{RC, disord}}(\Lambda_n)}{|B(\Lambda_n)|} . \quad (119)$$

The second step consisted in re-expressing the partition functions  $Z^{\text{RC.ord}}$  and  $Z^{\text{RC.disord}}$  in terms of the partition functions of two abstract contour models  $\mathcal{Z}(\cdot | \xi^\circ)$  and  $\mathcal{Z}(\cdot | \xi^{\text{d}})$  (Lemma 1) so that we could write

$$f^{\text{RC}}(\beta) = -e(\mathbb{B}) + g(\xi^\circ) = -e(\emptyset) + g(\xi^{\text{d}}), \quad (120)$$

where  $g(\xi^\circ)$  and  $g(\xi^{\text{d}})$  are the pressure functions for the two contour models (Eq. (69)).

If the weight function  $\chi$  of a contour model is sufficiently “damped” (i.e.,  $\chi(\gamma) \leq e^{-\tau|\gamma|}$  for  $\tau$  large enough), the corresponding pressure  $g(\chi)$  can be made arbitrarily small (Proposition 5). In order to exploit this result, in the third step, we truncated the weight functions  $\xi^\circ$  and  $\xi^{\text{d}}$  so as to render them artificially damped (Eq. (71)). We could then define two functions

$$f^\circ(\beta) = -e(\mathbb{B}) + g(\bar{\xi}^\circ), \quad (121)$$

$$f^{\text{d}}(\beta) = -e(\emptyset) + g(\bar{\xi}^{\text{d}}), \quad (122)$$

which approximate  $f^{\text{RC}}(\beta)$  from below, and which can be thought of (for sufficiently large  $\tau$ ) as perturbations of the functions  $-e(\mathbb{B})$  and  $-e(\emptyset)$ , respectively (see Figure 3(a)).

In the fourth step, we proved that, for  $(q+r)$  large (relative to  $\tau$ ), the pressure  $f^{\text{RC}}(\beta)$  is the maximum of these two approximations. This was achieved by proving that whenever  $f^\circ \geq f^{\text{d}}$ , the weight function  $\xi^\circ$  is “naturally” damped (i.e.,  $\xi^\circ = \bar{\xi}^\circ$ ) and vice versa (Lemma 6). Therefore,  $f^{\text{RC}}(\beta)$  can be closely approximated by the maximum between  $-e(\mathbb{B})$  and  $-e(\emptyset)$ . Due to the continuity of the pressure functions, the latter implies that the curves  $f^\circ(\beta)$  and  $f^{\text{d}}(\beta)$  intersect.

In the last step, we showed that for  $\tau$  sufficiently large, there exists a unique point  $\beta_c \in (0, \infty)$  at which the two function  $f^{\text{d}}(\beta)$  and  $f^\circ(\beta)$  cross each other with a positive angle, and such that for every  $\beta \leq \beta_c$ ,  $f^{\text{RC}}(\beta) = f^{\text{d}}(\beta)$  and for every  $\beta \geq \beta_c$ ,  $f^{\text{RC}}(\beta) = f^\circ(\beta)$  (Figure 3(b), Lemma 7 and the discussion afterwards). Hence, the  $(q, r)$ -Potts model undergoes a first-order phase transition at  $\beta_c$ .

We now prove the breaking of permutation symmetry at the transition temperature. Note that if  $\beta \geq \beta_c$ , we have  $f^\circ(\beta) \geq f^{\text{d}}(\beta)$  and therefore, in view of Lemma 6, the weights  $\xi^\circ$  satisfy  $\xi^\circ(\gamma) \leq e^{-\tau|\gamma|}$ . Since  $\tau$  was chosen large, for every finite set of sites  $A$ , the sum  $\sum_{\gamma \in \Gamma_A} \xi^\circ(\gamma)$  converges. Hence, it follows from Corollary 4 that

$$\phi_{p_\beta}^{\text{ord}} \left( \begin{array}{l} \exists \text{ unique infinite sea of order} \\ \text{with finite islands of disorder} \end{array} \right) = 1. \quad (123)$$

(The subscript  $p_\beta$  is added to emphasize the dependence on  $\beta$ .) Similarly, if  $\beta \leq \beta_c$ , the sum  $\sum_{\gamma \in \Gamma_A} \xi^{\text{d}}(\gamma)$  converges and thus

$$\phi_{p_\beta}^{\text{disord}} \left( \begin{array}{l} \exists \text{ unique infinite sea of disorder} \\ \text{with finite islands of order} \end{array} \right) = 1. \quad (124)$$

For a visible colour  $k$ , let  $\mu_\beta^k$  be, as introduced in Section 2, a weak limit of Boltzmann distributions with homogeneous boundary conditions  $\omega^k$ . Likewise, let  $\mu_\beta^{\text{free}}$  be a weak limit of free-boundary Boltzmann distributions.

Similar to the finite-volume couplings, there exists a coupling of  $\mu_\beta^k$  and  $\phi_{p_\beta}^{\text{ord}}$  with the property that with probability 1, every site incident to an infinite connected component of present bonds has colour  $k$  (see Appendix A.3). If  $\beta \geq \beta_c$ , we know that almost surely the bond configuration

consists of a unique sea of order with finite islands of disorder. In particular, the probability that a given site  $i$  takes a colour other than  $k$  is bounded by the probability that site  $i$  is surrounded by an order contour; that is,

$$\mu_\beta^k(\{\sigma : \sigma_i \neq k\}) \leq \phi_{p_\beta}^{\text{ord}}\{X : \bar{\partial}X \cap \Gamma_i \neq \emptyset\}. \quad (125)$$

In this region of  $\beta$ , Corollary 4 and Lemma 6 ensure that  $\phi_{p_\beta}^{\text{ord}}\{X : \bar{\partial}X \cap \Gamma_i \neq \emptyset\}$  can be made arbitrarily small by tuning  $\tau$ . Hence, for every  $\varepsilon > 0$ , choosing  $q + r$  large enough, we have  $\mu_\beta^k(\{\sigma : \sigma_i = k\}) > 1 - \varepsilon$ .

The measures  $\mu_\beta^{\text{free}}$  and  $\phi_{p_\beta}^{\text{disord}}$  can also be coupled, in such a way that, given a configuration of bonds, the colour of the isolated sites are chosen independently and uniformly among the  $q + r$  possibilities. Using this coupling, and conditioning on whether a given site  $i$  is isolated or not, we obtain

$$\mu_\beta^{\text{free}}(\{\sigma : \sigma_i = k\}) \leq \frac{1}{q + r} + \phi_{p_\beta}^{\text{disord}}\{X : i \text{ not isolated in } (\mathbb{S}, X)\}. \quad (126)$$

If  $\beta \leq \beta_c$ , the bond configuration almost surely consists of a unique sea of disorder with finite islands of order. Hence, the probability that site  $i$  is not isolated is bounded by the probability that site  $i$  is surrounded by a disorder contour; that is,

$$\phi_{p_\beta}^{\text{disord}}\{X : i \text{ not isolated in } (\mathbb{S}, X)\} \leq \phi_{p_\beta}^{\text{disord}}\{X : \bar{\partial}X \cap \Gamma_i \neq \emptyset\}. \quad (127)$$

As in the previous case, Corollary 4 and Lemma 6 guarantee that for every  $\varepsilon > 0$ , choosing  $q + r$  large enough,  $\mu_\beta^{\text{free}}(\{\sigma : \sigma_i = k\}) < \varepsilon$ .

It remains to show that for  $\beta < \beta_c$ , the  $(q, r)$ -Potts Gibbs measure is unique.

As in the standard random-cluster model, there exists a critical value  $0 < p_c < 1$  such that

- for  $p < p_c$ , almost surely with respect to  $\phi_p^{\text{ord}}$  and  $\phi_p^{\text{disord}}$ , there is no infinite connected component of bonds (order does not “percolate”), whereas
- for  $p > p_c$ , the event that a given site is in an infinite connected component happens with positive probability under both  $\phi_p^{\text{ord}}$  and  $\phi_p^{\text{disord}}$ .

(See Appendix A.3.) It follows that  $p_c = p_{\beta_c}$ . Namely, if  $p_\beta < p_c$ , then order does not percolate under  $\phi_{p_\beta}^{\text{ord}}$ . Therefore, equation (123) does not hold, implying that  $\beta < \beta_c$ . Conversely, if  $p_\beta > p_c$ , then order percolates with positive probability under  $\phi_{p_\beta}^{\text{disord}}$ , refuting (124). Hence, we must have  $\beta > \beta_c$ .

On the other hand, for every  $\beta$  at which

$$\phi_{p_\beta}^{\text{ord}}(\exists \text{ an infinite connected component of bonds}) = 0, \quad (128)$$

the measure  $\mu_\beta^{\text{free}}$  is the only Gibbs measure for the  $(q, r)$ -Potts model (see Appendix A.3). The latter condition is guaranteed whenever  $p_\beta < p_c$ , which is equivalent to  $\beta < \beta_c$ . Thus the uniqueness of Gibbs measure for  $\beta < \beta_c$  follows.  $\square$

## 7 Conclusion

In this paper, we presented a proof that the two-dimensional Potts model with  $q$  visible colours and  $r$  invisible colours undergoes a first-order phase transition in temperature accompanied by a  $q$ -fold symmetry breaking, provided the number of invisible colours is large enough. On the other hand, for  $r = 0$  (no invisible colours), the model reduces to the standard  $q$ -colour Potts model, for which it is known that if  $q = 2, 3, 4$ , the transition in two dimensions is second-order. Tamura, Tanaka and Kawashima [21, 22, 23] introduced the Potts model with  $r$  invisible colours as a simple two-dimensional example with short-range interactions in which, tuning the parameter  $r$ , the same symmetry breaking could accompany phase transitions of different orders.

The impossibility to infer the order of the phase transition from the broken symmetry was already noticed in other examples, such as the two-dimensional 3-colour Kac-Potts model [10]. For this model, Gobron and Merola proved that a 3-fold symmetry breaking might be accompanied with either a first-order or a second-order phase transition, by changing the finite range of the interactions. It is also worth mentioning that the Potts model with invisible colours can be embedded in the so-called Potts lattice gas model, studied in [2]. The Potts lattice gas model is another extension of the standard Potts model that, instead of invisible colours, allows for vacant sites. In this model, the role of the parameter  $r$  is played by a chemical potential that controls the density of the occupied sites. Depending on the value of the chemical potential, the Potts lattice gas model may undergo either a first or a higher order phase transition in temperature.

The first-order phase transition in the  $(q, r)$ -Potts model occurs as long as  $q + r$  is large enough. In particular, even for small values of  $q$  (say,  $q = 1, 2, 3, 4$ ), the presence of many invisible colours assures a first-order transition. The argument is very similar to the one for the standard  $q$ -Potts model, in which  $q$  is required to be large [17, 16]. The transition point is asymptotically (in  $q + r$ ) given by  $\beta_c \approx \frac{1}{2} \log(q + r)$ . For  $q + r$  large, the latent heat is approximately given by  $2 \left( -\frac{\partial e(\emptyset)}{\partial \beta} + \frac{\partial e(\mathbb{B})}{\partial \beta} \right) = 2 + \frac{2}{\sqrt{q+r}}$ , which tends to 2 as  $q + r \rightarrow \infty$ .

The proof relies on a formulation of the Potts model with invisible colours in terms of a variant of the random-cluster model, which we named the biased random-cluster model. The difference between this new model and the original random-cluster model is that it weights singleton connected components differently from non-singleton connected components. Such a disparity allows one to increase the entropy by increasing the number of invisible colours, while keeping the number of ground states (i.e., the number of visible colours) unchanged. The random-cluster representation allows for a clear formulation of order and disorder: order is associated with the presence of bonds while disorder with the absence of bonds. This leads to a simple notion of contours describing the interface between order and disorder. Hence, the random-cluster representation lends itself to a Pirogov-Sinai analysis, which is used to prove the existence of a first-order phase transition.

We remark that the above analysis extends to higher dimensions. In fact, in the performed analysis the dimension entered mainly in the counting arguments, which can be refined in higher dimensions. The two-dimensional case, however, has the asset of simplifying those counting arguments, and it allows for a much easier geometric visualization of the contours.

## A Appendix

### A.1 Derivation of the Biased Random-Cluster Representation

To derive the relation (18), we start from (17) and write

$$Z_\beta(\Lambda_n) = e^{\beta|B(\Lambda_n)|} \cdot Z_{p_\beta, q, r}^{\text{RC}}(\Lambda_n) \quad (129)$$

$$= \sum_{X \subseteq B(\Lambda_n)} (e^\beta - 1)^{|X|} (q+r)^{\kappa_0(S(\Lambda_n), X)} q^{\kappa_1(S(\Lambda_n), X)} \quad (130)$$

$$= (q+r)^{-|S(\Lambda_{n+1} \setminus \Lambda_n)|} \sum_{\substack{Y \subseteq B(\Lambda_{n+1}) \\ Y \cap (B(\Lambda_{n+1}) \setminus B(\Lambda_n)) = \emptyset}} (e^\beta - 1)^{|Y|} (q+r)^{\kappa_0(S(\Lambda_{n+1}), Y)} q^{\kappa_1(S(\Lambda_{n+1}), Y)} \quad (131)$$

$$= (q+r)^{-|S(\Lambda_{n+1} \setminus \Lambda_n)|} \cdot e^{\beta|B(\Lambda_{n+1})|} \cdot Z_{p_\beta, q, r}^{\text{RC, disord}}(\Lambda_{n+1}) . \quad (132)$$

To obtain the relation (21), we need to take the homogeneous boundary condition for the  $(q, r)$ -Potts model into account. Denoting the set of  $(q, r)$ -Potts configurations on  $\Lambda_n$  by  $\Omega_{\Lambda_n}$ , we start from the definition (5) and write

$$Z_\beta^{\omega^k}(\Lambda_n) = \sum_{\sigma \in \Omega_{\Lambda_n}} \exp \left\{ \beta \sum_{\{i, j\} \in B(\Lambda_n)} \delta(\sigma_i = \sigma_j \leq q) \quad + \quad \beta \sum_{\substack{\{i, j\} \in B(\Lambda_{n+1}) \\ i \in S(\Lambda_n), j \notin S(\Lambda_n)}} \delta(\sigma_i = \omega_j^k = k) \right\} \quad (133)$$

$$= \sum_{\sigma \in \Omega_{\Lambda_n}} \prod_{\{i, j\} \in B(\Lambda_n)} (1 + \delta(\sigma_i = \sigma_j \leq q)(e^\beta - 1)) \prod_{\substack{\{i, j\} \in B(\Lambda_{n+1}) \\ i \in S(\Lambda_n), j \notin S(\Lambda_n)}} (1 + \delta(\sigma_i = k)(e^\beta - 1)) . \quad (134)$$

Denoting

$$\partial\Lambda_n \triangleq \{\{i, j\} \in B(\Lambda_{n+1}) : i \in S(\Lambda_n) \text{ and } j \notin S(\Lambda_n)\} , \quad (135)$$

we can expand the products to obtain

$$Z_\beta^{\omega^k}(\Lambda_n) = \sum_{\sigma \in \Omega_{\Lambda_n}} \sum_{X_1 \subseteq B(\Lambda_n)} \sum_{X_2 \subseteq \partial\Lambda_n} (e^\beta - 1)^{|X_1| + |X_2|} \delta(\sigma \in \Xi_k(X_1, X_2)) , \quad (136)$$

where

$$\begin{aligned} \Xi_k(X_1, X_2) \\ \triangleq \{ \sigma \in \Omega_{\Lambda_n} : \sigma_i = \sigma_j \leq q \text{ for all } \{i, j\} \in X_1 \text{ and } \sigma_i = k \text{ for all } i \in S(\Lambda_n) \cap S(X_2) \} . \end{aligned} \quad (137)$$

To impose the ordered boundary condition, we multiply and divide by  $(e^\beta - 1)^{|B(\Lambda_{n+1} \setminus \Lambda_n)|}$  to emulate the presence of the bonds in  $B(\Lambda_{n+1} \setminus \Lambda_n)$ . This gives

$$Z_\beta^{\omega^k}(\Lambda_n) = (e^\beta - 1)^{-|B(\Lambda_{n+1} \setminus \Lambda_n)|} \sum_{\sigma \in \Omega_{\Lambda_n}} \sum_{\substack{X \subseteq B(\Lambda_{n+1}) \\ X \supseteq B(\Lambda_{n+1} \setminus \Lambda_n)}} (e^\beta - 1)^{|X|} \delta(\tilde{\sigma} \in \Theta_k(X)) , \quad (138)$$



where  $\tilde{\sigma}$  is the extension of  $\sigma$  to a configuration in  $\Omega_{\Lambda_{n+1}}$  with  $\tilde{\sigma}_i = k$  for  $i \in S(\Lambda_{n+1} \setminus \Lambda_n)$  and

$$\begin{aligned} & \Theta_k(X) \\ & \triangleq \{ \tilde{\sigma} \in \Omega_{\Lambda_{n+1}} : \tilde{\sigma}_i = \tilde{\sigma}_j \leq q \text{ for all } \{i, j\} \in X \text{ and } \tilde{\sigma}_i = k \text{ for all } i \in S(\Lambda_{n+1} \setminus \Lambda_n) \} . \end{aligned} \quad (139)$$

Changing the order of the sums gives

$$Z_{\beta}^{\omega^k}(\Lambda_n) = (e^{\beta} - 1)^{-|B(\Lambda_{n+1} \setminus \Lambda_n)|} \sum_{\substack{X \subseteq B(\Lambda_{n+1}) \\ X \supseteq B(\Lambda_{n+1} \setminus \Lambda_n)}} (e^{\beta} - 1)^{|X|} |\Theta_k(X)| . \quad (140)$$

Note that for  $X$  satisfying  $B(\Lambda_{n+1} \setminus \Lambda_n) \subseteq X \subseteq B(\Lambda_n)$ , the size of  $\Theta_k(X)$  is

$$q^{-1} \cdot (q + r)^{\kappa_0(S(\Lambda_{n+1}), X)} q^{\kappa_1(S(\Lambda_{n+1}), X)} \quad (141)$$

we obtain

$$Z_{\beta}^{\omega^k}(\Lambda_n) = q^{-1} (e^{\beta} - 1)^{-|B(\Lambda_{n+1} \setminus \Lambda_n)|} \sum_{\substack{X \subseteq B(\Lambda_{n+1}) \\ X \supseteq B(\Lambda_{n+1} \setminus \Lambda_n)}} (e^{\beta} - 1)^{|X|} \cdot (q + r)^{\kappa_0(S(\Lambda_{n+1}), X)} q^{\kappa_1(S(\Lambda_{n+1}), X)} \quad (142)$$

$$= q^{-1} \cdot (e^{\beta} - 1)^{-|B(\Lambda_{n+1} \setminus \Lambda_n)|} \cdot e^{\beta|B(\Lambda_{n+1})|} \cdot Z_{p_{\beta}, q, r}^{\text{RC.ord}}(\Lambda_{n+1}) . \quad (143)$$

## A.2 Derivation of the Contour Representation

We show that for  $\Lambda = \Lambda_{n+1}$ , the definitions (19) and (29) (resp., (22) and (30)) agree.

The weight of a configuration  $X \subseteq B(\Lambda_{n+1})$  is

$$p_{\beta}^{|X|} (1 - p_{\beta})^{|B(\Lambda_{n+1}) \setminus X|} (q + r)^{\kappa_0(S(\Lambda_{n+1}), X)} q^{\kappa_1(S(\Lambda_{n+1}), X)} . \quad (144)$$

For a configuration  $X$  in  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  or  $\mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$ , let  $\partial_X$  be the corresponding contour family in  $\Delta_{\Lambda_{n+1}}^{\text{disord}}$  or  $\Delta_{\Lambda_{n+1}}^{\text{ord}}$ . More precisely,  $\partial_X \triangleq \bar{\partial}X$  if  $\mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  and  $\partial_X \triangleq \partial(X \cup B(\Lambda_{n+1})^{\complement})$  if  $\mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$ .

**Claim.** For  $X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  we have the relation

$$2|B(\Lambda_{n+1}) \setminus X| = 4\kappa_0(S(\Lambda_{n+1}), X) + \sum_{\gamma \in \partial_X} |\gamma| - |\bar{\partial}B(\Lambda_{n+1})| , \quad (145)$$

and for  $X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$  we have

$$2|B(\Lambda_{n+1}) \setminus X| = 4\kappa_0(S(\Lambda_{n+1}), X) + \sum_{\gamma \in \partial_X} |\gamma| . \quad (146)$$

*Proof.* We first decompose the set  $\{(i, b) : b \in B(\Lambda_{n+1}) \setminus X \text{ and } i \sim b\}$  as

$$\{(i, b) : b \in B(\Lambda_{n+1}) \setminus X \text{ and } i \sim b \text{ and } i \in S(X)\} \quad (147)$$

\(\cup\)

$$\{(i, b) : b \in B(\Lambda_{n+1}) \setminus X \text{ and } i \sim b \text{ and } i \notin S(X)\} , \quad (148)$$

and furthermore note that the latter set can be expressed as

$$\{(i, b) : i \in S(\Lambda_{n+1}) \setminus S(X) \text{ and } i \sim b\} \quad (149)$$

$$\setminus \{(i, b) : i \in S(\Lambda_{n+1}) \setminus S(X) \text{ and } i \sim b \text{ and } b \notin B(\Lambda_{n+1})\} . \quad (150)$$

We have

$$|\{(i, b) : b \in B(\Lambda_{n+1}) \setminus X \text{ and } i \sim b\}| = 2 |B(\Lambda_{n+1}) \setminus X| , \quad (151)$$

$$|\{(i, b) : b \in B(\Lambda_{n+1}) \setminus X \text{ and } i \sim b \text{ and } i \in S(X)\}| = \sum_{\gamma \in \partial_X} |\gamma| , \quad (152)$$

$$|\{(i, b) : i \in S(\Lambda_{n+1}) \setminus S(X) \text{ and } i \sim b\}| = 4\kappa_0(S(\Lambda_{n+1}), X) , \quad (153)$$

and the cardinality of  $\{(i, b) : i \in S(\Lambda_{n+1}) \setminus S(X) \text{ and } i \sim b \text{ and } b \notin B(\Lambda_{n+1})\}$  equals

$$\begin{cases} |\bar{\partial}B(\Lambda_{n+1})| , & \text{if } X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{disord}} , \\ 0 , & \text{if } X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{ord}} . \end{cases} \quad (154)$$

□

Using the relations (145) and (146) the weight of  $X$  takes the form

$$(q+r)^{\frac{1}{4}|\bar{\partial}B(\Lambda_{n+1})|} e^{-e(\mathbb{B}) \cdot |X|} e^{-e(\emptyset) \cdot |B(\Lambda_{n+1}) \setminus X|} (q+r)^{-\frac{1}{4} \sum_{\gamma \in \partial_X} |\gamma|} q^{\kappa_1(S(\Lambda_{n+1}), X)} \quad (155)$$

if  $X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$  and

$$e^{-e(\mathbb{B}) \cdot |X|} e^{-e(\emptyset) \cdot |B(\Lambda_{n+1}) \setminus X|} (q+r)^{-\frac{1}{4} \sum_{\gamma \in \partial_X} |\gamma|} q^{\kappa_1(S(\Lambda_{n+1}), X)} \quad (156)$$

if  $X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$ . If  $X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}$ , every non-singleton connected component in  $(S(\Lambda_{n+1}), X)$  contains all the sites of a unique disorder contour in  $\partial_X$ , so that the number of disorder contours in  $\partial_X$  is the same as  $\kappa_1(S(\Lambda_{n+1}), X)$ , and we have

$$(q+r)^{-\frac{1}{4} \sum_{\gamma \in \partial_X} |\gamma|} q^{\kappa_1(S(\Lambda_{n+1}), X)} = \prod_{\gamma \in \partial_X} \rho(\gamma) . \quad (157)$$

On the other hand, if  $X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}$ , the outermost non-singleton connected component of  $(S(\Lambda_{n+1}), X)$  has no associated disorder contour in  $\partial_X$ , thus

$$(q+r)^{-\frac{1}{4} \sum_{\gamma \in \partial_X} |\gamma|} q^{\kappa_1(S(\Lambda_{n+1}), X)} = q \prod_{\gamma \in \partial_X} \rho(\gamma) . \quad (158)$$

In conclusion, summing over all configurations, we obtain

$$Z^{\text{RC.disord}}(\Lambda_{n+1}) = (q+r)^{\frac{|\bar{\partial}B(\Lambda_{n+1})|}{4}} \sum_{X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{disord}}} e^{-|X| \cdot e(\mathbb{B}) - |B(\Lambda_{n+1}) \setminus X| \cdot e(\emptyset)} \prod_{\gamma \in \partial} \rho(\gamma) , \quad (159)$$

$$Z^{\text{RC.ord}}(\Lambda_{n+1}) = q \sum_{X \in \mathcal{X}_{\Lambda_{n+1}}^{\text{ord}}} e^{-|X| \cdot e(\mathbb{B}) - |B(\Lambda_{n+1}) \setminus X| \cdot e(\emptyset)} \prod_{\gamma \in \partial} \rho(\gamma) . \quad (160)$$

We remark that the definitions (19) and (22) may also be extended to general finite volumes of the lattice in a compatible fashion. Namely, if for a volume  $\Lambda$ , we define

$$\mathcal{X}_\Lambda^{\text{disord}} \triangleq \{X \subseteq B(\Lambda) : \bar{\partial}X \in \Delta_\Lambda^{\text{disord}}\}, \quad (161)$$

$$\mathcal{X}_\Lambda^{\text{ord}} \triangleq \{X \subseteq B(\Lambda) : \bar{\partial}(X \cup B(\Lambda)^c) \in \Delta_\Lambda^{\text{ord}}\}, \quad (162)$$

the compatibility of the definitions can be verified similarly.

### A.3 A Few Properties of the Biased Random-Cluster Model

Much information about the standard Potts model can be detected by studying the corresponding random-cluster model (see [9, 11, 12]). Many of the properties of the standard random-cluster model can be extended to the biased random-cluster model. These properties, in turn, can be used, in a similar fashion, to obtain information about the Potts model with invisible colours. In this appendix, we briefly sketch some of these properties that we exploit in the proof of the main theorem. The proofs are straightforward modifications of the standard case, which can be found in [9, 11, 12].

Let  $\mathbb{G} = (S, B)$  be a finite graph. The configurations of the biased random-cluster model on  $\mathbb{G}$  can be ordered according to the inclusion ordering. A configuration  $X \subseteq B$  is considered to be *smaller than or equal to* a configuration  $Y \subseteq B$ , if and only if every bond present in  $X$  is also present in  $Y$ . An event  $\mathcal{E} \subseteq 2^B$  is *increasing* if for every two configurations  $X$  and  $Y$  such that  $X \in \mathcal{E}$  and  $Y \supseteq X$ , we have  $Y \in \mathcal{E}$ . We say that a probability distribution  $\nu$  on  $2^B$  is *positively correlated*, if  $\nu(\mathcal{E}_1 \cap \mathcal{E}_2) \geq \nu(\mathcal{E}_1)\nu(\mathcal{E}_2)$  for every two increasing events  $\mathcal{E}_1, \mathcal{E}_2 \subseteq 2^B$ . The inclusion ordering on the configuration space  $2^B$  induces an ordering on the space of probability distributions on  $2^B$ . If  $\nu_1$  and  $\nu_2$  are probability distributions on  $2^B$ , we write  $\nu_1 \preceq \nu_2$  if  $\nu_1(\mathcal{E}) \leq \nu_2(\mathcal{E})$  for every increasing event  $\mathcal{E} \subseteq 2^B$ . In this case, we say that  $\nu_1$  is *stochastically dominated* by  $\nu_2$ .

For every  $0 < p < 1$ ,  $q \geq 1$  and  $r \geq 0$ , the  $r$ -biased random-cluster distribution  $\phi_{p,q,r}$  on  $\mathbb{G}$  is positively correlated. This follows from the Fortuin-Kasteleyn-Ginibre theorem (Theorem 4.11 in [9]; see Corollary 6.7). It follows that, if  $\mathcal{E}$  is an increasing (resp., decreasing) event with  $\phi_{p,q,r}(\mathcal{E}) > 0$ , then the conditional distribution  $\phi_{p,q,r}(\cdot | \mathcal{E})$  stochastically dominates (resp., is dominated by)  $\phi_{p,q,r}$ . Furthermore, if  $0 < p_1 \leq p_2 < 1$ , it follows from Holley's theorem (Theorem 4.8 in [9]) that  $\phi_{p_1,q,r} \preceq \phi_{p_2,q,r}$  (see Corollary 6.7 in [9]).

Let  $\Lambda$  be a finite volume in the lattice and  $\phi_\Lambda$  the biased random-cluster distribution on  $\Lambda$  (as a graph, without boundary condition). Let us denote by  $\phi_\Lambda^{\text{ord}}$  and  $\phi_\Lambda^{\text{disord}}$  the biased random-cluster distributions on  $\Lambda$  with ordered and disordered boundary conditions, respectively. By an application of the positive correlation property of  $\phi_\Lambda$  we have

$$\phi_\Lambda^{\text{disord}} \preceq \phi_\Lambda \preceq \phi_\Lambda^{\text{ord}}. \quad (163)$$

Moreover, by a further application of the Fortuin-Kasteleyn-Ginibre theorem, the distributions  $\phi_\Lambda^{\text{disord}}$  and  $\phi_\Lambda^{\text{ord}}$  are also positively correlated. This implies that if  $\Lambda_1$  is a sub-volume of  $\Lambda_2$ , we have

$$\phi_{\Lambda_1}^{\text{disord}} \preceq \phi_{\Lambda_2}^{\text{disord}} \quad \text{and} \quad \phi_{\Lambda_1}^{\text{ord}} \succeq \phi_{\Lambda_2}^{\text{ord}}. \quad (164)$$

As in Lemma 6.8 in [9], this implies that the weak limits

$$\phi^{\text{disord}} \triangleq \lim_{\Lambda \uparrow \mathbb{L}} \phi_\Lambda^{\text{disord}} \quad \text{and} \quad \phi^{\text{ord}} \triangleq \lim_{\Lambda \uparrow \mathbb{L}} \phi_\Lambda^{\text{ord}} \quad (165)$$

exist, where the limit  $\Lambda \uparrow \mathbb{L}$  can be taken along the net of all finite volumes in  $\mathbb{L}$  with the inclusion ordering.

To emphasize the dependence on parameter  $p$ , let us write  $\phi_{\Lambda,p}^{\text{ord}}$  and  $\phi_{\Lambda,p}^{\text{disord}}$  for the biased random-cluster distributions with parameter  $p$ . Then, by an application of Holley's theorem, if  $0 < p_1 \leq p_2 < 1$ , we have

$$\phi_{\Lambda,p_1}^{\text{disord}} \preceq \phi_{\Lambda,p_2}^{\text{disord}} \quad \text{and} \quad \phi_{\Lambda,p_1}^{\text{ord}} \preceq \phi_{\Lambda,p_2}^{\text{ord}} . \quad (166)$$

Let  $i \overset{\circ}{\leftrightarrow} \infty$  denote the event that there exists an infinite path of bonds passing through site  $i$  (“order” *percolates* from site  $i$  to infinity). The latter stochastic inequalities imply that the probabilities  $\phi_p^{\text{ord}}(i \overset{\circ}{\leftrightarrow} \infty)$  and  $\phi_p^{\text{disord}}(i \overset{\circ}{\leftrightarrow} \infty)$  are increasing in  $p$ . This monotonicity assures the existence of *critical probabilities*  $0 \leq p_c^{\text{ord}}, p_c^{\text{disord}} \leq 1$  such that for every  $p < p_c^{\text{ord}}$  we have  $\phi_p^{\text{ord}}(i \overset{\circ}{\leftrightarrow} \infty) = 0$  while for every  $p > p_c^{\text{ord}}$  we have  $\phi_p^{\text{ord}}(i \overset{\circ}{\leftrightarrow} \infty) > 0$ , and similarly for  $p_c^{\text{disord}}$ . The critical probabilities are given by

$$p_c^{\text{ord}} \triangleq \sup \left\{ p : \phi_p^{\text{ord}}(i \overset{\circ}{\leftrightarrow} \infty) = 0 \right\} , \quad (167)$$

$$p_c^{\text{disord}} \triangleq \sup \left\{ p : \phi_p^{\text{disord}}(i \overset{\circ}{\leftrightarrow} \infty) = 0 \right\} . \quad (168)$$

It turns out that the two critical probabilities are actually the same, hence we define  $p_c \triangleq p_c^{\text{ord}} = p_c^{\text{disord}}$ . This follows from the fact that the probability measures  $\phi_p^{\text{ord}}$  and  $\phi_p^{\text{disord}}$  may differ for at most countably many values of  $p$ . The latter can be proved in a very similar manner as done in Theorem 8.17 in [12] for the standard random-cluster measures.

By means of the coupling, many properties of the  $(q, r)$ -Potts measures can be derived from the corresponding  $r$ -biased random-cluster measures. For instance, one can show that the thermodynamic limits  $\mu^k$  and  $\mu^{\text{free}}$  do not depend on the sequence  $\{\Lambda_n\}_n$  of volumes along which the limits are taken. In fact, the limits

$$\mu^k = \lim_{\Lambda \uparrow \mathbb{L}} \mu_{\Lambda}^k \quad \text{and} \quad \mu^{\text{free}} = \lim_{\Lambda \uparrow \mathbb{L}} \mu_{\Lambda}^{\text{free}} \quad (169)$$

can be taken along the net of all finite volumes in  $\mathbb{L}$ . In particular, this implies the translation-invariance of  $\mu^k$  and  $\mu^{\text{free}}$ . The proofs are similar to those of the standard case (Proposition 6.9 in [9]).

Uniqueness and multiplicity of the  $(q, r)$ -Potts measures are related to the percolation of “order” in the  $r$ -biased random-cluster model. More specifically, if  $\phi^{\text{ord}}(\text{order percolate}) = 0$ , then the  $(q, r)$ -Potts model admits a unique Gibbs measure (as in Theorem 6.10 in [9]). On the other hand, if

$$\phi^{\text{ord}} \left( \begin{array}{l} \exists \text{ unique infinite sea of order} \\ \text{with finite islands of disorder} \end{array} \right) = 1 , \quad (170)$$

then the measures  $\mu^1, \mu^2, \dots, \mu^q$  are distinct and satisfy

$$\mu^k \left( \begin{array}{l} \exists \text{ unique infinite uni-colour sea,} \\ \text{which has colour } k \end{array} \right) = 1 . \quad (171)$$

(Recall that a “sea” of order in a random-cluster configuration is simply a connected component of bonds. A “uni-colour sea” in a Potts configuration refers to a maximal connected subgraph of the lattice induced by sites having the same colour.)

The latter claim is a consequence of the existence of a coupling between  $\phi^{\text{ord}}$  and  $\mu^k$  (for visible  $k$ ), which can be constructed as follows:

- i) We first sample a bond configuration  $X$  according to  $\phi^{\text{ord}}$ .
- ii) For every site  $i$  that  $i \overset{\circ}{\leftrightarrow} \infty$  in  $(\mathbb{S}, X)$ , we colour  $i$  with colour  $k$ .
- iii) For every finite non-singleton connected component of  $(\mathbb{S}, X)$ , we choose a random visible colour uniformly among the  $q$  possibilities and colour all the sites in the component with this colour.
- iv) For every isolated site  $i$  in  $(\mathbb{S}, X)$ , we choose a random colour uniformly among the  $q + r$  possible colours.

The fact that the marginal of this construction on spin configurations is  $\mu^k$  is parallel to Theorem 4.91 in [11] and has a similar proof. An analogous coupling exists between  $\phi^{\text{disord}}$  and  $\mu^{\text{free}}$ .

## References

- [1] Rodney J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press, 1982.
- [2] A. N. Berker, S. Ostlund, and F. A. Putnam. Renormalization-group treatment of a Potts lattice gas for krypton adsorbed onto graphite. *Physical Review B*, 17(9):3650–3665, 1978.
- [3] Christian Borgs and John Z. Imbrie. A unified approach to phase diagrams in field theory and statistical mechanics. *Communications in Mathematical Physics*, 123(2):305–328, 1989.
- [4] J. Bricmont, K. Kuroda, and J. L. Lebowitz. First order phase transitions in lattice and continuous systems: Extension of Pirogov-Sinai theory. *Communications in Mathematical Physics*, 101(4):501–538, 1985.
- [5] Teunis C. Dorlas. *Statistical Mechanics: Fundamentals and Model Solutions*. Institute of Physics Publishing, 1999.
- [6] Aernout C. D. van Enter, Giulio Iacobelli, and Siamak Taati. First-order transition in Potts models with “invisible” states: Rigorous proofs. *Progress of Theoretical Physics*, 126(5):983–991, 2011.
- [7] C. M. Fortuin and P. W. Kasteleyn. On the random-cluster model: I. Introduction and relation to other models. *Physica*, 57(4):536–564, 1972.
- [8] Hans-Otto Georgii. *Gibbs Measures and Phase Transitions*. Walter de Gruyter, 1988.
- [9] Hans-Otto Georgii, Olle Häggström, and Christian Maes. The random geometry of equilibrium phases. In C. Domb and J. Lebowitz, editors, *Phase Transitions and Critical Phenomena*, volume 18, pages 1–142. Academic Press, 2000.
- [10] T. Gobron and I. Merola. First-order phase transition in Potts models with finite-range interactions. *Journal of Statistical Physics*, 126(3):507–583, 2007.
- [11] Geoffrey Grimmett. *The Random-Cluster Model*. Springer, 2006.

- [12] Geoffrey Grimmett. *Probability on Graphs*. Cambridge University Press, 2010.
- [13] O. Häggström. Random-cluster representation in the study of phase transitions. *Markov Processes and Related Fields*, 4(3):275–321, 1998.
- [14] Robert B. Israel. *Convexity in the Theory of Lattice Gases*. Princeton University Press, 1979.
- [15] R. Kotecký and S. B. Shlosman. First-order phase transitions in large entropy lattice models. *Communications in Mathematical Physics*, 83(4):493–515, 1982.
- [16] Roman Kotecký. Geometric representation of lattice models and large volume asymptotics. In G. Grimmett, editor, *Probability and Phase Transition*, pages 153–176. Kluwer, 1994.
- [17] Lahoussine Laanait, Alain Messager, Salvador Miracle-Sole, Jean Ruiz, and Senya Shlosman. Interfaces in the Potts model I: Pirogov-Sinai theory of the Fortuin-Kasteleyn representation. *Communications in Mathematical Physics*, 140(1):81–91, 1991.
- [18] D. H. Martirosian. Translation invariant Gibbs states in the  $q$ -state Potts model. *Communications in Mathematical Physics*, 105(2):281–290, 1986.
- [19] David Ruelle. *Thermodynamic Formalism*. Cambridge University Press, 2nd edition, 2004.
- [20] Ya. G. Sinai. *Theory of Phase Transitions: Rigorous Results*. Akadémiai Kiadó, 1982.
- [21] Ryo Tamura, Shu Tanaka, and Naoki Kawashima. Phase transition in Potts model with invisible states. *Progress of Theoretical Physics*, 124(2):381–388, 2010.
- [22] Shu Tanaka and Ryo Tamura. Dynamical properties of Potts model with invisible states. *Journal of Physics: Conference Series*, 320(1):012025, 2011.
- [23] Shu Tanaka, Ryo Tamura, and Naoki Kawashima. Phase transition of generalized ferromagnetic Potts model: Effect of invisible states. *Journal of Physics: Conference Series*, 297(1):012022, 2011.
- [24] Miloš Zahradník. An alternate version of Pirogov-Sinai theory. *Communications in Mathematical Physics*, 93(4):559–581, 1984.