

Marek Biskup¹ · Roman Kotecký²

Phase coexistence of gradient Gibbs states

Abstract. We consider the (scalar) gradient fields $\eta = (\eta_b)$ —with b denoting the nearest-neighbor edges in \mathbb{Z}^2 —that are distributed according to the Gibbs measure proportional to $e^{-\beta H(\eta)} \nu(d\eta)$. Here $H = \sum_b V(\eta_b)$ is the Hamiltonian, V is a symmetric potential, $\beta > 0$ is the inverse temperature, and ν is the Lebesgue measure on the linear space defined by imposing the loop condition $\eta_{b_1} + \eta_{b_2} = \eta_{b_3} + \eta_{b_4}$ for each plaquette (b_1, b_2, b_3, b_4) in \mathbb{Z}^2 . For convex V , Funaki and Spohn have shown that ergodic infinite-volume Gibbs measures are characterized by their tilt. We describe a mechanism by which the gradient Gibbs measures with non-convex V undergo a structural, order-disorder phase transition at some intermediate value of inverse temperature β . At the transition point, there are at least two distinct gradient measures with zero tilt, i.e., $E\eta_b = 0$.

1. Introduction

1.1. Gradient fields

One of the mathematical challenges encountered in the study of systems exhibiting phase coexistence is an efficient description of microscopic phase boundaries. Here various levels of detail are in general possible: The finest level is typically associated with a statistical-mechanical model (e.g., a lattice gas) in which both the interface and the surrounding phases are represented microscopically; at the coarsest level the interface is viewed as a macroscopic (geometrical) surface between two structureless bulk phases. An intermediate approach is based on effective (and, often, solid-on-solid) models, in which the interface is still microscopic—represented by a stochastic field—while the structural details of the bulk phases are neglected.

A simple example of such an effective model is a *gradient field*. To define this system, we consider a finite subset Λ of the d -dimensional hypercubic lattice \mathbb{Z}^d and, at each site of Λ and its external boundary $\partial\Lambda$, we consider the real-valued variable ϕ_x representing the height of the interface at x . The Hamiltonian is then given by

$$H_\Lambda(\phi) = \sum_{\langle x, y \rangle : x \in \Lambda, y \in \Lambda \cup \partial\Lambda} V(\phi_y - \phi_x), \quad (1.1)$$

¹ Department of Mathematics, UCLA, Los Angeles, California, USA

² Center for Theoretical Study, Charles University, Prague, Czech Republic

© 2006 by M. Biskup and R. Kotecký. Reproduction, by any means, of the entire article for non-commercial purposes is permitted without charge.

where the sum is over unordered nearest-neighbor pairs $\langle x, y \rangle$. A standard example is the quadratic potential $V(\eta) = \frac{1}{2}\kappa\eta^2$ with $\kappa > 0$; in general V is assumed to be a smooth, even function with a sufficient (say, quadratic) growth at infinity. The Gibbs measure takes the usual form

$$P_\Lambda(d\phi) = Z^{-1} e^{-\beta H_\Lambda(\phi)} d\phi, \quad (1.2)$$

where $d\phi$ is the $|\Lambda|$ -dimensional Lebesgue measure (the boundary values of ϕ remain fixed and implicit in the notation), $\beta > 0$ is the inverse temperature and Z is a normalization constant.

A natural question to ask is what are the possible limits of the Gibbs measures $P_\Lambda(d\phi)$ as $\Lambda \uparrow \mathbb{Z}^d$. Unfortunately, in dimensions $d = 1, 2$, the fields $(\phi_x)_{x \in \Lambda}$ are very “rough” no matter how tempered the boundary conditions are assumed to be. As a consequence, the family of measures $(P_\Lambda)_{\Lambda \subset \mathbb{Z}^d}$ is not tight and no meaningful object is obtained by taking the limit $\Lambda \uparrow \mathbb{Z}^d$ —i.e., the interface is *delocalized*. On the other hand, in dimensions $d \geq 3$ the fields are sufficiently smooth to permit a non-trivial thermodynamic limit—the interface is *localized*. These facts are established by combinations of Brascamp-Lieb inequality techniques and/or random walk representation (see, e.g., [16]) which, unfortunately, apply only for convex potentials with uniformly positive curvature. Thus, somewhat surprisingly, even for $V(\eta) = \eta^4$ the problem of localization in high-dimension is still open [24, Open Problem 1].

As it turns out, the thermodynamic limit of the measures P_Λ is significantly less singular once we restrict attention to the gradient variables $\eta = (\eta_b)$. These are defined by $\eta_b = \phi_y - \phi_x$ where b is the nearest-neighbor edge (x, y) oriented in one of the positive lattice directions. Indeed, the η -marginal of $P_\Lambda(d\phi)$ always has at least one (weak) limit “point” as $\Lambda \rightarrow \mathbb{Z}^d$. The limit measures satisfy a natural DLR condition and are therefore called *gradient Gibbs measures*. (Precise definitions will be stated below or can be found in [16, 23].) One non-standard aspect of the gradient variables is that they have to obey a host of constraints. Namely,

$$\eta_{b_1} + \eta_{b_2} = \eta_{b_3} + \eta_{b_4} \quad (1.3)$$

holds for each lattice plaquette (b_1, b_2, b_3, b_4) , where the edges b_j are listed counterclockwise and are assumed to be positively oriented. These constraints will be implemented at the level of *a priori* measure, see Sect. 2.

It would be natural to expect that the character (and number) of gradient Gibbs measures depends sensitively on the potential V . However, this is not the case for the class of uniformly strictly-convex potentials (i.e., the V 's such that $V''(\eta) \geq c_- > 0$ for all η). Indeed, Funaki and Spohn [17] showed that, in these cases, the translation-invariant, ergodic, gradient Gibbs measures are completely characterized by the *tilt* of the underlying interface. Here the tilt is a vector $u \in \mathbb{R}^d$ such that

$$E\eta_b = u \cdot b \quad (1.4)$$

for every edge b —which we regard as a vector in \mathbb{R}^d . Furthermore, the correspondence is one-to-one, i.e., for each tilt there exists precisely one gradient Gibbs

measure with this tilt. Alternative proofs permitting extensions to discrete gradient models have appeared in Sheffield's thesis [23].

It is natural to expect that a serious violation of the strict-convexity assumption on V may invalidate the above results. Actually, an example of a gradient model with multiple gradient Gibbs states of the same tilt has recently been presented [23]; unfortunately, the example is not of the type considered above because of the lack of translation invariance and its reliance on the discreteness of the fields. The goal of this paper is to point out a general mechanism by which the model (1.1) with a sufficiently non-convex potential V fails the conclusions of Funaki-Spohn's theorems.

1.2. Potentials of interest

The mechanism driving our example will be the occurrence of a structural surface phase transition. To motivate the forthcoming considerations, let us recall that phase transitions typically arise via one of two mechanisms: either due to the breakdown of an internal symmetry, or via an abrupt turnover between energetically and entropically favored states. The standard examples of systems with these kinds of phase transitions are the Ising model and the q -state Potts model with a sufficiently large q , respectively. In the former, at sufficiently low temperatures, there is a spontaneous breaking of the symmetry between the plus and minus spin states; in the latter, there is a first-order transition at intermediate temperatures between q ordered, low-temperature states and a disordered, high-temperature state.

Our goal is to come up with a potential V that would mimic one of the above situations. In the present context an analogue of the Ising model appears to be a *double-well potential* of the form, e.g.,

$$V(\eta) = \kappa(\eta^2 - \eta_*^2)^2. \quad (1.5)$$

Unfortunately, due to the underlying plaquette constraints (1.3), the symmetry between the wells cannot be completely broken and, even at the level of ground states, the system appears to be disordered. On \mathbb{Z}^2 this can be demonstrated explicitly by making a link to the *ice model*, which is a special case of the six vertex model [1]. A similar equivalence has been used [2] to study a roughening transition in an SOS interface.

To see how the equivalence works exactly, note that the ground states of the system (1.5) are such that all η 's equal $\pm\eta_*$. Let us associate a unit flow with each *dual bond* whose sign is determined by the value of η_b for its direct counterpart b . The plaquette constraint (1.3) then translates into a *no-source-no-sink* condition for this flow. If we mark the flow by arrows, the dual bonds at each plaquette are constrained to one of six zero-flux arrangements of the six vertex model; cf Fig. 1 and its caption. The weights of all zero-flux arrangements are equal; we thus have the special case corresponding to the ice model. The ice model can be "exactly solved" [1]: The ground states have a non-vanishing residual entropy [22] and are disordered with infinite correlation length [1, Sect. 8.10.III]. However, it is not clear how much of this picture survives to positive temperatures.

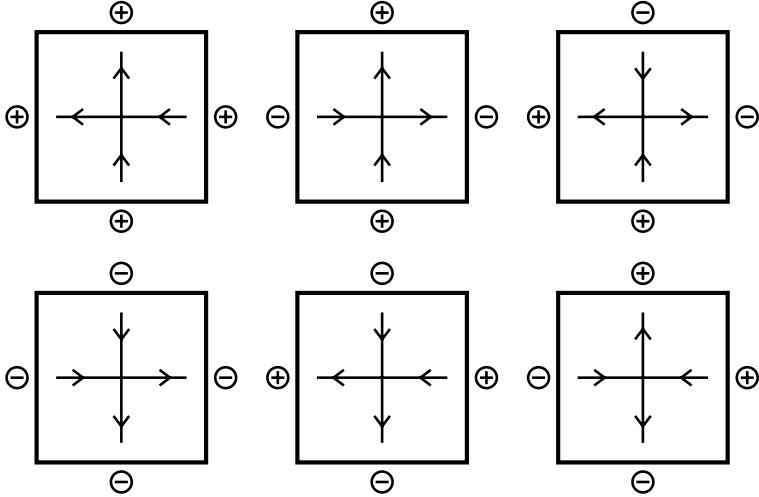


Fig. 1. The six plaquette configurations of minimal energy for the potential (1.5) on \mathbb{Z}^2 and their equivalent ice model configurations at the corresponding vertex on the dual grid. The sign marks represent the signs of η_b along the side of the plaquette (b_1, b_2, b_3, b_4)—with horizontal bonds b_1, b_3 oriented to the right and vertical bonds b_2, b_4 oriented upwards. The unit flow represented by the arrows runs upwards (downwards) through horizontal bonds with positive (negative) sign, and to the left (right) through vertical bonds with positive (negative) sign. The loop condition (1.3) makes the flow conserved (i.e., no sources or sinks).

The previous discussion shows that it will be probably quite hard to realize a symmetry-breaking transition in the context of the gradient model (1.1). It is the order-disorder mechanism for phase transitions that seems considerably more promising. There are two canonical examples of interest: a potential with *two centered wells* and a *triple-well potential*; see Fig. 2. Both of these lead to a gradient model which features a phase transition, at some intermediate temperature, from states with the η 's lying (mostly) within the thinner well to states whose η 's fluctuate on the scale of the thicker well(s).

Our techniques apply equally to these—as well as other similar—cases provided the widths of the wells are sufficiently distinct. Notwithstanding, the analysis becomes significantly cleaner if we abandon temperature as our principal parameter (e.g., we set $\beta = 1$) and consider potentials V that are simply *defined* by

$$e^{-V(\eta)} = p e^{-\kappa_O \eta^2/2} + (1-p) e^{-\kappa_D \eta^2/2}. \quad (1.6)$$

Here κ_O and κ_D are positive numbers and p is a parameter taking values in $[0, 1]$. For appropriate values of the constants, V defined this way will have a graph as in Fig. 2(a). To get the graph in part (b), we would need to consider V 's of the form

$$e^{-V(\eta)} = p e^{-\kappa_O \eta^2/2} + \frac{1-p}{2} e^{-\kappa_D(\eta-\eta_*)^2/2} + \frac{1-p}{2} e^{-\kappa_D(\eta+\eta_*)^2/2}, \quad (1.7)$$

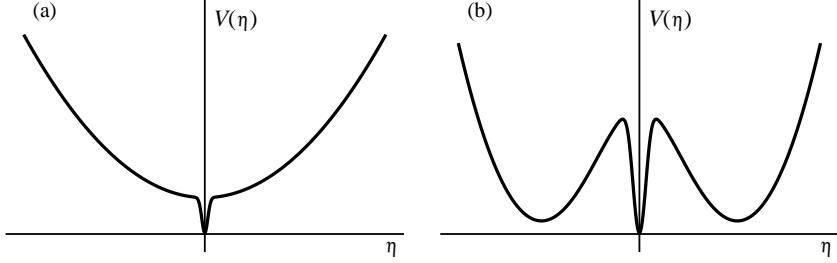


Fig. 2. Two canonical examples of potentials that will lead to a structural surface phase transition. The picture labeled (a) is obtained by superimposing—in the sense of (1.6)—two symmetric wells of (significantly) different widths. Part (b) of the figure represents the triple-well potential as defined in (1.7). For the application of our technique of proof, it only matters that the widths of the wells are sufficiently different.

where $\pm\eta_*$ are the (approximate) locations of the off-center wells.

The idea underlying the expressions (1.6) and (1.7) is similar to that of the Fortuin-Kasteleyn representation of the Potts model [12]. In the context of continuous-spin models similar to ours, such representation has fruitfully been used by Zahradník [25]. Focusing on (1.6), we can interpret the terms on the right-hand side of (1.6) as two distinct states of each bond. (We will soon exploit this interpretation in detail.) The indexing of the coupling constants suggests the names: “O” for *ordered* and “D” for *disordered*.

It is clear that the extreme values of p (near zero or near one) will be dominated by one type of bonds; what we intend to show is that, for κ_O and κ_D sufficiently distinct from each other, the transition between the “ordered” and “disordered” phases is (strongly) first order. Similar conclusions and proofs—albeit more complicated—apply also to the potential (1.7). However, for clarity of exposition, we will focus on the potential (1.6) for the rest of the paper (see, however, Sect. 2.5). In addition, we will also restrict ourselves to two dimensions, even though the majority of our results are valid for all $d \geq 2$.

2. Main results

2.1. Gradient Gibbs measures

We commence with a precise definition of our model. Most of the work in this paper will be confined to the lattice torus \mathbb{T}_L of $L \times L$ sites in \mathbb{Z}^2 , so we will start with this particular geometry. Choosing the natural positive direction for each lattice axis, let \mathbb{B}_L denote the corresponding set of positively oriented edges in \mathbb{T}_L . Given a configuration $(\phi_x)_{x \in \mathbb{T}_L}$, we introduce the gradient field $\eta = \nabla\phi$ by assigning the variable $\eta_b = \phi_y - \phi_x$ to each $b = (x, y) \in \mathbb{B}_L$. The product Lebesgue measure $\prod_{x \neq 0} d\phi_x$ induces a (σ -finite) measure ν_L on the space $\mathbb{R}^{\mathbb{B}_L}$ via

$$\nu_L(\mathcal{A}) = \int \left(\prod_{x \in \mathbb{T}_L \setminus \{0\}} d\phi_x \right) \delta(d\phi_0) \mathbf{1}_{\{\nabla\phi \in \mathcal{A}\}}, \quad (2.1)$$

where δ denotes the Dirac point-mass at zero.

We interpret the measure v_L as an *a priori* measure on *gradient* configurations $\eta \in \mathbb{R}^{\mathbb{B}_L}$. Since the η 's arise as the gradients of the ϕ 's it is easy to check that v_L is entirely supported on the linear subspace $\mathcal{X}_L \subset \mathbb{R}^{\mathbb{B}_L}$ of configurations determined by the condition that the sum of signed η 's—with a positive or negative sign depending on whether the edge is traversed in the positive or negative direction, respectively—vanishes around each closed circuit on \mathbb{T}_L . (Note that, in addition to (1.3), the condition includes also loops that wrap around the torus.) We will refer to such configurations as *curl-free*.

Next we will define gradient Gibbs measures on \mathbb{T}_L . For later convenience we will proceed in some more generality than presently needed: Let $(V_b)_{b \in \mathbb{B}_L}$ be a collection of measurable functions $V_b : \mathbb{R} \rightarrow [0, \infty)$ and consider the partition function

$$Z_{L,(V_b)} = \int_{\mathbb{R}^{|\mathbb{B}_L|}} \exp\left\{-\sum_{b \in \mathbb{B}_L} V_b(\eta_b)\right\} v_L(d\eta). \quad (2.2)$$

Clearly, $Z_{L,(V_b)} > 0$ and, under the condition that $\eta \mapsto e^{-V_b(\eta)}$ is integrable with respect to the Lebesgue measure on \mathbb{R} , also $Z_{L,(V_b)} < \infty$. We may then define $P_{L,(V_b)}$ to be the probability measure on $\mathbb{R}^{|\mathbb{B}_L|}$ given by

$$P_{L,(V_b)}(d\eta) = \frac{1}{Z_{L,(V_b)}} \exp\left\{-\sum_{b \in \mathbb{B}_L} V_b(\eta_b)\right\} v_L(d\eta). \quad (2.3)$$

This is the *gradient Gibbs measure* on \mathbb{T}_L corresponding to the potentials (V_b) . In the situations when $V_b = V$ for all b —which is the principal case of interest in this paper—we will denote the corresponding gradient Gibbs measure on \mathbb{T}_L by $P_{L,V}$.

It is not surprising that $P_{L,(V_b)}$ obeys appropriate DLR equations with respect to all connected $\Lambda \subset \mathbb{T}_L$ containing no topologically non-trivial circuit. Explicitly, if η_{Λ^c} in Λ^c is a curl-free boundary condition, then the conditional law of η_Λ given η_{Λ^c} is

$$P_{L,(V_b)}(d\eta_\Lambda | \eta_{\Lambda^c}) = \frac{1}{Z_\Lambda(\eta_{\Lambda^c})} \exp\left\{-\sum_{b \in \Lambda} V_b(\eta_b)\right\} v_\Lambda(d\eta_\Lambda | \eta_{\Lambda^c}). \quad (2.4)$$

Here $P_{L,(V_b)}(d\eta_\Lambda | \eta_{\Lambda^c})$ is the conditional probability with respect to the (tail) σ -algebra \mathcal{T}_Λ generated by the fields on Λ^c , $Z_\Lambda(\eta_{\Lambda^c})$ is the partition function in Λ , and $v_\Lambda(d\eta_\Lambda | \eta_{\Lambda^c})$ is the *a priori* measure induced by v_L on η_Λ given the boundary condition η_{Λ^c} .

As usual, this property remains valid even in the thermodynamic limit. We thus say that a measure on μ is an *infinite-volume gradient Gibbs measure* if it satisfies the DLR equations with respect to the specification (2.4) in any finite set $\Lambda \subset \mathbb{Z}^2$. (As is easy to check—e.g., by reinterpreting the η 's back in terms of the ϕ 's— $v_\Lambda(d\eta_\Lambda | \eta_{\Lambda^c})$ is independent of the values of η_{Λ^c} outside any circuit winding around Λ , and so it is immaterial that it originated from a measure on torus.)

An important aspect of our derivations will be the fact that our potential V takes the specific form (1.6), which can be concisely written as

$$e^{-V(\eta)} = \int \varrho(d\kappa) e^{-\frac{1}{2}\kappa\eta^2}, \quad (2.5)$$

where ϱ is the probability measure $\varrho = p\delta_{\kappa_0} + (1-p)\delta_{\kappa_D}$. It follows that the Gibbs measure $P_{L,V}$ can be regarded as the projection of the *extended gradient Gibbs measure*,

$$Q_L(d\eta, d\kappa) = \frac{1}{Z_{L,V}} \exp\left\{-\frac{1}{2} \sum_{b \in \mathbb{B}_L} \kappa_b \eta_b^2\right\} v_L(d\eta) \varrho_L(d\kappa), \quad (2.6)$$

to the σ -algebra generated by the η 's. Here ϱ_L is the product of measures ϱ , one for each bond in \mathbb{B}_L . As is easy to check, conditioning on $(\eta_b, \kappa_b)_{b \in \Lambda^c}$ yields the corresponding extension $Q_\Lambda(d\eta_\Lambda d\kappa_\Lambda | \eta_{\Lambda^c})$ of the finite-volume specification (2.4)—the result is independent of the κ 's outside Λ because, once η_{Λ^c} is fixed, these have no effect on the configurations in Λ .

The main point of introducing the extended measure is that, if conditioned on the κ 's, the variables η_b are distributed as gradients of a Gaussian field—albeit with a non-translation invariant covariance matrix. As we will see, the phase transition proved in this paper is manifested by a jump-discontinuity in the density of bonds with $\kappa_b = \kappa_0$ which at the level of η -marginal results in a jump in the characteristic scale of the fluctuations.

Remark 2.1. Notably, the extended measure Q_L plays the same role for $P_{L,V}$ as the so called Edwards-Sokal coupling measure [10] does for the Potts model. Similarly as for the Edwards-Sokal measures [3, 18], there is a one-to-one correspondence between the infinite-volume measures on η 's and the corresponding infinite-volume extended gradient Gibbs measures on (η, κ) 's. Explicitly, if μ is an infinite-volume gradient Gibbs measure for potential V , then $\tilde{\mu}$, defined by (extending the consistent family of measures of the form)

$$\tilde{\mu}\left((\eta_b, \kappa_b)_{b \in \Lambda} \in \mathcal{A} \times \mathcal{B}\right) = \int_{\mathcal{B}} \varrho_\Lambda(d\kappa) E_\mu\left(1_{\mathcal{A}} \prod_{b \in \Lambda} e^{-\frac{1}{2}\kappa_b \eta_b^2 + V(\eta_b)}\right), \quad (2.7)$$

is a Gibbs measure with respect to the extended specifications $Q_\Lambda(\cdot | \eta_{\Lambda^c})$. For the situations with only a few distinct values of κ_b , it may be of independent interest to study the properties of the κ -marginal of the extended measure, e.g., using the techniques of percolation theory. However, apart from some remarks in Sect. 2.3, we will not pursue these matters in the present paper.

2.2. Phase coexistence of gradient measures

Now we are ready to state our main results. Throughout we will consider the potentials V of the form (1.6) with $\kappa_0 \gg \kappa_D$. As a moment's thought reveals, the model is invariant under the transformation

$$\kappa_0 \rightarrow \kappa_0 \theta^2, \quad \kappa_D \rightarrow \kappa_D \theta^2, \quad \eta_b \rightarrow \eta_b / \theta \quad (2.8)$$

for any fixed $\theta \neq 0$. In particular, without loss of generality, one could assume from the beginning that $\kappa_O \kappa_D = 1$ and regard κ_O/κ_D as the sole parameter of the model. However, we prefer to treat the two terms in (1.6) on an equal footing, and so we will keep the coupling strengths independent.

Given a shift-ergodic gradient Gibbs measure, recall that its tilt is the vector u such that (1.4) holds for each bond. The principal result of the present paper is the following theorem:

Theorem 2.2. *For each $\epsilon > 0$ there exists a constant $c = c(\epsilon) > 0$ and, if*

$$\kappa_O \geq c\kappa_D, \quad (2.9)$$

a number $p_t \in (0, 1)$ such that, for interaction V with $p = p_t$, there are two distinct, infinite-volume, shift-ergodic gradient Gibbs measures μ_{ord} and μ_{dis} of zero tilt for which

$$\mu_{\text{ord}}\left(|\eta_b| \geq \frac{\lambda}{\sqrt{\kappa_O}}\right) \leq \epsilon + \frac{1}{4\lambda^2}, \quad \forall \lambda > 0, \quad (2.10)$$

and

$$\mu_{\text{dis}}\left(|\eta_b| \leq \frac{\lambda}{\sqrt{\kappa_D}}\right) \leq \epsilon + c_1 \lambda^{1/4}, \quad \forall \lambda > 0. \quad (2.11)$$

Here c_1 is a constant of order unity.

Remark 2.3. An inspection of the proof actually reveals that the above bounds are valid for any ϵ satisfying $\epsilon \geq c_2 (\kappa_D/\kappa_O)^{1/8}$, where c_2 is a constant of order unity.

As already alluded to, this result is a consequence of the fact that the density of ordered bonds, i.e., those with $\kappa_b = \kappa_O$, undergoes a jump at $p = p_t$. On the torus, we can make the following asymptotic statements:

Theorem 2.4. *Let R_L^{ord} denote the fraction of ordered bonds on \mathbb{T}_L , i.e.,*

$$R_L^{\text{ord}} = \frac{1}{|\mathbb{B}_L|} \sum_{b \in \mathbb{B}_L} \mathbf{1}_{\{\kappa_b = \kappa_O\}}. \quad (2.12)$$

For each $\epsilon > 0$ there exists $c = c(\epsilon) > 0$ such that the following holds: Under the condition (2.9), and for p_t as in Theorem 2.2,

$$\lim_{L \rightarrow \infty} Q_L(R_L^{\text{ord}} \leq \epsilon) = 1, \quad p < p_t \quad (2.13)$$

and

$$\lim_{L \rightarrow \infty} Q_L(R_L^{\text{ord}} \geq 1 - \epsilon) = 1, \quad p > p_t. \quad (2.14)$$

The present setting actually permits us to determine the value of p_t via a duality argument. This is the only result in this paper which is intrinsically two-dimensional (and intrinsically tied to the form (1.6) of V). All other conclusions can be extended to $d \geq 2$ and to more general potentials.

Theorem 2.5. *Let $d = 2$. If $\kappa_0/\kappa_D \gg 1$, then p_t is given by*

$$\frac{p_t}{1 - p_t} = \left(\frac{\kappa_D}{\kappa_0} \right)^{1/4}. \quad (2.15)$$

Theorem 2.4 is proved in Sect. 4.2, Theorem 2.2 is proved in Sect. 4.3 and Theorem 2.5 is proved in Sect. 5.3.

2.3. Discussion

The phase transition described in the above theorems can be interpreted in several ways. First, in terms of the extended gradient Gibbs measures on torus, it clearly corresponds to a transition between a state with nearly all bonds ordered ($\kappa_b = \kappa_0$) to a state with nearly all bonds disordered ($\kappa_b = \kappa_D$). Second, looking back at the inequalities (2.10–2.11), most of the η 's will be of order at most $1/\sqrt{\kappa_0}$ in the ordered state while most of them will be of order at least $1/\sqrt{\kappa_D}$ in the disordered state. Hence, the corresponding (effective) interface is significantly rougher at $p < p_t$ than it is at $p > p_t$ (both phases are rough according to the standard definition of this term) and we may thus interpret the above as a kind of *first-order roughening* transition that the interface undergoes at p_t . Finally, since the gradient fields in the two states fluctuate on different characteristic scales, the entropy (and hence the energy) associated with these states is different; we can thus view this as a standard energy-entropy transition. (By the energy we mean the expectation of $V(\eta_b)$; notably, the expectation of $\kappa_b \eta_b^2$ is the same in both measures; cf (4.35).) Energy-entropy transitions for spin models have been studied in [9, 20, 21] and, quite recently, in [11].

Next let us turn our attention to the conclusions of Theorem 2.4. We actually believe that the dichotomy (2.13–2.14) applies (in the sense of almost-sure limit of R_L^{ord} as $L \rightarrow \infty$) to all translation-invariant extended gradient Gibbs states with zero tilt. The reason is that, conditional on the κ 's, the gradient fields are Gaussian with uniformly positive stiffness. We rest assured that the techniques of [17] and [23] can be used to prove that the gradient Gibbs measure with zero tilt is unique for almost every configuration of the κ 's; so the only reason for multiplicity of gradient Gibbs measures with zero tilt is a phase transition in the κ -marginal. However, a detailed write-up of this argument would require developing the precise—and somewhat subtle—correspondence between the gradient Gibbs measures of a given tilt and the minimizers of the Gibbs variational principle (which we have, in full detail, only for convex periodic potentials [23]). Thus, to keep the paper at manageable length, we limit ourselves to a weaker result.

The fact that the transition occurs at p_t satisfying (2.15) is a consequence of a *duality* between the κ -marginals at p and $1 - p$. More generally, the duality links the marginal law of the configuration (κ_b) with the law of $(1/\kappa_b)$; see Theorem 5.3 and Remark 5.4. [At the level of gradient fields, the duality provides only a vague link between the flow of the weighted gradients ($\sqrt{\kappa_b} \eta_b$) along a given curve and its flux through this curve. Unfortunately, this link does not seem to be particularly useful.] The point $p = p_t$ is self-dual which makes it the most natural candidate for a transition point. It is interesting to ponder about what happens when κ_0/κ_D

decreases to one. Presumably, the first-order transition (for states at zero tilt) disappears before κ_0/κ_D reaches one and is replaced by some sort of critical behavior. Here the first problem to tackle is to establish the *absence* of first-order phase transition for small $\kappa_0/\kappa_D - 1$. Via a standard duality argument (see [8]) this would yield a power-law lower bound for bond connectivities at p_t .

Another interesting problem is to determine what happens with measures of non-zero tilt. We expect that, at least for moderate values of the tilt u , the first-order transition persists but shifts to lower values of p . Thus, one could envision a whole phase diagram in the p - u plane. Unfortunately, we are unable to make any statements of this kind because the standard ways to induce a tilt on the torus (cf [17]) lead to measures that are not reflection positive.

2.4. Outline of the proof

We proceed by an outline of the principal steps of the proof to which the remainder of this paper is devoted. The arguments are close in spirit to those in [9, 20, 21]; the differences arise from the subtleties in the setup due to the gradient nature of the fields.

The main line of reasoning is basically thermodynamical: Consider the κ -marginal of the extended torus state Q_L which we will regard as a measure on configurations of ordered and disordered bonds. Let $\chi(p)$ denote (the $L \rightarrow \infty$ limit of) the expected fraction of ordered bonds in the torus state at parameter p . Clearly $\chi(p)$ increases from zero to one as p sweeps through $[0, 1]$. The principal observation is that, under the assumption $\kappa_0/\kappa_D \gg 1$, the quantity $\chi(1 - \chi)$ is small, uniformly in p . Hence, $p \mapsto \chi(p)$ must undergo a jump from values near zero to values near one at some $p_t \in (0, 1)$. By usual weak-limiting arguments we construct two distinct gradient Gibbs measures at p_t , one with high density of ordered bonds and the other with high density of disordered bonds.

The crux of the matter is thus to justify the uniform smallness of $\chi(1 - \chi)$. This will be a consequence of the fact that the simultaneous occurrence of ordered and disordered bonds at any two given locations is (uniformly) unlikely. For instance, let us estimate the probability that a particular plaquette has two ordered bonds emanating out of one corner and two disordered bonds emanating out of the other. Here the technique of chessboard estimates [13–15] allows us to disseminate this pattern all over the torus via successive reflections (cf Theorem 4.2 in Sect. 4.1). This bounds the quantity of interest by the $1/L^2$ -power of the probability that every other horizontal (and vertical) line is entirely ordered and the remaining lines are disordered. The resulting “spin-wave calculation”—i.e., diagonalization of a period-2 covariance matrix in the Fourier basis and taking its determinant—is performed (for all needed patterns) in Sect. 3.

Once the occurrence of a “bad pattern” is estimated by means of various spin-wave free energies, we need to prove that these “bad-pattern” spin-wave free energies are always worse off than those of the homogeneous patterns (i.e., all ordered or all disordered)—this is the content of Theorem 3.3. Then we run a standard Peierls’ contour estimate whereby the smallness of $\chi(1 - \chi)$ follows. Extracting two distinct, infinite-volume, ergodic gradient Gibbs states μ_{ord} and μ_{dis} at $p = p_t$,

it remains to show that these are both of zero tilt. Here we use the fact that, conditional on the κ 's, the torus measure is symmetric Gaussian with uniformly positive stiffness. Hence, we can use standard Gaussian inequalities to show exponential tightness of the tilt, uniformly in the κ 's; cf Lemma 4.8. Duality calculations (see Sect. 5) then yield $p = p_t$.

2.5. Generalizations

Our proof of phase coexistence applies to any potential of the form shown in Fig. 2—even if we return to parametrization by β . The difference with respect to the present setup is that in the general case we would have to approximate the potentials by a quadratic well at each local minimum and, before performing the requisite Gaussian calculations, estimate the resulting errors.

Here is a sketch of the main ideas: We fix a scale Δ and regard η_b to be in a well if it is within Δ of the corresponding local minimum. Then the requisite quadratic approximation of β -times energy is good up to errors of order $\beta\Delta^3$. The rest of the potential “landscape” lies at energies of at least order Δ^2 and so it will be only “rarely visited” by the η 's provided that $\beta\Delta^2 \gg 1$. On the other hand, the same condition ensures that the spin-wave integrals are essentially not influenced by the restriction that η_b be within Δ of the local minimum. Thus, to make all approximations work we need that

$$\beta\Delta^3 \ll 1 \ll \beta\Delta^2 \quad (2.16)$$

which is achieved for $\beta \gg 1$ by, e.g., $\Delta = \beta^{-\frac{5}{12}}$. This approach has recently been used to prove phase transitions in classical [4, 5] as well as quantum [6] systems with highly degenerate ground states. We refer the reader to these references for further details.

A somewhat more delicate issue is the proof that both coexisting states are of zero tilt. Here the existing techniques require that we have some sort of uniform convexity. This more or less forces us to use the V 's of the form

$$V(\eta) = -\log\left(\sum_j e^{-V_j(\eta)}\right), \quad (2.17)$$

where the V_j 's are uniformly convex functions. Clearly, our choice (1.6) is the simplest potential of this type; the question is how general the potentials obtained this way can be. We hope to return to this question in a future publication.

3. Spin-wave calculations

As was just mentioned, the core of our proofs are estimates of the spin-wave free energy for various regular patterns of ordered and disordered bonds on the torus. These estimates are rather technical and so we prefer to clear them out of the way before we get to the main line of the proof. The readers wishing to follow the proof in linear order may consider skipping this section and returning to it only while reading the arguments in Sect. 4.2. Throughout this and the forthcoming sections we assume that L is an even integer.

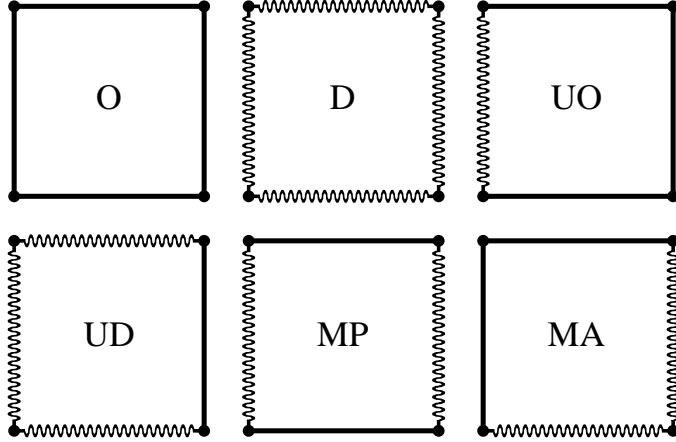


Fig. 3. Six possible arrangements of “ordered” and “disordered” bonds around a lattice plaquette. Here the “ordered” bonds are represented by solid lines and the “disordered” bonds by wavy lines. Each inhomogeneous pattern admits other rotations which are not depicted. The acronyms stand for: (top row) Ordered, Disordered and U-shape Ordered and (bottom row) U-shape Disordered, Mixed Periodic and Mixed Aperiodic, respectively.

3.1. Constrained partition functions

We will consider six partition functions $Z_{L,O}$, $Z_{L,D}$, $Z_{L,UO}$, $Z_{L,UD}$, $Z_{L,MP}$ and $Z_{L,MA}$ on \mathbb{T}_L that correspond to six regular configurations each of which is obtained by reflecting one of six possible arrangements of “ordered” and “disordered” bonds around a lattice plaquette to the entire torus. These quantities will be the “building blocks” of our analysis in Sect. 4. The six plaquette configurations are depicted in Fig. 3.

We begin by considering the homogeneous configurations. Here $Z_{L,O}$ is the partition function $Z_{L,(V_b)}$ for all edges of the “ordered” type:

$$V_b(\eta) = -\log p + \frac{1}{2}\kappa_O\eta^2, \quad b \in \mathbb{B}_L. \quad (3.1)$$

Similarly, $Z_{L,D}$ is the quantity $Z_{L,(V_b)}$ for

$$V_b(\eta) = -\log(1-p) + \frac{1}{2}\kappa_D\eta^2, \quad b \in \mathbb{B}_L, \quad (3.2)$$

i.e., with all edges “disordered.”

Next we will define the partition functions $Z_{L,UO}$ and $Z_{L,UD}$ which are obtained by reflecting a plaquette with three bonds of one type and the remaining bond of the other type. Let us split \mathbb{B}_L into the even $\mathbb{B}_L^{\text{even}}$ and odd $\mathbb{B}_L^{\text{odd}}$ horizontal and vertical edges—with the even edges on the lines of sites in the x direction with even y coordinates and lines of sites in y direction with even x coordinates.

Similarly, we will also consider the decomposition of \mathbb{B}_L into the set of horizontal edges $\mathbb{B}_L^{\text{hor}}$ and vertical edges $\mathbb{B}_L^{\text{vert}}$. Letting

$$V_b(\eta) = \begin{cases} -\log p + \frac{1}{2}\kappa_O\eta^2, & \text{if } b \in \mathbb{B}_L^{\text{hor}} \cup \mathbb{B}_L^{\text{even}}, \\ -\log(1-p) + \frac{1}{2}\kappa_D\eta^2, & \text{otherwise,} \end{cases} \quad (3.3)$$

the partition function $Z_{L,\text{UO}}$ then corresponds to the quantity $Z_{L,(V_b)}$. The partition function $Z_{L,\text{UD}}$ is obtained similarly; with the roles of “ordered” and “disordered” interchanged. Note that, since we are working on a square torus, the orientation of the pattern we choose does not matter.

It remains to define the partition functions $Z_{L,\text{MP}}$ and $Z_{L,\text{MA}}$ corresponding to the patterns with two “ordered” and two “disordered” bonds. For the former, we simply take $Z_{L,(V_b)}$ with the potential

$$V_b(\eta) = \begin{cases} -\log p + \frac{1}{2}\kappa_O\eta^2, & \text{if } b \in \mathbb{B}_L^{\text{hor}}, \\ -\log(1-p) + \frac{1}{2}\kappa_D\eta^2, & \text{if } b \in \mathbb{B}_L^{\text{vert}}. \end{cases} \quad (3.4)$$

Note that the two types of bonds are arranged in a “mixed periodic” pattern; hence the index MP. As to the quantity $Z_{L,\text{MA}}$, here we will consider a “mixed aperiodic” pattern. Explicitly, we define

$$V_b(\eta) = \begin{cases} -\log p + \frac{1}{2}\kappa_O\eta^2, & \text{if } b \in \mathbb{B}_L^{\text{even}}, \\ -\log(1-p) + \frac{1}{2}\kappa_D\eta^2, & \text{if } b \in \mathbb{B}_L^{\text{odd}}. \end{cases} \quad (3.5)$$

The “mixed aperiodic” partition function $Z_{L,\text{MA}}$ is the quantity $Z_{L,(V_b)}$ for this choice of (V_b) . Again, on a square torus it is immaterial for the values of $Z_{L,\text{MP}}$ and $Z_{L,\text{MA}}$ which orientation of the initial plaquette we start with.

As usual, associated with these partition functions are the corresponding free energies. In finite volume, these quantities can be defined in all cases by the formula

$$F_{L,\alpha}(p) = -\frac{1}{L^2} \log \frac{Z_{L,\alpha}}{(2\pi)^{\frac{1}{2}(L^2-1)}}, \quad \alpha = \text{O, D, UO, UD, MP, MA}, \quad (3.6)$$

where the factor $(2\pi)^{\frac{1}{2}(L^2-1)}$ has been added for later convenience and where the p -dependence arises via the corresponding formulas for V_b in each particular case.

3.2. Limiting free energies

The goal of this section is to compute the thermodynamic limit of the $F_{L,\alpha}$'s. For homogeneous and isotropic configurations, an important role will be played by the momentum representation of the lattice Laplacian $\widehat{D}(\mathbf{k}) = |1 - e^{i\mathbf{k}_1}|^2 + |1 - e^{i\mathbf{k}_2}|^2$ defined for all $\mathbf{k} = (k_1, k_2)$ in the corresponding Brillouin zone $\mathbf{k} \in [-\pi, \pi] \times [-\pi, \pi]$. Using this quantity, the “ordered” free energy will be simply

$$F_{\text{O}}(p) = -2 \log p + \frac{1}{2} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log\{\kappa_O \widehat{D}(\mathbf{k})\}, \quad (3.7)$$

while the disordered free energy boils down to

$$F_D(p) = -2 \log(1-p) + \frac{1}{2} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log\{\kappa_D \widehat{D}(\mathbf{k})\}. \quad (3.8)$$

It is easy to check that, despite the logarithmic singularity at $\mathbf{k} = \mathbf{0}$, both integrals converge. The bond pattern underlying the quantity $Z_{L,\text{MP}}$ lacks rotation invariance and so a different propagator appears inside the momentum integral:

$$\begin{aligned} F_{\text{MP}}(p) &= -\log[p(1-p)] \\ &\quad + \frac{1}{2} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log\{\kappa_O|1-e^{ik_1}|^2 + \kappa_D|1-e^{ik_2}|^2\}. \end{aligned} \quad (3.9)$$

Again, the integral converges as long as (at least) one of κ_O and κ_D is strictly positive.

The remaining partition functions come from configurations that lack translation invariance and are “only” periodic with period two. Consequently, the Fourier transform of the corresponding propagator is only block diagonal, with two or four different \mathbf{k} ’s “mixed” inside each block. In the UO cases we will get the function

$$F_{\text{UO}}(p) = -\frac{1}{2} \log[p^3(1-p)] + \frac{1}{4} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log\{\det \Pi_{\text{UO}}(\mathbf{k})\}, \quad (3.10)$$

where $\Pi_{\text{UO}}(\mathbf{k})$ is the 2×2 -matrix

$$\Pi_{\text{UO}}(\mathbf{k}) = \begin{pmatrix} \kappa_O|a_-|^2 + \frac{1}{2}(\kappa_O + \kappa_D)|b_-|^2 & \frac{1}{2}(\kappa_O - \kappa_D)|b_-|^2 \\ \frac{1}{2}(\kappa_O - \kappa_D)|b_-|^2 & \kappa_O|a_+|^2 + \frac{1}{2}(\kappa_O + \kappa_D)|b_-|^2 \end{pmatrix} \quad (3.11)$$

with a_{\pm} and b_{\pm} defined by

$$a_{\pm} = 1 \pm e^{ik_1} \quad \text{and} \quad b_{\pm} = 1 \pm e^{ik_2}. \quad (3.12)$$

The extra factor $1/2$ —on top of the usual $1/2$ —in front of the integral arises because $\det \Pi_{\text{UO}}(\mathbf{k})$ combines the contributions of two Fourier models; namely \mathbf{k} and $\mathbf{k} + \pi \hat{\mathbf{e}}_1$. A calculation shows

$$\det \Pi_{\text{UO}}(\mathbf{k}) \geq \kappa_O^2 |a_-|^2 |a_+|^2 + \kappa_O \kappa_D |b_-|^4, \quad (3.13)$$

implying that the integral in (3.10) converges. The free energy F_{UD} is obtained by interchanging the roles of κ_O and κ_D and of p and $(1-p)$.

In the MA-cases we will assume that $\kappa_O \neq \kappa_D$ —otherwise there is no distinction between any of the six patterns. The corresponding free energy is then given by

$$\begin{aligned} F_{\text{MA}}(p) &= -\log[p(1-p)] \\ &\quad + \frac{1}{8} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log\left\{\left(\frac{\kappa_O - \kappa_D}{2}\right)^4 \det \Pi_{\text{MA}}(\mathbf{k})\right\}. \end{aligned} \quad (3.14)$$

Here $\Pi_{\text{MA}}(\mathbf{k})$ is the 4×4 -matrix

$$\Pi_{\text{MA}}(\mathbf{k}) = \begin{pmatrix} r(|a_-|^2 + |b_-|^2) & |b_-|^2 & |a_-|^2 & 0 \\ |b_-|^2 & r(|a_+|^2 + |b_-|^2) & 0 & |a_+|^2 \\ |a_-|^2 & 0 & r(|a_-|^2 + |b_+|^2) & |b_+|^2 \\ 0 & |a_+|^2 & |b_+|^2 & r(|a_+|^2 + |b_+|^2) \end{pmatrix} \quad (3.15)$$

with the abbreviation

$$r = \frac{\kappa_O + \kappa_D}{\kappa_O - \kappa_D}. \quad (3.16)$$

Note that $r > 1$ in the cases of our interest.

Observe that $\det \Pi_{\text{MA}}(\mathbf{k})$ is a quadratic polynomial in r^2 , i.e., $\det \Pi_{\text{MA}} = Ar^4 + Br^2 + C$. Moreover, $\Pi_{\text{MA}}(\mathbf{k})$ annihilates $(1, -1, -1, 1)$ when $r = 1$, and so $r^2 = 1$ is a root of $Ar^4 + Br^2 + C$. Hence $\det \Pi_{\text{MA}}(\mathbf{k}) = (r^2 - 1)(Ar^2 - C)$, i.e.,

$$\begin{aligned} \det \Pi_{\text{MA}}(\mathbf{k}) = (r^2 - 1) & \left\{ -(|a_+|^2|a_-|^2 - |b_+|^2|b_-|^2)^2 \right. \\ & \left. + (|a_+|^2 + |b_+|^2)(|a_-|^2 + |b_+|^2)(|a_+|^2 + |b_-|^2)(|a_-|^2 + |b_-|^2)r^2 \right\}. \end{aligned} \quad (3.17)$$

Setting $r = 1$ inside the large braces yields

$$\det \Pi_{\text{MA}}(\mathbf{k}) \geq 4(r^2 - 1)|a_-|^2|a_+|^2|b_-|^2|b_+|^2, \quad (3.18)$$

implying that the integral in (3.14) is well defined and finite.

Remark 3.1. The fact that $\Pi_{\text{MA}}(\mathbf{k})$ has zero eigenvalue at $r = 1$ is not surprising. Indeed, $r = 1$ corresponds to $\kappa_D = 0$ in which case a quarter of all sites in the MA-pattern get decoupled from the rest. This indicates that the partition function blows up (at least) as $(r - 1)^{-|\mathbb{T}_L|/4}$ as $r \downarrow 1$ implying that there should be a zero eigenvalue at $r = 1$ per each 4×4 -block $\Pi_{\text{MA}}(\mathbf{k})$.

A formal connection between the quantities in (3.6) and those in (3.7–3.14) is guaranteed by the following result:

Theorem 3.2. *For all $\alpha = O, D, UO, UD, MP, MA$ and uniformly in $p \in (0, 1)$,*

$$\lim_{L \rightarrow \infty} F_{L,\alpha}(p) = F_\alpha(p). \quad (3.19)$$

Proof. This is a result of standard calculations of Gaussian integrals in momentum representation. We begin by noting that the Lebesgue measure $\prod_x d\phi_x$ can be regarded as the product of ν_L , acting only on the gradients of ϕ , and $d\phi_z$ for some fixed $z \in \mathbb{T}_L$. Neglecting temporarily the *a priori* bond weights p and $(1 - p)$, the partition function $Z_{L,\alpha}$, $\alpha = O, D, UO, UD, MP, MA$, is thus the integral of the

Gaussian weight $(2\pi|\mathbb{T}_L|)^{-1/2}e^{-\frac{1}{2}(\phi, C_\alpha^{-1}\phi)}$ against the measure $\prod_x d\phi_x$, where the covariance matrix C_α is defined by the quadratic form

$$(\phi, C_\alpha^{-1}\phi) = \sum_{b \in \mathbb{B}_L} \kappa_b^{(\alpha)} (\nabla_b \phi)^2 + \frac{1}{|\mathbb{T}_L|} \left(\sum_{x \in \mathbb{T}_L} \phi_x \right)^2. \quad (3.20)$$

Here $(\kappa_b^{(\alpha)})$ are the bond weights of pattern α . Indeed, the integral over $d\phi_z$ with the gradient variables fixed yields $(2\pi|\mathbb{T}_L|)^{1/2}$ which cancels the term in front of the Gaussian weight. The purpose of the above rewrite was to reinsert the “zero mode” $\widehat{\phi}_0 = |\mathbb{T}_L|^{-1/2} \sum_x \phi_x$ into the partition function; $\widehat{\phi}_0$ was not subject to integration due to the restriction to gradient variables.

To compute the Gaussian integral, we need to diagonalize C_α . For that we will pass to the Fourier components $\widehat{\phi}_{\mathbf{k}} = |\mathbb{T}_L|^{-1/2} \sum_{x \in \mathbb{T}_L} \phi_x e^{i\mathbf{k} \cdot x}$ with the result

$$(\phi, C_\alpha^{-1}\phi) = \sum_{\mathbf{k}, \mathbf{k}' \in \widetilde{\mathbb{T}}_L} \widehat{\phi}_{\mathbf{k}} \widehat{\phi}_{\mathbf{k}'}^* \left(\delta_{\mathbf{k}, 0} \delta_{\mathbf{k}', 0} + \sum_{\sigma=1,2} A_{\mathbf{k}, \mathbf{k}'}^{(\sigma)} (1 - e^{-ik_\sigma}) (1 - e^{ik'_\sigma}) \right), \quad (3.21)$$

where $\widetilde{\mathbb{T}}_L = \{\frac{2\pi}{L}(n_1, n_2) : 0 \leq n_1, n_2 < L\}$ is the reciprocal torus, $\delta_{\mathbf{p}, \mathbf{q}}$ is the Kronecker delta and

$$A_{\mathbf{k}, \mathbf{k}'}^{(\sigma)} = \frac{1}{|\mathbb{T}_L|} \sum_{x \in \mathbb{T}_L} \kappa_{(x, x + \hat{e}_\sigma)}^{(\alpha)} e^{i(\mathbf{k}' - \mathbf{k}) \cdot x}. \quad (3.22)$$

Now if the horizontal part of $(\kappa_b^{(\alpha)})$ is translation invariant in the γ -th direction, then $A_{\mathbf{k}, \mathbf{k}'}^{(1)} = 0$ whenever $k_\gamma \neq k'_\gamma$, while if it is “only” 2-periodic, then $A_{\mathbf{k}, \mathbf{k}'}^{(1)} = 0$ unless $k_\gamma = k'_\gamma$ or $k_\gamma = k'_\gamma + \pi \bmod 2\pi$. Similar statements apply to the vertical part of $(\kappa_b^{(\alpha)})$ and $A_{\mathbf{k}, \mathbf{k}'}^{(2)}$. Since all of our partition functions come from 2-periodic configurations, the covariance matrix can be cast into a block-diagonal form, with 4×4 blocks $\Theta_\alpha(\mathbf{k})$ collecting all matrix elements that involve the momenta $(\mathbf{k}, \mathbf{k} + \pi\hat{e}_1, \mathbf{k} + \pi\hat{e}_2, \mathbf{k} + \pi\hat{e}_1 + \pi\hat{e}_2)$. Due to the reinsertion of the “zero mode”—cf (3.20)—all of these blocks are non-singular (see also the explicit calculations below).

Hence we get that, for all $\alpha = O, D, UO, UD, MP, MA$,

$$\frac{Z_{L, \alpha}}{(2\pi)^{\frac{1}{2}(L^2-1)}} = \frac{1}{L} p^{N_O} (1-p)^{N_D} \prod_{\mathbf{k} \in \widetilde{\mathbb{T}}_L} \left[\frac{1}{\det \Theta_\alpha(\mathbf{k})} \right]^{1/8}, \quad (3.23)$$

where N_O and N_D denote the numbers of ordered and disordered bonds in the underlying bond configuration and where the exponent $1/8$ takes care of the fact that in the product, each \mathbf{k} gets involved in four distinct terms. Taking logarithms and dividing by $|\mathbb{T}_L|$, the sum over the reciprocal torus converges to a Riemann integral over the Brillouin zone $[-\pi, \pi] \times [-\pi, \pi]$ (the integrand has only logarithmic singularities in all cases, which are harmless for this limit).

It remains to justify the explicit form of the free energies in all cases under considerations. Here the situations $\alpha = O, D, MP$ are fairly standard, so we will

focus on $\alpha = \text{UO}$ and $\alpha = \text{MA}$ for which some non-trivial calculations are needed. In the former case we get that

$$A_{\mathbf{k},\mathbf{k}'}^{(1)} = \kappa_O \delta_{\mathbf{k},\mathbf{k}'}, \quad A_{\mathbf{k},\mathbf{k}}^{(2)} = \frac{\kappa_O + \kappa_D}{2}, \quad A_{\mathbf{k},\mathbf{k}+\pi\hat{\mathbf{e}}_1}^{(2)} = \frac{\kappa_O - \kappa_D}{2}, \quad (3.24)$$

with $A_{\mathbf{k},\mathbf{k}'}^{(\sigma)} = 0$ for all values that are not of this type. Plugging into (3.21) we find that the $(\mathbf{k}, \mathbf{k} + \pi\hat{\mathbf{e}}_1)$ -subblock of $\Theta_{\text{UO}}(\mathbf{k})$ reduces essentially to the 2×2 -matrix in (3.11). Explicitly,

$$\Theta_{\text{UO}}(\mathbf{k}) = \text{diag}(\delta_{\mathbf{k},\mathbf{0}}, \delta_{\mathbf{k},\pi\hat{\mathbf{e}}_1}, \delta_{\mathbf{k},\pi\hat{\mathbf{e}}_2}, \delta_{\mathbf{k},\pi\hat{\mathbf{e}}_1+\pi\hat{\mathbf{e}}_2}) + \begin{pmatrix} \Pi_{\text{UO}}(\mathbf{k}) & 0 \\ 0 & \Pi_{\text{UO}}(\mathbf{k} + \pi\hat{\mathbf{e}}_2) \end{pmatrix}. \quad (3.25)$$

Since $k'_\sigma = k_\sigma$ whenever $A_{\mathbf{k},\mathbf{k}'}^{(\sigma)} \neq 0$, the block matrix $\Theta_{\text{UO}}(\mathbf{k})$ will only be a function of moduli-squared of a_\pm and b_- . Using (3.25) in (3.23) we get (3.10).

As to the MA-case the only non-zero elements of $A_{\mathbf{k},\mathbf{k}'}^{(\sigma)}$ are

$$A_{\mathbf{k},\mathbf{k}}^{(1)} = A_{\mathbf{k},\mathbf{k}}^{(2)} = \frac{\kappa_O + \kappa_D}{2} \quad \text{and} \quad A_{\mathbf{k},\mathbf{k}+\pi\hat{\mathbf{e}}_2}^{(1)} = A_{\mathbf{k},\mathbf{k}+\pi\hat{\mathbf{e}}_1}^{(2)} = \frac{\kappa_O - \kappa_D}{2}. \quad (3.26)$$

So, again, $k'_\sigma = k_\sigma$ whenever $A_{\mathbf{k},\mathbf{k}'}^{(\sigma)} \neq 0$ and so $\Theta_{\text{MA}}(\mathbf{k})$ depends only on $|a_\pm|^2$ and $|b_\pm|^2$. An explicit calculation shows that

$$\Theta_{\text{MA}}(\mathbf{k}) = \text{diag}(\delta_{\mathbf{k},\mathbf{0}}, \delta_{\mathbf{k},\pi\hat{\mathbf{e}}_1}, \delta_{\mathbf{k},\pi\hat{\mathbf{e}}_2}, \delta_{\mathbf{k},\pi\hat{\mathbf{e}}_1+\pi\hat{\mathbf{e}}_2}) + \left(\frac{\kappa_O - \kappa_D}{2} \right) \Pi_{\text{MA}}(\mathbf{k}), \quad (3.27)$$

where $\Pi_{\text{MA}}(\mathbf{k})$ is as in (3.15). Plugging into (3.23), we get (3.14). \square

3.3. Optimal patterns

Next we establish the crucial fact that the spin-wave free energies corresponding to inhomogeneous patterns UO, UD, MP, MA exceed the smaller of F_O and F_D by a quantity that is large, independent of p , once $\kappa_O \gg \kappa_D$.

Theorem 3.3. *There exists $c_1 \in \mathbb{R}$ such that if $\kappa_D \leq \xi \kappa_O$ with $\xi \in (0, 1)$, then for all $p \in (0, 1)$,*

$$\min_{\alpha=\text{UO,UD,MP,MA}} F_\alpha(p) - \min_{\tilde{\alpha}=\text{O,D}} F_{\tilde{\alpha}}(p) \geq \frac{1}{8} \log \frac{\kappa_O}{\kappa_D} + \frac{1}{4} \log(1 - \xi) + c_1. \quad (3.28)$$

Proof. Let us use I and J to denote the integrals

$$I = \frac{1}{2} \int_{[-\pi,\pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log \{ \widehat{D}(\mathbf{k}) \}, \quad \text{and} \quad J = \int_{[-\pi,\pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log |a_-|. \quad (3.29)$$

We will prove (3.28) with $c_1 = J - I$.

First, we have

$$F_O(p) = -2 \log p + \frac{1}{2} \log \kappa_O + I \quad (3.30)$$

and

$$F_D(p) = -2 \log(1-p) + \frac{1}{2} \log \kappa_D + I, \quad (3.31)$$

while an inspection of (3.14) yields

$$\begin{aligned} F_{MA}(p) &= -\log[p(1-p)] \\ &\quad + \frac{1}{8} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log \left\{ \kappa_O \kappa_D (\kappa_O - \kappa_D)^2 |a_+ a_- b_+ b_-|^2 \right\} \\ &\geq -\log[p(1-p)] + \frac{3}{8} \log \kappa_O + \frac{1}{8} \log \kappa_D + \frac{1}{4} \log(1-\xi) + J. \end{aligned} \quad (3.32)$$

Using that

$$\min\{F_O, F_D\} \leq \frac{1}{2}(F_O + F_D), \quad (3.33)$$

we thus get

$$F_{MA}(p) - \min\{F_O(p), F_D(p)\} \geq \frac{1}{8} \log \frac{\kappa_O}{\kappa_D} + \frac{1}{4} \log(1-\xi) + J - I, \quad (3.34)$$

which agrees with (3.28) for our choice of c_1 .

Coming to the free energy F_{UO} , using (3.13) we evaluate

$$\det \Pi_{UO}(\mathbf{k}) \geq \kappa_O^2 |a_-|^2 |a_+|^2 = \left(\frac{\kappa_O}{\kappa_D} \right)^{1/2} \kappa_O^{3/2} \kappa_D^{1/2} |a_-|^2 |a_+|^2 \quad (3.35)$$

yielding

$$F_{UO}(p) \geq -\frac{1}{2} \log[p^3(1-p)] + \frac{1}{8} \log \frac{\kappa_O}{\kappa_D} + \frac{3}{8} \log \kappa_O + \frac{1}{8} \log \kappa_D + J. \quad (3.36)$$

Bounding

$$\min\{F_O, F_D\} \leq \frac{3}{4} F_O + \frac{1}{4} F_D \quad (3.37)$$

we thus get

$$F_{UO}(p) - \min\{F_O(p), F_D(p)\} \geq \frac{1}{8} \log \frac{\kappa_O}{\kappa_D} + J - I, \quad (3.38)$$

in agreement with (3.28). The computation for F_{UD} is completely analogous, interchanging only the roles of κ_O and κ_D as well as p and $(1-p)$. From the lower bound

$$\det \Pi_{UD}(\mathbf{k}) \geq \kappa_O \kappa_D |b_-|^4 = \left(\frac{\kappa_O}{\kappa_D} \right)^{1/2} \kappa_O^{1/2} \kappa_D^{3/2} |b_-|^4 \quad (3.39)$$

and the inequality

$$\min\{F_O, F_D\} \leq \frac{1}{4} F_O + \frac{3}{4} F_D, \quad (3.40)$$

we get again

$$F_{UD}(p) - \min\{F_O(p), F_D(p)\} \geq \frac{1}{8} \log \frac{\kappa_O}{\kappa_D} + J - I, \quad (3.41)$$

which is identical to (3.38).

Finally, for the free energy F_{MP} , we first note that

$$F_{\text{MP}}(p) \geq -\log[p(1-p)] + \frac{1}{2} \log \kappa_0 + J, \quad (3.42)$$

which yields

$$F_{\text{MP}}(p) - F_{\text{D}}(p) \geq \log \frac{1-p}{p} + \frac{1}{2} \log \frac{\kappa_0}{\kappa_D} + J - I. \quad (3.43)$$

Under the condition that $\log \frac{1-p}{p} \geq -\frac{3}{8} \log \frac{\kappa_0}{\kappa_D}$, we again get (3.28). For the complementary values of p , we will compare F_{MP} with F_{O} :

$$F_{\text{MP}}(p) - F_{\text{O}}(p) \geq \log \frac{p}{1-p} + J - I, \quad (3.44)$$

Since we now have $\log \frac{1-p}{p} \leq -\frac{3}{8} \log \frac{\kappa_0}{\kappa_D}$, this yields (3.28) with the above choice of c_1 .

4. Proof of phase coexistence

In this section we will apply the calculations from the previous section to the proof of Theorems 2.2 and 2.4. Throughout this section we assume that $\kappa_0 > \kappa_D$ and that L is even. We begin with a review of the technique of chessboard estimates which, for later convenience, we formulate directly in terms of extended configurations (κ_b, η_b) .

4.1. Review of RP/CE technology

Our principal tool will be chessboard estimates, based on reflection positivity. To define these concepts, let us consider the torus \mathbb{T}_L and let us split \mathbb{T}_L into two symmetric halves, \mathbb{T}_L^+ and \mathbb{T}_L^- , sharing a “plane of sites” on their boundary. We will refer to the set $\mathbb{T}_L^+ \cap \mathbb{T}_L^-$ as *plane of reflection* and denote it by P . The half-tori \mathbb{T}_L^\pm inherit the nearest-neighbor structure from \mathbb{T}_L ; we will use \mathbb{B}_L^\pm to denote the corresponding sets of edges. On the extended configuration space, there is a canonical map $\theta_P: \mathbb{R}^{\mathbb{B}_L} \times \{\kappa_0, \kappa_D\}^{\mathbb{B}_L} \rightarrow \mathbb{R}^{\mathbb{B}_L} \times \{\kappa_0, \kappa_D\}^{\mathbb{B}_L}$ —induced by the reflection of \mathbb{T}_L^+ into \mathbb{T}_L^- through P —which is defined as follows: If $b, b' \in \mathbb{B}_L$ are related via $b' = \theta_P(b)$, then we put

$$(\theta_P \eta)_b = \begin{cases} -\eta_{b'}, & \text{if } b \perp P, \\ \eta_{b'}, & \text{if } b \parallel P, \end{cases} \quad (4.1)$$

and

$$(\theta_P \kappa)_b = \kappa_{b'}. \quad (4.2)$$

Here $b \perp P$ denotes that b is orthogonal to P while $b \parallel P$ indicates that b is parallel to P . The minus sign in the case when $b \perp P$ is fairly natural if we recall that η_b represents the difference of ϕ_x between the endpoints of P . This difference changes sign under reflection through P if $b \perp P$ and does not if $b \parallel P$.

Let \mathcal{F}_P^\pm be the σ -algebras of events that depend only on the portion of (η_b, κ_b) -configuration on \mathbb{B}_L^\pm ; explicitly $\mathcal{F}_P^\pm = \sigma(\eta_b, \kappa_b; b \in \mathbb{B}_L^\pm)$. Reflection positivity is, in its essence, a bound on the correlation between events (and random variables) from \mathcal{F}_P^+ and \mathcal{F}_P^- . The precise definition is as follows:

Definition 4.1. Let \mathbb{P} be a probability measure on configurations $(\eta_b, \kappa_b)_{b \in \mathbb{B}_L}$ and let \mathbb{E} be the corresponding expectation. We say that \mathbb{P} is reflection positive if for any plane of reflection P and any two bounded \mathcal{F}_P^+ -measurable random variables X and Y the following inequalities hold:

$$\mathbb{E}(X\theta_P(Y)) = \mathbb{E}(Y\theta_P(X)) \quad (4.3)$$

and

$$\mathbb{E}(X\theta_P(X)) \geq 0. \quad (4.4)$$

Here, $\theta_P(X)$ denotes the random variable $X \circ \theta_P$.

Next we will discuss how reflection positivity underlines our principal technical tool: chessboard estimates. Consider an event \mathcal{A} that depends only on the (η_b, κ_b) -configurations on the plaquette with the lower-left corner at the torus origin. We will call such an \mathcal{A} a *plaquette event*. For each $x \in \mathbb{T}_L$, we define $\vartheta_x(\mathcal{A})$ to be the event depending only on the configuration on the plaquette with the lower-left corner at x which is obtained from \mathcal{A} as follows: If both components of x are even, then $\vartheta_x(\mathcal{A})$ is simply the translate of \mathcal{A} by x . In the remaining cases we first reflect \mathcal{A} along the side(s) of the plaquette in the direction(s) where the component of x is odd, and then translate the resulting event appropriately. (Thus, there are four possible “versions” of $\vartheta_x(\mathcal{A})$, depending on the parity of x .)

Here is the desired consequence of reflection positivity:

Theorem 4.2 (Chessboard estimate). Let \mathbb{P} be a reflection-positive measure on configurations $(\eta_b, \kappa_b)_{b \in \mathbb{B}_L}$. Then for any plaquette events $\mathcal{A}_1, \dots, \mathcal{A}_m$ and any distinct sites $x_1, \dots, x_m \in \mathbb{T}_L$,

$$\mathbb{P}\left(\bigcap_{j=1}^m \vartheta_{x_j}(\mathcal{A}_j)\right) \leq \prod_{j=1}^m \mathbb{P}\left(\bigcap_{x \in \mathbb{T}_L} \vartheta_x(\mathcal{A}_j)\right)^{\frac{1}{|\mathbb{T}_L|}}. \quad (4.5)$$

Proof. See [15, Theorem 2.2].

The moral of this result—whose proof boils down to the Cauchy-Schwarz inequality for the inner product $X, Y \mapsto \mathbb{E}(X\theta_P(Y))$ —is that the probability of any number of plaquette events factorizes, as a bound, into the product of probabilities. This is particularly useful for contour estimates (of course, provided that the word contour refers to a collection of plaquettes on each of which some “bad” event occurs). Indeed, by (4.5) the probability of a contour will be suppressed exponentially in the number of constituting plaquettes.

In light of (4.5), our estimates will require good bounds on probabilities of the so called *disseminated events* $\bigcap_{x \in \mathbb{T}_L} \vartheta_x(\mathcal{A})$. Unfortunately, the event \mathcal{A} is often a conglomerate of several, more elementary events which makes a direct estimate of $\bigcap_{x \in \mathbb{T}_L} \vartheta_x(\mathcal{A})$ complicated. Here the following subadditivity property will turn out to be useful.

Lemma 4.3 (Subadditivity). *Suppose that \mathbb{P} is a reflection-positive measure and let $\mathcal{A}_1, \mathcal{A}_2, \dots$ and \mathcal{A} be plaquette events such that $\mathcal{A} \subset \bigcup_j \mathcal{A}_j$. Then*

$$\mathbb{P}\left(\bigcap_{x \in \mathbb{T}_L} \vartheta_x(\mathcal{A})\right)^{\frac{1}{|\mathbb{T}_L|}} \leq \sum_j \mathbb{P}\left(\bigcap_{x \in \mathbb{T}_L} \vartheta_x(\mathcal{A}_j)\right)^{\frac{1}{|\mathbb{T}_L|}}. \quad (4.6)$$

Proof. This is Lemma 6.3 of [5].

Apart from the above reflections, which we will call *direct*, one estimate—namely (4.39)—in the proof of Theorem 2.2 requires the use of so called *diagonal reflections*. Assuming L is even, these are reflections in the planes P of sites of the form

$$P_{\pm} = \{y \in \mathbb{T}_L : \hat{e}_1 \cdot (y - x) \mp \hat{e}_2 \cdot (y - x) \in \{0, L/2\}\}. \quad (4.7)$$

Here x is a site that the plane passes through and \hat{e}_1 and \hat{e}_2 are the unit vectors in the x and y -coordinate directions. As before, the plane has two components—one corresponding to $\hat{e}_1 \cdot (y - x) = \pm \hat{e}_2 \cdot (y - x)$ and the other corresponding to $\hat{e}_1 \cdot (y - x) = \pm \hat{e}_2 \cdot (y - x) + L/2$ —and it divides \mathbb{T}_L into two equal parts. This puts us into the setting assumed in Definition 4.1. Some care is needed in the definition of reflected configurations: If b' is the bond obtained by reflecting b through P , then

$$(\theta_P \eta)_b = \begin{cases} \eta_{b'}, & \text{if } P = P_+, \\ -\eta_{b'}, & \text{if } P = P_-. \end{cases} \quad (4.8)$$

This is different compared to (4.1) because the reflection in P_+ preserves orientations of the edges, while that in P_- reverses them.

Remark 4.4. While we will only apply these reflections in $d = 2$, we note that the generalization to higher dimensions is straightforward; just consider all planes as above with (\hat{e}_1, \hat{e}_2) replaced by various pairs (\hat{e}_i, \hat{e}_j) of distinct coordinate vectors. These reflections will of course preserve the orientations of all edges in directions distinct from \hat{e}_i and \hat{e}_j .

4.2. Phase transitions on tori

Here we will provide the proof of phase transition in the form stated in Theorem 2.4. We follow pretty much the standard approach to proofs of order-disorder transitions which dates all the way back to [9, 20, 21]. A somewhat different approach (motivated by another perspective) to this proof can be found in [7].

In order to use the techniques described in the previous section, we have to determine when the extended gradient Gibbs measure Q_L on \mathbb{T}_L obeys the conditions of reflection positivity.

Proposition 4.5. *Let V be of the form (2.5) with any probability measure ϱ for which $Z_{L,V} < \infty$. Then Q_L is reflection positive for both direct and diagonal reflections.*

Proof. The proof is the same for both types of reflections so we proceed fairly generally. Pick a plane of reflection P . Let z be a site on P and let us reexpress the η_b 's back in terms of the ϕ 's with the convention that $\phi_z = 0$. Then

$$\nu_L(d\eta_b) = \delta(d\phi_z) \prod_{x \neq z} d\phi_x. \quad (4.9)$$

Next, let us introduce the quantity

$$W(\eta, \kappa) = \frac{1}{2} \sum_{b \in \mathbb{B}_L^+ \setminus P} \kappa_b \eta_b^2 + \frac{1}{4} \sum_{b \in P} \kappa_b \eta_b^2. \quad (4.10)$$

(We note in passing that the removal of P from the first sum is non-trivial even for diagonal reflections once $d \geq 3$.) Clearly, W is \mathcal{F}_P^+ -measurable and the full (η, κ) -interaction is simply $W(\eta, \kappa) + (\theta_P W)(\eta, \kappa)$. The Gibbs measure Q_L can then be written

$$Q_L(d\eta, d\kappa) = \frac{1}{Z_{L,V}} e^{-W(\nabla\phi, \kappa) - (\theta_P W)(\nabla\phi, \kappa)} \delta(d\phi_z) \left(\prod_{x \neq z} d\phi_x \right) \prod_{b \in \mathbb{B}_L} \rho(d\kappa_b) \quad (4.11)$$

Now pick a bounded, \mathcal{F}_P^+ -measurable function $X = X(\eta, \kappa)$ and integrate the function $X \theta_P X$ with respect to the torus measure Q_L . If \mathcal{G}_P is the σ -algebra generated by random variables ϕ_x and κ_b , with x and b “on” P , we have

$$E_{Q_L}(X \theta_P X | \mathcal{G}_P) \propto \left(\int X(\nabla\phi, \kappa) e^{-W(\nabla\phi, \kappa)} \prod_{x \in \mathbb{T}_L^+ \setminus P} d\phi_x \prod_{b \in \mathbb{B}_L^+ \setminus P} \rho(d\kappa_b) \right)^2 \geq 0, \quad (4.12)$$

where the values of (κ_b, ϕ_x) on P are implicit in the integral. This proves the property in (4.4); the identity (4.3) follows by the reflection symmetry of Q_L .

Let us consider two good plaquette events, \mathcal{G}_{ord} and \mathcal{G}_{dis} , that all edges on the plaquette are ordered and disordered, respectively. Let $\mathcal{B} = (\mathcal{G}_{\text{ord}} \cup \mathcal{G}_{\text{dis}})^c$ denote the corresponding bad event. Given a plaquette event \mathcal{A} , let

$$\mathfrak{z}_{L,p}(\mathcal{A}) = \left[Q_L \left(\bigcap_{x \in \mathbb{T}_L} \vartheta_x(\mathcal{A}) \right) \right]^{\frac{1}{|\mathbb{T}_L|}} \quad (4.13)$$

abbreviate the quantity on the right-hand side of (4.5) and define

$$\mathfrak{z}(\mathcal{A}) = \limsup_{L \rightarrow \infty} \sup_{0 \leq p \leq 1} \mathfrak{z}_{L,p}(\mathcal{A}). \quad (4.14)$$

The calculations from Sect. 3 then permit us to draw the following conclusion:

Lemma 4.6. *For each $\delta > 0$ there exists $c > 0$ such that if $\kappa_0 \geq c\kappa_D$, then*

$$\mathfrak{z}(\mathcal{B}) < \delta. \quad (4.15)$$

Moreover, there exist $p_0, p_1 \in (0, 1)$ such that

$$\limsup_{L \rightarrow \infty} \mathfrak{z}_{L,p}(\mathcal{G}_{\text{ord}}) < \delta, \quad p < p_0, \quad (4.16)$$

and

$$\limsup_{L \rightarrow \infty} \mathfrak{z}_{L,p}(\mathcal{G}_{\text{dis}}) < \delta, \quad p > p_1. \quad (4.17)$$

Proof. The event \mathcal{B} can be decomposed into a disjoint union of events \mathcal{B}_i each of which admits exactly one arrangement of ordered and disordered bonds around the plaquette; see Fig. 3 for the relevant patterns. If \mathcal{B}_i is an event of type $\alpha \in \{\text{O}, \text{D}, \text{UO}, \text{UD}, \text{MP}, \text{MA}\}$, then

$$\limsup_{L \rightarrow \infty} \mathfrak{z}_{L,p}(\mathcal{B}_i) \leq \exp\{-[F_{L,\alpha}(p) - \min_{\tilde{\alpha}=\text{O,D}} F_{\tilde{\alpha}}(p)]\}. \quad (4.18)$$

By Theorem 3.3, the right-hand side is bounded by $O(1)(\kappa_{\text{O}}/\kappa_{\text{D}})^{-1/8}$, uniformly in p . Applying Lemma 4.3, we conclude that $\mathfrak{z}_{L,p}(\mathcal{B})$ is small uniformly in $p \in [0, 1]$ once $L \gg 1$. (The values $p = 0, 1$ are handled by a limiting argument.)

The bounds (4.16–4.17) follow by the fact that

$$F_{\text{O}}(p) - F_{\text{D}}(p) = -2 \log \frac{p}{1-p} + \log \frac{\kappa_{\text{O}}}{\kappa_{\text{D}}}, \quad (4.19)$$

which is (large) negative for p near one and (large) positive for p near zero. \square

From $\mathfrak{z}(\mathcal{B}) \ll 1$ we immediately infer that the bad events occur with very low frequency. Moreover, a standard argument shows that the two good events do not like to occur in the same configuration. An explicit form of this statement is as follows:

Lemma 4.7. *Let R_L^{ord} be the random variable from (2.12). There exists a constant $C < \infty$ such that for all (even) $L \geq 1$ and all $p \in [0, 1]$,*

$$E_{Q_L}(R_L^{\text{ord}}(1 - R_L^{\text{ord}})) \leq C \mathfrak{z}_{L,p}(\mathcal{B}). \quad (4.20)$$

Proof. The claim follows from the fact that, for some constant $C' < \infty$,

$$Q_L(\vartheta_x(\mathcal{G}_{\text{ord}}) \cap \vartheta_y(\mathcal{G}_{\text{dis}})) \leq C' \mathfrak{z}_{L,p}(\mathcal{B})^4, \quad (4.21)$$

uniformly in $x, y \in \mathbb{T}_L$. Indeed, the expectation in (4.20) is the average of the probabilities $Q_L(\kappa_b = \kappa_{\text{O}}, \kappa_{\tilde{b}} = \kappa_{\text{D}})$ over all $b, \tilde{b} \in \mathbb{B}_L$. If x and y denotes the plaquettes containing the bonds b and \tilde{b} , respectively, then this probability is bounded by $Q_L(\vartheta_x(\mathcal{G}_{\text{ord}}) \cap \vartheta_y(\mathcal{G}_{\text{ord}}^c))$. But $\mathcal{G}_{\text{ord}}^c = \mathcal{B} \cup \mathcal{G}_{\text{dis}}$ and so by (4.21) the latter probability is bounded by $\mathfrak{z}_{L,p}(\mathcal{B}) + C' \mathfrak{z}_{L,p}(\mathcal{B})^4 \leq (C' + 1) \mathfrak{z}_{L,p}(\mathcal{B})$, where we used $\mathfrak{z}_{L,p}(\mathcal{B}) \leq 1$.

It remains to prove (4.21). Consider the event $\vartheta_x(\mathcal{G}_{\text{ord}}) \cap \vartheta_y(\mathcal{G}_{\text{dis}})$ where, without loss of generality, $x \neq y$. We claim that on this event, the good plaquettes at x and y are separated from each other by a $*$ -connected circuit of bad plaquettes. To see this, consider the largest connected component of good plaquettes containing x and note that no plaque neighboring on this component can be good, because

(by definition) the events \mathcal{G}_{ord} and \mathcal{G}_{dis} cannot occur at neighboring plaquettes (we are assuming that $\kappa_O \neq \kappa_D$). By chessboard estimates, the probability in \mathcal{Q}_L of any such (given) circuit is bounded by $\mathfrak{z}_{L,p}(\mathcal{B})$ to its size; a standard Peierls' argument in toroidal geometry (cf the proof of [5, Lemma 3.2]) now shows that the probability in (4.21) is dominated by the probability of the shortest possible contour—which is $\mathfrak{z}_{L,p}(\mathcal{B})^4$. (The contour argument requires that $\mathfrak{z}_{L,p}(\mathcal{B})$ be smaller than some constant, but this we may assume to be automatically satisfied because the left-hand side of (4.20) is less than one.) \square

Now we are in a position to prove our claims concerning the torus state:

Proof of Theorem 2.4. Let R_L^{ord} be the fraction of ordered bonds on \mathbb{T}_L (cf. (2.12)) and let $\chi_L(p)$ be the expectation of R_L^{ord} in the extended torus state \mathcal{Q}_L with parameter p . Since $(1-p)^{-|\mathbb{B}_L|} Z_{L,V}$ is log-convex in the variable $h = \log \frac{p}{1-p}$, and

$$|\mathbb{B}_L| \chi_L(p) = \frac{\partial \log((1-p)^{-|\mathbb{B}_L|} Z_{L,V})}{\partial h}, \quad (4.22)$$

we can conclude that the function $p \mapsto \chi_L(p)$ is non-decreasing. Moreover, as the thermodynamic limit of the torus free energy exists (cf Proposition 5.5 in Sect. 5.3), the limit $\chi(p) = \lim_{L \rightarrow \infty} \chi_L(p)$ exists at all but perhaps a countable number of p 's—namely the set $\mathcal{D} \subset [0, 1]$ of points where the limiting free energy is not differentiable.

Next we claim that $\mathcal{Q}_L(|R_L^{\text{ord}} - \chi(p)| > \epsilon_0)$ tends to zero as $L \rightarrow \infty$ for all $\epsilon_0 > 0$ and all $p \notin \mathcal{D}$. Indeed, if this probability stays uniformly positive along some subsequence of L 's for some $\epsilon_0 > 0$, then the boundedness of R_L^{ord} ensures that for some $\zeta > 0$ and some $\epsilon > 0$ we have $\mathcal{Q}_L(R_L^{\text{ord}} > \chi(p) + \epsilon) \geq \zeta$ and $\mathcal{Q}_L(R_L^{\text{ord}} < \chi(p) - \epsilon) \geq \zeta$ for all L in this subsequence. Vaguely speaking, this implies $p \in \mathcal{D}$ because one is then able to extract two infinite-volume Gibbs states with distinct densities of ordered bonds. A formal proof goes as follows: Consider the cumulant generating function $\Psi_L(h) = |\mathbb{B}_L|^{-1} \log E_{\mathcal{Q}_L}(e^{h|\mathbb{B}_L| R_L^{\text{ord}}})$ and note that its thermodynamic limit, $\Psi(h) = \lim_{L \rightarrow \infty} \Psi_L(h)$, is convex in h and differentiable at $h = 0$ whenever $p \notin \mathcal{D}$. But $\mathcal{Q}_L(R_L^{\text{ord}} > \chi(p) + \epsilon) \geq \zeta$ in conjunction with the exponential Chebyshev inequality implies

$$\Psi_L(h) - h(\chi(p) + \epsilon) \geq \frac{\log \zeta}{|\mathbb{B}_L|}, \quad (4.23)$$

which by taking $L \rightarrow \infty$ and $h \downarrow 0$ yields a lower bound on the right derivative at origin, $\frac{d}{dh_+} \Psi(h) \geq \chi(p) + \epsilon$. By the same token $\mathcal{Q}_L(R_L^{\text{ord}} < \chi(p) - \epsilon) \geq \zeta$ implies an upper bound on the left derivative, $\frac{d}{dh_-} \Psi(h) \leq \chi(p) - \epsilon$. Hence, both probabilities can be uniformly positive only if $p \in \mathcal{D}$.

To prove the desired claim it remains to show that χ jumps from values near zero to values near one at some $p_t \in (0, 1)$. To this end we first observe that

$$\lim_{L \rightarrow \infty} E_{\mathcal{Q}_L}(R_L^{\text{ord}}(1 - R_L^{\text{ord}})) = \chi(p)[1 - \chi(p)], \quad p \notin \mathcal{D}. \quad (4.24)$$

This follows by the fact that on the event $\{\chi(p) - \epsilon < R_L^{\text{ord}} < \chi(p) + \epsilon\}$ —whose probability tends to one as $L \rightarrow \infty$ —the quantity $R_L^{\text{ord}}(1 - R_L^{\text{ord}})$ is bounded

between $[\chi(p) + \epsilon](1 - \chi(p) + \epsilon)$ and $[\chi(p) - \epsilon](1 - \chi(p) - \epsilon)$ provided $\epsilon \leq \min\{\chi(p), 1 - \chi(p)\}$. Lemma 4.7 now implies

$$\chi(p)[1 - \chi(p)] \leq C\mathfrak{z}(\mathcal{B}), \quad (4.25)$$

with $\mathfrak{z}(\mathcal{B})$ defined in (4.14). By Lemma 4.6, for each $\delta > 0$ there is a constant $c > 0$ such that

$$\chi(p) \in [0, \delta] \cup [1 - \delta, 1], \quad p \notin \mathcal{D}, \quad (4.26)$$

once $\kappa_0/\kappa_D \geq c$. But the bounds (4.16–4.17) ensure that $\chi(p) \in [0, \delta]$ for $p \ll 1$ and $\chi(p) \in [1 - \delta, 1]$ for $1 - p \ll 1$. Hence, by the monotonicity of $p \mapsto \chi(p)$, there exists a unique value $p_t \in (0, 1)$ such that $\chi(p) \leq \delta$ for $p < p_t$ while $\chi(p) \geq 1 - \delta$ for $p > p_t$. In light of our previous reasoning, this proves the bounds (2.13–2.14). \square

4.3. Phase coexistence in infinite volume

In order to prove Theorem 2.2, we will need to derive a concentration bound on the tilt of the torus states. This is the content of the following lemma:

Lemma 4.8. *Let $\Lambda \subset \mathbb{T}_L$ and let \mathbb{B}_Λ be the set of bonds with both ends in Λ . Given a configuration $(\eta_b)_{b \in \mathbb{B}_L}$, we use $U_\Lambda = U_\Lambda(\eta)$ to denote the vector*

$$U_\Lambda = \left(\frac{1}{|\mathbb{B}_\Lambda|} \sum_{b \in \mathbb{B}_L^{\text{hor}} \cap \mathbb{B}_\Lambda} \eta_b, \frac{1}{|\mathbb{B}_\Lambda|} \sum_{b \in \mathbb{B}_L^{\text{vert}} \cap \mathbb{B}_\Lambda} \eta_b \right) \quad (4.27)$$

of empirical tilt of the configuration η_b in Λ . Suppose that $\kappa_{\min} = \inf \text{supp } \varrho > 0$. Then

$$P_L(|U_\Lambda| \geq \delta) \leq 4e^{-\frac{1}{8}\kappa_{\min}\delta^2|\mathbb{B}_\Lambda|} \quad (4.28)$$

for each $\delta > 0$, each $\Lambda \subset \mathbb{T}_L$ and each L .

Proof. We will derive a bound on the exponential moment of U_Λ . Let us fix a collection of numbers $(v_b)_{b \in \mathbb{B}_L} \in \mathbb{R}^{|\mathbb{B}_L|}$ and let $Q_{L,(\kappa_b)}$ be the conditional law of the η 's given a configuration of the κ 's. Let $Q_{L,\min}$ be the corresponding law when all $\kappa_b = \kappa_{\min}$. In view of the fact that $Q_{L,(\kappa_b)}$ and $Q_{L,\min}$ are Gaussian measures and $\kappa_b \geq \kappa_{\min}$, we have

$$\text{Var}_{Q_{L,(\kappa_b)}} \left(\sum_{b \in \mathbb{B}_L} v_b \eta_b \right) \leq \text{Var}_{Q_{L,\min}} \left(\sum_{b \in \mathbb{B}_L} v_b \eta_b \right). \quad (4.29)$$

(Note that both measures enforce the same loop conditions.) The right-hand side is best calculated in terms of the gradients. The result is

$$\text{Var}_{Q_{L,(\kappa_b)}} \left(\sum_{b \in \mathbb{B}_L} v_b \eta_b \right) \leq \frac{1}{\kappa_{\min}} \sum_{b \in \mathbb{B}_L} v_b^2. \quad (4.30)$$

The fact that $E_{Q_{L,(\kappa_b)}}(\eta_b) = 0$ and the identity $E(e^X) = e^{EX + \frac{1}{2}\text{Var}(X)}$, valid for any Gaussian random variable, now allow us to conclude

$$E_{Q_L} \left(\exp \left\{ \sum_{b \in \mathbb{B}_L} v_b \eta_b \right\} \right) \leq \exp \left\{ \frac{1}{2\kappa_{\min}} \sum_{b \in \mathbb{B}_L} v_b^2 \right\}. \quad (4.31)$$

Choosing $v_b = \lambda \cdot b / |\mathbb{B}_\Lambda|$ on \mathbb{B}_Λ and zero otherwise, we get

$$E_{Q_L}(e^{\lambda \cdot U_\Lambda}) \leq \exp \left\{ \frac{1}{2\kappa_{\min}} \frac{|\lambda|^2}{|\mathbb{B}_\Lambda|} \right\}. \quad (4.32)$$

Noting that $|U_\Lambda| > \delta$ implies that at least one of the components of U_Λ is larger (in absolute value) than $\delta/2$, the desired bound follows by a standard exponentiel-Chebyshev estimate.

Remark 4.9. We note that the symmetry of the law of the η 's in $Q_{L,(\kappa_b)}$ is crucial for the above argument. In particular, it is not clear how to control the tightness of the empirical tilt U_Λ in the measure obtained by normalizing $\exp\{-\sum_b V(\eta_b + h_b)\}v_L(d\eta)$, where $h_b = h \cdot b$ is a “built-in” tilt. In the strictly convex cases, these measures were used by Funaki and Spohn [17] to construct an infinite-volume gradient Gibbs state with a given value of the tilt.

Proof of Theorem 2.2. With Theorem 2.4 at our disposal, the argument is fairly straightforward. Consider a weak (subsequential) limit of the torus states at $p > p_t$ and then consider another weak limit of these states as $p \downarrow p_t$. Denote the result by $\tilde{\mu}_{\text{ord}}$. Next let us perform a similar limit as $p \uparrow p_t$ and let us denote the resulting measure by $\tilde{\mu}_{\text{dis}}$. As is easy to check, both measures are extended gradient Gibbs measures at parameter p_t .

Next we will show that the two measures are distinct measures of zero tilt. To this end we recall that, by (2.14) and the invariance of Q_L under rotations, $\liminf_{L \rightarrow \infty} Q_L(\kappa_b = \kappa_0) \geq 1 - \epsilon$ when $p > p_t$ while (2.13) implies that $\limsup_{L \rightarrow \infty} Q_L(\kappa_b = \kappa_0) \leq \epsilon$ when $p < p_t$. But $\{\kappa_b = \kappa_0\}$ is a local event and so

$$\tilde{\mu}_{\text{ord}}(\kappa_b = \kappa_0) \geq 1 - \epsilon \quad (4.33)$$

while

$$\tilde{\mu}_{\text{dis}}(\kappa_b = \kappa_0) \leq \epsilon, \quad (4.34)$$

for all b ; i.e., $\tilde{\mu}_{\text{ord}} \neq \tilde{\mu}_{\text{dis}}$. Moreover, the bound (4.28)—being uniform in p and L —survives the above limits unscathed and so the tilt is exponentially tight in volume for both $\tilde{\mu}_{\text{ord}}$ and $\tilde{\mu}_{\text{dis}}$. It follows that $U_\Lambda \rightarrow 0$ as $\Lambda \uparrow \mathbb{Z}^2$ almost surely with respect to both $\tilde{\mu}_{\text{ord}}$ and $\tilde{\mu}_{\text{dis}}$; i.e., both measures are supported entirely on configurations with zero tilt.

It remains to prove the inequalities (2.10–2.11) and thereby ensure that the η -marginals μ_{ord} and μ_{dis} of $\tilde{\mu}_{\text{ord}}$ and $\tilde{\mu}_{\text{dis}}$, respectively, are distinct as claimed in the statement of the theorem. The first bound is a consequence of the identity

$$\lim_{L \rightarrow \infty} E_{Q_L}(\kappa_b \eta_b^2) = \frac{1}{4}, \quad (4.35)$$

which extends via the aforementioned limits to $\tilde{\mu}_{\text{ord}}$ (as well as $\tilde{\mu}_{\text{dis}}$). Indeed, using Chebyshev's inequality and the fact that $\mu_{\text{ord}}(\kappa_b = \kappa_D) \leq \epsilon$ we get

$$\begin{aligned} \tilde{\mu}_{\text{ord}}(\kappa_O \eta_b^2 \geq \lambda^2) - \epsilon &\leq \tilde{\mu}_{\text{ord}}(\kappa_O \eta_b^2 \geq \lambda^2, \kappa_b = \kappa_O) \\ &\leq \tilde{\mu}_{\text{ord}}(\kappa_b \eta_b^2 \geq \lambda^2) \leq \frac{E_{\tilde{\mu}_{\text{ord}}}(\kappa_b \eta_b^2)}{\lambda^2} = \frac{1}{4\lambda^2}. \end{aligned} \quad (4.36)$$

To prove (4.35), the translation and rotation invariance of Q_L gives us

$$E_{Q_L}(\kappa_b \eta_b^2) = E_{Q_L}\left(E_{Q_{L,(\kappa_b)}}\left(\frac{1}{|\mathbb{B}_L|} \sum_{b \in \mathbb{B}_L} \kappa_b \eta_b^2\right)\right). \quad (4.37)$$

Let $Z_{L,(\kappa_b)}$ denote the integral of $\exp\{-\frac{1}{2} \sum_b \kappa_b \eta_b^2\}$ with respect to ν_L . Since we have $\nu_L(\beta d\eta) = \beta^{|\mathbb{T}_L|-1} \nu_L(d\eta)$, simple scaling of all fields yields $Z_{L,(\beta \kappa_b)} = \beta^{-\frac{1}{2}(|\mathbb{T}_L|-1)} Z_{L,(\kappa_b)}$. Interpreting the inner expectation above as the (negative) β -derivative of $|\mathbb{B}_L|^{-1} \log Z_{L,(\beta \kappa_b)}$ at $\beta = 1$, we get

$$E_{Q_{L,(\kappa_b)}}\left(\frac{1}{|\mathbb{B}_L|} \sum_{b \in \mathbb{B}_L} \kappa_b \eta_b^2\right) = \frac{|\mathbb{T}_L| - 1}{2|\mathbb{B}_L|}. \quad (4.38)$$

From here (4.35) follows by taking $L \rightarrow \infty$ on the right-hand side.

As to the inequality (2.11) for the disordered state, here we first use that the diagonal reflection allows us to disseminate the event $\{\kappa_b \eta_b^2 \leq \lambda^2\}$ around any plaquette containing b . Explicitly, if (b_1, b_2, b_3, b_4) is a plaquette, then

$$Q_L(\kappa_{b_1} \eta_{b_1}^2 \leq \lambda^2) \leq Q_L\left(\bigcap_{b=b_1, \dots, b_4} \{\kappa_b \eta_b^2 \leq \lambda^2\}\right)^{1/4}. \quad (4.39)$$

(We are using that the event in question is even in η and so the changes of sign of η_b are immaterial.) Direct reflections now permit us to disseminate the resulting plaquette event all over the torus:

$$Q_L(\kappa_b \eta_b^2 \leq \lambda^2) \leq Q_L\left(\bigcap_{\tilde{b} \in \mathbb{B}_L} \{\kappa_{\tilde{b}} \eta_{\tilde{b}}^2 \leq \lambda^2\}\right)^{\frac{1}{4|\mathbb{T}_L|}}. \quad (4.40)$$

Bounding the indicator of the giant intersection by

$$e^{(\beta-1)\lambda^2|B_L|} \exp\left\{-\frac{1}{2}(\beta-1) \sum_{b \in \mathbb{B}_L} \kappa_b \eta_b^2\right\}, \quad (4.41)$$

for $\beta \geq 1$, and invoking the scaling of the partition function $Z_{L,(\beta \kappa_b)}$, we deduce

$$Q_L(\kappa_b \eta_b^2 \leq \lambda^2) \leq \left[\frac{e^{(\beta-1)\lambda^2|B_L|}}{\beta^{\frac{1}{2}(|\mathbb{T}_L|-1)}}\right]^{\frac{1}{4|\mathbb{T}_L|}}. \quad (4.42)$$

Choosing $\beta = \lambda^{-2}$, letting $L \rightarrow \infty$ and $p \uparrow p_t$, we thus conclude

$$\tilde{\mu}_{\text{dis}}(\kappa_b \eta_b^2 \leq \lambda^2) \leq c_1 \lambda^{1/4}. \quad (4.43)$$

Noting that $\tilde{\mu}_{\text{dis}}(\kappa_b \eta_b^2 \leq \lambda^2) \geq \tilde{\mu}_{\text{dis}}(\kappa_D \eta_b^2 \leq \lambda^2) - \epsilon$, the bound (2.11) is also proved. \square

5. Duality arguments

The goal of this section is to prove Theorem 2.5. For that we will establish an interesting duality that relates the model with parameter p to the same model with parameter $1 - p$.

5.1. Preliminary considerations

The duality relation that our model (1.6) satisfies boils down, more or less, to an algebraic fact that the plaquette condition (1.3), represented by the delta function $\delta(\eta_{b_1} + \eta_{b_2} - \eta_{b_3} - \eta_{b_4})$, can formally be written as

$$\delta(\eta_{b_1} + \eta_{b_2} - \eta_{b_3} - \eta_{b_4}) = \int \frac{d\phi^*}{2\pi} e^{i\phi^*(\eta_{b_1} + \eta_{b_2} - \eta_{b_3} - \eta_{b_4})}. \quad (5.1)$$

We interpret the variable ϕ^* as the *dual field* that is associated with the plaquette $(\eta_{b_1}, \eta_{b_2}, \eta_{b_3}, \eta_{b_4})$. As it turns out (see Theorem 5.3), by integrating the η 's with the ϕ^* 's fixed a gradient measure is produced whose interaction is the same as for the η 's, except that the κ_b 's get replaced by $1/\kappa_b$'s. This means that if we assume that

$$\kappa_O \kappa_D = 1, \quad (5.2)$$

which is permissible in light of the remarks at the beginning of Section 2.2, then the duality simply exchanges κ_O and κ_D ! We will assume that (5.2) holds throughout this entire section.

The aforementioned transformation works nicely for the plaquette conditions which guarantee that the η 's can *locally* be integrated back to the ϕ 's. However, in two-dimensional torus geometry, two additional global constraints are also required to ensure the *global* correspondence between the gradients η and the ϕ 's. These constraints, which are by definition built into the *a priori* measure ν_L from Sect. 2, do not transform as nicely as the local plaquette conditions. To capture these subtleties, we will now define another *a priori* measure that differs from ν_L in that it disregards these global constraints.

Consider the linear subspace $\mathcal{X}_L^* \supset \mathcal{X}_L$ of $\mathbb{R}^{\mathbb{B}_L}$ that is characterized by the equations $\eta_{b_1} + \eta_{b_2} - \eta_{b_3} - \eta_{b_4} = 0$ for each plaquette (b_1, b_2, b_3, b_4) . This space inherits the Euclidean metric from $\mathbb{R}^{\mathbb{B}_L}$; we define ν_L^* as the corresponding Lebesgue measure on \mathcal{X}_L^* scaled by a constant C_L which will be determined momentarily. In order to make the link with ν_L , we define

$$\eta_{\text{vert}} = \sum_{x \in \mathbb{T}_L} \eta_{x+\hat{e}_1} \quad \text{and} \quad \eta_{\text{hor}} = \sum_{x \in \mathbb{T}_L} \eta_{x+\hat{e}_2}. \quad (5.3)$$

Clearly,

$$\mathcal{X}_L = \{\eta \in \mathcal{X}_L^*: \eta_{\text{vert}} = 0, \eta_{\text{hor}} = 0\}. \quad (5.4)$$

Consider also the projection $\Pi_L: \mathcal{X}_L^* \rightarrow \mathcal{X}_L$ which is defined, for any configuration $\eta \in \mathbb{R}^{\mathbb{B}_L}$, by

$$(\Pi_L \eta)_b = \begin{cases} \eta_b - \frac{1}{L^2} \eta_{\text{vert}}, & \text{if } b \in \mathbb{B}_L^{\text{vert}}, \\ \eta_b - \frac{1}{L^2} \eta_{\text{hor}}, & \text{if } b \in \mathbb{B}_L^{\text{hor}}. \end{cases} \quad (5.5)$$

Then we have:

Lemma 5.1. *There exist constants C_L such that, in the sense of distributions,*

$$v_L^*(d\eta) = \left(L^2 \int_{\mathbb{R}} d\theta \prod_{(b_1, b_2, b_3, b_4)} \delta(\eta_{b_1} + \eta_{b_2} - \eta_{b_3} - \eta_{b_4} - \theta) \right) \prod_{b \in \mathbb{T}_L} d\eta_b. \quad (5.6)$$

Moreover, we have

$$v_L(d\eta) = v_L^*(d\eta) \delta(\eta_{\text{hor}}) \delta(\eta_{\text{vert}}) \quad (5.7)$$

and

$$v_L^*(d\eta) = v_L \circ \Pi_L(d\eta) \lambda_{\Pi_L}(d\eta). \quad (5.8)$$

Here, $\lambda_{\Pi_L}(d\eta)$ is a multiple of the Lebesgue measure on the two-dimensional space $\Pi_L^{-1}(0) \cap \mathcal{X}_L^*$, which can be formally identified with $d\eta_{\text{hor}} d\eta_{\text{vert}}$.

Proof. We begin with (5.6). Consider the orthogonal decomposition $\mathbb{R}^{|\mathbb{B}_L|} = \mathcal{X}_L^* \oplus (\mathcal{X}_L^*)^\perp$. Clearly, $\dim \mathcal{X}_L^* = L^2 + 1$. Choosing an orthonormal basis w_1, \dots, w_n in $(\mathcal{X}_L^*)^\perp$ (where $n = \dim(\mathcal{X}_L^*)^\perp = L^2 - 1$) the measure v_L^* can be written as

$$v_L^*(d\eta) = C_L \left(\prod_{j=1}^n \delta(w_j \cdot \eta) \right) \prod_{b \in \mathbb{B}_L} d\eta_b. \quad (5.9)$$

Let ℓ_π denote the vectors in $\mathbb{R}^{|\mathbb{B}_L|}$ such that if $\pi = (b_1, b_2, b_3, b_4)$ then $\ell_\pi \cdot \eta = \eta_{b_1} + \eta_{b_2} - \eta_{b_3} - \eta_{b_4}$. Then $\ell_\pi \in (\mathcal{X}_L^*)^\perp$ with all but one of these vectors linearly independent. This means that we can replace the linear functionals $\eta \mapsto w_j \cdot \eta$ by the plaquette conditions. Fixing a particular plaquette, π_0 , we find that

$$v_L^*(d\eta) = \left(\prod_{\pi \neq \pi_0} \delta(\ell_\pi \cdot \eta) \right) \prod_{b \in \mathbb{T}_L} d\eta_b \quad (5.10)$$

provided that

$$C_L = |\det(w_j \cdot \ell_\pi)| = \sqrt{\det(\ell_\pi \cdot \ell_{\pi'})}. \quad (5.11)$$

The expression (5.10) is now easily checked to be equivalent to (5.6): Applying the constraints from the plaquettes distinct from π , we find that $\eta_{b_1} + \eta_{b_2} - \eta_{b_3} - \eta_{b_4} = (1 - L^2)\theta$. The corresponding δ -function becomes $\delta(L^2\theta)$, and so we can set $\theta = 0$ in the remaining δ -functions. Integration over θ yields an overall multiplier $\int_{\mathbb{R}} \delta(L^2\theta) d\theta = 1/L^2$.

In order to prove (5.7), pick a subtree \mathcal{T} of \mathbb{T}_L as follows: \mathcal{T} contains the horizontal bonds in $\{b_1 + \ell\hat{e}_1 : \ell = 0, \dots, L-2\}$ and the vertical bonds in $\{b_2 + \ell\hat{e}_1 + m\hat{e}_2 : \ell, m = 0, \dots, L-2\}$. As is easy to check, \mathcal{T} is a spanning tree. Denoting by γ_L the measure on the right-hand side of (5.7) pick a bounded, continuous function $f : \mathbb{R}^{|\mathbb{B}_L|} \rightarrow \mathbb{R}$ with bounded support and consider the integral $\int f(\eta) \gamma_L(d\eta)$. The complement of \mathcal{T} contains exactly $L^2 + 1$ edges and there are as many δ -functions in (5.10) and (5.7), in which all η_b , $b \notin \mathcal{T}$, appear with coefficient ± 1 . We may thus resolve these constraints and substitute for all $\{\eta_b : b \notin \mathcal{T}\}$ into f —call the result of this substitution $\tilde{f}(\eta)$. Then

we can integrate all of these variables which reduces our attention to the integral $\int \tilde{f}(\eta) \prod_{b \in \mathcal{T}} d\eta_b$.

As is easy to check, the transformation $\eta_b = \phi_y - \phi_x$ for $b = (x, y)$ with the convention $\phi_0 = 0$ turns the measure $\prod_{b \in \mathcal{T}} d\eta_b$ into $\delta(d\phi_0) \prod_{x \neq 0} d\phi_x$ and makes $\tilde{f}(\eta)$ into $f(\nabla\phi)$. We have thus deduced

$$\int_{\mathbb{R}^{|\mathbb{B}_L|}} f(\eta) \gamma_L(d\eta) = \int_{\mathbb{R}^{|\mathcal{T}|}} \tilde{f}(\eta) \prod_{b \in \mathcal{T}} d\eta_b = \int_{\mathbb{R}^{|\mathbb{T}_L|}} f(\nabla\phi) \delta(d\phi_0) \prod_{x \neq 0} d\phi_x. \quad (5.12)$$

From here we get (5.7) by noting that the latter integral can also be written as $\int f(\eta) v_L(d\eta)$.

To derive $v_L^* = v_L \circ \Pi_L d\eta_{\text{hor}} d\eta_{\text{vert}}$, let us write $\mathcal{X}_L^* = \mathcal{X}_L \oplus (\Pi_L^{-1}(0) \cap \mathcal{X}_L^*)$. Since v_L^* is the C_L -multiple of the Lebesgue measure on \mathcal{X}_L^* and since η_{hor} and η_{vert} represent orthogonal coordinates in $\Pi_L^{-1}(0) \cap \mathcal{X}_L^*$, we have

$$v_L^*(d\eta) = C_L \lambda_{\mathcal{X}_L} \circ \Pi_L(d\eta) L^{-2} d\eta_{\text{hor}} L^{-2} d\eta_{\text{vert}}, \quad (5.13)$$

where $\lambda_{\mathcal{X}_L}$ is the Lebesgue measure on \mathcal{X}_L . Plugging into (5.7) we find that $v_L = C_L L^{-4} \lambda_{\mathcal{X}_L}$ which in turn implies (5.8). \square

Remark 5.2. It is of some interest to note that the measure v_L^* is also reflection positive for direct reflections. One proof of this fact goes by replacing the δ -functions in (5.6) by Gaussian kernels and noting that the linear term in θ (in the exponent) exactly cancels. The status of reflection positivity for the diagonal reflections is unclear.

5.2. Duality for inhomogeneous Gaussian measures

Now we can state the principal duality relation. For that let \mathbb{T}_L^* denote the dual torus which is simply a copy of \mathbb{T}_L shifted by half lattice spacing in each direction. Let \mathbb{B}_L^* denote the set of dual edges. We will adopt the convention that if b is a direct edge, then its dual—i.e., the unique edge in \mathbb{B}_L^* that cuts through b —will be denoted by b^* . Then we have:

Theorem 5.3. *Given two collections $(\kappa_b)_{b \in \mathbb{B}_L}$ and $(\kappa_b^*)_{b \in \mathbb{B}_L}$ of positive weights on \mathbb{B}_L , consider the partition functions*

$$Z_{L,(\kappa_b)} = \int v_L(d\eta) \exp\left\{-\frac{1}{2} \sum_{b \in \mathbb{B}_L} \kappa_b \eta_b^2\right\} \quad (5.14)$$

and

$$Z_{L,(\kappa_b^*)}^* = \int v_L^*(d\eta) \exp\left\{-\frac{1}{2} \sum_{b \in \mathbb{B}_L} \kappa_b^* \eta_b^2\right\}. \quad (5.15)$$

If $(\kappa_b)_{b \in \mathbb{B}_L}$ and $(\kappa_b^*)_{b \in \mathbb{B}_L}$ are dual in the sense that

$$\kappa_{b^*}^* = \frac{1}{\kappa_b}, \quad b \in \mathbb{B}_L, \quad (5.16)$$

then

$$Z_{L,(\kappa_b^*)}^* = 2\pi L^2 \left[\prod_{b \in \mathbb{B}_L} \sqrt{\kappa_b} \right] Z_{L,(\kappa_b)}. \quad (5.17)$$

Proof. We will cast the partition function $Z_{L,(\kappa_b^*)}^*$ into the form on the right-hand side of (5.17). Let us regard this partition function as defined on the dual torus \mathbb{T}_L^* . The proof commences by rewriting the definition (5.6) with the help of (5.1) as

$$\nu_L^*(d\eta) = \left(L^2 \int_{\mathbb{R}} d\theta \int_{\mathbb{R}^{\mathbb{T}_L}} \prod_{x \in \tilde{\mathbb{T}}_L} \frac{d\phi_x}{2\pi} \exp \left\{ i \sum_{x \in \mathbb{T}_L} \phi_x (\eta_{\pi^*(x)} - \theta) \right\} \right) \prod_{b^* \in \mathbb{B}_L^*} d\eta_{b^*}, \quad (5.18)$$

where $\eta_{\pi^*(x)}$ is the plaquette curl for the dual plaquette $\pi^*(x)$ with the center at x . Rearranging terms and multiplying by the exponential (Gaussian) weight from (5.15), we are thus supposed to integrate the function

$$L^2 \int_{\mathbb{R}} d\theta \int_{\mathbb{R}^{\mathbb{T}_L}} \prod_{x \in \mathbb{T}_L} \frac{d\phi_x}{2\pi} \exp \left\{ -\frac{1}{2} \sum_{b^* \in \mathbb{B}_L^*} (\kappa_{b^*}^* \eta_{b^*}^2 - 2i\eta_{b^*} \nabla_b \phi) - i\theta \sum_{x \in \mathbb{T}_L} \phi_x \right\} \quad (5.19)$$

against the (unconstrained) Lebesgue measure $\prod_{b \in \mathbb{B}_L} d\eta_b$. Here $\nabla_b \phi = \phi_y - \phi_x$ if $b = (x, y)$ is dual to the bond b^* . Completing the squares and integrating over the η 's produces the function

$$L^2 \left[\prod_{b^* \in \mathbb{B}_L^*} \frac{1}{\sqrt{\kappa_{b^*}^*}} \right] \int_{\mathbb{R}} d\theta \int_{\mathbb{R}^{\mathbb{T}_L}} \prod_{x \in \mathbb{T}_L} d\phi_x \exp \left\{ -\frac{1}{2} \sum_{b \in \mathbb{B}_L} \frac{1}{\kappa_b^*} (\nabla_b \phi)^2 - i\theta \sum_{x \in \mathbb{T}_L} \phi_x \right\}. \quad (5.20)$$

Invoking (5.16), we can replace all $1/\kappa_{b^*}^*$ by κ_b . The integral over θ then yields 2π times the δ -function of $\sum_x \phi_x$ which—by the substitution $\phi_x \mapsto \phi_x + \frac{1}{|\mathbb{T}_L|}\phi_0$ that has no effect on the rest of the integral—can be converted to $2\pi\delta(\phi_0)$. Invoking the definition (2.1) of ν_L , this leads to the partition function (5.14). \square

Remark 5.4. Let $Q_{L,p}$ be the extended gradient Gibbs measure Q_L for $\varrho = p\delta_{\kappa_0} + (1-p)\delta_{\kappa_D}$ with parameter p and let $Q_{L,p}^*$ be the corresponding measure with the *a priori* measure ν_L replaced by ν_L^* . Then the above duality shows that the law of (κ_b) governed by $Q_{L,p}$ is the same as the law of its dual (κ_b^*) —defined via (5.16)—in measure Q_{L,p_*}^* , once p and p_* are related by

$$\frac{p}{1-p} \frac{p_*}{1-p_*} = \sqrt{\frac{\kappa_D}{\kappa_0}}. \quad (5.21)$$

Indeed, the probability in measure Q_{L,p_*}^* of seeing the configuration (κ_b^*) with N_O^* ordered bonds and N_D^* disordered bonds is proportional to $p_*^{N_O^*} (1-p_*)^{N_D^*} Z_{L,(\kappa_b^*)}^*$. Considering the dual configuration (κ_b) and letting $N_D = N_O^*$ denote the number of disordered bonds and $N_O = N_D^*$ the number of ordered bonds in (κ_b) , we thus have

$$p_*^{N_O^*} (1-p_*)^{N_D^*} Z_{L,(\kappa_b^*)}^* = 2\pi L^2 (p_* \sqrt{\kappa_0})^{N_D} ((1-p_*) \sqrt{\kappa_D})^{N_O} Z_{L,(\kappa_b)}. \quad (5.22)$$

For p and p_* related as in (5.21), the right-hand side is proportional to the probability of (κ_b) in measure $Q_{L,p}$. We believe that the difference between the two measures disappears in the limit $L \rightarrow \infty$ and so the κ -marginals of the states $\tilde{\mu}_{\text{ord}}$ and $\tilde{\mu}_{\text{dis}}$ at p_t can be considered to be dual to each other. However, we will not pursue this detail at any level of rigor.

5.3. Computing the transition point

In order to use effectively the duality relation from Theorem 5.3, we have to show that the difference in the *a priori* measure can be neglected. We will do this by showing that both partition functions lead to the same free energy. This is somewhat subtle due to the presence (and absence) of various constraints, so we will carry out the proof in detail.

Proposition 5.5. *Let $\varrho_L(d\kappa) = \prod_{b \in \mathbb{B}_L} [p\delta_{\kappa_0}(d\kappa_b) + (1-p)\delta_{\kappa_D}(d\kappa_b)]$ and recall that $Z_{L,V}$ denotes the integral of $Z_{L,(\kappa_b)}$ with respect to ϱ_L . Similarly, let $Z_{L,V}^*$ denote the integral of $Z_{L,(\kappa_b)}^*$ with respect to ϱ_L . Then (the following limits exist as $L \rightarrow \infty$ and)*

$$\lim_{L \rightarrow \infty} \frac{1}{|\mathbb{T}_L|} \log Z_{L,V} = \lim_{L \rightarrow \infty} \frac{1}{|\mathbb{T}_L|} \log Z_{L,V}^* \quad (5.23)$$

for all $p \in [0, 1]$.

Before we commence with the proof, let us establish the following variance bounds for homogeneous Gaussian measures relative to the *a priori* measure ν_L and ν_L^* :

Lemma 5.6. *Let μ_L be the (standard) Gaussian gradient measure*

$$\mu_L(d\eta) \propto \exp\left\{-\frac{1}{2} \sum_{b \in \mathbb{B}_L} \eta_b^2\right\} \nu_L(d\eta) \quad (5.24)$$

and μ_L^* be the measure obtained by replacing ν_L by ν_L^* . For $m = 1, \dots, L$, let

$$Y_m = \sum_{\ell=0}^{m-1} \eta_{(\ell\hat{e}_1, (\ell+1)\hat{e}_1)}. \quad (5.25)$$

There exists an absolute constant $c_3 > 0$ such that for all $L \geq 1$ and all $m = 1, \dots, L$,

$$\text{Var}_{\mu_L}(Y_m) \leq \text{Var}_{\mu_L^*}(Y_m) \leq c_3(1 + \log m). \quad (5.26)$$

Proof. In measure μ_L , we can reintroduce back the fields (ϕ_x) and Y_m then equals $\phi_{m\hat{e}_1}$. Discrete Fourier transform implies that

$$\text{Var}_{\mu_L}(\phi_x) = \frac{1}{L^2} \sum_{\mathbf{k} \in \tilde{\mathbb{T}}_L} \frac{|1 - e^{ix \cdot \mathbf{k}}|^2}{\widehat{D}(\mathbf{k})}, \quad (5.27)$$

where $\tilde{\mathbb{T}}_L$ is the reciprocal torus and $\widehat{D}(\mathbf{k}) = |1 - e^{ik_1}|^2 + |1 - e^{ik_2}|^2$ is the discrete (torus) Laplacian. Simple estimates show that the sum is bounded by a constant times $1 + \log|x|$, uniformly in L . Hence, $\text{Var}_{\mu_L}(Y_m) \leq \tilde{c}_3(1 + \log m)$ for some absolute constant \tilde{c}_3 .

As for the other measure, we recall the definitions (5.3) and use these to write $\eta_b = \nabla_b \phi + \frac{1}{L^2} \eta_{\text{hor}}$ if b is horizontal (and $\eta_b = \nabla_b \phi + \frac{1}{L^2} \eta_{\text{vert}}$ if b is vertical). The fact that the Gaussian field is homogeneous implies—via (5.8)—that the fields (ϕ_x) and the variables η_{vert} and η_{hor} are independent with (ϕ_x) distributed according to μ_L and η_{vert} and η_{hor} Gaussian with mean zero and variance $L^2/2$. In this case $Y_m = \phi_{m\hat{e}_1} + \frac{m}{L^2} \eta_{\text{hor}}$ and so we get

$$\text{Var}_{\mu_L^*}(Y_m) = \text{Var}_{\mu_L}(Y_m) + \frac{m^2}{2L^2}. \quad (5.28)$$

But $m \leq L$ and so the correction is bounded for all L . \square

Proof of Proposition 5.5. The proof follows the expected line: To compensate for the lack of obvious subadditivity of the torus partition function, we will first relate the periodic boundary condition to a “fixed” boundary condition. Then we will establish subadditivity—and hence the existence of the free energy—for the latter boundary condition.

Fix $M > 0$ and consider the partition function $Z_{L,V}^{(M)}$ defined as follows. Let Λ_L be a box of $L \times L$ sites and consider the set \mathbb{B}_L^0 of edges with *both* ends in Λ_L . Let $v_L^{(M)}(d\eta)$ be as in (2.1) subject to the restriction that $|\phi_x| \leq M$ for all x on the *internal* boundary of Λ_L . Let

$$Z_{L,V}^{(M)} = \int \varrho_L(d\kappa) \int v_L^{(M)}(d\eta) \exp\left\{-\frac{1}{2} \sum_{b \in \mathbb{B}_L^0} \kappa_b \eta_b^2\right\}. \quad (5.29)$$

We will now provide upper and lower bounds between the partition functions $Z_{L,V}$ (resp. $Z_{L,V}^*$) and $Z_{L,V}^{(M)}$, for a well defined range of values of M .

Comparing explicit expressions for $Z_{L,V}$ and $Z_{L,V}^{(M)}$ and using $\kappa_b \leq \kappa_0$, we get

$$Z_{L,V} \geq Z_{L,V}^{(M)} \exp\left\{-\frac{1}{2} \kappa_0 (2M)^2 (2L)\right\}. \quad (5.30)$$

To derive an opposite inequality, note that for $\kappa_b \geq \kappa_D$ we get that $\text{Var}_{Q_{L,(\kappa_b)}}(\phi_x) \geq \text{Var}_{\mu_L}(\phi_x)/\kappa_D$, where μ_L is as in (5.24). Invoking one more time the Gaussian identity $E(e^X) = e^{EX + \frac{1}{2}\text{Var}(X)}$ in conjunction with Lemma 5.6, yields

$$Q_L(\phi_x \geq M) \leq \exp\left\{-\frac{1}{2} \frac{M^2}{\kappa_D c_3 (1 + \log L)}\right\}. \quad (5.31)$$

Hence, if $M \gg \log L$ we have that with probability at least $1/2$ in measure Q_L , all variables ϕ_x are in the interval $[-M, M]$. Since the interaction that wraps Λ_L into the torus is of definite sign, it follows that

$$Z_{L,V} \leq 2Z_{L,V}^{(M)} \quad (5.32)$$

for all L and all $M \gg \log L$.

Concerning the star-partition function, Lemma 5.6 makes the proof of (5.32) exactly the same. As for the alternative of (5.30), we invoke (5.8) and restrict all $|\phi_x|$ on the internal boundary of Λ_L to values less than M and $|\eta_{\text{hor}}|$ and $|\eta_{\text{vert}}|$ to values less than ML^2 . Since $|\eta_b| = |\nabla_b \phi + \frac{1}{L^2} \eta_{\text{vert}}| \leq 2M + M = 3M$ for every vertical bond that wraps Λ_L into the torus (and similarly for the horizontal bonds), we now get

$$Z_{L,V}^* \geq Z_{L,V}^{(M)} (2ML)^2 \exp\left\{-\frac{1}{2}\kappa_0(3M)^2(2L)\right\}, \quad (5.33)$$

where the factor $(2ML)^2$ comes from the integration over η_{vert} and η_{hor} . We conclude that, for $\log L \ll M = o(\sqrt{L})$, the partition functions $Z_{L,V}$, $Z_{L,V}^*$ and $Z_{L,V}^{(M)}$ lead to the same free energy, provided at least one of these exists.

It remains to establish that the partition function $Z_{L,V}^{(M)}$ is (approximately) sub-multiplicative for some choice of $M = M_L$. Choose, e.g., $M_L = (\log L)^2$ and let $p \geq 1$ be an integer. If two neighbors have their ϕ 's between $-M_L$ and M_L , the energy across the bond is at most $\frac{1}{2}\kappa_0(4M_L)^2$. Splitting Λ_{pL} into p^2 boxes of size L , and restricting the ϕ 's to $[-M_L, M_L]$ on the internal boundaries of these boxes, we thus get

$$Z_{pL,V}^{(M_{pL})} \geq [Z_{L,V}^{(M_L)}]^{p^2} \exp\left\{-\frac{1}{2}\kappa_0(2M_L)^2 2(p-1)L\right\}. \quad (5.34)$$

The exponent can be bounded below by $(pL)^{3/2} - p^2 L^{3/2} = -(p^2 - p^{3/2})L^{3/2}$ for L sufficiently large which implies that $p \mapsto [Z_{pL,V}^{(M_{pL})} \exp\{-(pL)^{3/2}\}]^{1/(pL)^2}$ is increasing for all $p \geq 1$ and all $L \gg 1$. This proves the claim for limits along multiples of any fixed L ; to get the values “in-between” we just need to realize that, as before, $Z_{L+k,V}^{(M_{L+k})} \geq Z_{L,V}^{(M_L)} e^{O(kLM_L^2)}$, for any fixed k . \square

Now we finally prove our claim concerning the value of the transitional p :

Proof of Theorem 2.5. Let $Z_{L,V}^{(p)}$ denote the integral of $Z_{L,(\kappa_b)}$ with respect to the *a priori* measure $\varrho_L(d\kappa)$ with parameter p and let $Z_{L,V}^{*,(p)}$ denote the analogous quantity for $Z_{L,(\kappa_b)}^*$. The arguments leading up to (5.22) then yield

$$Z_{L,V}^{*,(p_\star)} = Z_{L,V}^{(p)} (2\pi L^2) (p_\star \sqrt{\kappa_0} + (1-p_\star) \sqrt{\kappa_D})^{|\mathbb{B}_L|} \quad (5.35)$$

whenever p_\star is dual to p in the sense of (5.21). Thus, using $F(p)$ to denote the limit in (5.23) with the negative sign, we have

$$F(p_\star) = F(p) - 2 \log(p_\star \sqrt{\kappa_0} + (1-p_\star) \sqrt{\kappa_D}). \quad (5.36)$$

Now, as a glance at the proof of Theorem 2.4 reveals, the value p_t is defined as the unique point where the derivative of $F(p)$, which at the continuity points of $p \mapsto \chi(p)$ is simply $F'(p) = 2\chi(p) - 1$, jumps from values near -1 to values near $+1$. Eq. (5.36) then forces the jump to occur at the self-dual point $p_\star = p$. In light of (5.21), this proves (2.15). \square

Acknowledgments

The research of M.B. was supported by the NSF grant DMS-0505356 and that of R.K. by the grants GAČR 201/03/0478, MSM 0021620845, and the Max Planck Institute for Mathematics in the Sciences, Leipzig. The authors are grateful to Scott Sheffield for discussions that ultimately led to the consideration of the model (1.6), and for valuable advice how to establish the zero-tilt property of the coexisting states. Discussions with Jean-Dominique Deuschel helped us understand the problems described in Remark 4.9.

References

1. J.R. Baxter, *Exactly Solved Models in Statistical Mechanics*, Reprint of the 1982 original. Academic Press, Inc., London, 1989.
2. H. van Beijeren, *Exactly solvable model for the roughening transition of a crystal surface*, Phys. Rev. Lett. **38** (1977) 993–996.
3. M. Biskup, C. Borgs, J.T. Chayes, and R. Kotecký, *Gibbs states of graphical representations of the Potts model with external fields*, J. Math. Phys. **41** (2000) 1170–1210.
4. M. Biskup, L. Chayes and S.A. Kivelson, *Order by disorder; without order, in a two-dimensional $O(2)$ antiferromagnet*, Ann. Henri Poincaré **5** (2004), no. 6, 1181–1205.
5. M. Biskup, L. Chayes and Z. Nussinov, *Orbital ordering in transition-metal compounds: I. The 120-degree model*, Commun. Math. Phys. **255** (2005) 253–292.
6. M. Biskup, L. Chayes and S. Starr, *Quantum spin systems at positive temperatures*, Commun. Math. Phys. (to appear).
7. M. Biskup and R. Kotecký, *Forbidden gap argument for phase transitions proved by means of chessboard estimates*, Commun. Math. Phys. (to appear).
8. L. Chayes and K. Shtengel, *Critical behavior for 2D uniform and disordered ferromagnets at self-dual points*, Commun. Math. Phys. **204** (1999), no. 2, 353–366.
9. R.L. Dobrushin and S.B. Shlosman, *Phases corresponding to minima of the local energy*, Selecta Math. Soviet. **1** (1981), no. 4, 317–338.
10. R.G. Edwards and A.D. Sokal, *Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm*, Phys. Rev. D **38** (1988) 2009–2012.
11. A.C.D. van Enter and S.B. Shlosman, *Provable first-order transitions for nonlinear vector and gauge models with continuous symmetries*, Commun. Math. Phys. **255** (2005) 21–32.
12. C.M. Fortuin and P.W. Kasteleyn, *On the random cluster model I. Introduction and relation to other models*, Physica **57** (1972) 536–564.
13. J. Fröhlich, R. Israel, E.H. Lieb and B. Simon, *Phase transitions and reflection positivity. I. General theory and long-range lattice models*, Commun. Math. Phys. **62** (1978), no. 1, 1–34.
14. J. Fröhlich, R. Israel, E.H. Lieb and B. Simon, *Phase transitions and reflection positivity. II. Lattice systems with short-range and Coulomb interactions*, J. Statist. Phys. **22** (1980), no. 3, 297–347.
15. J. Fröhlich and E.H. Lieb, *Phase transitions in anisotropic lattice spin systems*, Commun. Math. Phys. **60** (1978), no. 3, 233–267.
16. T. Funaki, *Stochastic Interface Models*, Lecture Notes for the International Probability School at Saint-Flour, 2003
17. T. Funaki and H. Spohn, *Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model*, Commun. Math. Phys. **185** (1997) 1–36.
18. G.R. Grimmett, *The stochastic random cluster process and the uniqueness of random cluster measures*, Ann. Probab. **23** (1995) 1461–1510.
19. G. Giacomin, S. Olla and H. Spohn, *Equilibrium fluctuations for $\nabla\phi$ interface model*, Ann. Probab. **29** (2001) 1138–1172.

20. R. Kotecký and S.B. Shlosman, *First-order phase transitions in large entropy lattice models*, Commun. Math. Phys. **83** (1982), no. 4, 493–515.
21. R. Kotecký and S.B. Shlosman, *Existence of first-order transitions for Potts models*, In: S. Albeverio, Ph. Combe, M. Sirigue-Collins (eds.), Proc. of the International Workshop—Stochastic Processes in Quantum Theory and Statistical Physics, Lecture Notes in Physics, vol. 173, pp. 248–253, Springer-Verlag, Berlin, 1982.
22. E.H. Lieb, *Residual entropy of square ice*, Phys. Rev. **162** (1967) 162–172.
23. S. Sheffield, *Random surfaces: large deviations principles and gradient Gibbs measure classifications*, PhD thesis, Stanford Univ., 2003; arXiv:math.PR/0304049.
24. Y. Velenik, *Localization and delocalization of random interfaces*, lecture notes for a minicourse at the meeting “Topics in Random Interfaces and Directed Polymers,” Leipzig 2005; arxiv: math.PR/0509695.
25. M. Zahradník, *Contour methods and Pirogov-Sinai theory for continuous spin lattice models*, In: On Dobrushin’s way. R. A. Minlos (ed.), pp. 197–220, Amer. Math. Soc. Transl. **198**, 2000.