

# True nature of long-range order in a plaquette orbital model

Marek Biskup<sup>1,2</sup> and Roman Kotecký<sup>3,4</sup>

<sup>1</sup>Department of Mathematics, UCLA, Los Angeles, CA 90024, U.S.A.

<sup>2</sup>School of Economics, University of South Bohemia, 370 05 České Budějovice, Czech Rep.

<sup>3</sup>Center for Theoretical Study, Charles University, 110 00 Praha 1, Czech Republic

<sup>4</sup>Mathematics Department, University of Warwick, Coventry, CV4 7AL, United Kingdom

E-mail: biskup@math.ucla.edu, R.Kotecky@warwick.ac.uk

**Abstract.** We analyze the classical version of a plaquette orbital model that was recently introduced and studied numerically by S. Wenzel and W. Janke. In this model, edges of the square lattice are partitioned into  $x$  and  $z$ -types that alternate along both coordinate directions and thus arrange into a checkerboard pattern of  $x$  and  $z$ -plaquettes; classical  $O(2)$ -spins are then coupled ferromagnetically via their first components over the  $x$ -edges and via their second components over the  $z$ -edges. We prove from first principles that, at sufficiently low temperatures, the model exhibits orientational long-range order (OLRO) in one of the two principal lattice directions. Magnetic order is precluded by the underlying symmetries. A similar set of results is inferred also for quantum systems with large spin although the spin- $1/2$  instance currently seems beyond the reach of rigorous methods. We point out that the Néel order in the plaquette energy distribution observed in numerical simulations is an artefact of the OLRO and a judicious choice of the plaquette energies. In particular, this order seems to disappear when the plaquette energies are adjusted to vanish at the ground-state level. We also discuss the specific role of the underlying symmetries in Wenzel and Janke's simulations and propose an enhanced method of numerical sampling that could in principle significantly increase the speed of convergence.

PACS numbers: 64.60.Bd – General theory of phase transitions, 05.30.Rt – Quantum phase transitions, 75.30.Ds – Spin waves, 75.47.Gk – Colossal magnetoresistance

AMS classification scheme numbers: 82B26, 82B10

*Keywords:* Rigorous results in statistical mechanics, Classical phase transitions, Quantum phase transitions, Classical Monte Carlo simulations, Sampling algorithms and rapid mixing

To appear in: *Journal of Statistical Mechanics: Theory and Experiment*

## 1. Introduction

The physics of transition-metal compounds [20] and of the underlying effects such as colossal magnetoresistance [8] has recently spawned a number of spin models of high theoretical and practical interest [26, 36]. Among the common features of these models is that (1) the degrees of freedom are represented by quantum or classical spins residing at the vertices of a regular lattice, typically, the square or cubic lattice, (2) the interactions are nearest neighbor and ferromagnetic but (3) only certain components — or projections — of the spins are coupled over each edge of the lattice. The spin variables actually represent effective degrees of freedom (pseudospins); typically, occupation characteristics of a partially filled atom orbital. The interaction is also effective and it is arrived at by considerations of crystal fields mediated by interlaced atoms [25], or by appealing to Jahn-Teller distortions [22] (or both methods, via different routes, [39]). Systems of a similar nature have sprung up independently in the field of topological quantum computation, e.g., the Kitaev model [28, 27], and so besides practical incentives to develop a theory for the behavior of these models, there are also strong theoretical reasons to understand their possible technical implementations [15, 23].

The definition of the aforementioned class of models starts by partitioning all edges of the lattice into families indexed by some  $\alpha$ ; a generic edge in the  $\alpha$ -th family is then denoted by  $\langle \mathbf{r}, \mathbf{r}' \rangle_\alpha$ . The Hamiltonian invariably takes the form

$$\mathcal{H} := - \sum_{\alpha} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle_\alpha} J_{\alpha} S_{\mathbf{r}}^{(\alpha)} S_{\mathbf{r}'}^{(\alpha)} \quad (1.1)$$

with positive coupling constants,  $J_{\alpha} > 0$ . The explicit meaning of the projections  $S_{\mathbf{r}}^{(\alpha)}$  is then a matter of what specific model one wishes to consider.

Two examples of interest have been studied earlier: the *orbital compass model* (e.g., [24, 31, 13, 33, 34, 9]), where  $S_{\mathbf{r}}^{(\alpha)}$ ,  $\alpha = 1, 2, 3$ , are the corresponding Cartesian components of the quantum spin and  $\langle \mathbf{r}, \mathbf{r}' \rangle_\alpha$  is an edge in the  $\alpha$ -th lattice direction, and the *120-degree model* (e.g., [32, 4, 5]), where the meaning of the edge  $\langle \mathbf{r}, \mathbf{r}' \rangle_\alpha$  is preserved but  $S_{\mathbf{r}}^{(\alpha)}$ ,  $\alpha = 1, 2, 3$ , now denotes the projections of the (three-component) spin  $\mathbf{S}_{\mathbf{r}}$  onto the vectors

$$\mathbf{v}_1 := (1, 0), \quad \mathbf{v}_2 := \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \mathbf{v}_3 := \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right), \quad (1.2)$$

i.e.,  $S_{\mathbf{r}}^{(\alpha)} := \mathbf{v}_\alpha \cdot \mathbf{S}_{\mathbf{r}}$ . Kitaev's model is defined similarly to the orbital compass model but the underlying graph is the honeycomb lattice.

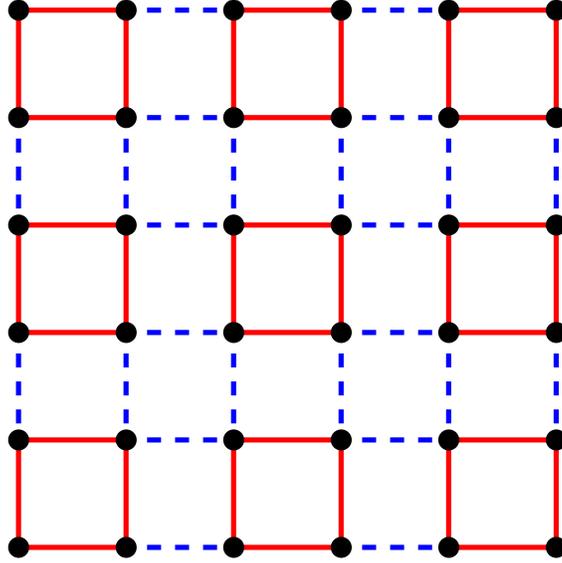
As usual, all models in the above class have a natural quantum version, where  $\mathbf{S}_{\mathbf{r}}$  is a three-component spin operator — with a distinct irreducible matrix representation for each non-negative half integer — and a classical version, where  $\mathbf{S}_{\mathbf{r}}$  is a vector *a priori* uniformly distributed on the unit sphere in  $\mathbb{R}^N$  (i.e., an  $O(N)$ -spin).

## 1.1. The plaquette orbital model

Recently, an interesting variant of the orbital compass model has been proposed and studied by Wenzel and Janke [37]. Their model, which they termed the *plaquette orbital model* (POM), is most naturally defined over the square lattice  $\mathbb{Z}^2$ , although generalizations to higher dimensions are straightforward. The index  $\alpha$  takes only two values,  $\alpha = 1, 2$ , and the spin projections are defined as follows:

$$S_r^{(1)} := S_r^x \quad \text{and} \quad S_r^{(2)} := S_r^z, \quad (1.3)$$

where, in agreement with Wenzel and Janke's notation,  $(S_r^x, S_r^z)$  denote the Cartesian components of the vector-valued  $O(2)$ -spin  $S_r$  in the classical version while, in the quantum version, it denotes the corresponding pair of operators for the quantum spin.



**Figure 1.** The configuration underlying the definition of the plaquette orbital model. Here the  $x$ -components of the spins are coupled over the red (solid) edges and the  $z$ -components are coupled over the blue (dashed) edges.

What distinguishes this model from the previously discussed counterparts is the partitioning of the lattice: edges are designated as  $x$  and  $z$ -types in an alternating fashion along each line of sites (copy of  $\mathbb{Z}$ ) inside  $\mathbb{Z}^2$  so that all of the lattice faces (plaquettes) with the lower-left corner on the even sublattice of  $\mathbb{Z}^2$  contain only edges of one type. Thus one gets the Hamiltonian in accord with the general form in Eq. (1.1) with the edges of two types arranged into to a checkerboard pattern as in Fig. 1,

$$\mathcal{H} := -J_1 \sum_{\langle r, r' \rangle_x} S_r^x S_{r'}^x - J_2 \sum_{\langle r, r' \rangle_z} S_r^z S_{r'}^z, \quad (1.4)$$

with the  $x$ -components of the spins coupled over the  $x$ -edges and  $z$ -components over the  $z$ -edges. The signs of the coupling constants are immaterial as they can always be absorbed into a sign change of the corresponding component on one of the sublattices.

In [37] Wenzel and Janke studied the POM numerically both in its classical and quantum version. For an order parameter they chose the plaquette energy,

$$\mathcal{E}_r := -\left\{ S_r^{\alpha_1} S_{r+e_1}^{\alpha_1} + S_r^{\alpha_2} S_{r+e_2}^{\alpha_2} + S_{r+e_1}^{\alpha_3} S_{r+e_1+e_2}^{\alpha_3} + S_{r+e_2}^{\alpha_4} S_{r+e_1+e_2}^{\alpha_4} \right\}, \quad (1.5)$$

indexed by the position of the lower-left corner of the plaquette, where  $e_1$  and  $e_2$  are the unit vectors in the first and second coordinate direction and  $\alpha_1, \dots, \alpha_4$  are either  $x$  or  $z$  depending on whether the corresponding edge is an  $x$  or  $z$ -type, respectively. The simulations of [37] indicated a clear onset of Néel order in the plaquette energies at low temperatures for both the classical two-component ( $O(2)$ -spin) model, and the quantum, spin- $1/2$  model. Explicitly, in one such state the plaquette energies on the  $x$ -plaquettes are low but those on the  $z$ -plaquettes are generally high and in another state the roles of  $x$  and  $z$ -plaquettes are interchanged. (The mixed-type plaquettes end up with an intermediate energy in both circumstances.) However, no specific conclusion was attempted for other possible ordering mechanisms (magnetic order, orientational order, etc).

## 1.2. Goals and outline

In the present paper we wish to address the true nature of the phase transition in the POM by means of rigorous mathematical (i.e., analytic) methods that draw from earlier work on models of this kind [32, 4, 1, 5]. We will predominantly focus on the classical  $O(2)$ -spin version of the model as the quantum spin- $1/2$  version poses technical difficulties that we do not yet know how to overcome. Nonetheless, thanks to the general theory [5], the conclusions for the classical system permit straightforward extensions to quantum systems once the magnitude of the quantum spin (i.e., the quantity  $\mathcal{S}$  such that  $\hat{S}_r \cdot \hat{S}_r = \mathcal{S}(\mathcal{S} + 1)$ ) is sufficiently large compared to the inverse temperature squared.

A key conceptual difference with the approach of Wenzel and Janke [37] is that instead of going via plaquette energies, we directly attempt (and succeed in) proving *orientational* long-range order (ORLO) of the spins in one of the two principal lattice directions. We then argue that the Néel ordering in plaquette energies found in [37] is actually an *artefact* of this ORLO: indeed, it is a direct consequence of the alignment of the spins along coordinate axes, the 2-periodicity of the interaction and the fact that the plaquette energies were not normalized to vanish in the ground states. In fact, with such normalization the Néel order seems to disappear altogether.

The remainder of this paper is organized as follows: In Sect. 2 we will discuss the ground states of the classical Hamiltonian and then state our principal results concerning the ORLO in the pure and diluted systems. In Sect. 3-5 we provide full and reasonably self-contained

mathematical proofs of these results. In Sect. 6 we discuss connections with the numerical findings and propose a potentially more efficient way to obtain samples of equilibrium configurations in this model.

## 2. Rigorous results

### 2.1. Ground states

Our discussion of the results opens up with the description of the ground states. We will focus on the situation in finite volumes with periodic boundary conditions. Specifically, let  $\mathbb{T}_N$  be the  $N \times N$  torus — obtained by periodizing the square  $\{0, \dots, N-1\} \times \{0, \dots, N-1\}$  — and assume that  $N$  is even to reflect the natural period-2 nature of the interaction. Let  $\mathcal{H}_N$  denote the Hamiltonian on  $\mathbb{T}_N$  which we define by (1.4) with the edges  $\langle \mathbf{r}, \mathbf{r}' \rangle_\alpha$  restricted to nearest neighbor pairs (of the proper type) on  $\mathbb{T}_N$ . As is common, we call a configuration  $\mathbf{S} = (S_{\mathbf{r}})_{\mathbf{r} \in \mathbb{T}_N}$  a *ground state* of  $\mathcal{H}_N$  if

$$\mathcal{H}_N(\mathbf{S}) = \min_{\mathbf{S}'} \mathcal{H}_N(\mathbf{S}'). \quad (2.1)$$

Here we note that the minimum is attained — and a ground state exists — by the sheer fact that  $\mathcal{H}_N$  is a bounded and continuous function on a compact metric space. The issue is how many ground states there are and how they can be concisely described.

A common feature of all models of the type (1.1) is abundance of symmetries with respect to simultaneous flips of (specific) components of the spins. An indisputable advantage of the POM over the other systems is that here the flips can be made locally. Explicitly, for  $\mathbf{r}$  with both coordinates even, i.e., the site designating a lower left corner of an  $x$ -plaquette, let  $\boldsymbol{\varphi}_{\mathbf{r}}(\mathbf{S})$  be the configuration defined by

$$[\boldsymbol{\varphi}_{\mathbf{r}}(\mathbf{S})]_{\mathbf{r}'}^x := \begin{cases} -S_{\mathbf{r}'}^x & \text{for } \mathbf{r}' = \mathbf{r}, \mathbf{r} + \mathbf{e}_1, \mathbf{r} + \mathbf{e}_2, \mathbf{r} + \mathbf{e}_1 + \mathbf{e}_2, \\ S_{\mathbf{r}'}^x & \text{otherwise,} \end{cases} \quad (2.2)$$

and

$$[\boldsymbol{\varphi}_{\mathbf{r}}(\mathbf{S})]_{\mathbf{r}'}^z := S_{\mathbf{r}'}^z. \quad (2.3)$$

For  $\mathbf{r}$  with both coordinates odd (lower left corner of a  $z$ -plaquette), the map  $\boldsymbol{\varphi}_{\mathbf{r}}$  is defined in a similar manner — with the  $z$ -components reflected instead of the  $x$ -components. A moment's thought now shows that

$$\mathcal{H}_N(\boldsymbol{\varphi}_{\mathbf{r}}(\mathbf{S})) = \mathcal{H}_N(\mathbf{S}), \quad (2.4)$$

i.e.,  $\mathbf{S} \mapsto \boldsymbol{\varphi}_{\mathbf{r}}(\mathbf{S})$  is a symmetry of the model.

**Theorem 2.1** *Suppose  $J_1 = J_2 > 0$ . Then every ground state of  $\mathcal{H}_N$  can be obtained from a constant configuration,  $\mathbf{S}_r \equiv \mathbf{e}$ , for some unit vector  $\mathbf{e} \in \mathbb{R}^2$ , by successive applications of a subset of the maps  $(\varphi_r)$ .*

*If  $J_1 > J_2 > 0$ , then all ground states arise (via applications of  $(\varphi_r)$ ) from  $\mathbf{S}_r \equiv \mathbf{e}_1$ , while if  $J_2 > J_1 > 0$ , then all ground states arise (again, via  $(\varphi_r)$ ) from  $\mathbf{S}_r \equiv \mathbf{e}_2$ .*

This statement is a precursor of the orientational LRO that we will establish for low (but non-zero) temperatures. The key problem there will be the fact that the ground-state degeneracy significantly increases at the symmetry point  $J_1 = J_2$ . This appears to be a common feature for all models covered by the equation (1.1).

We also remark that in infinite volume (i.e., the model defined on all of  $\mathbb{Z}^2$ ), the structure of ground states is considerably more complicated. (In infinite volume, a ground state is a configuration whose any local change will result in a non-negative change of energy.) It is still true that any configuration obtained from constant configurations by means of the maps  $(\varphi_r)$  is a ground state. However, further ground states can be constructed by imposing linear (or other) interfaces.

## 2.2. Orientational order

We proceed to discuss our results for positive temperatures. Let  $\beta := \frac{1}{k_B T}$  denote the inverse temperature. In the canonical ensemble, the spins on  $\mathbb{T}_N$  are distributed according to the Gibbs measure  $\mu_{N,\beta}$  that is defined by

$$\mu_{N,\beta}(\mathrm{d}\mathbf{S}) := \frac{e^{-\beta\mathcal{H}_N(\mathbf{S})}}{Z_{N,\beta}} \prod_{r \in \mathbb{T}_N} \nu(\mathrm{d}\mathbf{S}_r), \quad (2.5)$$

where  $\nu$  denotes the uniform (Haar) measure on the unit circle in  $\mathbb{R}^2$  normalized, for later convenience, to  $\sqrt{2\pi}$ . The normalization constant  $Z_{N,\beta}$  is the partition function.

In order to formulate the existence of a long range order in a mathematically precise way, one often considers Gibbs measures directly in infinite volume — i.e., as measures on infinite configurations  $(\mathbf{S}_r)_{r \in \mathbb{Z}^2}$ . These are defined by means of the so-called DLR-condition [11, 12, 30] stating that the conditional distribution in any finite  $\Lambda \subset \mathbb{Z}^2$  given  $(\mathbf{S}_r)_{r \in \mathbb{Z}^2 \setminus \Lambda}$  takes the above form with  $\mathcal{H}_N$  replaced by the Hamiltonian in  $\Lambda$  under the boundary conditions  $(\mathbf{S}_r)_{r \in \mathbb{Z}^2 \setminus \Lambda}$ .

A standard reference to the corresponding mathematical theory of these measures is Georgii [19]. We will now list the aspects that have a bearing on our problem. It is a standard fact that, for compactly-supported spins, infinite-volume Gibbs measures can be extracted (as weak limits) from sequences of finite-volume, or even torus, Gibbs measures. In particular, estimates on  $\mu_{N,\beta}$  that hold uniformly in  $N$  readily yield corresponding estimates for infinite volume limits of  $\{\mu_{N,\beta}\}$ .

In light of the period-2 nature of the interactions, a (Gibbs) measure  $\mu$  will be called *translation-invariant* if  $\mu(\sigma_\alpha(\mathcal{A})) = \mu(\mathcal{A})$  for any event  $\mathcal{A}$ , where

$$[\sigma_\alpha(\mathbf{S})]_r := \mathbf{S}_{r-2e_\alpha}, \quad \alpha = 1, 2, \quad (2.6)$$

are the shift operators. This measure is said to be *ergodic* if  $\mu(\mathcal{A})$  is either zero or one for any event  $\mathcal{A}$  that is invariant under the translations above, i.e., such that  $\sigma_\alpha^{-1}(\mathcal{A}) = \mathcal{A}$  for  $\alpha = 1, 2$ . Ergodic measures have the distinguished feature that all block averages converge to the corresponding expectations (i.e., expected value represents the average value in sufficiently large boxes) and that a typical sample from these measures thus has a fairly homogeneous structure at sufficiently large spatial scales.

We begin by ruling out magnetic ordering at any parameters of the model.

**Theorem 2.2** *Let  $J_1, J_2 \geq 0$  and  $\beta \geq 0$ . Then*

$$\mathbb{E}_\mu(\mathbf{S}_r) = 0 \quad (2.7)$$

for all infinite-volume Gibbs states  $\mu$ .

The absence of magnetic order is of course a consequence of the symmetry  $\mathbf{S} \mapsto \boldsymbol{\varphi}_r(\mathbf{S})$ . This result is prototypical for all models of the kind (1.1). It does not, however, exclude the existence of an orientational long-range order.

**Theorem 2.3** *Suppose  $J_1 = J_2 > 0$ . Then for each  $\delta > 0$ , there is  $\beta_0 = \beta_0(\delta)$  and a sequence  $(\varepsilon_N)$  with  $\varepsilon_N \rightarrow 0$  such that for each  $\beta \geq \beta_0(\delta)$ ,*

$$\mu_{N,\beta} \left( \sum_{r \in \mathbb{T}_N} [S_r^\alpha]^2 \geq |\mathbb{T}_N|(1 - \delta) \right) \geq \frac{1}{2} - \varepsilon_N, \quad \alpha = x, z. \quad (2.8)$$

In particular, for each  $\beta \geq \beta_0(\delta)$ , there exist two translation-invariant, ergodic infinite-volume Gibbs measures,  $\mu_\beta^x$  and  $\mu_\beta^z$ , such that

$$\mathbb{E}_{\mu_\beta^\alpha}([S_r^\alpha]^2) \geq 1 - \delta, \quad \alpha = x, z. \quad (2.9)$$

In addition, for any  $\beta \geq \beta_0(\delta)$  and any translation-invariant, ergodic infinite-volume Gibbs state  $\mu$  at inverse temperature  $\beta$ ,

$$\mathbb{E}_\mu([S_r^\alpha]^2) \in [0, \delta] \cup [1 - \delta, 1], \quad \alpha = x, z. \quad (2.10)$$

Notice that (2.8) states that typical configurations sampled from  $\mu_{N,\beta}$  on a large torus have a majority of all spins aligned either along direction  $\pm e_1$  or direction  $\pm e_2$  (with both orientations equally present thanks to the symmetries (2.4)). The inequality (2.9) is a version of this fact in infinite volume and it manifestly demonstrates the occurrence of a phase transition. The statement (2.10) in turn implies that only the ground states oriented along

the principal axes are stable under thermal perturbations. The infinite degeneracy at the symmetry point for  $\beta = \infty$  is thus reduced to a two-fold degeneracy† once  $\beta < \infty$ . Away from the symmetry point we have the following:

**Theorem 2.4** *If  $J_1 > J_2 > 0$  and  $\beta \geq \beta_0(\delta)$ , then in all translation-invariant, ergodic infinite-volume Gibbs states  $\mu$ ,*

$$\mathbb{E}_\mu([S_r^x]^2) \geq 1 - \delta, \quad \text{for all } \mathbf{r} \in \mathbb{Z}^2. \quad (2.11)$$

*Similarly for  $\mathbb{E}_\mu([S_r^z]^2)$  when  $J_2 > J_1 > 0$  and  $\beta \geq \beta_0(\delta)$ .*

The above results imply the existence of ORLO throughout the part of the quadrant in  $(J_1, J_2)$ -plane bounded away from zero.

### 2.3. Quantum systems

While our methods currently seem unable to treat the quantum spin- $1/2$  version of POM that was studied numerically in [37], the control of the classical system can be extended to the quantum problem with large spin. This is achieved essentially by plugging into the main result of [5]. We proceed to introduce the technical aspects of the quantum POM that are necessary to state the relevant theorem.

In the quantum POM, the spins  $\hat{S}_r$  are three-component operators  $\hat{S}_r = (\hat{S}_r^x, \hat{S}_r^y, \hat{S}_r^z)$  satisfying the commutation rules of the Lie algebra  $\mathfrak{su}(2)$ ,

$$[\hat{S}_r^j, \hat{S}_{r'}^k] = 2i\hat{S}_r^\ell \delta_{r,r'} \quad (2.12)$$

for any cyclic permutation  $(j, k, \ell)$  of  $(x, y, z)$ . We will work with the  $(2\mathcal{S} + 1)$ -dimensional irreducible representation of  $\mathfrak{su}(2)$ , where  $\mathcal{S} \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \dots\}$ . This representation is best discussed by means of the spin raising and lowering operators

$$\hat{S}_r^\pm := \hat{S}_r^x \pm i\hat{S}_r^y. \quad (2.13)$$

The Hilbert space is the linear span of vectors

$$|\dots M_r \dots\rangle := \bigotimes_r |M_r\rangle, \quad M_r = -\mathcal{S}, -\mathcal{S} + 1, \dots, \mathcal{S} - 1, \mathcal{S}, \quad (2.14)$$

on which the operators  $\hat{S}_r^z, \hat{S}_r^+, \hat{S}_r^-$  act as follows

$$\begin{aligned} \hat{S}_r^z |\dots M_r \dots\rangle &= M_r |\dots M_r \dots\rangle, \\ \hat{S}_r^+ |\dots M_r \dots\rangle &= \sqrt{\mathcal{S}(\mathcal{S} + 1) - M_r(M_r + 1)} |\dots M_r + 1 \dots\rangle, \\ \hat{S}_r^- |\dots M_r \dots\rangle &= \sqrt{\mathcal{S}(\mathcal{S} + 1) - M_r(M_r - 1)} |\dots M_r - 1 \dots\rangle. \end{aligned} \quad (2.15)$$

† A word of caution: Strictly speaking, the rigorous statement does not rule out the existence of additional extremal translation-invariant Gibbs states apart from those above. However, an overwhelming majority of the spins in a typical configuration in any such state will be close either to  $\pm e_1$  or to  $\pm e_2$ .

The Hamiltonian  $\hat{\mathcal{H}}_N$  on the torus  $\mathbb{T}_N$  is then the operator

$$\hat{\mathcal{H}}_N := -J_1 \mathcal{S}^{-2} \sum_{\langle r, r' \rangle_x} \hat{S}_r^x \hat{S}_{r'}^x - J_2 \mathcal{S}^{-2} \sum_{\langle r, r' \rangle_z} \hat{S}_r^z \hat{S}_{r'}^z. \quad (2.16)$$

The  $y$ -component of the spin does not enter the interaction. The scaling by  $\mathcal{S}^{-2}$  ensures that the Hamiltonian is, for each  $N$ , bounded uniformly in  $\mathcal{S} \geq 1/2$ .

The thermodynamical equilibrium is described by means of a linear functional  $\langle \cdot \rangle_{N, \beta}$  on the algebra  $\mathcal{A}_N$  of all bounded operators generated (via the spectral theorem) by the operators  $\hat{S}_r^z, \hat{S}_r^+, \hat{S}_r^-, \mathbf{r} \in \mathbb{T}_N$ . Explicitly,

$$\langle \hat{A} \rangle_{N, \beta} = \frac{\text{Tr}(\hat{A} e^{-\beta \hat{\mathcal{H}}_N})}{\text{Tr}(e^{-\beta \hat{\mathcal{H}}_N})}, \quad \hat{A} \in \mathcal{A}_N. \quad (2.17)$$

Our main result for the quantum system is now as follows:

**Theorem 2.5** *Suppose  $J_1 = J_2 > 0$ . Then for each  $\delta > 0$ , there are constants  $c = c(\delta) > 0$  and  $\beta_0 = \beta_0(\delta) < \infty$  such that for each  $\beta \geq \beta_0(\delta)$  and  $\mathcal{S} \geq c\beta^2$ , we have*

$$\mathcal{S}^{-4} \left\langle [\hat{S}_r^\alpha]^2 [\hat{S}_{r'}^\alpha]^2 \right\rangle_{N, \beta} \geq \frac{1}{2} - \delta, \quad \alpha = x, z, \quad (2.18)$$

while

$$\mathcal{S}^{-4} \left\langle [\hat{S}_r^y]^2 [\hat{S}_{r'}^y]^2 \right\rangle_{N, \beta} < \delta, \quad (2.19)$$

and

$$\mathcal{S}^{-4} \left| \left\langle [\hat{S}_r^x]^2 [\hat{S}_{r'}^z]^2 \right\rangle_{N, \beta} \right| < \delta, \quad (2.20)$$

uniformly in  $\mathbf{r}, \mathbf{r}' \in \mathbb{T}_N$  provided  $N$  is sufficiently large.

Note that the bound (2.20) yields the same bound on  $\langle [\hat{S}_r^z]^2 [\hat{S}_{r'}^x]^2 \rangle_{N, \beta}$  because the identity  $|\langle \hat{A} \hat{B} \rangle_{N, \beta}| = |\langle \hat{B} \hat{A} \rangle_{N, \beta}|$  holds for all self-adjoint operators  $\hat{A}, \hat{B}$ . It is standard that validity of such bounds implies nonanalyticity of the free energy in the appropriate conjugate variables. In our case, this will be the function

$$f(h) := \lim_{N \rightarrow \infty} \frac{1}{N^d} \log \text{Tr} \left( \exp \left\{ -\beta \hat{\mathcal{H}}_N + h \mathcal{S}^{-2} \sum_{\mathbf{r} \in \mathbb{T}_N} ([S_r^x]^2 - [S_r^z]^2) \right\} \right), \quad (2.21)$$

where the “external field”  $h$  couples to the natural order parameter  $[S_r^x]^2 - [S_r^z]^2$ . Based on this fact, we extract the corresponding result for the asymmetric situations as well:

**Theorem 2.6** Suppose  $J_1 > J_2 > 0$  and let  $c = c(\delta) > 0$  and  $\beta_0 = \beta_0(\delta) < \infty$  be as above. Then for each  $\delta$ , each  $\beta \geq \beta_0(\delta)$  and  $\mathcal{S} \geq c\beta^2$ , we have

$$\mathcal{S}^{-4} \left\langle [\hat{S}_r^x]^2 [\hat{S}_{r'}^x]^2 \right\rangle_{N,\beta} \geq 1 - 2\delta, \quad (2.22)$$

while

$$\mathcal{S}^{-4} \left\langle [\hat{S}_r^\alpha]^2 [\hat{S}_{r'}^\alpha]^2 \right\rangle_{N,\beta} < \delta, \quad \alpha = y, z, \quad (2.23)$$

uniformly in  $\mathbf{r}, \mathbf{r}' \in \mathbb{T}_N$  for  $N$  is sufficiently large. A similar result holds for  $J_2 > J_1 > 0$  with indices  $x$  and  $z$  interchanged.

We remark that the reason for assuming  $\mathcal{S} \geq c\beta^2$  is that, in the underlying proof, we use thermal fluctuations to dominate the quantum fluctuations (namely, the effects resulting from the non-commutative nature of the relevant variables).

### 3. Ground states

The goal of this section is to prove our characterization of the ground states on  $\mathbb{T}_N$ . As a starting point we note the following rewrite of the energy function:

**Lemma 3.1** Let  $\mathcal{H}_N$  denote the torus Hamiltonian. Then for all  $\mathcal{S}$ ,

$$\mathcal{H}_N(\mathcal{S}) = \frac{J_1}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle_x} (\mathcal{S}_r^x - \mathcal{S}_{r'}^x)^2 + \frac{J_2}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle_z} (\mathcal{S}_r^z - \mathcal{S}_{r'}^z)^2 - \sum_{\mathbf{r}} (J_1 [\mathcal{S}_r^x]^2 + J_2 [\mathcal{S}_r^z]^2). \quad (3.1)$$

*Proof.* Every vertex in  $\mathbb{T}_N$  has two  $x$ -edges and two  $z$ -edges coming out of it. Opening up the squares in the first two sums, the ‘‘diagonal’’ terms there are easily checked to exactly cancel the terms in the third sum.  $\square$

*Proof of Theorem 2.1.* First, we claim that

$$\min_{\mathcal{S}} \mathcal{H}_N(\mathcal{S}) = -\max\{J_1, J_2\} |\mathbb{T}_N|. \quad (3.2)$$

The inequality  $\leq$  is seen by taking  $\mathcal{S}_r \equiv \mathbf{e}_1$  or  $\mathcal{S}_r \equiv \mathbf{e}_2$ , depending on whether  $J_1 \geq J_2$  or not, and evaluating  $\mathcal{H}_N(\mathcal{S})$  for this choice. For the opposite bound we use  $J_1, J_2 > 0$  to drop the first two terms in (3.1) and conclude

$$\min_{\mathcal{S}} \mathcal{H}_N(\mathcal{S}) \geq -|\mathbb{T}_N| \max_{\mathcal{S} \in O(2)} (J_1 [\mathcal{S}^x]^2 + J_2 [\mathcal{S}^z]^2) \geq -\max\{J_1, J_2\} |\mathbb{T}_N|. \quad (3.3)$$

This also shows that the minimum is attained only by configurations for which

$$\mathcal{S}_r^\alpha = \mathcal{S}_{r'}^\alpha \quad \text{for the endpoints } \mathbf{r}, \mathbf{r}' \text{ of all } \alpha\text{-bonds } \langle \mathbf{r}, \mathbf{r}' \rangle_\alpha, \alpha = x, z, \quad (3.4)$$

and

$$J_1[S_r^x]^2 + J_2[S_r^z]^2 = \max\{J_1, J_2\} \quad \text{for all } r \in \mathbb{T}_N \quad (3.5)$$

hold true.

Let now  $S$  be a ground state of  $\mathcal{H}_N$ . The above findings guarantee that  $S_r^x$  is constant on any  $x$ -plaquette and  $S_r^z$  is constant on any  $z$ -plaquette. Applications of  $\varphi_r$  to the plaquettes where the corresponding value is negative defines a configuration  $\tilde{S}$  where  $\tilde{S}_r^\alpha \geq 0$  for  $\alpha = x, z$  and all  $r \in \mathbb{T}_N$ . (This is the desired modification of  $S$  by applications of the maps  $(\varphi_r)$ .) Let now  $r$  be a lower-left vertex of a  $z$ -plaquette and let  $r' := r - e_1$  be its neighbor to the left. Since  $\tilde{S}$  is also a ground state and the edge  $\langle r, r' \rangle$  is an  $x$ -edge, we have  $\tilde{S}_r^x = \tilde{S}_{r'}^x$ . However, the components  $\tilde{S}_r^z, \tilde{S}_{r'}^z$  are both non-negative and since one component of a spin determines the other up to a sign, we also have  $\tilde{S}_r^z = \tilde{S}_{r'}^z$ . It follows that  $\tilde{S}_r = \tilde{S}_{r'}$ . Proceeding similarly for all pairs of neighbors in  $\mathbb{T}_N$  we conclude that  $\tilde{S}_r = e$  for some unit vector  $e \in \{v \in \mathbb{R}^2: v \cdot e_i \geq 0, i = 1, 2\}$  and all  $r \in \mathbb{T}_N$ .

It remains to determine the set of vectors  $e$  that are admissible at given values of the parameters of the model. We have

$$\mathcal{H}_N(\tilde{S}) = -(J_1(e \cdot e_1)^2 + J_2(e \cdot e_2)^2) |\mathbb{T}_N|. \quad (3.6)$$

Thus, when  $J_1 > J_2$ , we must have  $e = e_1$  while  $e = e_2$  when  $J_2 > J_1$ . At the symmetry point,  $J_1 = J_2$ , any  $e$  will give the same value. We have thus shown that  $S$  is a modification of a constant configuration (namely  $\tilde{S}$ ) of the desired type in all cases of interest.  $\square$

## 4. Technical ingredients

In this section we assemble the technical ingredients needed for the proof of the main theorems concerning the phase transition in the model of interest. The proofs will come in Section 5.

### 4.1. Chessboard estimates

The proof of the positive-temperature part of the results will be based on the technique of *chessboard estimates*, based on *reflection positivity*, whose origins go to the seminal work of Dyson, Fröhlich, Israel, Lieb, Simon and Spencer from the late 1970s. This technique, along with a related *infrared-bound* technology, has proved extremely useful in establishing symmetry-breaking phase transitions in various classical and quantum systems with a continuous symmetry [18, 14, 16, 17], order-disorder transitions in the Potts and related models [29, 10, 3], low-temperature ordering in liquid-crystal models [21, 40]. More recently, this technique has also been used to prove phase transitions in systems with

highly degenerate ground states without an underlying symmetry [4, 1] including gradient fields with a non-convex interaction [7]. The theoretical foundations of this technique are well developed already in the original papers [16, 17]; the more recent developments are summarized in the lecture notes [2].

Consider the model with the Hamiltonian  $\mathcal{H}_N$  on the torus  $\mathbb{T}_N$  with  $N$  even. Consider an even integer  $B$  that divides  $N$  and let  $\Lambda_B := \{0, 1, \dots, B-1\} \times \{0, 1, \dots, B-1\}$  denote the block of  $B \times B$  vertices with the lower-left corner at the origin. Consider a regular partitioning of  $\mathbb{T}_N$  into pairwise disjoint translates of  $\Lambda_B$  by vectors from  $\{Br: r \in \mathbb{T}_{N/B}\}$ . For  $r \in \mathbb{T}_{N/B}$ , let  $\vartheta_r$  denote the translation by  $Br$ . On the configuration space,

$$[\vartheta_r(\mathcal{S})]_{r'} = \mathcal{S}_{r'-Br}, \quad r' \in \mathbb{T}_N, r \in \mathbb{T}_{N/B}. \quad (4.1)$$

We call an event  $\mathcal{A}$  a *B-block event* if  $\mathcal{A}$  depends only on  $\{S_r: r \in \Lambda_B\}$  and we use  $\mathcal{F}_B$  to denote the collection of all B-block events.

For each  $\mathcal{A} \in \mathcal{F}_B$  we now define a family  $\{\theta_r(\mathcal{A}): r \in \mathbb{T}_{N/B}\}$  of translations-reflections as follows. First, let  $\mathcal{A}_1$  denote the reflection of  $\mathcal{A}$  through the horizontal mid-line  $\{(y, \frac{B-1}{2}): y \in \mathbb{R}/(N\mathbb{R})\}$  halving the box  $\Lambda_B$ . Similarly, we use  $\mathcal{A}_2$  to denote the reflection of  $\mathcal{A}$  through the vertical mid-line  $\{(x, \frac{B-1}{2}): x \in \mathbb{R}/(N\mathbb{R})\}$  and  $\mathcal{A}_{12}$  to denote the reflection of  $\mathcal{A}$  through both lines (the two reflections commute and so the order in which they are taken is immaterial). Note that  $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_{12} \in \mathcal{F}_B$ . For  $r = (r_1, r_2) \in \mathbb{T}_{N/B}$ , we then set

$$\theta_r(\mathcal{A}) := \begin{cases} \vartheta_r^{-1}(\mathcal{A}), & \text{for } r_1, r_2 \text{ even,} \\ \vartheta_r^{-1}(\mathcal{A}_1), & \text{for } r_1 \text{ odd and } r_2 \text{ even,} \\ \vartheta_r^{-1}(\mathcal{A}_2), & \text{for } r_1 \text{ even and } r_2 \text{ odd,} \\ \vartheta_r^{-1}(\mathcal{A}_{12}), & \text{for } r_1, r_2 \text{ odd.} \end{cases} \quad (4.2)$$

Here  $\vartheta^{-1}(\mathcal{A}) := \{\vartheta_r(\mathcal{S}): \mathcal{S} \in \mathcal{A}\}$ . Notice that  $\theta_r(\mathcal{A})$  thus depends only on the part of the spin configuration in the block  $Br + \Lambda_B$ . We then have:

**Lemma 4.1 (Chessboard estimate)** *Suppose  $J_1, J_2 \geq 0$  and  $\beta \geq 0$ . Then for any events  $\mathcal{A}_1, \dots, \mathcal{A}_m \in \mathcal{F}_B$  and any distinct  $r_1, \dots, r_m \in \mathbb{T}_{N/B}$ ,*

$$\mu_{N,\beta} \left( \bigcap_{i=1}^m \theta_{r_i}(\mathcal{A}_i) \right) \leq \prod_{i=1}^m \left[ \mu_{N,\beta} \left( \bigcap_{r \in \mathbb{T}_{N/B}} \theta_r(\mathcal{A}_i) \right) \right]^{(B/N)^2}. \quad (4.3)$$

The punchline of this result is that the probability of a simultaneous occurrence of several (often undesirable) events on the torus is bounded by the product of the probabilities of events where the individual B-block events have been *disseminated* — using the maps  $\theta_r$  — throughout the entire torus. The latter quantities are often rather explicitly computable.

*Proof of Lemma 4.1.* As already alluded to, the key input for the chessboard estimates is reflection positivity of the interaction. We will now define the necessary concept and check

the validity of this property. Let  $P$  be a plane bisecting a horizontal or vertical line of edges of  $\mathbb{T}_N$ , i.e.,  $P$  is of the form either

$$\{(n + 1/2, y): y \in \mathbb{R}/(N\mathbb{R})\} \cup \{(n + N/2 + 1/2, y): y \in \mathbb{R}/(N\mathbb{R})\}, \quad (4.4)$$

with  $n = 0, 1, \dots, \frac{N}{2} - 1$ , or

$$\{(x, n + 1/2): x \in \mathbb{R}/(N\mathbb{R})\} \cup \{(x, n + N/2 + 1/2): x \in \mathbb{R}/(N\mathbb{R})\}, \quad (4.5)$$

with  $n = 0, 1, \dots, \frac{N}{2} - 1$ . The plane has two components and it splits the torus into a left half  $\mathbb{T}_N^-$  and the right half  $\mathbb{T}_N^+$ . Abusing the notation slightly, let  $\theta_P$  denote the map on the configuration space representing the reflection  $\mathbb{T}_N^+ \leftrightarrow \mathbb{T}_N^-$ .

A sufficient condition for the interaction to be reflection positive is that for each such a plane  $P$ , there are functions  $g$  and  $h = (h_i)$  depending only on  $\{\mathbf{S}_r: r \in \mathbb{T}_N^+\}$  so that

$$-\mathcal{H}_N = g + \theta_P(g) + \sum_i h_i \theta_P(h_i). \quad (4.6)$$

We will now demonstrate that  $\mathcal{H}_N$  is indeed of this form. Let

$$g(\mathbf{S}) := -J_1 \sum_{\substack{\langle r, r' \rangle_x \\ r, r' \in \mathbb{T}_N^+}} S_r^x S_{r'}^x - J_2 \sum_{\substack{\langle r, r' \rangle_z \\ r, r' \in \mathbb{T}_N^+}} S_r^z S_{r'}^z \quad (4.7)$$

and note that  $g$  depends only on the spins in  $\mathbb{T}_N^+$ . The collection of functions  $h$  will be parametrized by the vertices  $r \in \mathbb{T}_N^+$  that have an edge to a vertex  $r' \in \mathbb{T}_N^-$ . We will use  $P^+$  to denote the set of such vertices  $r$ . We set

$$h_r := \begin{cases} \sqrt{J_1} S_r^x, & \text{if } \langle r, r' \rangle \text{ is an } x\text{-edge,} \\ \sqrt{J_2} S_r^z, & \text{if } \langle r, r' \rangle \text{ is a } z\text{-edge,} \end{cases} \quad r \in P^+, \quad (4.8)$$

where  $r'$  stands for the reflection of  $r$  through plane  $P$ . Then  $h_r \theta_P(h_r) = J_\alpha S_r^\alpha S_{r'}^\alpha$ , with  $\alpha$  depending on the type of the edge  $\langle r, r' \rangle$ . A moment's thought then shows that

$$-\mathcal{H}_N = g + \theta_P(g) + \sum_{r \in P^+} h_r \theta_P(h_r) \quad (4.9)$$

and so the interaction is of the desired form. As the *a priori* measure on the spins has a product structure, standard theory (cf [16, Theorem 4.1] or [2, Theorem 5.8]) readily implies the desired claim.  $\square$

## 4.2. Gaussian calculations

Through the use of chessboard estimates, the proof of the phase transition will be reduced to some tedious but explicit computations of multivariable Gaussian integrals. Informally, these can be understood as calculations of spin-wave free energies corresponding to the

spin system at hand. The goal of this section is to carry out these calculations and derive the necessary estimates between actual partition functions and their Gaussian approximations. Throughout we will assume that

$$J_1 = J_2 =: J \quad (4.10)$$

with  $J > 0$ .

For a given unit vector  $\mathbf{e}(\theta) := (\cos \theta, \sin \theta) \in \mathbb{R}^2$  and  $\Delta > 0$ , we define the quantity

$$\mathcal{Z}_N(\theta, \Delta) := e^{-\beta J |\mathbb{T}_N|} \int e^{-\beta \mathcal{H}_N(\mathbf{S})} \prod_{\mathbf{r} \in \mathbb{T}_N} 1_{\{|\mathbf{S}_{\mathbf{r}} - \mathbf{e}(\theta)| < \Delta\}} \prod_{\mathbf{r} \in \mathbb{T}_N} \nu(d\mathbf{S}_{\mathbf{r}}). \quad (4.11)$$

This is the partition function restricted to configurations within  $\Delta$  of a constant configuration pointing in direction of the unit vector  $\mathbf{e}(\theta)$ . Next, we define a function  $\theta \mapsto F(\theta)$  as follows. For each  $\mathbf{k} := (k_1, k_2) \in [-\pi, \pi]^2$ , let us introduce the quantities

$$a_{\pm} := 1 \pm e^{-ik_1}, \quad b_{\pm} := 1 \pm e^{-ik_2}, \quad \text{and} \quad \rho := -\cos(2\theta). \quad (4.12)$$

Consider the matrix

$$M(\mathbf{k}, \theta) := \frac{1}{2} \begin{pmatrix} |a_-|^2 + |b_-|^2 & \rho a_- a_+^* & \rho b_- b_+^* & 0 \\ \rho a_-^* a_+ & |a_+|^2 + |b_-|^2 & 0 & \rho b_- b_+^* \\ \rho b_-^* b_+ & 0 & |a_-|^2 + |b_+|^2 & \rho a_- a_+^* \\ 0 & \rho b_-^* b_+ & \rho a_-^* a_+ & |a_+|^2 + |b_+|^2 \end{pmatrix}. \quad (4.13)$$

We will see in a moment that  $\det M(\mathbf{k}, \theta) \geq 0$  and so we may define

$$F(\theta) := \frac{1}{2} \log(\beta J) + \frac{1}{8} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log \det M(\mathbf{k}, \theta). \quad (4.14)$$

The key facts about the quantity  $F(\theta)$  and its relation to  $\mathcal{Z}_N(\theta, \Delta)$  are the subject of the following two claims:

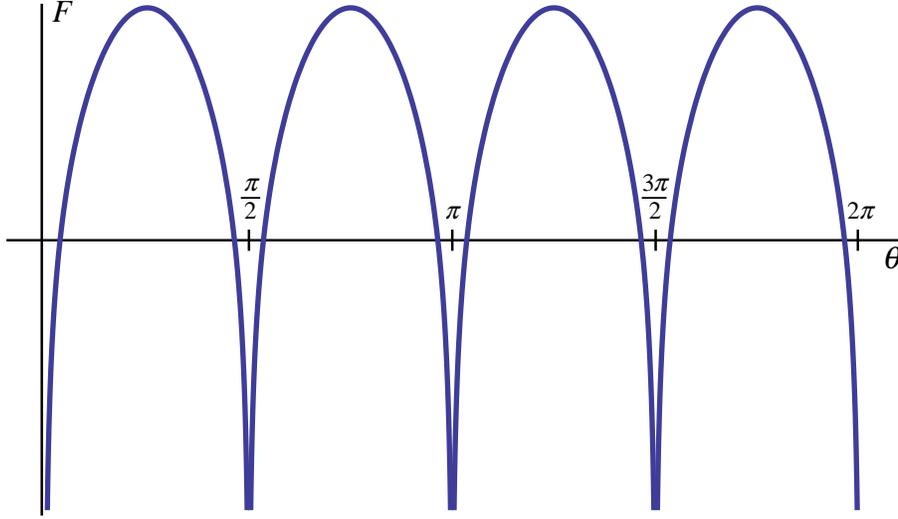
**Proposition 4.2** *For any  $\mathbf{k} := (k_1, k_2) \in [-\pi, \pi]^2$  and any  $\theta \in [0, 2\pi]$  we have*

$$\sin^2(2\theta) \sin^2(k_1) \sin^2(k_2) \leq \det M(\mathbf{k}, \theta) \leq 16 \sin^2(2\theta). \quad (4.15)$$

*In particular,  $F(\theta)$  is finite for all  $\theta$  with  $\sin(2\theta) \neq 0$ . The function  $\theta \mapsto F(\theta)$  is periodic with period  $\pi/2$ , symmetric and continuous on the interval  $(0, \pi/2)$ , increasing on  $(0, \pi/4)$  and decreasing on  $(\pi/4, \pi/2)$ . The infimum of  $F$  is  $-\infty$  and it is achieved exactly at  $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$ . See Fig. 2.*

**Proposition 4.3** *Suppose that  $J_1 = J_2 =: J > 0$ . For each  $\tau > 0$  there are numbers  $\delta > 0$  and  $N_0 < \infty$  such that if*

$$\beta J \Delta^2 > \frac{1}{\delta}, \quad \beta J \Delta^3 < \delta, \quad (4.16)$$



**Figure 2.** The plot of function  $\theta \mapsto F(\theta)$  for  $\theta$  ranging from 0 to  $2\pi$ . The key fact is that  $F$  tends to minus infinity as  $\theta$  tends to multiples of  $\pi/2$ . However, note that  $F$  is relevant for the approximation of  $N^{-2} \log \mathcal{Z}_N(\theta, \Delta)$  to within  $\tau$  only when  $\theta$  avoids the region where  $|\sin(2\theta)| \leq \tau$  — i.e., exactly the vicinity of its singularity points.

then

$$\left| \frac{1}{N^2} \log \mathcal{Z}_N(\theta, \Delta) + F(\theta) \right| < \tau \quad (4.17)$$

holds true whenever  $N \geq N_0$  and  $|\sin(2\theta)| > \tau$ .

We will first focus on Proposition 4.3 because its proof explains the reasons underlying the introduction of the quantity  $F$ . The proof consists of a sequence of approximations followed by a standard diagonalization of a multivariate Gaussian integral. Let us write each  $\mathbf{S}_r$  as

$$\mathbf{S}_r = (\cos(\theta + \vartheta_r), \sin(\theta + \vartheta_r)). \quad (4.18)$$

On the event entering the integral (4.11),  $\vartheta_r = O(\Delta)$ , so if  $\Delta$  is small (which we may assume since (4.16) forces  $\Delta < \delta^2$ ), then  $\boldsymbol{\vartheta} := (\vartheta_r)$  may be regarded as deviation variables. The conversion to a Gaussian integral is performed as follows:

**Lemma 4.4** Suppose  $\Delta < 1$  and define the function

$$G_{N,\theta}(\boldsymbol{\vartheta}) := \frac{1}{2} \sum_{\langle r,r' \rangle_x} (\vartheta_r - \vartheta_{r'})^2 \sin^2(\theta) + \frac{1}{2} \sum_{\langle r,r' \rangle_z} (\vartheta_r - \vartheta_{r'})^2 \cos^2(\theta). \quad (4.19)$$

There exists a constant  $c \in (0, \infty)$  such that for every  $\mathbf{S}$  that is related to  $\boldsymbol{\vartheta}$  via (4.18), with  $\vartheta_r \in (-\pi, \pi]$ , and obeys  $|\mathbf{S}_r - \mathbf{e}(\theta)| < \Delta$  at all  $\mathbf{r} \in \mathbb{T}_N$ ,

$$|\mathcal{H}_N(\mathbf{S}) + J|\mathbb{T}_N| - JG_{N,\theta}(\boldsymbol{\vartheta})| \leq cJ|\mathbb{T}_N|\Delta^3. \quad (4.20)$$

*Proof.* With the restriction  $\vartheta_r \in (-\pi, \pi]$  and  $\Delta < 1$ , the correspondence between  $\mathbf{S}$  and  $\boldsymbol{\vartheta}$  is one-to-one. The restriction  $|\mathbf{S}_r - \mathbf{e}(\theta)| < \Delta$  implies  $\vartheta_r = O(\Delta)$  uniformly in  $r \in \mathbb{T}_N$ . The claim now follows by writing  $\mathcal{H}_N$  in the form (3.1), and noting that

$$S_r^x - S_{r'}^x = (\vartheta_r - \vartheta_{r'}) \sin \theta + O(\Delta^2) \quad (4.21)$$

and similarly for  $S_r^z - S_{r'}^z$ .  $\square$

If we set  $\Delta' := 2 \arcsin(\Delta/2)$  for the maximal angle between  $\mathbf{S}_r$  and  $\mathbf{e}(\theta)$  allowed by the constraints, we thus have

$$\mathcal{Z}_N(\theta, \Delta) = e^{O(\Delta^3)\beta J|\mathbb{T}_N|} \int_{\mathbb{R}^{\mathbb{T}_N}} e^{-\beta J G_{N,\theta}(\boldsymbol{\vartheta})} \chi_{\Delta',N}(\boldsymbol{\vartheta}) \prod_{r \in \mathbb{T}_N} \frac{d\vartheta_r}{\sqrt{2\pi}}, \quad (4.22)$$

where  $\sqrt{2\pi}$  appears due to our normalization convention for  $\nu$  and where

$$\chi_{\Delta',N}(\boldsymbol{\vartheta}) := \prod_{r \in \mathbb{T}_N} 1_{(-\Delta', \Delta')}(\vartheta_r). \quad (4.23)$$

In order to evaluate the expression in (4.22), we will notice the following relation between the function  $G_{N,\theta}$  and the matrices  $M(\mathbf{k}, \theta)$ . Consider the Fourier-reciprocal torus,

$$\mathbb{T}_N^* := \left\{ \frac{2\pi}{N}(n_1, n_2) : n_1, n_2 = 0, \dots, N-1 \right\}, \quad (4.24)$$

and let  $(\mathbb{T}_N^*)_+$  be its first quadrant, i.e., the collection of  $\mathbf{k} = (k_1, k_2) \in \mathbb{T}_N^*$  satisfying  $0 \leq k_1, k_2 < \pi$ . Let  $M(\theta)_{\mathbf{k}, \mathbf{k}'}$  be the  $N^2 \times N^2$  matrix indexed by  $\mathbf{k}, \mathbf{k}' \in \mathbb{T}_N^*$  that is block diagonal and whose entry is zero unless  $\mathbf{k}, \mathbf{k}' \in \{\bar{\mathbf{k}}, \bar{\mathbf{k}} + \pi \mathbf{e}_1, \bar{\mathbf{k}} + \pi \mathbf{e}_2, \bar{\mathbf{k}} + \pi \mathbf{e}_1 + \pi \mathbf{e}_2\}$  for some  $\bar{\mathbf{k}} \in (\mathbb{T}_N^*)_+$ , and whose entries for these four values are collected (in the given order) in the  $4 \times 4$ -matrix  $M(\mathbf{k}, \theta)$  defined above.

**Lemma 4.5** *We have*

$$G_{N,\theta}(\boldsymbol{\vartheta}) = \sum_{r, r' \in \mathbb{T}_N} \widehat{M}(\theta)_{r, r'} \vartheta_r \vartheta_{r'}, \quad (4.25)$$

where

$$\widehat{M}(\theta)_{r, r'} := \frac{1}{|\mathbb{T}_N|} \sum_{\mathbf{k}, \mathbf{k}' \in \mathbb{T}_N^*} M(\theta)_{\mathbf{k}, \mathbf{k}'} e^{i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}}. \quad (4.26)$$

*Proof.* We may write  $G_{N,\theta}$  as

$$G_{N,\theta}(\boldsymbol{\vartheta}) = \sum_{r \in \mathbb{T}_N} \sum_{\mathbf{e} = \mathbf{e}_1, \mathbf{e}_2} J_{r, \mathbf{e}} (\vartheta_r - \vartheta_{r+\mathbf{e}})^2, \quad (4.27)$$

where  $J_{r, \mathbf{e}_1} := \sin^2(\theta)$  for vertices  $r$  with even first coordinate and  $J_{r, \mathbf{e}_1} := \cos^2(\theta)$  for vertices  $r$  with odd first coordinate, and same for  $J_{r, \mathbf{e}_2}$  and the second coordinate of  $r$ . Invoking the (discrete) Fourier representation  $\vartheta_r = |\mathbb{T}_N|^{-1/2} \sum_{\mathbf{k} \in \mathbb{T}_N^*} \widehat{\vartheta}_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{r}}$  we now get

$$G_{N,\theta}(\boldsymbol{\vartheta}) = \sum_{\mathbf{k}, \mathbf{k}' \in \mathbb{T}_N^*} \sum_{\mathbf{e} = \mathbf{e}_1, \mathbf{e}_2} \widehat{\vartheta}_{\mathbf{k}} \widehat{\vartheta}_{\mathbf{k}'}^* (1 - e^{-i\mathbf{k} \cdot \mathbf{e}}) (1 - e^{i\mathbf{k}' \cdot \mathbf{e}}) \widehat{J}_{\mathbf{k}, \mathbf{k}'}(\mathbf{e}), \quad (4.28)$$

where

$$\hat{J}_{k,k'}(\mathbf{e}) := \frac{1}{|\mathbb{T}_N|} \sum_{r \in \mathbb{T}_N} J_{r,e} e^{i(k'-k) \cdot r}. \quad (4.29)$$

Now, since the couplings  $J_{r,e}$  are 2-periodic in direction  $\mathbf{e}$  and translation invariant in the complementary direction,  $\hat{J}_{k,k'}(\mathbf{e})$  will be non-zero only when  $k' = k$  or  $k' = k + \pi\mathbf{e}$ . In these two cases we get

$$\hat{J}_{k,k}(\mathbf{e}) = \frac{1}{2} \quad \text{and} \quad \hat{J}_{k,k+\pi\mathbf{e}}(\mathbf{e}) = -\frac{1}{2} \cos(2\theta). \quad (4.30)$$

Consider a fixed vector  $\bar{\mathbf{k}} \in (\mathbb{T}_N^*)_+$  and consider the restriction of the sum in (4.28) to  $k, k' \in \{\bar{\mathbf{k}}, \bar{\mathbf{k}} + \pi\mathbf{e}_1, \bar{\mathbf{k}} + \pi\mathbf{e}_2, \bar{\mathbf{k}} + \pi\mathbf{e}_1 + \pi\mathbf{e}_2\}$ . Let  $a_{\pm}, b_{\pm}$  denote the quantities in (4.12) for this  $\bar{\mathbf{k}}$ . The quadratic form corresponding to  $\mathbf{e} := \mathbf{e}_1$  is then described by the matrix

$$\frac{1}{2} \begin{pmatrix} |a_-|^2 & \rho a_- a_+^* & 0 & 0 \\ \rho a_-^* a_+ & |a_+|^2 & 0 & 0 \\ 0 & 0 & |a_-|^2 & \rho a_- a_+^* \\ 0 & 0 & \rho a_-^* a_+ & |a_+|^2 \end{pmatrix}, \quad (4.31)$$

while the contribution corresponding to  $\mathbf{e} := \mathbf{e}_2$  is described by the matrix

$$\frac{1}{2} \begin{pmatrix} |b_-|^2 & 0 & \rho b_- b_+^* & 0 \\ 0 & |b_-|^2 & 0 & \rho b_- b_+^* \\ \rho b_-^* b_+ & 0 & |b_+|^2 & 0 \\ 0 & \rho b_-^* b_+ & 0 & |b_+|^2 \end{pmatrix}. \quad (4.32)$$

Adding these contributions together, we obtain (4.25)–(4.26) with (4.13).  $\square$

*Proof of Proposition 4.3.* Let  $I(\theta, \Delta)$  denote the integral in (4.22). Our goal is to evaluate  $I(\theta, \Delta)$  to within multiplicative correction of the order  $e^{O(\tau)|\mathbb{T}_N|}$ . If it were not for the indicator  $\chi_{\Delta', N}(\boldsymbol{\vartheta})$ , the integral would be Gaussian; unfortunately, as  $G_{N, \theta}(\boldsymbol{\vartheta})$  depends only on the differences  $\vartheta_r - \vartheta_{r'}$ , it would also diverge. We will therefore have to treat the indicator with some extra care by deriving suitable upper and lower bounds.

First, for any  $\lambda > 0$ ,

$$\chi_{\Delta', N}(\boldsymbol{\vartheta}) \leq \exp \left\{ \frac{\lambda}{2} \beta J c' \Delta^2 |\mathbb{T}_N| - \frac{\lambda}{2} \beta J \sum_{r \in \mathbb{T}_N} \vartheta_r^2 \right\}, \quad (4.33)$$

where  $c' := \sup_{|\Delta| < 2} (\Delta' / \Delta)^2 \in [1, \infty)$ . Substituting this bound into  $I(\theta, \Delta)$ , we now scale  $\vartheta_r$  by  $\sqrt{\beta J}$  and use Lemma 4.5 to diagonalize the quadratic form in the exponent to get

$$I(\theta, \Delta) \leq (\beta J)^{-\frac{1}{2}|\mathbb{T}_N|} e^{\frac{1}{2}\lambda\beta J c' \Delta^2 |\mathbb{T}_N|} \prod_{\mathbf{k} \in (\mathbb{T}_N^*)_+} \frac{1}{\sqrt{\det(\lambda + M(\mathbf{k}, \theta))}}. \quad (4.34)$$

The  $\lambda$  in the denominator now regularizes the contribution of the  $k = 0$  mode. We thus get the bound

$$\frac{\log \mathcal{Z}_N(\theta, \Delta)}{|\mathbb{T}_N|} \leq c\beta J\Delta^3 + \frac{\lambda}{2}\beta Jc'\Delta^2 - F_N(\theta, \lambda), \quad (4.35)$$

where

$$F_N(\theta, \lambda) := \frac{1}{2} \log(\beta J) + \frac{1}{2} \frac{1}{|\mathbb{T}_N|} \sum_{\mathbf{k} \in (\mathbb{T}_N^*)_+} \log \det[\lambda + M(\mathbf{k}, \theta)]. \quad (4.36)$$

Since  $F_N(\theta, \lambda)$  is, in the limit  $N \rightarrow \infty$ , finite and continuous in  $\theta$  and also obviously larger than  $F(\theta)$  for any  $\lambda > 0$ , it suffices to choose  $\lambda$  so that the first two terms on the right of (4.35) are strictly less than  $\tau$ . In particular, taking  $\lambda := \Delta$ , it suffices to choose  $\delta$  so that  $(c + \frac{1}{2}c')\delta \leq \tau$  and pick  $N_0$  so that  $F_N(\theta, \Delta) > F(\theta)$  for  $N \geq N_0$  and all  $\theta$ , to get  $N^{-2} \log \mathcal{Z}_N(\theta, \Delta) + F(\theta) \leq \tau$  whenever  $\beta J\Delta^3 \leq \delta$ .

The requisite lower bound is derived by a change-of-measure argument. Consider the Gaussian measure underlying the upper bound above:

$$\mathbb{P}_\lambda(d\boldsymbol{\vartheta}) := e^{|\mathbb{T}_N|F_N(\theta, \lambda) - \beta J G_{N, \theta}(\boldsymbol{\vartheta})} \exp\left\{-\frac{\lambda}{2}\beta J \sum_{r \in \mathbb{T}_N} \vartheta_r^2\right\} \prod_{r \in \mathbb{T}_N} \frac{d\vartheta_r}{\sqrt{2\pi}}. \quad (4.37)$$

Using  $\mathbb{E}_\lambda$  to denote the corresponding expectation, for any  $\lambda > 0$  we clearly have

$$\mathcal{Z}_N(\theta, \Delta) \geq e^{-c\beta J\Delta^3|\mathbb{T}_N| - |\mathbb{T}_N|F_N(\theta, \lambda)} \mathbb{E}_\lambda(\chi_{\Delta', N}). \quad (4.38)$$

Since  $F_N(\theta, \lambda)$  will tend to  $F(\theta)$  when  $N \rightarrow \infty$  and  $\lambda \downarrow 0$ , uniformly in  $N$  with  $|\sin(\theta)| > \tau$ , it will suffice to show that, for any  $\lambda > 0$  small enough,  $\mathbb{E}_\lambda(\chi_{\Delta', N})^{1/|\mathbb{T}_N|}$  is near one uniformly as  $N \rightarrow \infty$ . To this end, we first use chessboard estimates — exactly in the setting described above with  $B := 1$  — for the Gaussian measure  $\mathbb{P}_\lambda$  to derive

$$\mathbb{E}_\lambda(\chi_{\Delta', N}) = \mathbb{E}_\lambda\left(\prod_{r \in \mathbb{T}_N} 1_{\{|\vartheta_r| < \Delta'\}}\right) \geq \prod_{r \in \mathbb{T}_N} \mathbb{P}_\lambda(|\vartheta_r| < \Delta'). \quad (4.39)$$

Next the Chebyshev inequality yields

$$\mathbb{P}_\lambda(|\vartheta_r| \geq \Delta') \leq \mathbb{P}_\lambda(|\vartheta_r| \geq \Delta) \leq \frac{\mathbb{E}_\lambda(\vartheta_r^2)}{\Delta^2} = \frac{\text{Var}_\lambda(\vartheta_r)}{\Delta^2}, \quad (4.40)$$

where we used that  $\Delta' \geq \Delta$ . The variance of a Gaussian variable increases when we make the corresponding quadratic form in the exponent smaller (as a matrix). As this quadratic form is bounded below by  $\lambda\beta J$  times identity, the thus get

$$\mathbb{P}_\lambda(|\vartheta_r - \theta| < \Delta) \geq 1 - \frac{1}{\lambda\beta J\Delta^2}. \quad (4.41)$$

To derive the matching lower bound from (4.38), we thus need to choose  $\lambda$  small so that  $F_N(\theta, \lambda)$  is already close to  $F(\theta)$  but such that  $\lambda\beta J\Delta^2$  is still large. A specific choice that will

work is as follows: Pick  $\delta > 0$  so small that, for some  $N_0$ , we have  $|F_N(\theta, \sqrt{\delta}) - F(\theta)| < \frac{\tau}{2}$  for all  $N \geq N_0$  and all  $\theta$  with  $|\sin(2\theta)| > \tau$ . In addition, assume that also  $c\delta - \log(1 - \sqrt{\delta}) < \frac{\tau}{2}$ . The choice  $\lambda := \sqrt{\delta}$  and the bounds (4.38) and (4.41) then yield the desired lower bound on  $\log \mathcal{Z}_N(\theta, \Delta)$  once both conditions (4.16) are satisfied.  $\square$

*Proof of Proposition 4.2.* Assume that  $\mathbf{k}$  is such that  $a_{\pm}, b_{\pm} \neq 0$  and recall the definition of  $\rho$ . First, notice that  $\det M(\mathbf{k}, \theta) = 0$  when  $\rho = \pm 1$ . The case  $\rho = 1$  is easily checked by multiplying the matrix  $M(\mathbf{k}, \theta)$  by the vector  $(1, -1, -1, 1)$ ; the case  $\rho = -1$  is checked using the vector  $(1, 1, 1, 1)$ . Since  $\det M(\mathbf{k}, \theta)$  is an even quartic polynomial in  $\rho$  that is divisible by  $(1 - \rho^2)$ , it can be written in the form

$$\det M(\mathbf{k}, \theta) = (1 - \rho^2)(A - \rho^2 C). \quad (4.42)$$

An explicit computation yields

$$A := \frac{1}{16} (|a_-|^2 + |b_-|^2) (|a_+|^2 + |b_-|^2) (|a_-|^2 + |b_+|^2) (|a_+|^2 + |b_+|^2) \quad (4.43)$$

and

$$C := \frac{1}{16} (|a_-|^2 |a_+|^2 - |b_-|^2 |b_+|^2)^2. \quad (4.44)$$

Obviously, under the conditions  $a_{\pm}, b_{\pm} \neq 0$  we have  $A > 0$  and  $C \geq 0$ . Moreover, since  $|a_{\pm}|, |b_{\pm}| \leq 2$ , we have  $A \leq 16$ . Therefore

$$\frac{|a_-|^2 |a_+|^2 |b_-|^2 |b_+|^2}{16} \leq A - C \leq A - \rho^2 C \leq A \leq 16. \quad (4.45)$$

The left-hand side equals  $\sin^2(k_1) \sin^2(k_2)$  and so (4.15) follows by plugging the above inequalities into (4.42).

To get the second part of the claim, let  $g(x) := (1 - x)(A - Cx)$  be the function appearing on the right-hand side of (4.42). Then  $g'(0) = -A - C < 0$  and  $g'(1) = -A + C < 0$ . But  $g$  is quadratic and so it is strictly decreasing throughout  $[0, 1)$ . In particular,  $g(1) < g(x)$  for all  $x \in [0, 1)$ . Using this in (4.14), the desired claims follow.  $\square$

### 4.3. Good and bad events

For the proof of our key estimates, we will follow, as in the previous sections, the general scheme developed in [4, 1] and further discussed in [2, Sect. 6.4]. For a positive constant  $\eta$  and an even integer  $B$  that divides  $N$ , define the *good*  $B$ -block events  $\mathcal{G}_x$  and  $\mathcal{G}_z$  as follows. First set

$$\mathcal{G}_x^0 := \bigcap_{r \in \Lambda_B} \{S: |S_r - e_1| < \eta\}, \quad (4.46)$$

and similarly for  $\mathcal{G}_z^0$  with  $e_1$  replaced by  $e_2$ . Now let  $\Lambda_B^0$  denote the even-sublattice vertices in  $\Lambda_B$ ; these are the lower-left corners of either  $x$  or  $z$ -plaquettes. We use  $\Lambda_B^1$  to denote the set of those vertices in  $\Lambda_B^0$  that are lower-left corners of  $x$ -plaquettes,  $\Lambda_B^2 = \Lambda_B^0 \setminus \Lambda_B^1$ . For any set  $\Lambda \subset \Lambda_B^0$ , let  $\varphi_\Lambda(\mathbf{S})$  be the composition of  $\varphi_r$  for all  $r \in \Lambda$ . As these maps commute, the order of the composition is immaterial. With the help of these notations, we now set

$$\mathcal{G}_\alpha := \bigcup_{\Lambda \subset \Lambda_B^0} \varphi_\Lambda(\mathcal{G}_\alpha^0), \quad \alpha = 1, 2. \quad (4.47)$$

These are the *good*  $B$ -block events; the requisite *bad* event is defined by

$$\mathcal{B} := (\mathcal{G}_x \cup \mathcal{G}_z)^c. \quad (4.48)$$

For a  $B$ -block event  $\mathcal{A}$  define the quantity

$$\mathfrak{p}_\beta(\mathcal{A}) := \lim_{N \rightarrow \infty} \left[ \mu_{N,\beta} \left( \bigcap_{r \in \mathbb{T}_{N/B}} \theta_r(\mathcal{A}) \right) \right]^{(B/N)^2}, \quad (4.49)$$

where  $N$  is taken to infinity along the even multiples of  $B$ . The limit exists by standard subadditivity arguments. Note that this quantity is a limiting version of the objects on the right-hand side of the chessboard estimates. A key input for our proofs is the observation that  $\mathfrak{p}_\beta(\mathcal{B})$  is small:

**Proposition 4.6** *For each  $\eta > 0$  and  $\varepsilon > 0$  there exists  $\beta_0 < \infty$  such that for any  $\beta \geq \beta_0$  there exists  $B$  for which the bad event  $\mathcal{B}$  defined using these  $\eta$  and  $B$  satisfies*

$$\mathfrak{p}_\beta(\mathcal{B}) < \varepsilon. \quad (4.50)$$

In order to show that the bad event is unlikely to occur, we will need to further partition it into several subevents. First, consider a number  $\Delta > 0$  — to be eventually chosen in dependence of  $\beta$  so that the conditions (4.16) hold true — and let us isolate the configurations where the corresponding component of two neighbouring spins somewhere in  $\Lambda_B$  differ by more than a quantity proportional to  $\Delta$ :

$$\mathcal{B}_E := \bigcup_{\alpha=x,z} \bigcup_{\substack{\langle r,r' \rangle_\alpha \\ r,r' \in \Lambda_B}} \left\{ \mathbf{S} \in \mathcal{B} : |S_r^\alpha - S_{r'}^\alpha| \geq \frac{\Delta}{16B} \right\}. \quad (4.51)$$

These will be the configurations with too much energy — hence the subscript “E”. The remaining “bad” configurations will be collected in the set

$$\mathcal{B}_{SW} := \mathcal{B} \setminus \mathcal{B}_E, \quad (4.52)$$

where “SW” designates the method — a *spin-wave* calculation — that will be used to estimate the contributions to this event. The following claim provides a key structural information on the configurations contained in  $\mathcal{B}_{SW}$ :

**Lemma 4.7** *Suppose that  $\eta$  and  $\Delta$  satisfy the inequalities*

$$\Delta < \eta^2 < 1. \quad (4.53)$$

*Then for each  $\mathbf{S} \in \mathcal{B}_{\text{SW}}$ , there is a unit vector  $\mathbf{v} \in \mathbb{R}^d$  and a set  $\Lambda \subset \Lambda_B^0$  such that*

$$|(\boldsymbol{\varphi}_\Lambda(\mathbf{S}))_{\mathbf{r}} - \mathbf{v}| \leq \frac{\Delta}{2}, \quad \mathbf{r} \in \Lambda_B. \quad (4.54)$$

*Proof.* Fix  $\mathbf{S} \in \mathcal{B}_{\text{SW}}$  and define  $\Lambda \subset \Lambda_B^0$  by taking  $\Lambda \cap \Lambda_B^\alpha = \{\mathbf{r} \in \Lambda_B : S_{\mathbf{r}}^\alpha < 0\}$ ,  $\alpha = 1, 2$ . Consider the configuration  $\tilde{\mathbf{S}} := \boldsymbol{\varphi}_\Lambda(\mathbf{S})$ . Clearly,  $\tilde{S}_{\mathbf{r}}^\alpha \geq 0$  for all  $\mathbf{r} \in \Lambda_B^\alpha$ ,  $\alpha = x, z$ . The proof of the main claim now comes in two steps. First we will show that both components of all  $S_{\mathbf{r}}$ ,  $\mathbf{r} \in \Lambda_B$ , are at least  $\sqrt{\zeta}$  in absolute value, where

$$\zeta := \frac{\eta^2}{2} - \frac{\Delta}{4}. \quad (4.55)$$

This will be used to ensure that  $\tilde{S}_{\mathbf{r}}^\alpha \geq 0$  everywhere. Then we will rerun part of the argument to show that  $\tilde{\mathbf{S}}$  is, actually, to within  $O(\Delta)$  of a constant configuration.

Let  $\zeta$  be as above and abbreviate  $c := (16B)^{-1}$ . Notice that the bounds (4.53) imply

$$\zeta + \frac{\Delta}{4} \leq \frac{1}{2}, \quad c\Delta < \frac{1}{4} \quad \text{and} \quad \sqrt{\zeta} > 2c\Delta. \quad (4.56)$$

First, we claim that

$$\mathbf{S} \in \mathcal{B}_{\text{SW}} \quad \Rightarrow \quad [S_{\mathbf{r}}^\alpha]^2 \geq \zeta, \quad \mathbf{r} \in \Lambda_B, \quad \alpha = x, z. \quad (4.57)$$

Focusing our attention on  $\alpha = 1$ , assume that  $[S_{\mathbf{r}}^x]^2 < \zeta$  at some  $\mathbf{r} \in \Lambda_B$  and derive a contradiction with  $\mathbf{S} \in \mathcal{B}_{\text{SW}}$ . Notice that  $\mathbf{S} \in \mathcal{B}_{\text{SW}}$  implies that

$$|S_{\mathbf{r}}^\alpha - S_{\mathbf{r}'}^\alpha| < c\Delta \quad (4.58)$$

for any  $\mathbf{r}, \mathbf{r}' \in \Lambda_B$  connected by an  $\alpha$ -edge. If  $\mathbf{r}'$  is a neighbor of  $\mathbf{r}$  over a  $x$ -edge, then

$$[S_{\mathbf{r}'}^x]^2 < [\sqrt{\zeta} + c\Delta]^2 \leq \zeta + 2c\Delta, \quad (4.59)$$

where we used that  $\zeta \leq \frac{1}{2}$  and  $c\Delta \leq \frac{1}{2}$ . On the other hand, if  $\mathbf{r}'$  is connected to  $\mathbf{r}$  by a  $z$ -edge, we can proceed via the second components to get again

$$[S_{\mathbf{r}'}^x]^2 = 1 - [S_{\mathbf{r}'}^z]^2 \quad (4.60)$$

$$\leq 1 - [S_{\mathbf{r}}^z]^2 + 2|S_{\mathbf{r}'}^z - S_{\mathbf{r}}^z| \quad (4.61)$$

$$< \zeta + 2c\Delta. \quad (4.62)$$

Examining all pairs of nearest neighbors along a shortest path from  $\mathbf{r}$  to any  $\mathbf{r}' \in \Lambda_B$  and using that this path has at most  $2B$  edges, we conclude

$$[S_{\mathbf{r}'}^x]^2 < \zeta + (2B)2c\Delta = \zeta + \frac{\Delta}{4}, \quad \mathbf{r}' \in \Lambda_B. \quad (4.63)$$

Notice that to maintain the argument from (4.59) for the proof of the inequality  $[S_{r''}^x]^2 < [S_{r'}^x]^2 + 2c\Delta$  for every step  $(r', r'')$  of the iteration, we actually need to invoke that  $\zeta + \frac{\Delta}{4} \leq \frac{1}{2}$ . At the same time, since  $\sqrt{1 - (\zeta + \frac{\Delta}{4})} \geq \frac{1}{2} \geq 2c\Delta$ , the triangle inequality implies  $\tilde{S}_r^z \geq 0$  on each z-plaquette, and thus everywhere in  $\Lambda_B$ . In addition, this allows us to compute further:

$$|\tilde{S}_{r'} - e_2|^2 = [S_{r'}^x]^2 + \left(1 - \sqrt{1 - [S_{r'}^x]^2}\right)^2 \quad (4.64)$$

$$\leq [S_{r'}^x]^2 + [S_{r'}^x]^4 \leq 2[S_{r'}^x]^2 < 2\left(\zeta + \frac{\Delta}{4}\right) = 2\zeta + \frac{\Delta}{2} = \eta^2, \quad (4.65)$$

where we used that  $1 - \sqrt{1 - x} \leq x$  for  $0 < x < 1$ . This would imply that  $\tilde{S} \in \mathcal{G}_z^0$  and thus  $S \in \mathcal{G}_z$ , in contradiction with the assumption that  $S \in \mathcal{B}_{\text{SW}} \subset \mathcal{B}$ .

Having proven the bound (4.57), we can use the fact that  $\sqrt{\zeta} > 2c\Delta$  to imply that  $\tilde{S}_r^\alpha \geq 0$  for all  $r \in \Lambda_B$ . Combining now the positivity of components of  $\tilde{S}_r^\alpha$  with (4.57), we can improve the bound on the difference of the 1-components of spins over a z-edge  $\langle r, r' \rangle_z$  as follows: Suppose without loss of generality that  $S_{r'}^z \geq S_r^z$  and write

$$|S_{r'}^x - S_r^x| = \left| \sqrt{1 - [S_{r'}^z]^2} - \sqrt{1 - [S_r^z]^2} \right| \quad (4.66)$$

$$\leq \frac{|S_{r'}^z - S_r^z|}{\sqrt{1 - [S_{r'}^z]^2}} \leq \frac{c\Delta}{\sqrt{1 - \zeta}} \quad (4.67)$$

In conjunction with  $|S_{r'}^z - S_r^z| < c\Delta$ , this implies

$$|S_{r'} - S_r| \leq c\Delta \frac{2 - \zeta}{1 - \zeta} \leq 4c\Delta \quad (4.68)$$

where we used that  $1 - \zeta \geq \frac{1}{2}$ . As  $(2B)4c\Delta < \frac{\Delta}{2}$ , the desired claim now follows with  $v := \tilde{S}_0$  and  $\Lambda$  as above by examining a path of minimal length between the origin and any other site in  $\Lambda_B$ .  $\square$

The previous lemma allows us to further partition  $\mathcal{B}_{\text{SW}}$  as follows. Let  $v_1, \dots, v_n$  denote the unit vectors representing the complex  $n$ -th roots of unity,  $v_\ell := e^{(\ell \frac{2\pi}{n})}$ ,  $\ell = 1, \dots, n$ , where  $n := \lfloor \frac{2\pi}{\Delta} \rfloor + 1$  is the smallest integer for which  $n\Delta > 2\pi$ . Defining

$$\mathcal{B}_{\text{SW}}^{(\ell)} := \bigcup_{\Lambda \subset \Lambda_B^0} \varphi_\Lambda \left( \{S \in \mathcal{B}_{\text{SW}} : \forall r \in \Lambda_B \ |S_r - v_\ell| < \Delta\} \right), \quad (4.69)$$

it follows that

$$\mathcal{B}_{\text{SW}} = \bigcup_{\ell=1}^n \mathcal{B}_{\text{SW}}^{(\ell)}. \quad (4.70)$$

Having in mind that  $\Delta \ll \eta$  and that, by definition,  $\mathcal{B}_{\text{SW}} \cap (\mathcal{G}_x \cup \mathcal{G}_z) = \emptyset$ , we notice that  $\mathcal{B}_{\text{SW}}^{(\ell)} = \emptyset$  whenever the distance of  $v_\ell$  from the points  $\pm e_1, \pm e_2$ , is less than  $\eta - \Delta$ . The set function  $\mathcal{A} \mapsto \mathfrak{p}_\beta(\mathcal{A})$  is subadditive (see [2, Lemma 5.9]) and so we have

$$\mathfrak{p}_\beta(\mathcal{B}) \leq \mathfrak{p}_\beta(\mathcal{B}_E) + \sum_{\ell=1}^n \mathfrak{p}_\beta(\mathcal{B}_{\text{SW}}^{(\ell)}). \quad (4.71)$$

It remains to derive suitable estimates on  $\mathfrak{p}_\beta(\mathcal{B}_E)$  and  $\mathfrak{p}_\beta(\mathcal{B}_{SW}^{(\ell)})$ . An indispensable ingredient will be the following lower bound on the full partition function:

**Lemma 4.8** *Fix  $\tau > 0$  such that  $\sin(2\tau) > \tau$  and let  $\delta$  and  $N_0$  be as in Proposition 4.3. If  $\Delta$  and  $\beta$  satisfy the conditions (4.16) and  $N \geq N_0$ , then*

$$(Z_{N,\beta})^{1/|\mathbb{T}_N|} \geq e^{\beta J - F(\tau) - \tau}, \quad (4.72)$$

where  $F$  is the free energy introduced in (4.14).

*Proof.* By restricting the integrals over the spins to the set where  $|\mathbf{S}_r - e(\tau)| < \Delta$ , we get

$$Z_{N,\beta} \geq e^{\beta J |\mathbb{T}_N|} \mathcal{Z}_N(\tau, \Delta). \quad (4.73)$$

From here the claim follows by invoking the bound (4.17).  $\square$

**Lemma 4.9** *Suppose that  $\eta \in (0, \pi/4)$  and, given  $\tau > 0$  with  $\sin(2\tau) > \tau$ , let  $\delta$  be as in Proposition 4.3 and assume that  $\Delta$  and  $\beta$  satisfy the conditions (4.16) and that  $\Delta < \eta^2 < (\frac{\pi}{4})^2$ . Then*

$$\mathfrak{p}_\beta(\mathcal{B}_E) \leq (2B)^2 e^{c_1 B^2 - c_2 \beta J \Delta^2 / B^2}, \quad (4.74)$$

where  $c_1 := \frac{1}{2} \log(2\pi) + F(\tau) + \tau$  and  $c_2 := 1/512$ , and

$$\mathfrak{p}_\beta(\mathcal{B}_{SW}^{(\ell)}) \leq 2^{1/4} e^{-[F(\eta - \Delta) - F(\tau) - \tau] B^2}, \quad \ell = 1, \dots, n. \quad (4.75)$$

*Proof.* Consider the disseminated event  $\overline{\mathcal{B}}_E := \bigcap_{r \in \mathbb{T}_{N/B}} \theta_r(\mathcal{B}_E)$ . We need to derive an upper bound on the constrained partition function  $Z_{N,\beta}(\overline{\mathcal{B}}_E)$ , which is given by the same integral as the full partition function but only over configurations in the event  $\overline{\mathcal{B}}_E$ . Since on  $\mathcal{B}_E$ , each configuration has an “energetically charged” edge in each block in  $\{\Lambda_B + Bt : t \in \mathbb{T}_{N/B}\}$ , Lemma 3.1 tells us

$$\mathcal{H}_N(\mathbf{S}) \geq \frac{1}{2} \left( \frac{\Delta}{16B} \right)^2 J(N/B)^2 - J|\mathbb{T}_N|, \quad \mathbf{S} \in \overline{\mathcal{B}}_E. \quad (4.76)$$

To account for the entropy, we note that there are altogether  $2B(B-1)$  positions where the “energetically charged” edge can occur in each translate of  $\Lambda_B$ . Lemma 4.8 now gives

$$\mu_{N,\beta}(\overline{\mathcal{B}}_E)^{(B/N)^2} = \left( \frac{Z_{N,\beta}(\overline{\mathcal{B}}_E)}{Z_{N,\beta}} \right)^{(B/N)^2} \quad (4.77)$$

$$\leq (2\pi)^{B^2/2} (2B)^2 e^{-\beta J \Delta^2 / 512 + [F(\tau) + \tau] B^2}, \quad (4.78)$$

where the factor  $(2\pi)^{B^2/2}$  comes from performing (unconstrained) integrals of the spins. The limit  $N \rightarrow \infty$  along multiples of  $B$  now yields (4.74).

Next we consider the event  $\overline{\mathcal{B}_{\text{SW}}^{(\ell)}} := \bigcap_{r \in \mathbb{T}_{N/B}} \theta_r(\mathcal{B}_{\text{SW}}^{(\ell)})$ . We only need to focus on  $\ell$  satisfying  $|\frac{2\pi}{n}\ell - \frac{2\pi}{n}k| \geq \eta - \Delta$  for  $k = 0, 1, 2, 3$  because otherwise the event is void. Here we notice that, for each  $\mathcal{S} \in \overline{\mathcal{B}_{\text{SW}}^{(\ell)}}$ , there is a unique  $\Lambda \subset \bigcup_{t \in \mathbb{T}_{N/B}} (Bt + \Lambda_B^0)$  such that the configuration  $\tilde{\mathcal{S}} := \varphi_\Lambda(\mathcal{S})$  is within  $\Delta$  of the vector  $v_\ell$ . The number of distinct  $\Lambda$  associated with a single  $\tilde{\mathcal{S}}$  is  $2^{(N/B)^2/4}$ . As both the *a priori* measure and the Hamiltonian are invariant under  $\varphi_\Lambda$ , this yields

$$Z_{N,\beta}(\overline{\mathcal{B}_{\text{SW}}^{(\ell)}}) \leq 2^{(N/B)^2/4} Z_N(\ell \frac{2\pi}{n}, \Delta). \quad (4.79)$$

Hence,

$$\mu_{N,\beta}(\overline{\mathcal{B}_{\text{SW}}^{(\ell)}})^{(B/N)^2} = \left( \frac{Z_{N,\beta}(\overline{\mathcal{B}_{\text{SW}}^{(\ell)}})}{Z_{N,\beta}} \right)^{(B/N)^2} \quad (4.80)$$

$$\leq 2^{1/4} \exp\left\{ [-F(\ell \frac{2\pi}{n}) + F(\tau) + \tau] B^2 \right\}. \quad (4.81)$$

Now we invoke Proposition 4.2 and apply the monotonicity and periodicity properties of  $F$  to infer that  $F(\ell \frac{2\pi}{n}) \geq F(\eta - \Delta)$  for any  $\ell$  for which  $\mathcal{B}_{\text{SW}}^{(\ell)} \neq \emptyset$ . (This is where we need  $0 < \Delta < \eta < \pi/4$ .) The limit  $N \rightarrow \infty$  (along multiples of  $B$ ) finishes the claim.  $\square$

*Proof of Proposition 4.6.* Fix  $\epsilon > 0$  and  $\eta \in (0, \pi/4)$  and set  $\tau$  to be a positive number such that  $\sin(2\tau) > \tau$  and

$$F(\eta/2) > F(\tau) + \tau. \quad (4.82)$$

This is possible because  $\theta \mapsto F(\theta)$  tends to minus infinity as  $\theta \downarrow 0$ , see Fig. 2. Let  $\delta$  be related to this  $\tau$  as in Proposition 4.3. We will henceforth link  $\Delta$  and  $B$  to  $\beta > 0$  via

$$\Delta := \beta^{-\frac{5}{12}} \quad \text{and} \quad B := 2 \lceil \log \beta \rceil. \quad (4.83)$$

Notice that this choice of  $\Delta$  will make the bounds in Proposition 4.3 true once  $\beta$  is sufficiently large. Let  $\beta_0$  be now a value such that for all  $\beta \geq \beta_0$  these bounds hold, the condition (4.53) in Lemma 4.7 is satisfied, and the inequalities

$$\eta - \Delta > \frac{\eta}{2} \quad (4.84)$$

and

$$(2B)^2 e^{c_1 B^2 - c_2 \beta \Delta^2 / B^2} + \frac{4\pi}{\Delta} 2^{1/4} e^{-[F(\eta/2) - F(\tau) - \tau] B^2} < \epsilon \quad (4.85)$$

hold true. This is possible by (4.82) and our choices (4.83) — and the fact that the quantities  $c_1$  and  $c_2$  depend only on  $\tau$ .

We claim that (4.85) implies the desired bound  $\mathfrak{p}_\beta(\mathcal{B}) < \epsilon$ . Indeed, using that  $n \leq 4\pi/\Delta$  for sufficiently small  $\Delta$ , it suffices to insert into (4.71) the bound on  $\mathfrak{p}_\beta(\mathcal{B}_E)$  and the uniform

bound on  $\mathfrak{p}_\beta(\mathcal{B}_{\text{SW}}^{(\ell)})$  from (4.74–4.75). As a result we get the estimate of  $\mathfrak{p}_\beta(\mathcal{B})$  by the left-hand side of (4.85). (We also used that  $F$  is strictly increasing on  $(0, \pi/4)$  and applied (4.84).) By (4.85),  $\mathfrak{p}_\beta(\mathcal{B})$  is thus less than  $\varepsilon$  for any  $\beta$  in excess of  $\beta_0$  defined above.  $\square$

## 5. Proofs of main results

### 5.1. Classical model

We are now ready to prove our main results for the classical system. The key inputs are chessboard estimates alongside with the bound in Proposition 4.6.

*Proof of Theorem 2.3.* The standard line of reasoning leading to the proof of phase coexistence in the present context is based on the observation that distinct types of good events are unlikely to occur in the same configuration. A formal statement is as follows:

Assume that  $B \geq 4$  and  $\eta < 2 \sin(\pi/8)$ . For any  $\varepsilon > 0$  there exist  $\delta > 0$  and  $N_0$  such that if  $\mathfrak{p}_\beta(\mathcal{B}) < \delta$  and  $N \geq N_0$ , then  $\mu_{N,\beta}(\theta_r(\mathcal{G}_x) \cap \theta_{r'}(\mathcal{G}_z)) \leq \varepsilon$  for all  $\mathbf{r}, \mathbf{r}' \in \mathbb{T}_{N/B}$ .

The proof of this fact is an application of a Peierls-type contour argument. Fix  $\mathbf{r}, \mathbf{r}' \in \mathbb{T}_{N/B}$ ,  $\mathbf{r} \neq \mathbf{r}'$ , and let  $\mathcal{Y}(\mathbf{r})$  denote the set of all  $\Lambda \subset \mathbb{T}_{N/B}$  such that both  $\Lambda$  and  $\mathbb{T}_{N/B} \setminus \Lambda$  are connected and  $\Lambda \ni \mathbf{r}$ . We use  $\partial\Lambda$  to denote the set of all vertices outside, but adjacent to  $\Lambda$ . For  $\Lambda \in \mathcal{Y}(\mathbf{r})$ , consider the event

$$R_\Lambda := \left\{ \sum_{\mathbf{r}' \in \partial\Lambda} 1_B \circ \theta_{\mathbf{r}'} > \frac{1}{5} |\partial\Lambda| \right\}, \quad (5.1)$$

where  $\theta_r(\mathcal{B}) = \vartheta_r(\mathcal{B})$  with  $\theta_r$  denoting the shift by  $B\mathbf{r}$ . Recalling also the notation  $\sigma_1$ , resp.,  $\sigma_2$  for the shift by  $2\mathbf{e}_1$ , resp.,  $2\mathbf{e}_2$ , we now claim

$$\theta_r(\mathcal{G}_x) \cap \theta_{\mathbf{r}'}(\mathcal{G}_z) \subset \bigcup_{\substack{\Lambda \in \mathcal{Y}(\mathbf{r}) \\ \mathbf{r}' \notin \Lambda}} (R_\Lambda \cup \bigcup_{j=1,2} (\sigma_j(R_\Lambda) \cup \sigma_j^{-1}(R_\Lambda)))_{\mathbf{r}}. \quad (5.2)$$

To prove this, consider a configuration  $\mathbf{S} \in \theta_r(\mathcal{G}_x) \cap \theta_{\mathbf{r}'}(\mathcal{G}_z)$  and let  $\Lambda'$  denote the set of all  $\mathbf{s} \in \mathbb{T}_{N/B}$  for which there is a nearest-neighbor path that starts at  $\mathbf{r}$ , ends at  $\mathbf{s}$ , and visits only blocks  $B\mathbf{s}' + \Lambda_B$ , where  $\theta_{\mathbf{s}'}(\mathcal{G}_x)$  occurs. Clearly,  $\mathbf{r}' \notin \Lambda'$  so we may use  $\Lambda$  to denote the set of all  $\mathbf{s} \in \mathbb{T}_{N/B}$  such that every nearest neighbor path from  $\mathbf{s}$  to  $\mathbf{r}'$  contains at least one vertex of  $\Lambda'$  — i.e.  $\Lambda$  is  $\Lambda'$  with all of its “holes” filled.

Our task is to show that

$$\mathbf{S} \in R_\Lambda \cup \bigcup_{j=1,2} (\sigma_j(R_\Lambda) \cup \sigma_j^{-1}(R_\Lambda))_{\mathbf{r}}. \quad (5.3)$$

If  $S \in R_\Lambda$ , then we are done, so let us suppose that  $S \notin R_\Lambda$ . Under this condition more than  $4/5$  of all blocks corresponding to vertices in  $\partial\Lambda$  are good but, since they are not part of  $\Lambda$ , they are of type  $\mathcal{G}_z$ . Then, however, more than  $2/5$  of  $\mathcal{G}_z$ -blocks in  $\partial\Lambda$  are adjacent to  $\Lambda$  in one coordinate direction — say it is the first one — and thus  $1/5$  of them are adjacent in either positive or negative coordinate direction. Observing that  $S \in \theta_r(\mathcal{G}_x) \cap \theta_{r+e_j}(\mathcal{G}_z)$  implies that  $\sigma_1^{-1}(S) \in \theta_r(\mathcal{B})$  — and similarly for the opposite direction — we get (5.3).

With (5.3) on hand, we now perform a standard version of the Peierls argument combined with chessboard estimates. A key input is the fact that, for some  $c \in (1, \infty)$ ,

$$|\{\Lambda \in \mathcal{Y}(r) : |\partial\Lambda| = n\}| \leq c^n. \quad (5.4)$$

Using inclusion-exclusion and  $\sigma_j$ -invariance of  $\mu_{N,\beta}$ , it now suffices to show that for some  $\tilde{c} < \infty$  and  $N_0 < \infty$ ,

$$\mu_{N,\beta}(R_\Lambda) \leq \tilde{c}^{|\partial\Lambda|} \mathfrak{p}_\beta(\mathcal{B})^{|\partial\Lambda|/5}, \quad \Lambda \in \mathcal{Y}(r). \quad (5.5)$$

To this end, let  $N_0$  be such that, for all  $N \geq N_0$ , the probability on the right hand side of (4.49) for  $\mathcal{A} := \mathcal{B}$  is at most  $[2\mathfrak{p}_\beta(\mathcal{B})]^{\mathbb{T}^N}$ . Now, cover  $R_\Lambda$  by the union over all subsets of  $\partial\Lambda$  with  $|\partial\Lambda|/5$  elements where the bad event occurs. The number of such partitions is at most  $2^{|\partial\Lambda|}$ ; the probability of each occurrence is estimated by  $[2\mathfrak{p}_\beta(\mathcal{B})]^{|\partial\Lambda|/5}$ . This proves (5.5) with  $\tilde{c} := 2^{6/5}$  and thus the above claim.

Applying this claim alongside the fact that, on a good event, the spins are oriented along one of the coordinate directions proves (2.8) and (2.9). To get also (2.10), one just follows literally the argument proving the main result in [6].  $\square$

*Proof of Theorem 2.4.* This is a consequence of Theorem 2.3 and a convexity argument. Let  $\mathcal{E}^x = \mathcal{E}_0$  denote the Wenzel-Janke plaquette energy (1.5) for the type- $x$  plaquette at the origin. Let  $\mathfrak{G}(J_1, J_2)$  denote the set of all translation-invariant, ergodic Gibbs states. The convexity of  $J_1 \mapsto \log Z_N$  then implies that once  $J_1 < J'_1$ , for any  $\mu \in \mathfrak{G}(J_1, J_2)$  and  $\mu' \in \mathfrak{G}(J'_1, J_2)$  we have

$$\mathbb{E}_\mu(\mathcal{E}^x) \leq \mathbb{E}_{\mu'}(\mathcal{E}^x). \quad (5.6)$$

Now, at  $J_1 = J_2 =: J$  and  $\beta \geq \beta_0$ , Theorem 2.3 guarantees the existence of a  $\mu_\beta^x \in \mathfrak{G}(J, J)$ , such that

$$\mathbb{E}_{\mu_\beta^x}(\mathcal{E}^x) \geq 4 - \delta. \quad (5.7)$$

(Here  $\delta$  may differ from the one used in the statement of Theorem 2.3.) Thus, for all  $J_1 > J_2$  and  $\mu \in \mathfrak{G}(J_1, J_2)$ , we have

$$\mathbb{E}_\mu(\mathcal{E}^x) \geq 4 - \delta. \quad (5.8)$$

As  $\mathcal{E}^x$  is the sum of four terms of the type  $S_r^x S_{r'}^x$ , which are less than one, we must have

$$\mathbb{E}_\mu(S_r^x S_{r'}^x) \geq 1 - \delta \quad (5.9)$$

for all nearest-neighbour pairs  $\langle r, r' \rangle_x$  of type  $x$ . This forces  $\mathbb{E}_\mu(|S_r^x|) \geq 1 - \delta$  and thus  $\mathbb{E}_\mu([S_r^x]^2) \geq 1 - 2\delta$ . As  $\delta$  is arbitrary, this proves the claim.  $\square$

We will also give the formal proof of absence of magnetic order:

*Proof of Theorem 2.2.* It is easy to check that every Gibbs measure is invariant under the action of any  $\varphi_r$ . As each spin belongs to one plaquette of type- $x$  and one plaquette of type- $z$ , it follows that the distribution of  $S_r^\alpha$  is symmetric around zero, for any  $r$  and any  $\alpha = x, z$ . This implies the claim.  $\square$

## 5.2. Quantum model

Our set of results for the quantum model will be derived by an application of the general theory developed in [5] whose main conclusion can be found in Theorem 3.7 of [5]. This theorem says roughly the following: Whenever a quantum-spin model satisfies the conditions of (quantum) reflection positivity, and the classical system admits a proof of phase coexistence at a positive temperature by means of chessboard estimates, then the same phase coexistence occurs in the quantum system provided the magnitude of the quantum spin is sufficiently larger than the inverse temperature squared.

First let us check that the prerequisite concerning the quantum reflection positivity is satisfied. Since we are using reflections in planes bisecting edges of  $\mathbb{T}_N$ , this is proved by the same argument as in the classical case, except that we need to write all operators in a basis in which their matrix elements are all real valued. This is satisfied automatically in the representation (2.15) in which  $\hat{S}_r^z, \hat{S}_r^\pm$  are real and so is  $\hat{S}_r^x = \frac{1}{2}(\hat{S}_r^+ + \hat{S}_r^-)$ .

All we have to do is thus adapt the proof for the classical model to plug into Theorem 3.7 of [5]. We begin by introducing the formalism of *coherent states* that the whole connection is based on. Consider the space  $\mathbb{C}^{2S+1}$  that carries the corresponding  $2S + 1$ -dimensional representation of  $\mathfrak{su}(2)$ . Let  $\Omega$  be a vector on the unit sphere  $\mathbb{S}_2$  in  $\mathbb{R}^3$  that is described by the spherical angles  $\theta$  and  $\phi$ . Then we set

$$|\Omega\rangle := \sum_{M=-S}^S \binom{2S}{S+M}^{1/2} [\cos(\theta/2)]^{S+M} [\sin(\theta/2)]^{S-M} e^{i(S-M)\phi} |M\rangle. \quad (5.10)$$

Abusing the notation slightly, whenever  $\Omega := (\Omega_r)$  is a collection of such vectors, we will denote the corresponding product state by  $|\Omega\rangle := \otimes_r |\Omega_r\rangle$ .

The coherent states have a number of remarkable properties of which relevant for us are particularly those listed in Sect. 2.1 of [5]. Here we will only need the notions of the *lower* and

upper symbols. Given a linear operator  $\hat{A}$  on  $\bigotimes_{r \in \Lambda} \mathbb{C}^{2S+1}$ , the lower symbol is the function  $\Omega \mapsto \langle \hat{A} \rangle_\Omega$  on  $(\mathcal{S}_2)^\Lambda$  such that

$$\langle \hat{A} \rangle_\Omega := \langle \Omega | \hat{A} | \Omega \rangle. \quad (5.11)$$

The upper symbol is, in turn, a function  $\Omega \mapsto [\hat{A}]_\Omega$  such that

$$\hat{A} = \left( \frac{2S+1}{4\pi} \right)^{|\Lambda|} \int_{(\mathcal{S}_2)^\Lambda} d\Omega [\hat{A}]_\Omega |\Omega\rangle \langle \Omega|, \quad (5.12)$$

where  $d\Omega := \prod_{r \in \Lambda} d\Omega_r$  with  $d\Omega_r$  denoting the uniform measure on  $\mathcal{S}_2$  with total mass  $4\pi$ . The upper symbol is not necessarily unique, so we use the notation  $[\hat{A}]_\Omega$  to denote any version thereof. The upper and lower symbols are two natural classical approximations of the quantum Hamiltonian, so we need to check that they are reasonably close:

**Lemma 5.1** *Consider the operator  $\hat{\mathcal{H}}_N$  in (2.16). Then there is a constant  $c = c(J_1, J_2) < \infty$  and a version of the upper symbol  $[\hat{\mathcal{H}}_N]_\Omega$  such that for each  $\Omega \in (\mathcal{S}_2)^{\mathbb{T}_N}$ ,*

$$|\langle \hat{\mathcal{H}}_N \rangle_\Omega - [\hat{\mathcal{H}}_N]_\Omega| \leq \frac{c}{S} |\mathbb{T}_N|. \quad (5.13)$$

*Proof.* Let  $\mathcal{H}_N(\mathcal{S})$  denote the classical Hamiltonian corresponding to (2.16) — obtained by replacing the operators  $\hat{S}_r$  by vectors  $\mathbf{S}_r \in \mathcal{S}_2$  and dropping the normalization by  $S^2$ . Then the desired bound will hold once we verify that

$$\langle \hat{\mathcal{H}}_N \rangle_\Omega = \mathcal{H}_N(\Omega) \quad \text{and} \quad [\hat{\mathcal{H}}_N]_\Omega = (1 + 1/S)^2 \mathcal{H}_N(\Omega) \quad (5.14)$$

for some version of  $[\hat{\mathcal{H}}_N]_\Omega$ . This is in turn shown by noting that  $\hat{\mathcal{H}}_N$  is multilinear in the operators  $\hat{S}$  — which means that for each term in the sum the corresponding symbols take the form of a product — and by the fact that  $\langle \hat{S}_r \rangle_\Omega = S\Omega_r$  and that  $[\hat{S}_r]_\Omega := (S+1)\Omega_r$  is a version of the upper symbol for  $\hat{S}$ . The bound (5.13) now follows by the fact that  $\mathcal{H}_N(\Omega)$  is bounded by a constant times  $|\mathbb{T}_N|$ , uniformly in  $\Omega$  and  $S \geq 1/2$ .  $\square$

Our next step will be to verify the conditions required by [5, Theorem 3.7] for the corresponding classical model. The problem here is that the link through the coherent states naturally leads to *three*-component classical spins. Consider the events  $\tilde{\mathcal{G}}_x, \tilde{\mathcal{G}}_z$  and  $\mathcal{B}$  that are defined as follows. For a three-component spin configuration  $\mathbf{S} = (\mathbf{S}_r)$  with  $\mathbf{S}_r = (S_r^x, S_r^y, S_r^z)$ , let  $\mathbf{S}^{xz}$  denote its projection onto the  $xz$ -plane *scaled* to have a unit length. For a  $B$ -block event  $\mathcal{A} \subset (\mathcal{S}_2)^{\mathbb{T}_N}$ , let

$$\tilde{\mathcal{A}} := \{ \mathbf{S} : \mathbf{S}^{xz} \in \mathcal{A} \} \cap \bigcap_{r \in \Lambda_B} \{ \mathbf{S} : |S_r^y| \leq \Delta \} \quad (5.15)$$

denote its natural extension to three-component spin configurations. This immediately defines the events  $\tilde{\mathcal{G}}_x, \tilde{\mathcal{G}}_z$ , and  $\tilde{\mathcal{B}}$ . However,  $\tilde{\mathcal{B}}$  does not cover the complement  $\tilde{\mathcal{G}}_x \cup \tilde{\mathcal{G}}_z$ ; for that we will also need the event

$$\mathcal{B}_y := \bigcup_{r \in \Lambda_B} \{ \mathbf{S} : |S_r^y| > \Delta \}. \quad (5.16)$$

Then, obviously,  $\mathcal{B} := (\tilde{\mathcal{G}}_x \cup \tilde{\mathcal{G}}_z)^c$  satisfies  $\mathcal{B} = \tilde{\mathcal{B}} \cup \mathcal{B}_y$ . In order to plug into our calculations for the two-component spin, we will need to prove the two lemmas.

**Lemma 5.2** *Suppose  $\Delta < 1/2$ . Then there exists a constant  $c_3 < \infty$  such that, for any B-block event  $\mathcal{A}$  such that  $\mathcal{A} \cap \mathcal{B}_E = \emptyset$ , its counterpart  $\tilde{\mathcal{A}}$  defined as above satisfies*

$$\mathfrak{p}_\beta(\tilde{\mathcal{A}}) \leq \left( \pi \Delta^{-2} e^{\beta J c_3 \Delta^3} \right)^{B^2} \frac{\mathfrak{p}_\beta(\mathcal{A})}{1 - \mathfrak{p}_\beta(\mathcal{B}_E)} \quad (5.17)$$

*Proof.* Consider the expectation of the disseminated event  $\tilde{\mathcal{A}}$  in the torus measure  $\tilde{\mu}_{N,\beta}$  for the three-component model. We will write this expectation as  $\tilde{Z}_{N,\beta}(\tilde{\mathcal{A}})/\tilde{Z}_{N,\beta}$ . Pick  $\tilde{\mathcal{S}} \in \tilde{\mathcal{A}}$  and define  $\mathcal{S} := \tilde{\mathcal{S}}^{xz}$ . First, by the very definition  $\mathcal{S}$  belongs to the disseminated event  $\mathcal{A}$ . So in order to compare the expectation for the three-component spins with that for the two-component spins, we will need to control the change in the Hamiltonian and the *a priori* measure under the map  $\tilde{\mathcal{S}} \mapsto \mathcal{S}$ .

To treat both  $\tilde{Z}_{N,\beta}(\tilde{\mathcal{A}})$  and  $\tilde{Z}_{N,\beta}$  in a unified fashion, assume that  $|\tilde{\mathcal{S}}_r^y| \leq \delta$  for some  $\delta \leq \Delta$  and all  $r \in \mathbb{T}_N$ . Let  $\mathcal{S} := \tilde{\mathcal{S}}^{xz}$ . Note that

$$\tilde{\mathcal{S}}_r^\alpha = \frac{S_r^\alpha}{\sqrt{1 - [\tilde{\mathcal{S}}_r^y]^2}} = S_r^\alpha + O(\delta^2), \quad \alpha = x, z. \quad (5.18)$$

From the fact that  $\mathcal{A} \cap \mathcal{B}_E = \emptyset$  it follows that

$$(\tilde{\mathcal{S}}_r^\alpha - \tilde{\mathcal{S}}_{r'}^\alpha)^2 = (S_r^\alpha - S_{r'}^\alpha)^2 + O(\delta^2 \Delta) \quad (5.19)$$

and so we have

$$\mathcal{H}_N(\tilde{\mathcal{S}}) = \mathcal{H}_N(\mathcal{S}) + J \sum_{r \in \mathbb{T}_N} [\tilde{\mathcal{S}}_r^y]^2 + O(\delta^2 \Delta) J |\mathbb{T}_N|. \quad (5.20)$$

The *a priori* measures are related by  $\tilde{\nu}(d\tilde{\mathcal{S}}_r) = \sqrt{1 - [\tilde{\mathcal{S}}_r^y]^2} \nu(dS_r) d\tilde{\mathcal{S}}_r^y$ .

We will now derive bounds on the partition functions  $\tilde{Z}_{N,\beta}(\tilde{\mathcal{A}})$  and  $\tilde{Z}_{N,\beta}$  in terms of their two-component spin counterparts  $Z_{N,\beta}(\mathcal{A})$  and  $Z_{N,\beta}$ . For  $\tilde{Z}_{N,\beta}(\tilde{\mathcal{A}})$  we set  $\delta := \Delta$  and note that  $\mathcal{H}_N(\tilde{\mathcal{S}}) \geq \mathcal{H}_N(\mathcal{S}) + O(\Delta^3) \beta J |\mathbb{T}_N|$ . Integrating out the components  $\tilde{\mathcal{S}}_r^y$  gives

$$\tilde{Z}_{N,\beta}(\tilde{\mathcal{A}}) \leq Z_{N,\beta}(\mathcal{A}) (\pi e^{O(\Delta^3) \beta J})^{|\mathbb{T}_N|}. \quad (5.21)$$

For a lower bound on  $\tilde{Z}_{N,\beta}$ , we first note that  $\tilde{Z}_{N,\beta} \geq \tilde{Z}_{N,\beta}(\mathcal{B}_E^c \cap \mathcal{G}'_y)$ , where  $\mathcal{G}'_y$  is the event

$$\mathcal{G}'_y = \bigcap_{r \in \Lambda_B} \{\mathcal{S} : |S_r^y| < \delta\} \quad (5.22)$$

with  $\delta := \Delta^2$ . Applying (5.20), the Hamiltonians now differ by  $O(\Delta^3) J |\mathbb{T}_N|$  and the integral over  $\tilde{\mathcal{S}}_r^y, r \in \mathbb{T}_N$ , now yields a term  $\Delta^2$  per site. This shows

$$\tilde{Z}_{N,\beta} \geq (\Delta^2 e^{O(\Delta^3) \beta J})^{|\mathbb{T}_N|} Z_{N,\beta}(\mathcal{B}_E^c). \quad (5.23)$$

Combining the upper and lower bounds and applying the subadditivity bound  $\mathfrak{p}_\beta(\mathcal{B}_E^c) \geq 1 - \mathfrak{p}_\beta(\mathcal{B}_E)$  we get the desired claim.  $\square$

**Lemma 5.3** *There are constants  $c_4, c_5 \in (0, \infty)$  such that*

$$\max\{\mathfrak{p}_\beta(\widetilde{\mathcal{B}}_E), \mathfrak{p}_\beta(\mathcal{B}_y)\} \leq \left(c_4 \Delta^{-2} e^{\beta J c_3 \Delta^3}\right)^{B^2} \frac{B^2 e^{-c_5 \beta J \Delta^2}}{1 - \mathfrak{p}_\beta(\mathcal{B}_E)} \quad (5.24)$$

*Proof (Sketch).* We will instead estimate the objects  $\mathfrak{p}_\beta(\widetilde{\mathcal{B}}_E \setminus \mathcal{B}_y)$  and  $\mathfrak{p}_\beta(\mathcal{B}_y \setminus \widetilde{\mathcal{B}}_E)$ ; from these the claim will follow by invoking subadditivity  $\mathcal{A} \mapsto \mathfrak{p}_\beta(\mathcal{A})$ . Consider the partition function  $\widetilde{Z}_{N,\beta}(\widetilde{\mathcal{B}}_E \setminus \mathcal{B}_y)$  and let  $\widetilde{\mathcal{S}}$  be a configuration contributing to the sum. As the second component of all spins are small, the partition function is bounded as in the two-component spin case. The result is

$$\widetilde{Z}_{N,\beta}(\widetilde{\mathcal{B}}_E \setminus \mathcal{B}_y) \leq e^{-\beta J |\mathbb{T}_N|} [2\sqrt{2\pi}]^{|\mathbb{T}_N|} [B^2 e^{-\beta J \Delta^2}]^{(N/B)^2}. \quad (5.25)$$

Combining this with the lower bound on  $Z_{N,\beta}$  from the previous proof now gives the desired bound for  $\mathfrak{p}_\beta(\widetilde{\mathcal{B}}_E \setminus \mathcal{B}_y)$ .

Concerning the event  $\mathcal{B}_y \setminus \widetilde{\mathcal{B}}_E$ , as this is disjoint from  $\mathcal{B}_E$ , we just note that, in this case, that we can bound

$$\mathcal{H}_N(\widetilde{\mathcal{S}}) \geq \mathcal{H}_N(\mathcal{S}) + J \Delta^2 \left(\frac{N}{B}\right)^2 + O(\Delta^3) J |\mathbb{T}_N| \quad (5.26)$$

and then proceed as in the previous lemma. The desired suppression now comes from the second term on the right hand side of the last display.  $\square$

We are now ready to prove our main result on quantum systems with large spins:

*Proof of Theorem 2.5.* Key to the formalism of [5] is the operator  $\hat{Q}_A$  associated with the event  $\mathcal{A}$  as follows:

$$\hat{Q}_A := \left(\frac{2\mathcal{S} + 1}{4\pi}\right)^{\mathbb{T}_N} \int_A d\Omega |\Omega\rangle \langle \Omega|. \quad (5.27)$$

Notice that if  $\mathcal{A}$  is a  $B$ -block event, then  $\hat{Q}_A$  behaves as identity on the part of the Hilbert space outside  $\Lambda_B$ . The map  $\mathcal{A} \mapsto \hat{Q}_A$  is countably additive; in particular, if  $\mathcal{A}_1, \dots, \mathcal{A}_n$  form a partition of the probability space, then  $\hat{Q}_{\mathcal{A}_1} + \dots + \hat{Q}_{\mathcal{A}_n} = \mathbb{1}$ .

Consider now the operators  $\hat{Q}_{\widetilde{\mathcal{G}}_x}$ ,  $\hat{Q}_{\widetilde{\mathcal{G}}_z}$  and  $\hat{Q}_B$ . Notice that these operators are invariant under reflection of the box through any mid-plane (because so are the events they arise from). Let  $\vartheta_t$  denote the shift on  $\mathbb{T}_N$  by  $Bt$ , where  $t \in \mathbb{T}_{N/B}$ . Let  $\xi := \frac{c}{\mathcal{S}}$ , where  $c$  is as in Lemma 5.1. Since  $\beta \leq c_1 \sqrt{\mathcal{S}}$  and  $\mathfrak{p}_\beta(\mathcal{B}) e^{\xi + c_2 \beta / \sqrt{\mathcal{S}}}$  is small — of course, for  $\beta$  large, thanks to Lemmas 5.2 and 5.3, Proposition 4.6 and the choices (4.83) — for some absolute constant  $c_1, c_2 \in (0, \infty)$ , Theorem 3.7 of [5] tells us that for some  $\epsilon > 0$  small

$$\langle \hat{Q}_B \rangle_{N,\beta} < \epsilon \quad (5.28)$$

and

$$\langle \hat{Q}_{\theta_{t_1} \tilde{\mathcal{G}}_x} (1 - \hat{Q}_{\theta_{t_2} \tilde{\mathcal{G}}_z}) \rangle_{N,\beta} < \epsilon. \quad (5.29)$$

for any  $t_1, t_2 \in \mathbb{T}_{N/B}$ .

In order to process these to the bounds in the statement of the theorem, let us note that, if  $\mathcal{O}$  is a bounded operator generated by  $\hat{S}_r^\alpha$  with  $\alpha = x, y, z$ , and  $r \in (tB + \Lambda_B) \cup (t'B + \Lambda_B)$  and  $[\mathcal{O}]_\Omega$  is its upper index for which we set

$$\gamma := \sup_{\Omega} |[\mathcal{O}]_\Omega| \quad (5.30)$$

and

$$\gamma_\alpha := \sup \left\{ |[\mathcal{O}]_\Omega| : \Omega \in \theta_t \mathcal{G}_\alpha \cup \theta_{t'} \mathcal{G}_\alpha \right\}, \quad \alpha = x, z, \quad (5.31)$$

then (5.29) and (5.31) imply

$$\langle \mathcal{O} \rangle_{N,\beta} \leq 3\epsilon\gamma + \max\{\gamma_x, \gamma_z\}. \quad (5.32)$$

Indeed, writing

$$\langle \mathcal{O} \rangle_{N,\beta} = \frac{1}{Z_{N,\beta}} \left( \frac{2\mathcal{S} + 1}{4\pi} \right)^{|\mathbb{T}_N|} \int_{(\mathcal{S}_2)^\Lambda} d\Omega [\mathcal{O}]_\Omega \langle \Omega | e^{-\beta \hat{H}_N} | \Omega \rangle, \quad (5.33)$$

we can estimate  $[\mathcal{O}]_\Omega$  by  $\gamma_\alpha$  on  $\theta_t \mathcal{G}_\alpha \cup \theta_{t'} \mathcal{G}_\alpha$ , and by  $\gamma$  on the complement of these. The positivity of  $\langle \Omega | e^{-\beta \hat{H}_N} | \Omega \rangle$  permits us to convert the resulting integrals to expectations of the kind (5.28–5.29).

To see how this applies in a specific situation, consider  $\mathcal{O} := \mathcal{S}^{-4} [\hat{S}_r^y]^2 [\hat{S}_{r'}^y]^2$  for  $r, r'$  belonging to distinct translates of  $\Lambda_B$ . We then get

$$[\mathcal{O}]_\Omega = [\Omega_r^y]^2 [\Omega_{r'}^y]^2 + O(\mathcal{S}^{-1}), \quad (5.34)$$

which is  $O(\eta + \mathcal{S}^{-1})$  on  $\cup_{\alpha=x,z} (\theta_t \mathcal{G}_\alpha \cup \theta_{t'} \mathcal{G}_\alpha)$  and  $O(1)$  otherwise. As a consequence, for some  $c < \infty$  and  $N$  is sufficiently large,

$$\mathcal{S}^{-4} \left| \langle [\hat{S}_r^y]^2 [\hat{S}_{r'}^y]^2 \rangle_{N,\beta} \right| \leq c(\eta + \mathcal{S}^{-1} + \epsilon). \quad (5.35)$$

The other cases needed to establish (2.19–2.20) are checked analogously. Once (2.19–2.20) are proved, the bound (2.18) follows by the symmetries of the model and the fact that  $\mathcal{S}^{-2} \sum_\alpha [S_r^\alpha]^2 = 1 + O(1/\mathcal{S})$ .  $\square$

*Proof of Theorem 2.6.* Fix  $J > 0$  and throughout this proof let  $\hat{\mathcal{H}}_N$  denote the Hamiltonian for  $J_1 = J_2 = J$ . Consider the operator

$$\hat{\mathcal{E}}_N^\alpha := \mathcal{S}^{-2} \sum_{\substack{\langle r, r' \rangle_\alpha \\ r, r' \in \mathbb{T}_N}} S_r^x S_{r'}^x, \quad \alpha = x, y, z \quad (5.36)$$

and notice that increasing  $J_1$  above the common value  $J$  amounts to adding the term  $(J - J_1)\hat{\mathcal{E}}_N^x$  to  $\hat{\mathcal{H}}_N$ . Key to the proof is to show that

$$f(h) := \lim_{N \rightarrow \infty} \frac{1}{N^2} \log \frac{\text{Tr}(e^{-\beta\hat{\mathcal{H}}_N + h\hat{\mathcal{E}}_N^x})}{\text{Tr}(e^{-\beta\hat{\mathcal{H}}_N})} \quad (5.37)$$

satisfies

$$\frac{d}{dh^+} f(h)|_{h=0} > 2(1 - \delta). \quad (5.38)$$

Indeed,  $h \mapsto f(h)$  is convex (by Hölder inequality) and so (5.38) implies the same bound for all  $h > 0$ . This in turn shows that, for  $N$  sufficiently large and  $r$  with both coordinate even,

$$\mathcal{S}^{-2} \langle \hat{\mathcal{E}}_r \rangle_{N,\beta} \geq 2(1 - \delta) \quad (5.39)$$

where  $\hat{\mathcal{E}}_r$  is quantum counterpart of the plaquette energy (1.5). From here the desired claims follow along the same argument as in the classical case.

We thus have to show (5.38). To this end consider the torus events

$$\mathcal{A}_\alpha := \left\{ \Omega: \sum_{\langle r,r' \rangle_\alpha} \Omega_r^\alpha \Omega_{r'}^\alpha \geq (2 - \delta) |\mathbb{T}_N| \right\}, \quad \alpha = x, z. \quad (5.40)$$

A straightforward application of chessboard estimates shows that, for each  $\delta > 0$ , there is  $\beta_1 = \beta_1(\delta)$  such that for  $\beta \geq \beta_1$  and  $\mathcal{S} \geq c\beta^2$ , we have

$$\langle \hat{Q}_{\mathcal{A}_\alpha} \rangle_{N,\beta} \geq \frac{1}{2} - \epsilon_N, \quad \alpha = x, z, \quad (5.41)$$

where  $\epsilon_N \rightarrow 0$  as  $N \rightarrow \infty$ . Next we note that, by Jensen's inequality

$$\text{Tr}(e^{-\beta\hat{\mathcal{H}}_N + h\hat{\mathcal{E}}_N^\alpha}) \geq \left( \frac{2\mathcal{S} + 1}{4\pi} \right)^{|\mathbb{T}_N|} \int_{\mathcal{A}_\alpha} d\Omega e^{-\beta\langle \hat{\mathcal{H}}_N \rangle_\Omega + h(2-\delta)|\mathbb{T}_N|}, \quad (5.42)$$

where we already applied that  $\langle \hat{\mathcal{E}}_N^\alpha \rangle_\Omega \geq (2 - \delta)|\mathbb{T}_N|$  for  $\Omega \in \mathcal{A}_\alpha$ . Theorem 3.1 of [5] now shows that

$$e^{-\beta\langle \hat{\mathcal{H}}_N \rangle_\Omega} \geq \langle \Omega | e^{-\beta\hat{\mathcal{H}}_N} | \Omega \rangle e^{-c\beta|\mathbb{T}_N|/\sqrt{\mathcal{S}}} \quad (5.43)$$

for some  $c < \infty$ . Substituting this into (5.42) yields

$$\frac{\text{Tr}(e^{-\beta\hat{\mathcal{H}}_N + h\hat{\mathcal{E}}_N^\alpha})}{\text{Tr}(e^{-\beta\hat{\mathcal{H}}_N})} \geq e^{h(2-\delta)|\mathbb{T}_N| - c\beta|\mathbb{T}_N|/\sqrt{\mathcal{S}}} \langle \hat{Q}_{\mathcal{A}_\alpha} \rangle_{N,\beta}. \quad (5.44)$$

Setting  $\alpha := x$  and applying (5.41), the bound (5.38) follows once  $c\beta/\sqrt{\mathcal{S}} < \delta$ .  $\square$

## 6. Concluding remarks

### 6.1. Néel vs orientational order

As already mentioned, Wenzel and Janke determined in their numerical experiments that the model exhibits a Néel ordering of the plaquette energies; see Figs. 2(a-b) of [37]. Explicitly, energy was found low on the  $z$ -plaquettes and high on the  $x$ -plaquettes in a sample obtained by multiple updates of a configuration using Glauber dynamics with the Metropolis rule. We wish to point out that this ordering is a direct consequence of the ORLO established rigorously in the present work. Indeed, if the spins are with high probability aligned with  $e_2$  — i.e., pointing north or south — then the  $z$ -plaquettes will have energy  $\mathcal{E}_r \approx -4$  while the  $x$ -plaquettes will have energy  $\mathcal{E}_r \approx 0$ . In addition, the mixed plaquettes will settle at energy  $\mathcal{E}_r \approx -2$ . This is consistent with Fig. 2(d) of [37].

The state with the spins aligned with  $e_1$  will have the roles of the  $x$  and  $z$ -plaquettes interchanged, giving the energy distribution again a Néel type order resemblance. Notwithstanding, the physical significance of a Néel order is unclear given the period-2 nature of the interaction. And, in fact, matters seem to look quite different when instead of  $\mathcal{E}_r$  we work with more natural quantity,

$$\begin{aligned} \tilde{\mathcal{E}}_r(\mathbf{S}) := & [S_r^{\alpha_1} - S_{r+e_1}^{\alpha_1}]^2 + [S_r^{\alpha_2} - S_{r+e_2}^{\alpha_2}]^2 \\ & + [S_{r+e_1}^{\alpha_3} - S_{r+e_1+e_2}^{\alpha_3}]^2 + [S_{r+e_2}^{\alpha_4} - S_{r+e_1+e_2}^{\alpha_4}]^2, \end{aligned} \quad (6.1)$$

which is the plaquette energy normalized to *vanish* in all ground states. Although we do not see any reason why a *strong* Néel ordering should be exhibited by these plaquette energies, it would be perhaps of some interest to rerun the numerical experiments at higher precision to check this fact numerically.

### 6.2. Enhanced sampling

The samples of actual configurations shown in Fig.2(d) of [37] deserve one more comment. A cursory look at the figure reveals some level of orientational order in the  $e_2$ -spin direction — which is consistent with our mathematical results — but a more careful analysis uncovers an apparent statistical discrepancy. Indeed, most of the spins point down in the figure although the plaquette-flip symmetries ( $\varphi_r$ ) of the Hamiltonian, which can be applied independently at all even locations, indicate that about half of the  $z$ -plaquettes should be pointing up! It is thus somewhat surprising that in Fig.2(d) of [37] only 7 such plaquettes out of the total of 25 have upward-pointing spins; the others are clearly pointing down. As the number of upward-pointing plaquettes is well approximated by a binomial distribution, the probability that this happens is

$$\binom{25}{7} \frac{1}{2^{25}} \approx 0.014, \quad (6.2)$$

i.e., the configuration in the figure will typically appear only once in about 70 samples! We take this as a possible indication that, despite judicious methods of simulations, the configuration may not have fully equilibrated at the time the snapshot was taken.

Turning this observation into a positive statement, one can try to use the plaquette-spin flips to accelerate the convergence of the computer sample to equilibrium. Notice that, even at moderate temperatures, any single-spin update rule will have considerable difficulties to overcome the energy barrier associated with changing the orientation of an entire plaquette. The dynamics would naturally mix faster if an occasional flip of an entire plaquette — by an application of one of the maps  $(\varphi_r)$  — is incorporated into the stochastic dynamics. This would result in an algorithm reminiscent of the Swendsen-Wang method for sampling configurations in the  $q$ -state Potts model [35]. It could be expected that this enhancement would result in a substantially better performance of the simulations. An algorithm of this sort has been recently attempted in the context of the orbital compass model [38], although there the cluster flips have to be performed along entire lines of sites which makes them very non-local.

### 6.3. Correlation decay

Our mathematical argument establishes rigorously long-range order in the system. Nonetheless, we do so without giving any bound on the decay of (truncated) correlations. Since our argument is based on contour methods and suppression of long-wavelength part of the spin-wave decomposition, we tend to believe that the correlations generally decay exponentially fast in any translation-invariant, ergodic Gibbs measure for this system. However, we have not been able to find a rigorous argument in the vain. Again, it would be of interest to see if this question could be addressed by numerical methods.

## Acknowledgments

The research of M.B. was partially supported by the NSF grant DMS-0949250. The research of R.K. was partially supported by the grants GAČR 201-09-1931 and MSM 0021620845.

## References

- [1] M. Biskup, L. Chayes, and S. A. Kivelson, *Order by disorder, without order, in a two-dimensional spin system with  $O(2)$  symmetry*, *Annales Henri Poincaré* **5** (2004), no. 6, 1181–1205.
- [2] M. Biskup, *Reflection Positivity and Phase Transitions in Lattice Spin Models*. In: R. Kotecký (ed.), “Methods of Contemporary Mathematical Statistical Physics,” *Lecture Notes in Mathematics*, vol. **1970**. Springer-Verlag, 2009, pp. 1–86.

- [3] M. Biskup, L. Chayes, and R. Kotecký, *Coexistence of partially disordered/ordered phases in an extended Potts model*, J. Statist. Phys. **99** (2000), no. 5-6, 1169–1206.
- [4] M. Biskup, L. Chayes, and Z. Nussinov, *Orbital ordering in transition-metal compounds: I. the 120-degree model*, Commun. Math. Phys. **255** (2005), no. 2, 253–292.
- [5] M. Biskup, L. Chayes, and S. Starr, *Quantum spin systems at positive temperature*, Commun. Math. Phys. **269** (2007), no. 2, 611–657.
- [6] M. Biskup M and R. Kotecký, *Forbidden gap argument for phase transitions proved by means of chessboard estimates*, Commun. Math. Phys. (2006) **264**no. 3, 631-656.
- [7] M. Biskup and R. Kotecký, *Phase coexistence of gradient Gibbs states*, Probab. Theory Rel. Fields **139** (2007), no. 1-2, 1–39.
- [8] J. van de Brink, G. Khaliulin, and D. Khomskii, *Orbital effects in manganites*, in “Colossal Magnetoresistive Manganites”. **2002**. Editor T. Chatterij. Kluwer Academic Publishers, Dordrecht; cond-mat/0206053.
- [9] W. Brzezicki and A. M. Oleś, *Compass model on a ladder and square clusters*, J. Phys.: Conf. Ser. **200** (2010), no. 1, 012017.
- [10] L. Chayes, R. Kotecký, and S. B. Shlosman, *Aggregation and intermediate phases in dilute spin systems*, Commun. Math. Phys. **171** (1995), no. 1, 203–232.
- [11] R.L. Dobrushin, *Gibbsian random fields for lattice systems with pairwise interactions*, Funct. Anal. Appl. **2** (1968) 292–301.
- [12] R.L. Dobrushin, *Gibbsian random fields. The general case*, Funct. Anal. Appl. **3** (1969) 22–28.
- [13] J. Dorier, F. Becca, and F. Mila, *Quantum compass model on the square lattice*, Phys. Rev. B **72** (2005), no. 2, 24448.
- [14] F. J. Dyson, E. H. Lieb, and B. Simon, *Phase transitions in quantum spin systems with isotropic and nonisotropic interactions*, J. Statist. Phys. **18** (1978) 335–383.
- [15] S. Dusuel, K. P. Schmidt, and J. Vidal, *Creation and manipulation of anyons in the Kitaev model* Phys. Rev. Lett. **100** (2008), no. 17, 177204.
- [16] J. Fröhlich, R. B. 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, 134.
- [17] J. Fröhlich, R. B. 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.
- [18] J. Fröhlich, B. Simon, and T. Spencer, *Infrared bounds, phase transitions and continuous symmetry breaking*, Commun. Math. Phys. **50** (1976), no. 1, 7995.
- [19] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter Studies in Mathematics, vol. 9, Walter de Gruyter & Co., Berlin, 1988.
- [20] J. B. Goodenough, *Theory of the role of covalence in the perovskite-type manganites  $[La,M(II)]MnO_3$* , Phys. Rev. **100** (1955), 564.
- [21] O. J. Heilmann and E. H. Lieb, *Lattice models for liquid crystals*, Jour. Statist. Phys. **20** (1979), no. 6, 679–693.
- [22] H. A. Jahn and E. Teller, *Stability of polyatomic molecules in degenerate electronic states. I. Orbital degeneracy*, Proc. Roy. Soc. Series A, Math. and Phys. Sciences **161** (1937), no. 905, 220–235.
- [23] L. Jiang, G.K. Brennen, A.V. Gorshkov, K. Hammerer, M. Hafezi, D. Demler, M.D. Lukin and P. Zoller, *Anyonic interferometry and protected memories in atomic spin lattices*, Nature Phys. **4** (2008), no. 6, 482–488.
- [24] G. Khaliullin, *Order from disorder: Quantum spin gap in magnon spectra of  $LaTiO_3$* , Phys. Rev. B **64** (2003), 212405.
- [25] D. I. Khomskii and M. V. Mostovoy, *Orbital ordering and frustrations*, J. Phys. A: Math. Gen. **36** (2003), no. 35, 9197–9208.
- [26] K. I. Kugel’ and B. I. Khomskii, *Heisenberg model in a magnetic field, and metamagnetism of Jahn-Teller systems*, JETP Letters **23** (1976), no 5, 237-240.
- [27] A. Kitaev, *Anyons in an exactly solved model and beyond*, Ann. Phys. **321** (2006), no. 1, 2–111.
- [28] A. Kitaev, *Fault-tolerant quantum computation by anyons*, Ann. Phys. **303** (2003), no. 1, 2–30.

- [29] R. Kotecký and S. B. Shlosman, *First-order phase transitions in large entropy lattice models*, Commun. Math. Phys. **83** (1982), no. 4, 493–515.
- [30] O.E. Lanford and D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Commun. Math. Phys. **13** (1969) 194–215.
- [31] A. Mishra, M. Ma, F.C. Zhang, S. Guertler, L. H. Tang, and S. Wan, *Directional ordering of fluctuations in a two-dimensional compass model*, Phys. Rev. Lett. **93** (2004), no. 20, 207201.
- [32] Z. Nussinov, M. Biskup, L. Chayes, and J. van den Brink, *Orbital order in classical models of transition-metal compounds*, Europhys. Lett. **67** (2004), 990–996.
- [33] Z. Nussinov and E. Fradkin, *Discrete sliding symmetries, dualities, and self-dualities of quantum orbital compass models and  $p + ip$  superconducting arrays*, Phys. Rev. B **71** (2005), no. 19, 195120.
- [34] R. Orús, A. Doherty, and G. Vidal, *First order phase transition in the anisotropic quantum orbital compass model*, Phys. Rev. Lett. **102** (2009), no. 7, 077203.
- [35] R.H. Swendsen and J.-S. Wang, *Nonuniversal critical dynamics in Monte Carlo simulations*, Phys. Rev. Lett. **58** (1987), no. 2, 86–88.
- [36] J. van den Brink, G. Khaliullin, and D. Khomskii, *Charge and orbital order in half-doped manganites*, Phys. Rev. Lett. **83** (1999), no. 24, 5118–5121.
- [37] S. Wenzel and W. Janke, *Finite-temperature Néel ordering of fluctuations in a plaquette orbital model*, Phys. Rev. B **80** (2009), no. 5, 054403.
- [38] S. Wenzel, W. Janke W and A.M. Läuchli, *Re-examining the directional-ordering transition in the compass model with screw-periodic boundary conditions*, Phys. Rev. E **81** (2010) 066702.
- [39] W.-L. You, G.-S. Tian, and H.-Q. Lin, *Existence of long-range orbital order in a two-dimensional orbital-only model*, Phys. Rev. B **75** (2007), no. 19, 195118.
- [40] V. A. Zagrebnov, *Long-range order in a lattice-gas model of nematic liquid crystals*, Physica A, Statist. Mech. Appl. **232** (1996), 737–746.