DISCONTINUITY OF THE SPIN–WAVE STIFFNESS IN THE TWO–DIMENSIONAL XY MODEL

L. Chayes
Department of Mathematics
University of California, Los Angeles

Abstract. Using a graphical representation based on the Wolff algorithm, the (classical) d–dimensional XY model and some related spin–systems are studied. It is proved that in $d = 2$, the predicted discontinuity in the spin–wave stiffness indeed occurs. Further, the critical properties of the spin–system are related to percolation properties of the graphical representation. In particular, a suitably defined notion of percolation in the graphical representation is proved to be the necessary and sufficient condition for positivity of the spontaneous magnetization.

Key words and phrases. Kosterlitz–Thouless transition, Wolff representation.

Introduction. Among the most noted early achievements of the renormalization group was the analysis of the defect (vortex) unbinding transition in two–dimensional systems with Abelian symmetries [B], [KT]. The definitive (and experimentally accessible) prediction of this analysis is the occurrence of discontinuities at the edge of the low–temperature phase. Such a phenomenon is remarkable in and of the fact that the transition itself, by any other criterion is continuous. In the language of superfluid systems, the above mentioned discontinuity occurs in the superfluid density; for spin–systems, it is the spin–wave stiffness; sometimes known as the helicity modulus. This prediction has been born out by theoretical, numerical and experimental (and analog/experimental) tests; cf. the review articles [N] and [M] and references therein. In this note, a complete mathematical proof for the (classical) $2d$–XY model is provided.

The method of proof employs the graphical representation – or cluster representation – due to Wolff [W]. (More precisely, the graphical representation that is implicit in the Wolff algorithm.) The importance of understanding this representation was stressed in [PS] and this representation was exploited in [A] in the study of the "vortex–free" XY model. In [$\mathrm{CM_{II}}$], critical properties of the spin–system and the graphical representation were shown to be related. Here some characterizations are presented: Up to constant factors constant factors the magnetization in the spin–system is equal to the percolation density in the Wolff–representation and the susceptibility is "equal" to the average size of the connected clusters. Of more immediate relevance is the fact that the spin–wave stiffness tested in finite volume is directly related to crossing probabilities in the graphical representation and in particular, a small stiffness implies and is implied by a small crossing probability. If this probability is "too small" then, using elementary rescaling ideas borrowed from rigorous percolation theory, it tends to zero exponentially at larger scales (which furthermore implies exponential decay of correlations). Thus, the stiffness is either uniformly positive at all scales or it is zero. The existence of a low temperature phase with power law decay of correlations (proved in [FS]) thus implies a discontinuity of the stiffness at a positive temperature. A related class of problem – in the sense that the RG equations turn out to be nearly identical – are the one dimensional long–range discrete models, e.g. $1/r^2$ Ising model. In this context, the magnetization at the critical point plays the role of the Spin–Wave stiffness and it was predicted in [T] to be discontinuous at T_c (the Thouless effect). This was rigorously established in [ACCN] by vaguely similar methods: graphical representations and "real space renormalization group" inequalities. However, in the rigorous as well as in the renormalization group arenas the deeper relationship between these two problems is still unclear.

The remainder of this paper is organized along the following lines: Below, the definition of the spin–wave stiffness used in this note is provided. In the next section, the Wulff representation is developed. Here, the key relationship between the spin–wave stiffness and appropriate crossing probabilities is derived. This will be followed by the section in which the main result – the discontinuity of the spin–wave stiffness in $d = 2$ – is established. In the final section, some auxiliary results will be stated (but not proved) and in the appendix, complete proofs of these results and various properties of the Wolff representation will be provided.

Spin–Wave Stiffness. The spin–wave stiffness is the appropriate notion of a leading correction to the bulk free energy when the surface tension is zero. It may be defined as follows: Consider a regular finite volume d– dimensional shape V with two (separated) boundary components. Let V_L denote the lattice approximation to this shape at scale L i.e. the intersection of $\mathbb{Z}^d$ with the image of V that has been uniformly scaled by a factor of L. The general strategy is to consider the difference in free energies of the system with uniform boundary conditions and twisted boundary conditions on V_L. For typical ferromagnetic spin–systems, "uniform" means that all the boundary spin are aligned and "twisted" means that the two boundary components are individually aligned but are anti–parallel. For the purposes of this note, the above is sufficient. In more generality, one may consider cylindrical or even toroidal geometries which, in other contexts, are arguably a better choice. C.f. the discussion in [FJB]. Modulo constants, for $L \gg 1$, the log of the ratio of the twisted and uniform partition functions serves to "define" the spin–wave stiffness K. Let us proceed more cautiously and define this ratio as $e^{-\beta K_L(V,\beta)g(V)L^{d-2}}$ with β the inverse temperature and $g(V)$ a geometric constant (which is essentially the capacitance) to be described below. A spin–wave stiffness may be defined via the limiting behavior of $K_L(V,\beta)$; since there is no general proof that the limit exists, let alone is independent of V, the matter will be left as it stands. Suffice it to say that if for any V of a roughly annular shape, $K_L(V,\beta)$ tends to zero then all possible K_L's tend to zero (And similarly, in $d > 2$, if any $K_L(V,\beta)L^{d-2} \to 0$, then they all do.)

Let us tend to the constant $g(V)$. The models under consideration will have spins with bounded values in $\mathbb{R}^2$; let us assume that the bound is one. Furthermore (and here rather vaguely) let us assume that if the Hamiltonian is expressed in "deviation" variables, the leading non–constant term is quadratic with coefficient 1/2. Let ϕ_V be the solution to Laplaces' equation with boundary values ± 1 on the two components. Then

$$g = \int_V |\nabla \phi_V|^2 d^d x. \tag{1}$$

With this definition, it is an elementary exercise to show, for the standard XY model on $\mathbb{Z}^d$ (e.g. as defined in Equation (3.a) with unit couplings between neighboring sites) that

$$\lim_{L\to\infty}\lim_{\beta\to\infty} K_L(V,\beta) = 1. \tag{2}$$

In this paper, all that is needed is the simplest of annular shapes: Consider, in $d = 2$, the square of size 3, $S_{(3)} = \{x_1, x_2 \mid -\frac{3}{2} \le x_1 \le +\frac{3}{2}, -\frac{3}{2} \le x_2 \le +\frac{3}{2}\}$ and $S_{(1)}$ defined accordingly. The shape of interest is $A \equiv S_{(3)} \setminus S_{(1)}$. In $d > 2$ the corresponding generalization is used: a hypercube of side 3 with the central hypercube of side 1 removed.

The Representation: Notation and Definitions. Although the primarily concern is with the behavior of uniform systems on regular d–dimensional lattices, the cluster representation is just as easily formulated on an arbitrary (finite) graph. Indeed, there is a need for these sorts of generalities in order to formulate the representation of these systems in the presence of boundary conditions. Thus, let $\mathcal{G}$ denote finite graph with sites $\mathbb{S}_{\mathcal{G}}$ and bonds $\mathbb{B}_{\mathcal{G}}$. For each $i \in \mathbb{S}_{\mathcal{G}}$, let $\vec{s}_i$ denote a $2d$ spin of length one and for each $\langle i,j\rangle \in \mathbb{B}_{\mathcal{G}}$, let $J_{i,j} > 0$ denote the couplings. The XY–Hamiltonian is given by

$$H^{XY}_{\mathcal{G}} = -\sum_{\langle i,j\rangle} J_{i,j}\vec{s}_i\cdot\vec{s}_j. \tag{3.a}$$

Writing a_i and b_i for the magnitude of the Y and X components respectively, (here $0 \le a_i, b_i \le 1$) and allowing $\tau_i = \pm 1$ and $\sigma_i = \pm 1$, $H^{XY}_{\mathcal{G}}$ may be read

$$H^{XY}_{\mathcal{G}} = -\sum_{\langle i,j\rangle} J_{i,j}[a_ia_j\tau_i\tau_j + b_ib_j\sigma_i\sigma_j]. \tag{3.b}$$

For most of what remains, we will have little use for the specifics of the XY–model itself. Indeed, we might just as well allow the right hand side of Equation (3.b) to define the model along with some constraint on the (a_i, b_i) that makes one a decreasing function of the other and an *a priori* distribution, f_i, for the b_i (which need not be continuous). For the purposes of brevity we will, however assume complete symmetry between the a's and the b's and that these objects are bounded.

The idea behind the Wolff representation is to develop one (or both) of the Ising systems in an FK [FK] random cluster representation.[1] The partition is given by the usual

$$Z(\mathcal{G}, \underline{J}, \beta) = \sum_{\underline{\sigma},\ \underline{\tau}} \int \prod_i df_i(b_i) e^{\beta \sum_{\langle i,j \rangle} J_{i,j}[a_i a_j \tau_i \tau_j + b_i b_j \sigma_i \sigma_j]}. \tag{4}$$

In the above, $\underline{\sigma}$ and $\underline{\tau}$ are notation for the Ising configurations on $\mathcal{G}$ while $\underline{J}$ denotes the collection of couplings. And similarly, $\underline{a}$ and $\underline{b}$ will be notation for configurations of the magnitude of the spin components with the a_i understood to be a function of the b_i.

Let us start by writing the Ising portion of the Hamiltonian in Potts form: $\sigma_i \sigma_j = 2\delta_{\sigma_i \sigma_j} - 1$, etc. For fixed $\underline{b}$, let us trace over the $\underline{\tau}$ variables and then trade the $\underline{\sigma}$ degrees of freedom for those of an FK expansion. Thus let $Z^I_{\underline{a}}(\beta)$ denote the Ising partition function according to an Ising Hamiltonian written in Potts form:

$$H^I_{\underline{a}} = -\sum_{\langle i,j \rangle} J_{i,j}\ a_i a_j (\delta_{\tau_i \tau_j} - 1) \tag{5.a}$$

$$Z^I_{\underline{a}}(\beta) = \sum_{\underline{\tau}} e^{-\beta H^I_{\underline{a}}}. \tag{5.b}$$

Here, the dependence of these quantities on $\mathcal{G}$, and the $(\underline{J})$ has been temporarily suppressed. Unfortunately, the relevant β is twice what appears in Equation (5.b) so to avoid confusion, this parameter will stay with us. Performing the afore mentioned trace and expansion, we

[1] In typical simulations one does this for only one of the Ising variables – as will most often be the case here – but picking a direction at random. However, as argued in [CM$_{\text{II}}$], it may be advantageous to use the full expansion in conjunction with the *Invaded Cluster* algorithm.

arrive at the weights (or density function) of a joint distribution for the $\underline{b}$ and bond configurations $\omega \subset \mathbb{B}_{\mathcal{G}}$:

$$V_{\beta}^{W}(\underline{b}, \omega) = Z_{\underline{a}}^{I}(2\beta) \prod_{\langle i,j \rangle} e^{\beta J_{i,j}(a_i a_j + b_i b_j)} \; W_{\underline{b};2\beta}(\omega) \tag{6}$$

where $W_{\underline{b};2\beta}(\omega)$ are the usual $(q = 2)$ FK weights with couplings $J_{i,j} b_i b_j$ and inverse temperature 2β:

$$W_{\underline{b};2\beta}(\omega) = q^{C(\omega)} \prod_{\langle i,j \rangle \in \omega} p_{i,j} \prod_{\langle i,j \rangle \notin \omega} (1 - p_{i,j}), \tag{7}$$

$p_{i,j} = 1 - e^{2\beta J_{i,j} b_i b_j}$ and $C(\omega)$ the number of connected components of ω. The measures defined by the weights in Equation (6) will be denoted by $\nu_{\beta}^{W}(-)$

Let us consider the two marginal distributions: (i) Integrate out the $\underline{b}$ degrees of freedom to obtain a measure on the bond configurations ω. These will be denoted by $\mathbb{M}_{\beta}(-)$ – or $\mathbb{M}_{\beta,\mathcal{G}\dots}^{*}(-)$, with $*$ signifying possible boundary conditions to be discussed later. (ii) Integrate out the ω degrees of freedom (i.e. skip the FK step and trace the $\underline{\sigma}$ variables). The associated density will be denoted by $\rho_{\beta}(-)$ – or $\rho_{\beta,\mathcal{G}\dots}^{*}(-)$ when the need arises. Finally, let us consider the conditional FK measures, $\mu_{\underline{b}}^{FK}(-)$ determined by the weights in Equation (7). These distributions allow for a convenient decomposition of $\mathbb{M}_{\beta}(-)$

$$\mathbb{M}_{\beta}(-) = \int_{\underline{b}} d\rho_{\beta}(\underline{b}) \mu_{\underline{b}}^{FK}(-). \tag{8}$$

Some immediate applications of these measures have been discussed in [A] and [$\mathrm{CM_{II}}$]. For example, in the usual isotropic XY case, if $T_{i,j}$ is the (bond) event that i is connected to j then, e.g. in free boundary conditions,

$$2\mathbb{M}_{\beta,\mathcal{G}}(T_{i,j}) \geq \langle \vec{s}_i \cdot \vec{s}_j \rangle_{\beta,\mathcal{G}} \tag{9}$$

with $\langle - \rangle_{\beta,\mathcal{G}}$ denoting expectation with respect to the canonical distribution. This has been supplemented by a lower bound proportional to a power of $\mathbb{M}_{\beta,\mathcal{G}}(T_{i,j})$. Here we will obtain a lower bound of a constant times $\mathbb{M}_{\beta,\mathcal{G}}(T_{i,j})$. Of direct relevance to the present work is the following:

Let $K_L(A,\beta)$ denote the spin wave stiffness as discussed in the introduction. Explicitly, let $Z^{\imath^+ o^+}(A_L,\beta)$ denote the partition function on the annulus A_L with boundary conditions obtained by setting all boundary spins on the inner boundary ($\imath$) and the outer boundary (o) to the X–direction. (Or, in the language of Equation (3.b), all the b_i's are set to their maximum values and $\sigma_i \equiv 1$ on the boundary.) Similarly let $Z^{\imath^- o^+}(A_L,\beta)$ be the partition function for the setup in A_L where the spins on the outer boundary are pointing in the positive X–direction and the spins on the inner boundary pointing in the negative X–direction. Thus

$$e^{-\beta g(A) K_L(A,\beta) L^{d-2}} \equiv Z^{\imath^- o^+}(A_L,\beta)/Z^{\imath^+ o^+}(A_L,\beta).$$

Concerning the "$\imath^+ o^+$" system, it is clear that we can treat this setup along the lines already described: the boundary spins act as a single spin albeit with a concentrated distribution. Let us denote by $\mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_L}(-)$ the bond measure associated with these boundary conditions and let $T_{\imath,o}$ denote the event of a connection between the inner and outer boundaries of A_L. The first claim is

Proposition 1.

$$1 - Z^{\imath^- o^+}(A_L,\beta)/Z^{\imath^+ o^+}(A_L,\beta) = \mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_L}(T_{\imath,o}).$$

In particular, the spin–wave stiffness is related in a simple way to the probability of a connection between the boundary components of A_L.

Proof. As is well known, in random cluster measures corresponding to Potts systems with spins on the boundary set to some fixed value, the weights for the graphical configurations are given by the standard one with the interpretation that $C(\omega)$ counts only the components that are disconnected from the boundary. (Equivalently, up to an irrelevant constant, one counts all the sites that are attached to the boundary as part of the *same* component.) Thus if we write

$$Z^{\imath^+ o^+}(A_L, \beta) = \sum_{\omega} \int_{\underline{b}} dV_{\beta}^{W,\mathbf{1}^{++}}(\underline{b}, \omega), \tag{10}$$

the sum contains terms both with and without connections between the boundary. On the other hand, in an situation where two separate boundary components in the Potts system are set to different values, the rule for counting clusters is the same but now bond configurations containing connections between these components are assigned zero weight. Thus for fixed $\underline{b}$, the formula for the Wolff weights $V_{\beta}^{W,\mathbf{1}^{+-}}(\underline{b}, \omega)$ corresponding to the twisted boundary condition is seen to be identical except for the proviso that ω does not connect $\imath$ with o – and here these configurations are discounted. The desired result is established. □

It is plausible that these measures enjoy various monotonicity properties but in any case, this will not be easy to prove. In particular it turns out that the joint measure is not strong FKG. What can be proved is that for a certain class of boundary conditions – that are called the $\odot$–boundary conditions – the ρ–measures *do* have the FKG property. The precise definition of a $\odot$–boundary condition is somewhat intricate but this class includes every boundary condition of physical interest where one could expect the FKG property to hold e.g. free, periodic and setting all the boundary spins to the positive X–direction. Furthermore, among all boundary specifications in the $\odot$–class, this latter mentioned is *maximal* in the sense of FKG. The same dominance therefore holds over the $\overline{\odot}$–class of specifications which is defined as superpositions of specifications from the $\odot$–class. This larger class has the property that its restrictions to smaller sets are also in the $\overline{\odot}$–class relative to the "larger" boundary component. The relevant consequences of the above is summarized in the form of a Lemma:

Lemma 2. *Let $\mathcal{G}$ denote a graph. Then for every $\mathbb{L} \subset \mathbb{S}_{\mathcal{G}}$, there is a class of specifications on $\mathbb{L}$ called the $\overline{\odot}$–class such that: (1) If $\mathbb{K} \supset \mathbb{L}$ and $*$ is a $\overline{\odot}$–specification on $\mathbb{L}$ then the restriction of the various measures, $\nu_{\beta}^{W,*}(-)$, $\mathbb{M}^{*}_{\beta,\mathcal{G}}(-)$, etc. to the complement of $\mathbb{K}$ is itself a $\overline{\odot}$–class specification on $\mathbb{K}$. (2) Setting all spins of $\mathbb{L}$ to the X–direction constitutes a $\overline{\odot}$–class specification on $\mathbb{L}$; this is denoted by the $\mathbf{1}^{+}$ boundary conditions on $\mathbb{L}$. If $*$ is any other $\overline{\odot}$–specification on $\mathbb{L}$ then*

$$\mathbb{M}^{\mathbf{1}^{+}}_{\beta,\mathcal{G}}(-) \underset{\mathrm{FKG}}{\geq} \mathbb{M}^{*}_{\beta,\mathcal{G}}(-).$$

A proof (including relevant definitions) will be supplied in the appendix. The important point is that among all possible relevant boundary conditions, on A_L, the one that maximizes the probability of $T_{\imath,o}$ is precisely $\mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_L}(-)$.

Main Results. With the identity of Proposition 1 and the inequalities of Lemma 2, the main argument reduces to a standard routine in percolation theory:

Theorem 3. *There is an $\epsilon_0 = \epsilon_0(d)$ such that if for any L_0, $\mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_{L_0}}(T_{\imath,o}(L_0)) < \epsilon_0$ then*

$$\lim_{L\to\infty} \mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_L}(T_{\imath,o}(L)) = 0.$$

In particular, under these conditions, $\mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_L}(T_{\imath,o})$ tends to zero exponentially fast in L.

Proof. Suppose that $\mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_{L_0}}(T_{\imath,o}(L_0)) \leq \epsilon < \epsilon_0$ with ϵ_0 to be specified below. Let $N \gg 1$ and consider the event $T_{\imath,o}(NL_0)$ for the annulus A_{NL_0}. Divide A_{NL_0} into a grid of scale L_0 so as to have the appearance of an A_N on the large scale lattice. If $\mathcal{P} : \imath \to o$ is a path in A_{NL_0}, each "site" on the large scale lattice that is visited by $\mathcal{P}$ has achieved an event like $T_{\imath,o}(L_0)$ – with the possible exception of the sites next to the boundary. Let us denote a "site" of A_N to be "occupied" if the analog of the $T_{\imath,o}(L_0)$ occurs and vacant otherwise. For the sake of being definitive, let us deem all sites neighboring the boundary of A_N to be occupied. It is clear that $\mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_{NL}}(T_{\imath,o}(NL))$ does not exceed the probability of a connection between the $\imath$ and the o of A_N in the large–scale problem.

Now of course, these site variables are not independent. However let us regard a sublattice consisting of a fraction – $1/3^d$ – of these sites as sitting in the center of a translate of A_{L_0} with these translates of A_{L_0} situated in such a way that they tile the lattice. With the maximizing boundary conditions on these translates of A_{L_0}, the sublattice of site occupation variables *are* independent and their probability is bounded above by ϵ. There are 3^d possible ways to design such sublattices (depending on which sites are chosen as the centers) such that each site of A_N is a central site on one of these 3^d sublattices. Thus an "occupied cluster" of consisting of K interior sites of A_N must have at least $1/3^d$ of these sites on (at least) one of the sublattices. Therefore, the probability of a given occupied cluster with K interior sites is less than $(\epsilon)^{K/3^d}$. The minimum sized cluster that permits the possibility of an actual path is essentially N and there are only of the order of N^{d-1} starting points on the inner boundary. Hence

$$\mathbb{M}^{\mathbf{1}^{++}}_{\beta,A_N L}(T_{\imath,o}) \leq C_2 N^{d-1} \sum_{K > N - C_1} [\lambda(d)\epsilon^{1/3^d}]^K \tag{11}$$

with C_1 and C_2 constants of the order of unity and $\lambda(d) < (d-1)$ the connectivity constant. It is evident that if $\epsilon < \epsilon_0 = 1/\lambda^{3^d}$, the stated result follows. $\square$

Corollary. *For the 2d models, the spin–wave stiffness does not go continuously to zero at any temperature. In any dimension, if the conditions of Theorem 3 hold for some finite L_0, there is exponential decay of correlations in any limiting $\overline{\odot}$–state.*

Proof. According to Lemma 2, the $\overline{\odot}$–state that maximizes the probability of $T_{i,j}$ is always the $\mathbf{1}^+$–state. Under the conditions stated in Theorem 3, it is clear that the probability of $T_{i,j}$ tends to zero exponentially in any limiting $\overline{\odot}$–state. (Later we will show that under these conditions there is in fact a unique limiting $\odot$–state.) Using a bound along the lines of Equation (9), exponential decay for the 2–point function is readily established: The factor of 2 in this inequality is for the X and Y–component pieces of $\vec{s}_i \cdot \vec{s}_j$. Indeed, in *any* boundary condition $*$,

$$\langle s_i^{[X]} s_j^{[X]} \rangle^*_{\beta,\mathcal{G}} \equiv \langle b_i \sigma_i b_j \sigma_j \rangle^*_{\beta,\mathcal{G}} \leq \mathbb{M}^*_{\beta,\mathcal{G}}(T_{i,j}) \tag{12}$$

with connections through the boundary included in the definition of $T_{i,j}$. Since, among limiting $\overline{\odot}$–states this is maximized in the limiting $\mathbf{1}^+$–state, the correlation among the X–components goes to zero. The correlations between the Y components (in $\overline{\odot}$–states) would be maximized in the analog of the $\mathbf{1}^+$–state and hence, by the symmetry between X and Y components, is also (in any $\overline{\odot}$–state) always bounded by the probability of $T_{i,j}$ in the $\mathbf{1}^+$–state. Thus we actually recover Equation (9) for the $\mathbf{1}^+$–states and the conclusion about exponential decay is immediate. The statement concerning the spin wave stiffness is a tautology, however c.f. Remark 2 below. □

Remark 1. If β_c is *defined* by the infimum over temperatures at which $K_\infty(\beta)$ is zero, then, by an obvious continuity argument, $K_\infty(\beta_c) > 0$ in $d = 2$. For the XY–model, the results of [FS] (concerning the existence of a region of power law decay of correlations) rather easily imply that such a discontinuity occurs at a finite β.

Remark 2. Starting with [NK], detailed renormalization group studies of this "class" of problems predicts a *universal* value of $\beta_c K_\infty(\beta_c)$. Although the present derivation is far cry from a proof of any such statement, it is worth observing that the same set of results proved in Theorem 3 hold for a variety of models with "O(2)" characteristics – e.g. the $\mathbb{Z}_{4n}$–clock models – using the *same* value of ϵ_0. Thus we have a universal lower bound on $\beta_c K_\infty(\beta_c)$. This is analogous to (and borrowed from) the current situation in percolation theory: various crossing probabilities – even the one used here – which at the critical point are believed to converge to universal values at large length scale, can at least be shown to satisfy uniform bounds with universal constants.

Additional Results. Some further results will be stated below but all the remaining proofs have been relegated to the appendix.

The usual definition of *percolation* in correlated models starts, in finite volume, with the probability of a connection to the boundary in the boundary conditions that optimize this probability. (C.f. [$\mathrm{CM_I}$], definition following Equation (II.11).) Here, let us define:

Definition. Let $\Lambda \subset \mathbb{Z}^d$ be a finite connected set that contains the origin and let $T_{0,\partial\Lambda}$ denote the event that the origin is connected to the boundary. Let

$$\Pi_\Lambda(\beta) = \mathbb{M}^{\mathbf{1}^+}_{\beta,\Lambda}(T_{0,\partial\Lambda}) \equiv \max_{*\in\odot} \mathbb{M}^{*}_{\beta,\Lambda}(T_{0,\partial\Lambda}) \tag{13.a}$$

and

$$\Pi_\infty(\beta) = \lim_{\Lambda\nearrow\mathbb{B}^d} \Pi_\Lambda(\beta). \tag{13.b}$$

(In light of Lemma 2, the existence of this limit is not hard to establish.) The actual *percolation* probabilities, denoted by P's instead of Π's is defined as in Equations (13) but with the maximum taken over all boundary conditions.

Theorem 4. *[A] Let $m(\beta)$ denote the spontaneous magnetization. Then there are finite non–zero constants, c_1 and c_2 (that depend only on minor the details of the model) such that*

$$c_2\Pi_\infty(\beta) \leq m(\beta) \leq c_1\Pi_\infty(\beta).$$

[B] If $m(\beta) = 0$, there is a unique limiting $\odot$–state.

Proof. A proof will be provided in the appendix.

Remark. The results concerning uniqueness are hardly an improvement over the existing results which apply to most of these models considered here – uniqueness among translation invariant states when the magnetization vanishes [MMPf]. Of greater concern (to the author) is the connection between phase transitions in the spin–systems and percolation in the corresponding graphical representation. This is further underscored by the final result:

Theorem 5. *Let $*$ denote any finite volume $\overline{\odot}$–measure or infinite volume limit thereof and let $\langle s_i^{[X]} s_j^{[X]}\rangle_\beta^* \equiv \langle b_i\sigma_i b_j\sigma_j\rangle_\beta^*$ denote the (untruncated) correlation function for the X–components. Then,*

$$c_1^2\mathbb{M}_\beta^*(T_{i,j}) \geq \langle s_i^{[X]} s_j^{[X]}\rangle_\beta^* \geq c_2^2\mathbb{M}_\beta^*(T_{i,j})$$

with c_1 and c_2 as in Theorem 4. In particular, if $m(\beta) = 0$ and $\mathcal{X}$ is defined by

$$\mathcal{X}(\beta) = \sum_j \langle s_0^{[X]} s_j^{[X]}\rangle_\beta$$

evaluated in the unique limiting $\odot$–state then

$$c_1^2\mathbb{E}_\beta(|C_0|) \geq \mathcal{X}(\beta) \geq c_2^2\mathbb{E}_\beta(|C_0|)$$

where $\mathbb{E}_\beta(|C_0|)$ is the expected size of the connected cluster of the origin in the graphical representation.

Proof. The upper bound for the correlation function was derived in [A], the rest will be proved in the appendix.

Theorems 4 and 5 provide complete justification for the use of "percolation" as the critical criterion in the Wolff algorithm [W] or the Invaded Cluster version of this algorithm [CM_{II}].

Appendix: Monotonicity Properties of the Wolff Measures. For reasons that are primarily of a technical nature, this appendix will be concerned with generalizations of the types of models already discussed (even though such generalizations are "unphysical" from the perspective of systems with $O(2)$ symmetry). Thus consider a graph $\mathcal{G}$ and let $H_{\mathcal{G}}$ denote the Hamiltonian

$$H_{\mathcal{G}} = -\sum_{\langle i,j\rangle}(K_{i,j}a_ia_j\tau_i\tau_j + J_{i,j}b_ib_j\sigma_i\sigma_j) \tag{A.1}$$

with $K_{i,j}, J_{i,j} \geq 0$. As discussed previously, the single site *a priori* measures and the range of the a_i and b_i as well as the constraint between them may be regarded as fairly arbitrary: It is enough to assume that they are non–negative, uniformly bounded and that a_i goes down when b_i goes up. Finally, it will be assumed that if b_i achieves its maximum value then the corresponding a_i is zero. Most of these assumptions can be removed but with an unreasonable cost of labor and space. To avoid spurious notational provisos, let us assume that the single site measures are discrete. (Indeed, since we will always start in finite volume, the "general" case can be recovered from the discrete by a limiting procedure.) Thus we let $\rho_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(-)$ denote the measure on configurations $\underline{b} = (b_i \mid i \in \mathbb{S}_{\mathcal{G}})$ defined by the weights

$$R_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(\underline{b}) = Z_{\underline{a},\underline{K}}^{I}(2\beta) Z_{\underline{b},\underline{J}}^{I}(2\beta) \prod_{\langle i,j\rangle \in \mathbb{B}_{\mathcal{G}}} e^{\beta[K_{i,j}a_i a_j + J_{i,j} b_i b_j]} \prod_{i \in \mathbb{S}_{\mathcal{G}}} f_i(b_i) \tag{A.2}$$

where $f_i(b_i)$ is the *a priori* probability of b_i, $\underline{f} \equiv (f_i \mid i \in \mathbb{S}_{\mathcal{G}})$, $\underline{K} \equiv (K_{i,j} \mid \langle i,j\rangle \in \mathbb{B}_{\mathcal{G}})$ and all other notation has been defined elsewhere.

Proposition A.1. *The measures $\rho_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(-)$ are (strong) FKG.*

Proof. Let $\underline{b}$ denote a fixed configuration and let u and v denote any distinct pair of sites in $\mathcal{G}$. Let $\Delta_u > 0$ and δ_u denote the configuration that is zero except at the site u where it is equal to $b_u + \Delta_u$. Similarly for δ_v with some $\Delta_v > 0$. It may as well be assumed that $f_u(b_u + \Delta_u)$ and $f_u(b_v + \Delta_v)$ are positive. Thus, the configuration $\underline{b} \vee \delta_u \vee \delta_v$ has been "raised" at the sites u and v while $\underline{b} \vee \delta_u$ has been raised only at u, etc. Similarly, if Γ_u is the corresponding amount that a_u has to be lowered (determined by the constraint at u, the value of b_u and Δ_u) then let $\underline{a} \wedge \gamma_u$ denote the configuration of $\underline{a}$'s that has been lowered at u etc. (Formally, γ_u is $a_u - \Gamma_u$ at the site u and infinite elsewhere.) To prove the desired claim, it is sufficient (and necessary) to show

$$R_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(\underline{b} \vee \delta_u \vee \delta_v) R_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(\underline{b}) \geq R_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(\underline{b} \vee \delta_u) R_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(\underline{b} \vee \delta_v) \tag{A.3}$$

After cancellation of all manifestly equal terms (assumed non-zero) the purported inequality boils down to

$$\begin{aligned} &[e^{\beta\Delta_u\Delta_v} Z_{\underline{b}\vee\delta_u\vee\delta_v,\underline{J}}^{I}(2\beta) Z_{\underline{b},\underline{J}}^{I}(2\beta)](e^{\beta\Gamma_u\Gamma_v} Z_{\underline{a}\wedge\gamma_u\wedge\gamma_v,\underline{K}}^{I}(2\beta) Z_{\underline{a},\underline{K}}^{I}(2\beta)) \geq \\ &\qquad \geq [Z_{\underline{b}\vee\delta_u,\underline{J}}^{I}(2\beta) Z_{\underline{b}\vee\delta_v,\underline{J}}^{I}(2\beta)](Z_{\underline{a}\wedge\gamma_u,\underline{K}}^{I}(2\beta) Z_{\underline{a}\wedge\gamma_v,\underline{K}}^{I}(2\beta)). \end{aligned} \tag{A.4}$$

It is claimed that the term in the square bracket on the rhs does not exceed the corresponding term on the left and similarly for the terms in the round bracket. Indeed, a moments reflection will show that these two inequalities are of an identical form. Let us therefore focus on proving

$$[e^{\beta\Delta_u\Delta_v}Z^I_{\underline{b}\vee\delta_u\vee\delta_v,\underline{J}}(2\beta)Z^I_{\underline{b},\underline{J}}(2\beta)] \geq [Z^I_{\underline{b}\vee\delta_u,\underline{J}}(2\beta)Z^I_{\underline{b}\vee\delta_v,\underline{J}}(2\beta)]. \tag{A.5}$$

and the same derivation will hold for the $\underline{a}$–pairs.

It turns out that the derivation is far easier without the annoyance of the $\Delta_u\Delta_v$ cross terms. Let us thus define

$$H^{(0)} = -\sum_{\langle i,j\rangle} J_{i,j}(\delta_{\sigma_i,\sigma_j} - 1)b_ib_j, \tag{A.6a}$$

$$H^{(U)} = -\sum_{\langle i,u\rangle} J_{i,j}(\delta_{\sigma_i,\sigma_u} - 1)\Delta_u b_i, \tag{A.6b}$$

and similarly for $H^{(V)}$. In these terms $Z^I_{\underline{b}\vee\delta_u\vee\delta_v,\underline{J}}(2\beta)$ is given by

$$Z^I_{\underline{b}\vee\delta_u\vee\delta_v,\underline{J}}(2\beta) = Tr[e^{-2\beta H^{(0)}}e^{-2\beta H^{(U)}}e^{-2\beta H^{(V)}}e^{2\beta J_{u,v}\Delta_u\Delta_v(\delta_{\sigma_u,\sigma_v}-1)}]. \tag{A.7}$$

To get rid of the cross terms, it will be shown that

$$e^{\beta\Delta_u\Delta_v J_{u,v}}Tr[e^{-2\beta H^{(0)}}e^{-2\beta H^{(U)}}e^{-2\beta H^{(V)}}e^{2\beta J_{u,v}\Delta_u\Delta_v(\delta_{\sigma_u,\sigma_v}-1)}] \geq$$
$$\geq Tr[e^{-2\beta H^{(0)}}e^{-2\beta H^{(U)}}e^{-2\beta H^{(V)}}]. \tag{A.8a}$$

Indeed, dividing both sides of the purported inequality (A.8a) by the right hand side and denoting by $\mathbb{E}^I_{H,\beta}(-)$ the expectation with respect to the Ising Hamiltonian H at inverse temperature β, the desired (8.Aa) reads

$$e^{\beta\Delta_u\Delta_v J_{u,v}}\mathbb{E}^I_{H^{(0)}+H^{(U)}+H^{(V)},2\beta}(e^{2\beta J_{u,v}\Delta_u\Delta_v(\delta_{\sigma_u,\sigma_v}-1)}) \geq 1. \tag{A.8b}$$

Expanding the integrand in the usual FK fashion, this reduces to showing that

$$e^{-\beta\Delta_u\Delta_v J_{u,v}} + 2\mathrm{sh}(\beta\Delta_u\Delta_v J_{u,v})\mathbb{E}^I_{H^{(0)}+H^{(U)}+H^{(V)},2\beta}(\delta_{\sigma_u,\sigma_v}) \geq 1 \tag{A.8c}$$

Here is one of the few places where the fact that the underlying model has an Ising structure is used: $\mathbb{E}^I_{H,\beta}(\delta_{\sigma_i,\sigma_u}) \geq 1/2$ so the left hand side of (A.8) is at least as big as $\mathrm{ch}\beta\Delta_u\Delta_v J_{u,v}$. For the remainder of the proof, it might just as well be assumed that the underlying model is the q–state Potts model.

The remainder of this proof reduces to showing

$$\mathbb{E}^I_{H^{(0)}+H^{(U)},2\beta}(e^{-2\beta H^{(V)}}) \geq \mathbb{E}^I_{H^{(0)},2\beta}(e^{-2\beta H^{(V)}}). \tag{A.9}$$

This is, very similar to the sorts of inequalities that were established in [C] so here the derivation will be succinct. Let $\epsilon_{i,v} = 1 - e^{2\beta J_{i,v}b_i\Delta_v}$ and let $\mathcal{N}_v$ denote the collection of sets in $\mathbb{S}_{\mathcal{G}}$ each of which contains v and some subset of the sites in $\mathcal{G}$ that are connected to v. Expanding $e^{-2\beta H^{(V)}}$ in the usual FK fashion, it is seen that

$$e^{-2\beta H^{(V)}} = \sum_{\mathcal{F}\in\mathcal{N}_v} r_{\mathcal{F}}\delta_{\sigma_{\mathcal{F}}} \tag{A.10}$$

with $r_{\mathcal{F}} = \prod_{i\in\mathcal{F}}\epsilon_{i,v}\prod_{j\notin\mathcal{F}}(1-\epsilon_{j,v})$ and where $\delta_{\sigma_{\mathcal{F}}}$ is one if all the spins in $\mathcal{F}$ agree and zero otherwise. However, using an FK expansion of the q–state Potts system with Hamiltonian H, it is not hard to show

$$\mathbb{E}^I_{H,\beta}(\delta_{\sigma_{\mathcal{F}}}) = \mathbb{E}^{FK\ (q=2)}_{H,\beta}((\frac{1}{q})^{C_{\mathcal{F}}-1}) \tag{A.10}$$

where $C_{\mathcal{F}}$ is the number of connected components of the set $\mathcal{F}$. This is the expectation of an FKG increasing function and thus the desired inequality follows – term by term – from the fact that random cluster model that comes from the "bigger" Hamiltonian (i.e. $H^{(0)} + H^{(U)}$) is FKG dominant. $\square$

Corollary I. *Consider two systems on the same graph $\mathcal{G}$ with parameters $\underline{J}$, $\underline{J}'$ and single site measures determined by the collections $\underline{f}$ and $\underline{f}'$ respectively. Suppose that $\underline{J} \succ \underline{J}'$ meaning that for each $\langle i, j\rangle \in \mathbb{B}_{\mathcal{G}}$, $J_{i,j} \geq J'_{i,j}$ and further suppose that $\underline{f} \succ \underline{f}'$ in the sense that for each i, $f_i(b_i)/f'_i(b_i)$ is an increasing function of b_i. Then*

$$\rho_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(-) \underset{\text{FKG}}{\geq} \rho_{\beta,\mathcal{G}}^{\underline{J}',\underline{K},\underline{f}'}(-)$$

Proof. This is an immediate consequence of the FKG properties of these measures and the previous derivation. First, if $f' \prec f$, then

$$\prod_{i \in \mathbb{S}_{\mathcal{G}}} f_i(b_i) = [\frac{\prod_{i\in\mathbb{S}_{\mathcal{G}}} f_i(b_i)}{\prod_{i\in\mathbb{S}_{\mathcal{G}}} f'_i(b_i)}] \prod_{i \in \mathbb{S}_{\mathcal{G}}} f'_i(b_i) \tag{A.11}$$

so the $\underline{f}$–weights are of the form [increasing function]$\times$ $\underline{f}'$–weights. To establish the desired result for $\underline{J} \succ \underline{J}'$ it is sufficient to consider one bond at a time. Thus let $\langle u, v\rangle \in \mathbb{B}_{\mathcal{G}}$ and suppose that $J_{u,v} = J'_{u,v} + L_{u,v}$ (with $L_{u,v} > 0$) and all other J's equal. Then

$$R_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(\underline{b})/R_{\beta,\mathcal{G}}^{\underline{J}',\underline{K},\underline{f}}(\underline{b}) = e^{\beta L_{u,v} b_u b_v} \mathbb{E}^I_{H^I_{\underline{b}},2\beta}[e^{2\beta L_{u,v} b_u b_v (\delta_{\sigma_u,\sigma_v}-1)}] \tag{A.12a}$$

where the Ising Hamiltonian $H^I_{\underline{b}}$ was defined in Equation (5.a) – and the $\underline{J}$ dependence has been suppressed. After a few manipulations along the lines of those in the previous proposition, Equation (A.12a) reduces to

$$R_{\beta,\mathcal{G}}^{\underline{J},\underline{K},\underline{f}}(\underline{b})/R_{\beta,\mathcal{G}}^{\underline{J}',\underline{K},\underline{f}}(\underline{b}) = \text{ch}(\beta L_{u,v} b_u b_v) + \text{sh}(\beta L_{u,v} b_u b_v) \mathbb{E}^{FK\ (q=2)}_{H^I_{\underline{b}},2\beta}(\mathcal{X}_{T_{u,v}}) \tag{A.12b}$$

where $\mathcal{X}_{T_{u,v}}$ is indicator of the event that u is connected to v. The sines and cosines are manifestly (non–negative) increasing functions of $\underline{b}$, while the random cluster term is the expectation of a *positive* event and is therefore an increasing function of all couplings in the Hamiltonian – including the $\underline{b}$'s $\square$

Let us now turn to a discussion of boundary conditions. Let $\mathcal{G}$ denote a graph and let $\mathbb{L} \subset \mathbb{S}_{\mathcal{G}}$. The starting point will be the consideration of conditional measures for $\nu_{\beta,\mathcal{G}}^{W \underline{J},\underline{K},\underline{f}}(-)$ – the measures corresponding to the weights in Equation (6) cast in the more general framework – subject to specifications on $\mathbb{L}$ and the consequence of these specifications on the $\underline{b}$ marginals. A specification $*$ will be called a $\tilde{\odot}$–specification if (i) the values $(b_i \mid i \in \mathbb{L})$ are specified: $b_i = b_i^*$; $i \in \mathbb{L}$ and (ii) $\mathbb{L}$ is divided into disjoint components $\ell_1^*, \ell_2^*, \dots \ell_k^*$ such that the counting rule in the FK expansion deems all the sites in and connected to each ℓ_n^* to be part of the same cluster.

Remark. Back in the spin–system, one interpretation of a $\tilde{\odot}$–specification is obvious: having determined the b_i on $\mathbb{L}$, the signs of the X–components of the spins – the σ_i's – are locked together within each component and they take on both values with equal probability. On the other hand, the same graphical weights emerge if one (and only one) of the components is deemed to represent spins pointing in the positive X–direction. The reader is cautioned that at this stage, the signs of the Y components of the boundary spins still have all their *a priori* degrees of freedom.

There is a natural partial order on the set of all possible $\tilde{\odot}$–specifications: $* \succ *'$ if (1) $\mathbb{L} \supset \mathbb{L}'$ and each b_i on $\mathbb{L} \setminus \mathbb{L}'$ is set to the maximum value, (2) each $b_i^* \geq b_i^{*'}$, $i \in \mathbb{L} \cap \mathbb{L}'$ and (3), the components of $*$, $\ell_1^*, \ell_2^*, \dots \ell_k^*$ "contain" the $*'$–components $\ell_1^{*'}, \ell_2^{*'}, \dots \ell_{k'}^{*'}$ in the sense that if $\ell_{j'}^{*'} \cap \ell_j^* \neq \emptyset$ then $\ell_{j'}^{*'} \subset \ell_j^*$. The following is easily seen:

Corollary II. *If $*$ is a g–specification and $\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-)$ is the associated measure on the remaining $\underline{b}$'s then $\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-)$ is (strong) FKG. Furthermore if $* \succ *'$ in the sense described* *above,* $\underline{J} \succ \underline{J}'$ *and* $\underline{f} \succ \underline{f}'$ *then*

$$\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-) \underset{\text{FKG}}{\geq} \rho_{\beta,\mathcal{G}}^{*'\underline{J}',\underline{K},\underline{f}'}(-).$$

Proof. The above is clear given the following mechanism to create a $\tilde{\odot}$–specification: to fix the values of b_i on $\mathbb{L}$, concentrate the *a priori* measures. To lock the components, introduce artificial J–type couplings between all pairs of sites in a given component and send these couplings to infinity; the desired measure is recovered in the limit. If $* \succ *'$ this procedure involves higher J's and higher b's. $\square$

Proposition A.2. *Let* $\nu_{\beta,\mathcal{G}}^{W\ *\underline{J},\underline{K},\underline{f}}(-)$ *and* $\nu_{\beta,\mathcal{G}}^{W\ *'\underline{J}',\underline{K},\underline{f}'}(-)$ *denote two Wolff measures with all primed quantities below unprimed quantities in the sense described. Let* $\mathbb{M}_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-)$ *and* $\mathbb{M}_{\beta,\mathcal{G}}^{*'\underline{J}',\underline{K},\underline{f}'}(-)$ *denote the corresponding bond measures. Then*

$$\mathbb{M}_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-) \underset{\text{FKG}}{\geq} \mathbb{M}_{\beta,\mathcal{G}}^{*'\underline{J}',\underline{K},\underline{f}'}(-)$$

Proof. Let $\mathcal{A}$ denote an increasing bond event. Let us write as in Equation (8)

$$\mathbb{M}_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(\mathcal{A}) = \sum_{\underline{b}} \rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(\underline{b})\mu_{\underline{J},\underline{b}}^{FK*}(\mathcal{A}) \tag{A.13}$$

and similarly for $\mathbb{M}_{\beta,\mathcal{G}}^{*'\underline{J}',\underline{K},\underline{f}'}(\mathcal{A})$. The desired result follows immediately from the FKG properties of the usual random cluster measures: both $\mu_{\underline{J},\underline{b}}^{FK*}(\mathcal{A})$ and $\mu_{\underline{J}',\underline{b}}^{FK*'}(\mathcal{A})$ are increasing functions of $\underline{b}$ and furthermore, if $* \succ *'$ and $\underline{J} \succ \underline{J}'$ then $\mu_{\underline{J},\underline{b}}^{FK*}(\mathcal{A}) \geq \mu_{\underline{J}',\underline{b}}^{FK*'}(\mathcal{A})$. $\square$

Thus far, the Y degrees of freedom have been left completely unspecified. Now the same sorts of specifications will be considered for these objects and this defines a $\odot$–specification: In addition to a $\tilde{\odot}$ specification, $\mathbb{L}$ is divided into disjoint components $\jmath_1, \dots \jmath_m$ on which the τ–variables act in unison. A recapitulation of the previous arguments yields:

Proposition A.2. *Let $*$ denote a $\odot$ specification and let $\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-)$ denote the corresponding measure. Then $\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-)$ if FKG. Further, if $* \succ *'$ meaning the same as above regarding the $\underline{J}$'s, the $\underline{f}$'s and the ℓ–components while $\underline{K}' \succ \underline{K}$ and the $\jmath_1', \dots, \jmath_m'$ contain the $\jmath_1, \dots, \jmath_m$ then*

$$\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-) \underset{\mathrm{FKG}}{\geq} \rho_{\beta,\mathcal{G}}^{*'\underline{J}',\underline{K}',\underline{f}'}(-)$$

and accordingly

$$\mathbb{M}_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-) \underset{\mathrm{FKG}}{\geq} \mathbb{M}_{\beta,\mathcal{G}}^{*'\underline{J}',\underline{K}',\underline{f}'}(-).$$

In particular, the FKG maximizing boundary condition (on $\mathbb{L}$) in the $\odot$–class is the b_i set to the maximum value, $\sigma_i \equiv 1$ and the $\jmath_1, \dots, \jmath_m$ being the individual sites of $\mathbb{L}$. The latter is, of course automatic if b_i maximized $\Rightarrow a_i = 0$

Proof. Follows the lines of the previous arguments along with the observation that any increasing function of $\underline{a}$ is a decreasing function of $\underline{b}$. $\square$

Superpositions of $\odot$–specifications do not constitute a $\odot$–class boundary condition nor, in general, are they FKG measures. This is the usual situation in ferromagnetic systems and is of no serious consequence since we have knowledge of the maximizing measure in the $\odot$–class. In any case, let us define the $\overline{\odot}$–class as that which consists of superpositions from the $\odot$–class. The following is pivotal:

Lemma A.3. *Let $\mathbb{L} \subset \mathbb{S}_G$ and let $*$ denote a $\overline{\odot}$–specification on $\mathbb{L}$. Let $\mathbb{K} \supset \mathbb{L}$ and consider $\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-)_{||\mathbb{S}_{\mathcal{G}}\setminus\mathbb{K}}$, the restriction of $\rho_{\beta,\mathcal{G}}^{*\underline{J},\underline{K},\underline{f}}(-)$ to the remaining sites. Then this restricted measure is of the $\overline{\odot}$–class.*

Proof. It is sufficient to discuss the case where $*$ is itself a pure $\odot$–specification. Consider the full Wolff measures $w^{*\underline{J},\underline{K},\underline{f}}_{\beta,\mathcal{G}}(-)$ on configurations $(\omega,\eta,\underline{b})$ where ω and $\underline{b}$ are as have been described and η denotes configurations of FK bonds in the random cluster expansion of the τ–system. Thus, e.g. the $\nu^{\underline{J},\underline{K},\underline{f}}_{\beta,\mathcal{G}}(-)$ measures are obtained by integrating out the η–bonds. Now, to study the restricted measure, let us may condition on an $(\omega,\eta,\underline{b})$ configuration on $\mathbb{K}$ and sum over all η–configurations (and, if desired, ω–configurations) pertaining to the bonds of $\mathbb{S}_{\mathcal{G}}\setminus\mathbb{K}$. Having done so, a sum must be performed over all the external configurations with the appropriate weights assigned by $w^{*\underline{J},\underline{K},\underline{f}}_{\beta,\mathcal{G}}(-)$. But, since $*$ is a $\odot$–specification, it is clear that each $(\underline{\omega},\underline{\eta},\underline{b})$ configuration on $\mathbb{K}$ provides a $\odot$–specification on $\mathbb{S}_{\mathcal{G}}\setminus\mathbb{K}$: Indeed, the b–values are fixed, the components $\ell_1,\dots\ell_k$ are just the ω–components while the η–components constitute the $\jmath_i,\dots,\jmath_m$. $\square$

It is now straightforward to establish the various results claimed in Theorems 4 and 5. Indeed everything except the statements concerning uniqueness follow immediately from the existing machinery. Here, to simplify matters notationally, let us again assume that β, $\underline{J}$, and $\underline{K}$ and the graph $\mathcal{G}$ are fixed and omit any further explicit reference. All of Theorem 5 amounts to the stated bound of the correlation function in terms of the connectivity function. Recalling that in a $\odot$–state, the event $T_{i,j}$ includes connections via the boundary component, these bounds are easily proved:

Proof of Theorem 5. If $*$ denotes a $\odot$ state, it is claimed that

$$\langle s_i^{[X]} s_j^{[X]}\rangle = \mathbb{E}^*_\rho[b_i b_j \mu^*_{\underline{b}}(T_{i,j})] \tag{A.14}$$

where $\mathbb{E}^*_\rho[-]$ denotes expectation with respect to the $\rho^*(-)$ measure on the $\underline{b}$–configurations. Indeed, fixing $\underline{b}$ and ω, the Ising spins are equal if i is connected to j – either directly or via one of the boundary components – and are uncorrelated with at least one of them having equal probability of ± 1 otherwise. Summing over all ω with $\underline{b}$ fixed and then summing over $\underline{b}$ yields the identity displayed in Equation (A.14). But obviously, b_i and b_j cannot exceed their maximum values and this provides the upper bound with c_1 equal any uniform bound on these values. On the other hand, $\mu^*_{\underline{b}}(T_{i,j})$, b_i and b_j are all increasing functions of $\underline{b}$ and hence, the FKG inequality, provides the bound

$$\langle s_i^{[X]} s_j^{[X]} \rangle \geq \mathbb{E}^*_\rho[\mu^*_{\underline{b}}(T_{i,j})]\mathbb{E}^*_\rho[b_i]\mathbb{E}^*_\rho[b_j]. \tag{A.15}$$

The quantities $\mathbb{E}^*_\rho[b_i]$ and $\mathbb{E}^*_\rho[b_j]$ may be estimated by considering the worst case $\odot$–boundary conditions on the neighborhoods of i and j which yields the uniformly positive constant c_2. For the d–dimensional XY–model, we have $c_1 = 1$ and $c_2 = (2/\pi)(e^{-2d\beta})$

Proof of Theorem 4 [A]. First observe that the lower bound follows because the magnetization can be estimated from below by the average of the $s^{[X]}$'s in any state, and by using the $\mathbf{1}^+$–state, this is obtained. In fact, for the XY model, and several other of the models under consideration, both of these bounds follows because it can be proved, via correlation inequalities, that the $\mathbf{1}^+$ state is exactly the state that produces the magnetization. For the general case, consider the addition of the usual magnetic term:

$$\sum_i h s_i^{[X]} \equiv \sum_i 2hb_i(\delta_{\sigma_i,+} - 1) + hb_i \tag{A.16}$$

to the Hamiltonian. The effect of this additional term may be incorporated into the present analysis by the addition of a single "ghost" spin connected to all other spins with coupling h. (Here the ghost spin plays more the rôle of a boundary site than a full blown XY–degree of freedom.) Now for a.e. h, the (thermodynamic) magnetization can be defined by evaluating the actual magnetization (the average of the $s^{[X]}$'s) in any convenient state. Thus, using the limiting state constructed from $\mathbf{1}^+$ boundary conditions, it is clear that for a.e. positive h, the magnetization is bounded above by the (limiting) average fraction of sites connected to the ghost site or the boundary. Let Λ_L denote the box of scale L and define

$$\pi_L(h,\beta) = \frac{1}{\Lambda_L}\sum_{i\in\Lambda}\mathbb{M}^{\mathbf{1}^+}_{\beta,h,\Lambda_L}(T_{i,\mathbf{B}}) \tag{A.17}$$

where $T_{i,\mathbf{B}}$ is the event that the site i is connected to the boundary or the ghost site and the sum includes the contribution from the boundary sites themselves. The desired result follows from two elementary facts: First, by continuity in finite volume,

$$\lim_{h\to 0}\pi_L(h,\beta) = \pi_L(0,\beta). \tag{A.18}$$

Second, by a sequence of fairly standard manipulations,

$$\Pi_\infty(\beta) \equiv \lim_{L\to\infty}\Pi_{\Lambda_L}(\beta) = \lim_{L\to\infty}\pi_L(0,\beta). \tag{A.19}$$

Now, for $h > 0$ suppose we were to evaluate $m(h,\beta)$ starting on Λ_{NL} using $\mathbf{1}^+$ boundary conditions and letting $N\to\infty$. Since, for finite N, this is a certified finite volume $\odot$–state, we increase the value by conditioning on the event that the grid that divides Λ_{NL} into small copies of Λ_L is fully occupied. Thus, at each stage it is learned that

$$m_{\Lambda_{NL}}(h,\beta) \equiv \frac{1}{|\Lambda_{NL}|}\sum_{i\in\Lambda_{NL}}\langle s_i^{[X]}\rangle^{\mathbf{1}^+}_{\beta,h,\Lambda_L} \le \pi_L(h,\beta). \tag{A.20}$$

Taking $h\downarrow 0$ (along a sequence of points of continuity) the desired result follows from Equations (A.18) and (A.19).

Proof of Theorem 4 [B]. Let $\mathcal{G}$ denote a graph, $\mathbb{I}\subset\mathbb{S}_G$ and $\mathbb{K} = \mathbb{S}_G\setminus\mathbb{I}$. Let $\gamma = \{\langle i,k\rangle\in\mathbb{B}_{\mathcal{G}} \mid i\in\mathbb{I}, k\in\mathbb{K}\}$ denote the connecting bonds and let $\Gamma(\gamma)$ denote the contour event that every

ω–bond in γ is vacant. In what follows, it is assumed that is there is any specification on $\mathcal{G}$, it is of the $\odot$–type and involves only the sites of $\mathbb{K}$.

It is claimed that if $\Gamma(\gamma)$ occurs then the measure on the $(b_i \mid i \in \mathbb{I})$ lies below, in the sense of FKG, the "free measure" on $\mathbb{I}$ that would be obtained if all the $J_{i,k}$ on γ were zero. Indeed, for any fixed $\underline{b}$ on $\mathbb{K}$ and η–configuration the weights for the configurations $(b_i \mid i \in \mathbb{I})$ are given by $Z_{\underline{a}}^{I,\eta}(2\beta) \prod_{\langle i,k\rangle \in\gamma} e^{\beta J_{i,k}(a_i a_k - b_i b_k)} Z_{\underline{b}}^{I,f}(2\beta)$ where $Z^{I,f}$ denotes the free boundary partition function and $Z^{I,\eta}$ denotes the partition function with ($\odot$–type) boundary specification provided by η. On the other hand, the free weights are given simply by $Z_{\underline{a}}^{I,f}(2\beta) Z_{\underline{b}}^{I,f}(2\beta)$. Thus it is clear that irrespective of the information on the outside, the conditional weights are a decreasing function times the free weights. Now, supposing that $\Pi_\infty(\beta) = 0$, it is easy to establish uniqueness of the limiting ρ–measures among $\odot$–states: Let $\Lambda \subset \mathbb{Z}^d$ be a finite connected set. Let $\Xi \supset \Lambda$ with Ξ so large that the probability of an ω–connection between Λ and $\partial\Xi$ in the $\mathbf{1}^+$ state on Ξ is negligible. Under these circumstances, there are contours separating Λ from $\partial\Xi$; let γ denote such a contour and let $\tilde{\Gamma}(\gamma)$ the event that γ is the *outermost* such separating contour. These contour events form a disjoint partition so, up to the negligible probability of a connection between Λ and $\partial\Xi$, the restriction of the maximal measure in Ξ to Λ is below a superposition of free measures on various separating contours.

Now consider the lowest boundary condition on Ξ: setting all the boundary a_i to one and locking *their* spin directions. By $a \leftrightarrow b$ symmetry, the same outermost contours (in the η expansion) appear with the same probabilities and we find – again up to negligible terms – that this worst measure in Ξ restricted to Λ lies above the previously discussed superposition. Evidently the two restricted measures coincide in the $\Xi \nearrow \mathbb{Z}^d$ limit and hence all the limiting

$\odot$–measures coincide at least as far as the distributions of $\underline{b}$'s are concerned. However, the same argument implies uniqueness for the various other Wolff–measures in the $\odot$–class and, given the fact that all bond clusters are finite, uniqueness among all Gibbs measures of in the $\odot$–class follows easily. $\square$

I would like to thank Jon Machta who provided help and encouragement in all stages of this work and without whom this project would not have existed.

References

[A] M. Aizenman, *On the Slow Decay of $O(2)$ Correlations in the Absence of Topological Excitations: Remark on the Patrascioiu–Seiler Model*, Jour. Stat. Phys. **77** (1994), 351–359.

[ACCN] M. Aizenman, J.T. Chayes, L. Chayes and C.M. Newman, *Discontinuity of the Magnetization in One–Dimensional $1/|x-y|^2$ Ising and Potts Models*, Jour. Stat. Phys. **50** (1988), 1–40.

[B] Berezinskii, *Destruction of Long–range Order in One–Dimensional and Two–Dimensional Systems Having a Continuous Symmetry Group I. Classical Systems*, Soviet Phys. JETP **32** (1971), 493–500.

[C] L. Chayes, *Percolation and Ferromagnetism on $\mathbb{Z}^2$: the q–State Potts Cases*, Stoc. Proc. Appl **65** ((1996)), 209–216.

[$\mathrm{CM_I}$] L. Chayes and J. Machta, *Graphical representations and Cluster Algorithms Part I: Discrete Spin Systems*, to appear in Physica A..

[$\mathrm{CM_{II}}$] L. Chayes and J. Machta, *Graphical representations and Cluster Algorithms Part II*, submitted for publication.

[FK] C.M. Fortuin and P.W. Kasteleyn, *On the Random Cluster Model I. Introduction and Relation to Other Models*, Physica **57** (1972), 536–564.

[FJB] M. E. Fisher, D. Jasnow and M.N. Barber, Phys. Rev. A **8** (1973), 1111–1124.

[FS] J. Fröhlich and T. Spencer, *The Kosterlitz–Thouless phase transition in Two–Dimensional Abelian Spin–Systems and the Coulomb Gas*, Commun. Math. Phys. **81** (1981), 527–602.

[KT] J. M. Kosterlitz and D. J. Thouless, *Ordering, Metastability and Phase Transitions in Two–Dimensional Systems*, J. Phys. C **6** (1973), 1181–1203.

[MMPf] A. Messager, S. Miracle–Sole and C. Pfister, *Correlation Inequalities and Uniqueness of the Equilibrium State for the Plane Rotator Ferromagnetic Model*, Commun. Math. Phys. **58** (1978), 19–29.

[M] P. Minnhagen, *The Two – Dimensional Coulomb Gas, Vortex Unbinding and Superfluid–Superconduction Films*, Rev. Mod. Phys. **59** (1987), 1001 – 1066.

[N] D. Nelson, *Defect mediated Phase Transitions*, Phase Transitions and Critical Phenomena, Vol. 7 (C. Domb and J. L. Lebowitz Eds.), Academic Press, London, Boston, Tokyo, etc., 1983.

[NK] D. R. Nelson and J.M. Kosterlitz, Phys. Rev. Lett. **39** (1977), 1201.

[PS] A. Patrascioiu and E. Seiler, *Phase Structure of Two–Dimensional Spin Models and Percolation*, J. Stat. Phys. **69** (1992), 573.

[T] D. J. Thouless, Phys. Rev. **187** (1969), 86.

[W] U. Wolff, *Collective Monte Carlo Updating for Spin–Systems*, Phys. Rev. Lett. **62** (1989), 361.