Robust Phase Transitions for Heisenberg and Other Models on General Trees

Robin Pemantle[*] Jeffrey E. Steif[†]

University of Wisconsin-Madison and Chalmers University of Technology

Abstract

We study several statistical mechanical models on a general tree. Particular attention is devoted to the classical Heisenberg models, where the state space is the d–dimensional unit sphere and the interactions are proportional to the cosines of the angles between neighboring spins. The phenomenon of interest here is the classification of phase transition (non-uniqueness of the Gibbs state) according to whether it is *robust*. In many cases, including all of the Heisenberg and Potts models, occurrence of robust phase transition is determined by the geometry (branching number) of the tree in a way that parallels the situation with independent percolation and usual phase transition for the Ising model. The critical values for robust phase transition for the Heisenberg and Potts models are also calculated exactly. In some cases, such as the $q \geq 3$ Potts model, robust phase transition and usual phase transition do not coincide, while in other cases, such as the Heisenberg models, we conjecture that robust phase transition and usual phase transition are equivalent. In addition, we show that symmetry breaking is equivalent to the existence of a phase transition, a fact believed but not known for the rotor model on $\mathbb{Z}^2$.

AMS 1991 subject classifications. Primary 60K35, 82B05, 82B26.
Key words and phrases. phase transitions, symmetry breaking, Heisenberg models.
Running head: Phase transitions for Heisenberg models.

[*]Research partially supported by a Presidential Faculty Fellowship and a Sloan Foundation Fellowship.
[†]Research supported by grants from the Swedish Natural Science Research Council and from the Royal Swedish Academy of Sciences.

1 Definition of the model and main results

Particle systems on trees have produced the first and most tractable examples of certain qualitative phenomena. For example, the contact process on a tree has multiple phase transitions, ([19, 12, 22]) and the critical temperature for the Ising model on a tree is determined by its branching number or Hausdorff dimension ([13, 8, 20]), which makes the Ising model intimately related to independent percolation whose critical value is also determined by the branching number (see [14]). In this paper we study several models on general infinite trees, including the classical Heisenberg and Potts models. Our aim is to exhibit a distinction between two kinds of phase transitions, *robust* and *non-robust*, as well as to investigate conditions under which robust phase transitions occur.

In many cases, including the Heisenberg and Potts models, the existence of a robust phase transition is determined by the branching number. However, in some cases (including the $q > 2$ Potts model), the critical temperature for the existence of usual phase transition is not determined by the branching number. Thus robust phase transition behaves in a more universal manner than non-robust phase transition, being a function of the branching number alone, as it is for usual phase transition for independent percolation and the Ising model. Although particle systems on trees do not always predict the qualitative behavior of the same particle system on high-dimensional lattices, it seems likely that there is a lattice analogue of non-robust phase transition, which would make an interesting topic for further research. Another unresolved question is whether there is ever a non-robust phase transition for the Heisenberg models (see Conjecture 1.9).

We proceed to define the general statistical ensemble on a tree and to state the main results of the paper. Let G be a compact metrizable group acting transitively by isometries on a compact metric space $(\mathbf{S}, d)$. It is well known that there exists a unique G–invariant probability measure on $\mathbf{S}$, which we denote by $d\mathbf{x}$. An **energy function** is any nonconstant function $H : \mathbf{S} \times \mathbf{S} \to \mathbb{R}$ that is symmetric, continuous, and d–invariant in that $H(x, y)$ depends only on $d(x, y)$. This implies that

$$H(x, y) = H(gx, gy) \;\forall\, x, y \in \mathbf{S},\; g \in G.$$

$\mathbf{S}$ together with its G–action and the function H will be called a **statistical ensemble**. Several examples with which we will be concerned are as follows.

Example 1 *The Ising model. Here* $\mathbf{S} = \{1, -1\}$ *acted on by itself (multiplicatively), d is the usual discrete metric,* $d\mathbf{x}$ *is uniform on* $\mathbf{S}$*, and* $H(x, y) = -xy$.

Example 2 *The Potts model. Here* $\mathbf{S} = \{0, 1, \ldots, q-1\}$ *for some integer* $q > 1$*,* G *is the symmetric group* S_q *with its natural action, d is the usual discrete metric,* $d\mathbf{x}$ *is uniform on* $\mathbf{S}$*, and* $H(x, y) = 1 - 2\delta_{x,y}$*. This reduces to the Ising model when* $q = 2$.

Example 3 *The rotor model. Here* $\mathbf{S}$ *is the unit circle, acted on by itself by translations,* $d(\theta, \phi) = 1 - \cos(\theta - \phi)$*,* $d\mathbf{x}$ *is normalized Lebesgue measure, and* $H(\theta, \phi) = -\cos(\theta - \phi)$.

Example 4 *The Heisenberg models for* $d \geq 1$*. In the* d*–dimensional Heisenberg model,* $\mathbf{S}$ *is the unit sphere* S^d*,* G *is the special orthogonal group with its natural action;* $d(x, y)$ *is* $1 - x \cdot y$*,* $d\mathbf{x}$ *is normalized surface measure, and* $H(x, y)$ *is again the negative of the dot product of* x *and* y*. When* $d = 1$*, we recover the rotor model.*

Let A be any finite graph, with vertex and edge sets denoted by $V(A)$ and $E(A)$ respectively, and let $\mathcal{J} : E(A) \to \mathbb{R}^+$ be a function mapping the edge set of A to the nonnegative reals which we call **interaction strengths**. We now assume that $\mathbf{S}$, G and H are given and fixed.

Definition 1 *The* **Gibbs measure** *with interaction strengths* $\mathcal{J}$ *is the probability measure* $\mu = \mu^{\mathcal{J}}$ *on* $\mathbf{S}^{V(A)}$ *whose density with respect to product measure* $d\mathbf{x}^{V(A)}$ *is given by*

$$\frac{\exp(-H^{\mathcal{J}}(\eta))}{Z}, \quad \eta \in \mathbf{S}^{V(A)}$$

where

$$H^{\mathcal{J}}(\eta) = \sum_{e=\overline{xy} \in E(A)} \mathcal{J}(e) H(\eta(x), \eta(y)),$$

and $Z = \int \exp(-H^{\mathcal{J}}(\eta))\, d\mathbf{x}^{V(A)}$ *is a normalization.*

In statistical mechanics, one wants to define Gibbs measures on infinite graphs A in which case the above definition of course does not make sense. We follow the usual approach (see [10]), in which one introduces boundary conditions and takes a weak limit of finite subgraphs increasing to A. Since the precise nature of the boundary conditions play a role here (we know this to be true at least for the Potts model with $q > 2$), we handle boundary conditions with extra care and, unfortunately, notation. We give definitions in the case of a rooted tree, though the extensions to general locally finite graphs are immediate. By a **tree**, we mean any connected loopless graph Γ where every vertex has finite degree. One fixes a vertex o of Γ which we call the **root**, obtaining a **rooted tree**. The vertex set of Γ is denoted by $V(\Gamma)$. If x is a vertex, we write $|x|$ for the number of edges on the shortest path from o to x and for two vertices x and y, we write $|x-y|$ for the number of edges on the shortest path from x to y. For vertices x and y, we write $x \leq y$ if x is on the shortest path from o to y, $x < y$ if $x \leq y$ and $x \neq y$, and $x \to y$ if $x \leq y$ and $|y| = |x| + 1$. For $x \in V(\Gamma)$, the tree $\Gamma(x)$ denotes the subtree of Γ rooted at x consisting of x and all of its descendents. We also define $\partial\Gamma$, which we refer to as the boundary of Γ, to be the set of infinite self-avoiding paths starting from o. Throughout the paper, the following assumption is in force.

ASSUMPTION: For all trees considered in this paper, the number of children of the vertices will be assumed bounded and we will denote this bound by B.

A **cutset** C is a finite set of vertices not including o such that every self-avoiding infinite path from o intersects C and such that there is no pair $x, y \in C$ with $x < y$. Given a cutset C, $\Gamma \backslash C$ has one finite component (which contains o) which we denote by C^i ("i" for inside) and we let C^o ("o" for outside) denote the union of the infinite components of $\Gamma \backslash C$. We say that a sequence $\{C_n\}$ of cutsets approaches ∞ if for all $v \in \Gamma$, $v \in C_n^i$ for all sufficiently large n.

Boundary conditions will take the form of specifications of the value of η at some cutset C. Let δ be any element of $\mathbf{S}^C$. The Gibbs measure with boundary condition δ is the probability measure $\mu_C^\delta = \mu_C^{\mathcal{J},\delta}$ on $\mathbf{S}^{C^i}$ whose density with respect to product measure $d\mathbf{x}^{C^i}$ is given by

$$\frac{\exp(-H_C^{\mathcal{J},\delta}(\eta))}{Z}, \quad \eta \in \mathbf{S}^{C^i} \tag{1.1}$$

where

$$H_C^{\mathcal{J},\delta}(\eta) = \sum_{\substack{e=\overline{xy}\in E(\Gamma)\\ x,y\in C^i}} \mathcal{J}(e) H(\eta(x), \eta(y)) + \sum_{\substack{e=\overline{xy}\in E(\Gamma)\\ x\in C^i, y\in C}} \mathcal{J}(e) H(\eta(x), \delta(y))$$

and $Z = \int \exp(-H_C^{\mathcal{J},\delta}(\eta))\, d\mathbf{x}^{C^i}$ is a normalization. When we don't include the second summand above, we call this the *free* Gibbs measure on C^i, denoted by μ_C^{free}, where $\mathcal{J}$ is suppressed in the notation. As we will see in Lemma 1.1, the free measure does not depend on C except for its domain of definition, so we can later also suppress C in the notation.

Definition 2 *A probability measure μ on $\mathbf{S}^{V(\Gamma)}$ is called a* **Gibbs state** *for the interactions $\mathcal{J}$ if for each cutset C, the conditional distribution on C^i given the configuration δ' on $C \cup C^o$ is given by $\mu_C^{\mathcal{J},\delta}$ where δ is the restriction of δ' to C. (A similar definition is used for general graphs.) Both in the case of lattices and trees (or for any graph), we say that a statistical ensemble* **exhibits a phase transition (PT) for the interaction strengths** $\mathcal{J}$ *if there is more than one Gibbs state for the interaction strengths $\mathcal{J}$.*

In the next section we will prove

Lemma 1.1 *Fix interaction strengths $\mathcal{J}$ and let C and D be any two cutsets of Γ. Then the projections of μ_C^{free} and μ_D^{free} to $\mathbf{S}^{C^i \cap D^i}$ are equal. Hence the measures μ_C^{free} have a weak limit as $C \to \infty$, denoted μ^{free}.*

For general graphs, the measures μ_C^{free} are not compatible in this way. Also, one has the following fact, which follows from Theorems 4.17 and 7.12 in [10].

Lemma 1.2 *If $\{C_n\}$ is a sequence of cutsets approaching ∞ and if for each n, $\delta_n \in \mathbf{S}^{C_n}$, then any weak subsequential limit of the sequence $\{\mu_{C_n}^{\mathcal{J},\delta_n}\}_{n\geq 1}$ is a Gibbs state for the interactions $\mathcal{J}$. In addition, if all such possible limits are the same, then there is no phase transition. (A similar statement holds for graphs other than trees.)*

We pause for a few remarks about more general graphs, before restricting our discussion to trees for the rest of the paper. Lemma 1.1 does not apply to graphs with cycles, so the existence of a unique weak limit μ^{free} is not guaranteed there, but Lemma 1.2 together with compactness tells us that there always is at least one Gibbs state. The state of knowledge about the rotor model (Example 3) on more general graphs is somewhat interesting. It is known (see [10], p.178 and p.434) that for $\mathbb{Z}^d$, $d \leq 2$, all Gibbs states are rotationally invariant when $\mathcal{J} \equiv J$ for any J (and it is believed but not known that there is a unique Gibbs state for the rotor model in this case) while for $d \geq 3$, there are values of J for which the rotor model with $\mathcal{J} \equiv J$ has a Gibbs state whose distribution at the origin is not rotationally invariant (and hence there is more than one Gibbs state). In statistical mechanics, this latter phenomenon is referred to as a *continuous symmetry breaking* since we have a continuous state space (the circle) where the interactions are invariant under a certain continuous symmetry (rotations) but there are Gibbs states which are not invariant under this symmetry. We also mention that it is proved in [6] that for the rotor model with $\mathcal{J} \equiv J$ for any J on any graph of bounded degree for which simple random walk is recurrent, all the Gibbs states are rotationally invariant. (This was then extended in [15] where the condition of boundedness of the degree is dropped and the group involved is allowed to be more general than the circle.) This however is not a sharp criterion: in [7], a graph (in fact a tree) is constructed for which simple random walk is transient but such that there is no phase transition in the rotor model when $\mathcal{J} \equiv J$ for any J. (This will also follow from Theorem 1.10 below together with the easy fact that there are trees with branching number 1 for which simple random walk is transient.) However, Y. Peres has conjectured a sharp criterion, Conjecture 1.12 below, for which our Corollary 1.11 together with the discussion following it provides some corroboration.

For the rest of this paper, we will restrict to trees. It is usually in this context that the most explicit results can be obtained and our basic goal is to determine whether there is a phase transition by comparing the interaction strengths with the "size" (branching number) of our tree. It turns out that we can only partially answer this question but the question which we can answer more completely is whether there is a *robust* phase transition, a concept which we will introduce shortly.

Definition 3 *Given $\mathcal{J}, C$ and δ defined on C, let $f^{\mathcal{J},\delta}_{C,o}$ (or $f^{\delta}_{C,o}$ if $\mathcal{J}$ is understood) denote the marginal density of $\mu^{\mathcal{J},\delta}_{C}$ at the root o.*

For any tree, recall that $\Gamma(v)$ denotes the subtree rooted at v, so that the tree $\Gamma(v)$ has vertex set $\{w \in \Gamma : v \leq w\}$. If $v \in C^i$ and we intersect C with $\Gamma(v)$, we obtain a cutset $C(v)$ for $\Gamma(v)$. We now extend Definition 3 to other marginals as follows.

Definition 4 *With $\mathcal{J}, C$ and δ as in Definition 3 and $v \in C^i$, define $f^{\mathcal{J},\delta}_{C,v}$ by replacing Γ by $\Gamma(v)$, C with $C(v)$, $\mathcal{J}$ with $\mathcal{J}$ restricted to $E(\Gamma(v))$, δ with δ restricted to $C(v)$ and o with v in Definition 3.*

It is important to note that $f^{\mathcal{J},\delta}_{C,v}$ is not the density of the projection of $\mu^{\mathcal{J},\delta}_{C}$ onto vertex v, but rather the density of a Gibbs measure with similar boundary conditions on the smaller graph $\Gamma(v)$.

Definition 5 *A statistical ensemble on a tree* Γ *exhibits a* **symmetry breaking (SB) for the interactions** $\mathcal{J}$ *if there exists a Gibbs state such that the marginal distribution at some vertex* v *is not* G*–invariant (or equivalently is not* $d\mathbf{x}$*).*

The following proposition which will be proved in Section 2 is interesting since it establishes the equivalence of PT and SB for general trees and general statistical ensembles, something not known for general graphs, see the remark below.

Proposition 1.3 *Consider a statistical ensemble on a tree* Γ *with interactions* $\mathcal{J}$*. The following four conditions are equivalent.*
(i) There exists a vertex v *such that for any sequence of cutsets* $C_n \to \infty$*, there exist boundary conditions* δ_n *on* C_n *such that*

$$\inf_n \|f^{\delta_n}_{C_n,v} - 1\|_\infty \neq 0.$$

(ii) There exists a vertex v*, a sequence of cutsets* $C_n \to \infty$ *and boundary conditions* δ_n *on* C_n *such that*

$$\inf_n \|f^{\delta_n}_{C_n,v} - 1\|_\infty \neq 0.$$

(iii) The system satisfies SB.
(iv) The system satisfies PT.

We now fix a distinguished element in $\mathbf{S}$, hereafter denoted $\hat{0}$. The notation $\mu_C^{\mathcal{J},+}$ denotes $\mu_C^{\mathcal{J},\delta}$ when δ is the constant function $\hat{0}$. In the case $\mathcal{J} \equiv J$, we denote this simply $\mu_C^{J,+}$. We will be particularly concerned about whether $\mu_C^{\mathcal{J},+} \to \mu^{\text{free}}$ weakly, as $C \to \infty$.

Definition 6 *A statistical ensemble on a tree* Γ *exhibits a* **symmetry breaking with plus boundary conditions (SB+) for the interactions** $\mathcal{J}$ *if there exists a vertex* v *and a sequence of cutsets* $C_n \to \infty$ *such that*

$$\inf_n \|f^{\mathcal{J},+}_{C_n,v} - 1\|_\infty \neq 0.$$

Note that by symmetry, SB+ does not depend on which point of $\mathbf{S}$ is chosen to be $\hat{0}$.

In Section 4.1 we will prove:

Proposition 1.4 *For the rotor model on a tree, SB is equivalent to SB+.*

We conjecture but cannot prove the stronger statement:

Conjecture 1.5 *For any Heisenberg model on any graph, SB is equivalent to SB+.*

Remarks: (i) By Proposition 1.3, we have that SB+ implies SB for any statistical ensemble on a tree. While Proposition 1.3 tells us that PT and SB are equivalent for any statistical ensemble on a tree, we note that such a result is not even known for the rotor model on $\mathbb{Z}^2$ where it has been established that for all J, all Gibbs states are rotationally invariant for $\mathcal{J} \equiv J$ but where it has not been established that there is no phase transition. A weaker form of the above conjecture would be that SB+ and SB are equivalent for all Heisenberg models on trees. This is Problem 4.1 in Section 4. An extension to graphs with cycles would seem to entail a different kind of reasoning, perhaps similar to the inequalities of Monroe and Pearce [16] which fall just short of proving Conjecture 1.5 for the rotor model. (ii) The fact that PT and SB+ are equivalent when the rotor model is replaced by the Ising model is an immediate consequence of the fact that the probability measure is stochastically increasing in the boundary conditions. More generally, it is also the case that PT and SB+ are equivalent for the Potts models (see [2]).

We now consider the idea of a *robust phase transition* where we investigate if the boundary conditions on a cutset have a nontrivial effect on the root even when the interactions along the cutset are made arbitrarily small but fixed.

Given parameters $J > 0$ and $J' \in (0, J]$ and a cutset C of Γ, let $\mathcal{J}(J', J, C)$ be the function on $E(\Gamma)$ which is J on edges in C^i and J' on edges connecting C^i to C (the values elsewhere being irrelevant). Let $f^{J',J,+}_{C_n,o}$ denote the marginal at the root o of the measure $\mu^{J',J,+}_C := \mu^{\mathcal{J}(J',J,C),+}_C$.

Definition 7 *The statistical ensemble on the tree* Γ *has a* **robust phase transition (RPT) for the parameter** $J > 0$ *if for every* $J' \in (0, J]$

$$\inf_C \|f^{J',J,+}_{C,o} - 1\|_\infty \neq 0$$

where the inf *is taken over all cutsets* C*.*

Remarks: In the case $\mathcal{J} \equiv J$, by taking $J' = J$, it is clear that a RPT implies SB+ (which in turn implies SB and PT). Note that in this case, RPT is stronger than SB+ not only because J' can be any number in $(0, J]$ and the root o must play the role of v but also because in SB+, we only require that for *some* sequence of cutsets going to infinity, the marginal at the vertex v stays away from uniform while in RPT, we require this for *all* cutsets going to infinity. We note also that with some care, this definition makes sense for general graphs, and that the issue of robustness of phase transition on general graphs is worth investigating, although we do not do so here.

Our first theorem gives criteria based on J and the branching number of Γ (which will now be defined) for robust phase transition to occur for the Heisenberg models. A little later on, we will have an analogous result for the Potts models. In [9], Furstenberg introduced the notion of the Hausdorff dimension of a tree (or more accurately of the boundary of the tree). This was further investigated by Lyons ([14]) using the term branching number instead. The **branching number** of a tree Γ, denoted $br(\Gamma)$, is a real number greater than or equal to one that measures the average number of branches per vertex of the tree. More precisely, the **branching number** of Γ is defined by

$$br\,\Gamma := \inf\left\{\lambda > 0; \inf_C \textstyle\sum_{x\in C} \lambda^{-|x|} = 0\right\}$$

where the second infimum is over all cutsets C. The branching number is a measure of the average number of branches per vertex of Γ. It is less than or equal to $\liminf_{n\to\infty} M_n^{1/n}$, where $M_n := |\{x \in \Gamma; |x| = n\}|$, and takes more of the structure of Γ into account than does this latter growth rate. For sufficiently regular trees, such as homogeneous trees or, more generally, Galton-Watson trees, $br\,\Gamma = \lim_{n\to\infty} M_n^{1/n}$ ([14]). We also mention that the branching number is the exponential of the Hausdorff dimension of $\partial\Gamma$ where the latter is endowed with the metric which gives distance e^{-k} to two paths which split off after k steps. As indicated earlier, the branching number has been an important quantity in previous investigations. More specifically, in [13] and [14], the critical values for independent percolation and for phase transition in the Ising model on general trees are explicitly computed in terms of the branching number.

For each $J \geq 0$, define a continuous strictly positive probability density function $K_J : \mathbf{S} \to \mathbb{R}^+$ by

$$K_J(u) := C(J)^{-1} \exp(-JH(u, \hat{0})) \tag{1.2}$$

where $C(J) = \int \exp(-JH(w, \hat{0}))\, d\mathbf{x}(w)$ is a normalizing constant, and more generally let $K_{J,y} : \mathbf{S} \to \mathbb{R}^+$ be given by

$$K_{J,y}(u) := C(J)^{-1} \exp(-JH(u, y)) \tag{1.3}$$

(noting that $K_{J,\hat{0}} = K_J$). Let $\mathcal{K}_J$ denote the convolution operator on the space $L^2(\mathbf{S}, d\mathbf{x})$ given by the formula

$$\mathcal{K}_J f(u) := \int_{\mathbf{S}} f(x) K_{J,x}(u) d\mathbf{x}(x) \; . \tag{1.4}$$

Note that by the assumed invariance $\int_{\mathbf{S}} \exp(-JH(w,y))\, d\mathbf{x}(w)$ is independent of y and that $f \geq 0$ and $\int_{\mathbf{S}} f(x) d\mathbf{x}(x) = 1$ imply that $\mathcal{K}_J f \geq 0$ and $\int_{\mathbf{S}} \mathcal{K}_J f(x) d\mathbf{x}(x) = 1$. We extend the above notation to cover the case where f is a pointmass δ_y at y by defining in that case

$$\mathcal{K}_J \delta_y(u) := K_{J,y}(u). \tag{1.5}$$

We will now give the exact critical parameter J for RPT for the Heisenberg models. For any $d \geq 1$, let

$$\rho^d(J) := \frac{\int_{-1}^{1} r e^{Jr}(1-r^2)^{\frac{d}{2}-1} dr}{\int_{-1}^{1} e^{Jr}(1-r^2)^{\frac{d}{2}-1} dr}.$$

When $d = 1$ (rotor model), this is (by a change of variables) the first Fourier coefficient of K_J ($\int_{\mathbf{S}} K_J(\theta)\cos(\theta) d\theta$) which is perhaps more illustrative. When $d = 2$, this is the first Legendre coefficient of e^{Jr} (properly normalized) and for $d \geq 3$, this is the first so-called ultraspherical coefficient of e^{Jr} (properly normalized).

Theorem 1.6 *Let $d \geq 1$.*
(i) If $br(\Gamma)\rho^d(J) < 1$, then the d–dimensional Heisenberg model on Γ with parameter J does not exhibit a robust phase transition.
(ii) If $br(\Gamma)\rho^d(J) > 1$, then the d–dimensional Heisenberg model on Γ with parameter J exhibits a robust phase transition.

Remark: It is easy to see that $\lim_{d\to\infty} \rho^d(J) = 0$ which says that it is harder to obtain a robust phase transition on higher dimensional spheres. This is consistent with the fact that it is in some sense harder to have a phase transition for the rotor model than in the Ising model (0-dimensional sphere); this latter fact can be established using the ideas in [18].

A simple computation shows that the derivative of $\rho^d(J)$ with respect to J is the variance of a random variable whose density function is proportional to $e^{Jr}(1-r^2)^{d/2-1}$ on $[-1, 1]$, thereby obtaining the following lemma.

Lemma 1.7 *For any $d \geq 1$, we have that $\rho^d(J)$ is a strictly increasing function of J.*

Theorem 1.6 and Lemma 1.7 together with the fact that for any $d \geq 1$, $\rho^d(J)$ is a continuous function of J which approaches 0 as $J \to 0$ and approaches 1 as $J \to \infty$ give us the following corollary.

Corollary 1.8 *For any Heisenberg model with $d \geq 1$ and any tree Γ with branching number larger than 1, let $J_c = J_c(\Gamma, d)$ be such that $br(\Gamma)\rho^d(J_c) = 1$. Then there is a robust phase transition for the d–dimensional Heisenberg model on Γ if $J > J_c$ and there is no such robust phase transition for $J < J_c$.*

For the Heisenberg models, we believe that phase transition and robust phase transition coincide and therefore we have the following conjecture.

Conjecture 1.9 *For any $d \geq 1$, if $br(\Gamma)\rho^d(J) < 1$, then the d–dimensional Heisenberg model on Γ with parameter J does not exhibit a phase transition.*

We can however obtain the following weaker form of this conjecture which is valid for all statistical ensembles.

Theorem 1.10 *If $br(\Gamma) = 1$, then there is no phase transition for any statistical ensemble on Γ with bounded $\mathcal{J}$.*

Theorems 1.6(ii) and 1.10 together with the facts that RPT implies PT and that for any $d \geq 1$, $\lim_{J\to\infty} \rho^d(J) = 1$ immediately yield the following corollary.

Corollary 1.11 *For any Heisenberg model with $d \geq 1$ and for any tree Γ, there is a phase transition for the tree Γ for some value of the parameter J if and only if $br(\Gamma) > 1$.*

Since it is known (see [14]) that $br(\Gamma) > 1$ if and only if there is some $p < 1$ with the property that when performing independent percolation on Γ with parameter p, there exists a.s. an infinite cluster on which simple random walk is transient, the above corollary yields the following conjecture of Y. Peres for the special case of trees of bounded degree.

Conjecture 1.12 *For any graph A, the rotor model exhibits a phase transition for some J if and only if there is some $p < 1$ with the property that performing independent bond percolation on A with parameter p, there exists a.s. an infinite cluster on which simple random walk is transient.*

Recall that the rotor model on the graph A exhibits no SB for any parameter J if A is recurrent for simple random walk, which is of course consistent with the above conjecture. Note that, on the other hand, the standard Ising model does exhibit a phase transition on $\mathbb{Z}^2$, a graph which is recurrent (as are its subgraphs) for simple random walk.

The next result states the critical value for RPT for the Potts models.

Theorem 1.13 *Consider the Potts model with $q \geq 2$ and let*

$$\alpha_J = \frac{e^J - e^{-J}}{e^J + (q-1)e^{-J}}.$$

(i) If $br(\Gamma)\alpha_J < 1$, then the Potts model on Γ with parameter J does not exhibit a robust phase transition.
(ii) If $br(\Gamma)\alpha_J > 1$, then the Potts model on Γ with parameter J exhibits a robust phase transition.

Remarks:
(i) $d\alpha_J/dJ > 0$ and so there is a critical value of J depending on $br(\Gamma)$ analogous to in Corollary 1.8 for the Heisenberg models.
(ii) Note that when $q = 2$ (the Ising model), this formula agrees with the formula for the Heisenberg models when one formally sets $d = 0$ in the formula

$$\rho^d(J) = \int_{S^d} (x \cdot \hat{0}) K_J(x)\, d\mathbf{x}(x),$$

the latter being obtained by a change of variables.

To point out the subtlety involved in Conjecture 1.9, we continue to discuss the Potts model, a case in which the analogue of Conjecture 1.9 fails. Our final result tells us that phase transitions (unlike robust phase transitions) in the Potts model with $q > 2$ cannot be determined by the branching number.

Theorem 1.14 *Given any integer $q > 2$, there exist trees Γ_1 and Γ_2 and a nontrivial interval I such that $br(\Gamma_1) < br(\Gamma_2)$ and for any $J \in I$, there is a phase transition for the q–state Potts model with parameter J on Γ_1 but no such phase transition on Γ_2.*

Remarks:
(i) Γ_1 and Γ_2 can each be taken to be spherically symmetric which means that for all k, all vertices at the kth generation have the same number of children.
(ii) In the case $q = 2$, more is known. In [13], the critical value for phase transition in the Ising model is found and corresponds to what is obtained in Theorem 1.13 above. It follows that there is never a non-robust phase transition except possibly at the critical value. However, a sharp capacity criterion exists [20] for phase transition for the Ising model (settling the issue of phase transition at the critical parameter) and using this criterion, one can show that phase transition and robust phase transition correspond even at criticality. The arguments of [20] cannot be extended to the Potts model for $q > 2$ because the operator $\mathcal{K}_J$, acting on a certain likelihood function, when conjugated by the logarithm is not concave in this case. Theorems 1.13 and 1.14 together tell us that there is indeed a non-robust phase transition when $q > 2$ for a nontrivial interval of J.

The rest of the paper is devoted to the proofs of the above results. In Section 2, we collect several lemmas that apply to general statistical ensembles, including the basic recursion formula (Lemma 2.2) that allows us to analyze general statistical ensembles on trees, prove Lemma 1.1 and Proposition 1.3 as well as provide some background concerning Heisenberg models (showing that they satisfy the more general hypotheses of Theorems 3.1 and 3.2 given later on) and the more general notion of distance regular spaces. Section 3 is devoted to the proofs of Theorems 3.1 and 3.2. In Section 4, we use these theorems to find the critical parameters for robust phase transition in the Heisenberg and Potts models, Theorems 1.6 and 1.13, as well as prove Proposition 1.4. Section 5 discusses the special case of trees of branching number 1, proving Theorem 1.10. Finally, in Section 6, Theorem 1.14 is proved.

2 Basic background results

In this section, we collect various background results which will be needed to prove the results described in the introduction. We begin with a subsection describing results pertaining to trees that hold for general statistical ensembles. After discussing the concept of a distance regular space in Section 2.2, we specialize to Heisenberg models (the most relevant family of continuous distance regular models) in Section 2.3 and then to distance regular graphs in Section 2.4.

2.1 The fundamental recursion and other lemmas

We start off with two lemmas exploiting the recursive structure of trees.

Let $\mathbf{S}, G$ and H be a statistical ensemble. Let A_1 and A_2 be two disjoint finite graphs, with distinguished vertices $v_1 \in V(A_1)$ and $v_2 \in V(A_2)$. Let $\mathcal{J}_1$ and $\mathcal{J}_2$ be interaction functions for A_1 and A_2, i.e., positive functions on $E(A_1)$ and $E(A_2)$ respectively. For any $C_1 \subseteq V(A_1) \setminus \{v_1\}$ (possibly empty) and any $C_2 \subseteq V(A_2)$, and for any $\delta_1 \in \mathbf{S}^{C_1}$ and $\delta_2 \in \mathbf{S}^{C_2}$, we have measures $\mu_i := \mu_{C_i}^{\mathcal{J}_i, \delta_i}$, $i = 1, 2$ on $\mathbf{S}^{V(A_i) \setminus C_i}$ defined (essentially) by (1.1). Abbreviate $H_{C_i}^{\mathcal{J}_i, \delta_i}$ (which has the obvious meaning) by H_i. Let A be the union of A_1 and A_2 together with an edge connecting v_1 and v_2. Let $C = C_1 \cup C_2$, $\mathcal{J}$ extend each $\mathcal{J}_i$ and the value of the new edge be given the value J, δ extend each δ_i and denote $\mu_C^{\mathcal{J}, \delta}$ (a probability measure on $\mathbf{S}^{(V(A_1) \setminus C_1) \cup (V(A_2) \setminus C_2)}$) by μ and $H_C^{\mathcal{J}, \delta}$ (again having the obvious meaning) by H. The identity

$$H = H_1 + H_2 + JH(\eta(v_1), \eta(v_2)) \tag{2.1}$$

leads to the following lemma.

Lemma 2.1 *The measure μ satisfies*

$$\frac{d\mu}{d(\mu_1 \times \mu_2)} = c \exp[-JH(\eta_1(v_1), \eta_2(v_2))], \tag{2.2}$$

where

$$c = \left[\int \int \exp(-JH(\eta_1(v_1), \eta_2(v_2))) \, d\mu_1(\eta_1) \, d\mu_2(\eta_2) \right]^{-1}$$

is a normalizing constant. Let f_i denote the marginal density of μ_i at v_i, $i = 1, 2$, and f denotes the marginal density of μ at v_1. Then the projection $\mu^{(1)}$ of μ onto $\mathbf{S}^{V(A_1) \setminus C_1}$ satisfies

$$\mu^{(1)} = c \int \int \mu_{1,y} f_1(y) f_2(z) \exp(-JH(y, z)) \, d\mathbf{x}(z) \, d\mathbf{x}(y) \tag{2.3}$$

for some normalizing constant c, where $\mu_{1,y}$ denotes the conditional distribution of μ_1 given $\eta(v_1) = y$. Consequently,

$$f(y) = c f_1(y) \int f_2(z) \exp(-JH(y, z)) \, d\mathbf{x}(z), \tag{2.4}$$

where c normalizes f to be a probability density.

Proof. The relation (2.2) follows from (2.1) and the defining equation (1.1). From this it follows that the measure μ on pairs (η_1, η_2) makes η_1 and η_2 conditionally independent given $\eta_1(v_1)$ and $\eta_2(v_2)$. Hence the conditional distribution of $\mu^{(1)}$ given $\eta_1(v_1) = y$ and $\eta_2(v_2) = z$ is just $\mu_{1,y}$. Next, (2.2) and the last fact yield (2.3). The marginal of $\mu_{1,y}$ at v_1 is just δ_y, and so (2.3) yields (2.4). $\square$

A tree Γ may be built up from isolated vertices by the joining operation described in the previous lemma. The decompositions in Lemma 2.1 may be applied inductively to derive a fundamental recursion for marginals. This recursion, Lemma 2.2 below, expresses the marginal distribution at the root of Γ as a pointwise product of marginals at the roots of each of the generation 1 subtrees, each convolved with a kernel K_J. The normalized pointwise product will be ubiquitous throughout what follows, so we introduce notation for it.

Definition 8 *If $f_1, \ldots, f_k$ are nonnegative functions on* $\mathbf{S}$ *with $\int f_i \, d\mathbf{x} = 1$ for each i, let $\bigodot_k(f_1, \ldots, f_k)$ denote the normalized pointwise product,*

$$\bigodot_k(f_1, \ldots, f_k)(x) = \frac{\prod_{i=1}^k f_i(x)}{\int \prod_{i=1}^k f_i(y) \, d\mathbf{x}(y)}$$

whenever this makes sense, e.g., when each f_i is in $L^k(d\mathbf{x})$ and the product is not almost everywhere zero. Let $\bigodot$ denote the operator which for each k is $\bigodot_k$ on each k-tuple of functions. There is an obvious associativity property, namely $\bigodot(\bigodot(f, g), h) = \bigodot(f, g, h)$, which may be extended to arbitrarily many arguments.

Lemma 2.2 (Fundamental recursion) *Given a tree Γ, a cutset C, interactions $\mathcal{J}$, boundary condition δ and $v \in C^i$, let $\{w_1, \ldots, w_k\}$ be the children of v. Let $J_1, \ldots, J_k$ denote the values of $\mathcal{J}(v, w_1), \ldots, \mathcal{J}(v, w_k)$. Then*

$$f_{C,v}^{\mathcal{J},\delta} = \bigodot(\mathcal{K}_{J_1} f_{C,w_1}^{\mathcal{J},\delta}, \ldots, \mathcal{K}_{J_k} f_{C,w_k}^{\mathcal{J},\delta}), \tag{2.5}$$

where when $w_i \in C$, $f_{C,w_i}^{\mathcal{J},\delta}$ is taken to be the point mass at $\delta(w_i)$ and convention (1.5) is in effect.

Proof. Passing to the subtree $\Gamma(v)$, we may assume without loss of generality that $v = o$. Also assume without loss of generality that $w_1, \ldots, w_k$ are numbered so that for some s, $w_i \in C^i$ for $i \leq s$ and $w_i \in C$ for $i > s$. For $i \leq s$, let $C(w_i) = C \cap \Gamma(w_i)$. For such i, by definition, $f_i := f^{\mathcal{J},\delta}_{C,w_i}$ is the marginal at w_i of the measure $\mu_i := \mu^{\mathcal{J},\delta}_{C(w_i),w_i}$ on configurations on $\Gamma(w_i) \cap C^i$, where $\mathcal{J}$ and δ are restricted to $E(\Gamma(w_i))$ and $C(w_i)$ respectively. Let Γ_r denote the induced subgraph of Γ whose vertices are the union of $\{o\}$, $\Gamma(w_1), \ldots, \Gamma(w_r)$. We prove by induction on r that the density g_r at the root of Γ_r of the analogue of $\mu^{\mathcal{J},\delta}_C$ for Γ_r is equal to

$$\bigodot(\mathcal{K}_{J_1} f^{\mathcal{J},\delta}_{C,w_1}, \ldots, \mathcal{K}_{J_r} f^{\mathcal{J},\delta}_{C,w_r})\,;$$

The case $r = k$ is the desired conclusion.

To prove the $r = 1$ step, use (2.4) with $v_1 = o$, $A_1 = \{o\}$, $C_1 = \emptyset$, $v_2 = w_1$, $A_2 = \Gamma(w_1)$ and $C_2 = C(w_1)$. If $w_1 \in C$, the $r = 1$ case is trivially true, so assume $s \geq 1$. The measure μ_1 is uniform on $\mathbf{S}$ since $C(v) = \emptyset$. Thus from (2.4) we find that

$$g_1(y) = c \int e^{-J_1 H(y,z)} f_1(z)\, dz = (\mathcal{K}_{J_1} f_1)(y)$$

which proves the $r = 1$ case.

For $1 < r \leq s$, use (2.4) with $A_1 = \Gamma_{r-1}$, $v_1 = o$, $C_1 = \Gamma_{r-1} \cap C$, $A_2 = \Gamma(w_r)$, $v_2 = w_r$ and $C_2 = \Gamma(w_r) \cap C$. Using (2.4) we find that

$$\begin{aligned} g_r(y) &= c g_{r-1}(y) \int e^{-J_r H(y,z)} f_r(z)\, d\mathbf{x}(z) \\ &= c g_{r-1}(y) (\mathcal{K}_{J_r} f_r)(y) \\ &= (\bigodot(g_{r-1}, \mathcal{K}_{J_r} f_r))(y)\,. \end{aligned}$$

By associativity of $\bigodot$ the induction step is completed for $r \leq s$.

Finally, if $r > s$, then the difference between $H(\eta)$ on Γ_{r-1} and $H(\eta)$ on Γ_r is just $-J_r H(\eta(o), \delta(w_r))$, so

$$g_r(y) = c g_{r-1}(y) \exp(-J_r H(y, \delta(w_r))) = \left(\bigodot(g_{r-1}, \mathcal{K}_{J_r} f_r)\right)(y)$$

by the convention (1.5), and associativity of $\bigodot$ completes the induction as before. $\square$

Another consequence of Lemma 2.1 is Lemma 1.1, giving the existence of a natural and well defined free boundary measure.

Proof of Lemma 1.1. Observe that in (2.3), if $f_2 \equiv 1$ then the integral against z is independent of y, so one has $\mu^{(1)} = \mu_1$. Let F be any cutset and $w \in F^i$ be chosen so each of its children $v_1, \ldots, v_k$ is in F. Applying our observation inductively to eliminate each child of w in turn, we see that the projection of $\mu_F^{\rm free}$ onto $\mathbf{S}^{F^i \setminus \{w\}}$ is just $\mu_{F'}^{\rm free}$ where $F' = F \cup \{w\} \setminus \{v_1, \ldots, v_k\}$.

Given cutsets C and D with $D \cap C^i \neq \emptyset$, choose $v \in D \cap C^i$ and $w \geq v$ maximal in C^i. Then all children of w are in C. Applying the previous paragraph with $F = C$, we see that $\mu_C^{\rm free}$ agrees with $\mu_{F'}^{\rm free}$. Continually reducing in this way, we conclude that on $C^i \cap D^i$ $\mu_C^{\rm free}$ agrees with $\mu_Q^{\rm free}$ where Q is the exterior boundary of $C^i \cap D^i$. The same argument shows that $\mu_D^{\rm free}$ agrees with $\mu_Q^{\rm free}$, which finishes the proof of the lemma. $\square$

According to Lemma 2.2, if, for $J > 0$, we define $\mathcal{P}(J)$ to be the smallest class of densities containing each $K_{J',y}$ for $J' \in (0, J]$ and $y \in \mathbf{S}$ and closed under $\mathcal{K}_{J'}$ for $J' \in (0, J]$ and $\bigodot$, then, when $\mathcal{J}$ is strictly positive and bounded by J, each density $f_{C,v}^{\mathcal{J},\delta}$ is an element of $\mathcal{P}(J)$. Similarly, if $\mathcal{P}_+(J)$ is taken to be the smallest class of densities containing each $K_{J'}$ for $J' \in (0, J]$ and closed under $\mathcal{K}_{J'}$ for $J' \in (0, J]$ and $\bigodot$, then, when $\mathcal{J}$ is strictly positive and bounded by J, each density $f_{C,v}^{\mathcal{J},+}$ is an element of $\mathcal{P}_+(J)$. We also let $\mathcal{P} := \bigcup_{J>0} \mathcal{P}(J)$ and $\mathcal{P}_+ := \bigcup_{J>0} \mathcal{P}_+(J)$.

This leads to the following lemma whose proof is left to the reader.

Lemma 2.3 *Suppose the interaction strengths $\{\mathcal{J}(e)\}$ are bounded above by some constant. Then there exist constants $0 < B_{\rm min} < B_{\rm max}$ such that for every C, δ and $v \in C^i$, the one-dimensional marginal of μ_C^δ at v is absolutely continuous with respect to $d\mathbf{x}$ with a density function in $[B_{\rm min}, B_{\rm max}]$. It follows, since the above properties are closed under convex combinations, that all one-dimensional marginals of any Gibbs state have densities in $[B_{\rm min}, B_{\rm max}]$. Similarly, the k-dimensional marginals have densities in the interval $[B_{\rm min}^{(k)}, B_{\rm max}^{(k)}]$ for some constants $0 < B_{\rm min}^{(k)} < B_{\rm max}^{(k)}$. In addition, the family of all one-dimensional densities which arise as above is an equicontinuous family.*

The usefulness of the equicontinuity property is that the following easily proved lemma (whose proof is also left to the reader) tells us that in determining weak convergence to $d\mathbf{x}$, it is equivalent to look to see if there is convergence in L^∞ of the associated densities to 1.

Lemma 2.4 *Let (X, d) be a compact metric space and μ a probability measure on X with full support. If $\{f_n\}$ is an equicontinuous family of probability densities (with respect to μ), then*

$$\lim_{n\to\infty} \|f_n - 1\|_\infty = 0 \textit{ if and only if } \lim_{n\to\infty} f_n d\mu = \mu \textit{ weakly }.$$

Using this, we can prove the equivalence of phase transition and symmetry breaking on trees (Proposition 1.3).

Proof of Proposition 1.3. (i) implies (ii) is trivial. For (ii) implying (iii), assume we have a vertex v, a sequence of cutsets $C_n \to \infty$ and boundary conditions δ_n on C_n such that

$$\inf_n \|f^{\delta_n}_{C_n,v} - 1\|_\infty \neq 0.$$

Clearly we obtain the same result if we change δ_n on $C_n \setminus \Gamma(v)$ to anything, in particular, if we take no (i.e., free) boundary condition there. We then take any weak limit of these measures as $n \to \infty$. This will yield a Gibbs state and by the first line of the proof of Lemma 1.1, together with Lemma 2.4, the marginal density at v of this Gibbs state is not 1, which proves (iii). (iii) implies (iv) is also trivial of course. To see that (iv) implies (i), note that if there is PT, then there exists an extremal Gibbs state $\mu \neq \mu^{\text{free}}$. Choose a cutset C such that $\mu \neq \mu^{\text{free}}$ when restricted to C^i. If (i) fails, then for all $v \in C$, there exists a sequence of cutsets $C_n \to \infty$ such that for all boundary conditions δ_n on C_n we have that

$$\inf_n \|f^{\delta_n}_{C_n,v} - 1\|_\infty = 0. \tag{2.6}$$

Clearly, because of the geometry, $\{C_n\}$ can be chosen independent of v. Since μ is extremal, it is known (see Theorem 7.12(b) in [10], p. 122) that there exist boundary conditions δ'_n on C_n so that $\mu^{\delta'_n}_{C_n} \to \mu$ weakly. However, by (2.6) and Lemma 2.2, μ must equal μ^{free} on C^i, a contradiction. □

2.2 Distance regular spaces

Our primary interest in this paper is in the Heisenberg models. Nevertheless, it turns out that many of the properties of the Heisenberg model hold in the more general context of distance regular spaces. A **distance regular graph** is a finite graph for which the size of the set $\{z : d(x,z) = a, d(y,z) = b\}$ depends on x and y only through the value of $d(x,y)$ where $d(x,y)$ is the usual graph distance between x and y. We generalize this by saying that the metric space $(\mathbf{S}, d)$ with probability measure $d\mathbf{x}$ is **distance regular** if the law of the pair $(d(x,Z), d(y,Z))$ when Z has law $d\mathbf{x}$ depends only on $d(x,y)$. In particular, when the action of G on $\mathbf{S}$ is distance transitive (in addition to preserving d and $d\mathbf{x}$), meaning that (x,y) can be mapped to any (x',y') with $d(x,y) = d(x',y')$, it follows easily that $(\mathbf{S}, d, d\mathbf{x})$ is distance regular. All the examples we have mentioned so far are distance transitive (and hence distance regular) except for the rotor model which is still distance regular. (For an example of a graph showing that the full automorphism group acting distance transitively is strictly stronger than the assumption of distance regularity, see [1] or *Additional Result* **23b** of [5].)

We present some of the background in this generality not because we are fond of gratuitous generalization but because we find the reasoning clearer, and because it seems reasonable that someone in the future might study a particle system whose spin states are elements of some distance regular space, such as real projective space or the discrete n-cube. The primary consequence of distance regularity is that it allows one to define a commutative convolution on a certain subspace of L^2.

Definition 9 *Let $L^2(\mathbf{S})$ denote the space $L^2(d\mathbf{x})$, and let $L^2(\mathbf{S}/\hat{0})$ denote the space of functions $f \in L^2(\mathbf{S})$ for which $f(x)$ depends only on $d(x,\hat{0})$. For $f \in L^2(\mathbf{S}/\hat{0})$, define a function $\overline{f}$ on $\{d(\hat{0},y)\}_{y\in\mathbf{S}}$ by $\overline{f}(r) := f(x)$ where x is such that $d(\hat{0},x) = r$.*

Definition 10 *If $(\mathbf{S}, d\mathbf{x})$ is distance regular, define a commutative convolution operation on $L^2(\mathbf{S}/\hat{0}) \times L^2(\mathbf{S}/\hat{0})$ by*

$$f * h(x) := \int_{\mathbf{S}} h(y)\overline{f}(d(x,y))\, d\mathbf{x}(y) = \int_{[0,\infty)^2} \overline{f}(u)\overline{h}(v)\, d\pi_x(u,v)$$

where π_x is the law of $(d(x,Z), d(\hat{0},Z))$ for a variable Z with law $d\mathbf{x}$. It is clear from the definition of a distance regular space that $(d(x,Z), d(\hat{0},Z))$ and $(d(\hat{0},Z), d(x,Z))$ are equal in distribution implying that $f*h = h*f$ and that, since π_x only depends on $d(x,\hat{0})$, $f, h \in L^2(\mathbf{S}/\hat{0})$ implies that $f * h \in L^2(\mathbf{S}/\hat{0})$.

The following lemma is straightforward and left to the reader.

Lemma 2.5 *For all $J \geq 0$, $K_J \in L^2(\mathbf{S}/\hat{0})$ and for all $h \in L^2(\mathbf{S})$, $\mathcal{K}_J(h)(x)$ (defined in (1.4)) is equal to $\int_{\mathbf{S}} h(y)\overline{K_J}(d(x,y))\, d\mathbf{x}(y)$. In particular, if $(\mathbf{S}, d\mathbf{x})$ is distance regular, then the operators $\mathcal{K}_J$ map $L^2(\mathbf{S}/\hat{0})$ into itself and $\mathcal{K}_J(h) = K_J * h$ for all $h \in L^2(\mathbf{S}/\hat{0})$.*

We believe that for most distance regular spaces, one can verify the necessary hypotheses of Theorems 3.1 and 3.2 below in the same way as we will do for the Heisenberg models in detail in the next section. Doing this however would take us too far afield and so we content ourselves with pointing out to the reader that much of this probably can be done, and after analyzing the Heisenberg models in Section 2.3, explain how to carry much of this out in the context of distance regular graphs in Section 2.4.

2.3 Heisenberg models

In this subsection, we consider Example 4 in Section 1 and so we have $\mathbf{S} = S^d$, $d \geq 1$, the unit sphere in $(d+1)$–dimensional Euclidean space with the corresponding $G, d, d\mathbf{x}$ and H. Recall that this is distance transitive for $d \geq 2$ (and hence distance regular) and distance regular for $d = 1$. The following lemma allows us to set up coordinates in which our bookkeeping will be manageable. It is certainly well known.

Lemma 2.6 *For any $d \geq 1$, there exist real–valued functions $\psi_0, \psi_1, \psi_2, \ldots \in L^2(\mathbf{S}/\hat{0})$ ($\mathbf{S} = S^d$), orthogonal under the inner product $\langle f, g\rangle = \int_{\mathbf{S}} f\overline{g}\, d\mathbf{x}$, such that ψ_n is a polynomial of degree exactly n in $x \cdot \hat{0}$, and such that the following properties hold.*
(1) $\psi_0(x) \equiv 1$ and $\psi_1(x) = x \cdot \hat{0}$.
(2) $1 = \psi_j(\hat{0}) = \sup_{x\in\mathbf{S}} |\psi_j(x)|$, for all j.
(3) $\psi_i\psi_j = \sum_{r\geq 0} q_{ij}^r \psi_r$, where the coefficients q_{ij}^r are nonnegative and $\sum_r q_{ij}^r = 1$.
(4) $\psi_i * \psi_j = \gamma_j \delta_{ij}\psi_j$, where $\gamma_j := \psi_j * \psi_j(\hat{0}) = \int \psi_j^2(x)\, d\mathbf{x}(x)$.
(5) The functions ψ_j are eigenfunctions of any convolution operator, that is, $f * \psi_j = c\psi_j$ for any $f \in L^2(\mathbf{S}/\hat{0})$.
(6) Any $f \in L^2(\mathbf{S}/\hat{0})$ can be written as a convergent series $f(x) = \sum_{j\geq 0} a_j(f)\psi_j(x)$ (in the L^2 sense), where the complex numbers $a_j(f)$ are given by $a_j(f) := \gamma_j^{-1} \int f(x)\psi_j(x)\, d\mathbf{x}(x)$.
(7) For $f, g \in L^2(\mathbf{S}/\hat{0})$, we have $a_j(f * g) = \gamma_j a_j(f) a_j(g)$.

Proof. For each $\alpha, \beta > -1$, define the Jacobi polynomials $\{\mathbf{P}_n^{(\alpha,\beta)}(r)\}_{n\geq 0}$ by

$$(1-r)^\alpha (1+r)^\beta \mathbf{P}_n^{(\alpha,\beta)}(r) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dr^n}\left[(1-r)^{n+\alpha}(1+r)^{n+\beta}\right]. \qquad (2.7)$$

(The Jacobi polynomials are usually defined differently in which case (2.7) becomes what is known as Rodrigues' formula but we shall use (2.7) as our definition; when $\alpha = \beta$, which is the case relevant to us, these are the ultraspherical polynomials.)

For any given $d \geq 1$, we let, for $n \geq 0$,

$$\psi_n(x) := \frac{\mathbf{P}_n^{(\frac{d}{2}-1,\frac{d}{2}-1)}(x \cdot \hat{0})}{\mathbf{P}_n^{(\frac{d}{2}-1,\frac{d}{2}-1)}(1)}.$$

By p.254 in [21], $\mathbf{P}_n^{(\alpha,\beta)}$ is a polynomial of degree exactly n. By p.259 in [21], the collection $\{\mathbf{P}_n^{(\alpha,\beta)}\}_{n\geq 0}$ are orthogonal on $[-1,1]$ with respect to the weight function $(1-r)^\alpha(1+r)^\beta$. A change of variables then shows that the ψ_n's are orthogonal in $L^2(\mathbf{S})$.

(1) is then an easy calculation, the first equality in (2) is trivial while the second equality is in [21], p.278 and 281. (3) is in [3], p.41. (4) and (5) follow from the Funk–Hecke Theorem ([17], p.195) (the calculation of γ_j being trivial). Since the subspace generated by the $\{\mathbf{P}_n^{(\frac{d}{2}-1,\frac{d}{2}-1)}(r)\}$'s are uniformly dense in $C([-1,1])$ by the Stone-Weierstrass Theorem, it easily follows that the subspace generated by the ψ_n's are uniformly dense in $L^2(\mathbf{S}/\hat{0})\cap C(\mathbf{S})$. Hence the ψ_n's are a basis for $L^2(\mathbf{S}/\hat{0})$ and (6) follows. Finally, (4) and (6) together yield (7). $\square$

Note that for all $f,g\in L^2(\mathbf{S}/\hat{0})$, we have that $fg\in L^2(\mathbf{S}/\hat{0})$ provided $fg\in L^2(\mathbf{S})$. Since ψ_n is a polynomial of degree exactly n in $x\cdot\hat{0}$, the greatest r for which $q_{ij}^r\neq 0$ must be $i+j$. From this and the nonnegativity of the q_{ij}^r's, it follows that for $\lambda>0$ the function $e^{\lambda\psi_1(x)}=\sum_{n\geq 0}\lambda^n\psi_1(x)^n/n!$ has

$$a_j(e^{\lambda\psi_1})>0,\ \text{for all } j\geq 0. \tag{2.8}$$

It follows from Lemmas 2.2, 2.5 and 2.6(3,4) that $\mathcal{P}_+\subseteq L^2(\mathbf{S}/\hat{0})$ and that for all $g\in\mathcal{P}_+$,

$$a_j(g)>0,\ \text{for all } j\geq 0. \tag{2.9}$$

Definition 11 *Define the A norm on $L^2(\mathbf{S}/\hat{0})$ by*

$$||f||_A=\sum_{j\geq 0}|a_j(f)|,$$

provided it is finite.

From the fact that $\sum_{r\geq 0}q_{ij}^r=1$, one can easily show that for all $f,g\in L^2(\mathbf{S}/\hat{0})$ with $fg\in L^2(\mathbf{S}/\hat{0})$,

$$||fg||_A\leq ||f||_A||g||_A, \tag{2.10}$$

and that equality holds if $f,g\in\mathcal{P}_+$. An easy computation also shows that $||e^{\lambda\psi_1(x)}||_A=e^\lambda<\infty$ for all $\lambda\geq 0$ and hence by Lemmas 2.2 and 2.6(4) and (2.10), $||f||_A<\infty$ for all $f\in\mathcal{P}_+$. Also, it follows from (2.9), Lemma 2.6(2,6), the fact that $\int f\,d\mathbf{x}=1$ for all $f\in\mathcal{P}_+$ and the fact that $\mathcal{P}_+\subseteq L^2(\mathbf{S}/\hat{0})$ that for $f\in\mathcal{P}_+$,

$$1+||f-1||_A=||f||_A=f(\hat{0})=||f||_\infty=1+||f-1||_\infty. \tag{2.11}$$

The last equality is obtained by observing that $\leq$ is clear while $||g||_\infty \leq ||g||_A$ for all $g \in L^2(\mathbf{S}/\hat{0})$ is also clear.

Lemma 2.7 *There exists a function o with $\lim_{h\to 0} \frac{o(h)}{h} = 0$ such that for all $h_1, \ldots, h_k \in \mathcal{P}_+$ with $k \leq B$,*

$$||\bigodot(h_1, \ldots, h_k) - 1 - \sum_{i=1}^{k}(h_i - 1)||_A \leq o(\max_i ||h_i - 1||_A), \tag{2.12}$$

provided $\max_i ||h_i - 1||_A \leq 1$.

Proof. Write

$$||\prod_{i=1}^{k} h_i - 1 - \sum_{i=1}^{k}(h_i - 1)||_A = ||\sum_{\substack{A\subseteq\{1,\ldots,k\}\\ |A|\geq 2}} \prod_{i\in A}(h_i - 1)||_A. \tag{2.13}$$

Then $\max_i ||h_i - 1|| \leq 1$ and submultiplicativity (2.10) of $||\cdot||_A$ implies this is at most

$$2^k(\max_i ||h_i - 1||_A)^2.$$

Next, since $\int(h_i - 1)\, d\mathbf{x} = 0$ for $1 \leq i \leq k$, we similarly obtain

$$\left|\int \prod_{i=1}^{k} h_i - 1\right| \leq 2^k(\max_i ||h_i - 1||_A)^2.$$

We then have

$$||\bigodot(h_1, \ldots, h_k) - \prod_{i=1}^{k} h_i||_A = \frac{1}{\int\prod_{i=1}^{k} h_i}\left|\int \prod_{i=1}^{k} h_i - 1\right| ||\prod_{i=1}^{k} h_i||_A \leq 4^k(\max_i ||h_i - 1||_A)^2,$$

since $||\prod_{i=1}^{k} h_i||_A \leq 2^k$ and $\int\prod_{i=1}^{k} h_i \geq 1$ by the positivity of the q_{ij}^r and (2.9). A use of the triangle inequality completes the proof. $\square$

We note five facts that follow easily from the above, but which will be useful later on in generalizing our results. Let $\langle\mathcal{P}_+\rangle$ be the linear subspace of $L^2(\mathbf{S}/\hat{0})$ spanned by $\mathcal{P}_+$, $\langle\mathcal{P}_+(J)\rangle$ be the linear subspace of $L^2(\mathbf{S}/\hat{0})$ spanned by $\mathcal{P}_+(J)$ and $||\mathcal{K}_{J'}||_A$ denote the operator norm of $\mathcal{K}_{J'}$ on $(\langle\mathcal{P}_+\rangle, ||\ ||_A)$.

$$\lim_{J'\to 0} ||K_{J'} - 1||_A = 0; \tag{2.14}$$

$$c_1 := \sup_{f \in \langle \mathcal{P}_+ \rangle, f \neq 1} \frac{||f-1||_\infty}{||f-1||_A} < \infty; \qquad (2.15)$$

$$c_2 := \inf_{f \in \mathcal{P}_+, f \neq 1} \frac{||f-1||_\infty}{||f-1||_A} > 0; \qquad (2.16)$$

$$\text{For all } J' \geq 0, \ \ ||\mathcal{K}_{J'}||_A \leq 1; \qquad (2.17)$$

There exist $a, b \in \mathbf{S}$ such that for all $f \in \mathcal{P}_+$,

$$f(a) = \sup_{x \in \mathbf{S}} f(x) \text{ and } f(b) = \inf_{x \in \mathbf{S}} f(x). \qquad (2.18)$$

(2.17), for example, follows immediately from Lemmas 2.5 and 2.6(7) and the fact that $|\gamma_n a_n(g)| \leq 1$ for any probability density function $g \in L^2(\mathbf{S}/\hat{0})$.

The results on Heisenberg models presented thus far are parallel to the results obtainable for any finite distance regular graph (see the next subsection). One useful result that is not true for general distance regular models depends on the following obvious geometric property of the sphere:

$$|\{z : d(x,z) \leq a, d(y,z) \leq b\}|$$

is a nonincreasing function of $d(x,y)$ for any fixed a and b where $|\ |$ denotes surface measure. [Proof: For S^1, this is obvious. For S^d, $d \geq 2$, by symmetry, we can assume that $x = (0, \ldots, 0, 1)$ and $y = (\cos\theta, 0, \ldots, 0, \sin\theta)$ (both vectors with $d+1$ coordinates). Write S^d as

$$\cup_{u \in [-1,1]^{d-1}} A_u$$

where

$$A_u := S^d \cap \{(a_1, \ldots, a_{d+1}) : (a_2, \ldots, a_d) = u\}.$$

Each A_u is a circle (or is empty) and so essentially by the 1–dimensional case, we have the desired behaviour on each A_u (using 1–dimensional Lebesgue measure) and by Fubini's Theorem, we obtain the desired result on S^d.]

Calling a function $f \in L^2(\mathbf{S}/\hat{0})$ nonincreasing if the corresponding $\overline{f}$ is nonincreasing, the latter can be seen to be equivalent to the property that $\mathbf{1}_{d(x,\hat{0}) \leq a} * \mathbf{1}_{d(x,\hat{0}) \leq b}$ is nonincreasing, and by taking linear combinations, this is equivalent to $f * g$ being nonincreasing for all nonincreasing f and g in $L^2(\mathbf{S}/\hat{0})$. Since K_J is nonincreasing for all J, it follows from the fundamental recursion that

$$f \in \mathcal{P}_+ \Rightarrow f \text{ is nonincreasing.} \tag{2.19}$$

Lemma 2.8 *For any positive nonincreasing $f \in L^2(\mathbf{S}/\hat{0})$,*

$$\left| \int_{\mathbf{S}} f \psi_n \, d\mathbf{x} \right| \leq \int_{\mathbf{S}} f \psi_1 \, d\mathbf{x}$$

for all $n \geq 1$.

Proof. It suffices to prove this for functions of the form $f(x) = \mathbf{1}_{\{x \cdot \hat{0} \geq t\}}$ with $t \in [-1, 1]$. We rely on explicit formulae for the functions $\{\psi_n\}$. Letting $\alpha = d/2 - 1$, a change of variables yields

$$\int_{\mathbf{S}} f \psi_n \, d\mathbf{x} = s_d^{-1} \int_t^1 \frac{P_n^{(\alpha,\alpha)}(r)}{P_n^{(\alpha,\alpha)}(1)} (1 - r^2)^\alpha \, dr,$$

where

$$s_d = \int_{-1}^1 (1 - r^2)^\alpha dr$$

and $\mathbf{P}_n^{(\alpha,\alpha)}$ is the Jacobi polynomial defined earlier.

Taking the indefinite integral of each side in (2.7) with $\beta = \alpha$ yields

$$\begin{aligned} \int (1 - r)^\alpha (1 + r)^\alpha P_n^{(\alpha,\alpha)}(r) \, dr &= \frac{(-1)^n}{2^n n!} \frac{d^{n-1}}{dr^{n-1}} \left[(1 - r)^{n+\alpha} (1 + r)^{n+\alpha} \right] \\ &= \frac{-1}{2n} (1 - r^2)^{\alpha+1} P_{n-1}^{(\alpha+1,\alpha+1)}(r) \, . \end{aligned}$$

Evaluating at 1 and t gives

$$\begin{aligned} \int_{\mathbf{S}} f \psi_n \, d\mathbf{x} &= s_d^{-1} \int_t^1 \frac{P_n^{(\alpha,\alpha)}(r)(1 - r^2)^\alpha}{P_n^{(\alpha,\alpha)}(1)} \, dr \\ &= s_d^{-1} \frac{P_{n-1}^{(\alpha+1,\alpha+1)}(t)(1 - t^2)^{\alpha+1}}{2n P_n^{(\alpha,\alpha)}(1)} . \end{aligned}$$

When $n = 1$, using (2.7), this is just $s_d^{-1}(1 - t^2)^{\alpha+1}/2(1 + \alpha)$. Dividing, we get

$$\frac{\int_{\mathbf{S}} f \psi_n \, d\mathbf{x}}{\int_{\mathbf{S}} f \psi_1 \, d\mathbf{x}} = \frac{P_{n-1}^{(\alpha+1,\alpha+1)}(t)(1 + \alpha)}{n P_n^{(\alpha,\alpha)}(1)} = \frac{P_{n-1}^{(\alpha+1,\alpha+1)}(t)}{P_{n-1}^{(\alpha+1,\alpha+1)}(1)} \cdot \frac{P_{n-1}^{(\alpha+1,\alpha+1)}(1)}{n P_n^{(\alpha,\alpha)}(1)} \cdot (1 + \alpha).$$

The first term in the product is bounded in absolute value by 1. By [3], p.7,

$$P_n^{(\alpha,\alpha)}(1) = \binom{\alpha+n}{n},$$

and so we see that the second term is $1/(\alpha+1)$, completing the proof of the lemma. $\square$

Remark: The case $d = 1$ can also be handled by a rearrangement lemma.

Definition 12 *Define a linear functional L on $L^2(\mathbf{S}/\hat{0})$ by $L(g) := \int_{\mathbf{S}} g(x)\psi_1(x)d\mathbf{x}(x)$ $(= \gamma_1 a_1(g))$ and set $\mathbf{Op}_J = L(K_J)$. (Recall that ψ_1, γ_1 and a_1 are defined in Lemma 2.6.)*

It follows from Lemmas 2.5, 2.6(7) and 2.8, (2.19) and an easy computation that

$$||\mathcal{K}_J f - 1||_A \leq \mathbf{Op}_J ||f - 1||_A \text{ for all } f \in \mathcal{P}_+(J). \tag{2.20}$$

In the following inequalities, we denote $\rho := \mathbf{Op}_J$. For $f \in \mathcal{P}_+(J)$, it also follows easily that

$$L(\mathcal{K}_J f - 1) \geq \rho L(f - 1) \tag{2.21}$$

and that there is a constant c_3 such that for all $f \in \langle \mathcal{P}_+(J) \rangle$,

$$|L(f)| \leq c_3 ||f||_A. \tag{2.22}$$

(We can of course take c_3 to be 1, but we leave the condition written in this more general form for use as a hypothesis in Theorem 3.2.)

Putting together the results of Lemmas 2.6 and 2.8, as well as (2.8), (2.9) and (2.19), gives the following corollary.

Corollary 2.9 *For all $J \geq 0$, there is a constant $c_4 > 0$ such that for all $f \in \mathcal{P}_+(J)$,*

$$L(f) \geq c_4 ||f - 1||_A. \tag{2.23}$$

Proof. Fix $f \in \mathcal{P}_+(J)$. If $f = K_{J'}$ for some $J' \in (0, J]$, we argue as follows. As $||e^{\lambda\psi_1(x)}||_A = e^{\lambda}$ (which we mentioned earlier) and $K_{J'}(x) = e^{J'\psi_1(x)} / \int e^{J'\psi_1(x)} d\mathbf{x}(x)$, we have

$$\begin{aligned} ||K_{J'} - 1||_A &= ||K_{J'}||_A - 1 \\ &= \frac{e^{J'}}{\int e^{J'\psi_1(x)} d\mathbf{x}(x)} - 1 \\ &\leq e^{2J'} - 1. \end{aligned}$$

Next,

$$\begin{aligned} L(K_{J'}) &= \frac{1}{\int e^{J'\psi_1(x)} d\mathbf{x}(x)} \int e^{J'\psi_1(x)}\psi_1(x) d\mathbf{x}(x) \\ &= \frac{1}{\int e^{J'\psi_1(x)} d\mathbf{x}(x)} \sum_{k=0}^{\infty} \frac{(J')^k}{k!} \int \psi_1^{k+1}(x) d\mathbf{x}(x). \end{aligned}$$

By Lemma 2.6(3), all terms in the sum are nonnegative and by Lemma 2.6(4), the $k = 1$ term is $J'\gamma_1$. Hence $L(K_{J'}) \geq J'\gamma_1/e^{J'}$. Since

$$\inf_{J' \in (0,J]} \frac{J'\gamma_1}{e^{J'}(e^{2J'} - 1)} > 0,$$

we can find a c_4 in this case.

Otherwise, by the fundamental recursion, we may represent f as $\bigodot(\mathcal{K}_{J_1}h_1, \ldots, \mathcal{K}_{J_k}h_k)$ with each h_i either in $\mathcal{P}_+(J)$ or equal to $\delta_{\hat{0}}$ and each $J_i \in (0, J]$. Define $g_i = \mathcal{K}_{J_i}h_i - 1$. Let $m := \inf_{0<J'\leq J} a_1(K_{J'})/\sum_{n>0} a_n(K_{J'})$ which is strictly positive by the above. It follows that if $h_i \in \mathcal{P}_+(J)$ (the case $h_i = \delta_{\hat{0}}$ is already done),

$$\frac{L(g_i)}{||g_i||_A} = \frac{a_1(K_{J_i})a_1(h_i)\gamma_1^2}{\sum_{n>0} a_n(K_{J_i})a_n(h_i)\gamma_n} \geq m\gamma_1$$

by Lemma 2.6(7) and since $a_1(h_i)\gamma_1 \geq a_n(h_i)\gamma_n$ for all $n \geq 1$ by Lemma 2.8 and (2.19). Let $h = \prod_{i=1}^k \mathcal{K}_{J_i}h_i$. Then $L(h) = L(1 + \sum_{i=1}^k g_i + Q)$, where Q is a sum of monomials in $\{g_i\}$. Using $q_{ij}^r \geq 0$ and (2.9), we have that $L(Q) \geq 0$, and hence

$$L(h) \geq \sum_{i=1}^k L(g_i) \geq m\gamma_1 \sum_{i=1}^k ||g_i||_A. \tag{2.24}$$

On the other hand, for any B and M, there is $C = C(M, B)$ such that if $x_1, \ldots, x_k \in (0, M)$ with $k \leq B$, then

$$-1 + \prod_{i=1}^k (1 + x_i) \leq C \sum_{i=1}^k x_i.$$

Next, the positivity of the q_{ij}^r implies $\int_{\mathbf{S}} h(x)\, d\mathbf{x}(x) = a_0(h) \geq 1$. It follows that

$$||h - 1||_A = -1 + ||h||_A = -1 + \prod_{i=1}^k ||g_i + 1||_A \leq C \sum_{i=1}^k ||g_i||_A$$

for some constant C since $||g_i + 1||_A = ||g_i||_A + 1$ and $||g_i + 1||_A$ clearly has a universal upper bound. [To see the latter statement, one notes that

$$\sup_{0<J'\leq J} ||K_{J'}||_A < \infty,$$

$$||K_{J'} * f||_A \leq ||K_{J'}||_A$$

for any probability density function $f \in L^2(\mathbf{S}/\hat{0})$ (by Lemma 2.6(7)), (2.10) and the fact that we never have more than B terms in our pointwise products imply that

$$\sup_{f \in \mathcal{P}_+(J)} ||f||_A \leq \left(\sup_{0 < J' \leq J} ||K_{J'}||_A \right)^B < \infty.]$$

Putting this together with (2.24) gives

$$\frac{L(h)}{||h-1||_A} \geq \frac{m\gamma_1}{C}.$$

Finally, letting $f = h / \left(\int_{\mathbf{S}} h(x)\, d\mathbf{x}(x) \right)$, we obtain

$$\begin{aligned} ||h-1||_A &\geq \sum_{n \geq 1} a_n(h) \\ &= \sum_{n \geq 1} \left[\int_{\mathbf{S}} h(x) d\mathbf{x}(x) \right] a_n(f) \\ &= \left[\int_{\mathbf{S}} h(x) d\mathbf{x}(x) \right] ||f-1||_A. \end{aligned}$$

Hence

$$\frac{L(f)}{||f-1||_A} \geq \frac{L(h)}{||h-1||_A} \geq \frac{m\gamma_1}{C}$$

and we're done. □

2.4 Distance regular graphs

For the remainder of this section, we suppose that $\mathbf{S}$ is the vertex set of a finite, connected, distance regular graph, that $d(x,y)$ is the graph distance, and that the energy $H(x,y)$ depends only on $d(x,y)$. The Potts models fit into this framework, with the respective graphs being the complete graph K_q on q vertices. All the results we need follow in fact from an even weaker assumption, namely that $\mathbf{S}$ is an *association scheme*. For the definition of association schemes and the proofs of the relevant results, see [4] or [23]. By developing the analogue of Lemma 2.6 for distance regular graphs, we will illustrate the extent to which our results are independent of the special properties of the Heisenberg model.

We have a distinguished element $\hat{0} \in \mathbf{S}$ and the measure $d\mathbf{x}$ will of course be normalized counting measure $|\mathbf{S}|^{-1} \sum_{x \in \mathbf{S}} \delta_x$. The spaces $L^2(\mathbf{S})$ and $L^2(\mathbf{S}/\hat{0})$ are then simply finite dimensional vector spaces with respective dimensions $|\mathbf{S}|$ and $1 + D$, where D is the diameter of the graph $\mathbf{S}$.

Denote by $M(\mathbf{S})$ the space of matrices with rows and columns indexed by $\mathbf{S}$, thought of as linear maps from $L^2(\mathbf{S})$ to $L^2(\mathbf{S})$. Associated with each function $f \in L^2(\mathbf{S}/\hat{0})$ is the matrix $M_f \in M(\mathbf{S})$ whose (x, y) entry is $\overline{f}(d(x, y))$, whence the matrix M_f corresponds to the linear operator $h \mapsto h * f$ given in Section 2.2. The following analogue of Lemma 2.6 is derived from Section 2.4 of [23]; a published reference is Section 2.3 of [4].

Lemma 2.10 *There exists a basis of real-valued functions $\psi_0, \ldots, \psi_D$ of $L^2(\mathbf{S}/\hat{0})$ orthogonal under the inner product $\langle f, g\rangle = |\mathbf{S}|^{-1} \sum_x f(x)\overline{g(x)}$ with the following properties.*
(1) $\psi_0(x) \equiv 1$.
(2) $\psi_j(\hat{0}) = 1 = \sup_x |\psi_j(x)|$ for all j.
(3) $\psi_i \psi_j = \sum_{r=0}^{D} q_{ij}^r \psi_r$ for some nonnegative coefficients q_{ij}^r with $\sum_r q_{ij}^r = 1$.
(4) $\psi_i * \psi_j = \gamma_j \delta_{ij} \psi_j$, where $\gamma_j := \psi_j * \psi_j(\hat{0}) = |\mathbf{S}|^{-1} \sum_x \psi_j(x)^2$.
(5) The functions ψ_j are eigenfunctions of any convolution operator, that is, $M_f \psi_j = c\psi_j$ for any $f \in L^2(\mathbf{S}/\hat{0})$.
(6) For $f \in L^2(\mathbf{S}/\hat{0})$, we have $f = \sum_{j=0}^{D} a_j(f)\psi_j$, where $a_j := \gamma_j^{-1} |\mathbf{S}|^{-1} \sum_x f(x)\psi_j(x)$.
(7) For $f, g \in L^2(\mathbf{S}/\hat{0})$, we have $a_j(f * g) = \gamma_j a_j(f) a_j(g)$.
(8) For $f \in L^2(\mathbf{S}/\hat{0})$ which is positive and nonincreasing, $|\langle f, \psi_i\rangle| \leq \langle f, \psi_1\rangle$ for each $i \geq 1$.

If we place the norm $\sum_{j=0}^{D} |a_j(f)|$ on $\langle \mathcal{P}_+(J)\rangle$, essentially all of the hypotheses in Theorems 3.1 and 3.2 (to come later) are immediate noting that all norms are equivalent on finite dimensional spaces. If the analogue of (2.19) holds, then letting $L(g) := |S|^{-1} \sum_{x \in \mathbf{S}} g(x)\psi_1(x)$ and both $\mathbf{Op}_J$ and ρ to be $L(K_J)$, then one can easily show that *all* of the hypotheses in Theorems 3.1 and 3.2 hold. As far as (2.19), it trivially holds for the complete graph where the diameter D is equal to 1 and in any case, the reader is left with only one condition to check.

3 Two Technical Theorems

We now state two general results from which Theorems 1.6 and 1.13 will follow.

Theorem 3.1 *Let Γ be any tree (with bounded degree). For the d–dimensional Heisenberg model with $d \geq 1$, if $J > 0$ and*

$$br(\Gamma) \cdot \mathbf{Op}_J < 1,$$

then there is no robust phase transition for the parameter J, where $\mathbf{Op}_J$ is given in Definition 12 ($\mathbf{Op}_J$ implicitly depends on d). More generally, if $J > 0$ and if $(\mathbf{S}, G, H)$ is any statistical ensemble with a norm $||\cdot||$ on $\langle \mathcal{P}_+(J)\rangle$ satisfying (2.12), (2.14), (2.15) and (2.17) and there exists a number $\mathbf{Op}_J \in (0,1)$ satisfying (2.20) and $br(\Gamma)\cdot\mathbf{Op}_J < 1$, then there is no robust phase transition for the parameter J.

Theorem 3.2 *Let Γ be any tree (with bounded degree). For the d–dimensional Heisenberg model with $d \geq 1$, if $J > 0$ and*

$$br(\Gamma) \cdot \mathbf{Op}_J > 1,$$

then there is a robust phase transition for the parameter J, where $\mathbf{Op}_J$ is as above. More generally, if $J > 0$ and if $(\mathbf{S}, G, H)$ is any statistical ensemble with a norm $||\cdot||$ on $\langle \mathcal{P}_+(J)\rangle$ satisfying (2.12), (2.15), (2.16), (2.17) and (2.18), and if L is a linear functional on $\langle \mathcal{P}_+(J)\rangle$ which vanishes on the constants and satisfies (2.21), (2.22) and (2.23) for a constant $\rho > 0$, then $br(\Gamma)\cdot\rho > 1$ implies a robust phase transition for the parameter J.

To prove these results, we begin with a purely geometric lemma on the existence of cutsets of uniformly small content below the branching number.

Lemma 3.3 *Assume that $br(\Gamma) < d$. Then for all $\epsilon > 0$, there exists a cutset C such that*

$$\sum_{x\in C}(\frac{1}{d})^{|x|} \leq \epsilon$$

and for all $v \in C^i \cup C$,

$$\sum_{x\in C\cap\Gamma(v)}(\frac{1}{d})^{|x|-|v|} \leq 1. \tag{3.1}$$

Proof. Since $br(\Gamma) < d$, for any given $\epsilon > 0$, there exists a cutset C such that

$$\sum_{x\in C}(\frac{1}{d})^{|x|} \leq \epsilon.$$

We can assume that C is a minimal cutset with this property with respect to the partial order $C_1 \preceq C_2$ if for all $v \in C_1$, there exists $w \in C_2$ such that $v \leq w$. We claim that this cutset satisfies (3.1). If this property failed for some v, we let C' be the modified cutset obtained by replacing $C \cap \Gamma(v)$ by v (and leaving $C \cap \Gamma_v^c$ unchanged). As (3.1) clearly holds for $w \in C$, we must have that $v \notin C$ in which case $C' \neq C$. We then have

$$\begin{aligned}
\sum_{x \in C'} (\frac{1}{d})^{|x|} &= \sum_{x \in C \cap \Gamma(v)^c} (\frac{1}{d})^{|x|} + (\frac{1}{d})^{|v|} \\
&< \sum_{x \in C \cap \Gamma(v)^c} (\frac{1}{d})^{|x|} + (\frac{1}{d})^{|v|} \sum_{x \in C \cap \Gamma(v)} (\frac{1}{d})^{|v-x|} \\
&= \sum_{x \in C \cap \Gamma(v)^c} (\frac{1}{d})^{|x|} + \sum_{x \in C \cap \Gamma(v)} (\frac{1}{d})^{|x|} \\
&= \sum_{x \in C} (\frac{1}{d})^{|x|} \\
&\leq \epsilon,
\end{aligned}$$

contradicting the minimality of C since clearly $C' \preceq C$. $\square$

We now proceed with the proofs of Theorems 3.1 and 3.2.

Proof of Theorem 3.1. Since in Section 2.3 the Heisenberg models have been shown to satisfy all of the more general hypotheses of this theorem, we need only prove the last statement of the theorem where we have a given $J > 0$, a given $\| \ \|$ on $\langle \mathcal{P}_+(J) \rangle$ and a given $\mathbf{Op}_J$ satisfying the required conditions. By (2.12), for any $\epsilon > 0$, there is an $\epsilon_0 > 0$ such that for all $k \leq B$ and all $h_1, \ldots, h_k \in \mathcal{P}_+(J)$ with $\|h_i - 1\| \leq \epsilon_0$ for all i, we have that

$$\|\bigodot_k (h_1, \ldots, h_k) - 1\| \leq (1+\epsilon) \sum_{i=1}^{k} \|h_i - 1\| . \tag{3.2}$$

Choose $\epsilon > 0$ so that $(1+\epsilon)^{-1} > br(\Gamma) \cdot \mathbf{Op}_J$ and choose ϵ_0 as above. By (2.14), we can choose $J' > 0$ small enough so that $\|K_{J'} - 1\| \leq \epsilon_0 \mathbf{Op}_J$. Use Lemma 3.3 to choose a sequence of cutsets $\{C_n\}$ for which

$$\lim_{n \to \infty} \sum_{x \in C_n} [(1+\epsilon)\mathbf{Op}_J]^{|x|} = 0$$

and for all n and all $v \in C_n^i \cup C_n$,

$$\sum_{x \in C_n \cap \Gamma(v)} [(1+\epsilon)\mathbf{Op}_J]^{|x|-|v|} \leq 1. \tag{3.3}$$

We now show by induction that for all n and all $v \in C_n^i$,

$$\|f_{C_n,v}^{J',J,+} - 1\| \le \epsilon_0 \sum_{x \in C_n \cap \Gamma(v)} [(1+\epsilon)\mathbf{Op}_J]^{|x|-|v|} . \tag{3.4}$$

Indeed, from Lemma 2.2, letting $w_1, \ldots, w_k$ be the children of v,

$$\|f_{C_n,v}^{J',J,+} - 1\| = \|\bigodot(\mathcal{K}_{J_1''} f_{C_n,w_1}^{J',J,+}, \ldots, \mathcal{K}_{J_k''} f_{C_n,w_k}^{J',J,+}) - 1\|$$

where J_i'' is J if $w_i \in C_n^i$ and J' otherwise. When $w_i \in C_n$, the choice of J' guarantees that $\|\mathcal{K}_{J_i''} f_{C_n,w_i}^{J',J,+} - 1\| \le \epsilon_0 \mathbf{Op}_J \le \epsilon_0$, while when $w_i \notin C_n$, the induction hypothesis together with (3.3) guarantees that $\|f_{C_n,w_i}^{J',J,+} - 1\| \le \epsilon_0$ which implies that $\|\mathcal{K}_{J_i''} f_{C_n,w_i}^{J',J,+} - 1\| \le \epsilon_0$ by (2.17). Hence, from (3.2),

$$\|f_{C_n,v}^{J',J,+} - 1\| \le (1+\epsilon) \sum_{w_i \in C_n} \|\mathcal{K}_{J'} f_{C_n,w_i}^{J',J,+} - 1\| + (1+\epsilon) \sum_{w_i \notin C_n} \|\mathcal{K}_J f_{C_n,w_i}^{J',J,+} - 1\|.$$

The summands in the first sum are at most $\epsilon_0 \mathbf{Op}_J$ while those in the second sum are by (2.20) at most $\mathbf{Op}_J \|f_{C_n,w_i}^{J',J,+} - 1\|$. Therefore using the induction hypothesis on the second term, we obtain

$$\begin{aligned} \|f_{C_n,v}^{J',J,+} - 1\| &\le \sum_{i=1}^{k} (1+\epsilon)\epsilon_0 \mathbf{Op}_J \sum_{x \in C_n \cap \Gamma(w_i)} [(1+\epsilon)\mathbf{Op}_J]^{|x|-|w_i|} \\ &= \epsilon_0 \sum_{x \in C_n \cap \Gamma(v)} [(1+\epsilon)\mathbf{Op}_J]^{|x|-|v|} , \end{aligned}$$

completing the induction. Finally, the theorem follows by taking $v = o$, letting $n \to \infty$, and using (2.15). $\square$

For the proof of Theorem 3.2, it is easiest to isolate the following two lemmas.

Lemma 3.4 *Under the more general hypotheses of Theorem 3.2 (with a given $J > 0$, a given $\| \ \|$ on $\langle \mathcal{P}_+(J) \rangle$, a given L and a given ρ satisfying the required conditions), for all $\alpha > 0$, there exists $\beta > 0$ so that if $h_1, \ldots, h_k \in \mathcal{P}_+(J)$ with $k \le B$ and $\|h_i - 1\| < \beta$ for each i, then*

$$L\left[(\bigodot_k(\mathcal{K}_J h_1, \ldots, \mathcal{K}_J h_k)) - 1\right] \ge \frac{1}{1+\alpha} \sum_{i=1}^{k} L(\mathcal{K}_J h_i - 1)$$

Proof. In (2.12), choose $\beta < 1$ so that

$$o(h) \leq h\left(1 - \frac{1}{(1+\alpha)}\right)\frac{c_4}{c_3}$$

for all $h \in (0, \beta)$, with c_3 and c_4 as in (2.22) and (2.23). If $h_1, \ldots, h_k \in \mathcal{P}_+(J)$ are such that $\|h_i - 1\| < \beta$, then $\|\mathcal{K}_J h_i - 1\| < \beta$ by (2.17). We can now write

$$\bigodot_k(\mathcal{K}_J h_1, \ldots, \mathcal{K}_J h_k) - 1 - \frac{1}{(1+\alpha)}\sum_{i=1}^{k}(\mathcal{K}_J h_i - 1) \tag{3.5}$$

as

$$\left(1 - \frac{1}{(1+\alpha)}\right)\sum_{i=1}^{k}(\mathcal{K}_J h_i - 1) + U \tag{3.6}$$

where by assumption,

$$\begin{aligned} \|U\| &\leq o(\max_i \|\mathcal{K}_J h_i - 1\|) \\ &\leq \left(1 - \frac{1}{(1+\alpha)}\right)\frac{c_4}{c_3}\max_i \|\mathcal{K}_J h_i - 1\| \\ &\leq \left(1 - \frac{1}{(1+\alpha)}\right)\frac{c_4}{c_3}\sum_{i=1}^{k}\|\mathcal{K}_J h_i - 1\|. \end{aligned} \tag{3.7}$$

Letting a be the quantity (3.5), we see that

$$\begin{aligned} L(a) &= L\left[\left(1 - \frac{1}{(1+\alpha)}\right)\sum_{i=1}^{k}(\mathcal{K}_J h_i - 1)\right] + L(U) \\ &\geq \left(1 - \frac{1}{(1+\alpha)}\right)c_4\sum_{i=1}^{k}\|\mathcal{K}_J h_i - 1\| - c_3\|U\| \\ &\geq 0 \end{aligned}$$

by (2.22), (2.23) and (3.7), which is the conclusion of the lemma. $\square$

The next lemma tells us that in "one step", we can't move from being "far away" from uniform to being "very close" to uniform.

Lemma 3.5 *Under the more general hypotheses of Theorem 3.2 (with a given $J > 0$, a given $\|\ \|$ on $\langle\mathcal{P}_+(J)\rangle$, a given L and a given ρ satisfying the required conditions), for all $\beta > 0$ and $J' \in (0, J]$, there exists a $\gamma < \beta$ such that if $\|\bigodot_k(\mathcal{K}_{J_1''}h_1, \ldots, \mathcal{K}_{J_k''}h_k) - 1\| < \gamma$ with $h_1, \ldots, h_k \in \mathcal{P}_+(J) \cup \{\delta_{\hat{0}}\}$ and $k \leq B$ and with J_i'' being J if $h_i \in \mathcal{P}_+(J)$ and J' if $h_i = \delta_{\hat{0}}$, then each h_i is not $\delta_{\hat{0}}$ and $\sum_{i=1}^{k}\|h_i - 1\| < \beta$.*

Proof. Choose $\gamma \in (0, \min\{\beta, 1/c_1\})$ so that

$$\frac{2c_1c_3B\gamma}{\rho c_2c_4(1-c_1\gamma)} < \beta$$

and

$$\min\{||K_J - 1||, ||K_{J'} - 1||\} > \frac{2c_1\gamma}{(1-c_1\gamma)c_2}$$

where c_1, c_2, ρ, c_3 and c_4 come from (2.15), (2.16), (2.21), (2.22) and (2.23) respectively. We first show that if $h_1, \ldots, h_k \in \mathcal{P}_+(J)$, with $k \le B$, then $\|\bigodot_k(h_1, \ldots, h_k) - 1\| < \gamma < 1/c_1$ implies that for all i

$$\|h_i - 1\| < \frac{2c_1\gamma}{(1-c_1\gamma)c_2}.$$

[Proof:

$$||h_i - 1|| \le c_2^{-1}||h_i - 1||_\infty \le c_2^{-1}\left(\frac{\max h_i}{\min h_i} - 1\right)$$
$$\le c_2^{-1}\left(\frac{\max \prod_i h_i}{\min \prod_i h_i} - 1\right) = c_2^{-1}\left(\frac{\max \bigodot_k(h_1, \ldots, h_k)}{\min \bigodot_k(h_1, \ldots, h_k)} - 1\right)$$

where the second inequality is straightforward and the third inequality comes from (2.18). Next, $\|\bigodot_k(h_1, \ldots, h_k) - 1\| < \gamma < 1/c_1$ implies $||\bigodot_k(h_1, \ldots, h_k) - 1||_\infty \le c_1\gamma$ which implies the last expression is at most

$$c_2^{-1}\left(\frac{1+c_1\gamma}{1-c_1\gamma} - 1\right) = c_2^{-1}\frac{2c_1\gamma}{1-c_1\gamma}.]$$

It follows that if $\|\bigodot_k(\mathcal{K}_{J_1''}h_1, \ldots, \mathcal{K}_{J_k''}h_k) - 1\| < \gamma$, then

$$\|\mathcal{K}_{J_i''}h_i - 1\| < \frac{2c_1\gamma}{(1-c_1\gamma)c_2}$$

for each i which implies that $h_i \in \mathcal{P}_+(J)$ (as opposed to being $\delta_{\hat{0}}$). Hence J_i'' is J for all i.

Now from (2.21)–(2.23) we have

$$||\mathcal{K}_J h_i - 1|| \ge \frac{\rho c_4||h_i - 1||}{c_3}$$

and we obtain the conclusion of the lemma. □

Proof of Theorem 3.2. Since in Section 2.3 the Heisenberg models have been shown to satisfy all of the more general hypotheses of this theorem, we need only prove the last statement of the theorem, where we have a given $J > 0$, a given $\|\ \|$ on $\langle \mathcal{P}_+(J) \rangle$, a given L and a given ρ satisfying the required conditions. Choose an $\alpha > 0$ so that

$br(\Gamma) \cdot \rho > 1 + \alpha$. Choosing β from Lemma 3.4, we have, under our assumptions, that for all $h_1, \ldots, h_k \in \mathcal{P}_+(J)$ with $k \leq B$ and $\|h_i - 1\| < \beta$ for each i,

$$L\left[(\bigodot\nolimits_k(\mathcal{K}_J h_1, \ldots, \mathcal{K}_J h_k)) - 1\right] \geq \frac{1}{1+\alpha} \sum_{i=1}^{k} L(\mathcal{K}_J h_i - 1) \geq \frac{\rho}{1+\alpha} \sum_{i=1}^{k} L(h_i - 1). \quad (3.8)$$

Now, if there is no robust phase transition, then by (2.16) there must exist $J' \in (0, J]$ and a sequence of cutsets $\{C_n\}$ going to infinity such that $\lim_{n\to\infty} \|f^{J',J,+}_{C_n,o} - 1\| = 0$. Using Lemma 3.5, choose $\gamma < \beta$ corresponding to β and J'. Next, by our choice of α, we have

$$I := \inf_C \sum_{x \in C} \left(\frac{\rho}{1+\alpha}\right)^{|x|} > 0$$

where the infimum is over all cutsets. We now choose n so that

$$\|f^{J',J,+}_{C_n,o} - 1\| < \min\{\gamma, \frac{c_4 \gamma I}{c_3}\}.$$

where c_3 and c_4 come from (2.22) and (2.23) respectively. We then define Γ' to be the component of the set

$$\{v \in C_n^i : \|f^{J',J,+}_{C_n,v} - 1\| < \gamma\}$$

that contains o and let C be the exterior boundary of Γ' (that is, the set of $x \notin \Gamma'$ neighboring some $y \in \Gamma'$). By the choice of γ, $C \subseteq C_n^i$ and for each $v \in C^i \cup C$, the density $f^{J',J,+}_{C_n,v}$ is in

$$\mathcal{P}_+(J) \cap \{f : \|f - 1\| < \beta\}.$$

Using (3.8) and induction, we see that

$$L(f^{J',J,+}_{C_n,o} - 1) \geq \sum_{x \in C} \left(\frac{\rho}{1+\alpha}\right)^{|x|} L(f^{J',J,+}_{C_n,x} - 1).$$

By definition of Γ', C and I and the fact that $L(f-1) \geq c_4\|f-1\|$ on $\mathcal{P}_+(J)$, we see that

$$L(f^{J',J,+}_{C_n,o} - 1) \geq c_4 \gamma I.$$

Hence

$$\|f^{J',J,+}_{C_n,o} - 1\| \geq \frac{c_4}{c_3} \gamma I.$$

This contradicts the choice of n, proving that there is indeed a robust phase transition. $\square$

4 Analysis of specific models

4.1 Heisenberg models

For the Heisenberg models, recall that $\mathbf{S} = S^d$, $d \geq 1$, and $H(x,y) = -x \cdot y$. The operator $\mathcal{K}_J$ is convolution with the function $K_J(x) = ce^{Jx\cdot\hat{0}}$, where c is a normalizing constant.

Proof of Theorem 1.6. A change of variables shows that $L(K_J) = \rho^d(J)$ and so the result follows from Theorems 3.1 and 3.2. $\square$

For the rotor model, we now prove the equivalence of SB and SB+.

Proof of Proposition 1.4. We have already seen the representation

$$f = \sum_{n \geq 0} a_n(f)\psi_n,$$

for functions $f \in L^2(\mathbf{S}/\hat{0})$. In the case of the rotor model, where $\mathbf{S} = S^1$ and we take $\hat{0}$ to be $(1,0)$, the space $L^2(\mathbf{S}/\hat{0})$ is the space of even functions of $\theta \in [-\pi,\pi]$ and $\psi_n = \cos(n\theta)$. We now turn to the full Fourier decomposition $f = \sum_{n\in\mathbb{Z}} b_n(f)e^{in\theta}$, where $b_n(f) = \int_0^{2\pi} f(\theta)e^{-in\theta}\,d\theta/(2\pi)$.

Let C be any cutset and δ be a set of boundary conditions on C. Let $\mathcal{J}$ be any set of interaction strengths. It suffices to show that

$$||f_{C,w}^{\mathcal{J},\delta} - 1||_\infty \leq ||f_{C,w}^{\mathcal{J},+} - 1||_\infty$$

for all $w \in C^i$. For $v \in C$ and $n \in \mathbb{Z}$, let $x_{v,n} = b_n(K_{\mathcal{J}(x),\delta(v)})$ where e is the edge from v to its parent.

Claim: For all $y \in C^i$, the Fourier coefficients $\int_0^{2\pi} e^{in\theta}\,d\mu_{C,y}^{\mathcal{J},\delta}(\theta)$, which we denote by $\{u_{y,n} : n \in \mathbb{Z}\}$, are sums of monomials in $\{x_{v,n}\}_{v\in C, n\in\mathbb{Z}}$ with nonnegative coefficients. *Proof:* Let $w \in C^i$ have children $w_1,\ldots,w_r \in C^i$ and $w_{r+1},\ldots,w_k \in C$. Then the Fourier coefficients $\{u_{w,n} : n \in \mathbb{Z}\}$ are the convolution of the $k-r$ series $\{x_{v,n} : n \in \mathbb{Z}\}$ as v ranges over $w_{r+1},\ldots,w_k$, also convolved with the series $\{b_n(K_{\mathcal{J}(\overline{wv})})u_{v,n} : n \in \mathbb{Z}\}$ as v ranges over $w_1,\ldots,w_r$. Since $b_n(K_J) \geq 0$, this establishes the claim via induction and the fundamental recursion.

Now write $x_{v,n}^+$ for the Fourier coefficients $b_n(K_{\mathcal{J}(e)})$ where e is as before. Since $K_{J,e^{i\alpha}}(x) = K_J(e^{-i\alpha}x)$, it follows that

$$|x_{v,n}| = |x_{v,n}^+|.$$

But $x^+_{v,n}$ is real because K_J is even, and has been shown to be nonnegative. Thus

$$|x_{v,n}| = x^+_{v,n},$$

and it follows from the claim that each $u_{w,n}$ has modulus bounded above by the corresponding $u^+_{w,n}$ when plus boundary conditions are taken. Hence

$$||f^{\mathcal{J},\delta}_{C,w} - 1||_\infty \le ||f^{\mathcal{J},\delta}_{C,w} - 1||_A \le \sum_{n\neq 0} |u_{w,n}| \le \sum_{n\neq 0} u^+_{w,n} = ||f^{\mathcal{J},+}_{C,w} - 1||_A = ||f^{\mathcal{J},+}_{C,w} - 1||_\infty,$$

proving the lemma. □

Remark: Although we have used special properties of the Fourier decomposition on $L^2(S^1)$, there exist similar decompositions for S^d. We believe that a parallel argument can probably be constructed, bounding the modulus of the sum of the coefficients of spherical harmonics of a given order by the coefficients one obtains for the analogous monomials in the values $a_n(K_{\mathcal{J}(x)})$, whose coefficients are necessarily nonnegative by the nonnegativity of the connection coefficients q^r_{ij}. Thus we are led to state:

Problem 4.1 *Prove a version of Proposition 1.4 for general Heisenberg models on trees.*

4.2 The Potts model

Proof of Theorem 1.13. We will obtain this result from Theorems 3.1 and 3.2. For (i), letting $||\ ||$ be the L_∞ norm on $\langle \mathcal{P}_+(J)\rangle$ and $\mathbf{Op}_J = \alpha_J$, all of the hypotheses in Theorem 3.1 except (2.20) are clear. The function K_J is given by

$$K_J(x) = c\exp(J(2\delta_{x,0} - 1))$$

where $c = (e^J + (q-1)e^{-J})^{-1}$. The operator $\mathcal{K}_J$ is linear and

$$\mathcal{K}_J \delta_j = ce^J\delta_j + \sum_{i\neq j} ce^{-J}\delta_i\,.$$

Hence in the basis $\delta_0, \ldots, \delta_{q-1}$, the matrix representation of $\mathcal{K}_J$ is $c(e^J - e^{-J})I + ce^{-J}M$ where M is the matrix of all ones. On the orthogonal complement of the constant functions, $\mathcal{K}_J$ is $c(e^J - e^{-J})I$, and (2.20) follows, proving (i) by an application of Theorem 3.1.

For (ii), let $||\ ||$ be the same as above, $\rho = \alpha_J$ and $L(h) = h(0) - h(1)$. It is then immediate to check that all of the hypotheses in Theorem 3.2 hold and we may conclude (ii) by an application of Theorem 3.2. □

5 Proof of Theorem 1.10.

By Proposition 1.3 and the fact that any subtree of a tree with branching number 1 also has branching number 1, it suffices to show:

> For for any Γ with $br(\Gamma) = 1$, and any bounded $\mathcal{J}$, there is a sequence of cutsets $\{C_n\}$ such that for any sequence $\{\delta_n\}$ of boundary conditions on $\{C_n\}$,
>
> $$\lim_{n\to\infty} \|f^{\mathcal{J},\delta_n}_{C_n,o} - 1\|_\infty = 0.$$

It is convenient to work with a different measure of size, the *Max/Min* measure, defined as follows. (This arose already in the proof of Lemma 3.5.) For any continuous strictly positive function f on $\mathbf{S}$, let

$$\|f\|_M := \frac{\max_{x\in\mathbf{S}} f(x)}{\min_{x\in\mathbf{S}} f(x)}.$$

It is immediate to see:

Lemma 5.1 *For any sequence $\{h_n\}$ of continuous probability densities, $\|h_n - 1\|_\infty \to 0$ if and only if $\log \|h_n\|_M \to 0$.*

Next, we examine the effect of $\mathcal{K}_J$ on $\|f\|_M$.

Lemma 5.2 *For any statistical ensemble $(\mathbf{S}, G, H)$, any $J_{\max}$ and any $T > 0$ there is an $\epsilon > 0$ such that for any continuous strictly positive function f with $\|f\|_M \le T$, and any $J \le J_{\max}$,*

$$\log \|\mathcal{K}_J f\|_M \le (1-\epsilon) \log \|f\|_M .$$

Proof. Fix H, J and f and assume without loss of generality that $\int f\, d\mathbf{x} = 1$ since the *Max/Min* measure is unaffected by multiplicative constants. Let $[a, b]$ be the smallest closed interval containing the range of f and $[c, d]$ contain the range of K_J with $a, c > 0$. Since f is a probability density, $a < 1 < b$ (we rule out the trivial case $f \equiv 1$). Since $K_J = c + (1-c)g$ for some probability density g, it follows that for any $x \in \mathbf{S}$,

$$c + (1-c)a \le \mathcal{K}_J f(x) \le c + (1-c)b.$$

As J varies over $[0, J_{\max}]$, $\min_x K_J(x)$ is bounded below by some $c_0 > 0$, so for all such J,

$$c_0 + (1-c_0)a \le \mathcal{K}_J f(x) \le c_0 + (1-c_0)b$$

and so

$$\|\mathcal{K}_J f\|_M \leq \frac{c_0 + (1-c_0)b}{c_0 + (1-c_0)a}.$$

Setting $R = \|f\|_M - 1$, we have $b = (1+R)a$ and so

$$\|\mathcal{K}_J f\|_M \leq \frac{c_0 + (1-c_0)(1+R)a}{c_0 + (1-c_0)a} = 1 + R\frac{(1-c_0)a}{c_0 + (1-c_0)a} \leq 1 + R(1-c_0).$$

Thus

$$\|\mathcal{K}_J f\|_M \leq 1 + (1-c_0)\left(\|f\|_M - 1\right). \tag{5.1}$$

The function $\log(1+(1-c_0)u)/\log(1+u)$ is bounded above by some $1-\epsilon < 1$ as u varies over $(0, T-1]$, and setting $u = \|f\|_M - 1$ in (5.1) gives

$$\log \|\mathcal{K}_J f\|_M \leq \log(1+(1-c_0)(\|f\|_M - 1)) \leq (1-\epsilon)\log \|f\|_M,$$

proving the lemma. $\square$

Proceeding with the proof of Theorem 1.10, let C be a cutset with no vertices in the first generation,

$$\partial C = \{v \in C^i : \exists w \in C \text{ with } v \to w\},$$

and δ be defined on C. Clearly, for continuous strictly positive functions $h_1, \ldots, h_k$,

$$\|\bigodot(h_1, \ldots, h_k)\|_M \leq \prod_{i=1}^k \|h_i\|_M.$$

We have also previously seen (Lemma 2.3) that all densities that arise are uniformly bounded away from 0 and ∞ and hence there is a uniform bound on the $\| \; \|_M$ that arise. We can therefore choose ϵ from Lemma 5.2. Next for any $v \in C^i \setminus \partial C$, applying the fundamental recursion gives

$$\begin{aligned}
\log \|f_{C,v}^{\mathcal{J},\delta}\|_M &= \log \|\bigodot(\mathcal{K}_{\mathcal{J}(\overline{vw_1})} f_{C,w_1}^{\mathcal{J},\delta}, \ldots, \mathcal{K}_{\mathcal{J}(\overline{vw_k})} f_{C,w_k}^{\mathcal{J},\delta}\|_M \\
&\leq \sum_{i=1}^k \log \|\mathcal{K}_{\mathcal{J}(\overline{vw_i})} f_{C,w_i}^{\mathcal{J},\delta}\|_M \\
&\leq \sum_{i=1}^k (1-\epsilon) \log \|f_{C,w_i}^{\mathcal{J},\delta}\|_M.
\end{aligned}$$

Working backwards, we find that for any cutset C,

$$\log \|f_{C,o}^{\mathcal{J},\delta}\|_M \leq \sum_{w \in \partial C} (1-\epsilon)^{|w|} \log \|f_{C,w}^{\mathcal{J},\delta}\|_M.$$

Since $br(\Gamma) = 1$ one can choose a sequence of cutsets $\{C_n\}$ such that $\sum_{w \in \partial C_n} (1-\epsilon)^{|w|} \to 0$. The uniform bound on $\|f_{C,w}^{\mathcal{J},\delta}\|_M$ implies that for any sequence of functions δ_n on C_n,

$$\lim_{n \to \infty} \log \|f_{C_n,o}^{\mathcal{J},\delta_n}\|_M = 0,$$

which along with Lemma 5.1 proves the theorem. □

Olle Häggström pointed out to us that this result could also be obtained using ideas from disagreement percolation.

6 Proof of Theorem 1.14.

While we assume that q is an integer, the case of nonintegral q can be made sense of via the random cluster representation, and it is worth noting here that the break between $q = 2$ and $q = 3$ happens at $q = 2 + \epsilon$. See [11] for a discussion of the qualitative differences between the random cluster model on a tree when $q \leq 2$ as opposed to $q > 2$.

Lemma 6.1 *Assume that all of the hypotheses of Theorem 3.1 are in force (in particular, (2.20) and $br(\Gamma) \cdot \mathbf{Op}_J < 1$ hold and so there is no RPT for the parameter J) and in addition that $\sup_{y \in \mathbf{S}} \|K_{J,y}\| < \infty$ and (2.20) holds for all $f \in \mathcal{P}(J)$ (instead of just $\mathcal{P}_+(J)$). Then there is a tree Γ' with $br(\Gamma') = br(\Gamma)$ such that Γ' has no PT for the parameter J.*

Proof. We mimic the proof of Theorem 3.1. Choose ϵ, ϵ_0 and cutsets $\{C_n\}$ as in the proof of Theorem 3.1 where we can assume that the cutsets $\{C_n\}$ are disjoint. Choose an integer m sufficiently large so that the m-fold iterated convolution operator $\mathcal{K}_J^m$ satisfies $||\mathcal{K}_J^m \delta_y - 1|| \leq \epsilon_0 \mathbf{Op}_J$. For each increasing sequence $\{n(k) : k = 1, 2, \ldots\}$ of integers, define a tree Γ' by replacing each edge from an element of $C_{n(k)}$ to its parent by m edges in series, for all cutsets in the sequence $\{C_{n(k)}\}$. It is not too great an abuse of notation to let C_n denote the cutset of Γ' consisting of the same vertices as before. It is now possible to establish (3.4) for all $v \in D$, where D is the set of vertices in Γ' that are in C^i and in Γ (i.e., are not in a chain of parallel edges that was added). The only adjustment in the proof is as follows. Use Lemma 2.2 to represent $f_{C_n,v}^{J,+}$ in terms of $f_{C_n,w}^{J,+}$ where w are the children of v in Γ rather than in Γ', i.e., we leap the whole chain of m edges at once. Then the case $w \in C_n$ that was handled by the choice of J' is replaced by a case $w \in \Gamma' \backslash \Gamma$, which is handled by the choice of m. In fact, (3.4) holds when $+$ is replaced by any boundary condition as the exact same proof shows. By choosing $\{n(k)\}$ sufficiently sparse, we can ensure that $br(\Gamma') = br(\Gamma)$. Fixing any such choice of $\{n(k)\}$, it follows that there is no phase transition by the above together with Proposition 1.3. $\square$

We proceed now with the description of a counterexample. For Γ_1, we choose the homogeneous binary tree, where each vertex has precisely 2 children. Recall from Section 4.2 that under $+$ boundary conditions, the functions $f_{C,v}^{J,+}$ all lie in a one-dimensional set. The most convenient parameterization for the segment is by the log-likelihood ratio of state $\hat{0}$ to the other states. Thus the probability measure $a\delta_0 + \sum_{i=1}^{q-1}((1-a)/(q-1))\delta_i$ is mapped to the value $\log[(q-1)a/(1-a)]$. Let $g(v)$ denote the log-likelihood ratio at v under some interaction strength and boundary conditions. The recursion (2.5) of Lemma 2.2 boils down to

$$g(v) = \sum_{v\to w} \phi(g(w)); \quad \phi(z) := \log \frac{pe^z + 1 - p}{\frac{1-p}{q-1}e^z + (1 - \frac{1-p}{q-1})},$$

where

$$p := e^J/(e^J + (q-1)e^{-J}). \tag{6.1}$$

Taking a Taylor expansion to the second order gives

$$\phi(z) = \left(p - \frac{1-p}{q-1}\right) z + \frac{1-p}{2(q-1)^2}[p(q-1)^2 - (q-1) + (1-p)]z^2 + O(z^3).$$

To see that the second derivative is positive at 0 for $q > 2$, first take the q-derivative of the z^2 coefficient which is $[q + 2p - 3](1-p)/(2(q-1)^3)$. The definition of p and the fact that $J > 0$ imply that $p > 1/q \geq 1/(2(q-1))$. Since $x + 1/(x-1) - 3 > 0$ on $(2, \infty)$ and $2p > 1/(q-1)$, it follows that the z^2 coefficient has a positive q-derivative for $q \geq 2$, and is therefore positive for all $q > 2$. (This also implies that for $q \in (2-\delta, 2)$ for some δ, the function ϕ is concave (see [20] for a detailed analysis of the critical case $q = 2$).)

The Taylor expansion gives $\phi'(0) = p - (1-p)/(q-1)$. Note that $p_0 := (q+1)/(2q)$ satisfies $p_0 - (1-p_0)/(q-1) = 1/2$. The value of p_0 is chosen to make $\phi'(0) = 1/2$; by convexity of ϕ near zero, there is an interval $I := (p_0 - \epsilon, p_0)$ such that for $p \in I$, the equation $\phi(z) = z/2$ has a positive solution, call it $z(p)$. Take $\epsilon > 0$ so small that $p_0 - \epsilon > 1/q$. For any $1 > p > 1/q$ there is a unique $J > 0$ such that (6.1) holds. If $p \in I$, then $z(p)$ is a fixed point for the function 2ϕ and it is easy to see by induction that under $+$ boundary conditions on the binary tree, one will always have $g(v) \geq z(p)$. Thus we have shown that Γ_1 has a phase transition for any J such that $p \in I$.

To find Γ_2, we examine the connection between p_0 and $\|\mathcal{K}_J\|$ where for the rest of the proof, the operator norm refers to the L^∞ norm on the orthogonal complement of the constants. Observe that

$$p - \frac{1-p}{q-1} = \frac{e^J}{e^J + (q-1)e^{-J}} - \frac{e^{-J}}{e^J + (q-1)e^{-J}} = \|\mathcal{K}_J\|$$

by the computation in Section 4.2. Thus p_0 is chosen to make $\|\mathcal{K}_J\| = 1/2$ and for any $p \in I$, $\|\mathcal{K}_J\| < 1/2$. Fix any J so that $p \in I$, and let Γ be any tree with

$$2 = br(\Gamma_1) < br(\Gamma) < \|\mathcal{K}_J\|^{-1}.$$

Let Γ' be as in Lemma 6.1 and set $\Gamma_2 = \Gamma'$. Then there is no phase transition on Γ_2 for the chosen parameters, and since we have seen there is a phase transition for Γ_1, this completes the proof of Theorem 1.14. $\square$

Acknowledgements. We thank Richard Askey for discussions and showing us the proof of Lemma 2.8, Jöran Bergh, Yuval Peres and Paul Terwilliger for discussions, Anton Wakolbinger for providing us with reference [7] and the referee for a correction and some suggestions.

References

[1] Adel'son-Vel'skii, G., Veisfeiler, B., Leman, A. and Faradzev, I. (1969). Example of a graph without a transitive automorphism group. *Soviet Math. Dokl.* **10** 440–441.

[2] Aizenman, M., Chayes, J. T., Chayes, L. and Newman, C. M. (1988) Discontinuity of the magnetization in one–dimensional $1/|x-y|^2$ Ising and Potts models, *J. Stat. Phy.* **50** 1–40.

[3] Askey, R. (1974). *Orthogonal Polynomials and Special Functions.* S.I.A.M. Regional conferences in applied mathematics no. 21, J.W. Arrowsmith, Ltd.: Bristol, England.

[4] Brouwer., A., Cohen, A. and Neumaier, A. (1989). *Distance Regular Graphs.* Modern Surveys in Mathematics, Ser. 3, Bd. 18. Springer-Verlag: New York.

[5] Biggs, N. (1993). *Algebraic Graph Theory, 2nd Ed.* Cambridge University Press:Cambridge.

[6] Cassi, D. (1992). Phase transition and random walks on graphs: a generalization of the Mermin–Wagner theorem to disordered lattices, fractals, and other discrete structures. *Phys. Rev. Lett.* **68** 3631–3634.

[7] Eisele, M. (1994). *Phase transitions may be absent on graphs with transient random walks. Unpublished manuscript.*

[8] Evans, W., Kenyon, C., Peres, Y. and Schulman, L.J. (1998). Broadcasting on trees and the Ising model. *Preprint.*

[9] Furstenberg, H. (1970). Intersections of Cantor sets and transversality of semigroups. In *Problems in analysis. Sympos. in Honor of Salomon Bochner, Princeton Univ. (R. C. Gunning, ed.)* 41–59. Princeton Univ. Press, Princeton, N.J.

[10] Georgii, H.-O. (1988). *Gibbs Measures and Phase Transitions.* de Gruyter: New York.

[11] Häggström, O. (1996). The random-cluster model on a homogeneous tree. *Probab. Theory Related Fields* **104** 231–253.

[12] Liggett, T. M. (1996). Multiple transition points for the contact process on a binary tree. *Ann. Probab.* **24** 1675–1710.

[13] Lyons, R. (1989). The Ising model and percolation on trees and tree-like graphs. *Commun. Math. Phys.* **125** 337–353.

[14] Lyons, R. (1990). Random walks and percolation on trees. *Ann. Probab.* **18** 931–958

[15] Merkl, F., and Wagner, H. (1994). Recurrent random walks and the absence of continuous symmetry breaking on graphs. *J. Stat. Phy.* **75** 153–165.

[16] Monroe, J. L., and Pearce, P. A. (1979). Correlation inequalities for vector spin models. *J. Stat. Phy.* **21** 615–633.

[17] Natterer, F. (1986). *The Mathematics of Computerized Tomography.* John Wiley, Stuttgart.

[18] Patrascioiu A. and Seiler, E. (1992). Phase structure of two-dimensional spin models and percolation. *J. Stat. Phy.* **69** 573–595.

[19] Pemantle, R. (1992). The contact process on trees. *Ann. Probab.* **20** 2089–2116.

[20] Pemantle, R., and Peres, Y., Recursions on trees and the Ising model, *Preprint.*

[21] Rainville, E. D. (1960). *Special Functions.* MacMillan, New York.

[22] Stacey, A. (1996). The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24** 1711–1726.

[23] Terwilliger, P. (1998). *Unpublished lecture notes.*

Robin Pemantle
Department of Mathematics
University of Wisconsin-Madison
Van Vleck Hall
480 Lincoln Drive
Madison, WI 53706
pemantle@math.wisc.edu

Jeffrey E. Steif
Department of Mathematics
Chalmers University of Technology
S–41296 Gothenburg
Sweden
steif@math.chalmers.se