The weak limit of Ising models on locally tree-like graphs

Andrea Montanari[*], Elchanan Mossel[†] and Allan Sly[‡]

December 3, 2009

Abstract

We consider the Ising model with inverse temperature β and without external field on sequences of graphs G_n which converge locally to the k-regular tree. We show that for such graphs the Ising measure locally weak converges to the symmetric mixture of the Ising model with $+$ boundary conditions and the $-$ boundary conditions on the k-regular tree with inverse temperature β. In the case where the graphs G_n are expanders we derive a more detailed understanding by showing convergence of the Ising measure condition on positive magnetization (sum of spins) to the $+$ measure on the tree.

1 Introduction

An *Ising model on the finite graph* G (with vertex set V, and edge set E) is defined by the following distribution over $\underline{x} = \{x_i : i \in V\}$, with $x_i \in \{+1, -1\}$

$$\mu(\underline{x}) = \frac{1}{Z(\beta, B)} \exp\Big\{\beta \sum_{(i,j)\in E} x_i x_j + B \sum_{i\in V} x_i\Big\}. \tag{1.1}$$

The model is *ferromagnetic* if $\beta \geq 0$ and, by symmetry, we can always assume $B \geq 0$. Here $Z(\beta, B)$ is a normalizing constant (partition function).

The most important feature of the distribution $\mu(\,\cdot\,)$ is the 'phase transition' phenomenon. On a variety of large graphs G, for large enough β and $B = 0$, the measure decomposes into the convex combination of two well separated simpler components. This phenomenon has been studied in detail in the case of grids [2, 3, 4, 5], and on the complete graph [1]. In this paper we consider sequences of regular graphs $G_n = (V_n, E_n)$ with increasing vertex sets $V_n = [n] = \{1, \dots, n\}$ that converge locally to trees and prove a local characterization of the corresponding sequence of measures $\mu_n(\,\cdot\,)$, which corresponds to the phase transition phenomenon.

More precisely, consider the case in which G_n is a sequence of regular graphs of degree $k \geq 3$ with diverging girth. The neighborhood of B_i any vertex i in G_n converges to an infinite regular tree of degree k. It is natural to assume that the marginal distribution $\mu_{n,\mathsf{B}_i}(\,\cdot\,)$ converges to the marginal of a neighborhood of the root for an Ising Gibbs measure on the infinite tree. For large β, however, there are uncountably many Gibbs measures on the tree so it is natural to ask which is the limit

[*]Department of Electrical Engineering and Department of Statistics, Stanford University

[†]Faculty of Mathematics and Computer Science, Weizmann Institute and Departments of Statistics and Computer Science, UC Berkeley

[‡]Microsoft Research, Redmond, WA

A special role is played by the plus/minus boundary conditions Gibbs measures on the infinite tree, to be denoted, respectively, by $\nu^+(\,\cdot\,)$ and $\nu^-(\,\cdot\,)$. It was proved in [9] that, for any β, and any $B > 0$, $\mu_n(\,\cdot\,)$ converges locally to ν^+ as $n \to \infty$ and by symmetry when $B < 0$ $\mu_n(\,\cdot\,)$ converges locally to ν^- as $n \to \infty$.

In this paper we cover the remaining (and most interesting) case proving that

$$\mu_n(\,\cdot\,) \quad \underset{n}{\longrightarrow} \quad \frac{1}{2}\,\nu^+(\,\cdot\,) \;+\; \frac{1}{2}\,\nu^-(\,\cdot\,) \quad \text{for } B = 0 \text{ and any } \beta \geq 0 \;. \tag{1.2}$$

In fact, we prove a sharper result. If $\mu_{n,+}(\,\cdot\,)$ and $\mu_{n,-}(\,\cdot\,)$ denote the Ising measure (1.1) conditioned to, respectively, $\sum_{i\in V} x_i > 0$ and $\sum_{i\in V} x_i < 0$, then we have

$$\mu_{n,\pm}(\,\cdot\,) \quad \underset{n}{\longrightarrow} \nu_\pm(\,\cdot\,) \quad \text{for } B = 0 \text{ and any } \beta \geq 0\,, \tag{1.3}$$

and moreover the convergence above holds for almost all vertices of the graph. Since $\mu_n = \frac{1}{2}\mu_{+,n} + \frac{1}{2}\mu_{-,n}$ (exactly for n odd and approximately for even n), this result implies (1.2).

2 Definitions and main results

2.1 Locally tree-like graphs

We denote by $G_n = (V_n, E_n)$ a graph with vertex set $V_n \equiv [n] = \{1, \ldots, n\}$. The distance $d(i,j)$ between $i, j \in V_n$ is the length of the shortest path from i to j in G_n. Given a vertex $i \in V_n$, we denote by $\mathsf{B}_i(t)$ the set of vertices whose distance from i is at most t (and with a slight abuse of notation it will also denote the subgraph induced by those vertices). We will let I denote a vertex chosen uniformly from the vertices V_n, let U_n denote the measure induced by I and let J denote a uniformly random neighbor of I.

This paper is concerned by sequence of graphs $\{G_n\}_{n\in\mathbb{N}}$ of diverging size, that converge locally to T_k, the infinite rooted tree of degree k. Let $\mathsf{T}_k(t)$ be the subset of vertices of T_k whose distance from the root $\emptyset$ is at most t (and, by an abuse of notation, the induced subgraph). For a rooted tree T, we write $T \simeq \mathsf{T}_k(t)$ if there is a graph isomorphism between T and $\mathsf{T}_k(t)$ which maps the root of T to that of $\mathsf{T}_k(t)$. The following definition defines what we mean by convergence in the local weak topology.

Definition 2.1. *Consider a sequence of graphs $\{G_n\}_{n\in\mathbb{N}}$, and let U_n be the law of a uniformly random vertex I in V_n. We say that $\{G_n\}$* converges locally *to the degree-k regular tree T_k if, for any t,*

$$\lim_{n\to\infty} U_n\{\mathsf{B}_I(t) \simeq \mathsf{T}_k(t)\} = 1\,. \tag{2.1}$$

Part of our results hold for sequences of expanders (more precisely, edge expanders), whose definition we now recall. For a subset of vertices $S \subset V$, we will denote by ∂S the subset of edges $(i,j) \in E$ having only one endpoint in S.

Definition 2.2. *The k-regular graph $G = (V, E)$ is a (γ, λ) (edge) expander if, for any set of vertices $S \subseteq V$ with $|S| \leq n\gamma$, $|\partial S| \geq \lambda S$.*

2.2 Local weak convergence

In analogy with the definition of locally tree-like graph sequences, we introduce local weak convergence for Ising measures. This done in two different ways. First one can look at a random vertex and the random configuration in the neighbourhood of the vertex and examine its limiting measure. Alternatively, we may choose a random vertex and consider the marginal distribution of the variables in a neighborhood under the Ising model. This induces (via the random choice of the vertex) a distribution over probability measures. We can therefore ask whether this measure converges to a probability measure over Gibbs measures.

Recall that an Ising measure μ on the infinite tree T_k may be either defined as a weak limit of Gibbs measures on $\mathsf{T}_k(t)$ or in terms of the DLR conditions, see e.g. [11]. An Ising model is in particular a probability measure over $\{-1,+1\}^{\mathsf{T}_k}$ endowed with the σ-algebra generated by cylindrical sets. We let $\mathcal{G}_k$ denote the space of Ising Gibbs measures on T_k and let $\mathcal{H}_k$ denote the space of all probability measures on $\{+1,-1\}^{\mathsf{T}_k}$. We endow both these spaces with the topology of weak convergence. Since $\{+1,-1\}^{\mathsf{T}_k}$ is compact, $\mathcal{G}_k$ and $\mathcal{H}_k$ are also compact in the weak topology by Prohorov's theorem.

We define $\mathcal{M}_k$ (respectively $\mathcal{M}_k^{\mathcal{G}}$) to be the space of probability measures over $(\mathcal{H}_k, \mathcal{B}_{\mathcal{H}})$ (resp. $(\mathcal{G}_k, \mathcal{B}_{\mathcal{H}})$), with $\mathcal{B}_\Omega$ the Borel σ-algebra. Also $\mathcal{M}_k$, $\mathcal{M}_k^{\mathcal{G}}$ are compact in the weak topology.

We will use generically μ for Ising measures on G_n and ν for Ising measure on T_k. For a finite subset of vertices $S \subseteq V_n$, we let μ^S be the marginal of μ on the variables x_j, $j \in S$. We the shorthand μ^t for when $S = \mathsf{B}_i(t)$ is the ball of radius t about i (i should be clear from the context). For a measure $\nu \in \mathcal{G}_k$ we let ν^t denote its marginal over the variables x_j, $j \in \mathsf{T}_k(t)$. In other words ν^t is the projection of ν on $\{+1,-1\}^{\mathsf{T}_k(t)}$. For a measure $\mathfrak{m} \in \mathcal{M}_k$ we let $\mathfrak{m}^t$ denote the measure on the space of measures on $\{+1,-1\}^{\mathsf{T}_k(t)}$ induced by such projection.

Definition 2.3. *Consider a sequence of graphs/Ising measures pairs $\{(G_n, \mu_n)\}_{n\in\mathbb{N}}$ and let $\mathbb{P}_n^t(i)$ denote the law of the pair $(\mathsf{B}_i(t), \underline{x}_{\mathsf{B}_i(t)})$ when $\underline{x}$ is drawn with distribution μ_n and $i \in [n]$ is vertex in the graph. Let U_n denote the uniform measure over a random vertex $I \in [n]$. Let $\mathbb{P}_n^t = \mathbb{E}_{U_n}(\mathbb{P}_n^t(I))$ denote the average of $\mathbb{P}_n^t(I)$.*

A. *The first mode of convergence concerns picking a random vertex I and a random local configuration in the neighbourhood of I. Formally, for $\bar{\nu} \in \mathcal{G}_k$ we say that $\{\mu_n\}_{n\in\mathbb{N}}$* converges locally on average to $\bar{\nu}$ *if for any t and any $\epsilon > 0$ it holds that*

$$\lim_{n\to\infty} d_{\mathrm{TV}}\left(\mathbb{P}_n^t, \delta_{\mathsf{T}_k(t)} \times \bar{\nu}^t\right) = 0. \tag{2.2}$$

B. *A stronger form of convergence involves picking a random vertex I and the associated random local measure $\mathbb{P}_n^t(I)$ and asking if this distribution of distributions converges. Formally, we say that* the local distributions of $\{\mu_n\}_{n\in\mathbb{N}}$ converge locally to $\mathfrak{m} \in \mathcal{M}_k^{\mathcal{G}}$ *if it holds that the law of $\mathbb{P}_n^t(I)$ converges weakly to $\delta_{\mathsf{T}_k(t)} \times \mathfrak{m}^t$ for all t.*

C. *If $\mathfrak{m}$ is a point mass on $\bar{\nu} \in \mathcal{G}_k$ and if the local distributions of $\{\mu_n\}_{n\in\mathbb{N}}$ converge locally to $\mathfrak{m}$ then we say that $\{\mu_n\}_{n\in\mathbb{N}}$* converges in probability locally to $\bar{\nu}$. *Equivalently convergence in probability locally to $\bar{\nu}$ says that for any t and any $\epsilon > 0$ it holds that*

$$\lim_{n\to\infty} U_n\left(d_{\mathrm{TV}}(\mathbb{P}_n^t(I), \delta_{\mathsf{T}_k(t)} \times \nu^t) > \epsilon\right) = 0. \tag{2.3}$$

It is easy to verify that $C \Rightarrow B \Rightarrow A$.

Similar notions of the convergence has been studied before under the name metastates for Gibbs measures. Aizenman and Wehr [6], while investigating the quenched behaviour of lattice random field models, introduced the notion of a metastate which is a probability measures over Gibbs measures as a function of the disorder (the random field). Here, rather than taking a finite graph and choosing a random vertex they take a fixed random environment in $\mathbb{Z}^d$, and study the measure over increasing finite volumes. Rather than prove convergence (which depending on the model may not hold) they take subsequential limits and study the properties of these limiting distributions of Gibbs measures (metastates). Another, similar notion of convergence to metastates was developed by Newman and Stein [16] where they took the empirical measure over Gibbs measures at over increasing volumes to study spin-glasses. More references and discussions can be found in [13].

In order to state our main result formally, we recall that an Ising measure on T_k is Gibbs if, for any integer $t \geq 0$

$$\mu^{\mathsf{T}_k(t)|\mathsf{T}^c{}_k(t)}(\underline{x}_{\mathsf{T}_k(t)}|\underline{x}_{\mathsf{T}^c{}_k(t)}) = \frac{1}{Z_{t,\underline{x}}(\beta)} \exp\left\{\beta \sum_{(i,j)\in E(\mathsf{T}_k(t+1))} x_i x_j\right\}, \tag{2.4}$$

where $Z_t(\beta)$ is a normalization function that depends on the conditioning, namely on $\underline{x}_{\mathsf{T}_k(t+1)\setminus\mathsf{T}_k(t)}$.

It is well known that if $(k-1)\tanh\beta \leq 1$, there exist only one Gibbs measure on a k-regular tree while for $(k-1)\tanh\beta > 1$ the Gibbs measures form a non-trivial convex set (see e.g. [11]). Two of its extreme points, ν_+ and ν_- play a special role in the following. The 'plus-boundary conditions' measure ν_+ is defined as the monotone decreasing limit (with respect to the natural partial ordering on the space of configurations $\{+1,-1\}^{\mathsf{T}_k}$) of ν_+^t as $t \to \infty$, where ν_+^t is the measure on $\underline{x}_{\mathsf{T}_k(t)}$ defined by

$$\nu_+^t(\underline{x}_{\mathsf{T}_k(t)}) = \frac{1}{Z_{+,t}(\beta)} \exp\left\{\beta \sum_{(i,j)\in E(\mathsf{T}_k(t))} x_i x_j\right\} \prod_{i\in\mathsf{T}_k(t)\setminus\mathsf{T}_k(t-1)} \mathbb{I}(x_i = +1)\,. \tag{2.5}$$

The measure ν_- is defined analogously, by forcing spins on the boundary to take value -1 instead of $+1$. The two measures are obviously related through spin reversal. Further it well known (and easy to prove) that for any Gibbs measure ν we have $\nu_- \preceq \nu \preceq \nu_+$ (with $\preceq$ the stochastic ordering induced by the partial ordering on $\{+1,-1\}$ configurations, see e.g. [14]). Our main result may be now stated as follows

Theorem 2.4. *Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of k-regular graphs that converge locally to the tree T_k. For $(k-1)\tanh\beta > 1$, define the sequence $\{\mu_n\}_{n\in\mathbb{N}}$, $\{\mu_{n,+}\}_{n\in\mathbb{N}}$ by*

$$\mu_{n,+}(\underline{x}) = \frac{1}{Z_{n,+}(\beta)} \exp\Big\{\beta \sum_{(i,j)\in E_n} x_i x_j\Big\}\, \mathbb{I}\Big\{\sum_{i\in V_n} x_i > 0\Big\}, \tag{2.6}$$

$$\mu_n(\underline{x}) = \frac{1}{Z_n(\beta)} \exp\Big\{\beta \sum_{(i,j)\in E_n} x_i x_j\Big\}. \tag{2.7}$$

Then

I. μ_n converges locally in probability to $\frac{1}{2}(\nu_+ + \nu_-)$

II. If the graphs $\{G_n\}$ are $(1/2,\lambda)$ edge expanders for some $\lambda > 0$, then $\mu_{n,+}$ converges locally in probability to the plus-boundary Gibbs measure on the infinite tree ν_+.

This characterization has a number of useful consequences. In particular, 'spatial' averages of local functions are roughly constant under the conditional measure $\mu_{n,+}$. To be more precise, for each $i \in V_n$, let

$$f_{i,n} : \{+1,-1\}^{\mathsf{B}_i(\ell)} \to [-1,1],$$

be a function of its neighborhood $\mathsf{B}_i(\ell)$.

Theorem 2.5. *Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of k-regular $(1/2,\lambda)$ edge expanders, for some $\lambda > 0$, that converge locally to the tree T_k. For each n, let $\{f_{i,n}\}_{i=1}^n$ be a collection of local functions as above. Then, for any $\varepsilon > 0$*

$$\lim_{n\to\infty} \mu_{n,+}\Big\{\Big|\frac{1}{n}\sum_{i\in V_n}[f_{i,n}(\underline{x}_{\mathsf{B}_i(\ell)}) - \mu_{n,+}\left(\frac{1}{n}\sum_{i\in V_n} f_{i,n}(\underline{x}_{\mathsf{B}_i(\ell)})\right)]\Big| \geq \varepsilon\Big\} = 0\,. \tag{2.8}$$

The proof can be found in Section 5.

2.3 Examples and remarks

Notice that, for $(k-1)\tanh\beta \le 1$, the set of Ising Gibbs measures on T_k contains a unique element, that can be obtained as limit of free boundary measures. Therefore, the local limits of $\{\mu_n\}_{n\in\mathbb{N}}$, $\{\mu_{n,+}\}_{n\in\mathbb{N}}$ coincide trivially with this unique Gibbs measure.

Therefore, the claim I is proved under the weakest possible, hypothesis, namely local convergence of the graphs to T_k. An important class of graphs for which Theorem 2.4 is applicable are random k-regular graphs. These are known to converge locally to T_k [17].

The expansion condition (or an analogous 'connectedness' condition) is needed to obtain the convergence of the conditional measures $\mu_{n,+}$. For example consider r identical but disjoint graphs on n/r vertices. Then conditioning on the sum of the spins being positive the probability that the sum of spins in a specific component is positive is of order $r^{-1/2}$. Therefore in this case we have:

$$\mu_{n,+} \to (1-q)\nu_+ + q\nu_- \,,$$

with $q = 1/2 - O(r^{-1/2})$. A similar construction may be repeated with a small number of edges connecting different components, e.g., when the components are connected in a cyclic fashion.

In order to identify the limit for μ_n and obtain our results, there are a number of challenges that need to be overcome. First, while soft compactness arguments imply that subsequential limits exist, such arguments do not imply the existence of a proper limit. Second, recalling that there are uncountably many extremal Gibbs measures for T_k, it is remarkable we are able to identify precisely those that appear in the limit. Finally, for conditional measures such as $\mu_{n,+}$ it is not even a priori clear that (subsequential) limits are in fact Gibbs measures.

2.4 Proof strategy

The basic idea of the proof is the following. Look at a ball of radius t around a vertex i in G_n. Since G_n is tree like, the ball is with high probability a tree. The measure μ_n restricted to the ball is clearly a Gibbs measure on a tree of radius t. The same is true (although less obvious) for $\mu_{n,+}$.

In order to characterize the limit of this measure as $n \to \infty$,

1. The probability of agreement between neighboring spins in the ball is asymptotically the same as in the measure ν_+ on the infinite tree.

2. We further show that ν_+ maximizes the probability of agreement between neighboring spins among all Gibbs measures on the tree. These two facts together imply that any local limit must converge to a convex combination of ν_+ and ν_-.

3. By symmetry this already implies converges of μ_n to $\frac{1}{2}(\nu_+ + \nu_-)$. Note that this step does not require expansion, just the local weak convergence of the tree.

4. In order to deal with the conditional measure, we use expansion to show that it is unlikely that simultaneously a positive fraction of the vertices have their neighborhood "in the + state" and another positive fraction "in the − state".

3 Proof of the main theorem

We now proceed with the proof. For each of claims I and II we break the proof into 3 steps:

(i) We consider a subsequence of sizes $\{n(m)\}_{m\in\mathbb{N}}$ along which $\mu_{n(m)}$ or $\mu_{n(m),+}$ converge locally in average to a limit $\bar{\nu}$ or $\bar{\nu}_+$ (respectively).

(*ii*) We prove that any such limit is in fact always the same and is $\bar{\nu} = (1/2)(\nu_+ + \nu_-)$ for $\mu_{n(m)}$ and (using expansion) $\bar{\nu}_+ = \nu_+$ for $\mu_{n(m),+}$. As a consequence the sequences themselves converge.

(*iii*) Finally we show how is it possible to deduce local convergence from convergence in average.

3.1 Subsequential limits

The construction of subsequential weak limits is based on a standard diagonal argument, for similar results see [8]. For the sake of simplicity we refer to the measures $\mu_{n,+}$, and construct the subsequential limit $\bar{\nu}_+$, but the same procedure works for μ_n with limit $\bar{\nu}$. Let $\mathsf{B}_I(t)$ be the ball of radius t centered at a uniformly random vertex I in V_n, and $\underline{x}$ be an Ising configuration with distribution $\mu_{n,+}$. If $\mathbb{P}_n$ denotes the joint distribution of $(\mathsf{B}_I(t), \underline{x}_{\mathsf{B}_I(t)})$, we let

$$\mu^t_{+,n}(\underline{x}^*_{\mathsf{T}_k(t)}) \equiv \mathbb{P}_n\big\{(\mathsf{B}_I(t), \underline{x}_{\mathsf{B}_I(t)}) \simeq (\mathsf{T}_k(t), \underline{x}^*_{\mathsf{T}_k(t)})\big\}\,. \tag{3.1}$$

Since this is a sequence of measures over a finite state space, it converges over some subsequence $\{n_t(m)\}_{m\ge 0}$. Further, since by hypothesis $\mathbb{P}_n\{\mathsf{B}_i(t) \simeq \mathsf{T}_k(t)\} \to 1$, the limits of $\mu^t_{+,n_t(m)}$ and $\mu^t_{n_t(m)}$ are in fact probability measures. We call the limit $\bar{\nu}^t_+$.

Fix one of these subsequences $\{n_{t_0}(m)\}_{m\ge 0}$ for $t = t_0$, leading to the limit $\bar{\nu}^{t_0}_+$, and recursively refine it to $\{n_{t_0}(m)\}_{m\ge 0} \supseteq \{n_{t_0+1}(m)\}_{m\ge 0} \supseteq \{n_{t_0+2}(m)\}_{m\ge 0} \supseteq \dots$ leading to limits $\bar{\nu}^t_+$ for all $t \ge t_0$. Notice that, for any graph G_n, any vertex i and any t we have

$$\mu^t_{n,+}(\underline{x}_{\mathsf{B}_i(t)}) = \sum_{\underline{x}_{\mathsf{B}_i(t+1)\setminus\mathsf{B}_i(t)}} \mu^{t+1}_{n,+}(\underline{x}_{\mathsf{B}_i(t+1)})\,. \tag{3.2}$$

As a consequence, for any t, the measures limit $\bar{\nu}^{(t)}_+$ measure satisfies

$$\bar{\nu}^t_+(\underline{x}_{\mathsf{T}_k(t)}) = \sum_{\underline{x}_{\mathsf{T}_k(t+1)\setminus\mathsf{T}_k(t)}} \bar{\nu}^{t+1}_+(\underline{x}_{\mathsf{T}_k(t+1)})\,. \tag{3.3}$$

By Kolmogorov extension theorem, there exist measures $\bar{\nu}_+$ over $\{+1,-1\}^{\mathsf{T}_k}$ such that $\bar{\nu}^t_+$ are the marginals of $\bar{\nu}_+$ over the variables in the subtree $\mathsf{T}_k(t)$. By taking the diagonal subsequence $n(m) = n_m(m)$ we obtain the desired subsequence $\{n(m)\}_{m\in\mathbb{N}}$ such that $\mu_{n(m),+}$ converges locally on average to $\bar{\nu}_+$.

3.2 $\bar{\nu} = \frac{1}{2}(\nu_+ + \nu_-)$

In this section we carry out our program in the case of the unconditional measures μ_n. It is immediate that, since each of the measures μ^t_n is a Gibbs measure on T_k (although with a complicate boundary condition), the limit measure $\bar{\nu}$ is also a Gibbs measure on T_k (i.e. $\bar{\nu} \in \mathcal{G}_k$).

For proving convergence of the unconditional measure we need two lemmas. The first one establishes that the $+$ (equivalently $-$) Gibbs measure ν_+ has the correct expected number of edge disagreements (in physics terms, the correct energy density).

Lemma 3.1. *Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of k-regular graphs converging locally to T_k, let I be a uniformly random vertex in G_n, and J be chosen uniformly among its k neighbors. Then*

$$\lim_{n\to\infty} \mathbb{E}_{U_n}\, \mu_{n,+}(x_I \cdot x_J) = \lim_{n\to\infty} \mathbb{E}_{U_n}\, \mu_n(x_I \cdot x_J) = \nu_+(x_\emptyset \cdot x_1) = \nu_-(x_\emptyset \cdot x_1)\,, \tag{3.4}$$

where 1 is one of the neighbors of the root in T_k, and $\mathbb{E}_{U_n}$ denotes the expectation over the random edge (I,J) in G_n.

For the proof of this Lemma we refer to Section 4.2. Notice that ν_+ and ν_- have the same expectation of the product $x_\emptyset x_1$ by symmetry under inversion $\{x_i\} \to \{-x_i\}$. The probability that the spins at $\emptyset$ and 1 agree is simply $(1+\nu(x_\emptyset \cdot x_1))/2$. The second Lemma shows that ν_+, ν_- are uniquely characterized by this agreement probability among all Ising Gibbs measures on T_k.

Lemma 3.2. *Let ν be a Gibbs measure for the Ising model on T_k. Then*

$$\nu(x_\emptyset \cdot x_1) \le \nu_+(x_\emptyset \cdot x_1) = \nu_-(x_\emptyset \cdot x_1)\,, \tag{3.5}$$

and the inequality is strict unless ν is a convex combination of ν_+ and ν_-.

The proof of this Lemma can be found in Section 4.3. We can now prove the following:

Proposition 3.3. *Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of k-regular graphs that converge locally to the tree T_k. Then for $(k-1)\tanh\beta > 1$, it holds that μ_n converges locally in average to $(1/2)(\nu_+ + \nu_-)$.*

Proof. By Lemma 3.1 and weak convergence, we have $\bar\nu(x_\emptyset \cdot x_1) = \nu_+(x_\emptyset \cdot x_1)$. By Lemma 3.2, $\bar\nu = (1-q)\nu_+ + q\bar\nu_-$ for some $q \in [0,1]$. On the other hand $\mu_{n,+}$ is symmetric under spin inversion for each n, and therefore $\bar\nu$ must be symmetric as well, whence $q = 1/2$ □

We can now prove the first part of our main result.

Proof (Theorem 2.4, part I). By a similar construction to the one recalled in Section 3.1, and compactness of $\mathcal{M}_k$, we can construct a subsequence $\{n(m)\}_{m\in\mathbb{N}}$ such that $\mu_{n(m)}$ converges locally (not only in average) to a distribution $\mathfrak{m}$ over $\mathcal{H}_k$. By the arguments above, $\mathfrak{m}$ is in fact a measure over the space Ising Gibbs measures $\mathcal{G}_k$.

We claim that any such subsequential weak limit $\mathfrak{m}$ is in fact a point mass at $(1/2)(\nu_+ + \nu_-)$. Since $\nu \mapsto \nu(x_\emptyset \cdot x_1)$ is continuous in the weak topology it follows that

$$\lim_{m\to\infty} \mathbb{E}_{U_n} \mu_{n(m)}(x_I \cdot x_J) = \int \nu(x_\emptyset \cdot x_1)\, \mathfrak{m}(\mathrm{d}\nu)\,. \tag{3.6}$$

By Lemma 3.1, this implies

$$\int \nu(x_\emptyset \cdot x_1)\, \mathrm{d}\mathfrak{m}(\nu) = \nu_+(x_\emptyset \cdot x_1)\,, \tag{3.7}$$

and therefore, by Lemma 3.2, $\mathfrak{m}$ is supported on Ising Gibbs measures ν that are convex combinations of ν_+ and ν_-. Finally, μ_n is almost surely symmetric for any n. Here 'symmetric' means that, for any configuration $\underline{x}_{\mathsf{B}_i(t)}$, $\mu_n^t(\underline{x}_{\mathsf{B}_i(t)}) = \mu_n^t(-\underline{x}_{\mathsf{B}_i(t)})$. Therefore $\mathfrak{m}$ is supported on Ising Gibbs measures that are symmetric.

There is only one Ising Gibbs measure that is a convex combination of ν_+ and ν_- and is symmetric, namely $\nu = (1/2)(\nu_+ + \nu_-)$. Hence $\mathfrak{m}$ is a point mass on this distribution. □

3.3 $\bar\nu_+ = \nu_+$

We now turn to the subsequence of conditional measures $\{\mu_{n(m),+}\}_{m\in\mathbb{N}}$ converging locally in average to $\bar\nu_+$. The goal of this subsection is to show that $\bar\nu_+$ is equal to ν_+.

For this we repeat the previous proof with two additional ingredients. First we need to show that $\bar\nu_+$ is a Gibbs measure on the tree T_k. This requires proof since the conditioning on $\{\sum_{i\in V_n} x_i > 0\}$ implies that the measures $\mu_{n,+}^t$ are not Gibbs measures. The Gibbs property is only recovered in the limit.

Second even after we have established that $\bar{\nu}_+$ is a Gibbs measure, this measure is not symmetric with respect to spin flip. Therefore the argument above only implies that $\bar{\nu}_+ = (1-q)\nu_+ + q\nu_-$. It remains to show that $q = 0$. This is where the expansion assumption is used. The first lemma we prove is the following:

Lemma 3.4. *Any subsequential limit $\bar{\nu}_+$ constructed as above is an Ising-Gibbs measure on T_k.*

We defer the proof to Section 4.1. Given Lemma 3.4 the following lemma follows immediately from Lemmas 3.1 and 3.2.

Lemma 3.5. *For any subsequential limit $\bar{\nu}_+$ there exists a $q \in [0,1]$ such that*

$$\bar{\nu}_+ = (1-q)\,\nu_+ + q\,\nu_-\,. \tag{3.8}$$

Proof. By Lemma 3.4 the measure $\bar{\nu}_+$ is an Ising Gibbs measure on T_k. If it was not a convex combination of ν_+ and ν_- a contradiction to Lemma 3.2 would be derived. □.

The last step consists of arguing that $q = 0$. Given a vertex i (either in a graph G_n of the sequence or of T_k), an integer $\ell \geq 1$ and a random Ising configuration $\underline{x}$, let

$$\mathsf{F}_i(\ell, \delta, \underline{x}) \equiv \mathbb{I}\Big\{ \sum_{j \in \mathsf{B}_i(\ell)} x_j \leq -\delta\, |\mathsf{B}_i(\ell)| \Big\}\,, \tag{3.9}$$

where $\delta \in (0,1)$ will be chosen below. Roughly speaking F_i indicates which vertices are in the "$-$ state". We will drop reference to δ and to the configuration $\underline{x}$ when clear from the context. The following lemmas will be proven in Section 4.4.

Lemma 3.6. *Let $\{G_n\}$ be a sequence of graphs converging locally to T_k, and, for each n, $\underline{x} = \underline{x}(n)$ be a configuration in the support of $\mu_{n,+}$. Then there exists n_0, depending on δ, ℓ and the graph sequence, but not on $\underline{x}$, such that, for all $n \geq n_0$,*

$$\mathbb{E}_{U_n}(\mathsf{F}_I(\ell, \delta, \underline{x})) \leq \frac{1}{1+\delta/2}\,, \tag{3.10}$$

where $\mathbb{E}_{U_n}$ denotes expectation with respect to the uniformly random vertex I in V_n.

The following lemma is an immediate consequence of the definition of local weak convergence.

Lemma 3.7. *Consider a uniformly random vertex I in G_n, let J be one of its neighbors (again uniformly random), and let $\{n(m)\}_{m\in\mathbb{N}}$ a subsequence of graph sizes along which $\mu_{n(m),+}$ converges locally on average to $\bar{\nu}_+$. Then we have*

$$\lim_{m\to\infty} \mathbb{E}_{U_{n(m)}} \mu_{n(m),+}(\mathsf{F}_I(\ell)) \;=\; \bar{\nu}_+(\mathsf{F}_\emptyset(\ell))\,, \tag{3.11}$$

$$\lim_{m\to\infty} \mathbb{E}_{U_{n(m)}} \mu_{n(m),+}(\mathsf{F}_I(\ell) \neq \mathsf{F}_J(\ell)) \;=\; \bar{\nu}_+(\mathsf{F}_\emptyset(\ell) \neq \mathsf{F}_1(\ell))\,, \tag{3.12}$$

with $\mathbb{E}$ denoting expectation with respect to the law $U_{n(m)}$ of vertices I and J, and 1 one of the neighbors of $\emptyset$.

Now the limit quantities can be estimated as follows.

Lemma 3.8. *Assume $(k-1)\tanh\beta > 1$ and let $\nu = (1-q)\nu_+ + q\nu_-$ be a mixture of the plus and minus measures for the Ising model on T_k. Then there exist $\delta = \delta(\beta) > 0$ such that, letting $\mathsf{F}_i(\ell) = \mathsf{F}_i(\ell, \delta; \underline{x})$,*

$$\lim_{\ell\to\infty} \nu(\mathsf{F}_\emptyset(\ell) = 1) = q\,, \tag{3.13}$$

$$\lim_{\ell\to\infty} \nu(\mathsf{F}_\emptyset(\ell) \neq \mathsf{F}_1(\ell)) = 0\,. \tag{3.14}$$

We can now prove the following:

Proposition 3.9. *Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of k-regular graphs that are $(1/2,\lambda)$ expanders for some $\lambda>0$ and converge locally to the tree T_k. Then for $(k-1)\tanh\beta>1$, it holds that $\mu_{n,+}$ converges locally on average to ν_+*

Proof. Let $n(m)$ be a subsequence along which $\mu_{n,+}$ converges locally on average to some $\bar\nu_+$. By Lemma 3.5 we can write this in the form $\bar\nu_+=(1-q)\,\nu_+ + q\,\nu_-$. Then by Eqs. (3.11), (3.12), for any $\varepsilon>0$, there exists ℓ, such that for large enough $n(m)$,

$$\mathbb{E}\,\mu_{n(m),+}(\mathsf{F}_I(\ell)) \;\geq\; q-\varepsilon\,, \tag{3.15}$$

$$\mathbb{E}\,\mu_{n(m),+}(\mathbb{I}\{\mathsf{F}_I(\ell)\neq\mathsf{F}_J(\ell)\}) \;\leq\; \varepsilon\,. \tag{3.16}$$

On the other hand, since G_n is a $(1/2,\lambda)$ expander, and using Eq. (3.10), we have

$$\sum_{(i,j)\in E_n}\mathbb{I}\{\mathsf{F}_i(\ell)\neq\mathsf{F}_j(\ell)\} \;\geq\; \lambda\,\min(\sum_{i\in V_n}\mathsf{F}_i(\ell),\sum_{i\in V_n}(1-\mathsf{F}_i(\ell))) \tag{3.17}$$

$$\geq\; \lambda\min(\sum_{i\in V_n}\mathsf{F}_i(\ell), n\delta/(2+\delta)) \tag{3.18}$$

$$\geq\; \frac{\lambda\delta}{2+\delta}\sum_{i\in V_n}\mathsf{F}_i(\ell)\,. \tag{3.19}$$

Recalling (3.15), (3.16), taking expectation of both sides with respect to $\mu_{n,+}$ and representing the sums over E_n, V_n as expectations, we get

$$\frac{k}{2}\varepsilon\geq\frac{k}{2}\,\mathbb{E}\mu_{n(m),+}(\mathbb{I}\{\mathsf{F}_I(\ell)\neq\mathsf{F}_J(\ell)\})\geq\frac{\lambda\delta}{2+\delta}\mathbb{E}\mu_{n(m),+}(\mathsf{F}_I(\ell))\;\geq\frac{\lambda\delta}{2+\delta}(q-\varepsilon). \tag{3.20}$$

Since $\varepsilon>0$ is arbitrary, we derive a contradiction unless $q=0$. The proof follows. □

We can now complete the proof of Theorem 2.4.

Proof (Theorem 2.4, part II). Let $n(m)$ be a subsequence along which the local distributions of $\mu_{n,+}$ converge locally to some $\mathfrak{m}$ (by the same compactness arguments used in the previous section, one always exists). Now by Proposition 3.9 it follows that $\nu_+=\int_{\mathcal{G}_k}\nu\;\mathfrak{m}(\mathrm{d}\nu)$ which implies that $\mathfrak{m}$ is a point measure on ν_+ since it is extremal. This implies local convergence in probability to ν_+, which completes the proof. □

4 Proofs of Lemmas

4.1 Proof of Lemma 3.4

We start from a very general remark, which is implicit in [10] holding for a general Markov random field on a graph $G=(V,E)$

$$\mu(\underline{x})=\frac{1}{Z}\prod_{(i,j)\in E}\psi_{i,j}(x_i,x_j) \tag{4.1}$$

where $\underline{x}=\{x_i\}_{i\in V}\in\mathcal{X}^V$ for a finite spin alphabet $\mathcal{X}$, and $\psi_{ij}:\mathcal{X}\times\mathcal{X}\to\mathbb{R}$ is a collection of potentials. Recall that a subset S of the vertices of G is an independent set if, for any $i,j\in S$, $(i,j)\notin E$.

Lemma 4.1. *Assume $0 < \psi_{\min} \le \psi_{ij}(x_i, x_j) \le \psi_{\max}$, let k be the maximum degree of G, and $I(G)$ the maximum size of an independent set of G. Then there exists a constant $C = C(k, \psi_{\max}/\psi_{\min}) > 0$ such that, for any $x \in \mathcal{X}$ and any $\ell \in \mathbb{N}$,*

$$\mu\Big(\sum_{i\in V} \mathbb{I}_{x_i=x} = \ell\Big) \le \frac{C}{\sqrt{I(G)}}\,. \tag{4.2}$$

Proof. Let S be a maximum size independent set and $S^{\mathrm{c}} = V \setminus S$ its complement. Further, let $Y_U \equiv \sum_{i\in U} \mathbb{I}_{x_i=x}$ for $U \subseteq V$. Conditioning on $\underline{x}_{S^{\mathrm{c}}} = \{x_i : i \in S^{\mathrm{c}}\}$

$$\mu\Big(\sum_{i\in V} \mathbb{I}_{x_i=x} = \ell\Big) = \mathbb{E}_\mu\Big\{\mu\Big(Y_S = \ell - Y_{S^{\mathrm{c}}} | \underline{x}_{S^{\mathrm{c}}}\Big)\Big\}\,. \tag{4.3}$$

Conditional on $\underline{x}_{S^{\mathrm{c}}}$, the variables $\{x_i\}_{i\in S}$ are independent with $\delta \le \mu(x_i = x|\underline{x}_{S^{\mathrm{c}}}) \le 1-\delta$ for some $\delta > 0$ depending on k and $\psi_{\max}/\psi_{\min}$. As a consequence Y_S is the sum of $|S| = I(G)$ independent Bernoulli random variables with expectation bounded away from 0 and 1. By the Berry-Esseen Theorem

$$\mu\Big(Y_S = \ell - Y_{S^{\mathrm{c}}} | \underline{x}_{S^{\mathrm{c}}}\Big) \le \frac{C}{\sqrt{I(G)}}, \tag{4.4}$$

which implies the thesis. □

Proof. (Lemma 3.4) Recall that for T_k, the infinite rooted k-regular tree, we denote by $\mathsf{T}_k(t)$ the subtree induced by nodes with distance at most t from the root $\emptyset$. Also, denote $\mathsf{T}_k(t, t_+) = \mathsf{T}_k(t_+) \setminus \mathsf{T}_k(t)$, the subgraph induced by nodes i with distance $t+1 \le d(i,\emptyset) \le t_+$. Let $\bar{\nu}_+$ denote a subsequential limit of the measures $\mu_{n,+}$ constructed as in Section 2.4. For any $t \ge 1$ and $t_+ > t$ we will prove that the conditional distribution of $\underline{x}_{\mathsf{T}_k(t)}$ given $\underline{x}_{\mathsf{T}_k(t,t_+)}$ is given by (here and below we adopt the convention of writing $p(x|y) \cong f(x,y)$ for a conditional distribution p, whenever $p(x|y) = f(x,y)/\sum_{x'} f(x',y)$):

$$\bar{\nu}_+^{\mathsf{T}_k(t)|\mathsf{T}_k(t,t+)}(\underline{x}_{\mathsf{T}_k(t)}|\underline{x}_{\mathsf{T}_k(t,t_+)}) \cong \exp\left\{\beta \sum_{(i,j)\in E(\mathsf{T}_k(t+1))} x_i x_j\right\}\,. \tag{4.5}$$

This establishes the DLR conditions and implies that $\bar{\nu}_+$ is a Gibbs measure as required.

In analogy with the notation introduced above (and recalling that $\mathsf{B}_i(t)$ is the ball of radius t around vertex i in G_n), we let $\mathsf{B}_i(t,t_+) = \mathsf{B}_i(t_+) \setminus \mathsf{B}_i(t)$ be the subgraph induced by vertices j such that $t+1 \le d(i,j) \le t_+$. Also $\mathsf{E}_i(t)$ will be the set of edges in $\mathsf{B}_i(t)$, and $\mathsf{E}_i^{\mathrm{c}}(t) = E_n \setminus \mathsf{E}_i^{\mathrm{c}}(t)$. The marginal distribution of $\underline{x}_{\mathsf{B}_i(t_+)}$ under $\mu_{n,+}$ is given by

$$\mu_{n,+}^{t_+}(\underline{x}_{\mathsf{B}_i(t_+)}) \cong F_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)})\, Z_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)}) \tag{4.6}$$

$$F_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)}) \equiv \exp\Big\{\beta \sum_{(l,j)\in\mathsf{E}_i(t_+)} x_l x_j\Big\}\,, \tag{4.7}$$

$$Z_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)}) \equiv \sum_{\underline{x}_{V_n\setminus\mathsf{B}_i(t_+)}} \exp\Big\{\beta \sum_{(l,j)\in\mathsf{E}_i^{\mathrm{c}}(t_+)} x_l x_j\Big\}\, \mathbb{I}\Big(\sum_{j\in\mathsf{B}_i^{\mathrm{c}}(t_+)} x_j > -\sum_{j\in\mathsf{B}_i(t_+)} x_j\Big)\,. \tag{4.8}$$

We, therefore, have the following expression for the conditional distribution of $\underline{x}_{\mathsf{B}_i(t)}$, given $\underline{x}_{\mathsf{B}_i(t,t_+)}$:

$$\mu_{n,+}^{\mathsf{B}_i(t_+)|\mathsf{B}_i(t,t_+)}(\underline{x}_{\mathsf{B}_i(t_+)}|\underline{x}_{\mathsf{B}_i(t,t_+)}) = \frac{F_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)})\, Z_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)})}{\sum_{\underline{x}_{\mathsf{B}_i(t)}} F_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)})\, Z_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)})}\,. \tag{4.9}$$

On the other hand we have $Z^-_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)}) \le Z_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)}) \le Z^+_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)})$ where we define

$$Z^{\pm}_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)}) \equiv \sum_{\underline{x}_{V_n\setminus \mathsf{B}_i(t_+)}} \exp\Big\{\beta \sum_{(l,j)\in \mathsf{E}^c_i(t_+)} x_i x_j\Big\}\, \mathbb{I}\Big(\sum_{j\in \mathsf{B}^c_i(t_+)} x_j > \mp|\mathsf{B}_i(t_+)|\Big). \tag{4.10}$$

Notice that $Z^{\pm}_{\mathsf{B}_i(t_+)}(\underline{x}_{\mathsf{B}_i(t_+)})$ depend on $\underline{x}_{\mathsf{B}_i(t_+)}$ only through $\underline{x}_{\mathsf{B}_i(t,t_+)}$. Using the expression (4.9) for the conditional probability (and dropping subscripts on μ to lighten the notation), we have

$$\mu_{n,+}(\underline{x}_{\mathsf{B}_i(t_+)}|\underline{x}_{\mathsf{B}_i(t,t_+)}) \le \mu^*(\underline{x}_{\mathsf{B}_i(t_+)}|\underline{x}_{\mathsf{B}_i(t,t_+)}) \max_{\underline{x}\in\{+1,-1\}^{\mathsf{T}_k(t,t_+)}} \frac{Z^+_{\mathsf{B}_i(t_+)}(\underline{x})}{Z^-_{\mathsf{B}_i(t_+)}(\underline{x})}, \tag{4.11}$$

$$\mu_{n,+}(\underline{x}_{\mathsf{B}_i(t_+)}|\underline{x}_{\mathsf{B}_i(t,t_+)}) \ge \mu^*(\underline{x}_{\mathsf{B}_i(t_+)}|\underline{x}_{\mathsf{B}_i(t,t_+)}) \min_{\underline{x}\in\{+1,-1\}^{\mathsf{T}_k(t,t_+)}} \frac{Z^-_{\mathsf{B}_i(t_+)}(\underline{x})}{Z^+_{\mathsf{B}_i(t_+)}(\underline{x})}, \tag{4.12}$$

with

$$\mu^*(\underline{x}_{\mathsf{B}_i(t_+)}|\underline{x}_{\mathsf{B}_i(t,t_+)}) \cong \exp\Big\{\beta \sum_{(l,j)\in \mathsf{E}_i(t+1)} x_l x_j\Big\}. \tag{4.13}$$

The claim (4.5) thus follows from the fact that $\mathsf{B}_i(t_+) \simeq \mathsf{T}_k(t_+)$ with probability going to 1 as $n\to\infty$, if we can show that

$$\frac{Z^-_{\mathsf{B}_i(t_+)}(\underline{x})}{Z^+_{\mathsf{B}_i(t_+)}(\underline{x})} \to 1 \tag{4.14}$$

for all $\underline{x}\in\{+1,-1\}^{\mathsf{T}_k(t,t_+)}$ as $n\to\infty$.

Let $\hat\mu$ denote the Ising measure on $\underline{x}_{\mathsf{B}^c_i(t_+)}$ with boundary conditions $\underline{x}_{\mathsf{B}_i(t_+)}$

$$\hat\mu(\underline{x}_{\mathsf{B}^c_i(t_+)}) = \frac{1}{\hat Z(\underline{x}_{\mathsf{B}_i(t_+)})} \exp\Big\{\beta \sum_{(l,j)\in \mathsf{E}^c_i(t_+)} x_i x_j\Big\}. \tag{4.15}$$

Now

$$1 - \frac{Z^-_{\mathsf{B}_i(t_+)}(\underline{x})}{Z^+_{\mathsf{B}_i(t_+)}(\underline{x})} = \frac{\hat\mu\Big(\sum_{j\in\mathsf{B}^c_i(t_+)} x_j > -|\mathsf{B}_i(t_+)|\Big) - \hat\mu\Big(\sum_{j\in\mathsf{B}^c_i(t_+)} x_j > |\mathsf{B}_i(t_+)|\Big)}{\hat\mu\Big(\sum_{j\in\mathsf{B}^c_i(t_+)} x_j > -|\mathsf{B}_i(t_+)|\Big)}.$$

Observe that by the Gibbs construction of μ for any $\underline{x}_{\mathsf{B}^c_i(t_+)}$, we have that

$$\hat\mu(\underline{x}_{\mathsf{B}^c_i(t_+)}) \ge \exp(-2\beta k|\mathsf{B}_i(t^+)|)\mu_n(\underline{x}_{\mathsf{B}^c_i(t_+)})$$

as this is the maximum affect that conditioning on a set of size $|\mathsf{B}_i(t^+)|$ can have on the measure μ. By symmetry of the measure μ_n with respect to the sign of $\underline{x}$,

$$\begin{aligned}\hat\mu\Big(\sum_{j\in\mathsf{B}^c_i(t_+)} x_j > -|\mathsf{B}_i(t_+)|\Big) &\ge \hat\mu\Big(\sum_{j\in\mathsf{B}^c_i(t_+)} x_j \ge 0\Big)\\ &\ge \exp(-2\beta k|\mathsf{B}_i(t^+)|)\mu_n\Big(\sum_{j\in\mathsf{B}^c_i(t_+)} x_j \ge 0\Big)\\ &\ge \frac{1}{2}\exp(-2\beta k|\mathsf{B}_i(t^+)|).\end{aligned} \tag{4.16}$$

Now applying Lemma 4.1 to the measure $\hat{\mu}$ we have that

$$\hat{\mu}\Big(\sum_{j\in\mathsf{B}_i^c(t_+)} x_j > -|\mathsf{B}_i(t_+)|\Big) - \hat{\mu}\Big(\sum_{j\in\mathsf{B}_i^c(t_+)} x_j > |\mathsf{B}_i(t_+)|\Big) = \hat{\mu}\Big(\Big|\sum_{j\in\mathsf{B}_i^c(t_+)} x_j\Big| \le |\mathsf{B}_i(t_+)|\Big)$$
$$\le \frac{2C|\mathsf{B}_i(t_+)|}{\sqrt{n-|\mathsf{B}_i(t_+)|}} \to 0 \tag{4.17}$$

for some $C = C(k,\beta)$ as $n \to \infty$. Combining equations (4.17) and (4.16) we establish equation (4.14) which completes the proof. □

4.2 Proof of Lemma 3.1

For the convenience of the reader, we restate the main result of [9] in the case of k-regular graphs, with no magnetic field B. This provides an asymptotic estimate of the partition function

$$Z_n(\beta) = \sum_{\underline{x}} \exp\Big\{\beta \sum_{(i,j)\in E} x_i x_j + \sum_{i\in V} x_i\Big\}. \tag{4.18}$$

Theorem 4.2. *Let $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs that converges locally to the k-regular tree T_k. For $\beta > 0$, let h be the largest solution of*

$$h = (k-1)\tanh[\tanh(\beta)\tanh(h)]. \tag{4.19}$$

Then $\lim_{n\to\infty} \frac{1}{n}\log Z_n = \phi(\beta)$, where

$$\begin{aligned}\phi(\beta) &\equiv \frac{k}{2}\log\cosh(\beta) - \frac{k}{2}\log\{1+\tanh(\beta)\tanh(h)^2\}\\ &+ \log\Big\{[1+\tanh(\beta)\tanh(h)]^k + [1-\tanh(\beta)\tanh(h)]^k\Big\},\end{aligned} \tag{4.20}$$

For the proof of Lemma 3.1 we start by noticing that, by symmetry under change of sign of the x_i's, we have $\mu_{n,+}(x_i\cdot x_j) = \mu_n(x_i\cdot x_j)$. Simple calculus yields

$$\frac{1}{n}\frac{\partial}{\partial\beta}\log Z_n(\beta) = \frac{1}{n}\sum_{(i,j)\in\mathbb{E}_n} \mu_n(x_i\cdot x_j) = \frac{k}{2}\,\mathbb{E}\mu_n(x_I\cdot x_J), \tag{4.21}$$

where the expectation $\mathbb{E}$ is taken with respect to I uniformly random vertex, and J one of its neighbors taken uniformly at random.

On the other hand, differentiating Eq. (4.20) with respect to β, and using the fixed point condition (4.19), we get after some algebraic manipulations

$$\frac{\partial}{\partial\beta}\phi(\beta) = \frac{k}{2}\,\frac{\tanh\beta + (\tanh h)^2}{1+\tanh\beta(\tanh h)^2} = \frac{k}{2}\nu_+(x_\emptyset\cdot x_1). \tag{4.22}$$

The last identification comes from the fact that the joint distribution of $x_\emptyset$ and x_1 on a k-regular tree under the plus-boundary Gibbs measure is $\nu_+(x_\emptyset, x_1) \propto \exp\{\beta x_\emptyset x_1 + h x_\emptyset + h x_1\}$ (see [9]).

Further $\beta \mapsto \frac{1}{n}\log Z_n(\beta)$ is convex because its second derivative is proportional to the variance of $\sum_{(i,j)} x_i x_j$ with respect to the measure μ_n. Therefore, its derivative $(k/2)\mathbb{E}\mu_n(x_i\cdot x_j)$ converges to $(k/2)\nu_+(x_\emptyset\cdot x_1)$ for a dense subset of values of β. Since the limit $\beta \mapsto \nu_+(x_\emptyset\cdot x_1)$ is continuous, convergence takes place for every β.

4.3 Proof of Lemma 3.2

Recalling that T_k denotes the infinite k-regular tree rooted at $\emptyset$ let $\mathsf{T}^\emptyset$ and T^1 be the subtrees obtained by removing the edge $(\emptyset, 1)$ where 1 is a neighbor of $\emptyset$. It is sufficient to prove the claim when ν is an extremal Gibbs measure on T_k since of course we may decompose any Gibbs measure into a mixture of extremal measures. For $i \in \{\emptyset, 1\}$ define

$$m_i^\nu = \lim_{\ell\to\infty} \mathbb{E}_{\mathsf{T}^i}(x_i \mid \underline{x}_{\mathsf{B}_i^c(\ell)\cap\mathsf{T}^i})$$

where $\mathbb{E}_{\mathsf{T}^i}$ denotes expectation with respect to the Ising model on the tree T_i and the boundary condition $\underline{x}_{\mathsf{B}_i(\ell)\cap\mathsf{T}^i}$ is chosen according to ν. The limit exists by the Backward Martingale Convergence Theorem. Further it is a constant almost surely, because it is measurable with respect to the tail σ-field, and ν is extremal.

By the monotonicity of the Ising model if $\nu \preceq \nu'$, then $m_i^\nu \le m_i^{\nu'}$. Furthermore

$$\nu(x_\emptyset) = \frac{m_\emptyset^\nu + \tanh(\beta) m_1^\nu}{1 + \tanh(\beta) m_\emptyset^\nu m_1^\nu}\,. \tag{4.23}$$

Now if $\nu \neq \nu^+$ then $\nu(x_\emptyset = 1) < \nu^+(x_\emptyset = 1)$. Under the plus measure $m_\emptyset^{\nu+} = m_1^{\nu+} = m^+$ which by the monotonicity of the system is the maximal such value. Since the right hand side of Eq. (4.23) is increasing in $m_\emptyset, m_1$ it follows that $m_\emptyset^\nu = m_1^\nu = m^+$ if and only if $\nu = \nu^+$.

An easy tree calculation shows that the expectation of $x_\emptyset \cdot x_1$ is

$$\nu(x_\emptyset \cdot x_1) = \frac{\tanh(\beta) + m_\emptyset^\nu m_1^\nu}{1 + \tanh(\beta) m_\emptyset^\nu m_1^\nu}\,.$$

which is strictly increasing in $m_\emptyset^\nu$ when $m_1^\nu > 0$. By symmetry it is also strictly increasing in m_1^ν when $m_\emptyset^\nu > 0$. Hence amongst measures ν with $m_\emptyset^\nu \ge 0$, the expectation $\nu(x_\emptyset \cdot x_1)$ is uniquely maximized when $m_\emptyset^\nu = m_1^\nu = m^+$, that is when $\nu = \nu^+$. Similarly amongst measures ν with $m_\emptyset^\nu \le 0$ the agreement probability is uniquely maximized by ν_-, which completes the proof.

4.4 Proof of Lemma 3.6

Observe first by the local weak convergence of the graphs $\{G_n\}$ that all but $o(n)$ vertices appear in $|\mathsf{B}_i(\ell)|$ balls $\mathsf{B}_i(\ell)$. Hence given a configuration $\underline{x}$ with $\sum_i x_i \ge 0$, we have

$$\sum_{i\in V_n}\left(\frac{1}{|\mathsf{B}_i(\ell)|}\sum_{j\in\mathsf{B}_i(\ell)} x_j\right) \ge -o(n)\,. \tag{4.24}$$

By Markov's inequality (applied to the uniform choice of $i \in V_n$) we have

$$\frac{1}{n}\sum_{i\in V_n}\mathsf{F}_i(\ell) \le \frac{1}{1+\delta} + o_n(1) \le \frac{1}{1+\delta/2}\,, \tag{4.25}$$

where the second inequality holds for all n large enough

4.5 Proof of Lemma 3.8

Setting $\rho = \nu_+(x_\emptyset)$ note that by invariance of ν_+ under graph homomorphisms of T_k, we have

$$\nu_+\Big(\sum_{j\in\mathsf{B}_i(\ell)} x_j\Big) = \rho\,|\mathsf{B}_i(\ell)|\,.$$

Moreover, under ν_+, along any path of vertices in T_k the states are distributed as a 2-state homogenous Markov chain and hence

$$\nu_+(x_j \cdot x_{j'}) - \nu_+(x_j)\nu_+(x'_j) = A\, b^{d(j,j')}$$

where $d(j,j')$ is the graph distance between vertices i and j, and $b \in (0,1)$ is a constant depending on β.

This in particular implies that

$$\mathrm{Var}_{\nu_+}\Big(\sum_{j\in \mathsf{B}_i(\ell)} x_j\Big) = o\Big(\Big|\mathsf{B}_i(\ell)\Big|^2\Big)\,,$$

and therefore, using Chebychev inequality, $\frac{1}{\mathsf{B}_i(\ell)}\sum_{j\in\mathsf{B}_i(\ell)} x_j$ converges in probability to ρ as $\ell\to\infty$. Similarly under the measure ν_- we have that $\frac{1}{\mathsf{B}_i(\ell)}\sum_{j\in\mathsf{B}_i(\ell)} x_j$ converges in probability to $-\rho$. Now taking $0<\delta<\rho$ we have that

$$\lim_{\ell\to\infty} \nu_+(\mathsf{F}_\emptyset(\ell)=1) = 0\,,$$
$$\lim_{\ell\to\infty} \nu_-(\mathsf{F}_\emptyset(\ell)=1) = 1\,.$$

Therefore, for $\nu = (1-q)\nu_+ + q\nu_-$, we have $\nu_+(\mathsf{F}_\emptyset(\ell)=1)\to q$.

Moreover, by translation invariance

$$\nu_+(\mathsf{F}_\emptyset(\ell)\neq \mathsf{F}_1(\ell)) = 2\nu_+(\mathsf{F}_\emptyset(\ell)=1,\mathsf{F}_1(\ell)=0) \le 2\nu_+(\mathsf{F}_\emptyset(\ell)=1)\to 0\,.$$

By applying the same argument to ν_-, we deduce that the probability that $\mathsf{F}_\emptyset(\ell)$ and $\mathsf{F}_1(\ell)$ differ goes to 0 under any mixture of ν_+ and ν_-. Since ν is a mixture of ν_+ and ν_- this completes the lemma.

5 Proof of Theorem 2.5

To simplify notation we will write f_i or $f_i(\underline{x})$ for $f_{i,n}(\underline{x}_{\mathsf{B}_i(\ell)})$. We will prove that, denoting by $\mathrm{Var}_{n,+}$, $\mathrm{Cov}_{n,+}$ variance and covariance under $\mu_{n,+}$,

$$\lim_{n\to\infty}\mathrm{Var}_{n,+}\Big(\frac{1}{n}\sum_{i\in V_n} f_i(\underline{x}_{\mathsf{B}_i(\ell)})\Big) = \lim_{n\to\infty}\mathbb{E}_{U_n}\mathrm{Cov}_{n,+}(f_I(\underline{x}_{\mathsf{B}_I(\ell)}), f_L(\underline{x}_{\mathsf{B}_L(\ell)})) = 0\,.$$

Here $\mathbb{E}_{U_n}$ denotes expectation with respect to two independent and uniformly random vertices I, L in V_n. The thesis then follows by Chebyshev inequality.

Since the f_i's are bounded, we have for $r>\ell$,

$$\mathbb{E}_{U_n}\mathrm{Cov}_{n,+}(f_I,f_L) \le \mathbb{P}_{U_n}(d(I,L)\le 2r) + \mathbb{E}_{U_n}\Big\{\mathrm{Cov}_{n,+}(f_I,f_L); d(I,L)>2r\Big\}\,.$$

Since $\{G_n\}_{n\in\mathbb{N}}$ are k-regular, the probability $d(I,L)\le 2r$ vanishes as $n\to\infty$. It therefore suffices to show that

$$\lim_{r\to\infty}\lim_{n\to\infty}\mathbb{E}_{U_n}\Big\{\mathrm{Cov}_{n,+}(f_U,f_V); d(U,V)>2r\Big\} = 0\,.$$

Define

$$\hat{f}_i^+(r)(\underline{x}) = \mathbb{E}_{n,+}\{f(\underline{x}_{\mathsf{B}_i(\ell)})|x_{V_n\setminus \mathsf{B}_i(r)}\}\,,$$

the conditional expectation being taken with respect to $\mu_{n,+}$. Then we have for all i,j that

$$\mathbb{I}(d(i,j)>2r)\,\mathrm{Cov}_{n,+}(f_i,f_j) = \mathbb{I}(d(i,j)>2r)\,\mathrm{Cov}_{n,+}(\hat{f}_i^+(r),f_j) \le \sqrt{\mathrm{Var}_{n,+}(\hat{f}_i^+(r))}$$

and therefore

$$\lim_{r\to\infty}\lim_{n\to\infty}\mathbb{E}_{U_n}\Big\{\mathrm{Cov}_{n,+}(f_I,f_L);d(I,L)>2r\Big\} \quad\le\quad \lim_{r\to\infty}\lim_{n\to\infty}\mathbb{E}_{U_n}\sqrt{\mathrm{Var}_{n,+}(\hat f_I^+(r))} \tag{5.1}$$

$$\le\quad \lim_{r\to\infty}\lim_{n\to\infty}\sqrt{\mathbb{E}_{U_n}\mathrm{Var}_{n,+}(\hat f_U^+(r))}\,. \tag{5.2}$$

Define the modified function

$$\hat f_i(r)(\underline{x}) = \mathbb{E}_n\{f(\underline{x}_{\mathsf{B}_i(\ell)})|x_{V_n\setminus\mathsf{B}_i(r)}\}\,, \tag{5.3}$$

where the expectation is taken with respect to the measure μ_n. Since the latter is a Gibbs measure $\hat f_i(r)$ depends on $\underline{x}$ only through the variables x_j, $j\in\mathsf{B}_i(r)\setminus\mathsf{B}_i(r-1)$. Further $\hat f_i^+(r)$ and $\hat f_i(r)$ differ only if $|\sum_{j\in V_n\setminus\mathsf{B}_i(r)}x_j|\le|\mathsf{B}_i(r)|$. Therefore

$$\begin{aligned}\mathrm{Var}_{n,+}(\hat f_I^+(r)) &\le 2\mathrm{Var}_{n,+}(\hat f_I(r)) + 2\mathrm{Var}_{n,+}(\hat f_I^+(r)-\hat f_I(r))\\ &\le 2\mathrm{Var}_{n,+}(\hat f_I(r)) + 8\mu_{n,+}\Big(\Big|\sum_{j\in V_n\setminus\mathsf{B}_I(r)}x_j\Big|\le|\mathsf{B}_I(r)|\Big)\,.\end{aligned}$$

The last term vanishes as $n\to\infty$ by Lemma 4.1.

We are therefore left with the task of showing that $\lim_{r\to\infty}\lim_{n\to\infty}\mathbb{E}_{U_n}\mathrm{Var}_{n,+}(\hat f_I(r))=0$. For a function $f:\{-1,1\}^{\mathsf{T}_k(\ell)}\to[-1,1]$, let

$$\bar f(r)(\underline{x}) = \mathbb{E}_{\nu_+}\{f(\underline{x}_{\mathsf{T}_k(\ell)})|\underline{x}_{\mathsf{T}_k\setminus\mathsf{T}_k(r)}\}.$$

For all functions whose domain is not $\{-1,1\}^{\mathsf{T}_k(\ell)}$ we let $\bar f(r)=0$ by convention. Also, with an abuse of notation, we define $\bar f_i(r)=\bar g(r)$ for $g=\hat f_i$. Since $\hat f_I(r)$ depends on $\underline{x}$ only through $\underline{x}_{\mathsf{B}_I(r)}$, we obtain by Theorem 2.4 for every $\varepsilon>0$ that

$$\lim_{n\to\infty}\mathbb{E}_{U_n}|\mathrm{Var}_{n,+}(\hat f_I(r))-\mathrm{Var}_{\nu_+}(\bar f_I(r))|\le 2\varepsilon+\lim_{n\to\infty}U_n\left(d_{\mathrm{TV}}\left(\mathbb{P}_n^t(I),\delta_{\mathsf{T}_k(t)}\times\nu_+^t\right)>\varepsilon\right)=2\varepsilon\,,$$

and therefore

$$\lim_{r\to\infty}\lim_{n\to\infty}\mathbb{E}_{U_n}\mathrm{Var}_{n,+}(\hat f_I(r))\le\lim_{r\to\infty}\sup\Big\{\mathrm{Var}_{\nu_+}(\bar f(r))\ \ |\ \ f:\{-1,1\}^{\mathsf{T}_k(\ell)}\to[-1,1]\Big\}\,.$$

By extremality of ν_+, for each $f:\{-1,1\}^{\mathsf{T}_k(\ell)}\to[-1,1]$, $\bar f(r)$ converges to an almost sure constant as $r\to\infty$ and since f is bounded, $\lim_{r\to\infty}\mathrm{Var}_{\nu_+}(\bar f(r))=0$. For each r, the map $f\to\bar f(r)$ is a contraction in L^2 and therefore the map $f\to\sqrt{\mathrm{Var}_{\nu_+}(\bar f(r))}$ is a Lipchitz map with constant 1. Since the set of functions $f:\{-1,1\}^{\mathsf{T}_k(\ell)}\to[-1,1]$ is compact in L_2 and for each f we have $\lim_{r\to\infty}\mathrm{Var}_{\nu_+}(\bar f(r))=0$ we conclude that

$$\lim_{r\to\infty}\sup\Big\{\mathrm{Var}_{\nu_+}(\bar f(r))\ \ |\ \ f:\{-1,1\}^{\mathsf{T}_k(\ell)}\to[-1,1]\Big\}=0,$$

as needed.

Acknowledgements

A.M. was partially supported by by a Terman fellowship, the NSF CAREER award CCF-0743978 and the NSF grant DMS-0806211. E.M. was partially supported by the NSF CAREER award grant DMS-0548249, by DOD ONR grant (N0014-07-1-05-06), by ISF grant 1300/08 and by EU grant PIRG04-GA-2008-239317.

Part of this work was carried out while two of the authors (A.M. and E.M.) were visiting Microsoft Research.

References

[1] R. S. Ellis and C. M. Newman, *The statistics of Curie-Weiss models*, J. Stat. Phys. **19** (1978) 149-161

[2] M. Aizenman, *Translation invariance and instability of phase coexistence in the two-dimensional Ising system*, Comm. Math. Phys. **73** (1980) 8394

[3] R. Dobrushin and S. Shlosman, *The problem of translation invariance of Gibbs states at low temperatures*, Mathematical physics reviews, **5** (1985) 53195

[4] H.O. Georgii and Y. Higuchi, *Percolation and number of phases in the two-dimensional Ising model*, J. Math. Phys. **41** (2000) 11531169

[5] T. Bodineau, *Translation invariant Gibbs states for the Ising model*, Pr. Th. Rel. Fields **135** (2006), 153-168

[6] M. Aizenman and J. Wehr, *Rounding of first-order phase transitions in systems with quenched disorder*, Communications in Mathematical Physics, **130**, (1990) 489-530.

[7] D. Aldous and J. M. Steele, *The objective method: probabilistic combinatorial optimization and local weak convergence*, in "Probability on discrete structures", H. Kesten Ed., Springer Verlag, New York, 2003.

[8] D. Aldous and R Lyons, *Processes on unimodular random networks*, Electron. J. Probab., **12**, (2007) 1454-1508.

[9] A. Dembo and A. Montanari, *Ising models on locally tree-like graphs*, Ann. Appl. Prob. (2009), in press.

[10] R. L. Dobrushin and B. Tirozzi, *The central limit theorem and the problem of equivalence of ensembles.* Comm. Math. Phys. **54**, 173-192.

[11] H.O. Georgii, "Gibbs measures and phase transitions," Walter de Gruyter, Berlin, 1988.

[12] S. Janson, T. Luczak and A. Ruciński, "Random graphs," John Wiley and Sons, 2000.

[13] C. Külske, *Metastates in disordered mean-field models: random field and Hopfield models*, Journal of Statistical Physics, **88**, (1996), 1257–1293.

[14] T. M. Liggett "Interacting particle systems", Springer Verlag, New York, 1985.

[15] R. Lyons, *Phase transitions on nonamenable graphs* J. Math. Phys. 41, 1099 (2000)

[16] C.M. Newman and D.L. Stein, *Spatial inhomogeneity and thermodynamic chaos*, Physical review letters, **76**, (1996), 4821-4824.

[17] N.C. Wormald, *Models of random regular graphs*, Surveys in combinatorics, 1999 (Canterbury), London Math. Soc. Lecture Note Ser., 1999.