FERROMAGNETIC ISING MEASURES
ON LARGE LOCALLY TREE-LIKE GRAPHS

ANIRBAN BASAK* AND AMIR DEMBO†

Abstract. We consider the ferromagnetic Ising model on a sequence of graphs \( G_n \) converging locally weakly to a rooted random tree. Generalizing \[32\], under an appropriate “continuity” property, we show that the Ising measures on these graphs converge locally weakly to a measure, which is obtained by first picking a random tree, and then the symmetric mixture of Ising measures with + and − boundary conditions on that tree. Under the extra assumptions that \( G_n \) are edge-expanders, we show that the local weak limit of the Ising measures conditioned on positive magnetization, is the Ising measure with + boundary condition on the limiting tree. The “continuity” property holds except possibly for countably many choices of \( \beta \), which for limiting trees of minimum degree at least three, are all within certain explicitly specified compact interval. We further show the edge-expander property for (most of) the configuration model graphs corresponding to limiting (multi-type) Galton Watson trees.

1. Introduction

The ferromagnetic Ising model on a finite undirected graph \( G_n = (V_n, E_n) \), is the probability distribution over \( \mathbb{R} = \{x_i : i \in V_n\} \) with \( x_i \in \{-1, +1\} \), for some \( \beta \geq 0 \) (inverse temperature parameter), \( B \in \mathbb{R} \) (external magnetic field), given by

\[
\nu^\beta, B_n (x) = \frac{1}{Z_n(\beta, B)} \exp \left\{ \beta \sum_{(i,j) \in E_n} x_i x_j + B \sum_{i \in V_n} x_i \right\},
\]

where \( Z_n(\beta, B) \) is the normalizing constant (also known as partition function).

Ising model is a paradigm model in statistical physics [35], with much recent interest also in the Ising model on non-lattice complex networks (see \[33\], and the references therein). In this paper we focus on sparse graph sequences \{\( G_n \)\}_{n \in \mathbb{N}} converging locally weakly to (random) trees (see Definition 1.2). The study of statistical physics models on such graphs is motivated by numerous examples from combinatorics, computer science and statistical inference (c.f. \[11, 31\]). The key to such studies is the asymptotics of log partition function, appropriately scaled, as derived for example in \[10, 20, 37\]. In particular, \[12\] shows that for any sequence of graphs (with \( V_n \) of size \( n \)) that converges locally weakly to random trees, the asymptotic free entropy density of the ferromagnetic Ising model exists, i.e.,

\[
\phi(\beta, B) := \lim_{n \to \infty} \phi_n(\beta, B),
\]
where $\phi_n(\beta, B) := \frac{1}{n} \log Z_n(\beta, B)$. Beyond that, perhaps the most interesting feature of the distribution in (1.1) is its “phase transition” phenomenon. Namely, for a wide class of graphs, the Ising measure for large enough $\beta$ and $B = 0$ decomposes into convex combination of well-separated simple components. This has been shown for the complete graph [17], and for grids [1, 9, 14, 19].

In the context of tree-like graphs $G_n$, where the neighborhood of a typical vertex has, for large $n$, approximately the law of the neighborhood of the root of a randomly chosen limiting tree, this picture is only proven for a $k$-regular limit, see Montanari, Mossel and Sly [32]. We show here the universality of this phenomenon, applicable for a general sequence of locally tree-like graphs, including in particular, Erdős-Rényi graphs, random uniform $q$-partite graphs, and random graphs of a given degree distribution. More precisely, one expects that the marginal distribution of $\nu^\beta,B_n(\cdot)$ converges to the marginal distribution on a neighborhood of the root of some Ising Gibbs measure on the limiting tree $T$. Denoting by $\nu^{\beta,B}_{T}$ the Ising Gibbs measures on $T$, corresponding to plus and minus boundary conditions, for $B > 0$ it easily follows from [12] that, the limiting measure is given by first picking the random tree $T$, and then conditioned on $T$, using the Ising Gibbs measure $\nu^{\beta,B}_T$ (the same applies for $B < 0$ with $\nu^{\beta,B}_{T}$ replaced by $\nu^{\beta,B}_{T}$). Recall that for $B = 0$ and $\beta$ large, there are uncountably many Ising Gibbs measures, hence the convergence to a particular Gibbs measure is not at all clear, as is the choice of the correct Gibbs measure. As demonstrated in [32], for $k$-regular trees, the plus/minus boundary conditions play a special role. Indeed, it is shown in [32] that if $G_n$’s converge locally weakly to $k$-regular trees $T = T_k$ then, for any $\beta > 0$ and $B = 0$,

$$\nu^{\beta,0}_n(\cdot) \rightarrow \frac{1}{2} \nu^{\beta,0}_{+,T}(\cdot) + \frac{1}{2} \nu^{\beta,0}_{-,T}(\cdot). \quad (1.3)$$

It is further shown there that, when the graphs $\{G_n\}_{n \in \mathbb{N}}$ are edge-expanders ,

$$\nu^{\beta,0}_{n,\pm}(\cdot) \rightarrow \nu^{\beta,0}_{\pm,T}(\cdot), \quad (1.4)$$

where $\nu^{\beta,0}_{+,T}(\cdot)$ and $\nu^{\beta,0}_{-,T}(\cdot)$ are the measures (1.1) conditioned to, respectively, $\sum_i x_i \geq 0$ and $\sum_i x_i \leq 0$ (when $n$ is odd, see Remark 1.10 on slight modification usually taken for even $n$). The latter sharp result provides a better understanding of $\nu_n(\cdot)$, and is much harder to prove than (1.3).

For genuinely random limiting trees, one expects (1.3) and (1.4) to apply where now $T$ is chosen according to the limiting tree measure. As we focus on the case $B = 0$, hereafter we write $\nu^{\beta}_n(\cdot) := \nu^{\beta,0}_n(\cdot)$ and adopt the convention of using $\nu^{B}_n(\cdot)$ (or just $\nu_n$, in case $B = 0$), when the value of $\beta$ is either arbitrary, or clear from the context. Similar notations apply for Ising measures on the limiting trees.

It is well known (see [27]) that there exists a value of $\beta$, denoted here by $\beta_c$, such that for $\beta < \beta_c$ there is a unique Ising Gibbs measure, and for $\beta > \beta_c$ there are multiple Ising Gibbs measures. In the more interesting case of $\beta \geq \beta_c$, key estimates in the proof of (1.3) and (1.4) in [32], involve explicit calculations which crucially rely on the regularity of both graph sequence, and the limiting tree. Several new ideas are necessary in the absence of such regularity. For example, the key to the proof of (1.3) in [32] is the continuity, for $k$-regular infinite trees, of root magnetization under $\nu^{+}_k(\cdot)$, obtained there out of its representation as the largest zero of a real analytic function. While no such representation is known for any other possible limiting tree measure, in case it a.s. has minimum degree $d_* > 2$, we prove here the continuity of root magnetization under $\nu^{+}_k(\cdot)$ for all $\beta > \text{atanh}((d_* - 1)^{-1}) > \beta_c$ (see Section 5).\footnote{For $\beta = \beta_c$ one may use the equivalent capacity criterion provided in [36].}

The proof of (1.4) relies on choosing functionals $F_i(\cdot)$ of the spin configurations on $G_n$, which approximate the indicator on the vertices that are in “$-$ state”, and whose values concentrate as $n, l \to \infty$. The regularity of the graphs $G_n$, and that of...
their limit, provide for such functionals, and allows explicit computations involving them, both of which fail as soon as we move away from the regular regime. At the level of generality of our setting the only tools are unimodularity of the law of the limiting tree (see Definition 1.3), and properties of simple random walk on it. Hence, a completely different choice of functionals is required here. With $F_l(\cdot)$ defined via average occupation measure of the simple random walk on the tree, we show here that (1.4) holds under the same continuity property, for any edge-expander $G_n^*$'s (see Theorem 1.8). We also confirm the root magnetization continuity property at $\beta = \beta_*$ for Multitype Galton Watson (MGW) trees which arise as the limit of many natural locally tree-like graph ensembles, and show that subject to minimal degree at least 3, the corresponding configuration models are edge-expanders (see Section 5). Thus, our theorem applies for most naturally appearing locally tree-like graphs.

An interesting byproduct of our results is the continuity of percolation probability for Random Cluster Model, with $q = 2$, and wired boundary condition (see [22] for details on rcm, and its connection with Ising model). Another interesting byproduct of this work is the uniqueness of the splitting Gibbs measure (for a definition see [18, Chapter 12]), for large $\beta, B = 0$ and any boundary condition strictly larger than the free boundary condition (see Lemma 1.18 and Remark 1.19). Many of the techniques developed here should extend to more general settings, e.g. the Potts model.

### 1.1. Graph preliminaries and local weak convergence

In a connected undirected graph $G = (V,E)$ the distance between two vertices $v_1$ and $v_2$ is defined to be the length of the shortest path between them. For each vertex $v \in V$, we denote by $B_v(r)$ the ball of radius $r$ around $v$, i.e. the collection of all vertices whose distance from $v$ in $G$ is at most $r$. The set $B_v(1) \setminus \{v\}$ of all vertices adjacent to $v$ is also denoted by $\partial_v$, with $\Delta_v := |\partial_v|$, denoting its size, namely, the degree of $v$ in $G$. A rooted graph $(G,o)$ is a graph $G$ with a specified vertex $o \in V$, called the root, and a rooted network $(\overline{G},o)$ is a rooted graph $(G,o)$ with vector $\overline{\mathcal{X}}_G$ of $\mathcal{X}$-valued marks on each of its vertices (for Ising models $\mathcal{X} = \{-1,1\}$, more generally $\mathcal{X}$ assumed throughout to be a fixed finite set). A rooted isomorphism of rooted graphs (or networks) is a graph isomorphism which maps the root of one to that of another (while preserving the marks in case of networks), with $[G,o]$ denoting the collection of all rooted graphs that are isomorphic to $(G,o)$ (and $[\overline{G},o]$ denoting the collection of all rooted networks isomorphic to $(\overline{G},o)$).

Let $\mathcal{G}_s$ be the space of rooted isomorphism classes of rooted connected locally finite graphs. Similarly, for rooted networks let $\overline{\mathcal{G}}_s$ denote the space of rooted isomorphism classes of rooted connected locally finite networks. Setting the distance between $[G_1,o_1]$ and $[G_2,o_2]$ (and the same between $[\overline{G}_1,o_1]$ and $[\overline{G}_2,o_2]$) to be $1/(\alpha + 1)$, where $\alpha$ is the supremum over $r \in \mathbb{N}$ such that there is a rooted isomorphism of balls of radius $r$ around the roots of $G_i$ (and marks in those balls are same), results with $\mathcal{G}_s$ and $\overline{\mathcal{G}}_s$ which are complete separable metric spaces (see [5, 7]). Using hereafter this metric topology, we denote by $\mathcal{G}_s$ and $\overline{\mathcal{G}}_s$ the corresponding Borel $\sigma$-algebras on $\mathcal{G}_s$ and $\overline{\mathcal{G}}_s$, respectively. Similarly, we equip the spaces $\mathcal{T}_s$ and $\overline{\mathcal{T}}_s$ of all rooted isomorphism classes of locally finite trees (and marked trees, respectively), with the metric topology and Borel $\sigma$-algebra induced by $\mathcal{G}$ and $\overline{\mathcal{G}}_s$, respectively.

**Definition 1.1.** For $\zeta_n$ and $\mu$ Borel probability measures on $\mathcal{G}_s$ (or $\overline{\mathcal{G}}_s$), we write $\zeta_n \Rightarrow \mu$ when $\zeta_n$ converges weakly to $\mu$ with respect to the metric on $\mathcal{G}_s$ (or $\overline{\mathcal{G}}_s$). For any $G \in \mathcal{G}_s$ we denote by $\delta_G$ the probability measure on $\mathcal{G}_s$ assigning point mass at $G$, and for a finite graph $G$ let $U(G)$ be the probability measure on $\mathcal{G}_s$ obtained by choosing a uniform random vertex of $G$ as the root.
We proceed to define the local weak convergence of graphs.

**Definition 1.2.** With $I_n$ chosen uniformly over $[n] := \{1, 2, \ldots, n\}$, we call a sequence of graphs $\{G_n\}_{n \in \mathbb{N}}$ having vertex sets $[n]$ uniformly sparse, if the random variables $\{\Delta_{I_n}(G_n)\}$ are uniformly integrable. That is, if

$$\lim_{t \to \infty} \limsup_{n \to \infty} \frac{1}{n} \sum_{i \in [n]} \Delta_i(G_n) \mathbb{I}(\Delta_i(G_n) \geq t) = 0. \quad (1.5)$$

If in addition $U(G_n) \Rightarrow \mu$, a probability measure on $\mathcal{G}_s$, we say that the uniformly sparse collection $\{G_n\}$ converges locally weakly to $\mu$, denoted by $G_n \xrightarrow{\text{lwc}} \mu$. In particular, due to uniform sparseness $\text{deg}(\mu) := \mathbb{E}_\mu[\Delta_o]$ is finite for any such limit.

Similarly to the space $\mathcal{G}_s$, one defines $\mathcal{G}_{ss}$ as the space of all isomorphism classes of locally finite connected graphs with an ordered pair of distinguished vertices and the corresponding topology thereon, where a function $f$ on $\mathcal{G}_{ss}$ is written as $f(G, x, y)$, to indicate the distinguished pair of vertices $(x, y)$. In [7] it is shown that any LWC limit point must be involution invariant, a property that was found in [4] to be equivalent to the following property of unimodularity.

**Definition 1.3.** A Borel probability measure $\mu$ on $\mathcal{G}_s$ is called unimodular if for any Borel function $f : \mathcal{G}_{ss} \to [0, \infty]$,

$$\int \sum_{x \in V(G)} f(G, o, x) d\mu([G, o]) = \int \sum_{x \in V(G)} f(G, x, o) d\mu([G, o]) \quad (1.6)$$

and we denote by $\mathcal{U}$ the collection of all unimodular probability measures $\mu$ on $\mathcal{G}_s$ for which $\overline{\text{deg}(\mu)}$ is finite.

We consider throughout tree-like graphs, namely $G_n \xrightarrow{\text{lwc}} \mu$ with a limiting object which is a (random) tree, namely having $\mu(\mathcal{T}_o) = 1$. This assumption, and the fact that any LWC limit points is in $\mathcal{U}$ are both key for our results, with (1.6) being utilized in several proofs.

### 1.2. Local weak convergence of Ising measures.

The space of all probability measures on $(\mathcal{G}_s, \mathcal{E}_{ss})$ will be denoted by $\mathcal{P}(\mathcal{G}_s)$. For example, upon choosing a root, Ising measures on connected, locally finite graphs can be considered elements of $\mathcal{P}(\mathcal{G}_s)$, as is the probability measure $\mu \otimes \nu_G$ of $\mathcal{G}_s$ marginal $\mu$, having the conditional distribution $\nu_G$ on the mark space, given any $G \in \mathcal{G}_s$.

For any positive integer $t$, the subgraph $(G, o)(t)$ of $(G, o)$ induced by the vertices $B_o(t)$, is called the graph truncated at height $t$, with the corresponding definition for a rooted network. We further use the notations $G(t)$ and $\mathcal{G}(t)$, when the choice of root is clear from the context. For example, $T(t)$ denotes the first $t$ generations of a tree $T$ (i.e. the subtree induced by the vertices of $T$ of distance at most $t$ from its root). Accordingly, for each $t$ we let $\mathcal{G}_s(t)$ denote the space of rooted isomorphism classes of rooted connected locally finite networks truncated at height $t$, with $\mathcal{E}_{ss}(t)$ the corresponding Borel $\sigma$-algebra, yielding for each $\pi \in \mathcal{P}(\mathcal{G}_s)$ the probability measure $\pi^t$ induced on $(\mathcal{G}_s(t), \mathcal{E}_{ss}(t))$ by such truncation (of the network), and for each probability measure $\overline{\pi}$ on $\mathcal{P}(\mathcal{G}_s)$ the correspondingly induced probability measure $\overline{\pi}^t$ on $\mathcal{P}(\mathcal{G}_s(t), \mathcal{E}_{ss}(t))$.

We next adapt [32, Definition 2.3] to the case of non-deterministic graph limits.

**Definition 1.4.** Given a sequence of graphs $\{G_n\}_{n \in \mathbb{N}}$ having vertex sets $[n]$, and probability measures $\zeta_n$ on $\mathcal{X}^{V_n}$, for any positive integer $t$ let $\mathbb{P}_n(t) \in \mathcal{P}(\mathcal{G}_s(t), \mathcal{E}_{ss}(t))$ denote the law of the pair $(B_i(t), \mathcal{E}_B(t))$ for $x$ drawn according to $\zeta_n$ and $i \in [n]$ some vertex of $G_n$. 

---

A. BASAK AND A. DEMBO
When combined with the uniform measure $U_n$ over the choice of random vertex $I_n \in [n]$, this results with the random distributions $\mathcal{P}_n^\ell(I_n)$, and we say that $\{(G_n, \zeta_n)\}_{n \in \mathbb{N}}$ (or in short $\{\zeta_n\}$), converges locally weakly to a probability measure $\mathbf{m}$ on $\mathcal{P}(\mathbb{G}_\alpha)$, if the law of $\mathcal{P}_n^\ell(I_n)$ converges weakly to $\mathbf{m}^t$, as $n \to \infty$, for each $t \in \mathbb{N}$.

Notions of convergence similar to Definition 1.4, and the weaker form of convergence of Definition 4.1 were studied under the name of metastates for Gibbs measures (see [3, 24, 34]).

We proceed to formally define the relevant limiting Ising Gibbs measures $\nu_{+,T}^{\beta,B,t}$.

**Definition 1.5.** For each $t$, consider the following Ising measures on $T(t)$:

\[
\nu_{+,T}^{\beta,B,t}(x) := \frac{1}{Z_{+,T}} \exp \left\{ \beta \sum_{(i,j) \in E(T(t))} x_i x_j + B \sum_{i \in V(T(t))} x_i \right\} I_{\{\mathbb{Z}^T \setminus T(t-1) = (+)_{\mathbb{T} \setminus (t-1)}\}},
\]

\[
\nu_{-,T}^{\beta,B,t}(x) := \frac{1}{Z_{-,T}} \exp \left\{ \beta \sum_{(i,j) \in E(T(t))} x_i x_j + B \sum_{i \in V(T(t))} x_i \right\} I_{\{\mathbb{Z}^T \setminus T(t-1) = (-)_{\mathbb{T} \setminus (t-1)}\}},
\]

where for any $W \subseteq V(T)$, we denote by $(+)_{W}$ the vector $\{x_i = +1, i \in W\}$, and by $(-)_{W}$ the vector $\{x_i = -1, i \in W\}$, respectively. It is well known that as $t \to \infty$ both $\nu_{+,T}^{\beta,B,t}$ and $\nu_{-,T}^{\beta,B,t}$ converge to probability measures on $\{-1,+1\}^T$, denoted as $\nu_{+,T}^{\beta,B}$ (plus measure) and $\nu_{-,T}^{\beta,B}$ (minus measure), respectively (see [26, Chapter IV]).

For probability measure $\nu$ on $(\mathcal{A}_1, \mathcal{B}_1)$ and measurable map $f : (\mathcal{A}_1, \mathcal{B}_1) \to (\mathcal{A}_2, \mathcal{B}_2)$ we let $\nu \circ f^{-1}$ denote the probability measure on $(\mathcal{A}_2, \mathcal{B}_2)$ such that $\nu \circ f^{-1}(\cdot) = \nu(f^{-1}(\cdot))$, and in case $f$ is real-valued, use the shorthand $\nu(f)$ or $\langle \nu, f \rangle$, for the $\nu$-expected value of $f$ (i.e. $\sum_y \nu(y) f(y)$), using also $\langle f \rangle$ when the choice of $\nu$ is clear from the context. In particular, for any $\beta, B \geq 0$ and $\mu \in \mathcal{U}$ supported on the collection of rooted trees $(T, o) \in \mathcal{T}_\alpha$, let

\[
\mathbb{U}(\beta, B) := \frac{1}{2} \mathbb{E}_{\mu} \left[ \sum_{x \in \partial o} \nu_{+,T}^{\beta,B}(x_o \cdot x_i) \right].
\]  

(1.7)

Our first result generalizes [32, Theorem 2.4.I], namely the limit (1.3), to any limiting measure $\mu$ supported on $\mathcal{T}_\alpha$ subject to a mild continuity assumption on $\mathbb{U}(\cdot, 0)$.

**Theorem 1.6.** Suppose $G_n \overset{\text{LWC}}{\longrightarrow} \mu$ such that $\mu(\mathcal{T}_\alpha) = 1$. Then, at any continuity point $\beta \geq 0$ of the bounded, non-decreasing, right-continuous function $\mathbb{U}(\beta, 0)$, the Ising measures $\nu_{+,T}^{\beta,B}$ on $G_n$ converge locally weakly to $\mathbf{m} = \mu \circ \psi^{-1}$, where $\psi : \mathcal{T}_\alpha \to \mathcal{P}(\mathcal{T}_\alpha)$ with $\psi(T) = \delta_T \otimes (\frac{1}{2} \nu_{+,T}^{\beta} \delta_T + \frac{1}{2} \nu_{-,T}^{\beta})$.

Our generalization of (1.4), namely [32, Theorem 2.4.II], to all limiting tree measures, requires that the graph sequence has certain edge-expansion property related to the following definition.

**Definition 1.7.** A finite graph $G = (V, E)$ is a $(\delta_1, \delta_2, \lambda)$ edge-expander if, for any set of vertices $S \subseteq V$, with $\delta_1 |V| \leq |S| \leq \delta_2 |V|$, we have $|\partial S| \geq \lambda |S|$, where $\partial S := E(S, S^c)$, the edge set between $S$, and $S^c$, and $| \cdot |$ denotes the cardinality of a set.

**Theorem 1.8.** Suppose $\{G_n\}_{n \in \mathbb{N}}$ are $(\delta, 1/2, \lambda_\delta)$ edge-expanders for all $0 < \delta < 1/2$ and some $\lambda_\delta > 0$ (which is independent of $n$). If $G_n \overset{\text{LWC}}{\longrightarrow} \mu$ such that $\mu(\mathcal{T}_\alpha) = 1$. Then, at any continuity point of $\beta \mapsto \mathbb{U}(\beta, 0)$, the measures $\{\nu_{+,T}^{\beta}\}$ converge locally weakly to $\mathbf{m}_+ = \mu \circ \psi_+^{-1}$ where $\psi_+ : \mathcal{T}_\alpha \to \mathcal{P}(\mathcal{T}_\alpha)$ with $\psi_+(T) = \delta_T \otimes \nu_{+,T}^{\beta}$. 


Remark 1.9. Theorems 1.6 and 1.8 apply up to possibly countable set of discontinuity points of \( \beta \mapsto U(\beta,0) \). Note that \( U(\beta,B) \) is uniformly bounded for any \( \mu \in \mathcal{U} \), and while proving Lemma 3.3 we see that it is non-decreasing, right-continuous at all \( \beta, B \geq 0 \), and continuous whenever \( B > 0 \). Further, in proving both theorems, left-continuity of \( U(\beta,0) \) is only required for relating it to the limiting correlation \( \nu_n^{\beta,0}(x_i,x_j) \) across a uniformly chosen edge of \( G_n \) (see Lemma 3.3).

Remark 1.10. With \( B = 0 \), for \( n \) odd the probability measures \( \nu_n^{\beta,0} \) supported on \( \pm \sum_i x_i \geq 0 \) are uniquely determined by the identity \( \nu_n^{\beta} = \frac{1}{2} \nu_n^{\beta,-} + \frac{1}{2} \nu_n^{\beta,+} \). To circumvent non-essential technical issues, one slightly modifies \( \nu_n^{\beta,0} \) in case \( n \) is even to retain this property, as well as having \( \nu_n^{\beta,0}(x) = \nu_n^{\beta,-}(x) \) whenever \( \sum_i x_i = 0 \).

Remark 1.11. Recall the example in [32, §2.3], where it is shown that an expander-like condition is necessary to obtain the convergence of \( \nu_n^{\beta,0} \), even in case of \( k \)-regular tree limits.

1.3. Configuration models and Multi-type Galton-Watson (MGW) trees. We proceed to verify that our results apply for a general class of random graphs from the configuration model, for which the limiting tree follows a MGW distribution, starting with the definition of the configuration model we consider.

Definition 1.12. Fix a strictly positive probability measure \( \theta \) on some finite (type) space \( Q \). Let \( Z_{\geq} \) denotes the set of all non-negative integers and \( Z_{\geq}^{[Q]} := \{ k = (k_1, k_2, \ldots, k_{|Q|}) : k_j \in Z_{\geq}, j = 1, 2, \ldots, |Q| \} \). Consider a (finite) collection of probability measures \( P_k(\cdot), i \in Q \) on \( Z_{\geq}^{[Q]} \), such that for all \( i, j \in Q \),

\[
A(i,j) := \sum_k P_k(\cdot) k_j < \infty, \tag{1.8}
\]

\[
\theta(i) A(i,j) = \theta(j) A(j,i). \tag{1.9}
\]

Now for each \( n \) we define the random graph \( G_n = (V_n, E_n) \) as follows. For every \( i \in Q \) and \( k \in Z_{\geq}^{[Q]} \), we create \( [n\theta(i)P_k(\cdot)] \) many stars with types, such that the end of the star with one vertex has type \( i \in Q \) and the other end consists of \( \sum_j k_j \) vertices, of which exactly \( k_j \) have type \( j \), for each \( j \in Q \).

Edges in a star will be termed as half-edges, and we use the generic notation \((v,e_v)\) to denote a half-edge with \( v \) being the single vertex at one end of the star, and \( e_v \) being one of the vertices present in the other end of the star. The vertex \( v \) here will be called a permanent vertex, whereas the vertices like \( e_v \) will be termed as floating vertices. We denote half-edges \((v,e_v)\) having a permanent end \( v \) of type \( q(v) = i \) and a floating end \( e_v \) of type \( q(e_v) = j \) by \((i,j)\). Due to condition (1.9), if not for the integer truncation effects, for any \( i,j \in Q \) the number of half-edges of type \((i,j)\) would match that of type \((j,i)\). We thus achieve such equality between the numbers of \((i,j)\) and \((j,i)\) half-edges, upon adding to \( G_n \) at most

\[
2 \sum_{i,j} \sum_k [n\theta(i)P_k(\cdot)] k_j
\]

half-edges. This amounts to adding only \( o(n) \) half-edges to the stars (since \( \sum_{i,j} A(i,j) \) is finite, due to (1.8)).

Thereafter for every \( i,j \in Q \) we perform a uniform matching between half-edges with type \((i,j)\) and half-edges with type \((j,i)\). Once we have obtained a matching between these half-edges we throw out
is unimodular). One concrete example is the Erdős-Rényi random graph ensembles of a.s. measure coincides with the usual measure follows along the lines of \( \sum k P(k) \geq 1 \) the configuration model \( \{ (i, j) : A(i, j) > 0 \} \subseteq Q \times Q \) and \( \rho_{i,j} \) for \( (i, j) \in Q_A \), be the probability measures on \( \mathbb{Z}_{\geq 0}^{[Q]} \) given by
\[
\rho_{i,j}(k) = P_i(k + e_j) \frac{k_j + 1}{A(i,j)},
\]
where \( e_j \) denotes the vector with 1 at \( j^{th} \) co-ordinate and 0 elsewhere, and we assume that \( \rho_{i,j}(k) > 0 \) for some \( (i, j) \) and \( \| k \| := \sum_j k_j \neq 1 \) (in the branching processes literature this property is called non-singularity, c.f. [6, pp. 184]).
We assume that the mean matrix \( A_\rho \) for the kernel \( \rho \) over \( Q_A \), which is given by
\[
A_\rho((i_1, j_1), (i_2, j_2)) := \prod_{j_2 = i_1} \sum_k \rho_{i_1,j_1}(k)k_{i_2},
\]
is positive regular. That is, we require that for some finite positive integer \( r \) all entries of \( A_\rho^r \) be strictly positive (possibly infinite, and when multiplying matrices we adopt the convention that \( \infty \times 0 = 0 \)).

The umgw measure on the trees with types is the following: Type of the root is chosen according to \( \theta \), and conditional on the type of the root, say \( \theta \), it’s off-spring number and types are chosen according to \( P_{\theta}(\cdot) \). From the next generation onward, the off-spring numbers and types are chosen independently at each vertex according to \( \rho_{i,j} \) where \( i \) is the type of the current vertex and \( j \) being the type of its parent.

Remark 1.14. In the special case \( |Q| = 1 \), there are no types in the random graphs \( G_n \) of Definition 1.12, neither in the random ugw (umgw) tree of Definition 1.13. The condition (1.9) and positive regularity then trivially hold, while non-singularity and (1.8) amount to having \( P(1) < 1 \) and finite average degree \( \sum_k kP(k) \). In this setting \( G_n \) is the configuration model corresponding to uniformly chosen random graphs subject to given degree distribution \( P(\cdot) \) (c.f. [11, Section 1.2.4]), which is uniformly sparse and converges weakly to the corresponding umgw measure of Definition 1.13 (see [11, Proposition 2.5]). The latter is precisely the ugw tree measure of [4, Example 1.1], and [10, Section 2.1].
In particular, taking \( P(\cdot) \) a Poisson law of parameter \( 2\alpha \), results with \( \rho(k) = P(k) \) (i.e., here the ugw measure coincides with the usual GW law). The configuration model is then closely related to Erdős-Rényi random graph ensembles of \( n^{-1} |E_n| \rightarrow \alpha \) which also have the ugw measure as their a.s. lwc limit (see [11, Proposition 2.6 and Lemma 2.3]).
For \( |Q| > 1 \) the uniform sparseness of \( \{ G_n \} \) of Definition 1.12 is an immediate consequence of finiteness of \( Q \) and \( \sum_{i,j} \theta(i)A(i,j) \), while its local weak convergence to the corresponding umgw measure follows along the lines of [11, Proof of Proposition 2.5] (from the latter convergence we know that each umgw measure of Definition 1.13 is unimodular). One concrete example is the configuration model \( \{ G_n \} \) and umgw for random uniform \( q \)-partite, \( q \geq 2 \), graphs (of \( |\alpha n| \) edges), which fit within our framework upon taking \( \theta \) uniform on \( \{ 1, \ldots, q \} \) and \( P(k) = \prod_{i \neq j} P(k_\ell) \), with \( P(\cdot) \) the Poisson law of parameter \( 2\alpha q/(q - 1) \).
Lemma 1.15. If μ is any of the UMGW measures of Definition 1.13, with minimum degree \( d_\ast > 2 \), one has that \( \beta \mapsto \mathbb{U}(\beta, 0) \) is continuous except for possibly countably many values of \( \beta \in (\beta_c, \beta_\ast] \), where \( \beta_\ast \) = \( \text{atanh}[(d_\ast - 1)^{-1}] \).

Thus, upon applying Theorem 1.6 and Theorem 1.8 we immediately obtain that:

Corollary 1.16. Suppose \( G_n \overset{\text{UMGW}}{\rightarrow} \mu \) with \( \mu \) a UMGW measure as in Definition 1.13, having a.s. minimum degree \( d_\ast > 2 \). Then, except for a possibly countably many values of \( \beta \in (\beta_c, \beta_\ast] \),

(a) \( \nu_n \) converges locally weakly to \( m = \mu \circ \psi^{-1} \), for \( \psi \) as in Theorem 1.6.

(b) If in addition \( \{G_n\}_{n \in \mathbb{N}} \) are \( (\delta, 1/2, \lambda_\delta) \) edge-expanders for all \( 0 < \delta < 1/2 \) and some \( \lambda_\delta > 0 \) (independent of \( n \)), then \( \nu_{n, +} \) converges locally weakly to \( m_+ = \mu \circ \psi_+^{-1} \), for \( \psi_+ \) as in Theorem 1.8.

Examples of expander graphs are abundant in literature. Specifically, it is well-known that a uniformly chosen random \( d \)-regular graph is an expander with probability tending to 1 as its size \( n \to \infty \). Further, the edge-expander requirement of Corollary 1.16(b) holds for the configuration models of Definition 1.12, subject only to uniformly bounded degree and minimal degree at least three. That is,

Lemma 1.17. Suppose (1.9) holds for \( \theta \) strictly positive and \( \{P_i, i \in Q\} \) of bounded support, such that \( P_i(k) = 0 \) whenever \( |k| := \sum_j k_j \leq 2 \). Then, for any \( 0 < \delta < 1/2 \) there exists \( \lambda_\delta > 0 \), such that with probability tending to 1 as \( n \to \infty \), the random graph \( G_n \) of Definition 1.12 is an \( (\delta, 1/2, \lambda_\delta) \) edge-expander.

In particular, Corollary 1.16 holds for such configuration models without the edge-expander assumption.

The following by-product of our proof of Lemma 1.15 is of independent interest.

Lemma 1.18. Fix UGW measure with off-spring distribution \( P \) of finite mean, such that \( P([0, d_\ast)) = 0 \) for some \( d_\ast \geq 3 \). Let the positive integer random variable \( K \) follow its size-biased distribution. For any fixed \( \beta > \beta_c \) consider the recursion over \( t \geq 0 \),

\[
h^{(t+1)} = \sum_{\ell=1}^{K-1} \text{atanh}[\tanh(\beta) \tanh(h^{(t)}_{\ell})],
\]

where \( h^{(t)}_{\ell} \) are i.i.d. copies of \( h^{(t)} \) which are further independent of \( K \). Denote by \( h^{\beta, +} \) its limit in law when \( t \to \infty \) and starting at \( h^{(0)} = \infty \). Then, fixing any \( \beta \geq \beta_0 > \beta_\ast \) and starting this recursion at a stochastically dominating \( h^{(0)} \geq h^{\beta_0, +} \), yields a sequence \( \{h^{(t)}\} \) that converges in law to \( h^{\beta, +} \).

Remark 1.19. Fixing \( \beta > \beta_c \), recall that any Ising Gibbs measure arising out of a fixed point of (1.10) is a splitting Gibbs measure (see [12, Remarks 1.13 and 2.6]). Hence, Lemma 1.18 implies that there is only one Bethe Gibbs measure (see [12, Remark 2.6]), that corresponds to some \( h \geq h^{\beta_0, +} \), \( \beta_0 \in (\beta_\ast, \beta) \), with a similar conclusion for the UMGW measures of Definition 1.13. We expect both Lemma 1.15 and Lemma 1.18 to hold for UGW and UMGW measures at all \( \beta \) (and without a minimum degree assumption). However, the non-regularity of \( T \) under genuinely random UGW and UMGW measures yields for \( \beta \in (\beta_c, \beta_\ast] \) a technical difficulty which we can not overcome (c.f. Remark 5.7).

Outline of the paper.

- As shown in §2, weak convergence of \( U(G_n) \) implies that both \( \{\nu_n\} \) and \( \{\nu_{n, +}\} \) have subsequent local weak limit points (see Lemma 2.1), which subject to uniform sparseness are
supported on the set of Ising Gibbs measures (see Lemma 2.4). Both results neither require an Ising model nor tree-like graphs.

- Relying upon the lwc of $G_n$ to a law $\mu$ supported on $\mathcal{T}$, we find in Lemma 3.3 that at its continuity points $U(\beta, 0)$ is the limit of both the $\nu_n$-expected values and $\nu_{n,+}$-expected values, of certain functionals of $G$. Extending (in Lemma 3.4), the result of [32, Lemma 3.2], we deduce in Lemma 3.8 that the weak limit points of $\S 2$ must be convex combinations of $\nu_{\pm,T}$ and get Theorem 1.6 by the symmetry relation $\nu_n(x) = \nu_n(-x)$.

- In $\S 4$ we prove Theorem 1.8. First we deduce in Lemma 4.3 out of lwc of $G_n$ that the $\nu_{n,+}$-expected values of suitable functionals converge in expectation to the corresponding values for the limiting tree. Then, using in Lemma 4.5 and Lemma 4.6 properties of SRW on trees, the assumed edge-expander condition for $G_n$ eliminates all but one choice for the convex combination of $\nu_{\pm,T}$ (thus proving the theorem).

- In $\S 5$ we deal with continuity of $\beta \mapsto U(\beta, 0)$. Constructing in Lemma 5.4 a suitable sequence of random variables that increases to the root magnetization under $\nu_{+,T}$, we establish in Lemma 5.1 such continuity at any $\beta > \beta_*$. Further, Lemma 1.18 follows upon specializing Lemma 5.4 to the context of UGW measures, and we provide in Lemma 5.2 a capacity criterion for continuity of $\beta \mapsto U(\beta, 0)$ at $\beta = \beta_*$, which we verify for UMGW measures. Lastly, while Lemma 1.17 is well known, for completeness we outline its proof.

Acknowledgements. We thank Allan Sly for suggesting the weighted averages of (4.2), Andrea Montanari for a key idea in the proof of Lemma 5.1, and Yuval Peres for helpful discussions on the proof of Lemma 5.2. We also thank Noga Alon, Russell Lyons and Nike Sun for many helpful conversations. We thank the anonymous referee for his helpful suggestions on improving the presentation of the paper.

2. Convergence to Ising Gibbs Measure

We start with a general lemma about existence of sub-sequential local weak limits (based only on weak convergence of $U(G_n)$ and having marks from a finite set $\mathcal{X}$).

Lemma 2.1. Suppose $U(G_n)$ converges weakly on $\mathcal{G}_*$. Then for any probability measures $\zeta_n$ on $\mathcal{X}^{V_n}$ and any sub-sequence $\{n_m\}_{m \in \mathbb{N}}$ there exists a further sub-sequence $\{n_{m_k}\}_{k \in \mathbb{N}}$ such that $\{\zeta_{n_{m_k}}\}$ converges locally weakly to a limit $\mathfrak{m}$ (which may depend on $\{n_m\}$).

Proof: Fixing $\{\zeta_n\}$ and $t \in \mathbb{N}$, consider the probability measures $B_{I_n}(t)$ assigning to each $G \in \mathcal{G}_*(t)$ the probability

$$B_{I_n}(t)(G) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(B_i(t) \simeq G).$$

The assumed convergence of $\{U(G_n)\}_{n \in \mathbb{N}}$ in $\mathcal{P}(\mathcal{G}_*)$ implies the convergence of $\{B_{I_n}(t)\}$ in $\mathcal{P}(\mathcal{G}_*(t))$, so by Prohorov’s theorem $B_{I_n}(t)$ are uniformly tight. With $\mathcal{G}_*(t)$ a discrete space, any compact subset of $\mathcal{G}_*(t)$ is finite, hence for any $\varepsilon > 0$ we have a finite set $\mathcal{G}_*(t) \subset \mathcal{G}_*(t)$, such that

$$\liminf_{n \to \infty} B_{I_n}(t)(\mathcal{G}_*(t)) \geq 1 - \varepsilon. \quad (2.1)$$

Further, per $G \in \mathcal{G}_*(t)$ the space of marks $\mathcal{X}^G$ is finite, so the set $\mathcal{G}_*(t) := \{(G, \mu_G) : G \in \mathcal{G}_*(t), \mu_G \in \mathcal{X}^G\}$ is also finite, and by Prohorov’s theorem the collection of all probability measures on $\mathcal{G}_*(t)$ is compact. In particular, $\mathcal{M}_*(t) := \{\delta_G \otimes \nu_G : G \in \mathcal{G}_*(t), \nu_G \in \mathcal{P}(\{-1,1\}^G)\}$ is a pre-compact collection of probability measures on $\mathcal{P}(\mathcal{G}_*(t))$. Since $\mathbb{P}_n(I_n) \in \mathcal{M}_*(t)$ with probability $B_{I_n}(\mathcal{G}_*(t))$, it thus follows that for each $t \in \mathbb{N}$, the laws of $\mathbb{P}_n(I_n)$ are uniformly tight, hence relatively compact. Consequently, there exists a diagonal sub-sequence along which the random probability measures
\[ P'_n(I_n) \] converge in law, to say \( \bar{\pi}_t \), simultaneously for all \( t \in \mathbb{N} \). By the obvious embedding of \( \mathcal{G}_s(t) \) within \( \mathcal{G}_s(t+1) \), each \( \bar{\pi}_{t+1} \in \mathcal{P}(\mathcal{G}_s(t+1)) \) induces a marginal probability measure on \( \mathcal{G}_s(t) \), denoted \( \pi_t(\bar{\pi}_{t+1}) \). By definition, \( \pi_t(P'_{n+1}(I_n)) = P'_n(I_n) \) for all \( t, n \in \mathbb{N} \). This implies the relation \( \bar{\pi}_t = \bar{\pi}_{t+1} \circ \pi_t^{-1} \) between the corresponding weak limits. That is, the sequence \( \{\bar{\pi}_t\} \) of probability measures on the Polish spaces \( \mathcal{P}(\mathcal{G}_s(t)) \) is consistent with respect to the projections \( \pi_t \). This completes the proof, since by Kolmogorov’s extension theorem there exists a probability measure \( \bar{\pi} \) on \( \mathcal{P}(\mathcal{G}_s) \) such that \( \bar{\pi}_t = \bar{\pi} \) for all \( t \).

Fixing \( \beta \geq 0 \) and \( B = 0 \), with \( \{\nu_n\}_{n \in \mathbb{N}} \) being Ising Gibbs measures on finite graphs \( G_n \), we wish to identify their sub-sequential limits in terms of Ising Gibbs measures on \( \mathcal{G}_s \), which we define next. First recall that probability measure \( \nu_{G} \) on \( \mathcal{X}^G \) for a fixed infinite graph \( G \in \mathcal{G}_s \) is an Ising Gibbs measure iff \( \nu_{G} \) satisfies the relevant DLR condition. That is, setting \( G(\infty) = G, G(-1) = \emptyset \) and \( G(t, \overline{t}) = G(\overline{t}) \setminus G(t) \) for \( t < \overline{t} \leq \infty \), one requires that for \( \overline{t} = \infty \), any \( t \in \mathbb{N} \) and \( \nu_{G.t} \)-a.e. \( \nu_{G(t,\overline{t})} \),

\[ \nu_G \big( x_{G(t)} \mid x_{G(t,\overline{t})} \big) = \nu \big( x_{G(t)} \mid x_{G(t+1)}, G(t+1) \big), \]

where for any finite \( G' = (V', E') \in \mathcal{G}_s \) and \( W \subseteq V' \),

\[ \exp \left\{ \beta \sum_{(i,j) \in E'} x_i x_j \right\} \]

\[ \nu \big( x_{W} \mid x_{V' \setminus W}, G' \big) := \frac{\sum_{\{x'_{W} x_{V' \setminus W} = x_{V' \setminus W}\}} \exp \left\{ \beta \sum_{(i,j) \in E'} x'_i x'_j \right\} \}
\]

denotes the Ising measure on \( W \), given boundary values at \( V' \setminus W \) (see [18, Chapter 2]).

Next, for any \( r \in \mathbb{N} \), \( t \geq -1 \) and \( \overline{t} = t + 1, \ldots, \infty \), let \( B_{(G, x)}(r) \) denote the union over \( x_{G(t,\overline{t})} \) of balls \( B_{G(x)}(r) \) in \( \mathcal{G}_s \) and consider the corresponding sub-\( \sigma \)-algebras

\[ \mathcal{G}_{G_s(t,\overline{t})} := \sigma \left( B_{(G, x)}(r) \right), \quad (G, x) \in \mathcal{G}_s \),

which are non-decreasing in \( \overline{t} \) and non-increasing in \( t \). In particular, \( \mathcal{G}_{G_s(0,\infty)} = \mathcal{G}_{G_s(-1,0)} \) and \( \mathcal{G}_G = \mathcal{G}_{G_s(\infty, \infty)} \) (as a ball \( B_G(r) \subseteq G_s \) of radius \( r \) and center \( G \) is the \( G_s \)-projection of the union over all \( x_{G(t,\overline{t})} \) of the corresponding balls \( B_{G(x)}(r) \) in \( \mathcal{G}_s \)). Since \( \mathcal{G}_s \) is a Polish space, the regular conditional probability measure \( \nu(\cdot | \mathcal{G}_s) \) is thus well defined for any \( \nu \in \mathcal{P}(\mathcal{G}_s) \) (see [38, §9.2]), and we lift the notion of Ising Gibbs measure to \( \mathcal{P}(\mathcal{G}_s) \), by considering the DLR condition (2.2) with this conditional measure playing the role of \( \nu_G \). In this setting, per \( t \in \mathbb{N} \) what one has in the left-side of (2.2) amounts to the restriction to \( \mathcal{G}_{G_s(t)} \) of the regular conditional probability measure \( \nu(\cdot | \mathcal{G}_{G_s(t,\overline{t})}) \), resulting with the following definition.

**Definition 2.2.** A probability measure \( \nu \in \mathcal{P}(\mathcal{G}_s) \) is called an Ising Gibbs measure, denoted by \( \nu \in \mathcal{I} \), if for any \( t \in \mathbb{N} \), \( \nu \)-a.e.

\[ \nu \big( x_{G(t)} \mid \mathcal{G}_{G_s(t,\overline{t})} \big) = \nu \big( x_{G(t)} \mid x_{G(t+1)}, G(t+1) \big), \]

which we interpret as point-wise identities in the discrete countable space \( \mathcal{G}_s(t+1) \).

**Remark 2.3.** It is easy to verify from (2.4) that \( \mathcal{G}_{G_s(t,\overline{t})} \uparrow \mathcal{G}_{G_s(t,\infty)} \) as \( t \uparrow \infty \). Thus, from Lévy’s upward theorem (applied point-wise on \( \mathcal{G}_s(t+1) \)), we have that \( \nu \in \mathcal{I} \) iff for \( \nu \)-a.e. and any \( t < \overline{t} \in \mathbb{N} \),

\[ \nu \big( x_{G(t)} \mid \mathcal{G}_{G_s(t,\overline{t})} \big) = \nu \big( x_{G(t)} \mid x_{G(t+1)}, G(t+1) \big). \]
We focus hereafter on the subset $\mathcal{I}_s$ of all Ising Gibbs measures of the form $\nu = \delta_G \otimes \nu_G$, with $\nu_G$ being an Ising Gibbs measure for the fixed graph $G \in \mathcal{G}_s$. Denoting by $\mathcal{I}_{(t, \overline{t})}$ those $\nu = \delta_G \otimes \nu_G$ in $\mathcal{P}(\overline{\mathcal{G}}_s)$ with $\nu_G$ satisfying (2.2) per fixed $t < \overline{t}$ finite, we see that

$$\mathcal{I}_s = \bigcap_{t < \overline{t}} \mathcal{I}_{(t, \overline{t})}. \tag{2.7}$$

Further, since $\overline{\mathcal{G}}_s(\overline{t})$ is a discrete countable space, $\mathcal{G}_{\overline{t}}(\overline{t})$ being a subset of its Borel $\sigma$-algebra, is countably generated and the collection $\mathcal{I}_{(t, \overline{t})}$ is completely determined in terms of the marginals $\nu^t$ of probability measures $\nu$ on $\mathcal{G}_s$. For that reason we hereafter take the liberty of using $\mathcal{I}_{(t, \overline{t})}$ also for the subset of $\mathcal{P}(\overline{\mathcal{G}}_s(\overline{t}))$ consisting of the corresponding collection of marginals $\nu^t$.

Considering (2.2) at fixed $\overline{t} > t$ for $\nu_n$ and $\nu_{n,+}$, we next characterize the sub-sequential local weak limits of $\{\nu_n\}$ and $\{\nu_{n,+}\}$ in terms of certain Ising Gibbs measures.

**Lemma 2.4.** Suppose $U(G_n)$ converges to $\mu$ weakly on $\mathcal{G}_s$. Then,

(a) Any sub-sequential local weak limit $\tilde{m}$ of $\{\nu_n\}$ is supported on the collection $\mathcal{I}_s$ of Ising Gibbs measures and restricted to $\mathcal{P}(\mathcal{G}_s)$ it has the marginal $\tilde{m} = \mu \circ \tilde{\psi}^{-1}$, where $\tilde{\psi}(G) = \delta_G$ for any $G \in \mathcal{G}_s$.

(b) The same holds for sub-sequential limits $\tilde{m}_+$ of $\{\nu_{n,+}\}$, provided $\{G_n\}$ is uniformly sparse.

**Proof:** Fix a sub-sequence $n_m$ along which $\{\nu_n\}$ (or $\{\nu_{n,+}\}$), converges locally weakly to some $\tilde{m}$. Then, for each $t \in \mathbb{N}$ the $\mathcal{P}(\mathcal{G}_s(t))$-restriction $\mathbb{P}_n^t(I_n)$ of $\mathbb{P}_n(I_n)$ converges in law to $\tilde{m}^t$. Thus, for any fixed $G \in \mathcal{G}_s$,

$$\tilde{m}^t(\delta_G(t)) = \lim_{m \to \infty} \frac{1}{n_m} \sum_{i=1}^{n_m} \mathbb{P}_n(B_i(t) = \delta_G(t)) = \lim_{m \to \infty} \frac{1}{n_m} \sum_{i=1}^{n_m} \mathbb{P}_n(B_i(t) \simeq G(t)) = \mu^t(G(t)),$$

where, denoting by $\mu^t$ the probability measure on $\mathcal{G}_s(t)$ induced by $\mu$, the last equality follows from the weak convergence of $U(G_n)$ to $\mu$ in $\mathcal{G}_s$. Thus, for any $t \in \mathbb{N}$ the measure $\tilde{m}^t$ is supported on the set of atomic measures $\{\delta_G(t) : G \in \mathcal{G}_s\}$ and coincides with $\mu^t \circ \tilde{\psi}^{-1}$. Since any probability measure $\tilde{m}$ on $\mathcal{P}(\mathcal{G}_s)$ is uniquely determined by the collection $\{\tilde{m}^t : t \in \mathbb{N}\}$, we conclude that $\tilde{m} = \mu \circ \tilde{\psi}^{-1}$. As for proving that $\tilde{m} \in \mathcal{I}_s$, in view of (2.7) it suffices to show that for any finite $\overline{t} > t$,

$$\mathbb{P}^t(\mathcal{I}_{(t, \overline{t})}) = 1. \tag{2.8}$$

(a) Considering first the measures $\{\nu_n\}$, recall Definition 1.4 that $\mathbb{P}_n^t(I_n)$ is supported for each $n$ on the collection $\{\delta_{B_i(t)} \otimes \nu_{n,B_i(t)} : i \in [n]\}$, where the restriction $\nu_{n,B_i(t)}$ to $B_i(\overline{t})$ of the Ising Gibbs measure $\nu_n$, is also an Ising Gibbs measure. Next, per $\varepsilon > 0$ recall the finite set of graphs $\mathcal{G}_e(\overline{t} + 1)$ we defined while proving Lemma 2.1, and let $\mathcal{G}_e^+(\overline{t}) := \{G(\overline{t}) : G \in \mathcal{G}_e(\overline{t} + 1)\}$, denote the corresponding collection of one generation truncations. Based on it, define for each $\delta \in [0, 1]$,

$$\mathcal{I}^\delta_{(t, \overline{t})} := \{\delta_G \otimes \nu_G : G \in \mathcal{G}_e^+(\overline{t}), 1 - \delta \leq \frac{\nu_G(\mathcal{X}_G(t) | \mathcal{X}_G(\overline{t}))}{\nu(\mathcal{X}_G(t) | \mathcal{X}_G(\overline{t} + 1))} \leq \frac{1}{1 - \delta}\} \tag{2.9}$$

a closed subset of $\mathcal{P}(\overline{\mathcal{G}}_s(\overline{t}))$. Now, if $B_i(\overline{t} + 1) \simeq G$ for some $G \in \mathcal{G}_e(\overline{t} + 1)$, then

$$\nu_{n,B_i(t)}(\mathcal{X}_{B_i(t)} | \mathcal{X}_{B_i(\overline{t})}) = \nu(\mathcal{X}_G(t) | \mathcal{X}_G(\overline{t} + 1))$$

and consequently $\mathbb{P}_n^t(I_n) \subset \mathcal{I}^\varepsilon_{(t, \overline{t})}$. Clearly, for any $\varepsilon > 0$ fixed, $\mathcal{I}^\varepsilon_{(t, \overline{t})}$ is a subset of $\mathcal{I}_{(t, \overline{t})}$, hence from (2.1) and the assumed local weak convergence along the sub-sequence $n_m$, we deduce that

$$1 - \varepsilon \leq \limsup_{m \to \infty} U_n(\mathbb{P}_{n_m}^t(I_{n_m} \in \mathcal{I}^\varepsilon_{(t, \overline{t})} | \mathcal{I}^\varepsilon_{(t, \overline{t})}) \leq \mathbb{P}^t(\mathcal{I}^\varepsilon_{(t, \overline{t})}) \leq \mathbb{P}^t(\mathcal{I}_{(t, \overline{t})}). \tag{2.10}$$
Upon considering $\varepsilon \downarrow 0$, we conclude that (2.8) holds in this case.

(b). For odd $n \in \mathbb{N}$ and $i \in [n]$, let

$$Z_{n,i}^{0,t} = Z_{n,i}^{0,t}(\overline{t}, B_i(\overline{t}), \overline{x}_{B_i(\overline{t})}) := \nu_n\left( \sum_{j=1}^{n} x_j \geq 0 \mid \overline{x}_{B_i(\overline{t})} \right),$$

adopting also the notation $Z_{n,i}^0 := Z_{n,i}^{0,-1}$. While due to conditioning on $\{\sum_j x_j \geq 0\}$ the measures $\nu_{n,+}$ are not Ising Gibbs measures, it is not hard to verify that for any $i \in [n]$ and finite $\overline{t} > t$,

$$\nu_{n,+}\left( \overline{x}_{B_i(t)} \mid \overline{x}_{B_i(\overline{t})} \right) = \frac{Z_{n,i}^0}{Z_{n,i}^{0,t}} \nu_{n,B_i(t)}\left( \overline{x}_{B_i(t)} \mid \overline{x}_{B_i(\overline{t})} \right)$$

(for clarity of presentation we ignore the slight modification of $Z_{n,i}^{0,t}$ which is required for $n$ even, in accordance with Remark 1.10). Fixing $\varepsilon, \delta > 0$, we prove in the sequel that this effect of the conditioning eventually washes away, by showing that for all $n$ large enough, $i \in [n]$ and $\overline{x}_{B_i(\overline{t})}$,

$$B_i(\overline{t}) \in G^+_\varepsilon(\overline{t}) \quad \implies \quad 1 - \delta \leq \frac{Z_{n,i}^0}{Z_{n,i}^{0,t}} \leq \frac{1}{1 - \delta}. \tag{2.11}$$

Indeed, by the preceding identity, the right-side of (2.11) yields that the probability measures $\mathbb{P}^\varepsilon_{n,i}(i)$ corresponding to $\nu_{n,+}$ are then in $\mathcal{T}^{\varepsilon, \delta}_{(t, \overline{t})}$. Consequently, following the derivation of (2.10) we find that $1 - \varepsilon \leq \max_{i} \mathcal{T}^{\varepsilon, \delta}_{(t, \overline{t})}$ for any sub-sequential limit $\mathcal{M}$ of $\{\nu_{n,+}\}$ and all $\varepsilon, \delta > 0$. Since

$$\mathcal{T}_{(t, \overline{t})}^{\varepsilon,0} = \bigcap_{\delta > 0} \mathcal{T}_{(t, \overline{t})}^{\varepsilon, \delta},$$

considering $\delta \downarrow 0$ followed by $\varepsilon \downarrow 0$ completes the proof of (2.8).

Turning to prove (2.11), we clearly have that $Z_{n,i}^+ \leq Z_{n,i}^0 \leq Z_{n,i}^-$ for

$$Z_{n,i}^\pm := \nu_n\left( \sum_{j \notin B_i(\overline{t})} x_j \geq \pm |B_i(\overline{t})| \mid \overline{x}_{B_i(\overline{t})} \right).$$

Setting $\partial B_i(r) := B_i(r - 1, r)$, note that the value of $Z_{n,i}^\pm$ depends only on $\overline{x}_{\partial B_i(\overline{t})}$, hence we further have that $Z_{n,i}^+ \leq Z_{n,i}^{0,t} \leq Z_{n,i}^-$ whenever $t \leq \overline{t} - 1$ and consequently it suffices to show that for all $n$ large enough and $i \in [n]$,

$$B_i(\overline{t}) \in G^+_\varepsilon(\overline{t}) \quad \implies \quad \inf_{\overline{x}_{B_i(\overline{t})}} \left\{ \frac{Z_{n,i}^+}{Z_{n,i}^-} \right\} \geq 1 - \delta. \tag{2.12}$$

To this end, with $G_\varepsilon(\overline{t} + 1)$ a finite collection of finite graphs, necessarily,

$$K := \sup_{G \in G_\varepsilon(\overline{t} + 1)} |E(G)| < \infty.$$

Let $\nu_{E_n \setminus E(B_i(\overline{t} + 1))}$ denote the Ising measure on the sub-graph of $G_n$ in which all edges within $B_i(\overline{t} + 1)$ have been deleted and suppose hereafter that the left-side of (2.12) applies, namely, that $B_i(\overline{t} + 1) \simeq G$ for some $G \in G_\varepsilon(\overline{t} + 1)$. Then, at most $K$ edges of $G_n$ touch $B_i(\overline{t})$ and hence

$$\nu_n\left( \overline{x}_{G_n \setminus B_i(\overline{t})} \mid \overline{x}_{B_i(\overline{t})} \right) \geq e^{-2\beta K} \nu_{E_n \setminus E(B_i(\overline{t} + 1))}\left( \overline{x}_{G_n \setminus B_i(\overline{t})} \right),$$
out of which we deduce by the invariance at $B = 0$ of Ising Gibbs measures with respect to a global sign change, that

$$Z_{n,i}^- \geq \nu_n \left( \sum_{j \not\in B_i(\tilde{T})} x_j \geq 0 \mid \mathcal{Z}_{B_i(\tilde{T})} \right) \geq \frac{1}{2} e^{-2\beta K} \tag{2.13}$$

is uniformly bounded away from zero. Further, $|B_i(\tilde{T})| \leq K$ and by the assumed uniform sparseness of $\{G_n\}$, there exists $k \in \mathbb{N}$ and $n_0 \geq 3K$ large enough so that

$$\sum_{i=1}^{n} \Delta_i(G_n) I(\Delta_i(G_n) \geq k) \leq \frac{n}{3} \quad \forall n \geq n_0$$

(see (1.5)). Consequently, for any $n \geq n_0$ there are at least $n/3$ vertices in $G_n \setminus B_i(\tilde{T})$ of degree at most $k - 1$, out of which collection one can extract an independent set $S$ of $G_n$ whose size is at least $n/(3k)$. Thereby, one has as in the proof of [32, Lemma 4.1] that under $\nu_n$ and conditional on the values of $\mathcal{Z}_{SC}$, the ±-valued $\{x_j\}_{j \in S}$ are mutually independent, each having expectation within $(-\eta, \eta)$ for some $\eta = \eta(\beta, k) < 1$ and all $n$. As explained there, the Berry-Esseen theorem then implies that for some $C = C(k, \eta)$ finite and all $n \geq n_0$,

$$\sup_{r} \nu_n \left( \sum_{j \not\in B_i(\tilde{T})} x_j = r \mid \mathcal{Z}_{B_i(\tilde{T})} \right) \leq C n^{-1/2},$$

from which it follows that uniformly in $\mathcal{Z}_{B_i(\tilde{T})}$,

$$0 \leq Z_{n,i}^- - Z_{n,i}^+ \leq 2|B_i(\tilde{T})| C n^{-1/2} \leq 2KC n^{-1/2}.$$

Upon combining the preceding bound with (2.13), we conclude that (2.12) holds for all $n \geq n_0$. \hfill $\square$

3. Identifying the Limit Gibbs Measure

It helps to consider in the course of our proofs vertex dependent magnetic fields $B_i$. That is, to replace the model (1.1) by

$$\nu(x) = \frac{1}{Z(\beta, \mathcal{Z})} \exp \left\{ \beta \sum_{(i,j) \in E} x_i x_j + \sum_{i \in V} B_i x_i \right\}. \tag{3.1}$$

In this context, we often take advantage of Griffith’s inequality for ferromagnetic Ising models (which for completeness we state next, see also [26, Theorem IV.1.21]).

**Proposition 3.1.** [Griffith’s inequality] Consider two Ising models $\nu(\cdot)$ and $\nu'(\cdot)$ on finite graphs $G = (V, E)$ and $G' = (V, E')$, inverse temperatures $\beta$ and $\beta'$, and magnetic fields $\{B_i\}$ and $\{B'_i\}$, respectively. If $E \subseteq E'$, $\beta \leq \beta'$ and $0 \leq B_i \leq B'_i$, for all $i \in V$, then $0 \leq \langle \nu, \prod_{i \in U} x_i \rangle \leq \langle \nu', \prod_{i \in U} x_i \rangle$ for any $U \subseteq V$.

As we are having locally tree-like graphs, yielding local weak limit points supported on Ising Gibbs measures on trees, we often rely on the following representation for marginals of Ising measures on finite trees.

**Proposition 3.2.** [10, Lemma 4.1] For a subtree $U'$ of a finite tree $T$, let $\partial_s U'$ denote the subset of vertices $U'$ connected by an edge to $W \equiv T \setminus U'$ and for each $u \in \partial_s U'$ let $(x_u)_W$ denote the root magnetization of the Ising model on the maximal subtree $T_u$ of $W \cup \{u\}$ rooted at $u$. The marginal on $U'$ of an Ising measure $\nu$ on $T$, denoted $\nu^{T}_{U'}$, is then an Ising measure on $U'$ with magnetic field $B_u' = \tanh((x_u)_W) \geq B_u$ for $u \in \partial_s U'$ and $B_u' = B_u$ for $u \in U' \setminus \partial_s U'$. 
Adopting hereafter the notation $T_{x\rightarrow y}$ for the connected component of the sub-tree of $T$ rooted at $x$, after the path between $x$ and $y$ has been deleted, we start by relating $\mathbb{U}(\beta,0)$ to the limiting correlation $x; x_j$ across a uniformly chosen edge $(i,j) \in E_n$, under the measures $\nu_{n,\pm}$ and $\nu_n$.

**Lemma 3.3.** Suppose $G_n \overset{\text{strong}}{\longrightarrow} \mu$ such that $\mu(T_n) = 1$. Then, $(\beta, B) \mapsto \mathbb{U}(\beta, B)$ is bounded, non-decreasing, right-continuous at $\beta, B \geq 0$, continuous at any $B > 0$, and

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{(i,j) \in E_n} \nu_{n,+}^\beta(x_i \cdot x_j) = \lim_{n \to \infty} \frac{1}{n} \sum_{(i,j) \in E_n} \nu_n^\beta(x_i \cdot x_j) = \mathbb{U}(\beta,0),
$$

(3.2)

at any continuity point of $\beta \mapsto \mathbb{U}(\beta,0)$.

**Proof:** Since $\nu_n^\beta_0 = \frac{1}{2} \nu_{n,+}^\beta + \frac{1}{2} \nu_{n,-}^\beta$ and $\nu_n^{\beta_0}(x) = \nu_n^{\beta_0}(-x)$ for all $x$, clearly $\nu_n^{\beta_0}(x_i \cdot x_j) = \nu_n^{\beta_0}(x_i \cdot x_j)$ for any $(i,j) \in E_n$ and all $n$. It thus suffices to establish (3.2) in case of $\nu_n^{\beta_0}$, which since

$$
\frac{\partial}{\partial \beta} \phi_n(\beta, B) = \frac{1}{n} \sum_{(i,j) \in E_n} \nu_n^{\beta,B}(x_i \cdot x_j),
$$

(3.3)

for all $n$, $\beta$ and $B$, amounts to proving that

$$
\lim_{n \to \infty} \frac{\partial}{\partial \beta} \phi_n(\beta, B) = \mathbb{U}(\beta, B),
$$

(3.3)

for $B = 0$ and any $\beta \geq 0$ at which $\mathbb{U}(\beta,0)$ is continuous. To this end, applying [11, Lemma 2.12] for $A \equiv \{i,j\}$ and $U \equiv B_i(t)$, using Griffith’s inequality and local weak convergence, we obtain that per $\beta, B \geq 0$ and $t \geq 2$,

$$
\mathbb{E}_\mu \left[ \frac{1}{2} \sum_{i \in \partial o} \nu_{i,T}^{\beta,B,t}(x_o \cdot x_i) \right] \leq \liminf_{n \to \infty} \frac{\partial}{\partial \beta} \phi_n(\beta, B) \leq \limsup_{n \to \infty} \frac{\partial}{\partial \beta} \phi_n(\beta, B) \leq \mathbb{E}_\mu \left[ \frac{1}{2} \sum_{i \in \partial o} \nu_{i,T}^{\beta,B,t}(x_o \cdot x_i) \right],
$$

(3.4)

where $\nu_{i,T}^{\beta,B,t}$ is the Ising measure on $T(t)$ with free boundary condition on $\partial T(t) := T(t) \setminus T(t-1)$ (for more details, see [11, pp. 163-164]). Next, for probability measures

$$
\widehat{\nu}(x_1, x_2) = z^{-1} \exp \{ \beta x_1 x_2 + H_1 x_1 + H_2 x_2 \},
$$

(3.5)

on $\{ -1, +1 \}^2$ it is easy to check that

$$
\widehat{\nu}(x_1 \cdot x_2) = F(\tanh(\beta), m_1 m_2),
$$

(3.6)

with $m_j = \tanh(H_j), j = 1,2$. For $m_{\ell,t}(T^\prime) := \nu_{i,T^\prime}^{\beta,B,\ell}(x_o')$ for $\ell \in \{ f, + \}$ and the corresponding root-magnetization of the Ising measure on $(T^\prime(\ell), o') \in T_o(\ell)$, we note that for any $i \in \partial o$, the marginal on $U^\prime = (o,i)$ of the Ising measures $\nu_{i,T}^{\beta,B,t}$ is by Proposition 3.2 of the form (3.4), with $m_1 = m^{t+1}(T_{o-i})$ and $m_2 = m^{-1,1}(T_{i-o})$. Consequently, $\nu_{i,T}^{\beta,B,t}(x_o \cdot x_i) = F(\tanh(\beta), r_{\ell}(t))$ is a continuous function of $F(\beta, r(t))$ in the case $B > 0$, upon applying [15, Lemma 3.1] (which only requires local finiteness of the tree), first for $T = T_{o-i}$ and then for $T = T_{i-o}$, we deduce that $r_{\ell}(t) \sim r(t) \to 0$ and hence $\nu_{i,T}^{\beta,B,t}(x_o \cdot x_i) \sim 0$ as $t \to \infty$. This holds for all $i \in \partial o$, so recalling that $\deg(\mu)$ is finite (by uniform sparseness of $\{G_n\}$), we get by dominated convergence (DCT), that

$$
\lim_{t \to \infty} \mathbb{E}_\mu \left[ \frac{1}{2} \sum_{i \in \partial o} \nu_{i,T}^{\beta,B,t}(x_o \cdot x_i) \right] = \lim_{t \to \infty} \mathbb{E}_\mu \left[ \frac{1}{2} \sum_{i \in \partial o} \nu_{i,T}^{\beta,B,t}(x_o \cdot x_i) \right]
$$

(3.7)
and hence (3.3) holds, at any $B > 0$ and $\beta \geq 0$. While (3.7) is typically false at $B = 0$ and $\beta$ large enough, clearly for any $T \in T_*$ and finite $t \geq 0$, the function $\nu_{+,T}^{\beta,B,t}(x_o \cdot x_i)$ is jointly continuous in $\beta$ and $B$. These Ising measures of plus boundary condition correspond to taking $B_t \uparrow \infty$ at all $i \in T \setminus (t - 1)$ (see Definition 1.5). Hence, by Griffith’s inequality we have that $\frac{1}{2} \nu_{+,T}^{\beta,B,t}(x_o \cdot x_i)$ is non-increasing in $t$ and non-decreasing in $\beta, B \geq 0$. The same monotonicity properties apply for the sum of such functions over $i \in \partial_0$ and in so far as $(\beta, B)$ are concerned, retained by the expectation $\mathbb{U}(\beta, B)$ with respect to the law $\mu$ of $T$, of its limit as $t \uparrow \infty$. Since $\text{deg}(\mu)$ is finite we further deduce by DCT the joint continuity of

$$(\beta, B) \mapsto \mathbb{E}_\mu \left[ \sum_{i \in \partial_0} \nu_{+,T}^{\beta,B,t}(x_o \cdot x_i) \right]$$

which upon interchanging limits in $t$ and $\beta, B$, yields the right-continuity of $\mathbb{U}(\beta, B)$ at all $\beta, B \geq 0$. Since $\beta \mapsto \phi_n(\beta, B)$ are convex functions, so is their limit $\phi(\beta, B)$ (see [12, Theorem 1.8] for existence of such limit at any $\beta \geq 0$, $B \in \mathbb{R}$ fixed). Denoting hereafter by $f_n(\cdot) \stackrel{Q^c}{\rightarrow} f(\cdot)$ the convergence of $f_n$ to $f$ on some co-countable set, with $f(\cdot) \stackrel{Q^c}{\leq} g(\cdot)$ when $f$ and $g$ agree on a co-countable set, we thus have that $\frac{\partial}{\partial \beta} \phi_n(\beta, B) \stackrel{Q^c}{\rightarrow} \frac{\partial}{\partial \beta} \phi(\beta, B)$ per fixed $B \geq 0$, and consequently $\frac{\partial}{\partial \beta} \phi(\beta, B) \stackrel{Q^c}{\leq} \mathbb{U}(\beta, B)$ at any given $B > 0$. Fixing a sequence $B_m \downarrow 0$, by the convexity of $\beta \mapsto \phi(\beta, B)$ and the continuity of $B \mapsto \phi(\beta, B)$ we have $\frac{\partial}{\partial \beta} \phi(\beta, B_m) \stackrel{Q^c}{\rightarrow} \frac{\partial}{\partial \beta} \phi(\beta, 0)$. Further, $B \mapsto \mathbb{U}(\beta, B)$ is right continuous, hence $\mathbb{U}(\beta, B_m) \rightarrow \mathbb{U}(\beta, 0)$. From these two convergences we deduce that $\mathbb{U}(\beta, 0) \stackrel{Q^c}{\leq} \frac{\partial}{\partial \beta} \phi(\beta, 0)$. We have seen already that $\frac{\partial}{\partial \beta} \phi_n(\beta, 0) \stackrel{Q^c}{\leq} \frac{\partial}{\partial \beta} \phi(\beta, 0)$, hence also $\frac{\partial}{\partial \beta} \phi_n(\beta, 0) \stackrel{Q^c}{\leq} \mathbb{U}(\beta, 0)$. Since $\frac{\partial}{\partial \beta} \phi_n(\beta, 0)$ are non-decreasing continuous functions, this convergence extends to all continuity points of $\beta \mapsto \mathbb{U}(\beta, 0)$. \qed

The following extension of [32, Lemma 3.2] to arbitrary $T \in T_*$ allows us to utilize Lemma 3.3 for restricting the weak limit points of $\nu_{n,\pm}$ and $\nu_n$, to convex combinations of $\nu_{\pm,T}$.

**Lemma 3.4.** For any Ising Gibbs measure $\nu_T$ on some $T \in T^*$ and all $i \in V(T)$,

$$\sum_{j \in \partial_i} \nu_T \langle x_i \cdot x_j \rangle \leq \sum_{j \in \partial_i} \nu_{+,T} \langle x_i \cdot x_j \rangle = \sum_{j \in \partial_i} \nu_{-,T} \langle x_i \cdot x_j \rangle, \quad (3.8)$$

with strict inequality for some $i \in V(T)$ unless $\nu_T$ is a convex combination of $\nu_{+,T}$ and $\nu_{-,T}$.

**Proof.** The equality in (3.8) is an immediate consequence of the fact that under $\nu_{+,T}$ the random vector $-x_T$ admits the law $\nu_T$. Further, due to uniqueness of the Ising Gibbs measure for a finite $T$, we may and shall consider hereafter a fixed infinite tree $T$. There are only countably many edges in $T$ and the non-empty collection of Ising Gibbs measures on $T$ is convex, with each Ising Gibbs measure on $T$ being a mixture of the extremal Ising Gibbs measures on $T$ (see [18, Chapter 7]). Consequently, it suffices to find an extremal Ising Gibbs measure $\nu_T \neq \nu_{\pm,T}$ and show that for every edge $(i, j) \in E(T)$,

$$\nu_T \langle x_i \cdot x_j \rangle \leq \nu_{+,T} \langle x_i \cdot x_j \rangle \quad (3.9)$$

with a strict inequality for at least one $(i, j) \in E(T)$. To this end, for each $(i, j) \in E(T)$ let $m_{i \rightarrow j}^0 := \nu_T \langle x_i \rangle$ for the probability measure $\nu_T \langle x_i \rangle$ whose Radon-Nikodym derivative with respect to $\nu_T$ is proportional to $e^{-\beta x_i}$.

That is,

$$m_{i \rightarrow j}^0 = \frac{\nu_T \langle x_i e^{-\beta x_i} \rangle}{\nu_T \langle e^{-\beta x_i} \rangle} = \lim_{l \rightarrow \infty} \frac{\nu_T \langle x_i e^{-\beta x_i} \rangle |_{B_l(i)}}{\nu_T \langle e^{-\beta x_i} \rangle |_{B_l(i)}}.$$
where the limit exists by backward martingale convergence theorem and is a.e. constant by the tail triviality of the extremal measure \( \nu_T \) (see [18, Chapter 7]). Using the DLR condition (2.2) for \( \nu_T \) and the tree structure of \( T \), we deduce that \( \nu_T \)-a.e.

\[
m_{i\to j}^\nu = \lim_{l \to \infty} \tilde{\nu}(x_i \mid \mathcal{G}_{T_{i\to j}(l,l+1)}, T_{i\to j}(l+1)),
\]

with \( m_{i\to j}^- = -m_{i\to j}^+ \) denoting the values of \( m_{i\to j}^\nu \) in case of Ising Gibbs measures \( \nu_- \) and \( \nu_+ \), respectively. By the dct, the DLR condition (2.2) for \( \nu_T \), Proposition 3.2 and (3.10), for each \( t \in \mathbb{N} \) the marginal law of \( \mathcal{G}_{T(t)} \) under \( \nu_T \) is completely determined by \( \{m_{i\to j}^\nu, i \in \partial T(t), j \in \partial T(t-1)\} \).

In particular, considering the formula (3.5), we get by the same line of reasoning that

\[
\nu_T(x_i \cdot x_j) = F(\tanh(\beta), m_{i\to j}^\nu m_{j\to i}^\nu),
\]

for \( F(\theta, r) \) of (3.6) and any \( (i, j) \in E(T) \), with the analogous expression in case of \( \nu_+ \).

Further, from (3.10) and Griffith’s inequality we know that \( |m_{i\to j}^\nu| \leq m_{i\to j}^+ \) for all \( (i, j) \in E(T) \), out of which we get the inequality (3.9) by the strict monotonicity of \( r \to F(\theta, r) \) on \([-1, 1]\) (when \(|\theta| < 1\)). Turning to prove that having equality in (3.9) for all \( (i, j) \in E(T) \) implies either \( \nu_T = \nu_+ \) or \( \nu_T = \nu_- \), note that by the preceding such equalities translate into

\[
m_{i\to j}^\nu m_{j\to i}^\nu = m_{i\to j}^+ m_{j\to i}^- \quad \forall (i, j) \in E(T).
\]

From (3.10) one also have by an explicit calculation for Ising measures on trees, that

\[
m_{i\to j}^\nu = \tanh \left[ \sum_{k \in \partial i \setminus \{j\}} \text{atanh}(\tanh(\beta)m_{k\to i}^\nu) \right] \quad \forall (i, j) \in E(T),
\]

with the same recursion holding for the collections \( \{m_{i\to j}^\nu, (i, j) \in E(T)\} \). Suppose now that some \( (i, j) \in E(T) \) is a plus edge, namely both \( m_{i\to j}^\nu = m_{i\to j}^+ \) and \( m_{j\to i}^\nu = m_{j\to i}^+ \). Out of (3.13) we have that \( m_{i\to j}^\nu \) is strictly increasing in each \( m_{k\to i}^\nu \), \( k \in \partial i \setminus \{j\} \), so with \( |m_{k\to i}^\nu| \leq m_{k\to i}^+ \), the assumed equality \( m_{i\to j}^\nu = m_{i\to j}^+ \) at both directed edges \( e = \{i \to j\} \) and \( e = \{j \to i\} \), implies the same at all directed edges \( k \to i \), \( k \in \partial i \). Further, by (3.13) the values of \( m_{k\to i}^\nu \) and \( m_{i\to k}^\nu \) are given by the same function of \( \{m_{k\to i}^\nu\} \) and \( \{m_{i\to k}^\nu\} \) respectively, whose arguments are directed edges \( e \) where we already have \( m_{i\to j}^\nu = m_{i\to j}^- \). Hence, that equality holds also for all directed edges of the form \( e = \{i \to k\} \). That is, every edge of \( B_i(1) \) is a plus edge. This property extends in the same manner to \( B_i(t), t = 2, 3, \ldots \), and so we conclude that a single plus edge in \( T \) results with each edge being plus edge, and thereby with \( \nu_T = \nu_+ \). By the same line of reasoning, a single minus edge \( (i, j) \) where both \( m_{i\to j}^\nu = m_{i\to j}^- \) and \( m_{j\to i}^\nu = m_{j\to i}^- \) yields that all edges of \( T \) are minus edges and thereby \( \nu_T = \nu_- \). Suppose now that there are neither plus nor minus edges in \( T \). We then have by (3.12) that at each edge \( (i, j) \) either \( m_{i\to j}^+ > 0 \) and \( m_{j\to i}^+ = 0 \), or the same applies upon reversing the roles of \( i \) and \( j \), and we thus complete the proof by ruling out the possibility of \( \nu_+ \) having the latter property. Indeed, by (3.13) if some \( m_{i\to j}^+ > 0 \) then \( m_{i\to j}^+ \) is strictly positive for at least one edge \( (l, i) \) of \( T \).

The latter is neither plus nor minus edge, so \( m_{i\to l}^+ = 0 \), which with \( m^+ \) everywhere non-negative, implies by (3.13) that \( m_{k\to i}^+ = 0 \) at all \( k \in \partial i \setminus \{l\} \). That is, having \( m_{i\to j}^+ > 0 \) results with \( m_{i\to j}^+ \) strictly positive at exactly one edge \( e \) directed into \( i \). Continuing in this manner we find an infinite directed ray \( \{i_s \to i_{s-1} : (i_s, i_{s-1}) \in E(T), s \in \mathbb{N}\} \) (ending at \( i_1 = i \) and \( i_0 = j \), with \( m_{i_s\to i_{s-1}}^+ > 0 \) while \( m_{k\to i_{s-1}}^+ = 0 \) for all \( k \neq i_s, s \geq 1 \). That is, again by (3.13), \( m_{i_s\to i_{s-1}}^+ \leq \tanh(\beta)m_{i_{s-1}\to i_s}^+ \) for all \( s \geq 1 \). With \( \tanh(\beta) < 1 \) it is obviously impossible to have such an infinite sequence of strictly positive \( m_{i_s\to i_{s-1}}^+ \leq 1 \).
Remark 3.5. Unlike the case of $k$-regular trees $T_k$ considered in [32, Lemma 3.2], we may have
\[ \sum_{i \in \partial o} \nu_T(x_o \cdot x_i) = \sum_{i \in \partial o} \nu_{+T}(x_o \cdot x_i) \]
for some $T \in \mathcal{T}_s$ and an extremal Ising Gibbs measure $\nu_T \neq \nu_{\pm T}$ on it. Indeed, as the proof of Lemma 3.4 shows, this happens whenever $\beta > 0$ is such that for some $i \in \partial o$ there is a unique Ising Gibbs measure on the sub-tree $T_{o \rightarrow i}$ while $T' := T_{i \rightarrow o}$ admits an extremal Ising Gibbs measure other than $\nu_{\pm T'}$ (e.g. when $T_{i \rightarrow o}$ is $k_2$-regular, while $T_{o \rightarrow i}$ is finite or $k_1$-regular and $\beta_c(k_2) < \beta < \beta_c(k_1)$). Nevertheless, our next lemma utilizes the unimodularity of $\mu$ to circumvent this problem.

Lemma 3.6. Fixing $\mu \in \mathcal{U}$ such that $\mu(\mathcal{T}_s) = 1$, for any $m$ supported on the collection $\mathcal{I}_s$ of Ising Gibbs measures $\mathcal{I} = \delta_T \otimes \nu_T$ and having the law $\mu$ for $T$,
\[ \mathbb{E}_{\mu} \left[ \sum_{i \in \partial o} \nu_T(x_o \cdot x_i) \right] \leq \mathbb{E}_{\mu} \left[ \sum_{i \in \partial o} \nu_{+T}(x_o \cdot x_i) \right] = \mathbb{E}_{\mu} \left[ \sum_{i \in \partial o} \nu_{-T}(x_o \cdot x_i) \right] ; \tag{3.14} \]
with strict inequality unless $m$ is supported on the sub-collection $\mathcal{I}_+ \subset \mathcal{I}_s$ of $\mathcal{I} = \delta_T \otimes \nu_T$ where
\[ \nu_T = \alpha_T \nu_{+T} + (1 - \alpha_T) \nu_{-T} \tag{3.15} \]
for some Borel measurable function $\alpha : \mathcal{T}_s \rightarrow [0, 1]$, with $\alpha_T = \frac{1}{T}$ whenever $\nu_{+T} = \nu_{-T}$.

Remark 3.7. In the proof of Lemma 3.6 we take advantage of the Markov chain $\{Y_n\}_{n \geq 0}$ on the space $\mathcal{T}_s$ (commonly known as “walk from the point of view of the particle”), induced by setting the root of $T$ to follow the trajectory of discrete time simple random walk (srw) on $(T, o) \in \mathcal{T}_s$, starting at $o$. Specifically, associating with each $\mu \in \mathcal{U}$ for which $\text{deg}(\mu) > 0$, the “size-biased-root” probability measure $\sigma := \frac{\text{deg}}{\text{deg}(\mu)} \mu$, recall that choosing the initial state $(T, o) \in \mathcal{T}_s$ according to $\sigma$, yields a stationary and reversible joint law $\tilde{\mu}$ for the trajectory of $Y_n$ (c.f. [4, Theorem 4.1])

Proof: We get (3.14) by considering the expectation of (3.8) for $i = o$, over the law $m$ of $T$ and the Ising Gibbs measure $\nu_T$ on it. Further, there is only one Ising Gibbs measure on $T = \{o\}$. So, our claim about strictness of the inequality in (3.14) trivially holds in case $\text{deg}(\mu) = 0$, and assuming hereafter that $\text{deg}(\mu) > 0$, we consider the stationary Markov chain $\{Y_n\}_{n \geq 0}$ on $\mathcal{T}_s$, as in Remark 3.7, for $Y_0 = (T, o)$ of law $\sigma$. Let $\mathcal{I} = \delta_T \otimes \nu_T$ denote the expected value of $\nu_T$ under the probability measure $m$ conditional upon $T \in \mathcal{T}_s$, which up to some $\mu$-null set $N \subset \mathcal{T}_s$ is a uniquely defined Ising Gibbs measure on $T$ (due to convexity of the latter collection). Equality in (3.14) thus amounts to $\mathbb{E}[f(Y_0)] = 0$ for the $\mathcal{T}_s$-measurable, uniformly bounded and non-negative (see (3.8)),
\[ f((T, o)) := \frac{1}{\Delta_o} \sum_{j \in \partial o} \left[ \nu_{+T}(x_o \cdot x_j) - \nu_{T}(x_o \cdot x_j) \right] , \]
which by the stationarity of $\{Y_n\}$ (c.f. [4, Theorem 4.1]), implies that
\[ \mathbb{E}[f(Y_n)] = 0 \quad \forall n \in \mathbb{N} . \tag{3.16} \]
Conditional on $Y_0 = (T, o)$, the probability of $Y_\ell = (T, i)$ is strictly positive for each $\ell \in \mathbb{N}$ and $i \in \partial T(\ell)$, hence with $f(\cdot)$ non-negative, it follows from (3.16) that
\[ \sigma((T, o) \in \mathcal{T}_s : \exists i \in V(T), f((T, i)) > 0) = 0 . \]
We thus conclude that for $\mu$-a.e. $T$, equality holds in (3.8) for $\nu_T$ and all $i \in V(T)$, so by Lemma 3.4 the Ising Gibbs measure $\nu_T$ must then be a convex combination of $\nu_{+T}$ and $\nu_{-T}$. Now recall that to any Ising Gibbs measure $\nu_T$ on $T$ corresponds a unique probability measure $\Theta_{\nu_T}$ supported on the collection $\{\nu_T^c\}$ of extremal Ising Gibbs measures on $T$, such that $\nu_T(\cdot) = \int \nu_T^c(\cdot) d\Theta_{\nu_T}$ (c.f.
Therefore, by its definition, μ-a.e. \( \nabla_\tau(\cdot) = \int \nu^\tau_d(\cdot)d\Theta_\tau \) for the expected value \( \Theta_\tau(\cdot) \) of \( \Theta_{\nu_\tau}(\cdot) \) under the probability measure \( m \) conditional upon \( \tau \in \mathcal{T}_\tau \). We have just shown that μ-a.e. \( \Theta_{\nu_\tau}(\nu_{+\tau}, \nu_{-\tau})^c = 0 \), hence m-a.e. this holds for \( \Theta_{\nu_\tau} \). That is, up to some m-null set, \( \nu_\tau \) is of the form (3.15), as claimed.

The following lemma completes the proof of Theorem 1.6.

**Lemma 3.8.** Under the conditions of Lemma 3.3 we have that:
(a) Any sub-sequential local weak limit \( m_+ \) of \( \{\nu_{n+}\} \) is supported on the collection \( \mathcal{I}_+ \), with \( \tau \) distributed according to \( \mu \).
(b) Any sub-sequential local weak limit of \( \{\nu_n\} \) equals \( m = \mu \circ \psi^{-1} \) for \( \psi(\tau) = \delta_\tau \otimes (\frac{1}{2}\nu^\beta_{+\tau} + \frac{1}{2}\nu^\beta_{-\tau}) \).

**Proof:** (a). Recall Lemma 3.3, that
\[
\frac{2}{n} \sum_{(i,j) \in E_n} \nu_{n+i,j}(x_i \cdot x_j) = \mathbb{E}_{U_n}(F(P^2_n(I_n))) \rightarrow \mathbb{E}_\mu\left[ \sum_{i \in \partial o} \nu_{+\tau}(x_o \cdot x_i) \right], \tag{3.17}
\]
for \( P^2_n(i) \) corresponding to \( \nu_{n+} \) and the function \( F(\nu) := \nabla(\sum_{i \in \partial o} x_o \cdot x_i) \) on \( \mathcal{P}(\mathcal{G}_s(2)) \), which is bounded by \( \nabla(\Delta_\partial) \) and continuous with respect to weak convergence. By assumption, under \( U_n \) the law of \( P^2_n(I_{n,m}) \) converges weakly to \( m_+^2 \) along some sub-sequence \( n_m \rightarrow \infty \). Hence, by \( \text{dct} \) and the uniform integrability of \( \{\Delta_\partial\} \),
\[
\lim_{m \rightarrow \infty} \mathbb{E}_{U_{n,m}}(F(P^2_n(I_{n,m}))) = \mathbb{E}_{m_+}[F(\nu^2)]. \tag{3.18}
\]
Recall part (b) of Lemma 2.4 that \( m_+ \) is supported on the collection \( \mathcal{I}_s \) of Ising Gibbs measures of the form \( \delta_\tau \otimes \nu_\tau \), having the law \( \mu \in \mathcal{U} \) for \( \tau \in \mathcal{T}_s \). Thus, comparing the RHS of (3.17) with the RHS of (3.18), we deduce that
\[
\mathbb{E}_{m_+}\left[ \sum_{i \in \partial o} \nu_\tau(x_o \cdot x_i) \right] = \mathbb{E}_\mu\left[ \sum_{i \in \partial o} \nu_{+\tau}(x_o \cdot x_i) \right]
\]
out of which it follows by Lemma 3.6 that \( m_+ \) is supported on the sub-collection \( \mathcal{I}_+ \).

(b). Considering now part (a) of Lemma 2.4 we get by the preceding argument that any sub-sequential weak limit \( m \) of \( \{\nu_n\} \) is supported on \( \mathcal{I}_+ \) with \( \tau \) distributed according to \( \mu \). In particular, \( \mu \)-a.e.
\[
|\nu_\tau(x_o)| = |2\alpha_\tau - 1|\nu_{+\tau}(x_o).
\]
As in the proof of Lemma 3.4, if \( \nu_{+\tau}(x_o) = 0 \), then necessarily \( m_{i \rightarrow j}^+ = 0 \) for all \( (i, j) \in E(\tau) \), hence \( \nu_{+\tau} = \nu_{-\tau} \) and \( \alpha_\tau = \frac{1}{2} \). More generally, the bounded function \( \tilde{F}(\nu) := |\nu(x_o)| \) on \( \mathcal{P}(\mathcal{G}_s(1)) \) is continuous with respect to weak convergence. Since
\[
0 = n^{-1} \sum_{i=1}^n |\nu_n(x_i)| = \mathbb{E}_{U_n}(\tilde{F}(P^1_n(I_n)))
\]
for all \( n \), it thus follows that for any local weak limit point \( m \) of \( \{\nu_n\} \),
\[
0 = \mathbb{E}_m[\tilde{F}(\nu^1)] = \mathbb{E}_m[|\nu_\tau(x_o)|] = \mathbb{E}_m\left[ |2\alpha_\tau - 1|\nu_{+\tau}(x_o) \right],
\]
thereby forcing \( m \)-a.s. \( \alpha_\tau = \frac{1}{2} \). \( \square \)
4. Proof of Theorem 1.8

Given part (a) of Lemma 3.8 it remains only to show that \( m_+ -\)a.s. \( \alpha_T = 1 \) for any sub-sequential local weak limit point \( m_+ \) of \( \{ \nu_{n,+} \} \). To this end, we make use of the following definition.

**Definition 4.1.** Let \( \mathbb{P}^t_n := \mathbb{E}_{I_n}(\mathbb{P}^t_n(I_n)) \) denote the average of \( \mathbb{P}^t_n(i) \) of Definition 1.4. We say that \((G_n, \zeta_n)\), or in short, that probability measures \( \zeta_n \) on \( \mathcal{X}^V_n \), converge on average to \( \nu \), a probability measure on \((\mathcal{G}_n, \mathcal{G}_n^\alpha)\), if for any positive fixed integer \( t \),

\[
\mathbb{P}^t_n \Rightarrow \nu^t, \quad \text{as } n \to \infty. \tag{4.1}
\]

**Remark 4.2.** Note that if \( \{ \zeta_n \} \) converges locally weakly to \( \overline{\nu} \), then it also converges on average to \( \nu = \int \overline{\nu}(d\nu) \). In particular, if \( \overline{\nu} \) is supported on the subset \( \mathcal{I}_+ \) of Ising Gibbs measures then it follows by linearity of the conditional expectation that the corresponding limit on average \( \overline{\nu} \) of \( \{ \zeta_n \} \), is itself an Ising Gibbs measure, with \( \overline{\nu} \) distributed according to the \( \mathcal{P}(\mathcal{T}_s) \)-marginal of \( m \) and \( \overline{\nu}_T \) of the form (3.15) for some measurable \( \alpha : T_s \to [0, 1] \).

Next, for \( G \in \mathcal{G}_s, \ r \in \mathbb{N}, \ l \geq 0 \) and \( i \in V(G) \), let \( a_{l,i,j}^{G} \) denote the expected relative to \( l \) occupation time at \( j \in V(G) \) by the continuous time \( \text{S.RW} \) \( \{ X_s \} \) on \( G \) starting at \( X_0 = i \) and run till \( \min(l, \theta_r) \) for \( \theta_r := \inf\{ s \geq 0 : X_s \notin B_i(r) \} \). That is,

\[
a_{l,i,j}^{G} = \frac{1}{l} \int_0^l \mathbb{P}_i(X_s = j, s \leq \theta_r) \, ds,
\]

adopting the notations \( a_{l,i,j}^{G} \) for \( G = G_n \) and \( a_{l,i,j}^{G,T} \) in case \( G = T \in \mathcal{T}_s \). These non-negative weights induce for every \( \nu \in \mathcal{X}^V(G) \) the weighted averages

\[
y_{l,i,j}^{G} = \sum_j x_j a_{l,i,j}^{G} \tag{4.2}
\]

and our proof is based on the analysis of the corresponding functionals per \( \eta > 0 \) and \( t \in \mathbb{N} \),

\[
F_{l,i,j}^{G}(x, \eta) := J_{l,i,j}^{G}(x, \eta) A_{l,i,j}^{G}(\eta), \tag{4.3}
\]

\[
J_{l,i,j}^{G}(x, \eta) := I\left\{ y_{l,i,j}^{G}(x) \leq -\eta \right\}, \quad A_{l,i,j}^{G}(\eta) := I\left\{ \left| \nu_{l,i,j}^{G}(y_{l,i,j}^{G}) \right| \geq 2\eta \right\}, \tag{4.4}
\]

where \( \nu_{l,i,j}^{G} \) denotes the Ising measure on \((G, i)\) conditioned to have \( \mathbb{Z}_{B_i(t)}^{(t)} = (+)_{B_i(t)}^{(t)} \). We use \( y_{l,i,j}^{G} \) \((J_{l,i,j}^{G}, F_{l,i,j}^{G})\) when \( G = G_n \) (i.e. with weights \( a_{l,i,j}^{G} \)), and \( y_{l,i,j}^{G} \) \((J_{l,i,j}^{G}, F_{l,i,j}^{G})\) when \( G = T \in \mathcal{T}_s \) (i.e. weights \( a_{l,i,j}^{G} \)), omitting \( r \) and \( t \) in case \( r = \infty \) (respectively, \( t = \infty \), which for \( A_{l,i,j}^{G} \) means using \( \nu_{l,i,j}^{G} \)), and arguments \( \eta, G, x \) whose value is clear from the context.

**Lemma 4.3.** Suppose \( G_n \overset{\text{SWC}}{\Rightarrow} \mu \) for some \( \mu \in \mathcal{U} \) with \( \mu(T_s) = 1 \), and \( \{ \nu_{n,+} \} \) converges locally weakly to some \( m_+ \) supported on \( \mathcal{I}_+ \). Then, with \( \overline{\nu}_+ \in \mathcal{P}(\mathcal{T}_s) \) denoting the corresponding limit on average of \( \{ \nu_{n,+} \} \), for any fixed \( l \) and up to at most countably many \( \eta > 0 \),

\[
\lim_{t \to \infty} \lim_{n \to \infty} \mathbb{E}_{U_n} \left[ \nu_{n,+}(F_{l,t}^{G}) \right] = \lim_{t \to \infty} \overline{\nu}_+(F_{l,t}^{G}) = \overline{\nu}_+(F_{l,t}^{T}), \tag{4.5}
\]

\[
\lim_{t \to \infty} \lim_{n \to \infty} \mathbb{E}_{U_n} \left[ \sum_{i \in \partial I_n} \nu_{n,+}(F_{i,t}^{l}) \right] = \lim_{t \to \infty} \overline{\nu}_+ \left[ \sum_{i \in \partial \mathcal{I}_+} I(F_{i,t}^{l} \neq F_{i,t}^{G}) \right] = \overline{\nu}_+ \left[ \sum_{i \in \partial \mathcal{I}_+} I(F_{i,t}^{T} \neq F_{i,t}^{G}) \right]. \tag{4.6}
\]
Proof: We show that all functions considered here can be approximated well by local functions, upon which our conclusions follow from the local weak convergence of \( \nu_{n,+} \). Indeed, with \( |x_j| \leq 1 \), for any graph \( G \), all \( l, r \geq 0 \), and \( i \in V(G) \),
\[
|y^l,G_i - y^{l,r,G}_i| \leq \frac{1}{r} \int_0^r \mathbb{P}_i(\theta_r \leq s)ds \leq \mathbb{P}_i(\theta_r < l) := \gamma_i^{l,r,G}.
\]
Next, setting \( \varepsilon_r := \mathbb{E}_\mu[\gamma_0^{l,r,T}]^{1/2} \), by Markov’s inequality \( \mu(\gamma_0^{l,r,T} \geq \varepsilon_r) \leq \varepsilon_r \). We now show that \( \varepsilon_r \downarrow 0 \). Indeed, \( \theta_r \geq \tau_r \), the time of \( r \)-th jump by the srw \( \{X_t\} \) on \( T \), having the representation
\[
\tau_r = \sum_{k=1}^r \frac{E_k}{\Delta Y_{k-1}},
\]
for \( \{E_k\} \) i.i.d. standard Exponential random variables, independent of the discrete Markov chain \( \{Y_n\}_{n \geq 0} \) embedded in the continuous time srw \( \{X_t\}_{t \geq 0} \). Hence, with \( L^{r,T} \) denoting the cardinality of \( \{k \leq r : \Delta Y_{k-1} \geq \sqrt{r} \} \), observe that if \( L^{r,T} \leq r/2 \) then \( \tau_r \) stochastically dominates the sum \( r/2 \) among the \( E_k \)-s, divided by \( \sqrt{r} \), whereas by Markov’s inequality (and stationarity of \( \{\Delta Y_n\} \) for \( (T,0) \)), we have that
\[
\mu(L^{r,T} \geq \frac{r}{2}) \leq \frac{\text{deg}(\mu)}{\sigma(L^{r,T} \geq \frac{r}{2})} \leq \frac{\text{deg}(\mu)}{r} \mathbb{E}_\sigma[L^{r,T}] = 2\text{deg}(\mu)\sigma(\Delta_o \geq \sqrt{r}).
\]

So, with \( \mathbb{N}_b \) denoting a Poisson variable of parameter \( b \),
\[
\varepsilon_r^2 = \mu(\mathbb{P}_o(\theta_r < l)) \leq \mu(\mathbb{P}_o(\tau_r \leq l)) \leq \mathbb{P}(\mathbb{N}_b \mathbb{P}_r \geq \frac{r}{2}) + 2\text{deg}(\mu)\mathbb{P}(\Delta_o \geq \sqrt{r}),
\]

obviously decaying to zero when \( r \to \infty \). Now, for any graph \( G \), all \( i \in V(G) \) and \( l, t, r, \eta \geq 0 \)
\[
F_i^{l,t,r,G}(\eta + \varepsilon_r) - \mathbb{I}(\gamma_i^{l,r,G} \geq \varepsilon_r) \leq F_i^{l,t,G}(\eta) \leq F_i^{l,t,r,G}(\eta - \varepsilon_r) + \mathbb{I}(\gamma_i^{l,r,G} \geq \varepsilon_r).
\]

Further, if the balls \( B_i(t \vee r + 1) \) of \( G_1 \) and \( G_2 \) are isomorphic then \( a^{l,r,G}_i = a^{l,r,G_2}_i \) and, restricted to \( B_i(t) \), the measure \( \nu_{+,B_i(t)}^{\beta,0} \) is same for both graphs. Consequently,
\[
F_i^{l,t,r,G_1}(\mathbb{N}_b, \eta) = F_i^{l,t,r,G_2}(\mathbb{N}_b, \eta), \quad \gamma_i^{l,t,G_2} = \gamma_i^{l,r,G_2}.
\]

For \( \zeta_n = \nu_{n,+} \) we have assumed that \( \mathbb{P}^s_n \Rightarrow \mathbb{P}_+^s \) for any fixed \( s > t \vee r \), hence by (4.8) and (4.9),
\[
\mathbb{P}_+(F_0^{l,t,T}(\eta + 2\varepsilon_r) - 2\varepsilon_r) \leq \mathbb{P}_+(F_0^{l,t,r,T}(\eta + \varepsilon_r)) - \mu(\gamma_0^{l,r,T} \geq \varepsilon_r)
\leq \liminf_{n \to \infty} \mathbb{E}_{U_n}(\nu_{n,+}(F_0^{l,t}(\eta)^{i,j})) \leq \limsup_{n \to \infty} \mathbb{E}_{U_n}(\nu_{n,+}(F_0^{l,t}(\eta))
\]
\[
\leq \mathbb{P}_+^s(F_0^{l,t,r,T}(\eta - \varepsilon_r)) + \mu(\gamma_0^{l,r,T} \geq \varepsilon_r) \leq \mathbb{P}_+(F_0^{l,t,T}(\eta - 2\varepsilon_r) + 2\varepsilon_r).
\]

Taking \( r \to \infty \) and excluding for \( \eta > 0 \) the countably many points of discontinuity of the \( [0,1] \)-valued, non-increasing, left-continuous \( \mathbb{P}_+(F_0^{l,t,T}(\eta)) \), for any \( t \in \mathbb{N} \cup \{\infty\} \), we have that both lower and upper bounds in (4.10) converge to \( \mathbb{P}_+(F_0^{l,t,T}(\eta)) \), thus establishing the left identity of (4.5), as well as the bounds
\[
\mathbb{P}_+(F_0^{l,t,r,T}(\eta + \varepsilon_r)) - \varepsilon_r \leq \mathbb{P}_+(F_0^{l,t,T}(\eta)) \leq \mathbb{P}_+(F_0^{l,t,r,T}(\eta - \varepsilon_r)) + \varepsilon_r,
\]

for all \( t, r \in \mathbb{N} \). Further, recall that \( A_{i,t-1,r,T} = \mathbb{I}\left\{ \nu_{+,T}^{\beta,0}(y_0^{l,r,T}) \geq 2\eta \right\} \) for the Ising measures of Definition 1.5 which for any fixed \( i \in V(T) \) converge locally to \( \nu_{+,T}^{\beta,0} \) when \( t \to \infty \). Consequently,
upon taking $t \to \infty$ followed by $r \to \infty$, and further excluding for $\eta > 0$ the countable collection of points of discontinuity for any of $\{\nu_{+,T}(F_o^{r,T}(\eta \pm \varepsilon r)), r \in \mathbb{N}\}$, we deduce that

$$\nu_{+,T}(F_o^{r,T}(\eta)) \to \nu_{+,T}(F_o^{T}(\eta)),$$

which by (4.11) gives the RHS of (4.5). Turning to prove the left identity in (4.6), since

$$\mathbb{E}_{U_n}\left[ \sum_{i \in \partial I_n} \nu_{n,+}(F_{I_n}^{d,t,n} = 1, F_{I_n}^{d,t,n} = 0) \right] = \mathbb{E}_{U_n}\left[ \sum_{i \in \partial I_n} \nu_{n,+}(F_{I_n}^{d,t,n} = 0, F_{I_n}^{d,t,n} = 1) \right]$$

it suffices to prove that

$$\lim_{n \to \infty} \mathbb{E}_{U_n}\left[ \sum_{i \in \partial I_n} \nu_{n,+}(F_{I_n}^{d,t,n} = 1) \right] = \lim_{n \to \infty} \mathbb{E}_{U_n}\left[ \sum_{i \in \partial I_n} \nu_{n,+}(F_{I_n}^{d,t,n} = 0) \right] = \mathcal{H}_{11}^t(\eta),$$

and

$$\lim_{n \to \infty} \mathbb{E}_{U_n}\left[ \Delta_{I_n} \nu_{n,+}(F_{I_n}^{d,t,n}) \right] = \mathcal{H}_1^t(\eta).$$

Both $\mathcal{H}_{11}^t(\eta)$ and $\mathcal{H}_1^t(\eta)$ are bounded (by $\nu_{+,T}(\Delta_{I_n}) = \deg(\mu)$), non-negative, left-continuous, non-increasing functions of $\eta$. Thus, excluding the at most countably many points of discontinuity of $\eta \mapsto (\mathcal{H}_{11}^t(\eta), \mathcal{H}_1^t(\eta))$ over all choices of $t \in \mathbb{N} \cup \{\infty\}$, we establish (4.12) and (4.13) upon deriving inequalities analogous to (4.10) for $\sum_{i \in \partial I_n} F_{I_n}^{d,t}(\eta) F_{I_n}^{d,t}(\eta)$ and $\Delta_{I_n} F_{I_n}^{d,t}(\eta)$, respectively. These in turn also provide the analogs of (4.11) with $\nu_{+,T}(F_o^{T}(\eta \pm \varepsilon r))$ replaced by the non-increasing in $\eta$ and uniformly bounded $\mathcal{H}_{11}^{t,r}(\eta \pm 2\varepsilon r)$, $\mathcal{H}_1^{t,r}(\eta \pm 2\varepsilon r)$, respectively, out of which we get the RHS of (4.6) along the same lines we used for deriving the RHS of (4.5). We proceed to show that within the support of $\nu_{n,+}$ one has $F_{x,n}^{d,t}(\eta) = 0$ at least for a $\delta := \eta/(1 + \eta)$ fraction of the vertices $i \in G_n$.

**Lemma 4.4.** For any $\eta, l \geq 0$, $n \in \mathbb{N}$ and $x$ such that $\sum_j x_j \geq 0$, 

$$\mathbb{E}_{U_n}[1 - J_{I_n}^{l,n}(x,\eta)] \geq \frac{\eta}{1 + \eta}.$$

**Proof:** Since $\sum_k a_{i,j,k}^{l,n} = 1$ for any $n, l$, we have that $J_{I_n}^{l,n}(x,\eta) = \mathbb{1}_{\{z_i \geq 1 + \eta\}}$ for the non-negative $z_i := \sum_j (1 - x_j) a_{i,j}^{l,n}$. Further, due to reversibility of the SRW, $a_{i,j}^{l,n} = a_{j,i}^{l,n}$ for all $i, j \in V_n$. Hence,

$$\mathbb{E}_{U_n}(z_{I_n}) = \frac{1}{n} \sum_{k,j=1}^n (1 - x_j) a_{k,j}^{l,n} = \frac{1}{n} \sum_{j=1}^n (1 - x_j) \leq 1,$$

by our assumption that $\sum_j x_j \geq 0$ and applying Markov’s inequality to $z_{I_n}$ completes the proof. 

Clearly the class $U$ is convex (with $\mu \in U$ extremal, if it cannot be written as a convex combination of other elements in $U$). From [4, Theorem 4.6, Theorem 4.7] it follows that for extremal $\mu \in U$, the stationary joint law $\tilde{\mu}$ of Remark 3.7 is ergodic (i.e. every shift-invariant event is $\tilde{\mu}$ trivial). By a slight abuse of notation we term $\mu$ as ergodic (and denote it by $\mu^*$), when the corresponding $\tilde{\mu}$ is ergodic. The next lemma is key to our use of $y_{i}^{l,T}$, and more specifically, the functionals $A_{i}^{l,t,T}$, for determining the possible limit points $\mathbf{m}_+$. It shows that in an ergodic setting, for a.e. $T \in \mathcal{T}_*$ the mean

$$\rho_i^{l,T} := \nu_{+,T}(y_{i}^{l,T}),$$

of $y_{i}^{l,T}$ under the plus Ising Gibbs measure on $T$, converges as $l \to \infty$ (in case $i = o$), to the expected magnetization,

$$\rho_\mu := \mathbb{E}_\mu[\nu_{+,T}(x_o)]$$

(4.15)
under the plus measure (and by global sign-reversal, the same applies for the minus measure).

**Lemma 4.5.** Suppose \( \mu \in \mathcal{U} \) and \( \mu(T_*) = 1 \).

(a) If \( \mu = \mu^e \) is ergodic then

\[
\rho_{\mu}^{l,T} \to \rho_{\mu} \text{ as } l \to \infty, \quad \text{for } \mu^e\text{-a.e. } T \in T_*.
\]

(b) For any fixed \( \varepsilon > 0 \),

\[
E_{\mu} \left[ \sum_{i \in \partial o} \mathbb{I} \left( \left| \rho_{\mu}^{l,T} - \rho_{i}^{l,T} \right| > \varepsilon \right) \right] \to 0 \text{ as } l \to \infty.
\]

**Proof:** (a). By definition of \( a_{i,j}^{l,T} \) we have the representation,

\[
\rho_{\mu}^{l,T} = \frac{1}{l} \int_0^l \sum_j \nu_{+}^{\mu,T}(x_j) \mathbb{P}_o(X_t = j) dt = E_{\mu}^{\text{SRW}} \left[ \frac{1}{l} \int_0^l \nu_{+}^{\mu,T}(x_{X_t}) dt \right],
\]

where \( E_{\mu}^{\text{SRW}}[\cdot] \) denotes the expectation with respect to the continuous time SRW starting at \((T,o)\).

In case \( \mu^e \in \mathcal{U} \) supported on \( T_* \) is ergodic, we claim that a.s. with respect to the random \( T \) and the SRW,

\[
\frac{1}{l} \int_0^l \nu_{+}^{\mu,T}(x_{X_t}) dt \longrightarrow \rho_{\mu}, \quad (4.19)
\]

which by (4.18) and DCT for conditional expectation yields the stated \( \mu^e\text{-a.e.} \) convergence (4.16).

Proceeding to prove (4.19), consider the representation (4.7) of the jump times \( \{\tau_m\} \) in terms of the i.i.d. standard Exponential random variables \( \{E_k\} \), independent of the discrete time Markov chain \( \{Y_n\}_{n \geq 0} \) embedded in the continuous time SRW \( \{X_t\}_{t \geq 0} \). Then, with

\[
N_l := \max \left\{ m : \tau_m \leq l \right\}
\]

we have for all \( l > 0 \) the identity

\[
\frac{1}{l} \int_0^l \nu_{+}^{\mu,T}(x_{X_t}) dt = \frac{1}{l} \left[ \sum_{k=1}^{N_l} \nu_{+}^{\mu,T}(x_{Y_{k-1}}) \frac{E_k}{\Delta Y_{k-1}} \right] + \frac{l - \tau_{N_l}}{l} \nu_{+}^{\mu,T}(x_{Y_{N_l}}).
\]

Since \( (Y_n)_{n \geq 0} \) is ergodic for its stationary initial distribution \( \sigma^e \), by Birkhoff’s ergodic theorem we have that, for \( \sigma^e\text{-a.e.} \)

\[
\frac{1}{m} \sum_{k=1}^{m} \frac{1}{\Delta Y_{k-1}} \rightarrow E_{\sigma^e} \left[ \frac{1}{\Delta Y_o} \right] = \frac{1}{\deg(\mu^e)}.
\]

This obviously implies that

\[
\tau_m = \frac{1}{m} \sum_{k=1}^{m} \frac{E_k}{\Delta Y_{k-1}} \xrightarrow{\text{a.s.}} \frac{1}{\deg(\mu^e)}
\]

(see [23, Corollary 3.22]). Noting that \( N_l \to \infty \) it follows from (4.23) by standard renewal theorem techniques that,

\[
\frac{N_l}{l} \xrightarrow{\text{a.s.}} \frac{1}{\deg(\mu^e)}.
\]

Thus, by (4.20) and (4.23),

\[
\frac{\tau_{N_l}}{l} \xrightarrow{\text{a.s.}} 1.
\]
Therefore, upon using the ergodic theorem in combination with [23, Corollary 3.22], we deduce that the left-side of (4.21) converges a.s. to
\[
\overline{\text{deg}}(\mu) \mathbb{E}_{\sigma_{\mu}} \left[ \frac{\nu_{+,T}(x_o)}{\Delta_o} \right] = \mathbb{E}_{\mu} \left[ \nu_{+,T}(x_o) \right],
\]
thereby proving (4.19).

(b). Assuming without loss of generality that \( \overline{\text{deg}}(\mu) > 0 \), let \( \sigma \) denote the size-biased-root measure associated with \( \mu \in \mathcal{U} \) and note that by the triangle inequality, for any \( l \in \mathbb{N}, \varepsilon > 0 \),
\[
\mu(h_l^T) := \mathbb{E}_\mu \left[ \sum_{i \in \partial o} \mathbb{I} \left( |\rho_i^T - \rho_\mu| > 2\varepsilon \right) \right]
\leq \mathbb{E}_\mu \left[ \Delta_o \mathbb{I} \left( |\rho_o^T - \rho_\mu| > \varepsilon \right) \right] + \mathbb{E}_\mu \left[ \sum_{i \in \partial o} \mathbb{I} \left( |\rho_i^T - \rho_\mu| > \varepsilon \right) \right]
= \frac{1}{\overline{\text{deg}}(\mu)} \mathbb{E}_\sigma \left[ \mathbb{I} \left( |\rho_o^T - \rho_\mu| > \varepsilon \right) \right] + \mathbb{E}_\sigma \left[ \frac{1}{\Delta_o} \sum_{i \in \partial o} \mathbb{I} \left( |\rho_i^T - \rho_\mu| > \varepsilon \right) \right]
= \frac{2}{\overline{\text{deg}}(\mu)} \mathbb{P}_\sigma \left( |\rho_o^T - \rho_\mu| > \varepsilon \right). \tag{4.24}
\]

In case \( \mu = \mu^e \) is ergodic, we have, in view of (4.16), the convergence to zero of the bound (4.24). Hence, \( \mu^e(h_l^T) \to 0 \), namely (4.17) holds for ergodic measures. Recall that any fixed \( \mu \in \mathcal{U} \) can be written as a Choquet integral of extremal measures \([4, Lemma 6.8]\), and any extremal measure is ergodic. So, we have a probability measure \( \Theta \) on \( \{ \mu^e : \mu^e \in \mathcal{U} \text{ ergodic} \} \) such that \( \mu(h_l^T) = \int \mu^e(h_l^T) \, d\Theta(\mu^e) \) for all \( l \). The non-negative \( h_l^T \) are bounded by \( \Delta_o \) hence \( 0 \leq \mu^e(h_l^T) \leq \overline{\text{deg}}(\mu^e) \) for all \( l \). Further, \( \int \overline{\text{deg}}(\mu^e) \, d\Theta(\mu^e) = \overline{\text{deg}}(\mu) \) is finite, so by DCT we deduce from the fact that \( \mu^e(h_l^T) \to 0 \) for \( \Theta \)-a.e. \( \mu^e \) that \( \mu(h_l^T) \to 0 \) for all \( \mu \in \mathcal{U} \). \( \square \)

Equipped with Lemma 4.5 we proceed to identify the limit as \( l \to \infty \) of the relevant functionals from Lemma 4.3.

**Lemma 4.6.** Suppose probability measure \( \varpi_+ = \mu \otimes \nu_{+,T} \), with \( T \) distributed according to \( \mu \in \mathcal{U} \) and \( \nu_{+,T} = \alpha_T \nu_{+,T} + (1 - \alpha_T) \nu_{-,T} \) for some fixed, measurable \( \alpha : \mathcal{T}_s \to [0,1] \), with \( \alpha_T = 1 \) whenever \( \nu_{+,T} = \nu_{-,T} \). Then, for any \( \eta > 0 \),
\[
\lim_{l \to \infty} \left| \mathbb{E}_+ \left[ F_{o}^{l,T} \right] - \mathbb{E}_\mu \left[ (1 - \alpha_T) A_{o}^{l,T} \right] \right| = 0. \tag{4.25}
\]
Furthermore, up to at most countably many values of \( \eta > 0 \),
\[
\liminf_{l \to \infty} \mathbb{E}_+ \left[ \sum_{i \in \partial o} \mathbb{I} \left( F_{o}^{l,T} \neq F_{i}^{l,T} \right) \right] = 0. \tag{4.26}
\]

**Remark 4.7.** Recall the branching number of a rooted tree \( T \in \mathcal{T}_s \),
\[
\text{br} T := \left\{ \lambda > 0 : \inf \sum_{j \in \Pi} \lambda^{-|j|} = 0 \right\},
\]
where \( \Pi \subseteq V(T) \) is a cutset (i.e. a finite set of vertices that every infinite path from the root intersects), and \(|j|\) denotes the distance in \( T \) between \( j \) and the root. Our proof of Lemma 4.6 relies on connections between \( \text{br} T \) and recurrence/transience of SRW or phase transitions for Ising models on \( T \) (c.f. [27, 28]).
Proof: Recall that \( A_i^{\uparrow T} = \mathbb{1} (\rho_i^{\uparrow T} \geq 2 \eta) \) and \( \nu, \gamma (F_i^{\uparrow T}) = \mathbb{E}_\mu [ A_i^{\uparrow T} \nu, \gamma (J_i^{\uparrow T}) ] \), for all \( l \in \mathbb{N} \) and \( i \in V(T) \). So, with \( \alpha_T \in [0, 1] \) and \( J_0^{\uparrow T} = J_0^{\uparrow T} \in \{0, 1\} \), we get (4.25) upon showing that \( \mathbb{E}_\mu [D_0^{\uparrow T}] \to 0 \), where

\[
D_i^{\uparrow T} := A_i^{\uparrow T} \max \{ \nu, \gamma (J_i^{\uparrow T}), 1 - \nu, \gamma (J_i^{\uparrow T}) \} .
\]

Further, for any \( T, l, \eta > 0 \) and \( i \in V(T) \),

\[
1 - \nu, \gamma (J_i^{\uparrow T}) = \nu, \gamma (y_i^{\uparrow T}) > -\eta = \nu, \gamma (y_i^{\uparrow T}) < \eta \geq \nu, \gamma (y_i^{\uparrow T}) \leq -\eta = \nu, \gamma (J_i^{\uparrow T}) ,
\]

resulting with

\[
D_i^{\uparrow T} \leq A_i^{\uparrow T} \nu, \gamma (y_i^{\uparrow T}) < \eta := \bar{D}_i^{\uparrow T} (\eta) .
\]

Hence, by Markov’s inequality,

\[
\mathbb{E}_\mu [D_0^{\uparrow T}] \leq \mathbb{E}_\mu [\bar{D}_0^{\uparrow T}] \leq \mathbb{E}_\mu \left[ \nu, \gamma (y_0^{\uparrow T} - \rho_0^{\uparrow T} < -\eta) A_i^{\uparrow T} \right]
\]

\[
\leq \eta^{-2} \mathbb{E}_\mu \left[ \nu, \gamma (y_i^{\uparrow T} A_i^{\uparrow T}) \right] \leq \eta^{-2} \mathbb{E}_\mu \left[ \sum_j \text{Cov}_{\nu, \gamma}(x_o, x_j) \sum_i a_{o,i}^T a_{i,j}^T A_i^{\uparrow T} \right],
\]

with the latter identity obtained by expanding the variance of \( y_0^{\uparrow T} = \sum_j x_j a_{o,j}^T \), then using unimodularity of \( \mu \) as well as \( a_{o,i}^T = a_{i,o}^T \) (by reversibility of the continuous time SRW on \( T \)).

Fixing \( r \in \mathbb{N} \), we partition the sum over \( j \) in the RHS of (4.29) into Term I consisting of sum over all \( j \in B_o(r) \), and Term II for the sum over \( j \notin B_o(r) \). We then control Term II by confirming for \( \gamma := \tan(\beta) \in (0, 1) \) and all \( T \in T_* \) the uniform correlation decay

\[
0 \leq \text{Cov}_{\nu, \gamma}(x_o, x_j) \leq \gamma^{|j|} .
\]

Indeed, it follows from (4.30), by non-negativity of \( \{ a_{i,j}^T \} \) and the fact \( \sum_i a_{o,i}^T a_{i,j}^T = 1 \), that

\[
\text{Term II} \leq \left[ \sum_{k=r+1}^{\infty} \gamma^k \mathbb{E}_\mu \left( \sum_{j \in B_o(k)} \sum_i a_{o,i}^T a_{i,j}^T \right) \right] \leq \gamma^r \mathbb{E}_\mu \left( \sum_{j \notin B_o(r)} \sum_i a_{o,i}^T a_{i,j}^T \right) \leq \gamma^r .
\]

Turning to prove (4.30), note that for any tree \( T \) the marginal of \( \nu, \gamma \) on \( x_T \), with \( T' = (v_0, v_1, \ldots, v_k) \) a finite path in \( T \), is an Ising measure on \( T' \) or in turn a Markov chain of state space \( \{-1, 1\} \) (for finite \( T \) this follows by summation over all possible values of \( x_T \setminus T' \), hence holding also for infinite trees due to (2.5)). While this Markov chain is in general non-homogeneous, recall [8, Lemma 4.1] that for any \( v \neq w \in V(T') \) and Ising measure \( \nu \) on finite \( T' \) with \( \beta \geq 0 \) and any external magnetic field parameters, the value of

\[
\Phi[\nu](v, w) := \nu[x_w|x_v = 1] - \nu[x_w|x_v = -1] = \nu[x_v = 1] - \nu[x_v = -1]^{-1} (\nu[x_v = 1] - \nu[x_w])
\]

is non-negative (by Griffith’s inequality at \( 0 = B_v \leq B_v \uparrow \infty \)), and maximal at the measure \( \nu_f \) of zero external magnetic fields. Now, since \( x_v \in \{-1, 1\} \), we get that

\[
\text{Cov}_{\nu}(x_v, x_w) = 2\nu[x_v = 1]\nu[x_v = -1] = \nu[x_v = -1]^{-1} \Phi[\nu](v, w) \leq \frac{1}{2} \Phi[\nu_f](v, w) = \text{Cov}_{\nu_f}(x_v, x_w) .
\]

The Markov chain corresponding to \( \nu_f \) is homogeneous, of zero-mean and non-degenerate transition probabilities \( \pi(y|x) = \frac{1}{2} (1 + xy \gamma) \) on \( \{-1, 1\} \), from which we get by direct computation that \( \text{Cov}_{\nu_f}(x_v, x_v) = \gamma^k \), and (4.30) follows from (4.32).

As for Term I, recall that if \( \nu, \gamma \langle x_o \rangle = 0 \), then \( \nu, \gamma = \nu, -\gamma \) and \( A_i^{\uparrow T} = 0 \) for all \( i \in V(T) \) and \( l \in \mathbb{N} \). Therefore,

\[
0 \leq \text{Term I} \leq \mathbb{E}_\mu \left[ \sum_{j \in B_o(r)} \sum_i a_{o,i}^T a_{i,j}^T \mathbb{1}\{\nu, \gamma \langle x_o \rangle > 0\} \right]
\]
(the non-negativity of Term I is due to \(\text{Cov}_{\nu_{+,T}}(x_o, x_j) \geq 0\), per (4.30)). It is further known that for Ising model on tree \(T\) with zero external magnetic field, one has \(\nu_{+,T}^\beta(x_o) > 0\) only for \(\beta \geq \beta_c\), where \(|br T| \tanh(\beta_c) = 1\) (see [27, Theorem 1.1]). In particular, we bound \(I\{\nu_{+,T}(x_o) > 0\}\) in (4.33) by \(I\{|br T| > 1\}\), and note that

\[
\sum_{j \in B_o(r)} \sum_i a_{o,i}^T a_{i,j}^T = \sum_{j \in B_o(r)} \frac{1}{2^2} \int_0^1 \int_0^1 \mathbb{P}_o(X_t = i) \mathbb{P}_i(X_s = j) \, dt \, ds
\]

In case \(|br T| > 1\), the discrete time SRW on \(T\) is transient (see [28, Theorem 4.3]). Consequently, for such a tree \(\{X_t\}_{t \geq 0}\) is transient and in particular \(1 \geq \mathbb{P}_o(X_t \in B_o(r)) \rightarrow 0\) as \(t \rightarrow \infty\) for any fixed \(r \in \mathbb{N}\). By bounded convergence it thus follows that Term I goes to zero as \(l \rightarrow \infty\), thereby establishing (4.25).

Next, proceeding to deal with (4.26), for \(\{0, 1\}\)-valued random variables \(A_o = A_o^T, A_i = A_i^T, J_o = J_o^T, J_i = J_i^T\), we clearly have per \(T, l \in \mathbb{N}\) and \(i \in \partial o\), that

\[
A_o A_i \nu_{+,T}(J_o \neq J_i) \leq A_o \nu_{+,T}(J_o) + A_i \nu_{+,T}(J_i)
\]

Consequently, with \(\alpha_T \in [0, 1]\) and each \(F_j = J_j A_j\), we have per \(T, l \eta_0 > 0\) and \(i \in \partial o\) that

\[
\mathcal{F}_T(F_o \neq F_i) \leq I(A_o \neq A_i) + \alpha_T A_o A_i \nu_{+,T}(J_o \neq J_i) + (1 - \alpha_T) A_o A_i \nu_{-,T}(J_o \neq J_i)
\]

for \(D_i\) and \(\tilde{D}_i = \tilde{D}_i^T\) of (4.27) and (4.28), respectively. Taking the expectation with respect to \(T\) of unimodular law \(\mu\) we thus get that,

\[
\mathbb{E}_T + \left[\sum_{i \in \partial o} \mathbb{P}(\mathbb{F}_o^T \neq F_i^T)\right] \leq \mathbb{E}_T + \left[\sum_{i \in \partial o} \mathbb{P}(A_o^T \neq A_i^T)\right] + 2\mathbb{E}_T + \Delta_o \tilde{D}_o^T\]  

While proving (4.25) we saw that \(\mathbb{E}_T[\tilde{D}_o^T] \rightarrow 0\) as \(l \rightarrow \infty\). So, with \(\tilde{D}_o^T \in [0, 1]\) and \(\mathbb{E}_T[\Delta_o]\) finite, by DCT also \(\mathbb{E}_T[\Delta_o \tilde{D}_o^T] \rightarrow 0\). Turning to deal with the other term on the RHS of (4.34), note that for any \(\eta, \varepsilon > 0\), if \(A_o^T(\eta) \neq A_i^T(\eta)\), then either \(|\rho_o^T - \rho_i^T| > \varepsilon\) or \(\rho_o^T \in [2\eta - \varepsilon, 2\eta + \varepsilon]\). Hence, by part (b) of Lemma 4.5, it remains only to verify that

\[
\mathcal{E}(\eta) := \lim_{\varepsilon \rightarrow 0} \inf \lim_{l \rightarrow \infty} \mathbb{E}_T[\Delta_o \rho_o^T(\eta)]\]  

is non-zero for at most countably many values of \(\eta > 0\). Indeed, we get (4.35) by showing that for every positive integer \(K\), the cardinality of the set \(S_K := \{\eta > 0 : \mathcal{E}(\eta) \geq K^{-1}\}\) is at most \(2\mathbb{E}_T[\Delta_o]K\). To this end, define

\[
\mathcal{E}(\eta, \varepsilon, l) := \mathbb{E}_T[\Delta_o \rho_o^T(\eta)] \quad \text{and} \quad \mathcal{E}(\eta, \varepsilon) := \inf_{l \rightarrow \infty} \mathcal{E}(\eta, \varepsilon, l)\]  

Now if possible, let us assume that there exists a positive integer \(K\) such that \(|S_K| > 2\mathbb{E}_T[\Delta_o]K\). Then, there exists \(\{\eta_i\}_{i=1}^{K_0}\) such that \(\min_{i,j=1}^{K_0} \mathcal{E}(\eta_i) \geq K^{-1}\) for some finite \(K_0 > 2\mathbb{E}_T[\Delta_o]K\). From (4.35)-(4.36) we note that there exists \(\varepsilon_0\) small enough with \(0 < \varepsilon_0 < \min_{i,j=1}^{K_0} |\eta_i - \eta_j|\), such that
\[
\min_{i=1}^{K_0} \mathcal{E}(\eta_i, \varepsilon) > (\sqrt{2K})^{-1} \text{ for all } 0 < \varepsilon \leq \varepsilon_0. \text{ By a similar argument, there exists } l_0 = l_0(\varepsilon_0) \text{ such that for all } l \geq l_0, \text{ we have } \min_{i=1}^{K_0} \mathcal{E}(\eta_i, \varepsilon_0, l) > (2K)^{-1}. \text{ Since }
\]
\[
\mathbb{E}_\mu[\Delta_\alpha] < \frac{K_0}{2K} < \sum_{i=1}^{K_0} \mathcal{E}(\eta_i, \varepsilon_0, l) = \mathbb{E}_\mu\left[\Delta_\alpha \mathbb{I}\left(\rho^l_T \in \bigcup_{i=1}^{K_0} [2\eta - \varepsilon_0, 2\eta + \varepsilon_0]\right)\right] \leq \mathbb{E}_\mu[\Delta_\alpha],
\]
we arrive at a contradiction, thereby completing the proof of (4.26).

**Proof of Theorem 1.8.** Recall part (a) of Lemma 3.8 that any sub-sequential local weak limit point \( m_+ \) of \( \{\nu_{n,+}\} \), is effectively a distribution over random \( \alpha : \mathcal{T}_s \mapsto [0, 1] \). From Remark 4.2 we know that to such \( m_+ \) corresponds \( \nu_\alpha = \mu \otimes \nu_\alpha \) with \( \nu_\alpha = \alpha \nu_{+,T} + (1 - \alpha) \nu_{-,T} \) for some fixed measurable \( \alpha : \mathcal{T}_s \mapsto [0, 1] \), where without loss of generality \( \alpha \nu_{+,T} = 1 \) whenever \( \nu_{+,T} = \nu_{-,T} \) (i.e. \( \nu_{+,T}(x_0) = 0 \)), as done in Lemma 4.6. In particular, it suffices to show that the assumed edge-expansion property of \( \{G_n\}_{n \in \mathbb{N}} \)

\[
\mathbb{E}_\mu[(1 - \alpha \nu_0)\mathbb{I}\{\nu_{+,T} \neq \nu_{-,T}\}] = 0,
\]
for then also \( m_+ \)-a.e. \( \alpha \nu_0 = 1 \), as claimed. To this end, recall part (a) of Lemma 4.5, that for any extremal element \( \mu^e \) of \( \mathcal{U} \) and for \( \mu^e \)-a.e. \( \nu_0 \),

\[
\mathbb{E}_\mu[\nu_{+,T}(x_0)] = \lim_{l \to \infty} \rho^l_T.
\]
In particular, setting

\[
\mathcal{J}_\pm := \{ \mathcal{T} : \nu_{+,T} \neq \nu_{-,T}, \liminf_{l \to \infty} \rho^l_T = 0 \},
\]
we have that \( \mu(\mathcal{J}_\pm) = 0 \) for each extremal \( \mu \in \mathcal{U} \) and thus for all \( \mu \in \mathcal{U} \). Consequently, if (4.37) does not hold then

\[
\varepsilon_0 := \mathbb{E}_\mu[(1 - \alpha \nu_0)\mathbb{I}\{\liminf_{l \to \infty} \rho^l_T > 0\}] > 0,
\]
so there exists positive \( \eta_0(\varepsilon_0) \) such that for all \( \eta \in (0, \eta_0(\varepsilon_0)) \),

\[
\mathbb{E}_\mu[(1 - \alpha \nu_0)\mathbb{I}\{\rho^l_T \geq 2\eta \text{ eventually in } l\}] \geq \varepsilon_0/2.
\]
Fix such \( \eta \) small enough to have \( \delta := \eta/(1 + \eta) \leq \varepsilon_0/4 \), and one for which both Lemma 4.3 and Lemma 4.6 hold. For any \( l, t, n \) and \( \varepsilon \), let

\[
S_{n,l,t} := \{ i \in V_n : F^{l,t,n}_i(\varepsilon, \eta) = 1 \},
\]
and \( X_n := n^{-1}\sum_{i,j} F^{l,t,n}_i(X_n, \eta) = 1 \).

Recall Lemma 4.4 that whenever \( \sum_{i,j} x_{j} \geq 0 \)

\[
1 - X_n \geq 1 - \mathbb{E}_{U_n}[J^{l,n}_n] \geq \delta
\]
and since \( \{G_n\}_{n \in \mathbb{N}} \) are \( (\delta, 1/2, \lambda_\delta) \) edge-expanders, we have for such \( \delta \) that

\[
\{X_n \geq \delta \} \implies \frac{1}{n} \sum_{(i,j) \in E_n} \mathbb{I}(F^{l,t,n}_i \neq F^{l,t,n}_j) \geq \lambda_\delta \min \{X_n, 1 - X_n\} \geq \delta \lambda_\delta.
\]
Taking the expectation with respect to \( \nu_{n,+} \) we find that

\[
\frac{1}{2} \mathbb{E}_{U_n}\left[ \sum_{i \in \partial I_n} \nu_{n,+}(F^{l,t,n}_i) \mathbb{I}(F^{l,t,n}_i \neq F^{l,t,n}_j) \right] \geq \delta \lambda_\delta \nu_{n,+}\{X_n \geq \delta \} \geq \delta \lambda_\delta \left( \mathbb{E}_{U_n}[\nu_{n,+}(F^{l,t,n}_i)] - \delta \right),
\]
since \( \mathbb{P}(X \geq \delta) \geq \mathbb{E}[X] - \delta \) for any random variable \( X \leq 1 \) and \( \delta > 0 \). Considering first the limit over the sub-sequence \( n_m \) such that \( \nu_{n_m+} \) converges locally weakly to \( m_+ \), followed by the limit \( t \to \infty \) we deduce by Lemma 4.3, that
\[
\mathbf{\tau}_+ \left[ \sum_{i \in \partial o} I(F_{\nu}^{1,T} \neq F_{\nu}^{l,T}) \right] \geq 2 \delta \lambda \left( \mathbf{\tau}_+ (F_{\nu}^{1,T} - \delta) \right).
\]

Hence, taking \( l \to \infty \), by Lemma 4.6 and Fatou’s lemma we get that,
\[
\delta \geq \mathbb{E}_\mu \left[ (1 - \alpha_T) \liminf_{l \to \infty} A_{\nu}^{l,T} \right] = \mathbb{E}_\mu \left[ (1 - \alpha_T) \mathbb{I} \{ T : \rho_{\nu}^{l,T} \geq 2\delta \text{ eventually in } l \} \right].
\]
That is, we have chosen a positive \( \eta < \eta_0(\varepsilon_0) \) for which (4.39) holds with \( \delta \leq \varepsilon_0/4 \), in contradiction with our assumption (4.38) that (4.37) fails. \( \square \)

5. Continuity of \( U(\cdot, 0) \) in \( \beta \) and edge-expander property

With continuity of \( \beta \mapsto U(\beta, 0) \) at \( \beta < \beta_c \) being a consequence of uniqueness of the corresponding Ising Gibbs measure on \( T \), we prove here such continuity for any \( \mu \in \mathcal{U} \) supported on trees of minimum degree at least three and all \( \beta > \beta_c \), and also at \( \beta = \beta_c \) for all umgw measures, concluding the section with the proof of edge-expander property of the corresponding configuration models.

Lemma 5.1. Suppose \( \mu \in \mathcal{U} \) with \( \mu(T) = 1 \) such that \( \mu \)-a.e. the tree \( T \) has minimum degree at least \( d_* > 2 \) and set \( \beta_* := \text{atanh}(d_* - 1)^{-1} \). Then, \( \beta \mapsto U(\beta, 0) \) is continuous on \( (\beta_c, \infty) \).

In the next lemma we provide sufficient condition for continuity of \( U(\beta, 0) \) at \( \beta = \beta_c \), in case \( \beta_c(T) = \beta_c \) is constant for \( \mu \)-a.e. infinite \( T \).

Lemma 5.2. Suppose \( \mu \in \mathcal{U} \), with \( \mu(T) = 1 \) and \( \beta_c(T) = \beta_c \) finite, for \( \mu \)-a.e. infinite \( T \). If
\[
S_T(t) := \sum_{k=1}^{t} (br T)^{2k} |\partial T(k)|^{-2}
\]
diverges for \( \mu \)-a.e. infinite \( T \), then \( \beta \mapsto U(\beta, 0) \) is continuous at \( \beta = \beta_c \).

Remark 5.3. Same applies if \( |\partial T(k)| \) in (5.1) taken for size of subset of \( \partial T(k) \) connected to \( \partial T(t) \).

We defer the proof of these two lemmas to the sequel, proving first Lemma 1.15 by verifying that umgw measures satisfy the assumptions of Lemma 5.2.

Proof of Lemma 1.15: Since on any finite tree \( T \) there is only one Ising Gibbs measure, \( \beta \mapsto U(\beta, 0) \) is continuous for unimodular measures supported on finite trees. It thus suffices to prove the continuity of \( U(\cdot, 0) \) for super-critical umgw measures conditioned on non-extinction. Hence we merely need to verify the assumptions of Lemma 5.2 for such umgw measures conditioned on non-extinction. To this end, assume first that all entries of the mean matrix \( A_\rho \) of Definition 1.13 are finite.

• Branching number: We need to show that, for super-critical mgw conditioned on non-extinction, \( \beta_c(T) = \beta_c \) for almost every \( T \). By the one to one relation between br \( T \) and \( \beta_c(T) \) (c.f. [27, Theorem 1.1]), it suffices to show that conditioned on non-extinction, br \( T \) is constant umgw-a.e. This follows from having br \( T_{v \to o} \) constant, conditional on non-extinction of \( T_{v \to o} \) for umgw almost every \( T \) and \( v \in \partial o \) (since br \( T = \max_{v \in \partial o} \{br T_{v \to o} \} \), with zero branching number for finite trees and the non-extinction of \( T \) equivalent to non-extinction of some \( T_{v \to o} \)). Each \( T_{v \to o} \) has the same super-critical mgw law corresponding to probability kernels \( \rho_{i,j} \) over the extended type space \( Q_\lambda \), so
our claim follows from [28, Proposition 6.5] which says that for any super-critical, positive regular, non-singular MGW law of finite mean matrix $M$, regardless of the type of its root-vertex, conditional on its non-extinction the branching number of such MGW tree is a.s. the spectral radius $r(M)$ of $M$.

- $S_T$ diverges a.s.: Having finite, positive regular and non-singular mean matrix $A_ρ$, recall the Kesten-Stigum characterization of the a.s. finite limit of $r(A_ρ)^{-k}βT_{v→o}(k)$ conditional on non-extinction of $T_{v→o}$ (generated according to the MGW law with probability kernels $ρ_{i,j}$ and type space $Q_A$, for example, see [25, Theorem 1]). With $Δ_o$ finite a.s., by the preceding argument it follows that $S_T(t) → ∞$ a.s. conditional on non-extinction of the UMGW tree.

Turning to the case where some entry of $A_ρ$ is infinite, consider the following truncation of $ρ_{i,j}$,

$$ρ_{i,j}^\ell(k) := ρ_{i,j}(k)I(||k||≤\ell) + ∑_{||k||>\ell} ρ_{i,j}(k').$$

For all $\ell$ large enough, both positive regularity and non-singularity of $A_ρ$ are inherited by the finite mean matrices $A_ρ^\ell$. Further, positive regularity of the matrix $A_ρ$ having some infinite entries implies that $r(A_ρ^\ell) → ∞$ as $\ell → ∞$. Hence, by the preceding proof, upon choosing $\ell$ large enough, one can make $br T_{v→o}$ under the kernels $ρ_{i,j}^\ell$ uniformly arbitrarily large, conditioned on non-extinction of $T_{v→o}$. Since $br T_{v→o}$ under kernels $ρ_{i,j}^\ell$ is stochastically dominated by that for kernels $ρ_{i,j}$, it follows that conditioned on non-extinction of $T_{v→o}$, almost surely $br T_{v→o} = ∞$. Therefore, a.s. $br T = ∞$ conditional on non-extinction, and all assumptions of Lemma 5.2 are satisfied.

To prove Lemma 5.1 we identify functions $U(β,0) \leq U(β,0)$ that are non-decreasing in $\ell \in N$ and $β > 0$, so the left continuity of $U(β,0)$ follows by interchanging the order of limits in $β$ and $\ell$, provided that

$$U(β,0) = lim_{\ell→∞} U_\ell(β).$$

(5.2)

Indeed, for $T ∈ T_*$, non-negative $β$, $\ell$ and $\{H_v, v ∈ V(T)\}$, consider the Ising model $ν_{T(\ell)}$ of (3.1), for graph $T(\ell)$, inverse temperature parameter $β$ and external field $B_v = H_vI_v∈∂T(\ell)$, with $m_\ell\{H_v\} = ν_{T(\ell)}\{H_v\}\langle x_o \rangle$ denoting its root magnetization. Key to the proof of (5.2) is the joint continuity property (5.3) of $(β,\ell) → m_\ell\{H_v^\ell\}$, where

$$h_v^\ell := atanh(ν_{+T_{v→o}}^\ell\langle x_v \rangle), \ v ∈ V(T)$$

and $T_{v→o}$ denotes the connected component of the sub-tree of $T$ rooted at $v$, after the path between $v$ and $o$ has been deleted (so $T_{o→o} = T$).

**Lemma 5.4.** If $β > β_0$ such that $(d_1 - 1) tanh(β_0) > 1$, then there exists $M = M(β,β_0,d_1)$ finite such that for any $T ∈ T_*$ of minimum degree at least $d_⋆ > 2$ and all $\ell ≥ 1$, $0 ≤ lim_\ell [m_\ell\{h_v^\ell\} - m_\ell\{h_v^{β_0}\}] ≤ M$.

(5.3)

**Proof:** Fixing $β > β_0 > 0$, let $θ := tanh(β) = tanh(β_0)$. Using $v → w$ to denote that $v$ is the parent of $w$ in $T ∈ T_*$, the identity (3.13) becomes

$$h_v^\beta = ∑_{w : v → w} f_θ(h_w^\beta),$$

(5.4)

for $f_θ(h) := atanh(θ tanh(h))$. Since $g : [0,1] → (1,∞)$ given by

$$g(0) = \frac{θ}{θ_0}, \ g(r) = \frac{atanh(θ r)}{atanh(θ_0 r)}, \ ∀r ∈ [0,1],$$


is continuous, necessarily $g(r) \geq 1 + \varepsilon$ for some $\varepsilon = \varepsilon(\beta, \beta_0) > 0$ and all $r \in [0, 1]$. Hence, by Proposition 3.2, Griffith’s inequality and our uniform lower bound on $g(\cdot)$, for any $k \geq 0$ we have
\[
m_{k+1}(\{h^\beta_w\}) = m_k\left(\sum_{w:v \to w} f_\theta(h^\beta_w)\right) = m_k\left(\sum_{w:v \to w} \tanh(h^\beta_w)f_\theta(h^\beta_w)\right)
\geq m_k\left(\sum_{w:v \to w} (1 + \varepsilon)f_\theta(h^\beta_w)\right) = m_k\left((1 + \varepsilon)h^\beta_w\right),
\] (5.5)
with the last equality due to (5.4). The minimum degree of $T$ is at least $d_*$, so we have by Griffith’s inequality that $h^\beta_w \geq h^\beta_*$ for all $w \in V(T)$ and $h^\beta_* := \tanh^{-1}(r_\star^\beta)$ with $r_\star^\beta$ the positive root magnetization for Ising plus measure on the $(d_* - 1)$-ary tree, at parameter $\beta_0$ (which by assumption exceeds the critical parameter for Ising measure on the regular tree $T_{d_*}$). It then follows from (5.4) that moreover $h^\beta_v \geq \xi_\Delta v$, with $\xi := \frac{1}{2}f_\theta(h^\beta_\star)$ strictly positive. Using (5.4) once more, we see that $h^\beta_v \leq f_\theta(1)\Delta v = \beta\Delta v$ for all $v \in V(T)$. Thus, by Griffith’s inequality,
\[
m_{k+1}(\{h^\beta_\star\}) = m_k(\{h^\beta_\star\}) \leq m_k(\{\beta\Delta v\}) \leq m_k(\{(\beta/\xi)h^\beta_\star\}).
\] (5.6)
Choosing $\varepsilon > 0$ small enough, we have $\beta/\xi = 1 + M\varepsilon$ with $M > 1$ finite, hence by the concavity on $\mathbb{R}_+$ of $\lambda \mapsto m_k(\{\lambda H_v\})$, for each $k \geq 0$ and non-negative $\{H_v\}$ (which is a special case of the GHS inequality, see [21]), we get the inequality,
\[
m_k(\{(\beta/\xi)h^\beta_\star\}) - m_k(\{h^\beta_\star\}) \leq M\left[m_k((1 + \varepsilon)h^\beta_v) - m_k(\{h^\beta_v\})\right].
\] (5.7)
Combining (5.5), (5.6) and (5.7) we deduce that
\[
m_{k+1}(\{h^\beta_v\}) - m_{k+1}(\{h^\beta_\star\}) \leq M[m_{k+1}(\{h^\beta_\star\}) - m_k(\{h^\beta_v\})].
\]
Recall, for example from (5.5), that $k \mapsto m_k(\{h^\beta_v\}) \in [0, 1]$ is non-decreasing, and bounded above by $m_k(\{h^\beta_\star\})$ which is independent of $k$. Hence, summing the latter inequality over $k = 0, \ldots, \ell - 1$ results with
\[
0 \leq \ell\left[m_\ell(\{h^\beta_v\}) - m_\ell(\{h^\beta_\star\})\right] \leq \sum_{k=1}^{\ell} [m_k(\{h^\beta_v\}) - m_k(\{h^\beta_\star\})] \leq Mm_\ell(\{h^\beta_\star\}) \leq M,
\] as claimed. \qed

Remark 5.5. Fixing $i \in \partial o$ and keeping same choices of external field, the argument we used in proving Lemma 5.4 also establishes (5.3) when $m_\ell(\cdot)$ is replaced by the Ising root magnetization on $T(\ell) \cap T_{i \to o}$, as well as when it is replaced by the magnetization at $i$ for such Ising models on $T(\ell + 1) \cap T_{i \to o}$. Hereafter, we denote the former by $m_{\ell, i \to o}(\cdot)$ and the latter by $m_{\ell + 1, i \to o}(\cdot)$.

Remark 5.6. For UGW measure $\mu$ the variables $\{h^\beta_v, v \not= o\}$ are identically distributed, each having the law we called $h^\beta_{\star,+}$ in Lemma 1.18. Starting the recursion (1.10) with $h^{(0)} = h^{\beta_{0,+}}$ yields the sequence $h^{(\ell)}$ having the laws of atanh $(m_{\ell, i \to o}(\{h^\beta_\star\}))$. We have just coupled these with atanh $(m_{\ell + 1, i \to o}(\{h^\beta_{\star,+}\}))$ whose law equals $h^\beta_{\star,+}$, establishing the convergence in law of Lemma 1.18 (and by Griffith’s inequality this extends to starting laws which stochastically dominate $h^{\beta_{0,+}}$).

Proof of Lemma 5.1: As mentioned in Remark 1.9, fixing $\beta > \beta_0 > \beta_\star$ it suffices to show that $\cup(\beta, 0)$ is left continuous at $\beta$. To this end, for any infinite $T$ and integer $\ell \geq 1$, using the Ising
model $\nu^β\{h_0^β\}_{T(ℓ)}$ on $T(ℓ)$ with positive external field only at $∂T(ℓ)$, as in Lemma 5.4, we define

$$\mathbb{U}_ℓ(β) = \frac{1}{2} \mathbb{E}_μ \left[ \sum_{i \in ∂o} ν^β\{h_0^β\}_{T(ℓ)} (x_o \cdot x_i) \right].$$  \hspace{1cm} (5.8)

With $T(ℓ)$ a finite graph, fixing $β_0$ and $ℓ$, the function $β \mapsto \mathbb{U}_ℓ(β)$ is continuous and non-decreasing (by Griffith’s inequality). By Proposition 3.2 we further have that

$$\mathbb{U}_{ℓ+1}(β) = \frac{1}{2} \mathbb{E}_μ \left[ \sum_{i \in ∂o} ν^β\{H_r\}_{T(ℓ)} (x_o \cdot x_i) \right],$$

and since $β > β_0$, it follows from (5.4) and the monotonicity of $θ \mapsto f_θ(h)$, that for any $v ∈ ∂T(ℓ)$,

$$H_v := \sum_{\{w: v \rightarrow w\}} f_θ(h_0^β_v) ≥ \sum_{\{v: w \rightarrow v\}} f_θ(h_0^β_w) = h_0^β_v.$$

By yet another appeal to Griffith’s inequality we deduce that $ℓ \mapsto \mathbb{U}_ℓ(β)$ is also non-decreasing. Recall that $h_0^β_v ≥ h_0^β_w$ for all $v ∈ V(T)$, so by similar reasoning, $\mathbb{U}_ℓ(β) ≤ \mathbb{U}(β, 0)$ and as explained before it remains only to establish (5.2). To this end, in view of (3.11), we have that for any $i \in ∂o$ and $\{H_v, v ∈ V(T)\}$,

$$ν^β\{H_v\}_{T(ℓ)} (x_o \cdot x_i) = F(θ, m_{ℓ,i→o}(\{H_v\}); m_{ℓ,o→i}(\{H_v\})),$$

where $F(θ, r)$ of (3.6) is continuous and bounded on $[0, 1]^2$. Thus, with $Ψ(θ, δ) := \sup\{|F(θ, r) − F(θ, r')| \text{ over } r, r' ∈ [0, 1] \text{ such that } |r − r'| ≤ δ\}$ and $δ_ℓ := 2M/(ℓ − 1)$, clearly $Ψ(θ, δ_ℓ) → 0$ as $ℓ → ∞$. Now, in view of Remark 5.5, the expression (5.8) for $\mathbb{U}_ℓ(β)$ and the corresponding expression for $\mathbb{U}(β, 0)$, we deduce that

$$|\mathbb{U}(β, 0) − \mathbb{U}_ℓ(β)| ≤ \frac{1}{2} Ψ(θ, δ_ℓ)deg(μ),$$

from which (5.2) follows.

\[\square\]

**Remark 5.7.** It is easy to see that the proof of Lemma 5.1 applies at any $β ≥ 0$ and $μ ∈ \mathcal{U}$ with $μ(\mathcal{T}_s) = 1$ such that for some $β_0 < β$ one has a bound of the type (5.3). That is, as soon as $m_{ℓ}(\{h_0^β\}) − m_{ℓ}(\{h_0^{β_0}\}) → 0$ in probability, when $ℓ → ∞$. Further, the proof of (5.3) is completely general, except for requiring in (5.6) that $h_0^β_v / h_0^{β_0}$ (alternatively, $r_0^β_v / r_0^{β_0}$), be uniformly bounded over $v ∈ V(T)$. Unfortunately, while $h_0^{β_0}$ is strictly positive as soon as $β_0 > β_c(T)$, even for UGW $μ$, when $β \in (β_c, β_*)$ such ratios may be arbitrarily large (with small $μ$-probability, but nevertheless, they appear at some $v$ and a.e. infinite tree $T$). We did not find a way to by-pass this technical difficulty, hence our requirement of $β > β_*$.\[\] **Remark 5.8.** Lemma 5.4 and Lemma 1.18 are the analogs of [10, Lemma 4.3] and [10, Lemma 2.3], respectively, in case of zero external field and low temperature (i.e. $β > β_*$). While we do not pursue this here, utilizing the former one can establish similar conclusions as done in [10] based on [10, Lemma 2.3 and Lemma 4.3].

The proof of Lemma 5.2 builds on results from [36], to which end we introduce few relevant definitions and notations. First, for any finite $(T, o) ∈ T_s$ let $∂_o T$ denote the collection of rays emanating from $o$, namely finite non-backtracking paths in one-to-one correspondence with the leaves of $T$ other than $o$ (where each such ray terminates). Next, a flow $\varpi$ on such $(T, o)$ is a non-negative function on $E(T)$, of strength $|\varpi| := \sum_{y,o \rightarrow y} \varpi(ay)$, such that $\varpi(vw) = \sum_{y:w \rightarrow y} \varpi(wy)$,
whenever \( v \leftrightarrow w \) and \( w \notin \partial_* T \). Any given collection of resistances \( \{R(e) \geq 0 : e \in E(T)\} \), induces the functional
\[
V_\varpi := \sup \left\{ \sum_{e \in \varpi} (\varpi(e)R(e))^2 : y \in \partial_* T \right\},
\]
over flows \( \varpi \) on \( T \), in terms of which we define
\[
cap_3(T) := \sup\{ |\varpi| : \varpi \text{ a flow on } T \text{ with } V_\varpi = 1 \}.
\]

**Proof of Lemma 5.2:** For any \( (T,o) \in \mathcal{T} \) and \( e = vw \in E(T) \) let \( |e| = |v| \lor |w| \) where \( |v| \) denotes the graph distance between \( v \in V(T) \) and \( o \). From [36, Lemma 4.2] we know that for any \( \theta > 0 \) there exists \( \kappa > 0 \) such that
\[
f_\theta(h) \leq \frac{\theta h}{(1 + (\kappa h)^2)^{1/2}}
\]
for \( f_\theta(\cdot) \) of (5.4) and all \( h \geq 0 \). Futher, recall that for any finite \( t \geq 1 \), \( \theta = \tanh(\beta) > 0 \) and infinite tree \( (T,o) \in \mathcal{T} \) without leaves, the positive
\[
h^{(t)}_o(T) := \text{atanh} \left( \nu^{\beta,0,t}_{+,T(t)v \to o}(x_v) \right),
\]
satisfies the system of equations (5.4) at all \( |v| < t \), starting with \( h^{(t)}_w(T) = \infty \) when \( |w| = t \) (i.e. \( w \in \partial T(t) \)). More generally, in case \( (T,o) \) has leaves, let \( T_t \subseteq T(t) \) denote the union of all vertices and edges along rays of \( T(t) \) of length \( t \), emanating from \( o \). All non-root leaves of \( T_t \) are at distance \( t \) from \( o \) and it is easy to verify that \( h^{(t)}_v(T) = h^{(t)}_v(T_t) \) satisfy for \( v \in T_t \) the corresponding equations (5.4) on \( T_t \), starting with \( h^{(t)}_w(T_t) = \infty \) at \( w \in \partial T(t) \). In view of (5.9), it then follows from [36, Theorem 3.2] that
\[
h^{(t)}_o(T) \leq \kappa^{-1} \text{cap}_3(T_t),
\]
for \( \text{cap}_3(T_t) \) corresponding to resistances \( R(e) = \theta^{-|e|} \) on \( (T_t,o) \). Set \( \theta = \tanh(\beta) \) for \( \beta = \beta_c(T) \) finite, namely \( \theta = 1/(\text{br } T) \) (see [27, Theorem 1.1]). If such \( \text{cap}_3(T_t) \to 0 \) for \( t \to \infty \), then by (5.10) we deduce that
\[
\nu^{\beta,0,t}_{+,T}(x_o) = \lim_{t \to \infty} \text{tanh} \left( h^{(t)}_o(T) \right) \leq 0,
\]
so at \( \beta = \beta_c(T) \) there is then a unique Ising Gibbs measure on \( (T,o) \). Now, should this happen for \( \mu \)-a.e. infinite \( T \) at the same \( \beta_c(T) = \beta_c \), then necessarily \( \mathbb{U}(\beta_c,0) = 0 \) and in particular \( \beta \mapsto \mathbb{U}(\beta,0) \) is continuous at \( \beta = \beta_c \). With \( T_t \subseteq T(T) \), clearly
\[
S_{T_t} := \sum_{k=1}^t \theta^{-2k} |\partial T_t(k)|^{-2} \geq S_T(t)
\]
of (5.1), so it suffices to confirm that \( \text{cap}_3(T_t) \leq S^{-1/2}_{T_t} \) (see also Remark 5.3). To this end, fixing \( t \geq 1 \) let \( \varpi \) be any flow on \( T_t \) of strength \( |\varpi| = 1 \). Then, by the definition of \( V_\varpi \), for any probability measure \( p(\cdot) \) on \( \partial_* T_t \),
\[
V_\varpi \geq \sum_{y \in \partial_* T_t} \left[ \sum_{e \in \varpi} \varpi^2(e)\theta^{-2|e|} \right] p(y) = \sum_{k=1}^t \theta^{-2k} \sum_{|e|=k} \varpi^2(e) \sum_{y \in e} p(y).
\]
With slight abuse of notation, set \( p(e) := \sum_{y \in e} p(y) \). Note that the thus defined \( \{p(e), e \in E(T_t)\} \), constitutes a flow of strength \( |p| = 1 \). Further, \( \sum_{|e|=k} p(e) = 1 \) for any \( 1 \leq k \leq t \) since all non-root leaves of \( T_t \) are at \( \partial T_t(t) \). Applying Cauchy-Schwartz inequality and choosing \( p = \varpi \), we find that
\[
\left[ \sum_{|e|=k} \varpi^2(e)p(e) \right] \geq \left( \sum_{|e|=k} \varpi(e)p(e) \right)^2 = \left( \sum_{|e|=k} \varpi^2(e) \right)^2.
\]
Using Cauchy-Schwartz inequality once more,
\[
\left( \sum_{|e|=k} \varpi^2(e) \right) \geq \frac{1}{|\partial T_t(k)|} \left( \sum_{|e|=k} \varpi(e) \right)^2 = \frac{1}{|\partial T_t(k)|}.
\]  
(5.13)

Thus, from (5.11), (5.12) and (5.13), we see that \( V_\varpi \geq S_{T_t} \) for any flow \( \varpi \) on \( T_t \) such that \( \varpi = 1 \). By simple scaling, it then follows that \( \text{cap}_3(T_t) \leq S_{T_t}^{-1/2} \), as claimed. \( \square \)

**Proof of Lemma 1.17:** For each \( i \in Q \) and \( k \in \mathbb{Z}_{\geq}^{|Q|} \) let \( \alpha_{i,k} = \theta(i) P_{i,k} \), viewed as coordinates of the collection
\[
\alpha = (\alpha_{i,k})_{i \in Q, k \in \mathbb{Z}_{\geq}^{|Q|}}
\]
(which is finite by assumption of bounded support for all \( P_{i,k}, i \in Q \)). From Definition 1.12 it follows that for large \( n \) and all \( i, k \) there are with high probability \( n \alpha_{i,k} (1 + o(1)) \) vertices of type \( i \in Q \) and off-springs configuration \( k \) in the random graph \( G_n \). Fix \( \delta_0 \leq 1/2 \) and arbitrary vector \( \delta = (\delta_{i,k})_{i \in Q, k \in \mathbb{Z}_{\geq}^{|Q|}} \) such that \( \|\delta\| \in (\delta_0, 1/2) \). Let \( U_\delta \) denote a subset of \( n \|\delta\| (1 + o(1)) \) vertices from \( V_n \) such that for each \( i \) and \( k \) about \( n \delta_{i,k} (1 + o(1)) \) of the vertices of \( U_\delta \) are of type \( i \) and off-springs configuration \( k \). Let \( S_\delta \) denote the set of possible realizations of \( G_n \) for which there exists some \( U_\delta \) with \( \varepsilon(1 + o(1)) \) edges connecting \( U_\delta \) to \( U_\delta \). It suffices to show the existence of \( \varepsilon_0 := \varepsilon(\delta_0) > 0 \) such that for all \( \delta \in (\delta_0, 1/2) \) and \( \varepsilon \leq \varepsilon_0 \),
\[
\frac{1}{n} \log \mathbb{P}(S_\delta^\varepsilon) < -\varepsilon < 0,
\]  
(5.14)

for all large \( n \), uniformly over all choices of \( \delta \) and \( \varepsilon \). To this end, we first note that for \( \varepsilon = 0 \),
\[
\frac{1}{n} \log \mathbb{P}(S_\delta^0) = \frac{1}{n} \log \# \left\{ \text{choices possible for } U_\delta \right\} + \frac{1}{n} \log \mathbb{P}\left\{ \text{such choice matches with itself} \right\}
\]
\[
:= N_\delta + Q_\delta.
\]

We further define \( \hat{\alpha} := (\hat{\alpha}_{i,j})_{i,j \in Q} \), where \( \hat{\alpha}_{i,j} := \sum_k k_j \alpha_{i,k} \) for each \( i, j \in Q \), and let \( \hat{\delta} \) be similarly defined. Using the approximations,
\[
\frac{1}{n} \log n! = \log \left( \frac{n}{e} \right) + o(1) \text{ and } \frac{1}{n} \log n!! = \frac{1}{2} \log \left( \frac{n}{e} \right) + o(1),
\]
we have for \( H(p) := -p \log p - (1 - p) \log(1 - p), p \in [0,1] \), that
\[
N_\delta \approx \sum_{i,k} \alpha_{i,k} H\left( \frac{\delta_{i,k}}{\alpha_{i,k}} \right) = \sum_i \sum_j \sum_k \frac{k_j \alpha_{i,k}}{\|k\|} H\left( \frac{\delta_{i,k}}{\alpha_{i,k}} \right)
\]  
(5.15)

\[
Q_\delta \approx -\frac{1}{2} \sum_{i \in Q} \hat{\alpha}_{i,i} H\left( \frac{\hat{\delta}_{i,i}}{\hat{\alpha}_{i,i}} \right) - \sum_{i \neq j \in Q} \hat{\alpha}_{i,j} H\left( \frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}} \right).
\]  
(5.16)

By concavity of \( H(\cdot) \), upon noting that \( \|k\| \geq 3 \) we have for any \( i, j \in Q \), that
\[
\sum_k \frac{k_j \alpha_{i,k}}{\|k\|} H\left( \frac{\delta_{i,k}}{\alpha_{i,k}} \right) - \frac{1}{2} \hat{\alpha}_{i,j} H\left( \frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}} \right) \leq \frac{1}{3} \sum_k k_j \alpha_{i,k} H\left( \frac{\delta_{i,k}}{\alpha_{i,k}} \right) - \frac{1}{2} \hat{\alpha}_{i,j} H\left( \frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}} \right) \leq -\frac{1}{6} \hat{\alpha}_{i,j} H\left( \frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}} \right).
\]

Since \( \delta_{i,k} \leq \alpha_{i,k} \) for each \( i \in Q, k \in \mathbb{Z}_{\geq}^{|Q|} \) and \( \|\delta\| \leq 1/2 < \|\alpha\| = 1 \), we must have \( \hat{\delta}_{i,j} \leq \hat{\alpha}_{i,j} \) for at least one pair \((i,j)\). We thus get from (5.15) and (5.16) that
\[
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(S_\delta^\varepsilon) \leq -\frac{1}{6} \sum_{i,j \in Q} \hat{\alpha}_{i,j} H\left( \frac{\hat{\delta}_{i,j}}{\hat{\alpha}_{i,j}} \right),
\]  
(5.17)
with the rhs strictly negative (since $H(p) = 0$ only for $p \in \{0, 1\}$). Further, the approximations in (5.15) and (5.16) are uniform over $\delta$, because
\[
\sqrt{2\pi} \leq \frac{n!}{n^{n+1/2}e^{-n}} \leq e, \quad \text{for all } n.
\] (5.18)

The supremum of the upper bound of (5.17), over the compact set of all possible choices of $\delta$ is strictly negative, yielding (5.14) for $\varepsilon = 0$. Similar rational applies also for all $\varepsilon$ small enough. For example, in case $|Q| = 1$ we have that
\[
\limsup_{n \to \infty} \frac{1}{n} \log P(S_\delta^\varepsilon) \leq -\frac{\hat{\alpha}}{6} H\left(\frac{\delta}{\hat{\alpha}}\right) + \frac{\delta}{2} H\left(\frac{\varepsilon}{\delta}\right) + \frac{1}{2}(\hat{\alpha} - \delta) H\left(\frac{\varepsilon}{\hat{\alpha} - \delta}\right) \leq -\frac{\hat{\alpha}}{6} H\left(\frac{\hat{\delta}}{\hat{\alpha}}\right) + \frac{\hat{\alpha}}{2} H\left(\frac{2\varepsilon}{\hat{\alpha}}\right).
\]

The preceding bound is continuous in $\varepsilon$ and strictly negative at $\varepsilon = 0$. Consequently, there exists $\varepsilon_0 > 0$ small enough such that this bound is strictly negative at all $\varepsilon \leq \varepsilon_0$. Further, from (5.18) we get uniformity of the convergence in $n$, over all relevant $\delta$ and $\varepsilon \leq \varepsilon_0$, yielding (5.14) in case $|Q| = 1$. While we do not detail these, the computations in case $|Q| > 1$ and $\varepsilon > 0$ are similar. \qed

\begin{thebibliography}{99}

[1] M. Aizenman, Translation invariance and instability of phase coexistence in the two-dimensional Ising system. \textit{Comm. Math. Phys.}, 73, 839-944, 1980.

[2] M. Aizenman, D. Barsky, and R. Fernández. The phase transition in a general class of Ising-type models is sharp. \textit{J. Stat. Phys.} 47, 343-374, 1987.

[3] M. Aizenman, J. Wehr. Rounding of first-order phase transitions in systems with quenched disorder. \textit{Communications in Mathematical Physics}, 130 489-530, 1990.

[4] D. Aldous and R. Lyons. Processes on unimodular random networks. \textit{Electron. J. Probab.}, 12(54):1454–1508, 2007.

[5] D. Aldous and J. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. in H. Kesten, editor, \textit{Probability on Discrete Structures}, volume 110 of \textit{Encyclopedia Math Sci.}, pages 1-72. Springer, Berlin. Probability Theory, 1.

[6] K. B. Athreya, and P. E. Ney. \textit{Branching Process}. Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].

[7] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. \textit{Electron. J. Probab.}, 6(23):13 pp. (electronic), 2001.

[8] N. Berger, C. Kenyon, E. Mossel, and Y. Peres. Glauber dynamics on Trees and Hyperbolic graphs. \textit{Probab. Theory Related Fields}, Vol. 131 No. 3, 311-340, 2005.

[9] T. Bodineau. Translation invariant Gibbs states for the Ising model. \textit{Probability Theory and Related Fields}, 135, 153-186, 2006.

[10] A. Dembo, and A. Montanari. Ising models on locally tree-like graphs. \textit{Ann. Appl. Probab.}, 20(2):565–592, 2010.

[11] A. Dembo, and A. Montanari. Gibbs measures and phase transitions on random graphs. \textit{Brazilian Journal of Probability and Statistics}, 24(2): 137-211, 2010.

[12] A. Dembo, A. Montanari, and N. Sun. Factor models on locally tree like graphs. \textit{Ann. Probab.}, (to appear), 2013.

[13] A. Dembo, and N. Sun. Central limit theorem for biased random walk on Galton-Watson trees. \textit{Elect. J. Probab.}, 17, 75, 1–40, 2012.

[14] R. Dobrushin, and S. Shlosman. The problem of translation invariance of Gibbs states at low temperatures. \textit{Mathematical physics reviews}, 5, 53195, 1985.

[15] S. Dommers, C. Giardinà, and R. van der Hofstad. Ising models on power-law random graphs. \textit{J. Stat. Phys.}, 141(4):638–660, 2010.

[16] F. J. Dyson. Existence of a phase-transition in a one-dimensional Ising ferromagnet. \textit{Comm. Math. Phys.}, 12(2) :91-107, 1969.

[17] R. S. Ellis, and C. M. Newman. The statistics of Curie-Weiss models. \textit{Journal of Statistical Physics}, 19, 149-161, 1978.

[18] H. Georgii. \textit{Gibbs measures and phase transitions}, volume 9 of \textit{de Gruyter Studies in Mathematics}. Walter de Gruyter & Co., Berlin, 1988.
\end{thebibliography}
[19] H. O. Georgii, and Y. Higuchi. Percolation and number of phases in the two-dimensional Ising model. *Journal of Mathematical Physics*, 41, 11531169, 2000.

[20] A. Greschenfield, and A. Montanari. Reconstruction for models on random graphs. *48th Annual Symposiums of Foundations of Computer Science*, Providence, RI, 2007.

[21] R. B. Griffiths, C. A. Hurst, and S. Sherman. Concavity of magnetization of an Ising ferromagnet in a positive external field. *J. Math. Phys.*, 11, 790-795, 1970.

[22] G. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[23] O. Kallenberg. *Foundations of modern probability*. 2nd edition, Springer Series in Statistics, 2002.

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

[25] T. G. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the Kesten-Stigum theorem for multitype branching processes. *Classical and Modern Branching Processes*, Springer, New York, 1997, pp. 181-185.

[26] T. M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.

[27] R. Lyons. The Ising model and percolation on trees and tree-like graphs. *Commun. Math. Phys.*, 125(2), 337-353, 1989.

[28] R. Lyons. Random walk and percolation on trees. *Ann. Probab.*, 18(3), 931-958, 1990.

[29] R. Lyons, R. Pemantle, and Y. Peres. Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems*, 15, 593-619, 1995.

[30] R. Lyons, and Y. Peres. *Probability on Trees and Networks*. Preprint.

[31] M. Mézard, and A. Montanari. *Information, physics, and computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009.

[32] A. Montanari, E. Mossel, and A. Sly. The weak limit of Ising models on locally tree-like graphs. *Probab. Theory Related Fields*, 152, 31-51, 2012.

[33] M. E. J. Newmann. The structure and function of complex networks. *SIAM Review*, 45(2): 167-256, 2003.

[34] C. M. Newman, and D. L. Stein. Spatial inhomogeneity and thermodynamic chaos. *Physical Review letters*, 76, 4821-4824, 1996.

[35] M. Niss. History of Lenz-Ising model 1920-1950: from ferromagnetic to cooperative phenomena. *Archive for History of Exact Sciences*, 59(3): 267-318, 2005.

[36] R. Pemantle, and Y. Peres. The critical Ising model on trees, concave recursions and nonlinear capacity. *Ann. Probab.*, 38, 184-206, 2010.

[37] L. De Sanctis, and F. Guerra. Mean field dilute ferromagnet: high temperature and zero temperature behavior. *Journal of Statistical Physics*, 132: 759-785, 2008.

[38] D. W. Stroock. Probability Theory, an analytic view. *Cambridge University Press*, 1994.

†Department of Mathematics, Stanford University
Building 380, Sloan Hall, Stanford, California 94305

∗†Department of Statistics, Stanford University
Sequoia Hall, 390 Serra Mall, Stanford, California 94305