The Annals of Probability
2017, Vol. 45, No. 2, 780–823
DOI: 10.1214/15-AOP1075

FERROMAGNETIC ISING MEASURES ON LARGE LOCALLY TREE-LIKE GRAPHS

BY ANIRBAN BASAK AND AMIR DEMBO[1]

Duke University and Stanford University

We consider the ferromagnetic Ising model on a sequence of graphs G_n converging locally weakly to a rooted random tree. Generalizing [*Probab. Theory Related Fields* **152** (2012) 31–51], 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 countable many choices of β, 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 $\mathsf{G} = (V, E)$ is the probability distribution over $\underline{x} = \{x_i : i \in V\}$ with $x_i \in \{-1, +1\}$, for some $\beta \ge 0$ (*inverse temperature* parameter), $B \in \mathbb{R}$ (*external magnetic field*), given by

$$\nu_{\mathsf{G}}^{\beta,B}(\underline{x}) = \frac{1}{Z_{\mathsf{G}}(\beta, B)} \exp\left\{\beta \sum_{(i,j)\in E} x_i x_j + B \sum_{i\in V} x_i\right\}, \tag{1.1}$$

where $Z_{\mathsf{G}}(\beta, B)$ is the normalizing constant (also known as *partition function*).

The Ising model is a paradigm model in statistical physics [31] with much recent interest also in the Ising model on *nonlattice complex networks* (see [30], and the references therein). In this paper, we focus on sparse graph sequences $\{\mathsf{G}_n\}_{n\in\mathbb{N}}$ 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 (cf. [10, 27]). The key to such studies is the asymptotics of the log partition function, appropriately scaled, as derived, for example, in [9, 12, 18]. In particular, [11] shows that for any sequence of graphs $\mathsf{G}_n = (V_n, E_n)$, with V_n of size n, that converges locally weakly

Received December 2013; revised October 2015.

[1]Supported in part by NSF Grant DMS-11-06627.

MSC2010 subject classifications. 05C05, 05C80, 05C81, 60J80, 82B20, 82B26.

Key words and phrases. Ising model, random sparse graphs, Gibbs measures, local weak convergence.

to random trees, the asymptotic *free entropy density* of the ferromagnetic Ising models

$$\nu_n^{\beta,B}(\underline{x}) = \frac{1}{Z_n(\beta, B)} \exp\left\{\beta \sum_{(i,j)\in E_n} x_i x_j + B \sum_{i\in V_n} x_i\right\}, \tag{1.2}$$

exists, that is,

$$\phi(\beta, B) := \lim_{n\to\infty} \phi_n(\beta, B), \tag{1.3}$$

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 β and $B = 0$ decomposes into convex combination of well-separated simple components. This has been shown for the complete graph [15], and for grids [1, 8, 13, 17].

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 [28]. 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 $\nu_n^{\beta,B}(\cdot)$ converges to the marginal distribution on a neighborhood of the root for some Ising–Gibbs measure on the limiting tree T. Denoting by $\nu_{\pm,\mathsf{T}}^{\beta,B}$ the Ising–Gibbs measures on T, corresponding to plus and minus boundary conditions, for $B > 0$ it easily follows from [11] that the limiting measure is given by first picking the random tree T, and then conditioned on T, using the Ising–Gibbs measure $\nu_{+,\mathsf{T}}^{\beta,B}$ (the same applies for $B < 0$ with $\nu_{+,\mathsf{T}}^{\beta,B}$ is replaced by $\nu_{-,\mathsf{T}}^{\beta,B}$). Recall that for $B = 0$ and β 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 [28], for k-regular trees, the plus/minus boundary conditions play a special role. Indeed, it is shown in [28] that if G_n's converge locally weakly to k-regular trees $\mathsf{T} = \mathsf{T}_k$ then, for any $\beta > 0$ and $B = 0$,

$$\nu_n^{\beta,0}(\cdot) \to \tfrac{1}{2}\nu_{+,\mathsf{T}}^{\beta,0}(\cdot) + \tfrac{1}{2}\nu_{-,\mathsf{T}}^{\beta,0}(\cdot). \tag{1.4}$$

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

$$\nu_{n,\pm}^{\beta,0}(\cdot) \to \nu_{\pm,\mathsf{T}}^{\beta,0}(\cdot), \tag{1.5}$$

where $\nu_{n,+}^{\beta,0}(\cdot)$ and $\nu_{n,-}^{\beta,0}(\cdot)$ are the measures (1.2) 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 (1.4). For genuinely random limiting trees, one expects (1.4) and (1.5) to apply where now T is chosen according to the limiting tree measure. As we focus on the case $B = 0$, hereafter we write $\nu_n^\beta(\cdot) := \nu_n^{\beta,0}(\cdot)$ and adopt the convention of using $\nu_n^B(\cdot)$ (or just ν_n, in case $B = 0$), when the value of β is either arbitrary, or clear from the context. Similar notation apply for Ising measures on the limiting trees.

It is well known (see [25]) that there exists a value of β, denoted here by β_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 (1.4) and (1.5) in [28], 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 (1.4) in [28] is the continuity, for k-regular infinite trees, of root magnetization under $\nu_{+,\mathsf{T}_k}(\cdot)$, 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 $d_\star > 2$, we prove here the continuity of root magnetization under $\nu_{+,\mathsf{T}}(\cdot)$ for all $\beta > \operatorname{atanh}[(d_\star - 1)^{-1}] > \beta_c$ (see Section 5).[2] The proof of (1.5) relies on choosing functionals $\bar{\mathsf{J}}^l(\cdot)$ 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 $\bar{\mathsf{J}}^l(\cdot)$ defined via *average occupation measure* of the variable speed continuous time simple random walk (VSRW) on the tree, we show here that (1.5) 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 $\beta = \beta_c$ for *multi-type 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 [20] 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 [16], Chapter 12), for large β, $B = 0$ 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, for example, the Potts model.

[2] For $\beta = \beta_c$ one may use the equivalent capacity criterion provided in [32].

1.1. *Graph preliminaries and local weak convergence.* In a connected undirected graph $\mathsf{G} = (V, E)$, 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 $\mathsf{B}_v(r)$ the ball of radius r around v, that is, the collection of all vertices whose distance from v in G is at most r. The set $\mathsf{B}_v(1) \setminus \{v\}$ of all vertices adjacent to v is also denoted by ∂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* $(\overline{\mathsf{G}}, o)$ is a rooted graph (G, o) with vector $\underline{x}_{\mathsf{G}}$ 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 $[\mathsf{G}, o]$ denoting the collection of all rooted graphs that are isomorphic to (G, o) (and $[\overline{\mathsf{G}}, o]$ denoting the collection of all rooted networks isomorphic to $(\overline{\mathsf{G}}, o)$).

Let $\mathcal{G}_*$ be the space of rooted isomorphism classes of rooted *connected* locally finite graphs. Similarly, for rooted networks let $\overline{\mathcal{G}}_*$ denote the space of rooted isomorphism classes of rooted connected locally finite networks. Setting the distance between $[\mathsf{G}_1, o_1]$ and $[\mathsf{G}_2, o_2]$ (and the same between $[\overline{\mathsf{G}}_1, o_1]$ and $[\overline{\mathsf{G}}_2, o_2]$) to be $1/(\alpha + 1)$, where α 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 $\mathcal{G}_*$ and $\overline{\mathcal{G}}_*$ which are complete separable metric spaces (see [4, 6]). We use hereafter this metric topology, denoting by $\mathscr{C}_{\mathcal{G}_*}$ and $\mathscr{C}_{\overline{\mathcal{G}}_*}$ the corresponding Borel σ-algebras on $\mathcal{G}_*$ and $\overline{\mathcal{G}}_*$, respectively [but forgo the conversion $r \mapsto 1/(r + 1)$, letting $\mathsf{B}_{\mathsf{G}}(r)$ stand throughout for the $\mathcal{G}_*$-metric ball of radius $1/(r+1)$ around G, namely those rooted graphs (G', o') having $\mathsf{B}_{o'}(r)$ isomorphic to $\mathsf{B}_o(r) \subset \mathsf{G}$]. Similarly, we equip the spaces $\mathcal{T}_*$ and $\overline{\mathcal{T}}_*$ of all rooted isomorphism classes of locally finite trees (and marked trees, resp.), with the metric topology and Borel σ-algebra induced by $\mathcal{G}$ and $\overline{\mathcal{G}}_*$, respectively [while using as before $\mathsf{B}_{\overline{\mathsf{G}}}(r)$, $\mathsf{B}_{\mathsf{T}}(r)$ and $\mathsf{B}_{\overline{\mathsf{T}}}(r)$ for the metric balls of radius $1/(r + 1)$ in $\overline{\mathcal{G}}_*$, $\mathcal{T}_*$ and $\overline{\mathcal{T}}_*$, resp.].

DEFINITION 1.1. For ζ_n and μ Borel probability measures on $\mathcal{G}_*$ (or $\overline{\mathcal{G}}_*$), we write $\zeta_n \Rightarrow \mu$ when ζ_n converges weakly to μ with respect to the metric on $\mathcal{G}_*$ (or $\overline{\mathcal{G}}_*$) and for any $\mathsf{G} \in \mathcal{G}_*$ we denote by δ_{G} the probability measure on $\mathcal{G}_*$ assigning point mass at G.

For probability measure ν on $(\mathcal{X}_1, \mathcal{B}_1)$ and measurable map $f : (\mathcal{X}_1, \mathcal{B}_1) \mapsto (\mathcal{X}_2, \mathcal{B}_2)$, we let $\nu \circ f^{-1}$ denote the probability measure on $(\mathcal{X}_2, \mathcal{B}_2)$ such that $\nu \circ f^{-1}(\cdot) = \nu(f^{-1}(\cdot))$, and in case f is real-valued, use the shorthand $\nu[f]$ or $\nu\langle f\rangle$ for the ν-expected value of f (i.e., $\int f\,d\nu$), using also $\langle f\rangle$ when the choice of ν is clear form the context. Equipped with these notation, we proceed to define the *local weak convergence* of graphs.

DEFINITION 1.2. For a sequence of graphs $\{\mathsf{G}_n\}_{n\in\mathbb{N}}$ having vertex sets $[n]$, let μ_n denote the law of (G_n, I_n) in $\mathcal{G}_*$ for I_n chosen uniformly over $[n] := \{1, 2, \ldots, n\}$. We call such $\{\mathsf{G}_n\}$ *uniformly sparse*, if Δ_o is uniformly integrable under $\{\mu_n\}$. That is, if

$$\lim_{k\to\infty} \limsup_{n\to\infty} \frac{1}{n} \sum_{i\in[n]} \Delta_i(\mathsf{G}_n)\mathbb{I}\big(\Delta_i(\mathsf{G}_n) \geq k\big) = 0. \tag{1.6}$$

If in addition $\mu_n \Rightarrow \mu$, a probability measure on $\mathcal{G}_*$, we say that the uniformly sparse collection $\{\mathsf{G}_n\}$ converges locally weakly to μ, denoted by $\mathsf{G}_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$. In particular, due to uniform sparseness $\mu\langle\Delta_o\rangle$ is finite for any such limit.

Similarly to the space $\mathcal{G}_*$, one defines $\mathcal{G}_{**}$ 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 $\mathcal{G}_{**}$ is written as $f(\mathsf{G}, x, y)$, to indicate the distinguished pair of vertices (x, y). In [6], it is shown that any LWC limit point must be *involution invariant*, a property that was found in [3] to be equivalent to the following property of *unimodularity*.

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

$$\int \sum_{x\in V(\mathsf{G})} f(\mathsf{G}, o, x)\, d\mu\big([\mathsf{G}, o]\big) = \int \sum_{x\in V(\mathsf{G})} f(\mathsf{G}, x, o)\, d\mu\big([\mathsf{G}, o]\big). \tag{1.7}$$

We denote by $\mathcal{U}$ the collection of all unimodular probability measures μ on $\mathcal{G}_*$ for which $\mu\langle\Delta_o\rangle$ is finite and by $\mathcal{U}_*$ those $\mu \in \mathcal{U}$ having $\mu(\mathcal{T}_*) = 1$.

We consider throughout *tree-like graphs*, namely $\mathsf{G}_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$ with a limiting object which is a (random) tree, namely having $\mu \in \mathcal{U}_*$. This assumption, and the fact that any LWC limit points is in $\mathcal{U}$ are both key for our results, with (1.7) being utilized in several proofs.

1.2. *Local weak convergence of Ising measures*. The space of all probability measures on $(\overline{\mathcal{G}}_*, \mathscr{C}_{\overline{\mathcal{G}}_*})$ will be denoted by $\mathcal{P}(\overline{\mathcal{G}}_*)$. For example, upon choosing a root, Ising measures on connected, locally finite graphs can be considered elements of $\mathcal{P}(\overline{\mathcal{G}}_*)$. Considering the $\mathcal{G}_*$-projection $[\overline{\mathsf{G}}, o] \mapsto [\mathsf{G}, o]$ from rooted networks in $\overline{\mathcal{G}}_*$ to rooted graphs in $\mathcal{G}_*$, we let $\mu \otimes \nu_{\mathsf{G}}$ denote an element of $\mathcal{P}(\overline{\mathcal{G}}_*)$, whose marginal distribution on $\mathcal{G}_*$ is $\mu \in \mathcal{P}(\mathcal{G}_*)$, and given any fixed $\mathsf{G} \in \mathcal{G}_*$ has the (conditional) distribution ν_{G} on the corresponding mark space $\mathcal{X}^{\mathsf{G}}$.

For any positive integer t, the subgraph $(\mathsf{G}, o)(t)$ of (G, o) induced by the vertices $\mathsf{B}_o(t)$, is called the graph truncated at *height* t, with the corresponding definition for a rooted network. We further use the notation $\mathsf{G}(t)$ and $\overline{\mathsf{G}}(t)$, when the choice of root is clear from the context. For example, $\mathsf{T}(t)$ 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 $\overline{\mathcal{G}}_*(t)$ denote the space of rooted isomorphism classes of rooted connected locally finite networks truncated at height t, with $\mathscr{C}_{\overline{\mathcal{G}}_*(t)}$ the corresponding Borel σ-algebra, yielding for each $\overline{\nu} \in \mathcal{P}(\overline{\mathcal{G}}_*)$ the probability measure $\overline{\nu}^t$ induced on $(\overline{\mathcal{G}}_*(t), \mathscr{C}_{\overline{\mathcal{G}}_*(t)})$ by such truncation (of the network), and for each probability measure $\overline{\mathfrak{m}}$ *on* $\mathcal{P}(\overline{\mathcal{G}}_*)$ the correspondingly induced probability measure $\overline{\mathfrak{m}}^t$ on $\mathcal{P}(\overline{\mathcal{G}}_*(t), \mathscr{C}_{\overline{\mathcal{G}}_*(t)})$.

We next adapt [28], Definition 2.3, to the case of nondeterministic graph limits.

DEFINITION 1.4. Given a sequence of graphs $\{\mathsf{G}_n\}_{n\in\mathbb{N}}$ having vertex sets $[n]$, and probability measures ζ_n on $\mathcal{X}^{V_n}$, for any positive integer t let $\overline{\mathsf{P}}_n^t(i) \in \mathcal{P}(\overline{\mathcal{G}}_*(t), \mathscr{C}_{\overline{\mathcal{G}}_*(t)})$ denote the law of the pair $((\mathsf{B}_i(t), i), \underline{x}_{\mathsf{B}_i(t)})$ for $\underline{x}$ drawn according to ζ_n and $i \in [n]$ some vertex of G_n.

When combined with the uniform measure over the choice of random vertex $I_n \in [n]$, this results with the random distributions $\overline{\mathsf{P}}_n^t(I_n)$, and we say that $\{(\mathsf{G}_n, \zeta_n)\}_{n\in\mathbb{N}}$ (or in short $\{\zeta_n\}$), converges *locally weakly* to a probability measure $\overline{\mathfrak{m}}$ on $\mathcal{P}(\overline{\mathcal{G}}_*)$, if the law of $\overline{\mathsf{P}}_n^t(I_n)$ converges weakly to $\overline{\mathfrak{m}}^t$, 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 [2, 22, 29]).

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

DEFINITION 1.5. For each t, consider the following Ising measures on $\mathsf{T}(t)$:

$$\nu_{+,\mathsf{T}}^{\beta,B,t}(\underline{x}) := \frac{1}{Z^{t,+}} \exp\Biggl\{\beta \sum_{(i,j)\in E(\mathsf{T}(t))} x_i x_j + B \sum_{i\in V(\mathsf{T}(t))} x_i\Biggr\} \mathbb{I}\bigl(\underline{x}_{\mathsf{T}\setminus\mathsf{T}(t-1)} = (+)_{\mathsf{T}\setminus\mathsf{T}(t-1)}\bigr),$$

$$\nu_{-,\mathsf{T}}^{\beta,B,t}(\underline{x}) := \frac{1}{Z^{t,-}} \exp\Biggl\{\beta \sum_{(i,j)\in E(\mathsf{T}(t))} x_i x_j + B \sum_{i\in V(\mathsf{T}(t))} x_i\Biggr\} \mathbb{I}\bigl(\underline{x}_{\mathsf{T}\setminus\mathsf{T}(t-1)} = (-)_{\mathsf{T}\setminus\mathsf{T}(t-1)}\bigr),$$

where for any $W \subseteq V(\mathsf{T})$, 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 $\nu_{+,\mathsf{T}}^{\beta,B,t}$ and $\nu_{-,\mathsf{T}}^{\beta,B,t}$ converge to probability measures on $\{-1,+1\}^{\mathsf{T}}$, denoted as $\nu_{+,\mathsf{T}}^{\beta,B}$ (plus measure) and $\nu_{-,\mathsf{T}}^{\beta,B}$ (minus measure), respectively (see [24], Chapter IV).

For any $\beta, B \geq 0$ and $\mu \in \mathcal{U}$ supported on the collection of rooted trees $(\mathsf{T}, o) \in \mathcal{T}_*$, let

$$\mathbb{U}(\beta, B) := \frac{1}{2}\mu\Big[\sum_{i\in\partial o} \nu_{+,\mathsf{T}}^{\beta,B}\langle x_o x_i\rangle\Big]. \tag{1.8}$$

Our first result generalizes [28], Theorem 2.4.I, namely the limit (1.4), to any limiting measure μ supported on $\mathcal{T}_*$ subject to a mild continuity assumption on $\mathbb{U}(\cdot, 0)$.

THEOREM 1.6. *Suppose* $\mathsf{G}_n \overset{\text{LWC}}{\Longrightarrow} \mu$ *for some* $\mu \in \mathcal{U}_*$. *Then, at any continuity point* $\beta \geq 0$ *of the bounded, nondecreasing, right-continuous function* $\mathbb{U}(\beta, 0)$, *the Ising measures* ν_n^β *on* G_n *converge locally weakly to* $\overline{\mathfrak{m}} = \mu \circ \overline{\varphi}^{-1}$, *where* $\overline{\varphi} : \mathcal{T}_* \to \mathcal{P}(\overline{\mathcal{T}_*})$ *with* $\overline{\varphi}(\mathsf{T}) = \delta_\mathsf{T} \otimes (\frac{1}{2}\nu_{+,\mathsf{T}}^\beta + \frac{1}{2}\nu_{-,\mathsf{T}}^\beta)$.

Our generalization of (1.5), namely [28], 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 $\mathsf{G} = (V, E)$ 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 $|\cdot|$ denotes the cardinality of a set and ∂S denotes the collection of edges between S and S^c.

THEOREM 1.8. *Suppose* $\{\mathsf{G}_n\}_{n\in\mathbb{N}}$ *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* $\mathsf{G}_n \overset{\text{LWC}}{\Longrightarrow} \mu$ *for some* $\mu \in \mathcal{U}_*$, *then at any continuity point of* $\beta \mapsto \mathbb{U}(\beta, 0)$, *the measures* $\{\nu_{n,+}^\beta\}$ *converge locally weakly to* $\overline{\mathfrak{m}}_+ = \mu \circ \overline{\varphi}_+{}^{-1}$ *where* $\overline{\varphi}_+ : \mathcal{T}_* \to \mathcal{P}(\overline{\mathcal{T}}_*)$ *with* $\overline{\varphi}_+(\mathsf{T}) = \delta_\mathsf{T} \otimes \nu_{+,\mathsf{T}}^\beta$.

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

REMARK 1.10. With $B = 0$, for n odd the probability measures $\nu_{n,\pm}^\beta$ 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 nonessential technical issues, one slightly modifies $\nu_{n,\pm}^\beta$ in case n is even to retain this property, as well as having $\nu_{n,+}^\beta(\underline{x}) = \nu_{n,-}^\beta(\underline{x})$ whenever $\sum_i x_i = 0$.

REMARK 1.11. Recall the example in [28], Section 2.3, where it is shown that even in case of k-regular tree limits one cannot completely dispense of the expander-like condition when dealing with the convergence of $\nu^{\beta}_{n,+}$.

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 $p(\cdot)$ on some finite (type) space $\mathcal{Q}$. Let $\mathbb{Z}_{\geq}$ denotes the set of all nonnegative integers and $\mathbb{Z}_{\geq}^{|\mathcal{Q}|} := \{\underline{k} = (k_1, k_2, \ldots, k_{|\mathcal{Q}|}) : k_j \in \mathbb{Z}_{\geq}, j = 1, 2, \ldots, |\mathcal{Q}|\}$. Consider a (finite) collection of probability measures $P_i(\cdot)$, $i \in \mathcal{Q}$ on $\mathbb{Z}_{\geq}^{|\mathcal{Q}|}$, such that for all $i, j \in \mathcal{Q}$,

$$M(i, j) := \sum_{\underline{k}} P_i(\underline{k}) k_j < \infty, \tag{1.9}$$

$$p(i) M(i, j) = p(j) M(j, i). \tag{1.10}$$

For $m \in \mathbb{N}$, let an *m-star* denote the connected graph of $(m + 1)$ vertices, with one vertex of degree m and all others having degree one. Such m-star has two *ends*, one end being its vertex of degree m, with the other end consisting of the remaining m degree one vertices of the m-star. Now for each n we define the random graph $\mathsf{G}_n = (V_n, E_n)$ as follows. For every $i \in \mathcal{Q}$ and $\underline{k} \in \mathbb{Z}_{\geq}^{|\mathcal{Q}|}$, we create $\lfloor np(i) P_i(\underline{k}) \rfloor$ many $(\sum_j k_j + 1)$-stars with types, such that the end of each of the stars with one vertex has type $i \in \mathcal{Q}$ and the other end consists of $\sum_j k_j$ vertices, of which exactly k_j have type j, for each $j \in \mathcal{Q}$.

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 $\overrightarrow{(i, j)}$. Due to condition (1.10), if not for the integer truncation effects, for any $i, j \in \mathcal{Q}$ the number of half-edges of type $\overrightarrow{(i, j)}$ would match that of type $\overrightarrow{(j, i)}$. We thus achieve such equality between the numbers of $\overrightarrow{(i, j)}$ and $\overrightarrow{(j, i)}$ half-edges, upon adding to G_n at most

$$2 \sum_{i,j} \sum_{\underline{k}} \{np(i) P_i(\underline{k})\} k_j$$

half-edges. This amounts to adding only $O(1)$ half-edges to the stars [since $\sum_{i,j} M(i, j)$ is finite, due to (1.9)].

Thereafter for every $i, j \in \mathcal{Q}$, we perform a uniform matching between half-edges with type $\overrightarrow{(i, j)}$ and half-edges with type $\overrightarrow{(j, i)}$. Once we have obtained a matching between these half-edges, we throw out the floating vertices and join the permanent vertices of those half-edges, which have been matched, to get a graph with types [e.g., if in a matching the half-edge (v, e_v) of type $\overrightarrow{(i, j)}$ matches with the half-edge (w, e_w) of type $\overrightarrow{(j, i)}$ then we join v and w, and $q(v) = i, q(w) = j$]. This completes the recipe for generating the random graph $\mathsf{G}_n = (V_n, E_n)$.

We associate with each $p(\cdot)$ and collection of probability measures $P_i(\cdot)$ that satisfies the conditions of Definition 1.12, a unimodular version of the MGW law, to be denoted hereafter by UMGW.

DEFINITION 1.13. For each $p, \{P_i(\cdot), i \in \mathcal{Q}\},$ and $M(\cdot, \cdot)$ satisfying (1.10) and (1.9), let $\mathcal{Q}_M := \{(i, j) : M(i, j) > 0\} \subseteq \mathcal{Q} \times \mathcal{Q}$ and $\widehat{P}_{i,j}(\cdot)$ for $(i, j) \in \mathcal{Q}_M$, be the probability measures on $\mathbb{Z}_{\geq}^{|\mathcal{Q}|}$ given by

$$\widehat{P}_{i,j}(\underline{k}) = P_i(\underline{k} + e_j)\frac{k_j + 1}{M(i, j)},$$

where e_j denotes the vector with 1 at jth coordinate and 0 elsewhere, and we assume that $\widehat{P}_{i,j}(\underline{k}) > 0$ for some (i, j) and $\|\underline{k}\| := \sum_j k_j \neq 1$ (in the branching processes literature this property is called *nonsingularity*; cf. [5], page 184).

We assume that the mean matrix $\widehat{M}$ for the kernel $\widehat{P}$ over $\mathcal{Q}_M$, which is given by

$$\widehat{M}\big((i_1, j_1), (i_2, j_2)\big) := \mathbb{I}_{j_2 = i_1} \sum_{\underline{k}} \widehat{P}_{i_1, j_1}(\underline{k}) k_{i_2},$$

is *positive regular*. That is, we require that for some finite positive integer r all entries of $(\widehat{M})^r$ 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 $p(\cdot)$, and conditional on the type of the root, say i_0, it's off-spring number and types are chosen according to $P_{i_0}(\cdot)$. From the next generation onward, the off-spring numbers and types are chosen independently at each vertex according to $\widehat{P}_{i,j}$ where i is the type of the current vertex and j being the type of its parent.

REMARK 1.14. In the special case $|\mathcal{Q}| = 1$, 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 (1.10) and positive regularity then trivially hold, while nonsingularity and (1.9) 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 $P(\cdot)$ (cf. [10], Section 1.2.4), which is uniformly sparse and converges weakly to the corresponding UMGW measure of Definition 1.13 (see [10], Proposition 2.5). The latter is precisely the UGW tree measure of [3], Example 1.1, and [9], Section 2.1.

In particular, taking $P(\cdot)$ a Poisson law of parameter 2α, results with $\widehat{P}(k) = P(k)$ (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 $n^{-1}|E_n| \to \alpha$ which also have the UGW measure as their a.s. LWC limit (see [10], Proposition 2.6 and Lemma 2.3).

For $|\mathcal{Q}| > 1$, the uniform sparseness of $\{\mathsf{G}_n\}$ of Definition 1.12 is an immediate consequence of finiteness of $\mathcal{Q}$ and $\sum_{i,j} p(i)M(i,j)$, while its local weak convergence to the corresponding UMGW measure follows along the lines of [10], 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 $\{\mathsf{G}_n\}$ and UMGW for random uniform q-partite, $q \geq 2$, graphs (of $\lfloor \alpha n \rfloor$ edges), which fit within our framework upon taking p uniform on $\{1, \dots, q\}$ and $P_i(\underline{k}) = \prod_{\ell \neq i} P(k_\ell)$, 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 $d_\star > 2$, one has that $\beta \mapsto \mathbb{U}(\beta, 0)$ is continuous except for possibly countably many values of $\beta \in (\beta_c, \beta_\star]$, where $\beta_\star = \operatorname{atanh}[(d_\star - 1)^{-1}]$.*

Thus, upon applying Theorems 1.6 and 1.8 we immediately obtain the following.

COROLLARY 1.16. *Suppose $\mathsf{G}_n \stackrel{\text{LWC}}{\Longrightarrow} \mu$ with μ a* UMGW *measure as in Definition 1.13, having a.s. minimum degree $d_\star > 2$. Then, except for a possibly countably many values of $\beta \in (\beta_c, \beta_\star]$,*

(a) *ν_n converges locally weakly to $\overline{\mathfrak{m}} = \mu \circ \overline{\varphi}^{-1}$, for $\overline{\varphi}$ as in Theorem 1.6.*

(b) *If in addition $\{\mathsf{G}_n\}_{n \in \mathbb{N}}$ 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 $\overline{\mathfrak{m}}_+ = \mu \circ \overline{\varphi}_+^{-1}$, for $\overline{\varphi}_+$ 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, we have the following.

LEMMA 1.17. *Suppose (1.10) holds for $p(\cdot)$ strictly positive and $\{P_i, i \in \mathcal{Q}\}$ of bounded support, such that $P_i(\underline{k}) = 0$ whenever $\|\underline{k}\| := \sum_j k_j \leq 2$. 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 $P([0, d_\star)) = 0$ for some $d_\star \geq 3$ and let $\widehat{\Delta}$ be distributed on $\mathbb{N}$ according to $\widehat{P}_k := (k+1)P_{k+1}/\sum_j jP_j$, $k \geq 0$. For any fixed $\beta > \beta_c$ consider the recursion over $t \geq 0$,*

$$h^{(t+1)} \stackrel{d}{=} \sum_{\ell=1}^{\widehat{\Delta}} \operatorname{atanh}\bigl[\tanh(\beta)\tanh\bigl(h_\ell^{(t)}\bigr)\bigr], \tag{1.11}$$

where $h_\ell^{(t)}$ are i.i.d. copies of $h^{(t)}$ which are further independent of $\widehat{\Delta}$. Denote by $h^{\beta,+}$ its limit in law when $t \to \infty$ and starting at $h^{(0)} = \infty$. Then, fixing any $\beta \geq \beta_0 > \beta_\star$ and starting this recursion at a stochastically dominating $h^{(0)} \succeq h^{\beta_0,+}$, 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 (1.11) is a *splitting Gibbs measure* (see [11], Remarks 1.13 and 2.6). Hence, Lemma 1.18 implies that there is only one *Bethe–Gibbs measure* (see [11], Remark 2.6), that corresponds to some $h \succeq h^{\beta_0,+}$, $\beta_0 \in (\beta_\star, \beta)$, with a similar conclusion for the UMGW measures of Definition 1.13.

We expect both Lemmas 1.15 and 1.18 to hold for UGW and UMGW measures at all β (and without a minimum degree assumption). However, the nonregularity of T under genuinely random UGW and UMGW measures yields for $\beta \in (\beta_c, \beta_\star]$ a technical difficulty which we cannot overcome (cf. Remark 5.7).

Outline of the paper.

- As shown in Section 2, weak convergence of μ_n (of Definition 1.2) implies that the corresponding measures $\{\nu_n\}$ and $\{\nu_{n,+}\}$ have sub-sequential 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 μ supported on $\mathcal{T}_*$, we find in Lemma 3.3 that at its continuity points $\mathbb{U}(\beta, 0)$ is the limit of both the ν_n-expected values and $\nu_{n,+}$-expected values, of certain functionals of $\underline{x}$. Extending (in Lemma 3.4), the result of [28], Lemma 3.2, we deduce in Lemma 3.8 that the weak limit points of Section 2 must be convex combinations of $\nu_{\pm,\mathsf{T}}$ and get Theorem 1.6 by the symmetry relation $\nu_n(\underline{x}) = \nu_n(-\underline{x})$.
- In Section 4, we prove Theorem 1.8. First, we deduce in Lemma 4.4 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 Lemmas 4.5 and 4.7 properties of SRW on trees, the assumed edge-expander condition for G_n eliminates all but one choice for the convex combination of $\nu_{\pm,\mathsf{T}}$ (thus proving the theorem).

- In Section 5, we deal with continuity of $\beta \mapsto \mathbb{U}(\beta, 0)$. Constructing in Lemma 5.4 a suitable sequence of random variables that increases to the root magnetization under $\nu_{+,\mathsf{T}}$, we establish in Lemma 5.1 such continuity at any $\beta > \beta_\star$. 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 $\beta \mapsto \mathbb{U}(\beta, 0)$ at $\beta = \beta_c$, which we verify for UMGW measures. Lastly, while Lemma 1.17 is well known, for completeness we outline its proof.

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 μ_n and having marks from a finite set $\mathcal{X}$).

LEMMA 2.1. *Suppose $\{\mu_n\}$ of Definition 1.2 converges weakly in $\mathcal{P}(\mathcal{G}_*)$. Then for any probability measures ζ_n on $\mathcal{X}^{[n]}$ and any sub-sequence $\{n_\ell\}_{\ell\in\mathbb{N}}$ there exists a further sub-sequence $\{n_{\ell_k}\}_{k\in\mathbb{N}}$ such that $\{\zeta_{n_{\ell_k}}\}$ converges locally weakly to a limit $\overline{\mathfrak{m}}$ (which may depend on $\{n_{\ell_k}\}$).*

PROOF. Fixing $\{\zeta_n\}$ and $t \in \mathbb{N}$ recall that μ_n^t are such that

$$\mu_n^t(\mathsf{G}) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(\mathsf{B}_i(t) \simeq \mathsf{G}),$$

for each $\mathsf{G} \in \mathcal{G}_*(t)$ and the balls $\mathsf{B}_i(t)$ in G_n. The assumed convergence of $\{\mu_n\}_{n\in\mathbb{N}}$ in $\mathcal{P}(\mathcal{G}_*)$ implies the convergence of $\{\mu_n^t\}$ in $\mathcal{P}(\mathcal{G}_*(t))$, so by Prohorov's theorem μ_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 $\mathcal{G}_\varepsilon(t) \subset \mathcal{G}_*(t)$, such that

$$\liminf_{n\to\infty} \mu_n^t(\mathcal{G}_\varepsilon(t)) \ge 1 - \varepsilon. \tag{2.1}$$

Further, per $\mathsf{G} \in \mathcal{G}_*(t)$ the space of marks $\mathcal{X}^{\mathsf{G}}$ is finite, so the set $\overline{\mathcal{G}}_\varepsilon(t) := \{(\mathsf{G}, \underline{x}_{\mathsf{G}}) : \mathsf{G} \in \mathcal{G}_\varepsilon(t), \underline{x}_{\mathsf{G}} \in \mathcal{X}^{\mathsf{G}}\}$ is also finite, and by Prohorov's theorem the collection of all probability measures on $\overline{\mathcal{G}}_\varepsilon(t)$ is compact. In particular, $\mathcal{M}_\varepsilon(t) := \{\delta_{\mathsf{G}} \otimes \nu_{\mathsf{G}} : \mathsf{G} \in \mathcal{G}_\varepsilon(t), \nu_{\mathsf{G}} \in \mathcal{P}(\{-1, 1\}^{\mathsf{G}})\}$ is a pre-compact collection of probability measures on $\mathcal{P}(\overline{\mathcal{G}}_*(t))$. Since $\overline{\mathsf{P}}_n^t(I_n) \in \mathcal{M}_\varepsilon(t)$ with probability $\mu_n^t(\mathcal{G}_\varepsilon(t))$, it thus follows that for each $t \in \mathbb{N}$, the laws of $\overline{\mathsf{P}}_n^t(I_n)$ are uniformly tight, hence relatively compact. Consequently, there exists a diagonal sub-sequence along which the random probability measures $\overline{\mathsf{P}}_n^t(I_n)$ converge in law, to say $\overline{\mathfrak{m}}_t$, simultaneously for all $t \in \mathbb{N}$. By the obvious embedding of $\overline{\mathcal{G}}_*(t)$ within $\overline{\mathcal{G}}_*(t+1)$, each $\overline{\nu}_{t+1} \in \mathcal{P}(\overline{\mathcal{G}}_*(t+1))$ induces a marginal probability measure on $\overline{\mathcal{G}}_*(t)$, denoted

$\pi_t(\overline{\nu}_{t+1})$. By definition, $\pi_t(\overline{\mathsf{P}}_n^{t+1}(I_n)) = \overline{\mathsf{P}}_n^t(I_n)$ for all $t, n \in \mathbb{N}$. This implies the relation $\overline{\mathfrak{m}}_t = \overline{\mathfrak{m}}_{t+1} \circ \pi_t^{-1}$ between the corresponding weak limits. That is, the sequence $\{\overline{\mathfrak{m}}_t\}$ of probability measures on the Polish spaces $\mathcal{P}(\overline{\mathcal{G}}_*(t))$ is consistent with respect to the projections π_t. This completes the proof, since by Kolmogorov's extension theorem there exists a probability measure $\overline{\mathfrak{m}}$ on $\mathcal{P}(\overline{\mathcal{G}}_*)$ such that $\overline{\mathfrak{m}}_t = \overline{\mathfrak{m}}^t$ 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 $\overline{\mathcal{G}}_*$, which we define next. First, recall that probability measure ν_{G} on $\mathcal{X}^{\mathsf{G}}$ for a fixed infinite graph $\mathsf{G} \in \mathcal{G}_*$ is an Ising–Gibbs measure iff ν_{G} satisfies the relevant DLR condition. That is, setting $\mathsf{G}(\infty) = \mathsf{G}$, $\mathsf{G}(-1) = \varnothing$ and $\mathsf{G}(t, \overline{t}) = \mathsf{G}(\overline{t}) \setminus \mathsf{G}(t)$ for $t < \overline{t} \leq \infty$, one requires that for $\overline{t} = \infty$, any $t \in \mathbb{N}$ and ν_{G}-a.e. $\underline{x}_{\mathsf{G}(t,\overline{t})}$,

$$\nu_{\mathsf{G}}(\underline{x}_{\mathsf{G}(t)} | \underline{x}_{\mathsf{G}(t,\overline{t})}) = \widetilde{\nu}\big(\underline{x}_{\mathsf{G}(t)} | \underline{x}_{\mathsf{G}(t,t+1)}, \mathsf{G}(t+1)\big), \tag{2.2}$$

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

$$\widetilde{\nu}\big(\underline{x}_W | \underline{x}_{V'\setminus W}, \mathsf{G}'\big) := \frac{\exp\{\beta \sum_{(i,j)\in E'} x_i x_j\}}{\sum_{\{\underline{x}'_{V'}:\underline{x}'_{V'\setminus W} = \underline{x}_{V'\setminus W}\}} \exp\{\beta \sum_{(i,j)\in E'} x'_i x'_j\}} \tag{2.3}$$

denotes the Ising measure on W, given boundary values at $V' \setminus W$ (see [16], Chapter 2).

Next, for any $t \geq -1$ and $\overline{t} = t+1, \dots, \infty$, fixing $r \in \mathbb{N}$, $\mathsf{G} \in \mathcal{G}_*$ and the marks $\underline{x}_{\mathsf{G}(r\wedge\overline{t})\setminus\mathsf{G}(r\wedge t)}$ we denote by $\mathsf{B}^{(t,\overline{t})}_{(\mathsf{G},\underline{x}_{\mathsf{G}})}(r)$ the union over all possible mark values $\underline{x}_{\mathsf{G}(r\wedge t)}$ of the $\overline{\mathcal{G}}_*$-metric balls $\mathsf{B}_{\overline{\mathsf{G}}}(r \wedge \overline{t})$ centered at $\overline{\mathsf{G}} = (\mathsf{G}, \underline{x}_{\mathsf{G}})$. Considering the sub-$\sigma$-algebras

$$\mathscr{C}_{\overline{\mathcal{G}}_*(t,\overline{t})} := \sigma\big(\mathsf{B}^{(t,\overline{t})}_{(\mathsf{G},\underline{x}_{\mathsf{G}})}(r), (\mathsf{G}, \underline{x}_{\mathsf{G}}) \in \overline{\mathcal{G}}_*, r \in \mathbb{N}\big), \tag{2.4}$$

generated by these sets, note that $\mathscr{C}_{\overline{\mathcal{G}}_*(t,\overline{t})}$ are nondecreasing in $\overline{t}$ and nonincreasing in t, where in particular, $\mathscr{C}_{\overline{\mathcal{G}}_*(\overline{t})} = \mathscr{C}_{\overline{\mathcal{G}}_*(-1,\overline{t})}$ and $\mathscr{C}_{\mathcal{G}_*} = \mathscr{C}_{\overline{\mathcal{G}}_*(\infty,\infty)}$ [as a ball $\mathsf{B}_{\mathsf{G}}(r) \subset \mathcal{G}_*$ of radius r and center G is the $\mathcal{G}_*$-projection of the union over all $\underline{x}_{\mathsf{G}} \in \mathcal{X}^{\mathsf{G}}$ of the corresponding balls $\mathsf{B}_{(\mathsf{G},\underline{x}_{\mathsf{G}})}(r)$ in $\overline{\mathcal{G}}_*$]. Since $\overline{\mathcal{G}}_*$ is a Polish space, the regular conditional probability measure $\overline{\nu}(\cdot|\mathscr{C}_{\mathcal{G}_*})$ is thus well defined for any $\overline{\nu} \in \mathcal{P}(\overline{\mathcal{G}}_*)$ (see [33], Section 9.2), and we lift the notion of an Ising–Gibbs measure to $\mathcal{P}(\overline{\mathcal{G}}_*)$, by considering the DLR condition (2.2) with this conditional measure playing the role of ν_{G}. In this setting, per $t \in \mathbb{N}$ what one has in the left-hand side of (2.2) amounts to the restriction to $\underline{x}_{\mathsf{G}(t)}$ of the regular conditional probability measure $\overline{\nu}(\cdot|\mathscr{C}_{\overline{\mathcal{G}}_*(t,\infty)})$, resulting with the following definition.

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

$$\overline{\nu}(\underline{x}_{\mathsf{G}(t)} | \mathscr{C}_{\overline{\mathcal{G}}_*(t,\infty)}) = \widetilde{\nu}\big(\underline{x}_{\mathsf{G}(t)} | \underline{x}_{\mathsf{G}(t,t+1)}, \mathsf{G}(t+1)\big), \tag{2.5}$$

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

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

$$\overline{\nu}(\underline{x}_{\mathsf{G}(t)} | \mathscr{C}_{\overline{\mathcal{G}}_*(t,\overline{t})}) = \widetilde{\nu}\big(\underline{x}_{\mathsf{G}(t)} | \underline{x}_{\mathsf{G}(t,t+1)}, \mathsf{G}(t+1)\big). \tag{2.6}$$

We focus hereafter on the subset $\mathcal{I}_*$ of all Ising–Gibbs measures of the form $\overline{\nu} = \delta_{\mathsf{G}} \otimes \nu_{\mathsf{G}}$, with ν_{G} being an Ising–Gibbs measure for the *fixed* graph $\mathsf{G} \in \mathcal{G}_*$. Denoting by $\mathcal{I}_{(t,\overline{t})}$ those $\overline{\nu} = \delta_{\mathsf{G}} \otimes \nu_{\mathsf{G}}$ in $\mathcal{P}(\overline{\mathcal{G}}_*)$ with ν_{G} satisfying (2.2) per fixed $t < \overline{t}$ finite, we see that

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

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

Considering (2.2) at fixed $\overline{t} > t$ for ν_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 $\mu_n \Rightarrow \mu$, for μ_n as in Definition 1.2. Then*:

(a) *Any sub-sequential local weak limit $\overline{\mathfrak{m}}$ of $\{\nu_n\}$ is supported on the collection $\mathcal{I}_*$ of Ising–Gibbs measures and restricted to $\mathcal{P}(\mathcal{G}_*)$ it has the marginal $\widetilde{\mathfrak{m}} = \mu \circ \widetilde{\varphi}^{-1}$, where $\widetilde{\varphi}(\mathsf{G}) = \delta_{\mathsf{G}}$ for any $\mathsf{G} \in \mathcal{G}_*$.*

(b) *The same holds for sub-sequential limits $\overline{\mathfrak{m}}_+$ of $\{\nu_{n,+}\}$, provided $\{\mathsf{G}_n\}$ is uniformly sparse.*

PROOF. Fix a sub-sequence n_ℓ along which $\{\nu_n\}$ [or $\{\nu_{n,+}\}$], converges locally weakly to some $\overline{\mathfrak{m}}$. Then, for each $t \in \mathbb{N}$ the $\mathcal{P}(\mathcal{G}_*(t))$-restriction $\mathsf{P}_n^t(I_n)$ of $\overline{\mathsf{P}}_n^t(I_n)$ converges in law to $\widetilde{\mathfrak{m}}^t$. Thus, for any fixed $\mathsf{G} \in \mathcal{G}_*$,

$$\begin{aligned}\widetilde{\mathfrak{m}}^t(\delta_{\mathsf{G}(t)}) &= \lim_{\ell\to\infty} \frac{1}{n_\ell} \sum_{i=1}^{n_\ell} \mathbb{I}(\delta_{\mathsf{B}_i(t)} = \delta_{\mathsf{G}(t)}) = \lim_{\ell\to\infty} \frac{1}{n_\ell} \sum_{i=1}^{n_\ell} \mathbb{I}\big(\mathsf{B}_i(t) \simeq \mathsf{G}(t)\big) \\ &= \mu^t\big(\mathsf{G}(t)\big),\end{aligned}$$

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

$$\overline{\mathfrak{m}}^{\bar{t}}(\mathcal{I}_{(t,\bar{t})}) = 1. \tag{2.8}$$

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

$$\begin{aligned} \mathcal{I}_{(t,\bar{t})}^{\varepsilon,\delta} := \Bigg\{ & \delta_{\mathsf{G}} \otimes \nu_{\mathsf{G}} : \mathsf{G} \in \mathcal{G}_\varepsilon^+(\bar{t}), \\ & 1-\delta \le \frac{\nu_{\mathsf{G}}(\underline{x}_{\mathsf{G}(t)} | \underline{x}_{\mathsf{G}(t,\bar{t})})}{\widetilde{\nu}(\underline{x}_{\mathsf{G}(t)} | \underline{x}_{\mathsf{G}(t,t+1)}, \mathsf{G}(t+1))} \le \frac{1}{1-\delta} \Bigg\}, \end{aligned} \tag{2.9}$$

a closed subset of $\mathcal{P}(\overline{\mathcal{G}}_*(\bar{t}))$. Now, if $\mathsf{B}_i(\bar{t}+1) \simeq \mathsf{G}$ for some $\mathsf{G} \in \mathcal{G}_\varepsilon(\bar{t}+1)$, then

$$\nu_{n,\mathsf{B}_i(\bar{t})}(\underline{x}_{\mathsf{B}_i(t)} | \underline{x}_{\mathsf{B}_i(t,\bar{t})}) = \widetilde{\nu}(\underline{x}_{\mathsf{G}(t)} | \underline{x}_{\mathsf{G}(t,t+1)}, \mathsf{G}(t+1))$$

and consequently $\overline{\mathsf{P}}_n^{\bar{t}}(i) \in \mathcal{I}_{(t,\bar{t})}^{\varepsilon,0}$. Clearly, for any $\varepsilon > 0$ fixed, $\mathcal{I}_{(t,\bar{t})}^{\varepsilon,0}$ is a subset of $\mathcal{I}_{(t,\bar{t})}$, hence from (2.1) and the assumed local weak convergence along the subsequence n_ℓ, we deduce that

$$1-\varepsilon \le \limsup_{\ell\to\infty} \frac{1}{n_\ell} \sum_{i=1}^{n_\ell} \mathbb{I}\{\overline{\mathsf{P}}_{n_\ell}^{\bar{t}}(i) \in \mathcal{I}_{(t,\bar{t})}^{\varepsilon,0}\} \le \overline{\mathfrak{m}}^{\bar{t}}(\mathcal{I}_{(t,\bar{t})}^{\varepsilon,0}) \le \overline{\mathfrak{m}}^{\bar{t}}(\mathcal{I}_{(t,\bar{t})}). \tag{2.10}$$

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

$$Z_{n,i}^{0,t} = Z_{n,i}^{0,t}(\bar{t}, \mathsf{B}_i(\bar{t}), \underline{x}_{\mathsf{B}_i(\bar{t})}) := \nu_n\left(\sum_{j=1}^n x_j \ge 0 \Big| \underline{x}_{\mathsf{B}_i(t,\bar{t})}\right),$$

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

$$\nu_{n,+}(\underline{x}_{\mathsf{B}_i(t)} | \underline{x}_{\mathsf{B}_i(t,\bar{t})}) = \frac{Z_{n,i}^0}{Z_{n,i}^{0,t}} \nu_{n,\mathsf{B}_i(\bar{t})}(\underline{x}_{\mathsf{B}_i(t)} | \underline{x}_{\mathsf{B}_i(t,\bar{t})}) \tag{2.11}$$

(for clarity of presentation we ignore the slight modification of $Z^{0,t}_{n,i}$ which is required for n even, in accordance with Remark 1.10). The conditioning effect eventually washes away, since setting

$$Z^{\pm}_{n,i} := \nu_n\Big(\sum_{j\notin \mathsf{B}_i(\bar t)} x_j > \pm\big|\mathsf{B}_i(\bar t)\big|\Big|\underline{x}_{\mathsf{B}_i(t,\bar t)}\Big),$$

which are independent of $t < \bar t$, and fixing $\varepsilon, \delta > 0$ we show that for all n large enough and $i \in [n]$,

$$\mathsf{B}_i(\bar t) \in \mathcal{G}^+_\varepsilon(\bar t) \Longrightarrow \inf_{\underline{x}_{\mathsf{B}_i(\bar t)}}\left\{\frac{Z^+_{n,i}}{Z^-_{n,i}}\right\} \ge 1-\delta. \tag{2.12}$$

Indeed, clearly $Z^+_{n,i} \le Z^{0,t}_{n,i} \le Z^-_{n,i}$ and so by (2.11) the right-hand side of (2.12) yields that the probability measures $\mathbb{P}^{\bar t}_n(i)$ corresponding to $\nu_{n,+}$ are then in $\mathcal{I}^{\varepsilon,\delta}_{(t,\bar t)}$. Consequently, following the derivation of (2.10) we find that $1-\varepsilon \le \overline{\mathfrak{m}}^{\bar t}_+(\mathcal{I}^{\varepsilon,\delta}_{(t,\bar t)})$ for any sub-sequential limit $\overline{\mathfrak{m}}_+$ of $\{\nu_{n,+}\}$ and all $\varepsilon, \delta > 0$. Since

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

considering $\delta \downarrow 0$ followed by $\varepsilon \downarrow 0$ completes the proof of (2.8). As for (2.12), necessarily,

$$\kappa := \sup_{\mathsf{G}\in\mathcal{G}_\varepsilon(\bar t+1)} \big|E(\mathsf{G})\big| < \infty$$

[since $\mathcal{G}_\varepsilon(\bar t+1)$ is a finite collection of finite graphs]. Thus, assuming hereafter that $\mathsf{B}_i(\bar t+1) \simeq \mathsf{G}$ for some $\mathsf{G} \in \mathcal{G}_\varepsilon(\bar t+1)$, at most κ edges of G_n touch $\mathsf{B}_i(\bar t)$. Hence, by the invariance with respect to a global sign change of the Ising measure $\nu_{E_n\setminus E(\mathsf{B}_i(\bar t+1))}$ on the sub-graph of G_n in which all edges within $\mathsf{B}_i(\bar t+1)$ have been deleted, we conclude that

$$\begin{aligned} Z^-_{n,i} &\ge \nu_n\Big(\sum_{j\notin\mathsf{B}_i(\bar t)} x_j \ge 0\Big|\underline{x}_{\mathsf{B}_i(\bar t)}\Big) \ge e^{-2\beta\kappa}\nu_{E_n\setminus E(\mathsf{B}_i(\bar t+1))}\Big(\sum_{j\notin\mathsf{B}_i(\bar t)} x_j \ge 0\Big) \\ &\ge \frac{1}{2}e^{-2\beta\kappa}. \end{aligned} \tag{2.13}$$

Further, $|\mathsf{B}_i(\bar t)| \le \kappa$ and by the assumed uniform sparseness of $\{\mathsf{G}_n\}$, there exists $k \in \mathbb{N}$ and $n_0 \ge 3\kappa$ large enough so that

$$\sum_{i=1}^n \Delta_i(\mathsf{G}_n)\mathbb{I}\big(\Delta_i(\mathsf{G}_n) \ge k\big) \le \frac{n}{3} \qquad \forall n \ge n_0$$

[see (1.6)]. Consequently, for any $n \ge n_0$ there are at least $n/3$ vertices in $\mathsf{G}_n \setminus \mathsf{B}_i(\bar t)$ 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 [28], Lemma 4.1, that under ν_n and conditional on the values of $\underline{x}_{S^c}$, the $\pm$-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 $n\ge n_0$,

$$\sup_r \nu_n\biggl(\sum_{j\notin \mathsf{B}_i(\bar{t})} x_j = r \Big| \underline{x}_{\mathsf{B}_i(\bar{t})}\biggr) \le Cn^{-1/2},$$

from which it follows that uniformly in $\underline{x}_{\mathsf{B}_i(\bar{t})}$,

$$0\le Z^-_{n,i}-Z^+_{n,i}\le 2\big|\mathsf{B}_i(\bar{t})\big|Cn^{-1/2}\le 2\kappa Cn^{-1/2}.$$

Combining this bound with (2.13), we conclude that (2.12) holds for all $n\ge n_\delta$ sufficiently large. □

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

$$\nu(\underline{x})=\frac{1}{Z(\beta,\underline{B})}\exp\biggl\{\beta\sum_{(i,j)\in E} x_ix_j+\sum_{i\in V}B_ix_i\biggr\}. \tag{3.1}$$

In this context, we often take advantage of the Griffiths inequality for ferromagnetic Ising models (which for completeness we state next, see also [24], Theorem IV.1.21).

PROPOSITION 3.1 (The Griffiths inequality). *Consider two Ising models $\nu(\cdot)$ and $\nu'(\cdot)$ on finite graphs $\mathsf{G}=(V,E)$ and $\mathsf{G}'=(V,E')$, inverse temperatures β and β', and magnetic fields $\{B_i\}$ and $\{B_i'\}$, respectively. If $E\subseteq E'$, $\beta\le\beta'$ and $0\le B_i\le B_i'$, for all $i\in V$, then*

$$0\le \nu\biggl[\prod_{i\in W}x_i\biggr]\le \nu'\biggl[\prod_{i\in W}x_i\biggr]\qquad \forall W\subseteq V.$$

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 ([9], Lemma 4.1). *For a subtree T' of a finite tree T, let $\partial_\star\mathsf{T}'$ denote the subset of vertices T' connected by an edge to $\mathsf{W}:=\mathsf{T}\setminus\mathsf{T}'$ and for each $u\in\partial_\star\mathsf{T}'$ let $\langle x_u\rangle_{\mathsf{W}}$ denote the root magnetization of the Ising model on the maximal subtree T_u of $\mathsf{W}\cup\{u\}$ rooted at u. The marginal on T' of an Ising measure ν on T, denoted $\nu^{\mathsf{T}}_{\mathsf{T}'}$ is then an Ising measure on T' with magnetic field $B_u'=\operatorname{atanh}(\langle x_u\rangle_{\mathsf{W}})\ge B_u$ for $u\in\partial_\star\mathsf{T}'$ and $B_u'=B_u$ for $u\in\mathsf{T}'\setminus\partial_\star\mathsf{T}'$.*

Adopting hereafter the notation $\mathsf{T}_{x\to y}$ for the connected component of the subtree 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 $x_i x_j$ across a uniformly chosen edge $(i, j) \in E_n$, under the measures $\nu_{n,\pm}$ and ν_n.

LEMMA 3.3. *Suppose* $\mathsf{G}_n \stackrel{\mathrm{LWC}}{\Longrightarrow} \mu$ *for some* $\mu \in \mathcal{U}_*$. *Then*, $(\beta, B) \mapsto \mathbb{U}(\beta, B)$ *is bounded*, *nondecreasing*, *right-continuous at* $\beta, B \geq 0$, *continuous at any* $B > 0$, *and*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{(i,j)\in E_n} \nu_{n,+}^{\beta,0}\langle x_i x_j\rangle = \lim_{n\to\infty} \frac{1}{n} \sum_{(i,j)\in E_n} \nu_n^{\beta,0}\langle x_i x_j\rangle = \mathbb{U}(\beta, 0), \tag{3.2}$$

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

PROOF. Since $\nu_n^{\beta,0} = \frac{1}{2}\nu_{n,+}^{\beta,0} + \frac{1}{2}\nu_{n,-}^{\beta,0}$ and $\nu_{n,-}^{\beta,0}(\underline{x}) = \nu_{n,+}^{\beta,0}(-\underline{x})$ for all $\underline{x}$, clearly $\nu_{n,\pm}^{\beta,0}\langle x_i x_j\rangle = \nu_n^{\beta,0}\langle x_i x_j\rangle$ 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

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

for all n, β and B, amounts to proving that

$$\lim_{n\to\infty} \frac{\partial}{\partial\beta}\phi_n(\beta, B) = \mathbb{U}(\beta, B), \tag{3.3}$$

for $B = 0$ and any $\beta \geq 0$ at which $\mathbb{U}(\beta, 0)$ is continuous. To this end, we first establish (3.3) for all $\beta \geq 0$ and $B > 0$.

Applying [10], Lemma 2.12, for $A \equiv \{i, j\}$ and $\mathsf{U} \equiv \mathsf{B}_i(t)$, using the Griffiths inequality and local weak convergence, we obtain that per $\beta, B \geq 0$ and $t \geq 2$,

$$\begin{aligned} \mu\Biggl[\frac{1}{2}\sum_{i\in\partial o} \nu_{f,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle\Biggr] &\leq \liminf_{n\to\infty} \frac{\partial}{\partial\beta}\phi_n(\beta, B) \\ &\leq \limsup_{n\to\infty} \frac{\partial}{\partial\beta}\phi_n(\beta, B) \leq \mu\Biggl[\frac{1}{2}\sum_{i\in\partial o} \nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle\Biggr], \end{aligned} \tag{3.4}$$

where $\nu_{f,\mathsf{T}}^{\beta,B,t}$ is the Ising measure on $\mathsf{T}(t)$ with *free* boundary condition on $\partial\mathsf{T}(t) := \mathsf{T}(t) \setminus \mathsf{T}(t-1)$ (for more details, see [10], pages 163–164). Next, for probability measures

$$\widehat{\nu}(x_1, x_2) = z^{-1} \exp\{\beta x_1 x_2 + H_1 x_1 + H_2 x_2\}, \tag{3.5}$$

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

$$\widehat{\nu}\langle x_1 x_2\rangle = F\bigl(\tanh(\beta), m_1 m_2\bigr), \tag{3.6}$$

with $m_j = \tanh(H_j)$, $j = 1, 2$ and

$$F(\gamma, r) := \frac{\gamma + r}{1 + \gamma r}. \tag{3.7}$$

Setting $m^{\ell,\ddagger}(\mathsf{T}') := \nu_{\ddagger,\mathsf{T}'}^{\beta,B,\ell}\langle x_{o'}\rangle$ for $\ddagger \in \{f, +\}$ and the corresponding *root-magnetization* of the Ising measure on $(\mathsf{T}'(\ell), o') \in \mathcal{T}_*(\ell)$, we note that for any $i \in \partial o$, the marginal on $\mathsf{U}' = (o, i)$ of the Ising measures $\nu_{\ddagger,\mathsf{T}}^{\beta,B,t}$ is by Proposition 3.2 of the form (3.5), with $m_1 = m^{t,\ddagger}(\mathsf{T}_{o\to i})$ and $m_2 = m^{t-1,\ddagger}(\mathsf{T}_{i\to o})$. Consequently, $\nu_{\ddagger,\mathsf{T}}^{t}\langle x_o x_i\rangle = F(\tanh(\beta), r_\ddagger(t))$ is a continuous function of $r_\ddagger(t) := m^{t,\ddagger}(\mathsf{T}_{o\to i}) m^{t-1,\ddagger}(\mathsf{T}_{i\to o})$. In case $B > 0$, upon applying [14], Lemma 3.1 (which only requires local finiteness of the tree), first for $\mathsf{T} = \mathsf{T}_{o\to i}$ and then for $\mathsf{T} = \mathsf{T}_{i\to o}$, we deduce that $r_+(t) - r_f(t) \to 0$ and hence $\nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle - \nu_{f,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle \to 0$ when $t \to \infty$. This holds for all $i \in \partial o$, so recalling that $\mu\langle\Delta_o\rangle$ is finite (by uniform sparseness of $\{\mathsf{G}_n\}$), we get by dominated convergence (DCT), that

$$\lim_{t\to\infty} \mu\left[\frac{1}{2}\sum_{i\in\partial o} \nu_{f,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle\right] = \lim_{t\to\infty} \mu\left[\frac{1}{2}\sum_{i\in\partial o} \nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle\right], \tag{3.8}$$

for any $\beta \geq 0$ and $B > 0$. Now using (3.4), and recalling the definition of $\mathbb{U}(\beta, B)$, we note that (3.3) holds, at any $B > 0$ and $\beta \geq 0$.

While (3.8) is typically false at $B = 0$ and β large enough, clearly for any $\mathsf{T} \in \mathcal{T}_*$ and finite $t \geq 0$, the function $\nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle$ is jointly continuous in β and B. These Ising measures of plus boundary condition correspond to taking $B_i \uparrow \infty$ at all $i \in \mathsf{T} \setminus \mathsf{T}(t-1)$ (see Definition 1.5). Hence, by the Griffiths inequality we have that $\frac{1}{2}\nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle$ is nonincreasing in t and nondecreasing in β, B for $\beta, B \geq 0$. The same monotonicity properties apply for the sum of such functions over $i \in \partial o$ and in so far as (β, B) are concerned, retained by the expectation $\mathbb{U}(\beta, B)$ with respect to the law μ of T, of its limit as $t \uparrow \infty$. Since $\mu\langle\Delta_o\rangle$ is finite, we further deduce by DCT the joint continuity of

$$(\beta, B) \mapsto \mu\left[\sum_{i\in\partial o} \nu_{+,\mathsf{T}}^{\beta,B,t}\langle x_o x_i\rangle\right],$$

which upon interchanging limits in t and β, B, yields the right-continuity of $\mathbb{U}(\beta, B)$ at all $\beta, B \geq 0$.

We denote hereafter by $f_n(\cdot) \stackrel{\mathbb{Q}^c}{\to} f(\cdot)$ the convergence of f_n to f on some co-countable set, and $f(\cdot) \stackrel{\mathbb{Q}^c}{=} g(\cdot)$ when f and g agree on a co-countable set. Since $\beta \mapsto \phi_n(\beta, B)$ are convex functions, so is their limit $\phi(\beta, B)$ (see [11], Theorem 1.8, for existence of such limit at any $\beta \geq 0$, $B \in \mathbb{R}$ fixed). Such pointwise convergence of $\mathbb{R}$-valued convex functions yields that $\frac{\partial}{\partial\beta}\phi_n(\beta, B) \stackrel{\mathbb{Q}^c}{\to} \frac{\partial}{\partial\beta}\phi(\beta, B)$ per fixed $B \geq 0$, and consequently $\frac{\partial}{\partial\beta}\phi(\beta, B) \stackrel{\mathbb{Q}^c}{=} \mathbb{U}(\beta, B)$ 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 $\frac{\partial}{\partial\beta}\phi(\beta, B_m) \stackrel{\mathbb{Q}^c}{\to} \frac{\partial}{\partial\beta}\phi(\beta, 0)$. Further, $B \mapsto \mathbb{U}(\beta, B)$ is right continuous, hence $\mathbb{U}(\beta, B_m) \to \mathbb{U}(\beta, 0)$. From these two convergences, we deduce that $\mathbb{U}(\beta, 0) \stackrel{\mathbb{Q}^c}{=} \frac{\partial}{\partial\beta}\phi(\beta, 0)$. We have seen already that $\frac{\partial}{\partial\beta}\phi_n(\beta, 0) \stackrel{\mathbb{Q}^c}{\to} \frac{\partial}{\partial\beta}\phi(\beta, 0)$, hence also $\frac{\partial}{\partial\beta}\phi_n(\beta, 0) \stackrel{\mathbb{Q}^c}{\to} \mathbb{U}(\beta, 0)$. Since $\frac{\partial}{\partial\beta}\phi_n(\beta, 0)$ are nondecreasing continuous functions, this convergence extends to *all continuity points of* $\beta \mapsto \mathbb{U}(\beta, 0)$. □

The following extension of [28], Lemma 3.2, to arbitrary $\mathsf{T} \in \mathcal{T}_*$ allows us to utilize Lemma 3.3 for restricting the weak limit points of $\nu_{n,+}$ and ν_n, to convex combinations of $\nu_{\pm,\mathsf{T}}$.

LEMMA 3.4. *For any Ising–Gibbs measure* ν_{T} *on some* $\mathsf{T} \in \mathcal{T}_*$ *and all* $i \in V(\mathsf{T})$,

$$\sum_{j\in\partial i} \nu_{\mathsf{T}}\langle x_i x_j\rangle \le \sum_{j\in\partial i} \nu_{+,\mathsf{T}}\langle x_i x_j\rangle = \sum_{j\in\partial i} \nu_{-,\mathsf{T}}\langle x_i x_j\rangle, \tag{3.9}$$

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

PROOF. The equality in (3.9) is an immediate consequence of the fact that under $\nu_{+,\mathsf{T}}$ the random vector $-\underline{x}_{\mathsf{T}}$ admits the law $\nu_{-,\mathsf{T}}$. 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 nonempty 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 [16], Chapter 7). Consequently, it suffices to fix an *extremal* Ising–Gibbs measure $\nu_{\mathsf{T}} \neq \nu_{\pm,\mathsf{T}}$ and show that for every edge $(i, j) \in E(\mathsf{T})$,

$$\nu_{\mathsf{T}}\langle x_i x_j\rangle \le \nu_{+,\mathsf{T}}\langle x_i x_j\rangle \tag{3.10}$$

with a strict inequality for *at least one* $(i, j) \in E(\mathsf{T})$. To this end, for each $(i, j) \in E(\mathsf{T})$ let $m^{\nu}_{i\to j} := \nu^{(ij)}_{\mathsf{T}}\langle x_i\rangle$ for the probability measure $\nu^{(ij)}_{\mathsf{T}}$ whose Radon–Nikodym derivative with respect to ν_{T} is proportional to $e^{-\beta x_i x_j}$. That is,

$$m^{\nu}_{i\to j} = \frac{\nu_{\mathsf{T}}\langle x_i e^{-\beta x_i x_j}\rangle}{\nu_{\mathsf{T}}\langle e^{-\beta x_i x_j}\rangle} = \lim_{l\to\infty} \frac{\nu_{\mathsf{T}}\langle x_i e^{-\beta x_i x_j} | \underline{x}_{\mathsf{B}_i(l)^c}\rangle}{\nu_{\mathsf{T}}\langle e^{-\beta x_i x_j} | \underline{x}_{\mathsf{B}_i(l)^c}\rangle},$$

where the limit exists by backward martingale convergence theorem and is a.e. constant by the tail triviality of the extremal measure ν_{T} (see [16], Chapter 7). Using the DLR condition (2.2) for ν_{T} and the tree structure of T, we deduce that ν_{T}-a.e.

$$m^{\nu}_{i\to j} = \lim_{l\to\infty} \widetilde{\nu}\big\langle x_i | \underline{x}_{\mathsf{T}_{i\to j}(l,l+1)}, \mathsf{T}_{i\to j}(l+1)\big\rangle. \tag{3.11}$$

By the DCT, the DLR condition (2.2) for ν_{T}, Proposition 3.2 and (3.11), for each $t \in \mathbb{N}$ the marginal law of $\underline{x}_{\mathsf{T}(t)}$ under ν_{T} is completely determined by $\{m^{\nu}_{i\to j}, i \in \partial\mathsf{T}(t), j \in \partial\mathsf{T}(t-1)\}$. In particular, considering the formula (3.6), we get by the same line of reasoning that

$$\nu_{\mathsf{T}}\langle x_i x_j\rangle = F\big(\tanh(\beta), m^{\nu}_{i\to j} m^{\nu}_{j\to i}\big), \tag{3.12}$$

for $F(\gamma, r)$ of (3.7) and any $(i, j) \in E(\mathsf{T})$, with the analogous expression in case of $\nu_{+,\mathsf{T}}\langle x_i x_j\rangle$. Denoting by $m^{-}_{i\to j}$ and $m^{+}_{i\to j}$ the values of $m^{\nu}_{i\to j}$ for Ising–Gibbs measures $\nu_{-,\mathsf{T}}$ and $\nu_{+,\mathsf{T}}$, respectively, from (3.11) and the Griffiths inequality we know that $|m^{\nu}_{i\to j}| \le m^{+}_{i\to j}$ for all $(i, j) \in E(\mathsf{T})$, out of which we get the inequality (3.10) by the strict monotonicity of $r \mapsto F(\gamma, r)$ on $[-1, 1]$ (when $|\gamma| < 1$). Turning to prove that having equality in (3.10) for all $(i, j) \in E(\mathsf{T})$ implies either $\nu_{\mathsf{T}} = \nu_{+,\mathsf{T}}$ or $\nu_{\mathsf{T}} = \nu_{-,\mathsf{T}}$, note that by the preceding such an equality in (3.10) translates into

$$m^{\nu}_{i\to j} m^{\nu}_{j\to i} = m^{+}_{i\to j} m^{+}_{j\to i} \qquad \forall (i, j) \in E(\mathsf{T}). \tag{3.13}$$

From (3.11), one also have by an explicit calculation for Ising measures on trees, that

$$m^{\nu}_{i\to j} = \tanh\Bigg[\sum_{k\in\partial i\setminus\{j\}} \operatorname{atanh}\big(\tanh(\beta) m^{\nu}_{k\to i}\big)\Bigg] \qquad \forall (i, j) \in E(\mathsf{T}), \tag{3.14}$$

with the same recursion holding for the collections $\{m^{\pm}_{i\to j}, (i, j) \in E(\mathsf{T})\}$. Suppose now that some $(i, j) \in E(\mathsf{T})$ is a *plus edge*, namely both $m^{\nu}_{i\to j} = m^{+}_{i\to j}$ and $m^{\nu}_{j\to i} = m^{+}_{j\to i}$. Out of (3.14), we have that $m^{\nu}_{i\to j}$ is *strictly* increasing in each $m^{\nu}_{k\to i}$, $k \in \partial i \setminus \{j\}$, so with $|m^{\nu}_{k\to i}| \le m^{+}_{k\to i}$, the assumed equality $m^{\nu}_e = m^{+}_e$ 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 (3.14) the values of $m^{\nu}_{i\to k}$ and $m^{+}_{i\to k}$ are given by the same function of $\{m^{\nu}_e\}$ and $\{m^{+}_e\}$, respectively, whose arguments are directed edges e where we already have $m^{\nu}_e = m^{+}_e$. Hence, that equality holds also for all directed edges of the form $e = \{i \to k\}$, that is, every edge of $\mathsf{B}_i(1)$ is a plus edge. This property extends in the same manner to $\mathsf{B}_i(t)$, $t = 2, 3, \dots,$ and so we conclude that a single plus edge in T results with each edge being plus edge, and thereby with $\nu_{\mathsf{T}} = \nu_{+,\mathsf{T}}$. By the same line of reasoning, a single *minus edge* (i, j) where both $m^{\nu}_{i\to j} = m^{-}_{i\to j}$ and $m^{\nu}_{j\to i} = m^{-}_{j\to i}$ yields that all edges of T are minus edges and thereby $\nu_{\mathsf{T}} = \nu_{-,\mathsf{T}}$. Suppose now that there are neither plus nor minus edges in T. We then have by (3.13) 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 $\nu_{+,\mathsf{T}}$ having the latter property. Indeed, by (3.14) if some $m^{+}_{i\to j} > 0$ then $m^{+}_{l\to i}$ 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 nonnegative, implies by (3.14) 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^+_e strictly positive at exactly one edge e directed into i. Continuing in this manner, we find an infinite directed ray $\{i_s \to i_{s-1} : (i_s, i_{s-1}) \in E(\mathsf{T}), s \in \mathbb{N}\}$ (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 (3.14), $m^+_{i_s\to i_{s-1}} = \tanh(\beta) m^+_{i_{s+1}\to i_s}$ 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 [28], Lemma 3.2, we may have

$$\sum_{i\in\partial o} \nu_{\mathsf{T}}\langle x_o x_i\rangle = \sum_{i\in\partial o} \nu_{+,\mathsf{T}}\langle x_o x_i\rangle$$

for some $\mathsf{T} \in \mathcal{T}_*$ and an extremal Ising–Gibbs measure $\nu_{\mathsf{T}} \neq \nu_{\pm,\mathsf{T}}$ 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 $\mathsf{T}_{o\to i}$ while $\mathsf{T}' := \mathsf{T}_{i\to o}$ admits an extremal Ising–Gibbs measure other than $\nu_{\pm,\mathsf{T}'}$ [e.g., when $\mathsf{T}_{i\to o}$ is k_2-regular, while $\mathsf{T}_{o\to i}$ is finite or k_1-regular and $\beta_c(k_2) < \beta < \beta_c(k_1)$]. Nevertheless, our next lemma utilizes the unimodularity of μ to circumvent this problem.

LEMMA 3.6. *Fixing $\mu \in \mathcal{U}_*$, for any $\overline{\mathfrak{m}}$ supported on the collection $\mathcal{I}_*$ of Ising–Gibbs measures $\overline{\nu} = \delta_{\mathsf{T}} \otimes \nu_{\mathsf{T}}$ and having the law μ for* T,

$$\overline{\mathfrak{m}}\Big[\sum_{i\in\partial o} \nu_{\mathsf{T}}\langle x_o x_i\rangle\Big] \leq \mu\Big[\sum_{i\in\partial o} \nu_{+,\mathsf{T}}\langle x_o x_i\rangle\Big] = \mu\Big[\sum_{i\in\partial o} \nu_{-,\mathsf{T}}\langle x_o x_i\rangle\Big], \tag{3.15}$$

with strict inequality unless $\overline{\mathfrak{m}}$ is supported on the sub-collection $\mathcal{I}_\pm \subset \mathcal{I}_*$ of $\overline{\nu} = \delta_{\mathsf{T}} \otimes \nu_{\mathsf{T}}$ where

$$\nu_{\mathsf{T}} = \alpha_{\mathsf{T}} \nu^{\beta}_{+,\mathsf{T}} + (1 - \alpha_{\mathsf{T}})\nu^{\beta}_{-,\mathsf{T}} \tag{3.16}$$

for some Borel measurable function $\alpha : \mathcal{T}_* \mapsto [0, 1]$. *Further, w.l.o.g. we take hereafter $\alpha_{\mathsf{T}} = \frac{1}{2}$ on the set* $\{\mathsf{T} \in \mathcal{T}_* : \nu^{\beta}_{+,\mathsf{T}} = \nu^{\beta}_{-,\mathsf{T}}\}$.

REMARK 3.7. In the proof of Lemma 3.6, we take advantage of the $\mathcal{T}_*$-valued Markov chain $\{\widetilde{Y}_\ell\}$ commonly known as "walk from the point of view of the particle", induced by setting the root of T to follow the path of discrete time simple random walk (DSRW) $\{Y_\ell\}$ of law $\widehat{\mathbb{P}}^{\mathsf{T}}_o$ on $(\mathsf{T}, o) \in \mathcal{T}_*$, starting at $Y_0 = o$. Specifically, associating with each $\mu \in \mathcal{U}_*$ for which $\mu\langle\Delta_o\rangle > 0$, the "size-biased-root" probability measure $\widehat{\mu} := \frac{\Delta_o}{\mu\langle\Delta_o\rangle}\mu$ and choosing $\widetilde{Y}_0 \in \mathcal{T}_*$ according to $\widehat{\mu}$, yields the stationary and reversible joint law $\widehat{\mu} \otimes \widehat{\mathbb{P}}^{\mathsf{T}}_o$ for the trajectory $\{\widetilde{Y}_\ell\}$ (cf. [3], Theorem 4.1).

PROOF OF LEMMA 3.6. We get (3.15) by considering the expectation of (3.9) for $i = o$, over the law $\overline{\mathfrak{m}}$ of T and the Ising–Gibbs measure ν_{T} on it. Further, there is only one Ising–Gibbs measure on $\mathsf{T} = \{o\}$. So, our claim about strictness of the inequality in (3.15) trivially holds in case $\mu\langle\Delta_o\rangle = 0$, and assuming hereafter that $\mu\langle\Delta_o\rangle > 0$, we consider the $\mathcal{T}_*$-valued stationary Markov chain $\{\widetilde{Y}_\ell\}$, as in Remark 3.7. Let $\overline{\nu}_{\mathsf{T}}$ denote the expected value of ν_{T} under the probability measure $\overline{\mathfrak{m}}$ conditional upon $\mathsf{T} \in \mathcal{T}_*$, which up to some μ-null set $\mathcal{N} \subset \mathcal{T}_*$ is a uniquely defined Ising–Gibbs measure on T (due to convexity of the latter collection). Equality in (3.15) thus amounts to $\mathbb{E}[f(\widetilde{Y}_0)] = 0$ for the $\mathcal{T}_*$-measurable, uniformly bounded and nonnegative [see (3.9)],

$$f\big((\mathsf{T}, o)\big) := \frac{1}{\Delta_o} \sum_{j \in \partial o} \big[\nu_{+,\mathsf{T}}\langle x_o x_j\rangle - \overline{\nu}_{\mathsf{T}}\langle x_o x_j\rangle\big],$$

which by the stationarity of $\{\widetilde{Y}_\ell\}$ (cf. [3], Theorem 4.1), implies that

$$\mathbb{E}\big[f(\widetilde{Y}_\ell)\big] = 0 \qquad \forall \ell \in \mathbb{N}. \tag{3.17}$$

Conditional on $\widetilde{Y}_0 = (\mathsf{T}, o)$, the probability of $\widetilde{Y}_\ell = (\mathsf{T}, i)$ is strictly positive for each $\ell \in \mathbb{N}$ and $i \in \partial\mathsf{T}(\ell)$, hence with $f(\cdot)$ nonnegative, it follows from (3.17) that

$$\widehat{\mu}\big((\mathsf{T}, o) \in \mathcal{T}_* : \exists i \in V(\mathsf{T}), f\big((\mathsf{T}, i)\big) > 0\big) = 0.$$

We thus conclude that for μ-a.e. T, equality holds in (3.9) for $\overline{\nu}_{\mathsf{T}}$ and *all* $i \in V(\mathsf{T})$, so by Lemma 3.4 the Ising–Gibbs measure $\overline{\nu}_{\mathsf{T}}$ must then be a convex combination of $\nu_{+,\mathsf{T}}$ and $\nu_{-,\mathsf{T}}$. Now recall that to any Ising–Gibbs measure ν_{T} on T corresponds a unique probability measure $\Theta_{\nu_{\mathsf{T}}}$ supported on the collection $\{\nu_{\mathsf{T}}^e\}$ of extremal Ising–Gibbs measures on T, such that $\nu_{\mathsf{T}}(\cdot) = \int \nu_{\mathsf{T}}^e(\cdot)\, d\Theta_{\nu_{\mathsf{T}}}$ (cf. [16], Theorem 7.26). Therefore, by its definition, μ-a.e. $\overline{\nu}_{\mathsf{T}}(\cdot) = \int \nu_{\mathsf{T}}^e(\cdot)\, d\Theta_{\mathsf{T}}$ for the expected value $\Theta_{\mathsf{T}}(\cdot)$ of $\Theta_{\nu_{\mathsf{T}}}(\cdot)$ under the probability measure $\mathfrak{m}$ conditional upon $\mathsf{T} \in \mathcal{T}_*$. We have just shown that μ-a.e. $\Theta_{\mathsf{T}}(\{\nu_{+,\mathsf{T}}, \nu_{-,\mathsf{T}}\}^c) = 0$, hence $\mathfrak{m}$-a.e. this holds for $\Theta_{\nu_{\mathsf{T}}}$. That is, up to some $\mathfrak{m}$-null set, ν_{T} is of the form (3.16), 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 $\overline{\mathfrak{m}}_+$ of $\{\nu_{n,+}\}$ is supported on the collection $\mathcal{I}_\pm$, with T distributed according to μ.*

(b) *Any sub-sequential local weak limit of $\{\nu_n\}$ equals $\overline{\mathfrak{m}} = \mu \circ \overline{\varphi}^{-1}$ for $\overline{\varphi}(\mathsf{T}) = \delta_{\mathsf{T}} \otimes (\frac{1}{2}\nu_{+,\mathsf{T}}^\beta + \frac{1}{2}\nu_{-,\mathsf{T}}^\beta)$.*

PROOF. (a) Recall Lemma 3.3, that

$$\frac{2}{n} \sum_{(i,j) \in E_n} \nu_{n,+}^{\beta,0}\langle x_i x_j\rangle = \mu_n\big[F\big(\overline{\mathsf{P}}_n^2(I_n)\big)\big] \to \mu\Big[\sum_{i \in \partial o} \nu_{+,\mathsf{T}}\langle x_o x_i\rangle\Big], \tag{3.18}$$

for $\overline{\mathsf{P}}^t_n(i)$ corresponding to $\nu_{n,+}$ and the function $F(\overline{\nu}) := \overline{\nu}\langle \sum_{i\in\partial o} x_o x_i \rangle$ on $\mathcal{P}(\overline{\mathcal{G}}_*(2))$, which is bounded by $\overline{\nu}(\Delta_o)$ and continuous with respect to weak convergence. By assumption, under μ_n the law of $\overline{\mathsf{P}}^2_n(I_n)$ converges weakly to $\overline{\mathfrak{m}}^2_+$ along some sub-sequence $n_\ell \to \infty$. Hence, by DCT and the uniform integrability of $\{\Delta_{I_n}\}$,

$$\lim_{\ell\to\infty} \mu_{n_\ell}\big[F\big(\overline{\mathsf{P}}^2_{n_\ell}(I_{n_\ell})\big)\big] = \overline{\mathfrak{m}}_+\big[F\big(\overline{\nu}^2\big)\big]. \tag{3.19}$$

Recall part (b) of Lemma 2.4 that $\overline{\mathfrak{m}}_+$ is supported on the collection $\mathcal{I}_*$ of Ising–Gibbs measures of the form $\delta_{\mathsf{T}} \otimes \nu_{\mathsf{T}}$, having the law $\mu \in \mathcal{U}$ for $\mathsf{T} \in \mathcal{T}_*$. Thus, comparing the RHS of (3.18) with the RHS of (3.19), we deduce that

$$\overline{\mathfrak{m}}_+\bigg[\sum_{i\in\partial o} \nu_{\mathsf{T}}\langle x_o x_i\rangle\bigg] = \mu\bigg[\sum_{i\in\partial o} \nu_{+,\mathsf{T}}\langle x_o x_i\rangle\bigg]$$

out of which it follows by Lemma 3.6 that $\mathfrak{m}_+$ is supported on the sub-collection $\mathcal{I}_\pm$.

(b) Considering now part (a) of Lemma 2.4 we get by the preceding argument that any sub-sequential weak limit $\overline{\mathfrak{m}}$ of $\{\nu_n\}$ is supported on $\mathcal{I}_\pm$ with T distributed according to μ. In particular, $\overline{\mathfrak{m}}$-a.e.

$$\big|\nu_{\mathsf{T}}\langle x_o\rangle\big| = |2\alpha_{\mathsf{T}} - 1|\nu_{+,\mathsf{T}}\langle x_o\rangle.$$

As in the proof of Lemma 3.4, if $\nu_{+,\mathsf{T}}\langle x_o\rangle = 0$, then necessarily $m^+_{i\to j} = 0$ for all $(i,j) \in E(\mathsf{T})$, hence $\nu_{+,\mathsf{T}} = \nu_{-,\mathsf{T}}$ and by our convention $\alpha_{\mathsf{T}} = \frac{1}{2}$. More generally, the bounded function $\widetilde{F}(\overline{\nu}) := |\overline{\nu}\langle x_o\rangle|$ on $\mathcal{P}(\overline{\mathcal{G}}_*(1))$ is continuous with respect to weak convergence. Since

$$0 = n^{-1}\sum_{i=1}^n \big|\nu_n\langle x_i\rangle\big| = \mu_n\big[\widetilde{F}\big(\overline{\mathsf{P}}^1_n(I_n)\big)\big]$$

for all n, it thus follows that for any local weak limit point $\overline{\mathfrak{m}}$ of $\{\nu_n\}$,

$$0 = \overline{\mathfrak{m}}\big[\widetilde{F}\big(\overline{\nu}^1\big)\big] = \overline{\mathfrak{m}}\big[\big|\nu_{\mathsf{T}}\langle x_o\rangle\big|\big] = \overline{\mathfrak{m}}\big[|2\alpha_{\mathsf{T}} - 1|\nu_{+,\mathsf{T}}\langle x_o\rangle\big],$$

thereby forcing $\overline{\mathfrak{m}}$-a.s. $\alpha_{\mathsf{T}} = \frac{1}{2}$. $\square$

4. Proof of Theorem 1.8. Given part (a) of Lemma 3.8 it remains only to show that $\overline{\mathfrak{m}}_+$-a.s., we may take $\alpha_{\mathsf{T}} = 1$ for any sub-sequential local weak limit point $\overline{\mathfrak{m}}_+$ of $\{\nu_{n,+}\}$. To this end, we make use of the following definition.

DEFINITION 4.1. Given graphs $\{\mathsf{G}_n\}_{n\in\mathbb{N}}$ having vertex sets $V_n = [n]$ and probability measures ζ_n on $\mathcal{X}^{[n]}$, let $\overline{\mathsf{P}}^t_n \in \mathcal{P}(\overline{\mathcal{G}}_*(t), \mathscr{C}_{\overline{\mathcal{G}}_*(t)})$ denote the average over a uniformly chosen $I_n \in [n]$, of the law of $(\mathsf{B}_{I_n}(t), \underline{x}_{\mathsf{B}_{I_n}(t)})$, for a positive integer t and $\underline{x}$ drawn according to ζ_n [i.e., $\overline{\mathsf{P}}^t_n = \mu_n(\overline{\mathsf{P}}^t_n(I_n))$ for $\overline{\mathsf{P}}^t_n(i)$ of Definition 1.4].

We say that (G_n, ζ_n), or in short, that ζ_n, converge *on average* to $\overline{\nu}$, a probability measure on $(\overline{\mathcal{G}}_*, \mathscr{C}_{\overline{\mathcal{G}}_*})$, if for any fixed positive integer t,

$$\overline{\mathsf{P}}_n^t \Rightarrow \overline{\nu}^t \qquad \text{as } n \to \infty. \tag{4.1}$$

REMARK 4.2. Note that if $\{\zeta_n\}$ converges locally weakly to $\overline{\mathfrak{m}}$ then it also converges on average to $\overline{\nu} = \int \nu \overline{\mathfrak{m}}(d\nu)$. In particular, if $\overline{\mathfrak{m}}$ is supported on the subset $\mathcal{I}_\pm$ 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 T distributed according to the $\mathcal{P}(\mathcal{T}_*)$-marginal of $\overline{\mathfrak{m}}$ and $\overline{\nu}_{\mathsf{T}}$ of the form (3.16) for some measurable $\alpha : \mathcal{T}_* \mapsto [0, 1]$.

Given $\mathsf{G} \in \mathcal{G}_*$ and $X_0 \in V(\mathsf{G})$, let $\{X_s\}$ denote the variable speed continuous time simple random walk (VSRW) on G, that is, the Markov jump process of state space $V(\mathsf{G})$, which upon arriving at any $j \in V(\mathsf{G})$, jumps with unit rate to each possible $j' \in \partial j$. Now, for $r \in \mathbb{N}$, $l > 0$ and $i \in V(\mathsf{G})$, let $a_{i,j}^{l,r,\mathsf{G}}$ denote the expected relative to l occupation time at $j \in V(\mathsf{G})$ by such VSRW $\{X_s\}$ on G which starts at $X_0 = i$ and run till $\min(l, \theta_r)$ for $\theta_r := \inf\{s \ge 0 : X_s \notin \mathsf{B}_i(r)\}$. That is, with $\mathbb{P}_i^{\mathsf{G}}$ denoting the law of VSRW on the fixed G, starting at $X_0 = i$,

$$a_{i,j}^{l,r,\mathsf{G}} := \frac{1}{l} \int_0^l \mathbb{P}_i^{\mathsf{G}}(X_s = j, s \le \theta_r)\, ds. \tag{4.2}$$

These nonnegative weights induce for every $\underline{x} \in \mathcal{X}^{V(\mathsf{G})}$ the weighted averages

$$y_i^{l,r,\mathsf{G}}(\underline{x}) := \sum_j x_j a_{i,j}^{l,r,\mathsf{G}}, \tag{4.3}$$

having mean value

$$m_i^{l,t,r,\mathsf{G}} := \nu_{+,\mathsf{B}_i(t)}\langle y_i^{l,r,\mathsf{G}}\rangle, \tag{4.4}$$

under the Ising measure $\nu_{+,\mathsf{B}_i(t)}$ on (G, i), at parameters $(\beta, 0)$, conditioned to $\underline{x}_{\mathsf{B}_i(t)^c} = (+)_{\mathsf{B}_i(t)^c}$. Our proof is based on analyzing per $\eta \in (0, 1)$ and $t \in \mathbb{N}$, the functionals

$$\bar{\mathrm{J}}_i^{l,t,r,\mathsf{G}}(\underline{x}, \eta) := \mathrm{J}_i^{l,r,\mathsf{G}}(\underline{x}, \eta)\mathrm{M}_i^{l,t,r,\mathsf{G}}(\eta), \tag{4.5}$$

$$\mathrm{J}_i^{l,r,\mathsf{G}}(\underline{x}, \eta) := \mathbb{I}\{y_i^{l,r,\mathsf{G}}(\underline{x}) \le -\eta\}, \qquad \mathrm{M}_i^{l,t,r,\mathsf{G}}(\eta) := \mathbb{I}\{m_i^{l,t,r,\mathsf{G}} \ge 2\eta\}. \tag{4.6}$$

In doing so, we use $a_{i,j}^{l,r,n}$, $y_i^{l,r,n}$, $m_i^{l,t,r,n}$ ($\mathrm{J}_i^{l,r,n}$, $\mathrm{M}_i^{l,t,r,n}$, $\bar{\mathrm{J}}_i^{l,t,r,n}$), when $\mathsf{G} = \mathsf{G}_n$ and similarly $a_{i,j}^{l,r,\mathsf{T}}$, $y_i^{l,r,\mathsf{T}}$, $m_i^{l,t,r,\mathsf{T}}$ ($\mathrm{J}_i^{l,r,\mathsf{T}}$, $\mathrm{M}_i^{l,t,r,\mathsf{T}}$, $\bar{\mathrm{J}}_i^{l,t,r,\mathsf{T}}$) when $\mathsf{G} = \mathsf{T} \in \mathcal{T}_*$, omitting r and t in case $r = \infty$ (respectively, $t = \infty$, which for $\mathrm{M}_i^{l,r,\mathsf{T}}$ means using $\nu_{+,\mathsf{T}}$), and arguments η, G, $\underline{x}$ whose value is clear from the context.

To explain the role of the various quantities introduced in (4.3)–(4.6), recall that for k-regular graphs [28] fix $\eta > 0$ small so the indicators $\mathrm{J}_i^{l,n}$ identify vertices $i \in \mathsf{G}_n$ in the “-state” of each configuration $\underline{x}$, while the conditioning inherent to $\nu_{n,+}^{\beta}$ keeps at least $\frac{\eta}{2}n$ vertices $i \in \mathsf{G}_n$ out of this state. If we take $a_{i,j} = |\mathsf{B}_i(l)|^{-1}\mathbb{I}_{\{j \in \mathsf{B}_i(l)\}}$ in (4.3), as [28] do, then due to the variability of ball sizes $|\mathsf{B}_i(l)|$ across $i \in \mathsf{G}_n$, we would no longer find a clear relation between $\sum_i y_i$ and the value $\sum_j x_j$ on which we conditioned. We resolve this problem by using instead the weights of (4.2) and taking advantage of the reversibility of the VSRW. Indeed, as we show next, then within the support of $\nu_{n,+}$ one has at least $\frac{\eta}{2}n$ vertices $i \in \mathsf{G}_n$ for which $\mathrm{J}_i^{l,n}(\underline{x}, \eta) = 0$ [and hence $\bar{\mathrm{J}}_i^{l,t,n}(\underline{x}, \eta) = 0$].

LEMMA 4.3. *For any $\eta \in (0, 1)$, $l \geq 0$, $n \in \mathbb{N}$ and $\underline{x}$ such that $\sum_j x_j \geq 0$,*

$$\mu_n\big[1 - \mathrm{J}_{I_n}^{l,n}(\underline{x}, \eta)\big] \geq \frac{\eta}{2}.$$

PROOF. Since $\sum_k a_{j,k}^{l,n} = 1$ for any n, l, we have that $\mathrm{J}_i^{l,n}(\underline{x}, \eta) = \mathbb{I}_{\{z_i \geq 1+\eta\}}$ for the nonnegative $z_i := \sum_j (1 - x_j) a_{i,j}^{l,n}$. Further, due to reversibility of the VSRW, $a_{i,j}^{l,n} = a_{j,i}^{l,n}$ for all $i, j \in V_n$. Hence, by our assumption that $\sum_j x_j \geq 0$,

$$\mu_n[z_{I_n}] = \frac{1}{n}\sum_{k,j=1}^{n}(1 - x_j)a_{k,j}^{l,n} = \frac{1}{n}\sum_{j=1}^{n}(1 - x_j) \leq 1.$$

Thus, applying Markov’s inequality to z_{I_n} completes the proof. $\square$

In the regular case, [28] show that for $\beta > \beta_c$, if $\ell \gg 1$ then $\mathrm{J}_i^{l,n} = \mathrm{J}_j^{l,n}$ for most $(i, j) \in E_n$ which by the assumed edge-expander properties of G_n forces every limit point of $\nu_{n,+}^{\beta}$ to have $\frac{\eta}{2} \geq 1 - \alpha_{\mathsf{T}_k}$ (so taking $\eta \to 0$ completes their proof). To make this argument work, one needs that as $l \to \infty$ the means $m_i^{l,\mathsf{T}}$ be uniformly bounded away from zero, for μ-a.e. T. We have the latter property for $\beta > \beta_c$, provided that μ is an *extremal* element of $\mathcal{U}_*$, since then $m_o^{l,\mathsf{T}}$ converges as $l \to \infty$ to the strictly positive expected magnetization

$$m_\mu := \mu\big[\nu_{+,\mathsf{T}}\langle x_o\rangle\big] \tag{4.7}$$

(see Lemma 4.5 and Remark 4.6). However, for general $\mu \in \mathcal{U}_*$ we have no nontrivial uniform asymptotic lower bound on $m_i^{l,\mathsf{T}}$, so use the indicators $\mathrm{M}_i^{l,n}$ for masking out in (4.5) those $i \in \mathsf{G}_n$ for which $\mathsf{B}_i(t)$ converges to a tree T of too small mean (and we later dispense of this masking effect by taking $\eta \to 0$).

Both for utilizing the reversibility of VSRW and for masking the noise by $\mathrm{M}_o^{l,\mathsf{T}}$ we needed nonlocal functionals, so we in turn approximate these in (4.6) by the local functions corresponding to $r, t \in \mathbb{N}$. Indeed, our next order of business is to use such approximations in relating the relevant functions of $\bar{\mathrm{J}}_{I_n}^{l,t,n}$ to those of $\bar{\mathrm{J}}_o^{l,\mathsf{T}}$ (when $t, n \to \infty$).

LEMMA 4.4. *Suppose* $\mathsf{G}_n \overset{\text{LWC}}{\Longrightarrow} \mu$ *for some* $\mu \in \mathcal{U}_*$, *and* $\{\nu_{n,+}\}$ *converges locally weakly to some* $\overline{\mathfrak{m}}_+$ *supported on* $\mathcal{I}_*$. *Then, with* $\overline{\nu}_+ \in \mathcal{P}(\overline{\mathcal{T}}_*)$ *denoting the corresponding limit on average of* $\{\nu_{n,+}\}$, *for any fixed* l *and except for at most countably many* $\eta > 0$,

$$\lim_{t\to\infty}\lim_{n\to\infty}\mu_n\bigl[\nu_{n,+}\langle\bar{\mathsf{J}}_{I_n}^{l,t,n}\rangle\bigr] = \lim_{t\to\infty}\overline{\nu}_+\bigl[\bar{\mathsf{J}}_o^{l,t,\mathsf{T}}\bigr] = \overline{\nu}_+\bigl[\bar{\mathsf{J}}_o^{l,\mathsf{T}}\bigr], \tag{4.8}$$

$$\begin{aligned}\lim_{t\to\infty}\lim_{n\to\infty}\mu_n\Bigl[\sum_{i\in\partial I_n}\nu_{n,+}\bigl(\bar{\mathsf{J}}_{I_n}^{l,t,n} \neq \bar{\mathsf{J}}_i^{l,t,n}\bigr)\Bigr] &= \lim_{t\to\infty}\overline{\nu}_+\Bigl[\sum_{i\in\partial o}\mathbb{I}\bigl(\bar{\mathsf{J}}_o^{l,t,\mathsf{T}} \neq \bar{\mathsf{J}}_i^{l,t,\mathsf{T}}\bigr)\Bigr] \\ &= \overline{\nu}_+\Bigl[\sum_{i\in\partial o}\mathbb{I}\bigl(\bar{\mathsf{J}}_o^{l,\mathsf{T}} \neq \bar{\mathsf{J}}_i^{l,\mathsf{T}}\bigr)\Bigr].\end{aligned} \tag{4.9}$$

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 $|x_j| \le 1$, for any graph G, positive l, r, and $i \in V(\mathsf{G})$,

$$\bigl|y_i^{l,\mathsf{G}}(\underline{x}) - y_i^{l,r,\mathsf{G}}(\underline{x})\bigr| \le \frac{1}{l}\int_0^l \mathbb{P}_i^{\mathsf{G}}(\theta_r \le s)\,ds \le \mathbb{P}_i^{\mathsf{G}}(\theta_r < l) =: \bar{a}_i^{l,r,\mathsf{G}}. \tag{4.10}$$

In particular, for all $t \in \mathbb{N}$,

$$\bigl|m_i^{l,t,\mathsf{G}} - m_i^{l,t,r,\mathsf{G}}\bigr| \le \bar{a}_i^{l,r,\mathsf{G}} \tag{4.11}$$

and by (4.10)–(4.11), for any graph G, all $i \in V(\mathsf{G})$, positive l, t, r and $\eta > \varepsilon_r \ge 0$,

$$\begin{aligned}\bar{\mathsf{J}}_i^{l,t,r,\mathsf{G}}(\eta+\varepsilon_r) - \mathbb{I}\bigl(\bar{a}_i^{l,r,\mathsf{G}} \ge \varepsilon_r\bigr) &\le \bar{\mathsf{J}}_i^{l,t,\mathsf{G}}(\eta) \\ &\le \bar{\mathsf{J}}_i^{l,t,r,\mathsf{G}}(\eta-\varepsilon_r) + \mathbb{I}\bigl(\bar{a}_i^{l,r,\mathsf{G}} \ge \varepsilon_r\bigr).\end{aligned} \tag{4.12}$$

Further, if the balls $\mathsf{B}_i(t \vee r+1)$ of G_1 and G_2 are isomorphic then $a_{i,j}^{l,r,\mathsf{G}_1} = a_{i,j}^{l,r,\mathsf{G}_2}$ and, restricted to $\mathsf{B}_i(t)$, the Ising measures $\nu_{+,\mathsf{B}_i(t)}$ coincide for both graphs. Consequently,

$$\bar{\mathsf{J}}_i^{l,t,r,\mathsf{G}_1}(\underline{x},\eta) = \bar{\mathsf{J}}_i^{l,t,r,\mathsf{G}_2}(\underline{x},\eta), \qquad \bar{a}_i^{l,r,\mathsf{G}_1} = \bar{a}_i^{l,r,\mathsf{G}_2}. \tag{4.13}$$

Choosing $\varepsilon_r^2 = \mu[\bar{a}_o^{l,r,\mathsf{T}}]$, we get from Markov's inequality that

$$\mu\bigl(\bar{a}_o^{l,r,\mathsf{T}} \ge \varepsilon_r\bigr) \le \varepsilon_r. \tag{4.14}$$

Recall that for $\zeta_n = \nu_{n,+}$, as in Definition 4.1 we assumed that $\overline{\mathsf{P}}_n^s \Rightarrow \overline{\nu}_+^s$ for any fixed $s > t \vee r$, hence by (4.12)–(4.14), for $\eta > 2\varepsilon_r$

$$\begin{aligned}&\overline{\nu}_+\bigl[\bar{\mathsf{J}}_o^{l,t,\mathsf{T}}(\eta+2\varepsilon_r)\bigr] - 2\varepsilon_r \\ &\quad\le \overline{\nu}_+^s\bigl[\bar{\mathsf{J}}_o^{l,t,r,\mathsf{T}}(\eta+\varepsilon_r)\bigr] - \mu\bigl(\bar{a}_o^{l,r,\mathsf{T}} \ge \varepsilon_r\bigr) \\ &\quad\le \liminf_{n\to\infty}\mu_n\bigl[\nu_{n,+}\langle\bar{\mathsf{J}}_{I_n}^{l,t,n}(\eta)\rangle\bigr] \le \limsup_{n\to\infty}\mu_n\bigl[\nu_{n,+}\langle\bar{\mathsf{J}}_{I_n}^{l,t,n}(\eta)\rangle\bigr] \\ &\quad\le \overline{\nu}_+^s\bigl[\bar{\mathsf{J}}_o^{l,t,r,\mathsf{T}}(\eta-\varepsilon_r)\bigr] + \mu\bigl(\bar{a}_o^{l,r,\mathsf{T}} \ge \varepsilon_r\bigr) \le \overline{\nu}_+\bigl[\bar{\mathsf{J}}_o^{l,t,\mathsf{T}}(\eta-2\varepsilon_r)\bigr] + 2\varepsilon_r.\end{aligned} \tag{4.15}$$

Proceeding to show that $\varepsilon_r \downarrow 0$, recall that $\theta_r \geq \tau_r$, the time of the rth jump made by the VSRW $\{X_t\}$ on T. With $\mu\langle\Delta_o\rangle < \infty$, we have by [3], Corollary 4.4, that this continuous time Markov chain is a.s. nonexplosive. That is, $\tau_r \uparrow \infty$ a.s. and hence for $r \to \infty$,

$$\varepsilon_r^2 = \mu\big[\mathbb{P}_o^{\mathsf{T}}(\theta_r < l)\big] \leq \mu\big[\mathbb{P}_o^{\mathsf{T}}(\tau_r \leq l)\big] \to 0.$$

Taking $r \to \infty$ and excluding for $\eta > 0$ the union over $t \in \mathbb{N} \cup \{\infty\}$ of the countably many points of discontinuity for the [0, 1]-valued, nonincreasing, left-continuous $\overline{\nu}_+[\bar{\mathrm{J}}_o^{l,t,\mathsf{T}}(\eta)]$, we deduce that both lower and upper bounds in (4.15) converge to $\overline{\nu}_+[\bar{\mathrm{J}}_o^{l,t,\mathsf{T}}(\eta)]$, thus establishing the left identity of (4.8), as well as the bounds

$$\overline{\nu}_+\big[\bar{\mathrm{J}}_o^{l,t,r,\mathsf{T}}(\eta+\varepsilon_r)\big] - \varepsilon_r \leq \overline{\nu}_+\big[\bar{\mathrm{J}}_o^{l,t,\mathsf{T}}(\eta)\big] \leq \overline{\nu}_+\big[\bar{\mathrm{J}}_o^{l,t,r,\mathsf{T}}(\eta-\varepsilon_r)\big] + \varepsilon_r, \tag{4.16}$$

for all $t, r \in \mathbb{N}$. Further, $\mathrm{M}_i^{l,t,r,\mathsf{T}}(\eta) = \mathbb{I}\{m_o^{l,t,r,\mathsf{T}} \geq 2\eta\}$ and for any fixed $i \in V(\mathsf{T})$ the Ising measures $\nu_{+,(\mathsf{T},i)}^{\beta,0,t}$ of Definition 1.5 converge locally to $\nu_{+,\mathsf{T}}$ 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 $\{\overline{\nu}_+[\bar{\mathrm{J}}_o^{l,r,\mathsf{T}}(\eta \pm \varepsilon_r)], r \in \mathbb{N}\}$, we deduce that

$$\overline{\nu}_+\big[\bar{\mathrm{J}}_o^{l,t,r,\mathsf{T}}(\eta \pm \varepsilon_r)\big] \to \overline{\nu}_+\big[\bar{\mathrm{J}}_o^{l,\mathsf{T}}(\eta)\big],$$

which by (4.16) gives the RHS of (4.8). Turning to prove the left identity in (4.9), since

$$\mu_n\bigg[\sum_{i\in\partial I_n} \nu_{n,+}\big(\bar{\mathrm{J}}_{I_n}^{l,t,n} = 1, \bar{\mathrm{J}}_i^{l,t,n} = 0\big)\bigg] = \mu_n\bigg[\sum_{i\in\partial I_n} \nu_{n,+}\big(\bar{\mathrm{J}}_{I_n}^{l,t,n} = 0, \bar{\mathrm{J}}_i^{l,t,n} = 1\big)\bigg]$$

it suffices to prove that

$$\lim_{n\to\infty} \mu_n\bigg[\sum_{i\in\partial I_n} \nu_{n,+}\big\langle\bar{\mathrm{J}}_{I_n}^{l,t,n}\bar{\mathrm{J}}_i^{l,t,n}\big\rangle\bigg] = \overline{\nu}_+\bigg[\sum_{i\in\partial o} \bar{\mathrm{J}}_o^{l,t,\mathsf{T}}\bar{\mathrm{J}}_i^{l,t,\mathsf{T}}\bigg] =: F_{11}^t(\eta) \tag{4.17}$$

and

$$\lim_{n\to\infty} \mu_n\big[\Delta_{I_n}\nu_{n,+}\big\langle\bar{\mathrm{J}}_{I_n}^{l,t,n}\big\rangle\big] = \overline{\nu}_+\big[\Delta_o\bar{\mathrm{J}}_o^{l,t,\mathsf{T}}\big] =: F_1^t(\eta). \tag{4.18}$$

Both $F_{11}^t(\eta)$ and $F_1^t(\eta)$ are bounded (by $\overline{\nu}_+[\Delta_o] = \mu\langle\Delta_o\rangle$), nonnegative, left-continuous, nonincreasing functions of η. Thus, excluding the at most countably many points of discontinuity of $\eta \mapsto (F_{11}^t(\eta), F_1^t(\eta))$ over all choices of $t \in \mathbb{N} \cup \{\infty\}$, we establish (4.17) and (4.18) upon deriving inequalities analogous to (4.15) for $\sum_{i\in\partial I_n} \bar{\mathrm{J}}_{I_n}^{l,t,n}(\eta)\bar{\mathrm{J}}_i^{l,t,n}(\eta)$ and $\Delta_{I_n}\bar{\mathrm{J}}_{I_n}^{l,t,n}(\eta)$, respectively. These in turn also provide the analogs of (4.16) with $\overline{\nu}_+[\bar{\mathrm{J}}_o^{t,r}(\eta \pm \varepsilon_r)]$ replaced by the nonincreasing in η and uniformly bounded $F_{11}^{t,r}(\eta \pm 2\varepsilon_r)$, $F_1^{t,r}(\eta \pm 2\varepsilon_r)$, respectively, out of which we get the RHS of (4.9) along the same lines we used for deriving the RHS of (4.8). $\square$

In Lemma 4.7, we show that for generic $\eta > 0$, as $l \to \infty$ the RHS of (4.9) goes to zero, whereas the RHS of (4.8) has the limit point

$$g(\eta) := \mu\Big[(1-\alpha_{\mathsf{T}}) \liminf_{l\to\infty} \mathrm{M}_o^{l,\mathsf{T}}(\eta)\Big].$$

Utilizing the edge-expander property of G_n to control the LHS of the corresponding identities allows us to then deduce that $g(\eta) \to 0$ when $\eta \to 0$ out of which we reach the stated conclusion that μ-a.e. $\alpha_{\mathsf{T}} = 1$. To be able to carry this out, we next show that $\mathrm{M}_i^{l,\mathsf{T}}$ is sufficiently regular for $i \in \partial o$, and that $\liminf_l\{m_o^{l,\mathsf{T}}\}$ is uniformly (in T), bounded away from zero, at least μ^e-a.e. for each μ^e which is an extremal element of $\mathcal{U}_*$. For proving the latter result, we recall [3], Corollary 4.4, that every $\mu \in \mathcal{U}_*$ is invariant for the $\mathcal{T}_*$-valued Markov process $s \mapsto \widetilde{X}_s$, where $\widetilde{X}_s = (\mathsf{T}, X_s)$ (for the VSRW $\{X_s\}$ on T, starting at $X_0 = o$), and say that such μ is VSRW-ergodic if all the (continuous)-shift invariant events for $\widetilde{X}_\cdot$ are $\mu \otimes \mathbb{P}_o^{\mathsf{T}}$-trivial.

LEMMA 4.5. *If $\mu \in \mathcal{U}_*$ is VSRW*-ergodic *then*

$$m_o^{l,\mathsf{T}} \to m_\mu \qquad \text{as } l \to \infty, \text{ for } \mu\text{-a.e. } \mathsf{T} \in \mathcal{T}_*. \tag{4.19}$$

Further, for all $\mu \in \mathcal{U}_*$ and any fixed $\varepsilon > 0$,

$$\mu\Big[\sum_{i\in\partial o} \mathbb{I}\big(|m_o^{l,\mathsf{T}} - m_i^{l,\mathsf{T}}| > \varepsilon\big)\Big] \to 0 \qquad \text{as } l \to \infty. \tag{4.20}$$

REMARK 4.6. From [3], we easily deduce that all extremal elements μ^e of the convex set $\mathcal{U}_*$ are VSRW-ergodic. Indeed, this trivially holds if $\mu^e(\Delta_o = 0) = 1$. Otherwise, by extremality $\mu^e(\Delta_o = 0) = 0$, in which case by [3], Theorems 4.6, 4.7, the "size-biased-root" $\widehat{\mu}^e$ is DSRW-ergodic (i.e. all shift invariant events are $\widehat{\mu}^e \otimes \widehat{\mathbb{P}}_o^{\mathsf{T}}$-trivial for the corresponding stationary $\mathcal{T}_*$-valued Markov chain $\{\widetilde{Y}_\ell\}$ of Remark 3.7). Now if μ^e is not VSRW-ergodic then the corresponding stationary $\mathcal{T}_*$-valued Markov chain $\{\widetilde{X}_\ell\}_{\ell\in\mathbb{N}}$ must be nonergodic, hence has some $\mathcal{S} \subseteq \mathcal{T}_*$ with $\mu^e(\mathcal{S}) \in (0,1)$, as a trap set (namely, starting from $\widetilde{X}_0 \in \mathcal{S}$, w.p.1. $\{\widetilde{X}_\ell\} \subseteq \mathcal{S}$, cf. [21], Proposition 1.8). Clearly also $\widehat{\mu}^e(\mathcal{S}) \in (0,1)$, and by the same reasoning, due to the DSRW-ergodicity of $\widehat{\mu}^e$, with positive $\widehat{\mu}^e \otimes \widehat{\mathbb{P}}_o^{\mathsf{T}}$-probability $\widetilde{Y}_0 \in \mathcal{S}$ and the first exit time τ of $\mathcal{S}$ by $\{\widetilde{Y}_\ell\}$ is finite. Recall that the chain $\{\widetilde{Y}_\ell\}$ is embedded at the jump-times of $\{\widetilde{X}_s\}$, so applying the strong Markov property of $\{\widetilde{X}_s\}$ at the stopping time τ, we have that $\widetilde{X}_s = \widetilde{X}_\tau \notin \mathcal{S}$ for all $s \in [\tau, \tau+1]$ with positive $\mu^e \otimes \mathbb{P}_o^{\mathsf{T}}$-probability, in contradiction to $\mathcal{S}$ being a trap set for $\{\widetilde{X}_\ell\}$.

PROOF OF LEMMA 4.5. By definition of $a_{i,j}^{l,\mathsf{T}}$ we have the representation

$$m_o^{l,\mathsf{T}} = \frac{1}{l}\int_0^l \sum_j \nu_{+,\mathsf{T}}\langle x_j\rangle \mathbb{P}_o^{\mathsf{T}}(X_t = j)\,dt = \mathbb{P}_o^{\mathsf{T}}\Big[\frac{1}{l}\int_0^l \nu_{+,\mathsf{T}}\langle x_{X_t}\rangle\,dt\Big]. \tag{4.21}$$

Further, if $\mu \in \mathcal{U}_*$ is VSRW-ergodic then $\mu \otimes \mathbb{P}_o^{\mathsf{T}}$-a.s.

$$\frac{1}{l}\int_0^l \nu_{+,\mathsf{T}}\langle x_{X_t}\rangle\, dt \longrightarrow m_\mu$$

(cf. [21], pages 10–11), which by (4.21) and DCT for conditional expectation, yields the μ-a.e. convergence (4.19). Turning to (4.20), we assume w.l.o.g. that $\mu\langle\Delta_o\rangle > 0$ and setting

$$f^{l,\mathsf{T}} := \sum_{i\in\partial o} \mathbb{I}(|m_o^{l,\mathsf{T}} - m_i^{l,\mathsf{T}}| > 2\varepsilon),$$

note that by the triangle inequality, for any $l \in \mathbb{N}$, $\varepsilon > 0$,

$$\begin{aligned}\mu\langle f^{l,\mathsf{T}}\rangle &\le \mu[\Delta_o \mathbb{I}(|m_o^{l,\mathsf{T}} - m_\mu| > \varepsilon)] + \mu\Big[\sum_{i\in\partial o}\mathbb{I}(|m_i^{l,\mathsf{T}} - m_\mu| > \varepsilon)\Big] \\ &= \frac{1}{\mu\langle\Delta_o\rangle}\Big[\widehat{\mu}[\mathbb{I}(|m_o^{l,\mathsf{T}} - m_\mu| > \varepsilon)] + \widehat{\mu}\Big[\frac{1}{\Delta_o}\sum_{i\in\partial o}\mathbb{I}(|m_i^{l,\mathsf{T}} - m_\mu| > \varepsilon)\Big]\Big] \qquad (4.22) \\ &= \frac{2}{\mu\langle\Delta_o\rangle}\widehat{\mu}(|m_o^{l,\mathsf{T}} - m_\mu| > \varepsilon).\end{aligned}$$

For VSRW-ergodic μ we have, in view of (4.19), the convergence to zero of the bound (4.22). Hence, $\mu\langle f^{l,\mathsf{T}}\rangle \to 0$, namely (4.20) holds for VSRW-ergodic measures, and in particular for all extremal elements of $\mathcal{U}_*$ (by Remark 4.6). Recall that any fixed $\mu \in \mathcal{U}_*$ can be written as a *Choquet integral* of extremal measures [3], Lemma 6.8. So, we have a probability measure Θ on the collection of extremal measures of $\mathcal{U}_*$ such that $\mu\langle f^{l,\mathsf{T}}\rangle = \int \mu^e\langle f^{l,\mathsf{T}}\rangle\, d\Theta(\mu^e)$ for all l. The nonnegative $f^{l,\mathsf{T}}$ are bounded by Δ_o, hence $0 \le \mu^e\langle f^{l,\mathsf{T}}\rangle \le \mu^e\langle\Delta_o\rangle$ for all l. Further, $\int \mu^e\langle\Delta_o\rangle\, d\Theta(\mu^e) = \mu\langle\Delta_o\rangle$ is finite, so by DCT we deduce from the fact that $\mu^e\langle f^{l,\mathsf{T}}\rangle \to 0$ for Θ-a.e. μ^e that $\mu\langle f^{l,\mathsf{T}}\rangle \to 0$. That is, (4.20) holds for all $\mu \in \mathcal{U}_*$. □

Equipped with Lemma 4.5, we proceed to identify the limit as $l \to \infty$ of the relevant functionals from Lemma 4.4.

LEMMA 4.7. *Suppose probability measure* $\overline{\nu}_+ = \mu \otimes \overline{\nu}_{+,\mathsf{T}}$, *with* T *distributed according to* $\mu \in \mathcal{U}_*$ *and* $\overline{\nu}_{+,\mathsf{T}} = \alpha_{\mathsf{T}}\nu_{+,\mathsf{T}} + (1-\alpha_{\mathsf{T}})\nu_{-,\mathsf{T}}$ *for some fixed, measurable* $\alpha : \mathcal{T}_* \mapsto [0,1]$, *with* $\alpha_{\mathsf{T}} = 1$ *whenever* $\nu_{+,\mathsf{T}} = \nu_{-,\mathsf{T}}$. *Then, for any* $\eta > 0$,

$$\lim_{l\to\infty} |\overline{\nu}_+[\bar{\mathrm{J}}_o^{l,\mathsf{T}}] - \mu[(1-\alpha_{\mathsf{T}})\mathrm{M}_o^{l,\mathsf{T}}]| = 0. \qquad (4.23)$$

Furthermore, for Lebesgue a.e. $\eta \in (0,1)$,

$$\liminf_{l\to\infty} \overline{\nu}_+\Big[\sum_{i\in\partial o}\mathbb{I}(\bar{\mathrm{J}}_o^{l,\mathsf{T}} \ne \bar{\mathrm{J}}_i^{l,\mathsf{T}})\Big] = 0. \qquad (4.24)$$

REMARK 4.8. Recall the *branching number* of a rooted tree $\mathsf{T} \in \mathcal{T}_*$,

$$\operatorname{br}\mathsf{T} := \Big\{\lambda > 0 : \inf_{\Pi} \sum_{j \in \Pi} \lambda^{-|j|} = 0\Big\},$$

where $\Pi \subseteq V(\mathsf{T})$ 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 4.7 relies on connections between $\operatorname{br}\mathsf{T}$ and recurrence/transience of the VSRW or phase transitions for Ising models on T (cf. [25, 26]).

PROOF OF LEMMA 4.7. For any T, l, $\eta > 0$ and $i \in V(\mathsf{T})$,

$$\begin{aligned} 1 - \nu_{-,\mathsf{T}}\langle \mathrm{J}_i^{l,\mathsf{T}}(\eta)\rangle &= \nu_{-,\mathsf{T}}\big(y_i^{l,\mathsf{T}} > -\eta\big) \\ &= \nu_{+,\mathsf{T}}\big(y_i^{l,\mathsf{T}} < \eta\big) \ge \nu_{+,\mathsf{T}}\big(y_i^{l,\mathsf{T}} \le -\eta\big) = \nu_{+,\mathsf{T}}\langle \mathrm{J}_i^{l,\mathsf{T}}(\eta)\rangle \end{aligned}$$

and consequently

$$\begin{aligned} D_i^{l,\mathsf{T}}(\eta) &:= \mathrm{M}_i^{l,\mathsf{T}}(\eta)\nu_{+,\mathsf{T}}\big(y_i^{l,\mathsf{T}} < \eta\big) \\ &= \mathrm{M}_i^{l,\mathsf{T}}(\eta)\max\big\{\nu_{+,\mathsf{T}}\langle \mathrm{J}_i^{l,\mathsf{T}}\rangle, 1 - \nu_{-,\mathsf{T}}\langle \mathrm{J}_i^{l,\mathsf{T}}\rangle\big\}. \end{aligned} \tag{4.25}$$

Next, recall that $\overline{\nu}_+[\bar{\mathrm{J}}_i^{l,\mathsf{T}}] = \mu[\mathrm{M}_i^{l,\mathsf{T}}\overline{\nu}_{+,\mathsf{T}}\langle J_i^{l,\mathsf{T}}\rangle]$, for all $l \in \mathbb{N}$ and $i \in V(\mathsf{T})$. So, with $\alpha_{\mathsf{T}} \in [0,1]$ and $\mathrm{J}_o^l = \mathrm{J}_o^{l,\mathsf{T}} \in \{0,1\}$, fixing $\eta > 0$ we get (4.23) by showing that

$$\lim_{l\to\infty} \mu\big[D_o^{l,\mathsf{T}}(\eta)\big] = 0. \tag{4.26}$$

To this end, with $\mathrm{M}_i^{l,\mathsf{T}} = \mathbb{I}(m_i^{l,\mathsf{T}} \ge 2\eta)$ we get by Markov's inequality,

$$\begin{aligned} \mu\big[D_o^{l,\mathsf{T}}(\eta)\big] &\le \mu\big[\nu_{+,\mathsf{T}}\big(y_o^{l,\mathsf{T}} - m_o^{l,\mathsf{T}} < -\eta\big)\mathrm{M}_o^{l,\mathsf{T}}\big] \\ &\le \eta^{-2}\mu\big[\mathrm{Var}_{\nu_{+,\mathsf{T}}}\big(y_o^{l,\mathsf{T}}\big)\mathrm{M}_o^{l,\mathsf{T}}\big] \\ &= \eta^{-2}\mu\Big[\sum_j \mathrm{Cov}_{\nu_{+,\mathsf{T}}}(x_o, x_j)\sum_i a_{o,i}^{l,\mathsf{T}}a_{i,j}^{l,\mathsf{T}}\mathrm{M}_i^{l,\mathsf{T}}\Big], \end{aligned} \tag{4.27}$$

with the latter identity obtained by expanding the variance of $y_o^{l,\mathsf{T}} = \sum_j x_j a_{o,j}^{l,\mathsf{T}}$, then using unimodularity of μ as well as $a_{o,i}^{l,\mathsf{T}} = a_{i,o}^{l,\mathsf{T}}$ (by reversibility of the VSRW on T).

Fixing $r \in \mathbb{N}$, we partition the sum over j in the RHS of (4.27) into Term I consisting of sum over all $j \in \mathsf{B}_o(r)$, and Term II for the sum over $j \notin \mathsf{B}_o(r)$. We then control Term II by confirming for $\gamma := \tanh(\beta) \in (0,1)$ and all $\mathsf{T} \in \mathcal{T}_*$ the uniform correlation decay

$$0 \le \mathrm{Cov}_{\nu_{+,\mathsf{T}}}(x_o, x_j) \le \gamma^{|j|}. \tag{4.28}$$

Indeed, it follows from (4.28), by nonnegativity of $\{a^{l,\mathsf{T}}_{i,j}\}$ and the fact $\sum_{i,j} a^{l,\mathsf{T}}_{o,i} a^{l,\mathsf{T}}_{i,j} = 1$, that

$$\text{Term II} \le \sum_{k=r+1}^{\infty} \gamma^k \mu\Bigg[\sum_{j\in \mathsf{B}_o(k-1,k)} \sum_i a^{l,\mathsf{T}}_{o,i} a^{l,\mathsf{T}}_{i,j}\Bigg] \le \gamma^r. \tag{4.29}$$

Turning to prove (4.28), note that for any tree T the marginal of $\nu_{+,\mathsf{T}}$ on $\underline{x}_{\mathsf{T}'}$ with $\mathsf{T}' = (v_0, v_1, \dots, v_k)$ 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 $\underline{x}_{\mathsf{T}\setminus\mathsf{T}'}$, hence holding also for infinite trees due to (2.5)]. While this tree-indexed Markov chain is in general nonhomogeneous, recall [7], Lemma 4.1, that for any $v \ne w \in V(\mathsf{T}')$ and Ising measure ν on finite T' with $\beta \ge 0$ and *any* external magnetic field parameters, the value of

$$\begin{aligned}\Phi[\nu](v, w) &:= \nu[x_w | x_v = 1] - \nu[x_w | x_v = -1] \\ &= \nu[x_v = -1]^{-1}\big(\nu[x_w | x_v = 1] - \nu[x_w]\big)\end{aligned}$$

is nonnegative (by the Griffiths inequality at $0 = B_v \le B'_v \uparrow \infty$), and maximal at the measure ν_f of zero external magnetic fields. Now, since $x_v \in \{-1, 1\}$, we get that

$$\begin{aligned}\mathrm{Cov}_\nu(x_v, x_w) &= 2\nu[x_v = 1]\nu[x_v = -1]\Phi[\nu](v, w) \le \tfrac{1}{2}\Phi[\nu_f](v, w) \\ &= \mathrm{Cov}_{\nu_f}(x_v, x_w).\end{aligned} \tag{4.30}$$

The tree-indexed Markov chain corresponding to ν_f is homogeneous, of zero-mean and nondegenerate transition probabilities $\pi(y|x) = \frac{1}{2}(1 + xy\gamma)$ on $\{-1, 1\}$, from which we get by direct computation that $\mathrm{Cov}_{\nu_f}(x_{v_0}, x_{v_k}) = \gamma^k$, and (4.28) follows from (4.30).

As for Term I, recall that if $\nu_{+,\mathsf{T}}\langle x_o\rangle = 0$, then $\nu_{+,\mathsf{T}} = \nu_{-,\mathsf{T}}$ and $\mathrm{M}^{l,\mathsf{T}}_i \equiv 0$ for all $i \in V(\mathsf{T})$ and $l \in \mathbb{N}$. Therefore,

$$0 \le \text{Term I} \le \mu\Bigg[\Bigg(\sum_{j\in \mathsf{B}_o(r)} \sum_i a^{l,\mathsf{T}}_{o,i} a^{l,\mathsf{T}}_{i,j}\Bigg)\mathbb{I}\{\nu_{+,\mathsf{T}}\langle x_o\rangle > 0\}\Bigg] \tag{4.31}$$

[the nonnegativity of Term I is due to $\mathrm{Cov}_{\nu_{+,\mathsf{T}}}(x_o, x_j) \ge 0$, per (4.28)]. It is further known that for Ising model on tree T with zero external magnetic field, one has $\nu^{\beta,0}_{+,\mathsf{T}}\langle x_o\rangle > 0$ only for $\beta \ge \beta_c$, where $[\mathrm{br}\,\mathsf{T}]\tanh(\beta_c) = 1$ (see [25], Theorem 1.1). In particular, we bound $\mathbb{I}\{\nu_{+,\mathsf{T}}\langle x_o\rangle > 0\}$ in (4.31) by $\mathbb{I}\{[\mathrm{br}\,\mathsf{T}] > 1\}$, and note that

$$\begin{aligned}\sum_{j\in \mathsf{B}_o(r)} \sum_i a^{l,\mathsf{T}}_{o,i} a^{l,\mathsf{T}}_{i,j} &= \sum_{j\in \mathsf{B}_o(r)} \frac{1}{l^2}\int_0^l \int_0^l \sum_i \mathbb{P}^{\mathsf{T}}_o(X_t = i)\mathbb{P}^{\mathsf{T}}_i(X_s = j)\,dt\,ds \\ &= \frac{1}{l^2}\int_0^l \int_0^l \mathbb{P}^{\mathsf{T}}_o\big(X_{t+s} \in \mathsf{B}_o(r)\big)\,dt\,ds.\end{aligned}$$

In case $[\mathrm{br}\,\mathsf{T}] > 1$, the DSRW on T is transient (see [26], Theorem 4.3). Consequently, for such a tree also $\{X_t\}_{t\ge 0}$ is transient and in particular $1 \ge \mathbb{P}_o^{\mathsf{T}}(X_t \in \mathsf{B}_o(r)) \to 0$ as $t \to \infty$ for any fixed $r \in \mathbb{N}$. By bounded convergence it thus follows that Term I goes to zero as $l \to \infty$, for arbitrarily large (fixed) value of $r \in \mathbb{N}$. Taking $r \to \infty$ we conclude from (4.29) and (4.27) that $\mu[D_o^{l,\mathsf{T}}] \to 0$ as $l \to \infty$, thereby establishing (4.23).

Moving now to the proof of (4.24), for $\{0,1\}$-valued random variables $\mathrm{M}_o = \mathrm{M}_o^{l,\mathsf{T}}$, $\mathrm{M}_i = \mathrm{M}_i^{l,\mathsf{T}}$, $\mathrm{J}_o = \mathrm{J}_o^{l,\mathsf{T}}$ and $\mathrm{J}_i = \mathrm{J}_i^{l,\mathsf{T}}$, we clearly have per T, $l \in \mathbb{N}$ and $i \in \partial o$, that

$$\mathrm{M}_o\mathrm{M}_i \nu_{+,\mathsf{T}}(\mathrm{J}_o \neq \mathrm{J}_i) \le \mathrm{M}_o \nu_{+,\mathsf{T}}\langle \mathrm{J}_o\rangle + \mathrm{M}_i \nu_{+,\mathsf{T}}\langle \mathrm{J}_i\rangle,$$
$$\mathrm{M}_o\mathrm{M}_i \nu_{-,\mathsf{T}}(\mathrm{J}_o \neq \mathrm{J}_i) \le \mathrm{M}_o\bigl(1 - \nu_{-,\mathsf{T}}\langle \mathrm{J}_o\rangle\bigr) + \mathrm{M}_i\bigl(1 - \nu_{-,\mathsf{T}}\langle \mathrm{J}_i\rangle\bigr).$$

Consequently, with $\alpha_{\mathsf{T}} \in [0,1]$ and each $\bar{\mathrm{J}}_j = \mathrm{J}_j\mathrm{M}_j$, we have per T, l, $\eta > 0$ and $i \in \partial o$ that

$$\begin{aligned}\overline{\nu}_{+,\mathsf{T}}(\bar{\mathrm{J}}_o \neq \bar{\mathrm{J}}_i) &\le \mathbb{I}(\mathrm{M}_o \neq \mathrm{M}_i) + \alpha_{\mathsf{T}}\mathrm{M}_o\mathrm{M}_i\nu_{+,\mathsf{T}}(\mathrm{J}_o \neq \mathrm{J}_i) \\ &\quad + (1-\alpha_{\mathsf{T}})\mathrm{M}_o\mathrm{M}_i\nu_{-,\mathsf{T}}(\mathrm{J}_o \neq \mathrm{J}_i) \\ &\le \mathbb{I}(\mathrm{M}_o \neq \mathrm{M}_i) + D_o + D_i,\end{aligned}$$

for $D_i = D_i^{l,\mathsf{T}}$ of (4.25). Taking the expectation with respect to T of unimodular law μ, we thus get that,

$$(4.32)\quad \overline{\nu}_+\biggl[\sum_{i\in\partial o}\mathbb{I}\bigl(\bar{\mathrm{J}}_o^{l,\mathsf{T}} \neq \bar{\mathrm{J}}_i^{l,\mathsf{T}}\bigr)\biggr] \le \mu\biggl[\sum_{i\in\partial o}\mathbb{I}\bigl(\mathrm{M}_o^{l,\mathsf{T}} \neq \mathrm{M}_i^{l,\mathsf{T}}\bigr)\biggr] + 2\mu\bigl[\Delta_o D_o^{l,\mathsf{T}}\bigr].$$

Since $D_o^{l,\mathsf{T}} \in [0,1]$ and $\mu\langle\Delta_o\rangle$ finite, we have from (4.26) that $\mu[\Delta_o D_o^{l,\mathsf{T}}] \to 0$. Turning to deal with the other term on the RHS of (4.32), note that for any $\eta, \varepsilon > 0$, if $\mathrm{M}_o^{l,\mathsf{T}}(\eta) \neq \mathrm{M}_i^{l,\mathsf{T}}(\eta)$, then either $|m_o^{l,\mathsf{T}} - m_i^{l,\mathsf{T}}| > \varepsilon$ or $m_o^{l,\mathsf{T}} \in [2\eta - \varepsilon, 2\eta + \varepsilon)$. Further, with $\mu\langle\Delta_o\rangle$ finite, integrating the nonnegative

$$\mathfrak{E}(\eta) := \liminf_{\varepsilon\to 0}\liminf_{l\to\infty}\mu\bigl[\Delta_o\mathbb{I}\bigl(m_o^{l,\mathsf{T}} \in [2\eta-\varepsilon, 2\eta+\varepsilon)\bigr)\bigr]$$

over η, we get by Fatou's lemma and Fubini's theorem that

$$\int_0^1 \mathfrak{E}(\eta)\,d\eta \le \liminf_{\varepsilon\to 0}\liminf_{l\to\infty}\mu\biggl[\Delta_o\int_0^1\mathbb{I}\bigl(m_o^{l,\mathsf{T}} \in [2\eta-\varepsilon, 2\eta+\varepsilon)\bigr)\,d\eta\biggr] = 0.$$

Consequently, $\mathfrak{E}(\eta) = 0$ for a.e. $\eta \in (0,1)$, in which case the identity (4.20) of Lemma 4.5 completes the proof of (4.24). □

PROOF OF THEOREM 1.8. Recall part (a) of Lemma 3.8 that any subsequential local weak limit point $\overline{\mathfrak{m}}_+$ of $\{\nu_{n,+}\}$, is effectively a distribution over random $\alpha : \mathcal{T}_* \mapsto [0,1]$. From Remark 4.2, we know that to such $\overline{\mathfrak{m}}_+$ corresponds $\overline{\nu}_+ = \mu \otimes \overline{\nu}_{+,\mathsf{T}}$ with $\overline{\nu}_{+,\mathsf{T}} = \alpha_{\mathsf{T}}\nu_{+,\mathsf{T}} + (1-\alpha_{\mathsf{T}})\nu_{-,\mathsf{T}}$ for some fixed measurable

$\alpha : \mathcal{T}_* \to [0, 1]$, where without loss of generality $\alpha_{\mathsf{T}} = 1$ whenever $\nu_{+,\mathsf{T}} = \nu_{-,\mathsf{T}}$ (i.e., $\nu_{+,\mathsf{T}}\langle x_o \rangle = 0$), as done in Lemma 4.7. In particular, it suffices to show that the assumed edge-expansion property of $\{\mathsf{G}_n\}_{n\in\mathbb{N}}$ yields

$$\mu\big[(1-\alpha_{\mathsf{T}})\mathbb{I}\{\nu_{+,\mathsf{T}} \neq \nu_{-,\mathsf{T}}\}\big] = 0, \tag{4.33}$$

for then also $\overline{\mathfrak{m}}_+$-a.e. $\alpha_{\mathsf{T}} = 1$, as claimed. To this end, recall Lemma 4.5 (and Remark 4.6), that for any extremal element μ^e of $\mathcal{U}_*$ and for μ^e-a.e. T,

$$\mu^e\big[\nu_{+,\mathsf{T}}\langle x_o \rangle\big] = \lim_{l\to\infty} m_o^{l,\mathsf{T}}.$$

In particular, setting

$$\mathscr{S}_\pm := \Big\{\mathsf{T} : \nu_{+,\mathsf{T}} \neq \nu_{-,\mathsf{T}}, \liminf_{l\to\infty} m_o^{l,\mathsf{T}} = 0\Big\},$$

we have that $\mu(\mathscr{S}_\pm) = 0$ for each extremal $\mu \in \mathcal{U}_*$ and thus for all $\mu \in \mathcal{U}_*$. Consequently, (4.33) holds as soon as

$$\mu\Big[(1-\alpha_{\mathsf{T}})\mathbb{I}\Big\{\liminf_{l\to\infty} m_o^{l,\mathsf{T}} > 0\Big\}\Big] = 0. \tag{4.34}$$

Now for any l, t, n, η and $\underline{x}$, let

$$W_n := n^{-1}\sum_{i=1}^{n} \bar{\mathsf{J}}_i^{l,t,n},$$

that is, $W_n = n^{-1}|W^{l,t,n}|$ for the subset of vertices

$$W^{l,t,n}(\underline{x}, \eta) := \big\{i \in V_n : \bar{\mathsf{J}}_i^{l,t,n}(\underline{x}, \eta) = 1\big\}.$$

Setting $\delta := \eta/2$ for $\eta > 0$ such that both Lemmas 4.4 and 4.7 hold, recall Lemma 4.3 that whenever $\sum_j x_j \geq 0$

$$1 - W_n \geq 1 - \mu_n\big[\mathsf{J}_{I_n}^{l,n}\big] \geq \delta.$$

Further, since $\{\mathsf{G}_n\}_{n\in\mathbb{N}}$ are $(\delta, 1/2, \lambda_\delta)$ edge-expanders, we have for such $\underline{x}$ that

$$\{W_n \geq \delta\} \Longrightarrow \frac{1}{n}\sum_{(i,j)\in E_n} \mathbb{I}\big(\bar{\mathsf{J}}_i^{l,t,n} \neq \bar{\mathsf{J}}_j^{l,t,n}\big) \geq \lambda_\delta \min\{W_n, 1 - W_n\} \geq \delta\lambda_\delta.$$

Taking the expectation with respect to $\nu_{n,+}$, we find that

$$\frac{1}{2}\mu_n\bigg[\sum_{i\in\partial I_n} \nu_{n,+}\big(\bar{\mathsf{J}}_{I_n}^{l,t,n} \neq \bar{\mathsf{J}}_i^{l,t,n}\big)\bigg] \geq \delta\lambda_\delta \nu_{n,+}(W_n \geq \delta) \geq \delta\lambda_\delta\big(\mu_n\big[\nu_{n,+}\langle\bar{\mathsf{J}}_{I_n}^{l,t,n}\rangle\big] - \delta\big),$$

since $\mathbb{P}(W \geq \delta) \geq \mathbb{E}[W] - \delta$ for any random variable $W \leq 1$ and $\delta > 0$. Considering first the limit over the sub-sequence n_ℓ such that $\nu_{n_\ell,+}$ converges locally weakly to $\overline{\mathfrak{m}}_+$, followed by the limit $t \to \infty$, we deduce from Lemma 4.4 that

$$\overline{\nu}_+\bigg[\sum_{i\in\partial o} \mathbb{I}\big(\bar{\mathsf{J}}_o^{l,\mathsf{T}} \neq \bar{\mathsf{J}}_i^{l,\mathsf{T}}\big)\bigg] \geq 2\delta\lambda_\delta\big(\overline{\nu}_+\langle\bar{\mathsf{J}}_o^{l,\mathsf{T}}\rangle - \delta\big).$$

Hence, considering $l \to \infty$, by Lemma 4.7 and Fatou's lemma we get that,

$$\delta \geq \mu\Big[(1-\alpha_{\mathsf{T}}) \liminf_{l\to\infty} \mathrm{M}_o^{l,\mathsf{T}}(\eta)\Big] \geq \mu\Big[(1-\alpha_{\mathsf{T}})\mathbb{I}\Big\{\liminf_{l\to\infty} m_o^{l,\mathsf{T}} > 2\eta\Big\}\Big]. \tag{4.35}$$

Taking now $\eta \to 0$ along suitable sub-sequence, we arrive at (4.34) and complete the proof. □

5. Continuity of $\mathbb{U}(\cdot, 0)$ in β and edge-expander property. With continuity of $\beta \mapsto \mathbb{U}(\beta, 0)$ 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 $\beta > \beta_\star$, 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 $\mu \in \mathcal{U}_*$ such that μ-a.e. the tree T has minimum degree at least $d_\star > 2$ and set $\beta_\star := \operatorname{atanh}[(d_\star - 1)^{-1}]$. Then $\beta \mapsto \mathbb{U}(\beta, 0)$ is continuous on $(\beta_\star, \infty)$.*

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

LEMMA 5.2. *Suppose $\mu \in \mathcal{U}_*$ and $\beta_c(\mathsf{T}) = \beta_c$ finite, for μ-a.e. infinite T. If*

$$S_{\mathsf{T}}(t) := \sum_{k=1}^{t} (\operatorname{br} \mathsf{T})^{2k} \big|\partial \mathsf{T}(k)\big|^{-2} \tag{5.1}$$

diverges for μ-a.e. infinite T, then $\beta \mapsto \mathbb{U}(\beta, 0)$ is continuous at $\beta = \beta_c$.

REMARK 5.3. Same applies if $|\partial \mathsf{T}(k)|$ in (5.1) taken for size of subset of $\partial \mathsf{T}(k)$ connected to $\partial \mathsf{T}(t)$.

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, $\beta \mapsto \mathbb{U}(\beta, 0)$ is continuous for unimodular measures supported on finite trees. It thus suffices to prove the continuity of $\mathbb{U}(\cdot, 0)$ for super-critical UMGW measures conditioned on nonextinction. Hence, we merely need to verify the assumptions of Lemma 5.2 for such UMGW measures conditioned on nonextinction. To this end, assume first that all entries of the mean matrix $\widehat{M}$ of Definition 1.13 are finite.

• *Branching number*: We need to show that, for super-critical UMGW conditioned on nonextinction, $\beta_c(\mathsf{T}) = \beta_c$ for almost every T. By the one to one relation between $\operatorname{br} \mathsf{T}$ and $\beta_c(\mathsf{T})$ (cf. [25], Theorem 1.1), it suffices to show that conditioned on nonextinction, $\operatorname{br} \mathsf{T}$ is constant UMGW-a.e. This follows from having

$\operatorname{br}\mathsf{T}_{v\to o}$ constant, conditional on nonextinction of $\mathsf{T}_{v\to o}$, for UMGW almost every T and $v \in \partial o$ (since $\operatorname{br}\mathsf{T} = \max_{v\in\partial o}\{\operatorname{br}\mathsf{T}_{v\to o}\}$, with zero branching number for finite trees and the nonextinction of T equivalent to nonextinction of some $\mathsf{T}_{v\to o}$). Each $\mathsf{T}_{v\to o}$ has the same super-critical MGW law corresponding to probability kernels $\widehat{P}_{i,j}$ over the extended type space $\mathcal{Q}_M$, so our claim follows from [26], Proposition 6.5, which says that for any super-critical, positive regular, nonsingular MGW law of finite mean matrix M, regardless of the type of its root-vertex, conditional on its nonextinction 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 nonsingular mean matrix $\widehat{M}$, recall the Kesten–Stigum characterization of the a.s. finite limit of $r(\widehat{M})^{-k}|\partial\mathsf{T}_{v\to o}(k)|$ conditional on nonextinction of $\mathsf{T}_{v\to o}$ (generated according to the MGW law with probability kernels $\widehat{P}_{i,j}$ and type space $\mathcal{Q}_M$, e.g., see [23], Theorem 1,). With Δ_o finite a.s., by the preceding argument it follows that $S_{\mathsf{T}}(t) \to \infty$ a.s. conditional on nonextinction of the UMGW tree.

Turning to the case where some entry of $\widehat{M}$ is infinite, consider the following truncation of $\widehat{P}_{i,j}$:

$$\widehat{P}^{\ell}_{i,j}(\underline{k}) := \widehat{P}_{i,j}(\underline{k})\mathbb{I}_{\{\|\underline{k}\|\le\ell\}} + \mathbb{I}_{\underline{k}=\underline{0}} \sum_{\|\underline{k}'\|>\ell} \widehat{P}_{i,j}(\underline{k}').$$

For all ℓ large enough, both positive regularity and nonsingularity of $\widehat{M}$ are inherited by the finite mean matrices $\widehat{M}^{\ell}$ corresponding to the kernels $\widehat{P}^{\ell}$. Further, positive regularity of the matrix $\widehat{M}$ having some infinite entries implies that $r(\widehat{M}^{\ell}) \to \infty$ as $\ell \to \infty$. Hence, by the preceding proof, upon choosing ℓ large enough, one can make $\operatorname{br}\mathsf{T}_{v\to o}$ under the kernels $\widehat{P}^{\ell}_{i,j}$ uniformly arbitrarily large, conditioned on nonextinction of $\mathsf{T}_{v\to o}$. Since $\operatorname{br}\mathsf{T}_{v\to o}$ under kernels $\widehat{P}^{\ell}_{i,j}$ is stochastically dominated by that for kernels $\widehat{P}_{i,j}$, it follows that conditioned on nonextinction of $\mathsf{T}_{v\to o}$, almost surely $\operatorname{br}\mathsf{T}_{v\to o} = \infty$. Therefore, a.s. $\operatorname{br}\mathsf{T} = \infty$ conditional on nonextinction, and all assumptions of Lemma 5.2 are satisfied. □

To prove Lemma 5.1, we identify functions $\mathbb{U}_\ell(\beta) \le \mathbb{U}(\beta, 0)$ that are nondecreasing in $\ell \in \mathbb{N}$ and $\beta \ge 0$, so the left continuity of $\mathbb{U}(\beta, 0)$ follows by interchanging the order of limits in β and ℓ, provided that

$$\mathbb{U}(\beta, 0) = \lim_{\ell\to\infty} \mathbb{U}_\ell(\beta). \tag{5.2}$$

Indeed, for $\mathsf{T} \in \mathcal{T}_*$, nonnegative β, ℓ and $\{H_v, v \in V(\mathsf{T})\}$, consider the Ising model $\nu^{\beta,\{H_v\}}_{\mathsf{T}(\ell)}$ of (3.1), for graph $\mathsf{T}(\ell)$, inverse temperature parameter β and external field $B_v = H_v\mathbb{I}_{v\in\partial\mathsf{T}(\ell)}$, with $m_\ell(\{H_v\}) = \nu^{\beta,\{H_v\}}_{\mathsf{T}(\ell)}\langle x_o\rangle$ denoting its root magnetization. Key to the proof of (5.2) is the joint continuity property (5.3) of $(\beta, \ell) \mapsto m_\ell(\{h^{\beta'}_v\})$, where

$$h^{\beta'}_v := \operatorname{atanh}\big(\nu^{\beta',0}_{+,\mathsf{T}_{v\to o}}\langle x_v\rangle\big), \qquad v \in V(\mathsf{T})$$

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

LEMMA 5.4. *If $\beta > \beta_0$ such that $(d_\star - 1)\tanh(\beta_0) > 1$, then there exists $\kappa = \kappa(\beta, \beta_0, d_\star)$ finite such that for any $\mathsf{T} \in \mathcal{T}_*$ of minimum degree at least $d_\star > 2$ and all $\ell \geq 1$,*

$$0 \leq \ell\big[m_\ell(\{h_v^\beta\}) - m_\ell(\{h_v^{\beta_0}\})\big] \leq \kappa. \tag{5.3}$$

PROOF. Fixing $\beta > \beta_0 > 0$, let $\gamma := \tanh(\beta)$, $\gamma_0 := \tanh(\beta_0)$. Using $v \hookrightarrow w$ to denote that v is the parent of w in $\mathsf{T} \in \mathcal{T}_*$, the identity (3.14) becomes

$$h_v^\beta = \sum_{\{w:v\hookrightarrow w\}} f_\gamma(h_w^\beta), \tag{5.4}$$

for $f_\gamma(h) := \operatorname{atanh}(\gamma \tanh(h))$. Since $g : [0, 1] \to (1, \infty)$ given by

$$g(0) = \frac{\gamma}{\gamma_0}, \qquad g(r) = \frac{\operatorname{atanh}(\gamma r)}{\operatorname{atanh}(\gamma_0 r)} \qquad \forall r \in (0, 1],$$

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, the Griffiths inequality and our uniform lower bound on $g(\cdot)$, for any $k \geq 0$ we have

$$\begin{aligned} m_{k+1}(\{h_w^{\beta_0}\}) &= m_k\Big(\Big\{\sum_{\{w:v\hookrightarrow w\}} f_\gamma(h_w^{\beta_0})\Big\}\Big) \\ &= m_k\Big(\Big\{\sum_{\{w:v\hookrightarrow w\}} g(\tanh(h_w^{\beta_0}))f_{\gamma_0}(h_w^{\beta_0})\Big\}\Big) \\ &\geq m_k\Big(\Big\{\sum_{\{w:v\hookrightarrow w\}} (1+\varepsilon)f_{\gamma_0}(h_w^{\beta_0})\Big\}\Big) = m_k(\{(1+\varepsilon)h_v^{\beta_0}\}), \end{aligned} \tag{5.5}$$

with the last equality due to (5.4). The minimum degree of T is at least $d_\star$, so we have by the Griffiths inequality that $h_w^{\beta_0} \geq h_\star^{\beta_0}$ for all $w \in V(\mathsf{T})$ and $h_\star^{\beta_0} := \operatorname{atanh}(r_\star^{\beta_0})$ with $r_\star^{\beta_0}$ the positive root magnetization for Ising plus measure on the $(d_\star - 1)$-ary tree, at parameter β_0 (which by assumption exceeds the critical parameter for Ising measure on the regular tree $\mathsf{T}_{d_\star}$). It then follows from (5.4) that moreover $h_v^{\beta_0} \geq \xi\Delta_v$, with $\xi := \frac{1}{2}f_{\gamma_0}(h_\star^{\beta_0})$ strictly positive. Using (5.4) once more, we see that $h_v^\beta \leq f_\gamma(1)\Delta_v = \beta\Delta_v$ for all $v \in V(\mathsf{T})$. Thus, by the Griffiths inequality,

$$m_{k+1}(\{h_w^\beta\}) = m_k(\{h_v^\beta\}) \leq m_k(\{\beta\Delta_v\}) \leq m_k(\{(\beta/\xi)h_v^{\beta_0}\}). \tag{5.6}$$

Choosing $\varepsilon > 0$ small enough, we have $\beta/\xi = 1 + \kappa\varepsilon$ with $\kappa > 1$ finite, hence by the concavity on $\mathbb{R}_+$ of $\lambda \mapsto m_k(\{\lambda H_v\})$, for each $k \geq 0$ and nonnegative $\{H_v\}$ (which is a special case of the GHS inequality, see [19]), we get the inequality,

$$m_k(\{(\beta/\xi)h_v^{\beta_0}\}) - m_k(\{h_v^{\beta_0}\}) \leq \kappa\big[m_k(\{(1+\varepsilon)h_v^{\beta_0}\}) - m_k(\{h_v^{\beta_0}\})\big]. \tag{5.7}$$

Combining (5.5), (5.6) and (5.7), we deduce that

$$m_{k+1}(\{h_w^\beta\}) - m_{k+1}(\{h_w^{\beta_0}\}) \le \kappa[m_{k+1}(\{h_w^{\beta_0}\}) - m_k(\{h_w^{\beta_0}\})].$$

Recall, for example, from (5.5), that $k \mapsto m_k(\{h_v^{\beta_0}\}) \in [0, 1]$ is nondecreasing, and bounded above by $m_k(\{h_v^\beta\})$ which is independent of k. Hence, summing the latter inequality over $k = 0, \dots, \ell - 1$ results with

$$0 \le \ell[m_\ell(\{h_v^\beta\}) - m_\ell(\{h_v^{\beta_0}\})] \le \sum_{k=1}^{\ell}[m_k(\{h_w^\beta\}) - m_k(\{h_w^{\beta_0}\})] \le \kappa m_\ell(\{h_w^{\beta_0}\}) \le \kappa,$$

as claimed. □

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 $\mathsf{T}(\ell) \cap \mathsf{T}_{o\to i}$, as well as when it is replaced by the magnetization at i for such Ising models on $\mathsf{T}(\ell + 1) \cap \mathsf{T}_{i\to o}$. Hereafter, we denote the former by $m_{\ell,o\to i}(\cdot)$ and the latter by $m_{\ell+1,i\to o}(\cdot)$.

REMARK 5.6. For UGW measure μ, the variables $\{h_v^{\beta'}, v \neq o\}$ are identically distributed, each having the law we called $h^{\beta',+}$ in Lemma 1.18. Starting the recursion (1.11) with $h^{(0)} \stackrel{d}{=} h^{\beta_0,+}$ yields the sequence $h^{(\ell)}$ having the laws of $\operatorname{atanh}(m_{\ell+1,i\to o}(\{h_v^{\beta_0}\}))$. We have just coupled these with $\operatorname{atanh}(m_{\ell+1,i\to o}(\{h_v^\beta\}))$ whose law equals $h^{\beta,+}$, establishing the convergence in law of Lemma 1.18 (and by the Griffiths 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 $\mathbb{U}(\beta, 0)$ is left continuous at β. To this end, for any infinite T and integer $\ell \ge 1$, using the Ising model $\nu_{\mathsf{T}(\ell)}^{\beta,\{h_w^{\beta_0}\}}$ on $\mathsf{T}(\ell)$ with positive external field only at $\partial\mathsf{T}(\ell)$, as in Lemma 5.4, we define

$$\mathbb{U}_\ell(\beta) = \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{i\in\partial o} \nu_{\mathsf{T}(\ell)}^{\beta,\{h_v^{\beta_0}\}}\langle x_o x_i\rangle\Big]. \tag{5.8}$$

With $\mathsf{T}(\ell)$ a finite graph, fixing β_0 and ℓ, the function $\beta \mapsto \mathbb{U}_\ell(\beta)$ is continuous and nondecreasing (by the Griffiths inequality). By Proposition 3.2, we further have that

$$\mathbb{U}_{\ell+1}(\beta) = \frac{1}{2}\mu\Big[\sum_{i\in\partial o} \nu_{\mathsf{T}(\ell)}^{\beta,\{H_v\}}\langle x_o x_i\rangle\Big],$$

and since $\beta > \beta_0$, it follows from (5.4) and the monotonicity of $\gamma \mapsto f_\gamma(h)$, that for any $v \in \partial\mathsf{T}(\ell)$,

$$H_v := \sum_{\{w:v\hookrightarrow w\}} f_\gamma(h_w^{\beta_0}) \ge \sum_{\{w:v\hookrightarrow w\}} f_{\gamma_0}(h_w^{\beta_0}) = h_v^{\beta_0}.$$

By yet another appeal to the Griffiths inequality, we deduce that $\ell \mapsto \mathbb{U}_\ell(\beta)$ is also nondecreasing. Recall that $h_v^{\beta} \geq h_v^{\beta_0}$ for all $v \in V(\mathsf{T})$, so by similar reasoning, $\mathbb{U}_\ell(\beta) \leq \mathbb{U}(\beta, 0)$ and as explained before it remains only to establish (5.2). To this end, in view of (3.12), we have that for any $i \in \partial o$ and $\{H_v, v \in V(\mathsf{T})\}$,

$$\nu_{\mathsf{T}(\ell)}^{\beta,\{H_v\}}\langle x_o x_i\rangle = F\bigl(\gamma, m_{\ell,i\to o}\bigl(\{H_v\}\bigr) m_{\ell,o\to i}\bigl(\{H_v\}\bigr)\bigr),$$

where $F(\gamma, r)$ of (3.7) is continuous and bounded on $[0, 1]^2$. Thus, with $\Psi(\gamma, \delta) := \sup\{|F(\gamma, r) - F(\gamma, r')| \text{ over } r, r' \in [0, 1] \text{ such that } |r - r'| \leq \delta\}$ and $\delta_\ell := 2\kappa/(\ell - 1)$, clearly $\Psi(\gamma, \delta_\ell) \to 0$ as $\ell \to \infty$. Now, in view of Remark 5.5, the expression (5.8) for $\mathbb{U}_\ell(\beta)$ and the corresponding expression for $\mathbb{U}(\beta, 0)$, we deduce that

$$\bigl|\mathbb{U}(\beta, 0) - \mathbb{U}_\ell(\beta)\bigr| \leq \tfrac{1}{2}\Psi(\gamma, \delta_\ell)\mu\langle\Delta_o\rangle,$$

from which (5.2) follows. □

REMARK 5.7. It is easy to see that the proof of Lemma 5.1 applies at any $\beta \geq 0$ and $\mu \in \mathcal{U}_*$ such that for some $\beta_0 < \beta$ one has a bound of the type (5.3), that is, as soon as $m_\ell(\{h_v^{\beta}\}) - m_\ell(\{h_v^{\beta_0}\}) \to 0$ in probability, when $\ell \to \infty$. Further, the proof of (5.3) is completely general, *except* for requiring in (5.6) that $h_v^{\beta}/h_v^{\beta_0}$ (alternatively, $r_v^{\beta}/r_v^{\beta_0}$), be uniformly bounded over $v \in V(\mathsf{T})$. Unfortunately, while $h_v^{\beta_0}$ is strictly positive as soon as $\beta_0 > \beta_c(\mathsf{T})$, even for UGW μ, when $\beta \in (\beta_c, \beta_\star)$ 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 $\beta > \beta_\star$.

REMARK 5.8. Lemmas 5.4 and 1.18 are the analogs of [9], Lemma 4.3, and [9], Lemma 2.3, respectively, in case of zero external field and low temperature (i.e., $\beta > \beta_\star$). While we do not pursue this here, utilizing the former one can establish similar conclusions as done in [9] based on [9], Lemmas 2.3 and 4.3.

The proof of Lemma 5.2 builds on results from [32], to which end we introduce few relevant definitions and notation. First, for any finite $(\mathsf{T}, o) \in \mathcal{T}_*$ let $\partial_\star\mathsf{T}$ denote the collection of *rays* emanating from o, namely finite nonbacktracking paths in one-to-one correspondence with the leaves of T other than o (where each such ray terminates). Next, a *flow* ϖ on such (T, o) is a nonnegative function on $E(\mathsf{T})$, of *strength* $|\varpi| := \sum_{y:o\hookrightarrow y}\varpi(oy)$, such that $\varpi(vw) = \sum_{y:w\hookrightarrow y}\varpi(wy)$, whenever $v \hookrightarrow w$ and $w \notin \partial_\star\mathsf{T}$. Any given collection of *resistances* $\{R(e) \geq 0 : e \in E(\mathsf{T})\}$, induces the functional

$$V_\varpi := \sup\biggl\{\sum_{e\in y}\bigl(\varpi(e)R(e)\bigr)^2 : y \in \partial_\star\mathsf{T}\biggr\},$$

over flows ϖ on T, in terms of which we define

$$\operatorname{cap}_3(\mathsf{T}) := \sup\{|\varpi| : \varpi \text{ a flow on } \mathsf{T} \text{ with } V_\varpi = 1\}.$$

PROOF OF LEMMA 5.2. For any $(\mathsf{T}, o) \in \mathcal{T}_*$ and $e = vw \in E(\mathsf{T})$ let $|e| = |v| \vee |w|$ where $|v|$ denotes the graph distance between $v \in V(\mathsf{T})$ and o. From [32], Lemma 4.2, we know that for any $\gamma > 0$ there exists $\kappa > 0$ such that

$$f_\gamma(h) \le \frac{\gamma h}{(1 + (\kappa h)^2)^{1/2}} \tag{5.9}$$

for $f_\gamma(\cdot)$ of (5.4) and all $h \ge 0$. Further, recall that for any finite $t \ge 1$, $\gamma = \tanh(\beta) > 0$ and infinite tree $(\mathsf{T}, o) \in \mathcal{T}_*$ without leaves the positive

$$h_v^{(t)}(\mathsf{T}) := \operatorname{atanh}\big(\nu^{\beta,0,t}_{+,\mathsf{T}(t)_{v\to o}}\langle x_v\rangle\big),$$

satisfies the system of equations (5.4) at all $|v| < t$, starting with $h_w^{(t)}(\mathsf{T}) = \infty$ when $|w| = t$ [i.e. $w \in \partial\mathsf{T}(t)$]. More generally, in case (T, o) has leaves, let $\mathsf{T}_t \subseteq \mathsf{T}(t)$ denote the union of all vertices and edges along rays of $\mathsf{T}(t)$ of length t, emanating from o. All nonroot leaves of T_t are at distance t from o and it is easy to verify that $h_v^{(t)}(\mathsf{T}) = h_v^{(t)}(\mathsf{T}_t)$ satisfy for $v \in \mathsf{T}_t$ the corresponding equations (5.4) on T_t, starting with $h_w^{(t)}(\mathsf{T}_t) = \infty$ at $w \in \partial\mathsf{T}(t)$. In view of (5.9), it then follows from [32], Theorem 3.2, that

$$h_o^{(t)}(\mathsf{T}) \le \kappa^{-1} \operatorname{cap}_3(\mathsf{T}_t), \tag{5.10}$$

for $\operatorname{cap}_3(\mathsf{T}_t)$ corresponding to resistances $R(e) = \gamma^{-|e|}$ on (T_t, o). Set $\gamma = \tanh(\beta)$ for $\beta = \beta_c(\mathsf{T})$ finite, namely $\gamma = 1/(\operatorname{br}\mathsf{T})$ (see [25], Theorem 1.1). If such $\operatorname{cap}_3(\mathsf{T}_t) \to 0$ for $t \to \infty$, then by (5.10) we deduce that

$$\nu^{\beta,0}_{+,\mathsf{T}}\langle x_o\rangle = \lim_{t\to\infty} \tanh\big(h_o^{(t)}(\mathsf{T})\big) \le 0,$$

so at $\beta = \beta_c(\mathsf{T})$ there is then a *unique* Ising–Gibbs measure on (T, o). Now, should this happen for μ-a.e. infinite T at the same $\beta_c(\mathsf{T}) = \beta_c$, then necessarily $\mathbb{U}(\beta_c, 0) = 0$ and in particular $\beta \mapsto \mathbb{U}(\beta, 0)$ is continuous at $\beta = \beta_c$. With $\mathsf{T}_t \subseteq \mathsf{T}(t)$, clearly

$$S_{\mathsf{T}_t} := \sum_{k=1}^{t} \gamma^{-2k} |\partial\mathsf{T}_t(k)|^{-2} \ge S_{\mathsf{T}}(t)$$

of (5.1), so it suffices to confirm that $\operatorname{cap}_3(\mathsf{T}_t) \le S_{\mathsf{T}_t}^{-1/2}$ (see also Remark 5.3). To this end, fixing $t \ge 1$ let ϖ be any flow on T_t of strength $|\varpi| = 1$. Then, by the definition of V_ϖ, for any probability measure $p_\star(\cdot)$ on $\partial_\star\mathsf{T}_t$,

$$V_\varpi \ge \sum_{y \in \partial_\star \mathsf{T}_t} \Big[\sum_{e\in y} \varpi^2(e)\gamma^{-2|e|}\Big] p_\star(y) = \sum_{k=1}^{t} \gamma^{-2k} \sum_{|e|=k} \varpi^2(e) \sum_{y \ni e} p_\star(y). \tag{5.11}$$

With slight abuse of notation, set $p_\star(e) := \sum_{y \ni e} p_\star(y)$. Note that the thus defined $\{p_\star(e), e \in E(\mathsf{T}_t)\}$, constitutes a flow of strength $|p_\star| = 1$. Further, $\sum_{|e|=k} p_\star(e) = 1$ for any $1 \le k \le t$ since all nonroot leaves of T_t are at $\partial \mathsf{T}_t(t)$. Applying the Cauchy–Schwarz inequality and choosing $p_\star = \varpi$, we find that

$$\left[\sum_{|e|=k} \varpi^2(e) p_\star(e)\right] \ge \left(\sum_{|e|=k} \varpi(e) p_\star(e)\right)^2 = \left(\sum_{|e|=k} \varpi^2(e)\right)^2. \tag{5.12}$$

Using the Cauchy–Schwarz inequality once more,

$$\left(\sum_{|e|=k} \varpi^2(e)\right) \ge \frac{1}{|\partial \mathsf{T}_t(k)|}\left(\sum_{|e|=k} \varpi(e)\right)^2 = \frac{1}{|\partial \mathsf{T}_t(k)|}. \tag{5.13}$$

Thus, from (5.11), (5.12) and (5.13), we see that $V_\varpi \ge S_{\mathsf{T}_t}$ for any flow ϖ on T_t such that $|\varpi| = 1$. By simple scaling, it then follows that $\mathrm{cap}_3(\mathsf{T}_t) \le S_{\mathsf{T}_t}^{-1/2}$, as claimed. □

PROOF OF LEMMA 1.17. For each $i \in \mathcal{Q}$ and $\underline{k} \in \mathbb{Z}_{\ge}^{|\mathcal{Q}|}$ let $\alpha_{i,\underline{k}} = p(i) P_i(\underline{k})$, viewed as coordinates of the collection

$$\underline{\alpha} = (\alpha_{i,\underline{k}})_{i \in \mathcal{Q}, \underline{k} \in \mathbb{Z}_{\ge}^{|\mathcal{Q}|}}$$

[which is finite by assumption of bounded support for all $P_i(\cdot), i \in \mathcal{Q}$]. Fixing $\delta_0 \le 1/2$, for any vector $\underline{\delta} = (\delta_{i,\underline{k}})_{i \in \mathcal{Q}, \underline{k} \in \mathbb{Z}_{\ge}^{|\mathcal{Q}|}}$ such that $\|\underline{\delta}\| \in (\delta_0, 1/2)$, let $W_{\underline{\delta}}$ denote a subset of $[n\|\underline{\delta}\|]$ vertices from V_n where for each i and $\underline{k}$, about $n\delta_{i,\underline{k}}(1 + o(1))$ of the vertices of $W_{\underline{\delta}}$ are of type i and off-springs configuration $\underline{k}$. For any n and $\varepsilon \ge 0$ denote by $\mathcal{G}_{\underline{\delta}}^{\varepsilon,n}$ the event that within G_n there exists some $W_{\underline{\delta}}$ having precisely $[n\varepsilon]$ edges between $W_{\underline{\delta}}$ to $W_{\underline{\delta}}^c$. By Definition 1.12, with high probability, for all large n and each i, $\underline{k}$, there are $n\alpha_{i,\underline{k}}(1 + o(1))$ vertices of type $i \in \mathcal{Q}$ and off-springs configuration $\underline{k}$ in the random graph G_n. In particular, with high probability only events $\mathcal{G}_{\underline{\delta}}^{\varepsilon,n}$ having

$$\delta_{i,\underline{k}} \le \alpha_{i,\underline{k}} \qquad \forall i, \underline{k} \tag{5.14}$$

occur. We have the stated edge-expansion property upon the existence of $\varepsilon_0 := \varepsilon_0(\delta_0) > 0$ such that the probability of the union of all such $\mathcal{G}_{\underline{\delta}}^{\varepsilon,n}$ for which (5.14) holds, $\|\underline{\delta}\| \in (\delta_0, 1/2)$ and $\varepsilon \le \varepsilon_0$, goes to zero as $n \to \infty$. Vertex types and edge counts are integer valued, so with both the length of $\underline{\delta}$ and $\varepsilon \le n^{-1}|E_n|$ uniformly bounded, we have at most n^C such events to rule out. Consequently, it suffices to show that for any $\underline{\delta} \in (\delta_0, 1/2)$ satisfying (5.14) and $\varepsilon \le \varepsilon_0$,

$$\frac{1}{n} \log \mathbb{P}(\mathcal{G}_{\underline{\delta}}^{\varepsilon,n}) < -\varepsilon < 0, \tag{5.15}$$

for all large n, uniformly over all such choices of $\underline{\delta}$ and ε. To this end, we first note that for $\varepsilon = 0$,

$$\begin{aligned}\frac{1}{n}\log \mathbb{P}(\mathcal{G}_{\underline{\delta}}^{0,n}) &= \frac{1}{n}\log \#\{\text{choices possible for } W_{\underline{\delta}}\}\\ &\quad + \frac{1}{n}\log \mathbb{P}\{\text{such choice matches with itself}\}\\ &=: N_{\underline{\delta}} + Q_{\underline{\delta}}.\end{aligned}$$

We further define $\alpha_{i,j} := \sum_{\underline{k}} k_j \alpha_{i,\underline{k}}$ and $\delta_{i,j} := \sum_{\underline{k}} k_j \delta_{i,\underline{k}}$ for each $i, j \in \mathcal{Q}$. Using the approximations

$$\frac{1}{n}\log n! = \log\left(\frac{n}{e}\right) + o(1) \quad \text{and} \quad \frac{1}{n}\log n!! = \frac{1}{2}\log\left(\frac{n}{e}\right) + o(1),$$

we have for $H(q) := -q\log q - (1-q)\log(1-q)$, $q \in [0,1]$, that

$$N_{\underline{\delta}} \approx \sum_{i,\underline{k}} \alpha_{i,\underline{k}} H\left(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\right) = \sum_{i}\sum_{j}\sum_{\underline{k}} \frac{k_j \alpha_{i,\underline{k}}}{\|\underline{k}\|} H\left(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\right), \tag{5.16}$$

$$Q_{\underline{\delta}} \approx -\frac{1}{2}\sum_{i\in\mathcal{Q}} \alpha_{i,i} H\left(\frac{\delta_{i,i}}{\alpha_{i,i}}\right) - \sum_{i\neq j\in\mathcal{Q}} \alpha_{i,j} H\left(\frac{\delta_{i,j}}{\alpha_{i,j}}\right). \tag{5.17}$$

By concavity of $H(\cdot)$, upon noting that $\|\underline{k}\| \ge 3$ we have for any $i, j \in \mathcal{Q}$, that

$$\begin{aligned}\sum_{\underline{k}} \frac{k_j \alpha_{i,\underline{k}}}{\|\underline{k}\|} H\left(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\right) - \frac{1}{2}\alpha_{i,j} H\left(\frac{\delta_{i,j}}{\alpha_{i,j}}\right) &\le \frac{1}{3}\sum_{\underline{k}} k_j \alpha_{i,\underline{k}} H\left(\frac{\delta_{i,\underline{k}}}{\alpha_{i,\underline{k}}}\right) - \frac{1}{2}\alpha_{i,j} H\left(\frac{\delta_{i,j}}{\alpha_{i,j}}\right)\\ &\le -\frac{1}{6}\alpha_{i,j} H\left(\frac{\delta_{i,j}}{\alpha_{i,j}}\right).\end{aligned}$$

With $\|\underline{\delta}\| \le 1/2 < \|\underline{\alpha}\| = 1$ for $\underline{\delta}$ satisfying (5.14), we must have $\delta_{i,j} < \alpha_{i,j}$ for at least one pair (i, j). We thus get from (5.16) and (5.17) that

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\mathcal{G}_{\underline{\delta}}^{0,n}) \le -\frac{1}{6}\sum_{i,j\in\mathcal{Q}} \alpha_{i,j} H\left(\frac{\delta_{i,j}}{\alpha_{i,j}}\right), \tag{5.18}$$

with the RHS strictly negative [since $H(q) = 0$ only for $q \in \{0, 1\}$]. Further, the approximations in (5.16) and (5.17) are uniform over $\underline{\delta}$, because

$$\sqrt{2\pi} \le \frac{n!}{n^{n+1/2}e^{-n}} \le e \qquad \text{for all } n. \tag{5.19}$$

The supremum of the upper bound of (5.18), over the compact set of all possible choices of $\underline{\delta}$ is strictly negative, yielding (5.15) for $\varepsilon = 0$. Similar rational applies also for all ε small enough. For example, in case $|\mathcal{Q}| = 1$ we have for $\delta := \sum_k k\delta_k$

and $\alpha := \sum_k k\alpha_k$, that $\delta < \alpha$ and

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\mathcal{G}^{\varepsilon,n}_{\underline{\delta}}) \le -\frac{\alpha}{6}H\left(\frac{\delta}{\alpha}\right) + \frac{\delta}{2}H\left(\frac{\varepsilon}{\delta}\right) + \frac{1}{2}(\alpha-\delta)H\left(\frac{\varepsilon}{\alpha-\delta}\right)$$
$$\le -\frac{\alpha}{6}H\left(\frac{\delta}{\alpha}\right) + \frac{\alpha}{2}H\left(\frac{2\varepsilon}{\alpha}\right).$$

The preceding bound is continuous in ε and strictly negative at $\varepsilon = 0$. Consequently, there exists $\varepsilon_0 > 0$ small enough such that this bound is strictly negative at all $\varepsilon \le \varepsilon_0$. Further, from (5.19) we get uniformity of the convergence in n, over all relevant $\underline{\delta}$ and $\varepsilon \le \varepsilon_0$, yielding (5.15) in case $|\mathcal{Q}| = 1$. While we do not detail these, the computations in case $|\mathcal{Q}| > 1$ and $\varepsilon > 0$ are similar. □

Acknowledgments. We thank Allan Sly for suggesting the weighted averages of (4.3), Andrea Montanari for a key idea in the proof of Lemma 5.1 and Yuval Peres for helpful discussions about Remark 4.6 and 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 her/his helpful suggestions on improving the presentation of the paper.

Part of the work was done when the authors participated in the MSRI program on Random Spatial Processes.

REFERENCES

[1] AIZENMAN, M. (1980). Translation invariance and instability of phase coexistence in the two-dimensional Ising system. *Comm. Math. Phys.* **73** 83–94. MR0573615

[2] AIZENMAN, M. and WEHR, J. (1990). Rounding effects of quenched randomness on first-order phase transitions. *Comm. Math. Phys.* **130** 489–528. MR1060388

[3] ALDOUS, D. and LYONS, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.* **12** 1454–1508. MR2354165

[4] ALDOUS, D. and STEELE, J. M. (2004). The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures. Encyclopaedia Math. Sci.* **110** 1–72. Springer, Berlin. MR2023650

[5] ATHREYA, K. B. and NEY, P. E. (2004). *Branching Process*. Dover Publications, Mineola, NY.

[6] BENJAMINI, I. and SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** 13 pp. (electronic). MR1873300

[7] BERGER, N., KENYON, C., MOSSEL, E. and PERES, Y. (2005). Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Related Fields* **131** 311–340. MR2123248

[8] BODINEAU, T. (2006). Translation invariant Gibbs states for the Ising model. *Probab. Theory Related Fields* **135** 153–168. MR2218869

[9] DEMBO, A. and MONTANARI, A. (2010). Ising models on locally tree-like graphs. *Ann. Appl. Probab.* **20** 565–592. MR2650042

[10] DEMBO, A. and MONTANARI, A. (2010). Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.* **24** 137–211. MR2643563

[11] DEMBO, A., MONTANARI, A. and SUN, N. (2013). Factor models on locally tree-like graphs. *Ann. Probab.* **41** 4162–4213. MR3161472

[12] DE SANCTIS, L. and GUERRA, F. (2008). Mean field dilute ferromagnet: High temperature and zero temperature behavior. *J. Stat. Phys.* **132** 759–785. MR2430780

[13] DOBRUSHIN, R. L. and SHLOSMAN, S. B. (1985). The problem of translation invariance of Gibbs states at low temperatures. In *Mathematical Physics Reviews*, *Vol*. 5. *Soviet Sci*. *Rev*. *Sect*. *C Math*. *Phys*. *Rev*. **5** 53–195. Harwood Academic Publ., Chur. MR0852217

[14] DOMMERS, S., GIARDINÀ, C. and VAN DER HOFSTAD, R. (2010). Ising models on power-law random graphs. *J*. *Stat*. *Phys*. **141** 638–660. MR2733399

[15] ELLIS, R. S. and NEWMAN, C. M. (1978). The statistics of Curie–Weiss models. *J*. *Stat*. *Phys*. **19** 149–161. MR0503332

[16] GEORGII, H.-O. (1988). *Gibbs Measures and Phase Transitions*. *De Gruyter Studies in Mathematics* **9**. de Gruyter, Berlin. MR0956646

[17] GEORGII, H.-O. and HIGUCHI, Y. (2000). Percolation and number of phases in the two-dimensional Ising model. *J*. *Math*. *Phys*. **41** 1153–1169. MR1757954

[18] GRESCHENFIELD, A. and MONTANARI, A. (2007). Reconstruction for models on random graphs. In 48*th FOCS Symposium*, Providence, RI.

[19] GRIFFITHS, R. B., HURST, C. A. and SHERMAN, S. (1970). Concavity of magnetization of an Ising ferromagnet in a positive external field. *J*. *Math*. *Phys*. **11** 790–795. MR0266507

[20] GRIMMETT, G. (2006). *The Random-Cluster Model*. *Grundlehren der Mathematischen Wissenschaften* **333**. Springer, Berlin. MR2243761

[21] KRENGEL, U. (1985). *Ergodic Theorems*. *De Gruyter Studies in Mathematics* **6**. de Gruyter, Berlin. MR0797411

[22] KÜLSKE, C. (1997). Metastates in disordered mean-field models: Random field and Hopfield models. *J*. *Stat*. *Phys*. **88** 1257–1293. MR1478069

[23] KURTZ, T., LYONS, R., PEMANTLE, R. and PERES, Y. (1997). A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes* (*Minneapolis*, *MN*, 1994). *IMA Vol*. *Math*. *Appl*. **84** 181–185. Springer, New York. MR1601737

[24] LIGGETT, T. M. (2005). *Interacting Particle Systems*. Springer, Berlin. MR2108619

[25] LYONS, R. (1989). The Ising model and percolation on trees and tree-like graphs. *Comm*. *Math*. *Phys*. **125** 337–353. MR1016874

[26] LYONS, R. (1990). Random walks and percolation on trees. *Ann*. *Probab*. **18** 931–958. MR1062053

[27] MÉZARD, M. and MONTANARI, A. (2009). *Information*, *Physics*, *and Computation*. Oxford Univ. Press, Oxford. MR2518205

[28] MONTANARI, A., MOSSEL, E. and SLY, A. (2012). The weak limit of Ising models on locally tree-like graphs. *Probab*. *Theory Related Fields* **152** 31–51. MR2875752

[29] NEWMAN, C. M. and STEIN, D. L. (1996). Spatial inhomogeneity and thermodynamic chaos. *Phys*. *Rev*. *Lett*. **76** 4821–4824.

[30] NEWMAN, M. E. J. (2003). The structure and function of complex networks. *SIAM Rev*. **45** 167–256 (electronic). MR2010377

[31] NISS, M. (2005). History of the Lenz–Ising model 1920–1950: From ferromagnetic to cooperative phenomena. *Arch*. *Hist*. *Exact Sci*. **59** 267–318. MR2124728

[32] PEMANTLE, R. and PERES, Y. (2010). The critical Ising model on trees, concave recursions and nonlinear capacity. *Ann*. *Probab*. **38** 184–206. MR2599197

[33] STROOCK, D. W. (2011). *Probability Theory*: *An Analytic View*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2760872

DEPARTMENT OF MATHEMATICS
DUKE UNIVERSITY
DURHAM, NORTH CAROLINA 27708-0320
USA
E-MAIL: anirbanb@math.duke.edu

DEPARTMENT OF MATHEMATICS
AND DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-2125
USA
E-MAIL: adembo@stanford.edu