The Annals of Probability
2013, Vol. 41, No. 6, 4162–4213
DOI: 10.1214/12-AOP828

FACTOR MODELS ON LOCALLY TREE-LIKE GRAPHS

BY AMIR DEMBO[1], ANDREA MONTANARI[1,2] AND NIKE SUN[1,3]

Stanford University

We consider homogeneous factor models on uniformly sparse graph sequences converging locally to a (unimodular) random tree T, and study the existence of the free energy density ϕ, the limit of the log-partition function divided by the number of vertices n as n tends to infinity. We provide a new interpolation scheme and use it to prove existence of, and to explicitly compute, the quantity ϕ subject to uniqueness of a relevant Gibbs measure for the factor model on T. By way of example we compute ϕ for the independent set (or hard-core) model at low fugacity, for the ferromagnetic Ising model at all parameter values, and for the ferromagnetic Potts model with both weak enough and strong enough interactions. Even beyond uniqueness regimes our interpolation provides useful explicit bounds on ϕ.

In the regimes in which we establish existence of the limit, we show that it coincides with the Bethe free energy functional evaluated at a suitable fixed point of the belief propagation (Bethe) recursions on T. In the special case that T has a Galton–Watson law, this formula coincides with the nonrigorous "Bethe prediction" obtained by statistical physicists using the "replica" or "cavity" methods. Thus our work is a rigorous generalization of these heuristic calculations to the broader class of sparse graph sequences converging locally to trees. We also provide a variational characterization for the Bethe prediction in this general setting, which is of independent interest.

1. Introduction. Let $G = (V, E)$ be a finite undirected graph, and $\mathscr{X}$ a finite alphabet of *spins*. A *factor model* on G is a probability measure on the space of *(spin) configurations* $\underline{\sigma} \in \mathscr{X}^V$ of form

$$\nu_{G,\underline{\psi}}^{\beta,B}(\underline{\sigma}) = \frac{1}{Z_{G,\underline{\psi}}(\beta, B)} \prod_{(ij)\in E} \psi^{\beta}(\sigma_i, \sigma_j) \prod_{i\in V} \bar{\psi}^{B}(\sigma_i), \tag{1.1}$$

where $\psi \equiv \psi^\beta$ is a symmetric function $\mathscr{X}^2 \to \mathbb{R}_{\geq 0}$ parametrized by $\beta \in \mathbb{R}$, $\bar{\psi} \equiv \bar{\psi}^B$ is a positive function $\mathscr{X} \to \mathbb{R}_{\geq 0}$ parametrized by $B \in \mathbb{R}$ and $Z_{G,\underline{\psi}}(\beta, B)$ is the normalizing constant, called the *partition function* (with its logarithm called

Received October 2011; revised November 2012.

[1]Supported in part by NSF Grant DMS-11-06627.
[2]Supported in part by NSF Grant CCF-0743978.
[3]Supported in part by Dept. of Defense NDSEG Fellowship.
MSC2010 subject classifications. 05C80, 60K35, 82B20, 82B23.
Key words and phrases. Factor models, random graphs, belief propagation, Bethe measures, Potts model, independent set, Gibbs measures, free energy density, local weak convergence.

the *free energy*). The pair $\underline{\psi} \equiv (\psi, \bar{\psi})$ is called a *specification* for the factor model (1.1).

In this paper we study the asymptotics of the free energy for sequences of (random) graphs $G_n = (V_n = [n], E_n)$ in the *thermodynamic limit* $n \to \infty$. More precisely, with $Z_n(\beta, B) \equiv Z_{G_n,\underline{\psi}}(\beta, B)$ and $\mathbb{E}_n$ denoting expectation with respect to the law of G_n, we seek to establish the existence of the *free energy density*

$$(1.2) \qquad \phi(\beta, B) \equiv \lim_{n\to\infty} \phi_n(\beta, B), \qquad \text{where } \phi_n(\beta, B) \equiv \frac{1}{n}\mathbb{E}_n\big[\log Z_n(\beta, B)\big],$$

and to determine its value. [In the literature, $\phi(\beta, B)$ is also referred to as the "free entropy density" or "pressure."]

The primary example we consider is the Potts model for a system of interacting spins on a graph. Formally, the *q-Potts model* on G with inverse temperature β and magnetic field B is the probability measure on $\mathcal{X}^V = [q]^V$ (with $[q] \equiv \{1, \dots, q\}$) given by

$$(1.3) \qquad \nu_G^{\beta,B}(\underline{\sigma}) = \frac{1}{Z_G(\beta, B)} \exp\left\{\beta \sum_{(ij)\in E} \mathbf{1}\{\sigma_i = \sigma_j\} + B \sum_{i\in V} \mathbf{1}\{\sigma_i = 1\}\right\}.$$

For $\beta > 0$ the system favors monochromatic edges and is said to be *ferromagnetic*, while for $\beta < 0$ the system favors edge disagreements and is said to be *anti-ferromagnetic*; the magnetic field B biases vertices toward the distinguished spin 1. The q-Potts model generalizes the Ising model which corresponds to the case $q = 2$. In analogy with the Potts model, in the general factor model setting we continue to refer to β as the interaction or temperature parameter and to B as the magnetic field.

Potts models have been intensively studied in statistical mechanics because of their key role in the theory of phase transitions [45], critical phenomena [48] and conformally invariant scaling limits [37]. As demonstrated, for instance, in [34] for the Ising model, determining the limit (1.2) plays a key role in characterizing the asymptotic structure of the measures $\nu_{G_n}^{\beta,B}$ in the thermodynamic limit. Potts models are also of great interest in combinatorics: recall in fact that the partition function admits a random-cluster representation ([16, 24]; see also Section 4.2), which at $B = 0$ reads

$$Z_G(\beta, 0) = \sum_{F\subseteq E} (e^\beta - 1)^{|F|} q^{k(F)},$$

with $k(F)$ denoting the number of connected components induced by the subset of edges $F \subseteq E$; cf. (4.2). Up to a multiplicative constant this coincides with the Tutte polynomial $T_G(x, y)$ of G evaluated at $x = 1 + q(e^\beta - 1)^{-1}$, $y = e^\beta$; see, for example, [42].

Mathematical statistical mechanics has focused so far on specific graph sequences G_n, for example, on finite exhaustions of the rectangular grid or other regular lattices in d dimensions with d fixed. Under mild conditions on the sequence, existence of the free energy density is a consequence of the following well-known argument (see, e.g., [38], Proposition 2.3.2): each graph G_n can be decomposed into smaller blocks by deleting a collection of edges whose number is negligible in comparison with the volume. Consequently the sequence $\log Z_{G_n}$ is approximately sub-additive in n, implying existence of the limit; see [26].

In this paper we consider sparse graphs with a locally tree-like structure—formally, graph sequences G_n converging locally weakly to (random) trees; see Definition 1.1 below; see also [1, 6]. Although the study of statistical mechanics "beyond $\mathbb{Z}^d$" is not directly motivated by physics considerations, physicists have been interested in models on alternative graph structures for a long time (an early example being [14]). Moreover, the study of factor models on *sparse* graphs has many motivations coming from computer science and statistical inference; see [9, 33]. Indeed, another example we will consider is the hard-core model for random independent sets on a graph. In this model the configuration space is $\mathscr{X}^V = \{0, 1\}^V$, where 0 means unoccupied, and 1 means occupied, and the only configurations receiving positive measure are those for which no two neighboring vertices are occupied, that is, so that the occupied vertices form an independent set in the graph. Formally, the *independent set* or *hard-core* model on G with fugacity $\lambda > 0$ is the probability measure on $\{0, 1\}^V$ given by

$$\nu_G^\lambda(\underline{\sigma}) = \frac{1}{Z_G(\lambda)} \prod_{(ij)\in E} \mathbf{1}\{\sigma_i\sigma_j \neq 1\} \prod_{i\in V} \lambda^{\sigma_i}, \tag{1.4}$$

so that as λ increases the measure becomes more biased toward the larger independent sets (and we write $B \equiv \log\lambda$ for the magnetic field). Due to the hard constraint preventing neighboring 1s, this system always has anti-ferromagnetic interactions and is of significant interest in computer science. The independent set decision problem is NP-complete (via the clique decision problem [8, 28]). As λ increases the measure ν_G^λ becomes increasingly concentrated on the maximal independent sets; the optimization problem of finding such sets is NP-hard [30] and hard to approximate ([49] and references therein). The problem of counting independent sets [i.e., computing $Z_G(1)$] for graphs of maximum degree Δ is #P-complete for $\Delta \geq 3$ ([22] and references therein). Although there exists a PTAS (polynomial-time approximation scheme) for $Z_G(\lambda)$ for λ below a certain "uniqueness threshold" [44], a series of previous works (see [20, 35, 40] and references therein) gave strong evidence that computation is hard for any λ above this threshold. This question was resolved simultaneously in the subsequent works [19, 41], with [41] building on methods from this paper.

Since infinite trees are nonamenable, G_n cannot be decomposed by removing a vanishing fraction of edges, so the preceding argument no longer applies: in physics terms, *surface effects are nonnegligible even in the thermodynamic limit.* Despite this, statistical physicists expect the free energy density (1.2) to exist on a large class of locally tree-like graphs. Even more surprisingly, employing nonrigorous but mathematically sophisticated heuristics such as the "replica" or "cavity" methods, they derive exact formulas for this limit for a number of statistical mechanics models on locally tree-like graphs; see, for example, [33] and the references therein. The primary example considered in these works is the graph chosen uniformly at random from those with n vertices and $m = m(n)$ edges, with $m/n \to \gamma \in \mathbb{R}$; such graphs converge locally to the Galton–Watson tree with $\mathrm{Pois}(2\gamma)$ offspring distribution. The Galton–Watson tree with general offspring distribution can be obtained as the local weak limit of random graphs with specified degree profile corresponding to the offspring distribution; the physics heuristics extend to this and even more general settings.

There is no good argument for why the limit (1.2) exists; the heuristic replica or cavity methods compute this limit starting from the postulate that it exists. A significant breakthrough was achieved by the interpolation method first developed by Guerra and Toninelli [25] for the Sherrington–Kirkpatrick model from spin-glass theory, and then generalized to a number of statistical physics models on sparse graphs [17, 18, 36] and related constraint satisfaction problems [5]. This method establishes super-additivity of $\log Z_{G_n}$ which implies existence of the limit (1.2). Unfortunately, this approach appears limited to models with *repulsive* interactions, that is, in which higher weight is given to configurations in which neighboring vertices take different values. In particular, it does not apply to the ferromagnetic Potts model. This is especially puzzling because the heuristic physics predictions do not distinguish between the two cases, and there is no fundamental reason why the limit should be computable in one case and not in the other. Further, this interpolation method only applies to very restricted classes of graph sequences (typically, uniformly random given the degree sequence); notably, existence of the limit is not proved for *deterministic* graph sequences. Finally, the method gives no way to actually compute the limit, although interpolation has been used to prove upper bounds [17, 18, 36].

In this paper we follow a different approach relying only on local weak convergence of the graph sequence $(G_n)_{n\geq 1}$ to some limiting (random) tree. The general idea is that the corresponding factor models (1.1) must converge (passing to a subsequence as needed), to a Gibbs measure on the limiting tree; the task then "reduces" to the one of identifying the correct limit. This is still a substantial challenge because, in general, there is an uncountable number of "candidate" Gibbs measures for the limit. Nevertheless, this program was carried through in [10] for Ising models on graphs converging locally to a Galton–Watson tree, under a "uniform sparsity" assumption (Definition 1.3), on the degree distribution. (It is further assumed in [10] that the distribution has finite second moment; this condition was relaxed in [13], thereby handling the case of power law graphs.) The result of [10, 13] provides also a fairly explicit expression $\Phi(\beta, B)$ for the free energy density, defined solely in terms of the limiting tree. This expression coincides with the so-called "Bethe prediction" of statistical physics, derived earlier for random graphs with given degree distribution using the "replica" or "cavity" methods.

We develop this approach here in more generality. Rather than considering a specific model such as the Ising, we establish results for general abstract factor models satisfying mild regularity conditions [see (H1) below], covering in particular the Potts and independent set models. We also make no distributional assumptions on the graphs G_n or the limiting random tree, other than some integrability conditions [see Definition 1.3 and (H2) below]. In this setting we develop a general interpolation scheme (Theorem 1.15) which, under appropriate assumptions, bounds differences $\phi_n(\beta, B) - \phi_n(\beta_0, B_0)$ in the limit $n \to \infty$ by differences $\Phi(\beta, B) - \Phi(\beta_0, B_0)$ for Φ a functional defined solely in terms of the limiting tree; see (1.12). We refer the reader to [2] for a discussion of the computation of limits of finite large random structures through optimization procedures on the limiting infinite structure. Although we continue to refer to this $\Phi(\beta, B)$ as the "Bethe prediction," we remark that it is a considerable generalization of earlier formulas obtained in the special case of Galton–Watson trees by statistical physics methods. It is defined as the evaluation of the "Bethe free energy functional" (1.9) at a specific Gibbs measure on the limiting tree, and corresponds to what physicists call the "replica symmetric solution": whereas it is expected to hold in the high-temperature regime (i.e., with small enough interactions), for many factor models it is incorrect at low temperature. However, we will show that in "uniqueness regimes," where the set of Gibbs measures on the limiting tree corresponding to the factor model specification $\underline{\psi}$ is a singleton, the upper and lower bounds of Theorem 1.15 match to completely verify the Bethe prediction (Theorem 1.16).

We then apply our interpolation scheme to compute the free energy density in specific models. We verify the Bethe prediction for the independent set model with low fugacity (Theorem 1.12) as a consequence of Theorem 1.16. Further, by using monotonicity properties to restrict the set of relevant Gibbs measures, we obtain results for the Potts model going beyond the implications of Theorem 1.16: for $q = 2$ (Ising), we verify the Bethe prediction for all $\beta \geq 0$, $B \in \mathbb{R}$ (Theorem 1.9), extending the results of [10, 13] to general locally tree-like graph sequences. For general q, we verify the prediction in regimes of nonnegative (β, B) in which two specific Gibbs measures on the limiting tree coincide, namely, the Gibbs measures arising from free and 1 boundary conditions coincide, see Definition 1.8 below. This condition is satisfied throughout the range $\{\beta \geq 0, B > 0\}$ for $q = 2$; when $q \geq 3$ there are regimes of nonuniqueness in which it fails, but we will show that it is satisfied both at β sufficiently small and sufficiently large, that is, at high and low temperatures.

Theorem 1.15 can give useful bounds even beyond uniqueness regimes. As an illustration, we study the Potts model in the case that G_n converges locally to the d-regular tree T_d. In Theorem 1.11 we explicitly characterize the nonuniqueness regime of this model and use Theorem 1.15 to give bounds for $\phi_n(\beta, B)$ within this regime. In a subsequent work [11] we prove that in this setting, $\phi(\beta, B)$ exists and matches the lower bound of Theorem 1.11. We also compute there the asymptotic free energy $\phi(\lambda)$ (all $\lambda \geq 0$) for the independent set model on d-regular bipartite graphs. In contrast, for generic nonbipartite G_n the consensus in physics is for a full replica symmetry breaking for large enough λ, and consequently there does not exist even a heuristic prediction for the free energy density in this regime.

As mentioned above, the Bethe prediction $\Phi(\beta, B)$ is the evaluation of the Bethe free energy functional at a specific Gibbs measure on the limiting tree. This Gibbs measure has a characterization in terms of "messages" $h_{x\to y} \equiv h_{(T,x\to y)}$ defined on the directed edges $x \to y$ of each tree T, such that the entire collection of messages is a fixed point of a certain "belief propagation" or "Bethe recursion" (1.10). Motivated by the finite-graph optimization of [46], we provide a variational characterization of the Bethe prediction (Theorem 1.18) which is of independent interest. In particular, this formulation suggests nontrivial connections with large deviation principles.

1.1. *Local weak convergence and the Bethe prediction.* We study factor models on graphs which are "locally tree-like" in a sense which we now formalize, starting with a few notation and conventions. All graphs are taken to be *undirected and locally finite*. In a graph $G = (V, E)$, let d denote graph distance, and for $v \in V$ write $B_t(v)$ for the sub-graph of G induced by $\{w \in V : d(v, w) \leq t\}$. Write $v \sim w$ if v, w are neighbors in G, and write ∂v for the set of neighbors of v and $D_v \equiv |\partial v|$. Let $\mathcal{G}_\bullet$ denote the space of isomorphism classes of (finite or infinite) rooted, connected graphs (G, o). A metric on this space is given by defining the distance between (G_1, o_1) and (G_2, o_2) in $\mathcal{G}_\bullet$ to be $1/(1 + R)$ where R is the maximal $r \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$ such that $B_R(o_1) \cong B_R(o_2)$; with this definition $\mathcal{G}_\bullet$ is a complete separable metric space; see, for example, [1]. Let $\mathcal{T}_\bullet \subset \mathcal{G}_\bullet$ denote the closed subspace of (*rooted*) *trees* $T \equiv (T, o)$, the acyclic elements of $\mathcal{G}_\bullet$. We write T^t for $B_t(o)$ in T, and in particular we use T^0 to denote the single-vertex tree. We now define the precise notion of graph limits considered throughout this paper.

DEFINITION 1.1. Let $G_n = (V_n, E_n)$ $(n \geq 1)$ be a sequence of random graphs, and let I_n be a vertex chosen uniformly at random from V_n. We say G_n *converges locally* (*weakly*) to the random tree T if for each $t \geq 0$, $B_t(I_n)$ converges in law to T^t in the space $\mathcal{G}_\bullet$. We say in this case that the G_n are *locally tree-like*.

We will make repeated use of the fact that *any* local weak limit of graph sequences satisfies the "unimodularity" or "mass-transport" property whose definition we recall here; for a detailed account, see [1]. Let $\mathcal{G}_{\bullet\bullet}$ denote the space of isomorphism classes of *bi-rooted*, connected graphs with a distinguished ordered pair, denoted (G, i, j) (we do not require $i \sim j$); $\mathcal{G}_{\bullet\bullet}$ is metrizable in a similar manner as $\mathcal{G}_\bullet$.

DEFINITION 1.2. A Borel probability measure μ on $\mathcal{G}_\bullet$ is said to be *unimodular* if it obeys the *mass-transport principle*,

$$\mathbb{E}_\mu\Big[\sum_{x\in V(G)} f(G,o,x)\Big] = \mathbb{E}_\mu\Big[\sum_{x\in V(G)} f(G,x,o)\Big] \tag{1.5}$$
$$\forall f:\mathcal{G}_{\bullet\bullet}\to\mathbb{R}_{\ge 0} \text{ Borel.}$$

We say that μ is *involution invariant* if (1.5) holds when restricted to f supported only on those (G,x,y) with $x\sim y$.

A measure μ on $\mathcal{G}$ is involution invariant if and only if it is unimodular ([1], Proposition 2.2). Unimodularity corresponds to "indistinguishability of the root;" the concept first appeared in [6] where it was observed that local weak limits of graph sequences must be unimodular ([6], Section 3.2). The converse of this implication remains a well-known open question; see [1].

DEFINITION 1.3. The graph sequence G_n is *uniformly sparse* if the D_{I_n} are uniformly integrable, that is, if

$$\lim_{L\to\infty}\Big(\limsup_{n\to\infty}\mathbb{E}_n\big[D_{I_n}\mathbf{1}\{D_{I_n}\ge L\}\big]\Big)=0$$

(where $\mathbb{E}_n$ denotes expectation over the law of G_n and I_n).

We assume throughout that G_n $(n\ge 1)$ is a uniformly sparse graph sequence converging locally weakly to the random tree T of (unimodular) law μ such that the root degree D_o is nonzero with positive μ-probability; this entire setting is hereafter denoted $G_n\to_{lwc}\mu$. In this setting we will describe general conditions under which the asymptotic free energy $\phi(\beta,B)$ for the factor model (1.1) exists and agrees with the "Bethe energy prediction," which we now describe. [If the sequence of random graphs G_n is such that $G_n\to_{lwc}\mu$ for almost every realization of the sequence—as is the case for Erdös–Rényi random graphs or random graphs with given degree distribution (see, e.g., [9], Propositions 2.5 and 2.6)—then our results apply instead to the a.s. limit of $n^{-1}\log Z_n(\beta,B)$.]

Let $\Delta_{\mathscr{X}}$ denote the $(|\mathscr{X}|-1)$-dimensional simplex of probability measures on the finite alphabet of spins $\mathscr{X}$. Let $\mathcal{T}_\bullet^+$ denote $\mathcal{T}_\bullet$ without the single-vertex tree T^0, and let $\mathcal{T}_{\mathrm{e}}\subset\mathcal{G}_{\bullet\bullet}$ be the space of isomorphism classes of trees $T\in\mathcal{T}_\bullet^+$ rooted at a directed edge $x\to y$, written $(T,x\to y)$ or simply $x\to y$ for short. If T has law μ for μ a unimodular measure on $\mathcal{T}_\bullet$, we let $\mu^\uparrow$ and $\mu^\downarrow$ denote the laws of $(T,J\to o)$ and $(T,o\to J)$, respectively, for J chosen uniformly at random from ∂o conditioned on the event $\{T\in\mathcal{T}_\bullet^+\}$. Involution invariance of μ is then equivalent to

$$\begin{aligned}\mathbb{E}_{\mu^\downarrow}\big[D_x f(T,x\to y)\big] &= \frac{\mathbb{E}_\mu[\sum_{j\in\partial o} f(T,o\to j)]}{\mu(D_o>0)} = \frac{\mathbb{E}_\mu[\sum_{j\in\partial o} f(T,j\to o)]}{\mu(D_o>0)}\\ &= \mathbb{E}_{\mu^\uparrow}\big[D_y f(T,x\to y)\big]\end{aligned}$$

(where o corresponds to x on the left-hand side and to y on the right-hand side), so in particular $\mu^{\uparrow}$ and $\mu^{\downarrow}$ are mutually absolutely continuous.

DEFINITION 1.4. The *message space* is the space $\mathcal{H} \equiv \mathcal{H}_\mu$ of measurable functions

$$h : \mathcal{T}_{\mathrm{e}} \times \mathbb{R}^2 \to \Delta_{\mathscr{X}}, \qquad \big((T, x \to y), \beta, B\big) \mapsto \big(h^{\beta,B}_{x \to y}(\sigma)\big)_{\sigma \in \mathscr{X}},$$

taken up to $\mu^{\uparrow}$-equivalence.

REMARK 1.5. For $(T, x \to y) \in \mathcal{T}_{\mathrm{e}}$ let $T_{x \to y}$ denote the component sub-tree rooted at x which results from deleting edge (x, y) from T. The interpretation of $h_{x \to y}$ is that it is a message from x to y on the tree T, giving the distribution of σ_x for the factor model (1.1) on $T_{x \to y}$. Indeed, although we do not require it in general, in our concrete examples $h_{x \to y}$ depends only on this component sub-tree.

For $T \in \mathcal{T}_\bullet$ and $h \in \mathcal{H}$, let

$$\begin{aligned} \Phi_T(\beta, B, h) &\equiv \Phi^{\mathrm{vx}}_T(\beta, B, h) - \Phi^{\mathrm{e}}_T(\beta, B, h) \\ &\equiv \Phi^{\mathrm{vx}}_T(\beta, B, h) - \frac{1}{2} \sum_{j \in \partial o} \Phi^{(oj)}_T(\beta, B, h), \end{aligned} \tag{1.6}$$

where “vx” and “e” indicate vertex and edge terms, respectively:

$$\Phi^{\mathrm{vx}}_T(\beta, B, h) \equiv \log\Big\{\sum_{\sigma} \bar{\psi}(\sigma) \prod_{j \in \partial o} \Big(\sum_{\sigma_j} \psi(\sigma, \sigma_j) h_{j \to o}(\sigma_j)\Big)\Big\}, \tag{1.7}$$

the log-partition function of the star graph T^1 with boundary conditions h [see Figure 1(a)] and

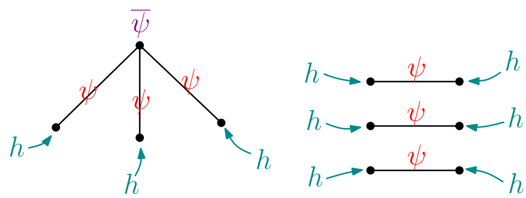
$$\begin{aligned} \Phi^{\mathrm{e}}_T(\beta, B, h) &\equiv \frac{1}{2} \sum_{j \in \partial o} \Phi^{(oj)}_T(\beta, B, h) \\ &= \frac{1}{2} \sum_{j \in \partial o} \log\Big\{\sum_{\sigma, \sigma_j} \psi(\sigma, \sigma_j) h_{j \to o}(\sigma_j) h_{o \to j}(\sigma)\Big\}, \end{aligned} \tag{1.8}$$

half the log-partition function on D_o disjoint edges with boundary conditions h; see Figure 1(b). (See Definition 1.8 below for a detailed discussion of boundary conditions.)

We take the usual convention that the empty sum is zero, and the empty product is one, so $\Phi_T = \log(\sum_\sigma \bar{\psi}(\sigma))$ in case $T = T^0$. Although we suppress it from the notation, in the above equations ψ and h are taken to be evaluated at (β, B). The *Bethe free energy functional* on $\mathcal{H}$ for the factor model (1.1) on $G_n \to_{lwc} \mu$ is defined by

$$\Phi_\mu(\beta, B, h) \equiv \mathbb{E}_\mu\big[\Phi_T(\beta, B, h)\big], \tag{1.9}$$

provided the expectation exists; see Lemma 2.2.



(a) Star graph T^1 (b) Edge graph

FIG. 1. *Φ_T^{vx} and $2\Phi_T^{\mathrm{e}}$ are log-partition functions of star and edge graphs.*

DEFINITION 1.6. The *belief propagation* or *Bethe recursion* is the mapping $\mathsf{BP} \equiv \mathsf{BP}^{\beta,B} : \mathcal{H} \to \mathcal{H}$,

$$(\mathsf{BP}^{\beta,B} h)_{x\to y}(\sigma) \equiv \frac{1}{z_{x\to y}(\beta,B)} \bar{\psi}^B(\sigma) \prod_{v\in\partial x\setminus y} \Big(\sum_{\sigma_v} \psi^\beta(\sigma,\sigma_v) h_{v\to x}(\sigma_v) \Big), \tag{1.10}$$

with $z_{x\to y}(\beta,B)$ normalizing constants. For μ a measure on $\mathcal{T}_\bullet$ and fixed (β,B), let $\mathcal{H}_\mu^\star(\beta,B)$ denote the space of measurable functions $h : \mathcal{T}_{\mathrm{e}} \to \Delta_{\mathscr{X}}$, again taken up to $\mu^\uparrow$-equivalence, which are fixed points of the Bethe recursion: that is, satisfying

$$h = \mathsf{BP}^{\beta,B} h, \qquad \mu^\uparrow\text{-a.s.} \tag{1.11}$$

The *Bethe prediction* is that the asymptotic free energy $\phi(\beta,B)$ of (1.2) exists and equals

$$\Phi_\mu^{\mathrm{Bethe}}(\beta,B) \equiv \Phi_\mu(\beta,B,h^\star) \tag{1.12}$$

for $h^\star$ a certain element of $\mathcal{H}_\mu^\star(\beta,B)$. We often drop the subscript μ when it is clear from context.

REMARK 1.7. In the case that the recursion (1.11) has multiple solutions ($|\mathcal{H}_\mu^\star(\beta,B)| > 1$), the *Bethe prediction* is defined to be the supremum of $\Phi(\beta,B,h^\star)$ over admissible fixed points $h^\star$. While in the abstract factor model setting all fixed points are admissible, in specific models typically there are "natural" criteria restricting the set of admissible fixed points. We will demonstrate this in the Ising and Potts models where restrictions are imposed by monotonicity and symmetry considerations.

The rationale behind the Bethe recursions and Bethe prediction is explained in detail in [9], Section 3; see also [33]. In brief, solutions to the Bethe recursions correspond to consistent "boundary laws" for the factor model on tree-like graphs; for further details, see Remark 1.13 below. When G is a finite tree, and μ_G is the law of (G,I) for I a uniform element of V (here μ_G is a measure on $\mathcal{T}_\bullet$, but not necessarily unimodular), the Bethe recursions have a unique solution, given by the so-called "standard message set;" see [9], Remark 3.5. In this setting it holds exactly (see [9], Proposition 3.7) that

$$|V|^{-1}\log Z_G = \Phi_{\mu_G} = |V|^{-1}\sum_{v\in G}\Phi_{(G,v)},$$

where $\Phi_{(G,v)}$ is as defined by (1.6) with $T = (G, v)$. The heuristic then is that for G_n locally like the random tree $T \sim \mu$, the (normalized) free energy ϕ_n is approximated by $\Phi_\mu \equiv \mathbb{E}_\mu[\Phi_T]$ for n large. We emphasize that *no* averaging over the vertices of the tree T takes place in the definition of Φ_T; indeed for $T \in \mathcal{T}_\bullet$ the sub-trees T^t typically do *not* converge locally weakly to T. For example, when T is the d-regular tree T_d, the subtrees T^t converge locally weakly to the so-called d-canopy tree; see, for example, [9], Lemma 2.8. Instead the averaging of $\Phi_{(G,v)}$ over the vertices $v \in G$ in the evaluation of Φ_{μ_G} corresponds to the averaging with respect to the law μ in the evaluation of the Bethe prediction Φ_μ.

The following is a terminology which we adopt throughout the paper:

DEFINITION 1.8. If G is any graph and U a sub-graph, the *external boundary* ∂U of U is the set of vertices of $G \setminus U$ adjacent to U. Let U^+ denote the sub-graph of G induced by the vertices in $V_U \cup \partial U$. For U finite (so U^+ is finite, since G is locally finite), and $\nu^\ddagger$ a measure on $\mathscr{X}^{\partial U}$, the factor model on U with $\nu^\ddagger$ *boundary conditions* is the probability measure on configurations $\underline{\sigma}_U \in \mathscr{X}^{V_U}$ given by

$$\nu^\ddagger_{U,G,\underline{\psi}}(\underline{\sigma}_U) \cong \int \prod_{(ij)\in E_{U^+}} \psi(\sigma_i,\sigma_j)\prod_{i\in U}\bar{\psi}(\sigma_i)\,d\nu^\ddagger(\underline{\sigma}_{\partial U}). \tag{1.13}$$

(Throughout, $\cong$ indicates equivalence up to a positive normalizing constant.) The case in which $\nu^\ddagger$ gives probability one to the identically-σ_0 spin configuration on ∂U ($\sigma_0 \in \mathscr{X}$) is referred to as σ_0 *boundary conditions* and denoted $\nu^\ddagger = \nu^{\sigma_0}$, while the case in which $\nu^\ddagger$ is uniform measure on $\mathscr{X}^{\partial U}$ is referred to as *free boundary conditions* and denoted $\nu^\ddagger = \nu^{\mathrm{f}}$.

1.2. *Application to Ising, Potts and independent set.* Before formally stating our main theorem for general factor models, we mention its consequences in some models of interest: we verify the Bethe prediction for the ferromagnetic Ising model at all temperatures, the ferromagnetic Potts model with field $B \geq 0$ in uniqueness regimes, and the independent set model with low fugacity λ.

1.2.1. *Ising model.* The *Ising model* is the Potts model (1.3) with $q = 2$. For convenience we use the equivalent formulation which takes $\mathscr{X} = \{\pm 1\}$ and defines the probability measure on $\mathscr{X}^V$

$$\nu_G^{\beta,B}(\underline{\sigma}) = \frac{1}{Z_G(\beta,B)}\exp\left\{\beta\sum_{(ij)\in E}\sigma_i\sigma_j + B\sum_{i\in V}\sigma_i\right\}. \tag{1.14}$$

For $T \in \mathcal{T}_\bullet$ let $\bar{h}_T^{t,+} \equiv \bar{h}_T^{t,+,\beta,B}$ denote the root marginal for the Ising model of parameters (β, B) on T^t with $+$ boundary conditions (i.e., with σ_v conditioned to be $+1$ for all v at level $t+1$), and similarly define $\bar{h}_T^{\mathrm{f}}$ corresponding to free boundary conditions. For $\ddagger \in \{\mathrm{f}, +\}$ let $\bar{h}_T^{\ddagger} \equiv \bar{h}_T^{\ddagger,\beta,B} \equiv \lim_{t\to\infty} \bar{h}_T^{t,\ddagger,\beta,B}$. (Existence of the limits $\bar{h}_T^{\mathrm{f}}, \bar{h}_T^{+}$ for the Ising model is an easy consequence of Griffiths's inequality; see Lemma 4.1.) We then define messages $h^{\ddagger} \in \mathcal{H}_\mu$ by

$$h^{\ddagger}_{x\to y} = \bar{h}^{\ddagger}_{T_{x\to y}}$$

for $T_{x\to y}$ as defined in Remark 1.5. For $G_n \to_{lwc} \mu$, the Bethe free energy prediction for the Ising model with $\beta \ge 0$, $B > 0$ is $\phi(\beta, B) = \Phi_\mu(\beta, B, h^+)$. This prediction was verified in [10], Theorem 2.4, for uniformly sparse graph sequences converging locally weakly to Galton–Watson trees subject to the second-moment condition $\mathbb{E}_\mu[D_o^2] < \infty$, which was relaxed in [13] to a $(1+\varepsilon)$-moment condition. We have the following generalization of this result to an arbitrary limiting law.

THEOREM 1.9. *For the Ising model* (1.14) *on* $G_n \to_{lwc} \mu$,

$$\phi(\beta, B) = \Phi_\mu(\beta, B, h^{\mathrm{f}}) = \Phi_\mu(\beta, B, h^+)$$

for $\beta \ge 0$, $B > 0$. *Also* $\phi(\beta, B) = \phi(\beta, -B)$ *and* $\phi(\beta, 0) = \lim_{B\to 0} \phi(\beta, B)$.

Note that in the Ising model we are able to characterize the free energy density for *all* $\beta \ge 0$. The underlying reason is that for $B > 0$, all boundary conditions dominating the free boundary condition give rise to the same Gibbs measure on the limiting tree, that is, $\bar{h}^{\mathrm{f}} = \bar{h}^+$. This phenomenon appears to be in line with physicists' intuition that the Ising model always undergoes a second-order phase transition. The physics argument suggests therefore that the zero-magnetization phase becomes unstable below the critical temperature. In other words, even with free boundary conditions, an arbitrarily small external field $B > 0$ is sufficient to drive the system into the "plus" phase.

1.2.2. *Potts model.* Throughout the remainder let $(\beta_0, B_0) \le (\beta_1, B_1)$, where $\le$ means coordinate-wise less than or equal to. An *interpolation path* is a piecewise linear path, with each piece parallel to a coordinate axis, increasing from (β_0, B_0) to (β_1, B_1) with respect to the partial order $\le$.

We restrict our attention to the Potts model with $\beta, B \ge 0$. In this regime we are able to use a random-cluster representation to extract important monotonicity properties. For $T \in \mathcal{T}_\bullet$ and $\ddagger \in \{\mathrm{f}, 1\}$ let $\bar{h}_T^{t,\ddagger} \equiv \bar{h}_T^{t,\ddagger,\beta,B}$ denote the root marginal for the Potts model on T^t with $\ddagger$ boundary conditions. Let $\bar{h}_T^{\ddagger} \equiv \bar{h}_T^{\ddagger,\beta,B} \equiv \lim_{t\to\infty} \bar{h}_T^{t,\ddagger,\beta,B}$ (existence of the limits $\bar{h}_T^{\mathrm{f}}, \bar{h}_T^{1}$ for the Potts model follows from monotonicity properties of the random-cluster representation; see Corollary 4.4). We then define messages $h^{\ddagger} \in \mathcal{H}_\mu$ by $h^{\ddagger}_{x\to y} = \bar{h}^{\ddagger}_{T_{x\to y}}$, and let

$$\mathcal{R}_\mu \equiv \{(\beta, B) : h^{\mathrm{f}} = h^1, \mu^{\uparrow}\text{-a.s.}\}.$$

We also define

$$\mathcal{R}_\infty \equiv (\{0\} \times \mathbb{R}_{\geq 0}) \cup (\mathbb{R}_{\geq 0} \times \{\infty\}) \cup (\{\infty\} \times \mathbb{R}_{>0}).$$

THEOREM 1.10. *For the Potts model* (1.3) *with* $q > 2$ *and* $\beta, B \geq 0$ *on* $G_n \to_{lwc} \mu$, *the following hold* (*with* $\Phi \equiv \Phi_\mu$, $\mathcal{R} \equiv \mathcal{R}_\mu$):

(a) *If there exists an interpolation path contained in* $\mathcal{R}$ *joining* (β, B) *and* $\mathcal{R}_\infty$, *then*

$$\phi(\beta, B) = \Phi(\beta, B, h^{\mathrm{f}}) = \Phi(\beta, B, h^1).$$

(b) *If there exists an interpolation path from* (β_0, B_0) *to* (β_1, B_1) *along which* h^{f} *is continuous* (*in the interpolation parameter*), *then*

$$\liminf_{n\to\infty}[\phi_n(\beta_1, B_1) - \phi_n(\beta_0, B_0)] \geq \Phi(\beta_1, B_1, h^{\mathrm{f}}) - \Phi(\beta_0, B_0, h^{\mathrm{f}}).$$

If h^{f} *is replaced with* h^1, *then we have instead*

$$\limsup_{n\to\infty}[\phi_n(\beta_1, B_1) - \phi_n(\beta_0, B_0)] \leq \Phi(\beta_1, B_1, h^1) - \Phi(\beta_0, B_0, h^1).$$

We obtain more explicit results when the limiting tree is the d-regular tree T_d.

THEOREM 1.11. *For the Potts model* (1.3) *with* $q > 2$ *and* $\beta, B \geq 0$ *on* $G_n \to_{lwc} \mathsf{T}_d$, *the following hold* (*with* $\Phi \equiv \Phi_{\mathsf{T}_d}$, $\mathcal{R} \equiv \mathcal{R}_{\mathsf{T}_d}$, *and* $\mathcal{R}_{\neq} \equiv \{\beta, B \geq 0\} \setminus \mathcal{R}$):

(a) *If* $d = 2$, $\mathcal{R}_{\neq} = \varnothing$. *If* $d > 2$ *and* $q = 2$, *there exists* $0 < \beta_- < \infty$ *such that* $\mathcal{R}_{\neq} = \{B = 0, \beta > \beta_-\}$. *If* $d > 2$ *and* $q > 2$, *there exists* $0 < B_+ < \infty$ *and smooth curves* $\beta_{\mathrm{f}}(B) \leq \beta_+(B)$ *defined on* $[0, B_+]$ *with* $\beta_{\mathrm{f}}(B_+) = \beta_+(B_+)$ *such that*

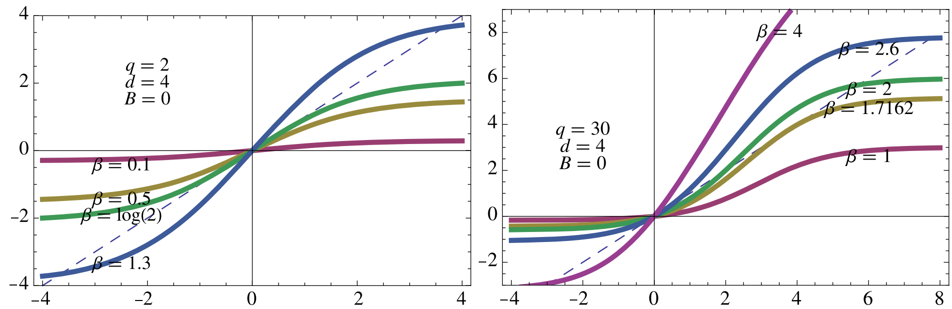
$$\mathcal{R}_{\neq} = \{B = 0, \beta \geq \beta_{\mathrm{f}}(0)\} \cup \{0 < B < B_+, \beta \in [\beta_{\mathrm{f}}(B), \beta_+(B)]\}.$$

(b) *For* $(\beta, B) \notin \mathcal{R}_{\neq}$, $\phi(\beta, B) = \Phi(\beta, B, h^{\mathrm{f}}) = \Phi(\beta, B, h^1)$. *If* $(\beta, B) \in \partial\mathcal{R}_{\neq}$ *with* $\beta = \beta_{\mathrm{f}}(B)$, *then* $\phi(\beta, B) = \Phi(\beta, B, h^{\mathrm{f}})$. *If* $(\beta, B) \in \partial\mathcal{R}_{\neq}$ *with* $\beta \geq \beta_+(B)$, *then* $\phi(\beta, B) = \Phi(\beta, B, h^1)$. *For* (β, B) *in the interior* $\mathcal{R}^\circ_{\neq}$ *of* $\mathcal{R}_{\neq}$,

$$\begin{aligned}\max\{\Phi(\beta, B, h^1), \Phi(\beta, B, h^{\mathrm{f}})\} &\leq \liminf_{n\to\infty} \phi_n(\beta, B) \\ &\leq \limsup_{n\to\infty} \phi_n(\beta, B) \leq \min\{\widetilde{\Phi}^{\mathrm{f}}(\beta, B), \widetilde{\Phi}^1(\beta, B)\},\end{aligned}$$

where

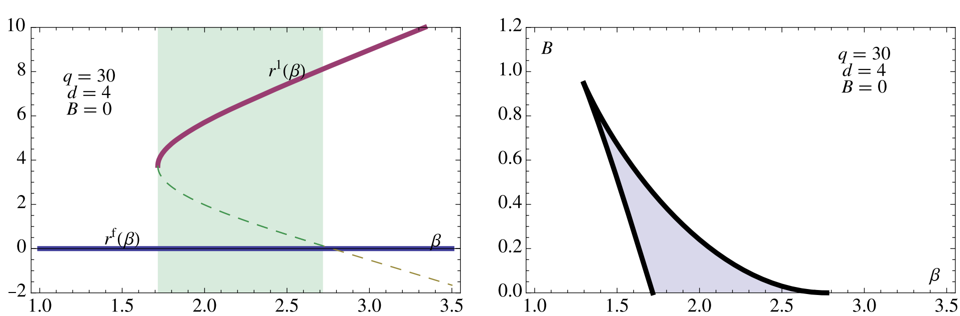
$$\begin{aligned}\widetilde{\Phi}^1(\beta, B) &\equiv \Phi(\beta_{\mathrm{f}}(B), B, h^{\mathrm{f}}) + [\Phi(\beta, B, h^1) - \Phi(\beta_{\mathrm{f}}(B), B, h^1)], \\ \widetilde{\Phi}^{\mathrm{f}}(\beta, B) &\equiv \Phi(\beta_+(B), B, h^1) - [\Phi(\beta_+(B), B, h^{\mathrm{f}}) - \Phi(\beta, B, h^{\mathrm{f}})].\end{aligned}$$



(a) Ising (b) Potts

FIG. 2. *Ising and Potts Bethe recursions.*

Figures 2–4 highlight the difficulty in analyzing the Potts model ($q > 2$) as opposed to the Ising model. Figure 2(a) shows the Ising Bethe recursion parametrized in terms of the log-likelihood ratio $r \equiv \log h(+) - \log h(-)$. For sufficiently large β the recursion has three fixed points, but in this case the $r = 0$ fixed point is unstable, and we will see in the proof of Theorem 1.9 that adding a small magnetic field resolves the nonuniqueness. The remaining plots were computed for the Potts model with $q = 30$ and $d = 4$. Figure 2(b) shows the Potts Bethe recursion at $B = 0$ restricted to those h which are symmetric among the spins $\neq 1$, and parametrized by $r \equiv \log h(1) - \log h(2)$. The fixed point at $r = 0$ corresponds to h^{f} while the uppermost fixed point corresponds to h^1; Figure 3(a) shows how the fixed points vary with β. In an intermediate regime of β-values [shaded in Figure 3(a)] both fixed points are stable, and perturbing by a magnetic field does not resolve the nonuniqueness: indeed, Figure 3(b) shows that there is a two-dimensional region $\mathcal{R}_{\neq}$ of (β, B) values for which $h^{\mathrm{f}} \neq h^1$, making the exact Bethe prediction inaccessible via our current interpolation scheme. Figure 4 shows the discrepancy between the upper and lower bounds of Theorem 1.11(b) inside $\mathcal{R}_{\neq}$.



(a) BP fixed points (b) Regime $\mathcal{R}_{\neq}$ (shaded)

FIG. 3. *Potts Bethe fixed points and the intermediate regime* $\mathcal{R}_{\neq}$.

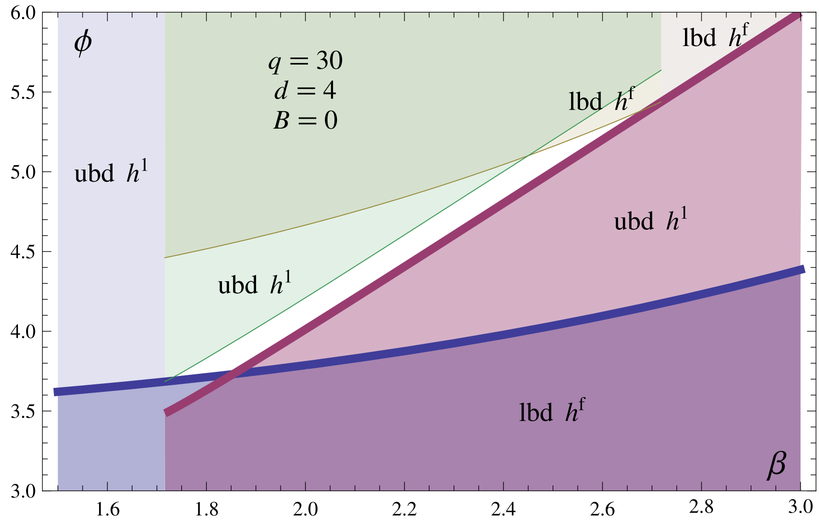


FIG. 4. *Potts Bethe interpolation: the heavy (light) shaded regions are the asymptotic lower (upper) bounds on ϕ_n given by Theorem 1.11; the bounds fail to match when $(\beta, B) \in \mathcal{R}_{\neq}$. The Bethe prediction is the upper envelope of the thick lines. In the figure, a shaded region marked "lbd $h^\dagger$" (resp., "ubd $h^\dagger$") means an asymptotic bound on ϕ_n obtained from interpolation using the asymptotic lower (resp., upper) bound on $a_n^{\mathrm{e}}(\beta, B)$ by $a^{\mathrm{e}}(\beta, B, h^\dagger)$, in the notation of Theorem 1.15. For example, the shaded region labeled "ubd h^1" is an asymptotic (lower) bound on ϕ_n obtained by interpolating from $\beta = \infty$ using the asymptotic* upper bound $\limsup a_n^{\mathrm{e}}(\beta, B) \le a^{\mathrm{e}}(\beta, B, h)^1$.

1.2.3. *Independent set model.* We consider the independent set model (1.4) in the regime of low fugacity. For $\ddagger \in \{0, 1\}$ let $\bar{h}_T^{t,\ddagger} \equiv \bar{h}_T^{t,\ddagger,\lambda}$ denote the root marginal on T^t with $\ddagger$ boundary conditions on ∂T^t: that is, $\bar{h}_T^{t,1}$ (resp., $\bar{h}_T^{t,0}$) is calculated conditional on the event of being fully occupied (unoccupied) at level $t+1$ of T. Let $\bar{h}_T^{\ddagger} \equiv \lim_{t\to\infty} \bar{h}_T^{2t-1,\ddagger}$ (existence of the limits $\bar{h}_T^0, \bar{h}_T^1$ for the independent set model follows from anti-monotonicity; see Section 2.4). We then define messages $h^\ddagger \in \mathcal{H}_\mu$ by $h^\ddagger_{x\to y} = \bar{h}^\ddagger_{T_{x\to y}}$, and let

$$\lambda_c \equiv \lambda_{c,\mu} \equiv \inf\{\lambda \ge 0 : \mu^\uparrow(h^{0,\lambda}_{x\to y} = h^{1,\lambda}_{x\to y}) < 1\}$$

denote the uniqueness threshold. For $T \in \mathcal{T}_\bullet$ we write

$$\begin{aligned} \operatorname{br} T &\equiv \inf\Big\{ y > 0 : \liminf_{|\Pi|\to\infty} \sum_{v\in\Pi} y^{-d(o,v)} = 0 \Big\} \\ &= \sup\Big\{ y > 0 : \liminf_{|\Pi|\to\infty} \sum_{v\in\Pi} y^{-d(o,v)} = \infty \Big\} \end{aligned} \tag{1.15}$$

(where the limit is taken over cutsets Π of T with distance $|\Pi|$ from the root tending to infinity) for the *branching number* of T; see [32], Section 2.

THEOREM 1.12. *Consider the independent set model* (1.4) *on* $G_n \to_{lwc} \mu$, *and write* $\lambda_c \equiv \lambda_{c,\mu}$.

(a) *If* $\lambda < \lambda_c$ *and the function* $\lambda \mapsto h^{0,\lambda} = h^{1,\lambda}$ *has total variation bounded by a deterministic constant on* $[0, \log \lambda]$, *then*

$$\phi(\lambda) = \Phi_\mu(\lambda, h^0) = \Phi_\mu(\lambda, h^1), \tag{1.16}$$

which converges to $\phi(\lambda_c)$ *as* $\lambda \uparrow \lambda_c$.

(b) *If* $\operatorname{br} T_{x\to y} \le \Delta - 1$ $\mu^{\uparrow}$*-a.s. for* Δ *a deterministic constant, then* (1.16) *holds for* $\lambda < \lambda_c$ *with* $\lambda(\Delta - 2) < 1$.

(c) *If* $\mu = \delta_{\mathsf{T}_d}$, *then* (1.16) *holds for* $\lambda \le \lambda_c$.

For the d-regular tree T_d, the uniqueness threshold $\lambda_c(d)$ is $(d-1)^{d-1}/(d-2)^d$ (see [29], Section 2), and [44], Theorem 2.3, shows that T_d has the lowest value of λ_c among trees with maximum degree at most d. The identity (1.16) has been proved in the case that the G_n are random d-regular graphs [3, 4]. It is also suggested by Weitz's PTAS for $Z_G(\lambda)$ on a finite graph G of maximum degree Δ and with $\lambda < \lambda_c(\Delta)$ ([44], Corollary 2.8). For μ a unimodular measure on $\mathcal{T}_\bullet$ giving a local tree approximation to G (in the sense of Definition 1.1), $\lambda_{c,\mu}$ is often an improvement over $\lambda_c(\Delta)$, making it possible to compute $\phi(\lambda)$ above $\lambda_c(\Delta)$ provided (H3^B) can be verified. In [41] the interpolation scheme of Theorem 1.12 is refined to give a verification of the Bethe prediction on locally tree-like d-regular *bipartite* graphs for all $\lambda > 0$; this result is then leveraged to show inapproximability of the hard-core partition function on d-regular graphs above $\lambda_c(d)$.

1.3. *Results for general factor models.* We now state our results for the factor model (1.1). With the convention $\log 0 \equiv -\infty$, let $\log \psi \equiv \xi$ and $\log \bar{\psi} \equiv \bar{\xi}$, and impose the following regularity condition:

(H1) The specification is *permissive*, that is, $\bar{\psi}(\sigma) > 0$ for all $\sigma \in \mathscr{X}$, and there exists a "permitted state" $\sigma^{\mathrm{p}} \in \mathscr{X}$ such that $\min_\sigma \psi(\sigma, \sigma^{\mathrm{p}}) > 0$.

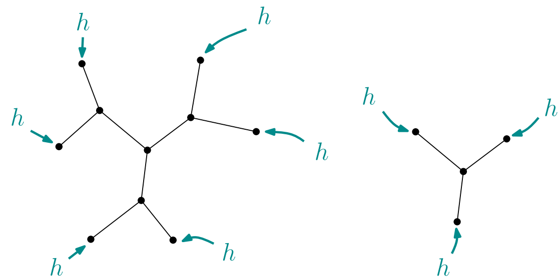
For any $\sigma \in \mathscr{X}$, $\bar{\xi}^B(\sigma)$ is continuously differentiable in B. For any $\sigma, \sigma' \in \mathscr{X}$, $\xi^\beta(\sigma, \sigma')$ is either identically $-\infty$ over all β, or finite and continuously differentiable in β.

Recalling Definition 1.4 of the message space $\mathcal{H} \equiv \mathcal{H}_\mu$, for $h \in \mathcal{H}_\mu$ we can define $\bar{h} : \mathcal{T}_\bullet \to \Delta_{\mathscr{X}}$ up to μ-equivalence by

$$\bar{h}_T(\sigma) \cong \bar{\psi}(\sigma) \prod_{j \in \partial o} \biggl(\sum_{\sigma_j} \psi(\sigma, \sigma_j) h_{j\to o}(\sigma_j) \biggr). \tag{1.17}$$

In particular, if $h \in \mathcal{H}^\star_\mu(\beta, B)$ and $T \in \mathcal{T}^+_\bullet$, then comparing (1.17) with (1.10) gives

$$\bar{h}_T(\sigma) \cong \sum_{\sigma_j} \psi(\sigma, \sigma_j) h_{o\to j}(\sigma) h_{j\to o}(\sigma_j) \tag{1.18}$$



(a) $U = T^2$ (b) $U = T^1$

FIG. 5. *A Bethe fixed point defines a consistent family of f.d.d.* $\nu^h_{U,T}$ (*Remark* 1.13).

independently of the choice of $j \in \partial o$. From now on, for $h \in \mathcal{H}_\mu$, we will write $h \in \mathcal{H}^\star$ to indicate that $h^{\beta,B} \in \mathcal{H}^\star_\mu(\beta, B)$ for (β, B) in the range being considered.

REMARK 1.13. The elements of $\mathcal{H}^\star$ are consistent with the recursion structure of the tree in the following precise sense: for $T \in \mathcal{T}_\bullet$ and U a finite connected sub-graph of T, consider the factor model $\nu^h_{U,T}$ on U with boundary conditions $\sigma_v \sim h_{v\to p(v)}$ independently for $v \in \partial U$, where $p(v)$ denotes the (necessarily unique) neighbor of v inside U. Then the marginal of $\nu^h_{T^t,T}$ on T^{t-1} is exactly the factor model $\nu^{\mathsf{BP}h}_{T^{t-1},T}$ on T^{t-1} with boundary conditions $\sigma_u \sim (\mathsf{BP}h)_{u\to p(u)}$ independently for $u \in \partial T^{t-1}$, including any u which are leaves of T^t. This statement remains valid if ∂T^t or even ∂T^{t-1} is empty, since if $\partial T^t = \varnothing$ then $\nu^h_{T^t,T}$ is simply ν_T as defined by (1.1). Continuing the recursion up the tree, we see that $h \in \mathcal{H}^\star$ implies that the marginal law of σ_o will be $\bar{h}_T$ as defined by (1.17). From this it is easy to see that the measures $\nu^h_{U,T}$ form a consistent family of finite-dimensional marginals (see Figure 5), so by the Kolmogorov consistency theorem they uniquely determine a probability measure $\nu_T \equiv \nu^h_T$ belonging to $\mathscr{G}_T$, the set of Gibbs measures (or Markov random fields) associated to the specification $\underline{\psi} \equiv (\psi, \bar{\psi})$ on T.[4] (In fact this mapping is one-to-one, e.g., by Remark 2.3 below.) Each ν_T belongs to a special class of measures in $\mathscr{G}_T$ which are called *Markov chains* or *splitting Gibbs measures* in the literature, and the entire collection $(\nu_T)_{T\in\mathcal{T}_\bullet}$ arising from $h \in \mathcal{H}^\star_\mu$ has a consistency property which leads us to term them "unimodular Markov chains" or "Bethe Gibbs measures;" see Section 2.3.

[4]Strictly speaking the term "Gibbs measures" refers to the case $\psi > 0$, but we will follow common practice and say Gibbs measures also for the general case. For the general theory of Gibbs measures see, for example, [21].

In this general setting, the *Bethe prediction* is the supremum of $\Phi_\mu(\beta, B, h)$ over $\mathcal{H}^\star_\mu(\beta, B)$; cf. Remark 1.7. (It will be shown in Lemma 2.2 that Φ_μ is uniformly bounded on $\mathcal{H}^\star_\mu(\beta, B)$ subject to $\mathbb{E}_\mu[D_o^2] < \infty$; if further $\psi > 0$, then Φ_μ is in fact uniformly bounded on $\mathcal{H}$ subject only to $\mathbb{E}_\mu[D_o] < \infty$.) We define the following integrability condition for unimodular measures μ on $\mathcal{T}_\bullet$ (not necessarily arising from a graph sequence):

(H2) The probability measure μ on $\mathcal{T}_\bullet$ satisfies $\mathbb{E}_\mu[D_o] < \infty$. If ψ is not everywhere positive, then furthermore $\mathbb{E}_\mu[e^{cD_o}] < \infty$ for all $c \in \mathbb{R}$.

Note that if $G_n \to_{lwc} \mu$ and $\psi > 0$, then (H2) holds trivially by the assumption of uniform sparsity. We will in fact justify our interpolation scheme under a weaker assumption than (H2); for the exact condition see (H2$^\beta$), (H2^B) in Section 2.2.

1.3.1. *Bethe interpolation.* We will deduce the results of Section 1.2 from the abstract interpolation method given by Theorem 1.15 below, which bounds differences of $\phi(\beta, B)$ by differences of $\Phi(\beta, B, h)$ ($h \in \mathcal{H}^\star$) when the limiting expectation of a certain edge or vertex functional in the finite graph (capturing resp. $\partial_\beta \phi_n$ or $\partial_B \phi_n$) is bounded by the expectation of an analogous functional on the infinite tree.

To be more precise, recall that I_n denotes a uniformly random vertex of V_n. Let $\langle\cdot\rangle_n^{\beta,B}$ denote expectation with respect to $\nu_{G_n,\underline{\psi}}$, conditioned on G_n. For $h \in \mathcal{H}^\star_\mu(\beta, B)$ and $T \in \mathcal{T}_\bullet$, let $[\![\,]\!]_T^{h,\beta,B}$ denote expectation with respect to ν_T^h (as defined in Remark 1.13), conditioned on T, and define

$$
\begin{aligned}
a_n^{\mathrm{e}}(\beta, B) &\equiv \frac{1}{2}\mathbb{E}_n\Bigg[\sum_{j\in\partial I_n} \langle \partial_\beta \xi(\sigma_{I_n}, \sigma_j)\rangle_n^{\beta,B}\Bigg],\\
a^{\mathrm{e}}(\beta, B, h) &\equiv \frac{1}{2}\mathbb{E}_\mu\Bigg[\sum_{j\in\partial o} [\![\partial_\beta \xi(\sigma_o, \sigma_j)]\!]_T^{h,\beta,B}\Bigg],\\
a_n^{\mathrm{vx}}(\beta, B) &\equiv \mathbb{E}_n\big[\langle \partial_B \bar{\xi}(\sigma_{I_n})\rangle_n^{\beta,B}\big],\\
a^{\mathrm{vx}}(\beta, B, h) &\equiv \mathbb{E}_\mu\big[[\![\partial_B \bar{\xi}(\sigma_o)]\!]_T^{h,\beta,B}\big].
\end{aligned}
$$

The left-hand side expressions are the derivatives $\partial_\beta \phi_n$, $\partial_B \phi_n$ (Lemma 2.1). The right-hand side expressions are the infinite-tree analogues, which, as we will show in Proposition 2.4, may be thought of as derivatives in β and B of Φ_μ.

EXAMPLE 1.14. For example in the Potts model (1.3) we have $\partial_B \bar{\xi}(\sigma) = \mathbf{1}\{\sigma = 1\}$, so $a_n^{\mathrm{vx}}(\beta, B)$ is the expected density of 1s in the graph while $a_n^{\mathrm{e}}(\beta, B)$ is $1/n$ times the expected number of edge agreements, both with respect to the Potts measure on G_n. The infinite tree analogues of a_n^{vx} and a_n^{e} are

$$a^{\mathrm{vx}}(\beta,B,h)=\mathbb{E}_\mu\Bigg[\Bigg(e^B\prod_{j\in\partial o}[(e^\beta-1)h_{j\to o}(1)+1]\Bigg)\Bigg/\Bigg(e^B\prod_{j\in\partial o}[(e^\beta-1)h_{j\to o}(1)+1]+\sum_{\sigma\neq 1}\prod_{j\in\partial o}[(e^\beta-1)h_{j\to o}(\sigma)+1]\Bigg)\Bigg],$$

the ν_T^h-probability (averaged over $T\sim\mu$) that the root spin takes value 1 and

$$a^{\mathrm{e}}(\beta,B,h)=\frac{1}{2}\mathbb{E}_\mu\Bigg[\sum_{j\in\partial o}\frac{\sum_\sigma e^\beta h_{o\to j}(\sigma)h_{j\to o}(\sigma)}{1+\sum_\sigma(e^\beta-1)h_{o\to j}(\sigma)h_{j\to o}(\sigma)}\Bigg],$$

the ν_T^h-expectation (averaged over $T\sim\mu$) of half the number of edge agreements incident to the root.

For interpolation in β on a compact interval $[\beta_0,\beta_1]$ using some particular $h\in\mathcal{H}^\star$, we require the following regularity condition on h:

(H3$^\beta$) On $[\beta_0,\beta_1]$, for all $\sigma\in\mathscr{X}$ it holds $\mu^\uparrow$-a.s. that the function $\beta\mapsto h^\beta_{x\to y}(\sigma)$ is continuous with total variation in β bounded by a deterministic constant depending only on β_0,β_1.

Likewise for interpolation in B on a compact interval $[B_0,B_1]$ using $h\in\mathcal{H}^\star$ we require

(H3^B) On $[B_0,B_1]$, for all $\sigma\in\mathscr{X}$ it holds $\mu^\uparrow$-a.s. that the function $B\mapsto h^B_{x\to y}(\sigma)$ is continuous with total variation in B bounded by a deterministic constant depending only on B_0,B_1.

The condition of boundedness in total variation is implied for example whenever the functions h are (anti-)monotone in the interpolation parameter.

THEOREM 1.15. *Let* $\underline{\psi}\equiv(\psi,\bar\psi)$ *specify a factor model* (1.1) *on* $G_n\to_{lwc}\mu$ *such that* (H1) *and* (H2) *are satisfied.*

(a) *If on* $[\beta_0,\beta_1]$ *we have* $h\in\mathcal{H}^\star$ *satisfying* (H3$^\beta$), *and*

$$\limsup_{n\to\infty}a_n^{\mathrm{e}}(\beta,B)\le a^{\mathrm{e}}(\beta,B,h), \tag{1.19}$$

then $\limsup_{n\to\infty}[\phi_n(\beta_1,B)-\phi_n(\beta_0,B)]\le\Phi(\beta_1,B,h)-\Phi(\beta_0,B,h)$.

(b) *If on* $[B_0,B_1]$, *we have* $h\in\mathcal{H}^\star$ *satisfying* (H3^B), *and*

$$\limsup_{n\to\infty}a_n^{\mathrm{vx}}(\beta,B)\le a^{\mathrm{vx}}(\beta,B,h), \tag{1.20}$$

then $\limsup_{n\to\infty}[\phi_n(\beta,B_1)-\phi_n(\beta,B_0)]\le\Phi(\beta,B_1,h)-\Phi(\beta,B_0,h)$.

The same results hold if all inequalities are reversed, replacing limit superior with inferior.

Conditions (1.19), (1.20) (and their reverses) are automatically verified in the following special case, where we recall that $\mathscr{G}_T$ denotes the set of Gibbs measures associated to the specification $\underline{\psi}$ on T; cf. Remark 1.13:

THEOREM 1.16. *Let* $\underline{\psi} \equiv (\psi, \bar{\psi})$ *specify a factor model* (1.1) *on* $G_n \to_{lwc} \mu$ *satisfying* (H1) *and* (H2). *We say that* uniqueness *holds if* $\mathscr{G}_T$ *at* (β, B) *consists of a single measure* ν_T, μ*-a.s. In this case,* $\mathcal{H}^\star_\mu(\beta, B)$ *is a singleton.*

(a) *If on* $[\beta_0, \beta_1] \times \{B\}$ *uniqueness holds and the unique element* $h \in \mathcal{H}^\star$ *satisfies* (H3$^\beta$), *then*

$$\lim_{n\to\infty}\big[\phi_n(\beta_1, B) - \phi_n(\beta_0, B)\big] = \Phi(\beta_1, B, h) - \Phi(\beta_0, B, h).$$

(b) *If on* $\{\beta\} \times [B_0, B_1]$ *uniqueness holds and the unique element* $h \in \mathcal{H}^\star$ *satisfies* (H3^B), *then*

$$\lim_{n\to\infty}\big[\phi_n(\beta, B_1) - \phi_n(\beta, B_0)\big] = \Phi(\beta, B_1, h) - \Phi(\beta, B_0, h).$$

Uniqueness for $\mathscr{G}_T$ corresponds to the vanishing effect of boundary conditions on ∂T^t as $t \to \infty$ ([21], Chapter 7). Dobrushin's uniqueness theorem (see, e.g., [39]) gives a sufficient condition for uniqueness to hold, together with a bound on the rate of convergence of the root marginal in T^t to the limit as $t \to \infty$. Note that if the convergence rate is uniform in β, B then the continuity required in (H3$^\beta$) and (H3^B) immediately follows. We will obtain continuity in uniqueness regimes via a different route, making use of certain monotonicity properties; see the proof of Theorem 1.9.

1.3.2. *Variational principle.* We further develop the theory by providing a variational principle for the Bethe prediction: we express $\Phi_\mu(\beta, B)$ as an optimum of a function $\Phi_\mu(\beta, B, \mathbf{h})$ defined for $\mathbf{h}$ in a larger space $\mathcal{H}_{\mathrm{loc}}$ which, unlike $\mathcal{H}^\star_\mu(\beta, B)$, is *independent* of β, B. This alternative characterization of Φ_μ is the infinite-tree analogue of the finite-graph optimization problem that is considered in [46]. Recall from Section 1.1 that $\mathcal{T}_{\mathrm{e}}$ denotes the space of trees rooted at a *directed* edge.

DEFINITION 1.17. The *local polytope* $\mathcal{H}_{\mathrm{loc}} \equiv \mathcal{H}_{\mathrm{loc},\mu}$ is the space of measurable functions

$$\mathbf{h} : \mathcal{T}_{\mathrm{e}} \to \Delta_{\mathscr{X}^2}, \qquad (T, x \to y) \mapsto \mathbf{h}_{(T,x\to y)} \equiv \mathbf{h}_{xy},$$

taken up to $\mu^\uparrow$-equivalence, such that:

(i) $\mathbf{h}_{xy}(\sigma, \sigma') = \mathbf{h}_{yx}(\sigma', \sigma)$ for all $\sigma, \sigma' \in \mathscr{X}$, and

(ii) for $T \in \mathcal{T}_\bullet^+$, the one-point marginal $\bar{h}_x(\sigma) \equiv \bar{h}_{(T,x)}(\sigma) \equiv \sum_{\sigma_y} \mathbf{h}_{xy}(\sigma, \sigma_y)$ is well-defined, that is, does not depend on the choice of $y \in \partial x$.

We also define

$$\mathcal{H}_{\mathrm{loc}}[\psi] \equiv \{\mathbf{h} \in \mathcal{H}_{\mathrm{loc}} : \mu^{\uparrow}(\operatorname{supp}\mathbf{h} \subseteq \operatorname{supp}\psi) = 1\},$$
$$\mathcal{H}^{\circ}_{\mathrm{loc}}[\psi] \equiv \{\mathbf{h} \in \mathcal{H}_{\mathrm{loc}} : \mu^{\uparrow}(\operatorname{supp}\mathbf{h} = \operatorname{supp}\psi) = 1\}.$$

In accordance with (1.17), we set

$$\bar{h}_T(\sigma) \cong \bar{\psi}(\sigma) \qquad \text{if } T = T^0. \tag{1.21}$$

For *fixed* (β, B), by symmetry of ψ^β and (1.18), the space $\mathcal{H}^\star_\mu(\beta, B)$ has a natural mapping into $\mathcal{H}_{\mathrm{loc}}$ given by

$$h \mapsto \mathbf{h}, \qquad \mathbf{h}_{xy}(\sigma, \sigma') \cong \psi(\sigma, \sigma') h_{x\to y}(\sigma) h_{y\to x}(\sigma'). \tag{1.22}$$

With $\underline{\psi}$ permissive this is in fact an embedding; see Remark 2.3. We define the *Bethe free energy functional on* $\mathcal{H}_{\mathrm{loc}}$ by

$$\begin{aligned}\Phi_\mu(\mathbf{h}) \equiv \mathbb{E}_\mu\Big[&\langle \bar{\xi}(\sigma_o)\rangle_{\bar{h}_o} - (D_o - 1) H(\bar{h}_o) \\ &+ \frac{1}{2}\sum_{j\in\partial o} \{\langle \xi(\sigma_o, \sigma_j)\rangle_{\mathbf{h}_{oj}} + H(\mathbf{h}_{oj})\}\Big],\end{aligned} \tag{1.23}$$

where $H(p)$ denotes the Shannon entropy $-\sum_k p_k \log p_k$ for p a probability measure on a finite space. This is an infinite-tree analogue of the definition of [46], (37)–(38), for finite graphs. With the usual conventions $\log 0 \equiv -\infty$, $0 \log 0 \equiv 0$ and $0 \log(0/0) \equiv 0$, Φ_μ is bounded above on $\mathcal{H}_{\mathrm{loc}}$ whenever $\mathbb{E}_\mu[D_o] < \infty$, and we show in Lemma 3.1 that for μ unimodular, this Φ_μ extends the previous definition (1.9) on $\mathcal{H}^\star$ [under the embedding (1.22)], provided the latter is finite. Furthermore, writing $H(q\|p)$ for the relative entropy $\sum_k q_k \log(q_k/p_k)$ between q and p (well defined for any nonnegative reference measure p), for μ unimodular we can alternatively express

$$\begin{aligned}\Phi_\mu(\mathbf{h}) &= -\mathbb{E}_\mu[H(\bar{h}_o\|\bar{\psi})] - \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{j\in\partial o} H(\mathbf{h}_{oj}\|\psi)\Big] - \mathbb{E}_\mu[D_o H(\bar{h}_o)] \\ &= -\mathbb{E}_\mu[H(\bar{h}_o\|\bar{\psi})] - \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{j\in\partial o} \{H(\mathbf{h}_{oj}\|\psi) + H(\bar{h}_o) + H(\bar{h}_j)\}\Big]\end{aligned} \tag{1.24}$$

$$= -\mathbb{E}_\mu[H(\bar{h}_o\|\bar{\psi})] - \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{j\in\partial o} H(\mathbf{h}_{oj}\|\bar{h}_o \times_\psi \bar{h}_j)\Big], \tag{1.25}$$

where $(\bar{h}_o \times_\psi \bar{h}_j)(\sigma_o, \sigma_j) \equiv \bar{h}_o(\sigma_o)\psi(\sigma_o, \sigma_j)\bar{h}_j(\sigma_j)$, and unimodularity is used in the second identity.

This extended definition of Φ_μ provides the following variational principle for the Bethe free energy:

THEOREM 1.18. *Let $\underline{\psi} \equiv (\psi, \bar{\psi})$ specify a factor model* (1.1) *satisfying* (H1), *and let μ be a unimodular measure on $\mathcal{T}_\bullet$ with $\mathbb{E}_\mu[D_o] < \infty$.*

(a) $\widetilde{\Phi}_\mu(\beta, B) \equiv \sup_{\mathbf{h} \in \mathcal{H}_{\text{loc}}} \Phi_\mu(\beta, B, \mathbf{h})$ *is continuous in* (β, B).

(b) *Any local maximizer of $\Phi_\mu(\beta, B)$ belongs to $\mathcal{H}^\circ_{\text{loc}}[\psi]$. Any stationary point of $\Phi_\mu(\beta, B)$ belonging to $\mathcal{H}^\circ_{\text{loc}}[\psi]$ is the image under* (1.22) *of an element of $\mathcal{H}^\star_\mu(\beta, B)$. In particular, if Φ_μ attains its supremum on $\mathcal{H}_{\text{loc}}$, then*

$$\widetilde{\Phi}_\mu(\beta, B) = \max_{h \in \mathcal{H}^\star_\mu(\beta, B)} \Phi_\mu(\beta, B, h) \equiv \Phi^{\text{Bethe}}_\mu(\beta, B),$$

so that the Bethe free energy is also continuous in (β, B).

Although we do not pursue this point, we mention that even in specific models where the abstract definition of Φ^{Bethe} is supplanted by $\Phi(\beta, B, h)$ for some "naturally" distinguished h, an adaptation of Theorem 1.18 [involving a restricted subspace of $\mathcal{H}_{\text{loc}}$ which is independent of (β, B)], may be relevant.

REMARK 1.19. In case $G_n \to_{lwc} \mathsf{T}_d$ the d-regular tree, $\mathcal{H}_{\text{loc}}$ is parametrized by a single measure $\mathbf{h}_{xy}$ on $\mathscr{X}^2$ whose one-point marginals are required to agree, and the formula (1.25) simplifies to

$$-\Phi_\mu(\mathbf{h}) = H(\bar{h}_0 \| \bar{\psi}) + \frac{d}{2} H(\mathbf{h}_{01} \| \bar{h}_0 \times_\psi \bar{h}_1). \tag{1.26}$$

For $\underline{\sigma} \in \mathscr{X}^{V_n}$ let $L^{\text{vx}}_n \equiv n^{-1} \sum_{i \in V_n} \delta_{\sigma_i}$ and $L^{\text{e}}_n \equiv (2|E_n|)^{-1} \sum_{(ij) \in E_n} [\delta_{(\sigma_i, \sigma_j)} + \delta_{(\sigma_j, \sigma_i)}]$ denote the induced empirical and pair empirical measures, respectively. If G_n is d-regular, then the one-point marginals of L^{e}_n coincide with L^{vx}_n, and

$$\phi_n = \log|\mathscr{X}| + \frac{1}{n} \mathbb{E}_n\left[\log \mathbb{E}_{\bar{u}_n} \exp\left\{n \langle \bar{\xi} \rangle_{L^{\text{vx}}_n} + \frac{nd}{2} \langle \xi \rangle_{L^{\text{e}}_n}\right\}\right],$$

where the law of $\underline{\sigma}$ is the uniform measure $\bar{u}_n$ on $\mathscr{X}^{[n]}$ and $\mathbb{E}_{\bar{u}_n}$ denotes expectation with respect to $\bar{u}_n$ (with G_n fixed).

If (G_n) is an independent sequence of uniformly random d-regular graphs and $\underline{\sigma}_n \sim \bar{u}_n$, one might guess that for a.e. (G_n) the induced sequence L^{e}_n satisfies a large deviation principle (LDP) with good rate function

$$I(\mathbf{h}_{01}) = H(\bar{h}_0 \| \bar{u}) + \frac{d}{2} H(\mathbf{h}_{01} \| \bar{h}_0 \times \bar{h}_1), \tag{1.27}$$

where $\bar{u} \equiv \bar{u}_1$. If this were the case, it would be an immediate consequence of Varadhan's lemma (see [12], Section 4.3.1) that $\phi_n \to \widetilde{\Phi}_\mu(\beta, B)$ (as defined in Theorem 1.18) for *any* factor model satisfying (H1). However, for many of these models the Bethe prediction is known to fail at low temperature for $d \geq 3$. So, while Theorem 1.18 suggests a potential connection to large deviations theory, such a connection would be highly nontrivial and applicable only in certain regimes of (β, B).

One special case in which everything trivializes is the (rooted) infinite line T_2, the local weak limit of the simple path G_n on n vertices. In this case $\bar{u}_n$ may be viewed as the law of a stationary reversible Markov chain on $\mathscr{X}$ with transitions $q(\sigma,\sigma') = \bar{u}(\sigma')$ and reversing measure $\bar{u}$, and it is well-known (see, e.g., [12], Theorem 3.1.13) that the associated pair empirical measure L_n^{e} satisfies an LDP with good rate function $I(\mathbf{h}_{01}) = H(\mathbf{h}_{01}(\sigma,\sigma') \| \bar{h}_0(\sigma) q(\sigma,\sigma'))$ which matches (1.27). The implication of Varadhan's lemma is also easy to see: a factor model on the simple path G_n with general positive specification $\underline{\psi}$ corresponds in the limit $n \to \infty$ to a reversible Markov chain with transition kernel p and positive reversing measure π given by

$$p(\sigma,\sigma') = \frac{1}{\rho} \widetilde{\psi}(\sigma,\sigma') \frac{m(\sigma')}{m(\sigma)}, \qquad \pi(\sigma) = m(\sigma)^2,$$

where ρ and m are the Perron–Frobenius eigenvalue and eigenvector of the symmetric positive $|\mathscr{X}|$-dimensional matrix with entries $\widetilde{\psi}(\sigma,\sigma') \equiv \psi(\sigma,\sigma') \times \bar{\psi}(\sigma)^{1/2} \bar{\psi}(\sigma')^{1/2}$. The Bethe free energy functional (1.26) is then maximized at $\mathbf{h}_{01}(\sigma,\sigma') = \widetilde{\psi}(\sigma,\sigma') m(\sigma) m(\sigma')/\rho$, where it takes the value $\Phi_\mu(\mathbf{h}) = \log \rho$ which coincides with ϕ by the Perron–Frobenius theorem; see, for example, [12], Theorem 3.1.1.

Outline of the paper.

- In Section 2 we prove the abstract interpolation results. Section 2.1 presents some preliminary lemmas which will be useful in our proofs. Our main result for abstract factor models, Theorem 1.15, is proved in Section 2.2. Section 2.3 contains the specialization of this theorem to the uniqueness case (Theorem 1.16) and also contains discussion on unimodular Markov chains (or Bethe Gibbs measures). Section 2.4 shows how to deduce our result for independent set (Theorem 1.12) from Theorem 1.15.
- In Section 3 we prove the variational characterization Theorem 1.18 for the Bethe free energy prediction, establishing in particular the correspondence between interior stationary points $\mathbf{h} \in \mathcal{H}^\circ_{\mathrm{loc}}[\psi]$ of Φ_μ and fixed points $h \in \mathcal{H}^\star$ of the Bethe recursion. We further provide in Proposition 3.4 a simple criterion for such stationary points to be local maximizers.
- Section 4 contains applications of our abstract results to the Ising and Potts models. In Section 4.1 we prove Theorem 1.9, generalizing the results of [10, 13]. In Section 4.2 we prove Theorem 1.10 by appealing to a random-cluster representation. Finally, Section 4.3 analyzes the d-regular case and proves Theorem 1.11.

2. Bethe interpolation for general factor models.

2.1. *Preliminaries*. We begin with some straightforward observations on the boundedness of the free energy ϕ_n and the Bethe free energy Φ_μ as defined on $\mathcal{H}$, and we prove that the mapping (1.22) of $\mathcal{H}^\star$ into $\mathcal{H}_{\mathrm{loc}}$ is in fact an embedding for permissive specifications.

LEMMA 2.1. *For the factor model* (1.1) *satisfying* (H1) *on* $G_n \to_{lwc} \mu$, *the functions* $\phi_n(\beta, B)$ *are uniformly bounded and equicontinuous on compact regions of* (β, B), *with*

$$\begin{aligned}\partial_\beta \phi_n(\beta, B) &= \frac{1}{n}\mathbb{E}_n\big[\partial_\beta \log Z_n(\beta, B)\big],\\ \partial_B \phi_n(\beta, B) &= \frac{1}{n}\mathbb{E}_n\big[\partial_B \log Z_n(\beta, B)\big].\end{aligned} \tag{2.1}$$

Further,

$$\frac{1}{n}\partial_\beta \log Z_n(\beta, B) = \frac{1}{2}\sum_{j\in\partial I_n}\big\langle \partial_\beta \xi(\sigma_{I_n}, \sigma_j)\big\rangle_n^{\beta,B},$$

$$\frac{1}{n}\partial_B \log Z_n(\beta, B) = \big\langle \partial_B \bar{\xi}(\sigma_{I_n})\big\rangle_n^{\beta,B},$$

with the convention $\partial_\beta \xi(\sigma, \sigma') \equiv 0$ *in case* $\xi(\sigma, \sigma') \equiv -\infty$.

PROOF. The expressions for $n^{-1}\partial_\beta \log Z_n(\beta, B)$ and $n^{-1}\partial_B \log Z_n(\beta, B)$ are obtained by a straightforward computation. Now note that if $G_n \to_{lwc} \mu$, then the uniform sparsity assumption gives

$$\frac{1}{n}\mathbb{E}_n\big[|E_n|\big] = \frac{1}{2}\mathbb{E}_n[D_{I_n}] \to \frac{1}{2}\mathbb{E}_\mu[D_o] < \infty. \tag{2.2}$$

Let (β, B) vary within a given compact region. By (H1) we have $\xi, \bar{\xi} \le \xi_{\max}$ as well as $\xi(\sigma^{\mathrm{p}}, \cdot), \bar{\xi} \ge \xi_{\min}$. Therefore,

$$\big(1 + |E_n|/n\big)\xi_{\min} \le n^{-1}\log Z_n(\beta, B) \le \log|\mathscr{X}| + \big(1 + |E_n|/n\big)\xi_{\max}$$

so $\phi_n = n^{-1}\mathbb{E}_n[\log Z_n(\beta, B)]$ is uniformly bounded by uniform sparsity. The exchange of differentiation and integration in (2.1) is justified by Vitali's convergence theorem, in view of the boundedness of $\partial_\beta \xi$, $\partial_B \bar{\xi}$ and the uniform integrability of $|E_n|/n$. It follows furthermore that $\partial_\beta \phi_n(\beta, B)$ and $\partial_B \phi_n(\beta, B)$ are bounded uniformly in n, from which equicontinuity follows. $\square$

LEMMA 2.2. *Let* $\underline{\psi} \equiv (\psi, \bar{\psi})$ *specify a factor model* (1.1) *satisfying* (H1), *and let* μ *be a unimodular measure on* $\mathcal{T}_\bullet$. *For any compact region of* (β, B) *there exists a deterministic constant* $C < \infty$ *such that*:

(a) $|\Phi_T(\beta, B, h)| \le C(D_o^2 + 1)$ *for any* $h \in \mathcal{H}^\star$, *and*
(b) *if further* $\psi > 0$, *then* $|\Phi_T(\beta, B, h)| \le C(D_o + 1)$ *for any* $h \in \mathcal{H}$.

Hence, *on any compact region of* (β, B), Φ_μ *is uniformly bounded on* $\mathcal{H}_\mu^\star$ *provided* $\mathbb{E}_\mu[D_o^2] < \infty$, *and if* $\psi > 0$, *uniformly bounded on* $\mathcal{H}_\mu$ *subject only to* $\mathbb{E}_\mu[D_o] < \infty$.

PROOF. Let $\xi_{\min}, \xi_{\max}$ be as in the proof of Lemma 2.1. Then, for any $h \in \mathcal{H}$,

$$\log|\mathscr{X}| + (D_o + 1)\xi_{\max} \ge \Phi_T^{\mathrm{vx}}(h) \ge (D_o + 1)\xi_{\min}.$$

If $\psi > 0$, then we also have

$$D_o\xi_{\max} \ge 2\Phi_T^{\mathrm{e}}(h) \ge D_o\xi_{\min},$$

so $|\Phi_T(\beta, B, h)| \le C(D_o + 1)$ on $\mathcal{H}$, which proves (b). For general permissive ψ, the preceding lower bound on $\Phi_T^{\mathrm{e}}(h)$ may fail, but (1.11) implies that for $h \in \mathcal{H}^\star$,

$$\log h_{o\to j}(\sigma^{\mathrm{p}}) \ge D_o(\xi_{\min} - \xi_{\max}) - \log|\mathscr{X}| \qquad \forall j \in \partial o. \tag{2.3}$$

Therefore,

$$\begin{aligned} D_o\xi_{\max} \ge 2\Phi_T^{\mathrm{e}}(h) &\ge \sum_{j\in\partial o} \bigl(\xi_{\min} + \log h_{o\to j}(\sigma^{\mathrm{p}})\bigr) \\ &\ge D_o\bigl(\xi_{\min} - \log|\mathscr{X}|\bigr) + D_o^2(\xi_{\min} - \xi_{\max}), \end{aligned}$$

which proves (a). $\square$

REMARK 2.3. It is now easy to see that the mapping (1.22) of $\mathcal{H}_\mu^\star(\beta, B)$ into $\mathcal{H}_{\mathrm{loc}}$ is injective: if $h, h' \in \mathcal{H}$ give rise to the same $\mathbf{h}$, then

$$h_{x\to y}(\sigma)h_{y\to x}(\sigma^{\mathrm{p}}) = z_{x,y}h'_{x\to y}(\sigma)h'_{y\to x}(\sigma^{\mathrm{p}}) \qquad \forall\sigma \in \mathscr{X}$$

for $z_{x,y}$ a positive scaling factor. If $h, h' \in \mathcal{H}_\mu^\star(\beta, B)$, then (2.3) implies that $\mu^\uparrow$-a.s. both $h_{y\to x}$ and $h'_{y\to x}$ give positive measure to σ^{p}. Therefore, $\mu^\uparrow$-a.s. the $|\mathscr{X}|$-dimensional vectors $h_{x\to y}$ and $h'_{x\to y}$ are equivalent up to scaling, and since both are probability measures on $\mathscr{X}$, we must have $h = h'$ $\mu^\uparrow$-a.s. as claimed.

2.2. *Bethe interpolation*. We now prove Theorem 1.15(a). The result is for fixed B, so we suppress it from the notation. The proof of Theorem 1.15(b) is very similar and will be given in brief at the end of this section.

Our interpolation procedure relies on the proposition below which expresses Φ_μ as the integral of its partial derivative with respect to β only, ignoring the dependence on β through the function h. Recall that although it is suppressed from the notation, ψ and h depend on β, and are taken to be evaluated at β in expressions such as $\overline{\Phi}_T(\beta, B)$. We will prove our result under the following integrability condition, which by (2.3) is a relaxation of (H2):

($\mathrm{H2}^\beta$) The probability measure μ on $\mathcal{T}_\bullet$ satisfies $\mathbb{E}_\mu[D_o] < \infty$. If ψ is not everywhere positive, then furthermore,

$$\mathbb{E}_\mu\biggl[\sum_{j\in\partial o}\sup_{\beta\in[\beta_0,\beta_1]}\frac{1}{h^\beta_{j\to o}(\sigma^{\mathrm{p}})}\biggr] < \infty.$$

We define the analogous condition ($\mathrm{H2}^B$) on an interval $[B_0, B_1]$.

PROPOSITION 2.4. *Let $\underline{\psi} \equiv (\psi, \bar{\psi})$ be a specification satisfying* (H1), *and μ a unimodular measure on $\mathcal{T}_\bullet$. If on $[\beta_0, \beta_1]$ we have $h \in \mathcal{H}^\star$ satisfying* ($\mathrm{H2}^\beta$) *and* ($\mathrm{H3}^\beta$), *then*

$$\Phi(\beta_1, h) - \Phi(\beta_0, h) = \int_{\beta_0}^{\beta_1} a^{\mathrm{e}}(\beta, h)\, d\beta.$$

PROOF. For fixed $T \in \mathcal{T}_\bullet$ we shall regard Φ_T simply as a function of a vector $(\beta, h_{x\to y})_{x\to y\in T^1}$ in $(1+2|\mathscr{X}|D_o)$-dimensional euclidean space (with h depending on β). We begin by computing the partial derivatives of this function with respect to β and h. We abbreviate $\widehat{h}^\beta_{o\to j}(\sigma) \equiv (\mathsf{BP}^\beta h)_{o\to j}(\sigma)$ for the belief propagation mapping of (1.10), which for fixed T and each $j \in \partial o$ is a well-defined function on the same euclidean space as Φ_T. Making use of (H1) we find

$$\frac{\partial \Phi^{\mathrm{vx}}_T}{\partial\beta}(\beta, h) = \sum_{j\in\partial o}\frac{\sum_{\sigma,\sigma_j}\partial_\beta\xi(\sigma,\sigma_j)\psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)\widehat{h}^\beta_{o\to j}(\sigma)}{\sum_{\sigma,\sigma_j}\psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)\widehat{h}^\beta_{o\to j}(\sigma)}, \tag{2.4}$$

$$\frac{\partial \Phi^{(oj)}_T}{\partial\beta}(\beta, h) = \frac{\sum_{\sigma,\sigma_j}\partial_\beta\xi(\sigma,\sigma_j)\psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)h_{o\to j}(\sigma)}{\sum_{\sigma,\sigma_j}\psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)h_{o\to j}(\sigma)}. \tag{2.5}$$

If $h \in \mathcal{H}^\star$, then $\widehat{h}^\beta = h$, therefore (recalling the notation $[\![]\!]^{h,\beta}_T$ from Section 1.3.1) we re-express the above as

$$\frac{\partial \Phi^{\mathrm{vx}}_T}{\partial\beta}(\beta, h) = \sum_{j\in\partial o}[\![\partial_\beta\xi(\sigma_o,\sigma_j)]\!]^{h,\beta}_T, \qquad \frac{\partial \Phi^{(oj)}_T}{\partial\beta}(\beta, h) = [\![\partial_\beta\xi(\sigma_o,\sigma_j)]\!]^{h,\beta}_T,$$

and combining gives

$$\frac{\partial \Phi_T}{\partial\beta}(\beta, h) = \frac{1}{2}\sum_{j\in\partial o}[\![\partial_\beta\xi(\sigma_o,\sigma_j)]\!]^{h,\beta}_T \equiv a^{\mathrm{e}}_T(\beta, h). \tag{2.6}$$

Likewise we compute that for $T \in \mathcal{T}^+_\bullet$,

$$\frac{\partial \Phi^{\mathrm{vx}}_T(\beta, h)}{\partial h_{o\to j}(\sigma)} = 0,$$

$$\frac{\partial \Phi_T^{\mathrm{vx}}(\beta,h)}{\partial h_{j\to o}(\sigma_j)} = \widehat{g}^{\beta}_{\sigma_j}(j\to o;h) \equiv \frac{\sum_{\sigma}\psi(\sigma,\sigma_j)\widehat{h}^{\beta}_{o\to j}(\sigma)}{\sum_{\sigma',\sigma'_j}\psi(\sigma',\sigma'_j)h_{j\to o}(\sigma'_j)\widehat{h}^{\beta}_{o\to j}(\sigma')},$$

$$\frac{\partial \Phi_T^{\mathrm{e}}(\beta,h)}{\partial h_{j\to o}(\sigma_j)} = \frac{1}{2}g^{\beta}_{\sigma_j}(j\to o;h), \qquad \frac{\partial \Phi_T^{\mathrm{e}}(\beta,h)}{\partial h_{o\to j}(\sigma)} = \frac{1}{2}g^{\beta}_{\sigma}(o\to j;h),$$

where g^{β}_{σ} is the same as $\widehat{g}^{\beta}_{\sigma}$ but with h in place of $\widehat{h}$. Note that for permissive ψ and any $\sigma \in \mathscr{X}$,

$$\widehat{g}^{\beta}_{\sigma}(x\to y;h) \le \frac{\sum_{\sigma'}\psi(\sigma',\sigma)\widehat{h}^{\beta}_{y\to x}(\sigma')}{\sum_{\sigma'}\psi(\sigma',\sigma^{\mathrm{p}})h_{x\to y}(\sigma^{\mathrm{p}})\widehat{h}^{\beta}_{y\to x}(\sigma')} \le \frac{\psi^{\beta}_{\max}}{\psi^{\beta}_{\min}h_{x\to y}(\sigma^{\mathrm{p}})}. \tag{2.7}$$

If further $\psi > 0$ everywhere, then $\widehat{g}^{\beta}_{\sigma}(x\to y;h) \le \psi^{\beta}_{\max}/\psi^{\beta}_{\min}$ is uniformly bounded on $[\beta_0,\beta_1]$.

Consider now a small sub-interval $[\beta,\beta+\delta]$ of $[\beta_0,\beta_1]$. Writing $\Delta_{\beta,\delta}h \equiv h^{\beta+\delta} - h^{\beta}$ and applying the mean value theorem to the differentiable function $t \mapsto \Phi_T(\beta+t\delta, h+t\Delta_{\beta,\delta}h)$ for $t\in[0,1]$ gives

$$\begin{aligned} &\Phi_T(\beta+\delta,h) - \Phi_T(\beta,h) \\ &\quad = \frac{\partial \Phi_T}{\partial \beta}(\beta+t\delta, h^{\beta}+t\Delta_{\beta,\delta}h)\delta + \Gamma_T(\beta,\delta) + E_T(\beta,\delta) \end{aligned} \tag{2.8}$$

for some $t \equiv t_{\beta,\delta} \in [0,1]$, where

$$\begin{aligned} \Gamma_T(\beta,\delta) \equiv \sum_{\sigma}\sum_{j\in\partial o}\biggl\{&\frac{\partial \Phi_T}{\partial h_{j\to o}(\sigma)}(\beta,h)\Delta_{\beta,\delta}h_{j\to o}(\sigma) \\ &+ \frac{\partial \Phi_T}{\partial h_{o\to j}(\sigma)}(\beta,h)\Delta_{\beta,\delta}h_{o\to j}(\sigma)\biggr\}, \\ E_T(\beta,\delta) \equiv \sum_{\sigma}\sum_{x\to y}{}^{*}\biggl\{&\frac{\partial \Phi_T}{\partial h_{x\to y}(\sigma)}(\beta+t\delta, h^{\beta}+t\Delta_{\beta,\delta}h) \\ &- \frac{\partial \Phi_T}{\partial h_{x\to y}(\sigma)}(\beta,h)\biggr\}\Delta_{\beta,\delta}h_{x\to y}(\sigma), \end{aligned}$$

and $\sum^{*}_{x\to y}$ indicates the sum over the $2D_o$ directed edges $x\to y$ within T^1.

Setting $\delta \equiv \delta_m \equiv (\beta_1-\beta_0)/m$, we now sum $\Phi(\beta+\delta_m,h) - \Phi(\beta,h)$ over $\beta \in \Pi_m \equiv \{\beta_0 + k\delta_m : 0 \le k < m\}$ and analyze separately the contribution of each term on the right-hand side of (2.8):

(a) First we show that $\mathbb{E}_{\mu}[\Gamma_T(\beta,\delta)] = 0$ for any $[\beta,\beta+\delta] \subseteq [\beta_0,\beta_1]$. Indeed, since $h \in \mathcal{H}^{\star}$ we have $\widehat{h}^{\beta} = h^{\beta}$ and $\widehat{g}^{\beta} = g^{\beta}$. Therefore,

$$\Gamma_T(\beta,\delta) = \frac{1}{2}\sum_{\sigma}\sum_{j\in\partial o}\bigl\{g^{\beta}_{\sigma}(j\to o;h)\Delta_{\beta,\delta}h_{j\to o}(\sigma) - g^{\beta}_{\sigma}(o\to j;h)\Delta_{\beta,\delta}h_{o\to j}(\sigma)\bigr\}.$$

The result then follows from unimodularity of μ, subject to μ-integrability of

$$\sum_{\sigma}\sum_{j\in\partial o}\left|g_{\sigma}^{\beta}(j\to o;h)\Delta_{\beta,\delta}h_{j\to o}(\sigma)\right|.$$

Clearly $|\Delta_{\beta,\delta}h_{x\to y}(\sigma)|\le 2$ so integrability certainly holds when $\psi>0$, since $\mathbb{E}_{\mu}[D_o]<\infty$ and g_{σ}^{β} is deterministically uniformly bounded on $[\beta_0,\beta_1]$ as noted above. More generally, for permissive ψ the required μ-integrability follows from (2.7) and (H2$^{\beta}$).

(b) The total contribution of the first term on the right-hand side of (2.8) is

$$A_m\equiv\delta\mathbb{E}_{\mu}\biggl[\sum_{\beta\in\Pi_m}\frac{\partial\Phi_T}{\partial\beta}(\beta+t_{\beta,\delta}\delta,h^{\beta}+t_{\beta,\delta}\Delta_{\beta,\delta}h)\biggr].$$

Observe that $A_m=\int Y_m\,d(\lambda\times\mu)$ where λ is Lebesgue measure on $[\beta_0,\beta_1]$ and

$$Y_m(\beta',T)\equiv\sum_{\beta\in\Pi_m}\mathbf{1}\{\beta\le\beta'<\beta+\delta\}\frac{\partial\Phi_T}{\partial\beta}(\beta+t_{\beta,\delta}\delta,h^{\beta}+t_{\beta,\delta}\Delta_{\beta,\delta}h).$$

For $(\lambda\times\mu)$-a.e. (β',T), this sum has at most one nonzero term, in which the argument of $\partial_{\beta}\Phi_T$ converges by (H3$^{\beta}$) to $(\beta',h^{\beta'})$ as $m\to\infty$. From (H1), (1.10) and the computation of $\partial_{\beta}\Phi_T$ in (2.4)–(2.5), we see that $\partial_{\beta}\Phi_T(\beta,h)$ is continuous in (β,h). Therefore, $Y_m(\beta',T)\to a_T^{\mathrm{e}}(\beta',h)$, $(\lambda\times\mu)$-a.e. Furthermore, (H1) implies that $|\partial_{\beta}\xi|\le C$ uniformly on $[\beta_0,\beta_1]$ for some deterministic constant C, so $|Y_m|\le 2CD_o$ for all m, $(\lambda\times\mu)$-a.e. see (2.4) and (2.5). Dominated convergence then gives

$$\lim_{m\to\infty}A_m=\int a_T^{\mathrm{e}}(\beta',h)\,d(\lambda\times\mu)=\int_{\beta_0}^{\beta_1}a^{\mathrm{e}}(\beta',h)\,d\beta'.$$

(c) The contribution of the final term in (2.8) is $\mathbb{E}_{\mu}[E_{T,m}]$ where

$$E_{T,m}\equiv\sum_{\beta\in\Pi_m}E_T(\beta,\delta),$$

and we conclude the proof by showing that $\lim_{m\to\infty}\mathbb{E}_{\mu}[E_{T,m}]=0$.

Indeed, it is not hard to see that $\lim_{m\to\infty}E_{T,m}=0$ μ-a.s.: by the uniform bound on total variation assumed in (H3$^{\beta}$), there exists deterministic C such that

$$\begin{aligned}|E_{T,m}|\le C\sum_{x\to y}{}^{*}\max_{\sigma}\sup_{\beta\in[\beta_0,\beta_1]}\sup_{t\in[0,1]}\biggl|&\frac{\partial\Phi_T}{\partial h_{x\to y}(\sigma)}(\beta+t\delta_m,h^{\beta}+t\Delta_{\beta,\delta_m}h)\\&-\frac{\partial\Phi_T}{\partial h_{x\to y}(\sigma)}(\beta,h)\biggr|\end{aligned}$$

μ-a.s., uniformly in m. It also follows from (H3$^{\beta}$) that μ-a.e. h^{β} is uniformly continuous on $[\beta_0,\beta_1]$. Using (H1), the partials $\partial_h\Phi_T$ computed above are uniformly continuous in (β, h) for $\beta \in [\beta_0, \beta_1]$ and $h^{\beta}_{j\to o}(\sigma^{\mathrm{p}})$ uniformly bounded away from zero. By (2.3) there exists deterministic c such that

$$\inf_{\beta\in[\beta_0,\beta_1]} h^{\beta}_{j\to o}(\sigma^{\mathrm{p}}) \geq e^{-c(D_j+1)} \qquad \forall j \in \partial o, \mu\text{-a.s.}$$

Combining these observations gives $\lim_{m\to\infty} E_{T,m} = 0$ μ-a.s.

To take the limit in μ-expectation, we argue similarly as in part (a): by (2.7) and (H1) there exists deterministic C' such that

$$\begin{aligned}\left|\frac{\partial \Phi_T}{\partial h_{x\to y}(\sigma)}(\beta + t\delta, h^{\beta} + t\Delta_{\beta,\delta}h)\right| &\leq \frac{C'}{h^{\beta}_{x\to y}(\sigma^{\mathrm{p}}) + t\Delta_{\beta,\delta}h_{x\to y}(\sigma^{\mathrm{p}})} \\ &\leq \sup_{\beta'\in[\beta_0,\beta_1]} \frac{C'}{h^{\beta'}_{x\to y}(\sigma^{\mathrm{p}})}\end{aligned}$$

for all $\beta \in [\beta_0, \beta_1 - \delta]$, $x \to y \in T^1$, $\sigma \in \mathscr{X}$ and $t \in [0, 1]$, hence

$$|E_{T,m}| \leq CC' \sum_{x\to y}^{*} \sup_{\beta\in[\beta_0,\beta_1]} \frac{1}{h^{\beta}_{x\to y}(\sigma^{\mathrm{p}})}.$$

This is integrable by (H2$^{\beta}$) and unimodularity of μ, so dominated convergence implies that $\lim_{m\to\infty} \mathbb{E}_{\mu}[E_{T,m}] = 0$ as claimed.

Combining (a)–(c) gives the result of the proposition. □

PROOF OF THEOREM 1.15(A). Recalling Lemma 2.1,

$$\begin{aligned}\limsup_{n\to\infty}[\phi_n(\beta_1) - \phi_n(\beta_0)] &= \limsup_{n\to\infty} \int_{\beta_0}^{\beta_1} a^{\mathrm{e}}_n(\beta, B)\, d\beta \\ &\leq \int_{\beta_0}^{\beta_1} \limsup_{n\to\infty} a^{\mathrm{e}}_n(\beta, B)\, d\beta \leq \int_{\beta_0}^{\beta_1} a^{\mathrm{e}}(\beta, h)\, d\beta,\end{aligned}$$

where the first inequality follows by (the reversed) Fatou's lemma and the second one by the hypothesis (1.19). By Proposition 2.4 the right-most expression equals to $\Phi(\beta_1, h) - \Phi(\beta_0, h)$, so the theorem is proved. □

The justification for interpolation in B is entirely similar:

PROOF OF THEOREM 1.15(B). Now β is fixed, so we suppress it from the notation. For $h \in \mathcal{H}$ and $T \in \mathcal{T}_{\bullet}^{+}$, then

$$\frac{\partial \Phi_T(B,h)}{\partial B} = \frac{\partial \Phi^{\mathrm{vx}}_T(B,h)}{\partial B} = \frac{\sum_{\sigma,\sigma_j} \partial_B \bar{\xi}(\sigma)\psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)\widehat{h}_{o\to j}(\sigma)}{\sum_{\sigma,\sigma_j} \psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)\widehat{h}_{o\to j}(\sigma)} \qquad \forall j \in \partial o,$$

while if $T = T^0$, then $\partial_B \Phi_T = \sum_\sigma \partial_B \bar{\xi}(\sigma)\bar{\psi}(\sigma)/\sum_\sigma \bar{\psi}(\sigma)$. If $h \in \mathcal{H}^\star$, then $\widehat{h}^B = h^B$, so

$$\mathbb{E}_\mu[\partial_B \Phi_T(B,h)] = \mathbb{E}_\mu[[\![\partial_B \bar{\xi}(\sigma_o)]\!]_T^{h,B}] \equiv a^{\mathrm{vx}}(B,h).$$

The result now follows by adapting the proofs of Proposition 2.4 and Theorem 1.15(a). □

2.3. *Discussion and first consequences.* We now prove Theorem 1.16 by considering an extended notion of local weak convergence. As discussed in [1], a graph $G = (V, E)$ together with a spin configuration $\underline{\sigma} \in \mathscr{X}^V$ on the graph can be regarded as a graph with *marks* in $\mathscr{X}$. Let $\mathcal{G}_\bullet^{\mathscr{X}}$ and $\mathcal{G}_{\bullet\bullet}^{\mathscr{X}}$ denote the spaces of marked isomorphism classes of connected, rooted and bi-rooted graphs, respectively, with marks in $\mathscr{X}$. These spaces are metrizable by the obvious generalizations of the metrics on $\mathcal{G}_\bullet, \mathcal{G}_{\bullet\bullet}$ defined in Section 2.1, giving rise to the notion of local weak convergence for pairs $(G_n, \underline{\sigma}_n)$ of graphs with spin configurations. Definition 1.2 generalizes naturally to this setting, and we show next that if $\underline{\sigma}_n$ is a random configuration on G_n with law $\nu_{G_n,\underline{\psi}}$ [as defined in (1.1)], then a local weak limit of $(G_n, \underline{\sigma}_n)$, if it exists, must be unimodular.

LEMMA 2.5. *If $G_n \to_{lwc} \mu$ and $\underline{\sigma}_n \sim \nu_{G_n,\underline{\psi}}$, then the laws of $(G_n, \underline{\sigma}_n)$ have subsequential local weak limits belonging to the space $\mathscr{U}$ of unimodular measures on $\mathcal{G}_\bullet^{\mathscr{X}}$.*

PROOF. For each fixed t, the laws of $B_t(I_n)$ are weakly convergent, hence by Prohorov's theorem form a uniformly tight sequence. Consequently, for each $\varepsilon > 0$ there exists $K_\varepsilon \subseteq \mathcal{G}_\bullet$ compact with $\sup_n \mathbb{P}_n(B_t(I_n) \notin K_\varepsilon) \le \varepsilon$. Further, K_ε may be taken to contain only graphs of depth at most t, whereby the minimal distance between any two graphs in K_ε is uniformly bounded below [by $1/(1+t)$], hence the compactness of K_ε implies that it must be a finite set. The collection of all marked graphs in $\mathcal{G}_\bullet^{\mathscr{X}}$ whose underlying graph is in K_ε must therefore be finite, hence compact as well. Thus, by yet another application of Prohorov's theorem, the joint laws of $(B_t(I_n), \underline{\sigma}_{B_t(I_n)})$ are uniformly tight in $\mathcal{G}_\bullet^{\mathscr{X}}$ and consequently have subsequential weak limits. By extracting successive subsequences for increasing t and taking the diagonal subsequence, it follows that the sequence $(G_n, \underline{\sigma}_n)$ admits subsequential local weak limits $\widehat{\mu} \in \mathscr{U}$. □

For $\widehat{\mu} \in \mathscr{U}$, the marginal μ of $\widehat{\mu}$ is a unimodular measure on $\mathcal{G}_\bullet$. If it is supported on a single tree T as in the d-regular case, then clearly $\widehat{\mu}$ may be represented as $\delta_T \times \nu$ where $\nu \in \mathscr{G}_T$, the space of Gibbs measures on T corresponding to specification $\underline{\psi}$. To make such a statement in the general setting, note that there is a continuous mapping π from $\mathcal{G}_\bullet$ to the space $\mathcal{N}_\bullet$ of graphs on $\mathbb{Z}_{\ge 0}$ rooted at 0, taking an isomorphism class to its canonical representative ([1], page 1461). Thus $\widehat{\mu}$

may be regarded as a measure on the product space $\mathcal{N}_\bullet \times \mathscr{X}^{\mathbb{Z}_{\geq 0}}$, and consequently $\widehat{\mu}$ has a representation as the measure $\mu \otimes \nu$ on pairs $(T, \underline{\sigma})$ where T has law μ and $\underline{\sigma}$ given T has law $\nu_T \in \mathscr{G}_T$. In particular, if $|\mathscr{G}_T| = 1$ μ-a.s., then $\mu \otimes \nu$ is uniquely determined.

Let μ be a unimodular measure on $\mathcal{T}_\bullet$. It was noted in Remark 1.13 that there is a mapping from $\mathcal{H}^\star_\mu(\beta, B)$ to collections $(\nu_T \in \mathscr{G}_T)_{T \in \mathcal{T}_\bullet}$. For such ν, $\mu \otimes \nu$ belongs to $\mathscr{U}$: if f is a nonnegative Borel function on $\mathcal{G}^{\mathscr{X}}_{\bullet\bullet}$, it follows from the $\mathcal{T}_{\mathrm{e}}$-measurability of elements of $\mathcal{H}_{\mathrm{loc}}$ that

$$\mathbb{E}_{\mu\otimes\nu}\biggl[\sum_{j\in\partial o} f\bigl((T,\underline{\sigma}), o, j\bigr)\biggr] = \mathbb{E}_\mu\biggl[\sum_{j\in\partial o} \bar{f}(T, o, j)\biggr],$$

where $\bar{f}$ is a nonnegative Borel function on $\mathcal{G}_{\bullet\bullet}$. The unimodularity of the underlying measure μ then gives

$$\mathbb{E}_{\mu\otimes\nu}\biggl[\sum_{j\in\partial o} f\bigl((T,\underline{\sigma}), o, j\bigr)\biggr] = \mathbb{E}_{\mu\otimes\nu}\biggl[\sum_{j\in\partial o} f\bigl((T,\underline{\sigma}), j, o\bigr)\biggr],$$

and therefore $\mu \otimes \nu \in \mathscr{U}$.

REMARK 2.6. An element $\nu \in \mathscr{G}_T$ is called a *Markov chain* (or *splitting Gibbs measure*) if for any finite connected sub-graph $U \subseteq T$, the marginal of ν on U is a Markov random field [47]; see also [21], Chapter 12, and [43]. A collection $\Lambda_T \equiv (\lambda_i^j)_{(ij)\in E_T}$ of probability measures on $\mathscr{X}$ is called an *entrance law* (or *boundary law*) for the specification $\underline{\psi} \equiv (\psi, \bar{\psi})$ on T if it satisfies the consistency requirement ([47], (3.4))

$$\lambda_i^j(\sigma_i) = \prod_{k\in\partial i\setminus j}\biggl(\sum_{\sigma_k} \phi_{ik}(\sigma_i, \sigma_k)\lambda_k^i(\sigma_k)\biggr),$$

where $\phi_{ij}(\sigma_i, \sigma_j) \equiv \bar{\psi}(\sigma)^{1/D_i}\psi(\sigma, \sigma')\bar{\psi}(\sigma')^{1/D_j}$, the pairwise interaction potential corresponding to $\underline{\psi}$. It is shown in [47], Theorem 3.2, that there is a one-to-one correspondence between Markov chains ν and entrance laws Λ_T, given by

$$\nu(\underline{\sigma}_U) \cong \prod_{(ij)\in E_U} \phi_{ij}(\sigma_i, \sigma_j) \prod_{i\in\partial U}\biggl(\sum_{\sigma_i} \phi_{ip(i)}(\sigma_i, \sigma_{p(i)})\lambda_i^{p(i)}(\sigma_i)\biggr)$$

for U any finite connected sub-graph of T, with $p(i)$ denoting the unique neighbor of i inside U for $i \in \partial U$. In particular, we see that the Gibbs measure ν_T arising from $h \in \mathcal{H}^\star_\mu(\beta, B)$ is precisely the Markov chain with entrance law $\lambda_j^i(\sigma) \cong h_{j\to i}(\sigma)\bar{\psi}(\sigma)^{-1/D_j}$. Extremal elements of $\mathscr{G}_T$ are Markov chains ([47], Theorem 2.1), but the converse is false; for example, the free-boundary Ising Gibbs measure is nonextremal at low temperature; see [15, 27]. The measures $\mu \otimes \nu$ arising from elements of $\mathcal{H}^\star_\mu(\beta, B)$ might naturally be termed "unimodular

Markov chains" or "Bethe Gibbs measures," in the sense that the entrance laws for the entire collection $(\nu_T)_{T\in\mathcal{T}_\bullet}$ are specified by a single measurable function $h:\mathcal{T}_{\mathrm{e}}\to\Delta_{\mathscr{X}}$ which is a Bethe fixed point. In the case $\mu=\delta_{\mathsf{T}_d}$ these correspond precisely to the *completely homogeneous Markov chains* studied in [47], Section 4.

PROOF OF THEOREM 1.16. Suppose uniqueness holds at (β,B), that is, $\mathscr{G}_T=\{\nu_T\}$ μ-a.s. Then $\mathcal{H}^\star_\mu(\beta,B)$ has size at most one by Remark 2.3. For μ-a.e. T, the measure ν_T is extremal, and so specifies a Markov chain on T with entrance law Λ_T; see Remark 2.6. If we define $h_{x\to y}(\sigma)\equiv h_{(T,x\to y)}(\sigma)\cong\lambda^y_x(\sigma)\bar\psi(\sigma)^{1/D_x}$, then $h\in\mathcal{H}^\star_\mu(\beta,B)$, which proves that $\mathcal{H}^\star_\mu(\beta,B)$ is a singleton.

Now consider interpolation in β or B. All the conditions of Theorem 1.15 are satisfied by assumption except (1.19) and (1.20). If uniqueness holds at (β,B), it follows from the preceding discussion that there is a unique $\mu\otimes\nu\in\mathscr{U}$ corresponding to the specification $(\psi^\beta,\bar\psi^B)$. Any local weak limit of $(G_n,\underline{\sigma}_n)$ must be such a measure, so $(G_n,\underline{\sigma}_n)\to_{lwc}\mu\otimes\nu$; likewise, any element of $\mathcal{H}^\star_\mu(\beta,B)$ gives rise to $\mu\otimes\nu$. Therefore,

$$\lim_{n\to\infty}a^{\mathrm{e}}_n(\beta,B)=\frac{1}{2}\mathbb{E}_{\mu\otimes\nu}\biggl[\sum_{j\in\partial o}\partial_\beta\xi(\sigma_o,\sigma_j)\biggr]=a^{\mathrm{e}}(\beta,B,h),$$

where the limit in expectation is justified by the boundedness of $\partial_\beta\xi$ on compacts and uniform sparsity (as in the proof of Lemma 2.1). This verifies (1.19), and the verification of (1.20) is entirely similar. The result therefore follows from Theorem 1.15. $\square$

REMARK 2.7. If uniqueness of Gibbs measures does not hold, one may consider extremal decomposition of the subsequential local weak limits $\widehat{\mu}$ of $(G_n,\underline{\sigma}_n)$, either in the spaces $\mathscr{G}_T$ (possibly losing unimodularity in the decomposition), or in the space $\mathscr{U}$. Extremal decomposition in $\mathscr{U}$ is discussed in [1], Section 4, but it is unclear whether extremal elements would be unimodular Markov chains in the sense described here. A decomposition of $\widehat{\mu}=\mu\otimes\nu$ into unimodular Markov chains $\mu\otimes\nu'$ would obviously yield a substantial generalization of Theorem 1.16.

2.4. *Application to independent set.* We now prove Theorem 1.12, our result for the independent set model (1.4), by verifying the conditions of Theorem 1.16 for the interpolation parameter $B\equiv\log\lambda$. In this setting a convenient parametrization for the messages $h\in\mathcal{H}$ is $u\equiv h(0)$, so that the BP mapping (1.10) becomes

$$(\mathsf{BP}^\lambda u)_{x\to y}=\frac{1}{1+\lambda\prod_{v\in\partial x\setminus y}u_{v\to x}}. \tag{2.9}$$

A single BP iteration is anti-monotone in the messages $u_{v\to x}$, so a double iteration is monotone. Since the root marginal for an independent set model in T^{2t-1} is obtained by an even number of BP iterations starting from level $2t$ (see Remark 1.13), it is monotone in the boundary conditions. Recalling from Section 1.2.3 the definition of $\bar{h}_T^{t,\ddagger} \equiv \bar{h}_T^{t,\ddagger,\lambda}$ for $\ddagger \in \{0, 1\}$ and writing $\bar{u}_T^{t,\ddagger} \equiv \bar{h}_T^{t,\ddagger}(0)$, the above implies that for $1 \le s \le t$,

$$\bar{u}_T^{2s-1,0} \ge \bar{u}_T^{2t-1,0} \ge \bar{u}_T^{2t-1,1} \ge \bar{u}_T^{2s-1,1} \ge \bar{u}_T^{1,1} = \frac{1}{1+\lambda}.$$

Thus the $t \to \infty$ limits $\bar{h}_T^0, \bar{h}_T^1$ are well-defined with $\bar{h}_T^0(1) \ge \bar{h}_T^1(1) \ge 1/(1+\lambda)$, and using these we define messages $h^\ddagger \in \mathcal{H}$, $h^\ddagger_{x\to y} = \bar{h}^\ddagger_{T_{x\to y}}$. The next lemma gives the boundary values for the interpolation.

LEMMA 2.8. *For the independent set model on $G_n \to_{lwc} \mu$,*

$$\lim_{\lambda\downarrow 0} \limsup_{n\to\infty} |\phi_n(\lambda)| = 0 = \lim_{\lambda\downarrow 0} \Phi(\lambda, h^\ddagger), \qquad \ddagger \in \{0, 1\}.$$

PROOF. The left limit follows from the trivial bounds $1 \le Z_n \le (1+\lambda)^n$. Next, for any $h \in \mathcal{H}$,

$$\Phi_T^{\mathrm{vx}}(\lambda, h) = \log\left\{1 + \lambda \prod_{j\in\partial o} h_{j\to o}(0)\right\},$$

so $\Phi_T^{\mathrm{vx}}(\lambda, h) \to 0$ both μ-a.s. and in μ-expectation as $\lambda \downarrow 0$, by bounded convergence. The same holds for $\Phi_T^{\mathrm{e}}(\lambda, h^\ddagger)$, $\ddagger \in \{0, 1\}$, using the bound $h^\ddagger_{x\to y}(0) \ge 1/(1+\lambda)$. $\square$

PROOF OF THEOREM 1.12. The independent set model (1.4) is of form (1.1) with $\mathscr{X} = \{0, 1\}$, $\psi(\sigma, \sigma') = \mathbf{1}\{\sigma\sigma' \ne 1\}$, and $\bar{\psi}(\sigma) = \lambda^\sigma \equiv e^{B\sigma}$, so (H1) is clearly satisfied with $\sigma^{\mathrm{p}} = 0$ the permitted state. By definition of λ_c, if $\lambda < \lambda_c$, then $h^0 = h^1 \equiv h$ in $\mathcal{H}$, and it then follows from the recursive structure of the tree that $h \in \mathcal{H}^\star_\mu(\lambda)$. Since $h^\ddagger_{x\to y}(0) \ge 1/(1+\lambda)$ as noted above, (H2^B) is satisfied on any compact interval of λ.

For $T \in \mathcal{T}_\bullet$, as noted above the root occupation probability on T^s for $s \ge 2t-1$ with *any* boundary conditions is sandwiched between $\bar{h}_T^{2t-1,0}(1)$ and $\bar{h}_T^{2t-1,1}(1)$, with the former increasing to $\bar{h}_T^0(1)$ and the latter decreasing to $\bar{h}_T^1(1)$. Since the $\bar{h}_T^{t,\ddagger}$ are clearly continuous in λ, it follows that $\bar{h}_T^0(1)$ and $\bar{h}_T^1(1)$ are, respectively, lower and upper semi-continuous in λ, so if they coincide, then their common value $\bar{h}_T(1)$ is continuous in λ. Applying this with $T = T_{x\to y}$ gives the $\mu^\uparrow$-a.s. continuity of $h^\ddagger_{x\to y}$ on $(0, \lambda_c)$.

For $T \in \mathcal{T}_\bullet$, $\bar{h}_T^\ddagger$ for $\ddagger \in \{0, 1\}$ is a function of $(h^{1-\ddagger}_{j\to o})_{j\in\partial o}$, so for $\lambda < \lambda_c$ we have that $\bar{h}_T^0 = \bar{h}_T^1$, μ-a.s. It then follows from the preceding observations and Remark 1.13 that the boundary effect vanishes and $|\mathscr{G}_T| = 1$ μ-a.s. Thus, we are in the setting of Theorem 1.16(b), and it remains only to complete the verification of (H3^B), that is, the boundedness in total variation of the messages $h_{x\to y}$:

(a) No verification is needed since boundedness in total variation is simply assumed.

(b) For $T \in \mathcal{T}_\bullet$, $\bar{u}_T^{t,\ddagger} \equiv \bar{h}_T^{t,\ddagger}(0)$ satisfies

$$\log \bar{u}_T^{2t+1,\ddagger} = -\log\Bigl(1+\lambda \prod_{j\in\partial o} \frac{1}{1+\lambda\prod_{k\in\partial j\setminus o} \bar{u}_{T_{k\to j}}^{2t-1,\ddagger}}\Bigr).$$

Differentiating with respect to λ, we find that $r_T^{t,\ddagger} \equiv (1+\lambda)\partial_\lambda \log \bar{u}_T^{t,\ddagger}$ satisfies

$$|r_T^{2t+1,\ddagger}| \le 1 + \frac{\lambda}{1+\lambda} D_o + \Bigl(\frac{\lambda}{1+\lambda}\Bigr)^2 \sum_{j\in\partial o}\sum_{k\in\partial j\setminus o} |r^{2t-1,\ddagger}|_{T_{k\to j}}. \tag{2.10}$$

Since $\bar{u}_T^{1,1} = 1/(1+\lambda)$ for any $T \in \mathcal{T}_\bullet$, we find that

$$\sup_{t\ge 0} |r_T^{2t-1,1}| \le 1 + \sum_{\ell\ge 1}\Bigl(\frac{\lambda}{1+\lambda}\Bigr)^\ell |\partial T^{\ell-1}|.$$

If $\lambda_1/(1+\lambda_1) < 1/\operatorname{br} T$, then this is finite and uniformly bounded on $[\lambda_0,\lambda_1]$ (see (1.15) or [32], Section 2), and consequently $\bar{u}_T^1 \equiv \lim_{t\to\infty} \bar{u}_T^{2t-1,1}$ has deterministically bounded total variation on $[\lambda_0,\lambda_1]$. If $\lambda_1 < \lambda_{c,\mu}$, then $h_{x\to y}^{2t-1,1} \to h_{x\to y}$ on $[\lambda_0,\lambda_1]$, so if $\operatorname{br} T_{x\to y} \le \Delta - 1$ $\mu^\uparrow$-a.s. and $\lambda_1/(1+\lambda_1) < 1/(\Delta-1)$ [i.e., $\lambda_1(\Delta-2) < 1$], then h has deterministically bounded total variation on $[\lambda_0,\lambda_1]$.

(c) Since the limiting measure is supported on T_d, only $h \equiv h_{(\mathsf{T}_{d,x\to y})}$ is of relevance, and (2.9) reduces to $\mathsf{BP}^\lambda u = (1+\lambda u^{d-1})^{-1}$. For $\lambda \le \lambda_c = \lambda_c(d)$ there is a unique fixed point (see [29], Section 2), which is then easily seen to be monotone in λ.

Thus (H3^B) is verified in parts (a)–(c). Also, $\phi(\lambda_c) = \lim_{\lambda\uparrow\lambda_c}\phi(\lambda)$ as an immediate consequence of Lemma 2.1. The rest of the theorem follows by applying Theorem 1.16 and then taking $B_0 = \log\lambda_0 \to -\infty$, relying on the boundary value given by Lemma 2.8. $\square$

3. Bethe prediction as optimization over local polytope. Throughout this section we assume that $\underline{\psi} \equiv (\psi,\bar{\psi})$ satisfying (H1) specifies a factor model (1.1), and that μ is a unimodular measure on $\mathcal{T}_\bullet$ with $\mathbb{E}_\mu[D_o] < \infty$. We study the Bethe prediction as the optimization of the Bethe free energy functional Φ_μ on $\mathcal{H}_{\mathrm{loc}}$ as defined by (1.23). We first verify that this agrees with the previous definition (1.9) of Φ_μ on $\mathcal{H}_\mu^\star(\beta,B)$, which we always regard as being embedded into $\mathcal{H}_{\mathrm{loc}}$ via (1.22). Recall from Definition 1.17 that for $\mathbf{h} \in \mathcal{H}_{\mathrm{loc}}$, the one-point marginals of $\mathbf{h}_{xy}$ are denoted $\bar{h}_x$ and $\bar{h}_y$, and are measurable functions $\mathcal{T}_\bullet \to \Delta_{\mathscr{X}}$.

LEMMA 3.1. *The functional Φ_μ on $\mathcal{H}_{\mathrm{loc}}$ given by (1.25) agrees with the previous definition (1.9) on $\mathcal{H}_\mu^\star(\beta,B)$, subject to finiteness of $\mathbb{E}_\mu[\Phi_T^{\mathrm{e}}]$.*

PROOF. If $\mathbf{h}$ corresponds to $h \in \mathcal{H}_\mu^\star(\beta, B)$, then (1.22) and (1.11) imply that

$$\mathbf{h}_{xy}(\sigma,\sigma')\exp\{\Phi_T^{(xy)}(h)\} = \psi(\sigma,\sigma')h_{x\to y}(\sigma)h_{y\to x}(\sigma'),$$

$$\bar{h}_o(\sigma)\exp\{\Phi_T^{\mathrm{vx}}(h)\} = \bar{\psi}(\sigma)\prod_{j\in\partial o}\biggl(\sum_{\sigma_j}\psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)\biggr).$$

Letting $\Phi^{(i)}(\mathbf{h})$ $(1 \le i \le 3)$ denote the three terms on the right-hand side of (1.24), it follows from the above that

$$\Phi^{(1)}(\mathbf{h}) = \mathbb{E}_\mu[\Phi_T^{\mathrm{vx}}(h)] - \mathbb{E}_\mu\biggl[\sum_{j\in\partial o}\sum_{\sigma}\bar{h}_o(\sigma)\log\biggl(\sum_{\sigma_j}\psi(\sigma,\sigma_j)h_{j\to o}(\sigma_j)\biggr)\biggr],$$

$$\begin{aligned}\Phi^{(2)}(\mathbf{h}) &= \mathbb{E}_\mu[\Phi_T^{\mathrm{e}}(h)] - \frac{1}{2}\mathbb{E}_\mu\biggl[\sum_{j\in\partial o}\sum_{\sigma,\sigma_j}\mathbf{h}_{oj}(\sigma,\sigma_j)\log(h_{o\to j}(\sigma)h_{j\to o}(\sigma_j))\biggr]\\ &= \mathbb{E}_\mu[\Phi_T^{\mathrm{e}}(h)] - \mathbb{E}_\mu\biggl[\sum_{j\in\partial o}\sum_{\sigma}\bar{h}_o(\sigma)\log h_{o\to j}(\sigma)\biggr],\end{aligned}$$

$$\begin{aligned}\Phi^{(3)}(\mathbf{h}) &= \mathbb{E}_\mu\biggl[\sum_{j\in\partial o}\sum_{\sigma}\bar{h}_o(\sigma)\log\biggl(\sum_{\sigma_j}\psi(\sigma,\sigma_j)h_{o\to j}(\sigma)h_{j\to o}(\sigma_j)\biggr)\biggr]\\ &\quad - 2\mathbb{E}_\mu[\Phi_T^{\mathrm{e}}(h)],\end{aligned}$$

where unimodularity was used in the simplification of $\Phi^{(2)}$. Adding these three identities gives $\Phi_\mu(\mathbf{h}) = \mathbb{E}_\mu[\Phi_T^{\mathrm{vx}}(h) - \Phi_T^{\mathrm{e}}(h)]$, as claimed. $\square$

As mentioned in Section 1.3.2, our definition Φ_μ of the Bethe free energy functional on $\mathcal{H}_{\mathrm{loc}}$ is an infinite-tree analogue of the definition of [46] for finite graphs. It is proved in [46], Proposition 6, that when $\psi > 0$, all local maxima of the Bethe free energy lie in the *interior* of the local polytope. We now prove an analogous result for infinite unimodular trees, assuming only permissivity of $\underline{\psi}$.

PROPOSITION 3.2. *For permissive* $\underline{\psi}$, *if* $\mathbf{h}$ *is a local maximizer of* Φ_μ *over* $\mathcal{H}_{\mathrm{loc}}$, *then* $\mathbf{h} \in \mathcal{H}_{\mathrm{loc}}^\circ[\psi]$.

PROOF. Assume without loss that $\mathbf{h} \in \mathcal{H}_{\mathrm{loc}}[\psi]$, since otherwise clearly $\Phi_\mu(\mathbf{h}) = -\infty$. If $\mathbf{u} \in \mathcal{H}_{\mathrm{loc}}[\psi]$, then it follows by convexity of $\mathcal{H}_{\mathrm{loc}}$ that $\mathbf{h}^\eta \equiv \mathbf{h} + \eta(\mathbf{u} - \mathbf{h}) \equiv \mathbf{h} + \eta\boldsymbol{\delta}$ belongs to $\mathcal{H}_{\mathrm{loc}}[\psi]$ for any $\eta \in (0, 1]$. Letting

$$R^\eta(\boldsymbol{\delta}) \equiv \frac{2}{\eta}[\Phi_\mu(\mathbf{h}^\eta) - \Phi_\mu(\mathbf{h})], \qquad \widehat{R}^\eta(\boldsymbol{\delta}) \equiv \frac{R^\eta(\boldsymbol{\delta})}{|\log\eta|},$$

our claim will follow upon showing that if $h \notin \mathcal{H}_{\mathrm{loc}}^\circ[\psi]$, then there exists such $\mathbf{u}$ for which

$$\lim_{\eta\downarrow 0}\widehat{R}^\eta(\boldsymbol{\delta}) = \widehat{R}^0(\boldsymbol{\delta}) > 0.$$

To this end, note that by an easy computation $[H(\mathbf{h}^\eta) - H(\mathbf{h})]/\eta = -\langle \log \mathbf{h}^\eta \rangle_{\boldsymbol{\delta}} - \langle f^\eta(\boldsymbol{\delta}/\mathbf{h}) \rangle_{\mathbf{h}}$, where $f^\eta(r) \equiv \eta^{-1}\log(1+\eta r)$ and $(\boldsymbol{\delta}/\mathbf{h})(\sigma,\sigma')$ is defined to be $\boldsymbol{\delta}(\sigma,\sigma')/\mathbf{h}(\sigma,\sigma')$ if $\mathbf{h}(\sigma,\sigma') > 0$, zero otherwise; note that $\mathbf{u} = \mathbf{h} + \boldsymbol{\delta} \ge 0$ implies $\boldsymbol{\delta}/\mathbf{h} \ge -1$. Thus from (1.23) we obtain $R^\eta(\boldsymbol{\delta}) = R_1^\eta(\boldsymbol{\delta}) + R_2^\eta(\boldsymbol{\delta})$ where

$$
\begin{aligned}
R_1^\eta(\boldsymbol{\delta}) \equiv \mathbb{E}_\mu\bigg[& 2\langle \bar{\xi} \rangle_{\bar{\delta}_o} + 2(D_o - 1)\langle f^\eta(\bar{\delta}_o/\bar{h}_o)\rangle_{\bar{h}_o} \\
& + \sum_{j \in \partial o} \big(\langle \xi \rangle_{\boldsymbol{\delta}_{oj}} - \langle f^\eta(\boldsymbol{\delta}_{oj}/\mathbf{h}_{oj})\rangle_{\mathbf{h}_{oj}} \big) \bigg],
\end{aligned} \tag{3.1}
$$

$$
R_2^\eta(\boldsymbol{\delta}) \equiv \mathbb{E}_\mu\bigg[2(D_o-1)\langle \log \bar{h}_o^\eta \rangle_{\bar{\delta}_o} - \sum_{j \in \partial o} \langle \log \mathbf{h}_{oj}^\eta \rangle_{\boldsymbol{\delta}_{oj}} \bigg].
$$

Since for $r \ge -1$ and $\eta \in (0,1)$, we have $\eta^{-1}\log(1-\eta) \le f^\eta(r) \le r$, it follows from dominated convergence (and the boundedness of ξ on $\operatorname{supp}\boldsymbol{\delta}$) that $R_1^\eta(\boldsymbol{\delta})$ converges to a finite limit as $\eta \downarrow 0$, and so converges to zero upon rescaling by $|\log \eta|$. Again by dominated convergence, $R_2^\eta(\boldsymbol{\delta})/|\log \eta|$ converges as $\eta \downarrow 0$ to

$$
\begin{aligned}
\widehat{R}^0(\boldsymbol{\delta}) = \mathbb{E}_\mu\bigg[& (2 - 2D_o)\bar{u}_o\big(\{\sigma : \bar{h}_o(\sigma) = 0\}\big) \\
& + \sum_{j \in \partial o} \mathbf{u}_{oj}\big(\{\sigma,\sigma' : \mathbf{h}_{oj}(\sigma,\sigma') = 0\}\big) \bigg].
\end{aligned} \tag{3.2}
$$

Let $A_o \equiv A_{(T,o)} \equiv \{\sigma \in \mathscr{X} : \bar{h}_o(\sigma) = 0\}$. Since $\mathbf{h}_{xy}(\sigma,\sigma') = 0$ whenever either $\sigma \in A_x$ or $\sigma' \in A_y$, we have by unimodularity of μ that

$$
\begin{aligned}
\widehat{R}^0(\boldsymbol{\delta}) &\ge \mathbb{E}_\mu\bigg[(2 - 2D_o)\bar{u}_o(A_o) + \sum_{j \in \partial o} \{ \bar{u}_o(A_o) + \bar{u}_j(A_j) - \mathbf{u}_{oj}(A_o \times A_j) \} \bigg] \\
&= \mathbb{E}_\mu\bigg[2\bar{u}_o(A_o) - \sum_{j \in \partial o} \mathbf{u}_{oj}(A_o \times A_j) \bigg] = \mathbb{E}_\mu\bigg[\sum_{j \in \partial o} \widehat{R}_{o \to j} \bigg],
\end{aligned}
$$

where $\widehat{R}_{o \to j} \equiv \mathbf{1}\{D_o > 0\}[D_o^{-1}\bar{u}_o(A_o) + D_j^{-1}\bar{u}_j(A_j) - \mathbf{u}_{oj}(A_o \times A_j)]$ [by (1.21), necessarily $A_o = \varnothing$ when $D_o = 0$].

Noting that $A_o^c \ne \varnothing$, consider the measurable function $\bar{u} : \mathcal{T}_\bullet^+ \to \Delta_{\mathscr{X}}$ defined (up to μ-equivalence) by

$$
\begin{aligned}
\bar{u}_o(\sigma) &\equiv \bar{u}_{(T,o)}(\sigma) \\
&\equiv \begin{cases} \mathbf{1}\{\sigma = \sigma^{\mathrm{p}}\}, & A_o^c = \{\sigma^{\mathrm{p}}\}, \\ \dfrac{1}{2}\bigg(\mathbf{1}\{\sigma = \sigma^{\mathrm{p}}\} + \dfrac{\mathbf{1}\{\sigma \in A_o^c \setminus \{\sigma^{\mathrm{p}}\}\}}{|A_o^c \setminus \{\sigma^{\mathrm{p}}\}|} \bigg), & \text{else.} \end{cases}
\end{aligned} \tag{3.3}
$$

Among those $\mathbf{u} \in \mathcal{H}_{\mathrm{loc}}$ with support contained in $\{(\sigma,\sigma') : \sigma^{\mathrm{p}} \in \{\sigma,\sigma'\}\}$, there is a unique one with marginals (3.3). On the event $\{D_o > 0\}$, we have the following:

– If $\sigma^{\mathrm{p}} \in A_o \cap A_j$, then

$$\mathbf{u}_{oj}(\sigma, \sigma') = \frac{1}{2}\left(\mathbf{1}\{\sigma = \sigma^{\mathrm{p}}\}\frac{\mathbf{1}\{\sigma' \in A_j^c \setminus \{\sigma^{\mathrm{p}}\}\}}{|A_j^c \setminus \{\sigma^{\mathrm{p}}\}|} + \mathbf{1}\{\sigma' = \sigma^{\mathrm{p}}\}\frac{\mathbf{1}\{\sigma \in A_o^c \setminus \{\sigma^{\mathrm{p}}\}\}}{|A_o^c \setminus \{\sigma^{\mathrm{p}}\}|}\right),$$

so $\widehat{R}_{o\to j} = (2D_o)^{-1} + (2D_j)^{-1}$.
– If $\sigma^{\mathrm{p}} \in A_o \cap A_j^c$, then $\bar{u}_o(A_o) \geq 1/2$ while $\bar{u}_j(A_j) = 0 = \mathbf{u}_{oj}(A_o \times A_j)$, so $\widehat{R}_{o\to j} \geq (2D_o)^{-1}$. Symmetrically if $\sigma^{\mathrm{p}} \in A_o^c \cap A_j$, then $\widehat{R}_{o\to j} \geq (2D_j)^{-1}$.
– If $\sigma^{\mathrm{p}} \notin A_o \cup A_j$, then $\widehat{R}_{o\to j} = 0$.

Thus $\widehat{R}^0(\boldsymbol{\delta}) \geq 0$, with strict inequality unless $\sigma^{\mathrm{p}} \notin A_o \cup A_j$ $\mu^{\uparrow}$-a.s., in which case we take A_o, A_j in place of A_o^c, A_j^c in (3.3). Then

$$\widehat{R}_{o\to j} = (2D_o)^{-1}\mathbf{1}\{A_o \neq \varnothing\} + (2D_j)^{-1}\mathbf{1}\{A_j \neq \varnothing\},$$

so $\widehat{R}^0(\boldsymbol{\delta}) > 0$ unless $\mu(A_o = \varnothing) = 1$. But in this case taking $\mathbf{u} \in \mathcal{H}_{\mathrm{loc}}$ identically equal to the uniform measure on $\operatorname{supp}\psi$ gives

$$\widehat{R}^0(\boldsymbol{\delta}) = \frac{1}{|\operatorname{supp}\psi|}\mathbb{E}_\mu\left[\sum_{j\in\partial o}\big|(\operatorname{supp}\psi) \setminus (\operatorname{supp}\mathbf{h}_{oj})\big|\right].$$

If $\mathbf{h} \notin \mathcal{H}_{\mathrm{loc}}^{\circ}[\psi]$, then this is positive, completing the proof of our claim. □

Our main result in this section is the following infinite-tree analogue of [46], Theorem 2, characterizing the interior stationary points of Φ_μ as fixed points of the Bethe recursion.

PROPOSITION 3.3. *For $\underline{\psi}$ permissive, any stationary point of Φ_μ inside $\mathcal{H}_{\mathrm{loc}}^{\circ}[\psi]$ belongs to $\mathcal{H}^\star$.*

PROOF. Let $\mathcal{H}_{\mathrm{loc}}^{\pm}[\psi]$ denote the space of measurable functions $\boldsymbol{\delta}:\mathcal{T}_{\mathrm{e}} \to \mathbb{R}^{\mathscr{X}^2}$ (defined up to $\mu^{\uparrow}$-equivalence) such that $\operatorname{supp}\boldsymbol{\delta}_{xy} \subseteq \operatorname{supp}\psi$, $\boldsymbol{\delta}_{xy}(\sigma, \sigma') = \boldsymbol{\delta}_{yx}(\sigma', \sigma)$, the one-point marginals $\bar{\delta}_x(\sigma) \equiv \sum_{\sigma'}\boldsymbol{\delta}_{xy}(\sigma, \sigma')$ do not depend on the choice of $y \in \partial x$, and $\sum_\sigma \bar{\delta}(\sigma) = \sum_{\sigma,\sigma'}\boldsymbol{\delta}(\sigma, \sigma') \equiv 0$.

Step 1. We first show that if $\mathbf{h} \in \mathcal{H}_{\mathrm{loc}}^{\circ}[\psi]$ is a stationary point of Φ_μ, then there exists $\lambda:\mathcal{T}_{\mathrm{e}} \to \mathbb{R}^{\mathscr{X}}$ measurable such that

$$\mathbf{h}_{xy}(\sigma, \sigma') = \psi(\sigma, \sigma')\exp\{\lambda_{x\to y}(\sigma) + \lambda_{y\to x}(\sigma')\}, \qquad \mu^{\uparrow}\text{-a.s.} \tag{3.4}$$

Since $\mathbf{h} \in \mathcal{H}_{\mathrm{loc}}^{\circ}[\psi]$, if $\boldsymbol{\delta} \in \mathcal{H}_{\mathrm{loc}}^{\pm}[\psi]$ with $|\boldsymbol{\delta}| \leq \mathbf{h}$ $\mu^{\uparrow}$-a.s., then $\mathbf{h}^\eta \equiv \mathbf{h} + \eta\delta$ belongs to $\mathcal{H}_{\mathrm{loc}}[\psi]$ for all $|\eta| \leq 1$. Taking $\eta \to 0$ in (3.1) gives (by stationarity of Φ_μ at $\mathbf{h}$)

$$0 = R^0(\boldsymbol{\delta}) = \mathbb{E}_\mu\left[2\langle\bar{\kappa}'_o\rangle_{\bar{\delta}_o} + \sum_{j\in\partial o}\langle\boldsymbol{\kappa}'\rangle_{\boldsymbol{\delta}}\right],$$

where $\bar{\kappa}'_x \equiv \bar{\xi} + (D_x - 1)\log\bar{h}_x$, $\boldsymbol{\kappa}'_{xy} \equiv (\xi - \log\mathbf{h}_{xy})\mathbf{1}_{\operatorname{supp}\psi}$.

Consider now $\boldsymbol{\delta}$ with one-point marginals $\bar{\delta} \equiv 0$, so that the value of $\bar{\kappa}'$ becomes irrelevant: in this case the value of $R^0(\boldsymbol{\delta})$ is unchanged upon replacing $\boldsymbol{\kappa}'$ by

$$\boldsymbol{\kappa}_{xy}(\sigma, \sigma') \equiv \mathbf{1}_{\operatorname{supp} \psi}(\sigma, \sigma')[\boldsymbol{\kappa}'_{xy}(\sigma, \sigma') + \lambda_{x\to y}(\sigma) + \lambda_{y\to x}(\sigma')].$$

We claim it is possible to choose λ such that $\boldsymbol{\kappa}$ has one-point marginals $\bar{\kappa} \equiv 0$, $\mu^{\uparrow}$-a.s. This amounts to solving the linear system

$$\begin{pmatrix} a_{x\to y} \\ a_{y\to x} \end{pmatrix} = \begin{pmatrix} I & Q \\ Q & I \end{pmatrix} \begin{pmatrix} \lambda_{x\to y} \\ \lambda_{y\to x} \end{pmatrix} \equiv \mathbf{Q} \begin{pmatrix} \lambda_{x\to y} \\ \lambda_{y\to x} \end{pmatrix}, \tag{3.5}$$

where, writing $r(\sigma) \equiv |\{\sigma' : \psi(\sigma, \sigma') > 0\}|$,

$$a_{x\to y}(\sigma) \equiv -\frac{\sum_{\sigma'} \boldsymbol{\kappa}'_{xy}(\sigma, \sigma')}{r(\sigma)}, \qquad Q(\sigma, \sigma') \equiv \frac{\mathbf{1}_{\operatorname{supp} \psi}(\sigma, \sigma')}{r(\sigma)}.$$

For $\underline{\psi}$ permissive, the Markov kernel Q is irreducible and aperiodic, with stationary distribution $\underline{r} \equiv (r(\sigma))_\sigma$ (by symmetry of ψ). By the Perron–Frobenius theorem, Q, Q^2 both have unique left eigenvector $\underline{r}$ corresponding to eigenvalue 1. Therefore $\dim \ker(I - Q^2) = 1$, from which it is easy to see that $\ker \mathbf{Q}^t = (\operatorname{im} \mathbf{Q})^{\perp}$ is the linear span of $(\underline{r}, -\underline{r})$. Since the assumed symmetry properties of ψ and $\mathbf{h}$ imply that

$$\langle (\underline{r}, -\underline{r}), (a_{x\to y}, a_{y\to x}) \rangle = \sum_{\sigma,\sigma'} (-\boldsymbol{\kappa}'_{xy}(\sigma, \sigma') + \boldsymbol{\kappa}'_{yx}(\sigma, \sigma')) = 0, \qquad \mu^{\uparrow}\text{-a.s.},$$

there is a unique solution $(\lambda_{x\to y}, \lambda_{y\to x})$ to the system (3.5) giving the required solution to (3.4).

For this choice of $\boldsymbol{\kappa}$, $\boldsymbol{\delta} = c\boldsymbol{\kappa}$ belongs to $\mathcal{H}^{\pm}_{\mathrm{loc}}[\psi]$ for any measurable $c : \mathcal{T}_{\mathrm{e}} \to \mathbb{R}_{>0}$ with $c_{xy} = c_{yx}$. We can choose c small enough so that $|\boldsymbol{\delta}| < |\mathbf{h}|$ on $\operatorname{supp} \psi$ $\mu^{\uparrow}$-a.s. With this choice, $0 = R^0(\boldsymbol{\delta})$ becomes the μ-expectation of a (weighted) sum of squares, so $\boldsymbol{\kappa} \equiv 0$, and rearranging gives (3.4).

Step 2. Returning now to general $\boldsymbol{\delta} \in \mathcal{H}^{\pm}_{\mathrm{loc}}[\psi]$ with $|\boldsymbol{\delta}| \le \mathbf{h}$ $\mu^{\uparrow}$-a.s., we obtain from (3.4) the simplification

$$\begin{aligned} 0 = R^0(\boldsymbol{\delta}) &= \mathbb{E}_\mu\Big[2\langle \bar{\kappa}'_o \rangle_{\bar{\delta}_o} - \sum_{j\in\partial o} (\langle \lambda_{o\to j} \rangle_{\bar{\delta}_o} + \langle \lambda_{j\to o} \rangle_{\bar{\delta}_j})\Big] \\ &= 2\mathbb{E}_\mu\Big[\Big\langle \bar{\kappa}'_o - \sum_{j\in\partial o} \lambda_{o\to j} \Big\rangle_{\bar{\delta}_o}\Big], \end{aligned} \tag{3.6}$$

using unimodularity of μ for the last identity. We claim that

$$\begin{aligned} \bar{\delta}'_x(\sigma) &\equiv \bar{\kappa}'_x(\sigma) - \sum_{y\in\partial x} \lambda_{x\to y}(\sigma) - \frac{1}{|\mathscr{X}|} \sum_{\sigma'} \Big(\bar{\kappa}'_x(\sigma') - \sum_{y\in\partial x} \lambda_{x\to y}(\sigma')\Big) \\ &= 0, \qquad \mu^{\uparrow}\text{-a.s.} \end{aligned} \tag{3.7}$$

Indeed, for any $\bar{\delta}' : \mathcal{T}_\bullet \to \mathbb{R}^{\mathscr{X}}$ measurable with $\sum_\sigma \bar{\delta}'_o(\sigma) \equiv 0$ μ-a.s.,

$$\boldsymbol{\delta}'_{xy}(\sigma, \sigma') \equiv \bar{\delta}'_x(\sigma)\mathbf{1}\{\sigma' = \sigma^{\mathrm{p}}\} + \bar{\delta}'_y(\sigma')\mathbf{1}\{\sigma = \sigma^{\mathrm{p}}\}$$

defines an element of $\mathcal{H}^{\pm}_{\mathrm{loc}}[\psi]$. By considering (3.6) with $\boldsymbol{\delta} = c\boldsymbol{\delta}'$ where $c_{xy} = c_{yx}$ is small enough so that $|c\boldsymbol{\delta}'| < |\mathbf{h}|$, we obtain the claim (3.7).

Step 3. Rearranging (3.7) we find that $\mathbf{h}$ satisfies $\mu^{\uparrow}$-a.s.

$$\mathbf{h}_{oj}(\sigma, \sigma') = \psi(\sigma, \sigma') \exp\{\lambda_{o\to j}(\sigma) + \lambda_{j\to o}(\sigma')\}, \tag{3.8}$$

$$\bar{h}_o(\sigma) \cong \exp\left\{\frac{\sum_{j\in\partial o} \lambda_{o\to j}(\sigma) - \bar{\xi}(\sigma)}{D_o - 1}\right\}. \tag{3.9}$$

If we then re-parametrize

$$\lambda_{o\to j} \equiv \bar{\xi} + \sum_{k\in\partial o\setminus j} \log \widehat{m}_{k\to o}, \qquad \mu^{\uparrow}\text{-a.s.} \tag{3.10}$$

(well defined, for each T and $\sigma \in \mathscr{X}$, by invertibility of the D_o-dimensional matrix $\mathbf{1}\mathbf{1}^t - I$), then formula (3.9) for $\bar{h}_o$ becomes

$$\bar{h}_o(\sigma) \cong \bar{\psi}(\sigma) \prod_{k\in\partial o} \widehat{m}_{k\to o}(\sigma), \qquad \mu^{\uparrow}\text{-a.s.}$$

On the other hand, $\bar{h}_o$ is the first marginal of $\mathbf{h}_{oj}$, and setting the above equal to the sum of (3.8) over σ' gives [making use of (3.10)]

$$\widehat{m}_{j\to o}(\sigma) \cong \sum_{\sigma'} \psi(\sigma, \sigma') e^{\lambda_{j\to o}(\sigma')}, \qquad \mu^{\uparrow}\text{-a.s.}$$

Thus, if we define $m : \mathcal{T}_{\mathrm{e}} \to \Delta_{\mathscr{X}}$, $m_{x\to y}(\sigma) \cong e^{\lambda_{x\to y}(\sigma)}$, then (3.10) can be written in terms of m as

$$m_{o\to j}(\sigma) \cong \bar{\psi}(\sigma) \prod_{k\in\partial o\setminus j} \left(\sum_{\sigma_k} \psi(\sigma, \sigma_k) m_{j\to o}(\sigma_k)\right), \qquad \mu^{\uparrow}\text{-a.s.},$$

that is, $m \in \mathcal{H}^\star$. Then (3.8) is precisely the statement that m maps to $\mathbf{h}$ via (1.22), which completes the proof. $\square$

PROOF OF THEOREM 1.18. By (H1) the set $\mathcal{H}^{\mathrm{fin}}_{\mathrm{loc}}$ of $\mathbf{h} \in \mathcal{H}_{\mathrm{loc}}$ for which $\Phi(\beta, B, \mathbf{h}) > -\infty$ is nonempty and does not depend on (β, B), so without loss we will restrict to $\mathbf{h} \in \mathcal{H}^{\mathrm{fin}}_{\mathrm{loc}}$.

Again by (H1), the functions $(\beta, B) \mapsto \Phi_\mu(\beta, B, \mathbf{h})$ indexed by $\mathbf{h} \in \mathcal{H}^{\mathrm{fin}}_{\mathrm{loc}}$ are uniformly equicontinuous on compact regions of (β, B): for any $\varepsilon > 0$ there exists $\delta > 0$ sufficiently small so that if (β, B) and (β', B') are within distance δ, then $|\Phi_\mu(\beta, B, \mathbf{h}) - \Phi_\mu(\beta', B', \mathbf{h})| < \varepsilon$ for all $\mathbf{h} \in \mathcal{H}^{\mathrm{fin}}_{\mathrm{loc}}$. Let $\mathbf{h} \in \mathcal{H}^{\mathrm{fin}}_{\mathrm{loc}}$ such that $\Phi_\mu(\beta, B, \mathbf{h}) \ge \widetilde{\Phi}_\mu(\beta, B) - \varepsilon$. Then

$$\widetilde{\Phi}_\mu(\beta', B') \ge \Phi_\mu(\beta', B', \mathbf{h}) \ge \widetilde{\Phi}_\mu(\beta, B) - 2\varepsilon$$

for all (β', B') within distance δ of (β, B). Reversing the roles of (β, B) and (β', B') completes the proof of part (a). The statement of part (b) is a summary of the results of Lemma 3.1, Propositions 3.2 and 3.3. $\square$

We supplement Proposition 3.3 by computing the second derivatives $\partial_\eta^2 \Phi_\mu(\mathbf{h} + \eta\boldsymbol{\delta})$ at interior stationary points $\mathbf{h}$, giving a criterion to verify that such points are local maximizers.

PROPOSITION 3.4. *For permissive $\underline{\psi}$, let $\mathbf{h} \in \mathcal{H}^\circ_{\text{loc}}[\psi]$ be a stationary point of Φ_μ, and let $\boldsymbol{\delta} \in \mathcal{H}^\pm_{\text{loc}}[\psi]$ with $|\delta| \le |\mathbf{h}|$. Then $\mathbf{h}$ is a local maximizer of Φ on the one-dimensional space $\mathcal{H}_{\text{loc}} \cap \{\mathbf{h} + \eta\boldsymbol{\delta} : \eta \in \mathbb{R}\}$ if and only if*

$$
\begin{aligned}
4\partial_\eta^2 \Phi_\mu(\mathbf{h} + \eta\boldsymbol{\delta})|_{\eta=0} &= \mathbb{E}_\mu\Big[2(D_o - 1)\big\langle(\bar{\delta}_o/\bar{h}_o)^2\big\rangle_{\bar{h}_o} - \sum_{j\in\partial o}\big\langle(\boldsymbol{\delta}_{oj}/\mathbf{h}_{oj})^2\big\rangle_{\mathbf{h}_{oj}}\Big] \\
&\le 0,
\end{aligned} \tag{3.11}
$$

or equivalently

$$
\begin{aligned}
&\mathbb{E}_\mu\big[\big\langle(\bar{\delta}_o/\bar{h}_o)^2\big\rangle_{\bar{h}_o}\big] \\
&\quad \ge \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{j\in\partial o}\big(\big\langle(\bar{\delta}_o/\bar{h}_o)^2\big\rangle_{\bar{h}_o} + \big\langle(\bar{\delta}_h/\bar{h}_h)^2\big\rangle_{\bar{h}_h} - \big\langle(\boldsymbol{\delta}_{oj}/\mathbf{h}_{oj})^2\big\rangle_{\mathbf{h}_{oj}}\big)\Big].
\end{aligned} \tag{3.12}
$$

It is a strict local maximizer if (3.11) *and* (3.12) *hold with strict inequality.*

PROOF. For $\mathbf{h} \in \mathcal{H}^\circ_{\text{loc}}[\psi]$ and $\boldsymbol{\delta} \in \mathcal{H}^\pm_{\text{loc}}[\psi]$ with $|\boldsymbol{\delta}| \le |\mathbf{h}|$, arguing as in the proof of Proposition 3.3 gives

$$
\begin{aligned}
&2\partial_\eta \Phi_\mu(\mathbf{h} + \eta\boldsymbol{\delta})|_{\eta=0} \\
&\quad = \lim_{\eta\to 0} R^\eta(\boldsymbol{\delta}) = R^0(\delta) \\
&\quad \equiv \mathbb{E}_\mu\Big[2\langle\bar{\xi}\rangle_{\bar{\delta}_o} + 2(D_o - 1)\langle\log\bar{h}_o\rangle_{\bar{\delta}_o} + \sum_{j\in\partial o}\big(\langle\xi\rangle_{\boldsymbol{\delta}_{oj}} - \langle\log\mathbf{h}_{oj}\rangle_{\boldsymbol{\delta}_{oj}}\big)\Big].
\end{aligned}
$$

If $\mathbf{h}$ is further a stationary point of Φ_μ, then, for $\eta < 1$,

$$
\begin{aligned}
T^\eta(\boldsymbol{\delta}) \equiv \frac{2}{\eta}R^\eta(\boldsymbol{\delta}) &= \frac{2}{\eta}\big[R^\eta(\boldsymbol{\delta}) - R^0(\boldsymbol{\delta})\big] \\
&= 2\mathbb{E}_\mu\Big[2(D_o - 1)\big\langle f^\eta(\bar{\delta}_o/\bar{h}_o)\big\rangle_{\bar{\delta}_o} - \sum_{j\in\partial o}\big\langle f^\eta(\boldsymbol{\delta}_{oj}/\mathbf{h}_{oj})\big\rangle_{\boldsymbol{\delta}_{oj}} \\
&\qquad + 2(D_o - 1)\big\langle g^\eta(\bar{\delta}_o/\bar{h}_o)\big\rangle_{\bar{h}_o} - \sum_{j\in\partial o}\big\langle g^\eta(\boldsymbol{\delta}_{oj}/\mathbf{h}_{oj})\big\rangle_{\mathbf{h}_{oj}}\Big],
\end{aligned}
$$

where $g^\eta(r) \equiv [f^\eta(r) - r]/\eta$, with $\lim_{\eta\to 0} g^\eta(r) = -r^2/2$. Since $|\boldsymbol{\delta}/\mathbf{h}| \le 1$, it follows by dominated convergence that

$$4\partial_\eta^2 \Phi_\mu(\mathbf{h} + \eta\boldsymbol{\delta})|_{\eta=0} = \lim_{\eta\to 0} T^\eta(\boldsymbol{\delta}) = T^0(\boldsymbol{\delta})$$
$$\equiv \mathbb{E}_\mu\Big[2(D_o - 1)\langle(\bar{\delta}_o/\bar{h}_o)^2\rangle_{\bar{h}_o} - \sum_{j\in\partial o}\langle(\boldsymbol{\delta}_{oj}/\mathbf{h}_{oj})^2\rangle_{\mathbf{h}_{oj}}\Big].$$

The stationary point $\mathbf{h}$ is a local maximizer on $\mathcal{H}_{\text{loc}} \cap \{\mathbf{h} + \eta\boldsymbol{\delta} : \eta \in \mathbb{R}\}$ if and only if $\partial_\eta^2 \Phi_\mu(\mathbf{h} + \eta\boldsymbol{\delta})|_{\eta=0} \le 0$, which gives (3.11). Condition (3.12) is equivalent by an application of unimodularity. □

4. Application to Ising and Potts models. In this section we apply Theorem 1.15 to prove our results for the ferromagnetic Ising and Potts models, Theorems 1.9–1.11. Although both models have regimes of multiple fixed points, monotonicity arguments allow us to restrict the space of fixed points. In the Ising model we can restrict to a unique fixed point and give a complete verification of the Bethe free energy prediction; in the Potts model with $q > 2$ there remain regimes of nonuniqueness where we can only provide bounds.

4.1. *Ising model.* We first prove Theorem 1.9. Recall definition (1.14) for the Ising measure $\nu_G^{\beta,B}$ for a finite graph $G = (V, E)$, and more generally (from Definition 1.8) the Ising measures $\nu_{U,G}^{\mathrm{f},\beta,B}$ and $\nu_{U,G}^{+,\beta,B}$ for a finite sub-graph U of a (possibly infinite) graph G with free and $+$ boundary conditions. We will make use of the following direct consequence of the classical Griffiths's inequality; see, for example, [31], Theorem IV.1.21.

LEMMA 4.1. *For the Ising model with parameters $\beta, B \ge 0$ on U a finite sub-graph of a graph G with boundary conditions $\ddagger \in \{\mathrm{f}, +\}$, the* magnetization $\langle\sigma_v\rangle_{U,G}^{\ddagger,\beta,B}$ *at vertex $v \in U$ is nonnegative, nondecreasing in β, B, nondecreasing in U for $\ddagger = \mathrm{f}$ and nonincreasing in U for $\ddagger = +$.*

Recall from Section 1.2.1 the definitions of $\bar{h}_T^{t,\ddagger}$ for $\ddagger \in \{\mathrm{f}, +\}$; the measure $\bar{h}_T^{t,\ddagger}$ is parametrized by the corresponding magnetization $\bar{m}_T^{t,\ddagger} \equiv \bar{h}_T^{t,\ddagger}(+) - \bar{h}_T^{t,\ddagger}(-)$. By Lemma 4.1, $\bar{m}_T^{t,\mathrm{f}}$ is nondecreasing in t while $\bar{m}_T^{t,+}$ is nonincreasing, so there exist well-defined limits $\bar{m}_T^{\ddagger}(\beta, B) \equiv \lim_{t\to\infty} \bar{m}_T^{t,\ddagger}(\beta, B)$. The following result from [13], an extension of [10], Lemma 4.3, shows that these limits agree on *any* $T \in \mathcal{T}_\bullet$.

LEMMA 4.2 ([13], Lemma 3.1). *For the Ising model (1.14) on an infinite tree T with $\beta, B > 0$, there exists a constant $C \equiv C(\beta, B)$ such that*

$$\bar{m}_T^{t,+} - \bar{m}_T^{t,\mathrm{f}} \le C/t \qquad \forall t \ge 1.$$

By this result we can define $h \in \mathcal{H}$ by $h_{x\to y} = \bar{h}^{\mathrm{f}}_{T_{x\to y}} = \bar{h}^{+}_{T_{x\to y}}$, and we now proceed to verify the Bethe prediction $\phi(\beta, B) = \Phi_\mu(\beta, B, h)$.

PROOF OF THEOREM 1.9. The Ising model (1.14) is of form (1.1) with $\mathscr{X} = \{\pm 1\}$, $\xi(\sigma, \sigma') = \beta\sigma\sigma'$ and $\bar{\xi}(\sigma) = B\sigma$, so (H1) and (H2) are clearly satisfied (with no additional moment conditions on D_o, since $\psi > 0$). It follows directly from the recursive structure of the tree that $h \in \mathcal{H}^\star$. It will be shown in Lemma 4.5 that for $\beta \geq 0$ fixed,

$$\lim_{B\to\infty} \limsup_{n\to\infty} \left|\phi_n(\beta, B) - \Phi_\mu(\beta, B, h)\right| = 0,$$

so to prove the theorem we will interpolate from (β, B) to (β, B_1), then take $B_1 \to \infty$.

It follows from Lemmas 4.1 and 4.2 that for $T \in \mathcal{T}_\bullet$, $\bar{m}^{\mathrm{f}}_T(\beta, B) = \bar{m}^{+}_T(\beta, B) \equiv m_T(\beta, B)$ is the increasing limit of $\bar{m}^{t,\mathrm{f}}_T(\beta, B)$ and the decreasing limit of $\bar{m}^{t,+}_T(\beta, B)$. The $\bar{m}^{t,\ddagger}(\beta, B)$ are continuous and nondecreasing in β, B, so m inherits these properties by the same argument as in the proof of Theorem 1.12, and so (since it takes values in $[-1, 1]$) is of uniformly bounded total variation. This verifies both (H3$^\beta$) and (H3^B) (though we will use only the latter).

We conclude by showing [cf. (1.20)] that

$$\lim_{n\to\infty} \mathbb{E}_n\big[\langle \partial\bar{\xi}(\sigma_{I_n})\rangle_n^{\beta,B}\big] = a^{\mathrm{vx}}(\beta, B) \equiv \mathbb{E}_\mu\big[[\![\partial_B\bar{\xi}(\sigma_o)]\!]_T^{h,\beta,B}\big].$$

Here $\partial_B\bar{\xi}(\sigma) = \sigma$, and it follows from Lemma 4.1, our assumption of $G_n \to_{lwc} \mu$ and Fatou's lemma that

$$\begin{aligned}
\mathbb{E}_\mu\big[[\![\sigma_o]\!]_T^{h^{\mathrm{f}},\beta,B}\big] &\leq \liminf_{t\to\infty} \mathbb{E}_\mu\big[\langle\sigma_o\rangle^{\mathrm{f},\beta,B}_{T^t,T}\big] \leq \liminf_{n\to\infty} \mathbb{E}_n\big[\langle\sigma_{I_n}\rangle_n^{\beta,B}\big] \\
&\leq \limsup_{n\to\infty} \mathbb{E}_n\big[\langle\sigma_{I_n}\rangle_n^{\beta,B}\big] \leq \limsup_{t\to\infty} \mathbb{E}_\mu\big[\langle\sigma_o\rangle^{+,\beta,B}_{T^t,T}\big] \\
&\leq \mathbb{E}_\mu\big[[\![\sigma_o]\!]_T^{h^{+},\beta,B}\big].
\end{aligned}$$

The left-most and right-most expressions coincide by Lemma 4.2 so equality holds throughout.

By Theorem 1.15(b), $\phi(\beta, B) = \Phi(\beta, B, h^+) = \Phi(\beta, B, h^{\mathrm{f}})$ for $\beta \geq 0$, $B > 0$. Since ϕ_n is symmetric in B and continuous at $B = 0$ (uniformly in n), we have $\phi(\beta, B) = \phi(\beta, -B)$ and $\phi(\beta, 0) = \lim_{B\to 0}\phi(\beta, B)$. $\square$

4.2. *Potts model.* We now apply Theorem 1.15 to deduce our result (Theorem 1.10) for the Potts model (1.3) with $\beta, B \geq 0$. From now on we let $\mathscr{X} \equiv [q]$ with $q \geq 2$. It will be convenient to generalize (1.3) to the *inhomogeneous* Potts model

$$\nu_G^{\underline{\beta},\underline{B}}(\underline{\sigma}) \cong \exp\left\{\sum_{(ij)\in E} \beta_{ij} \cdot \mathbf{1}\{\sigma_i = \sigma_j\} + \sum_{i\in V} B_i \cdot \mathbf{1}\{\sigma_i = 1\}\right\}, \qquad \underline{\sigma} \in \mathscr{X}^V.$$

We now introduce the coupling of the Potts model with a random-cluster model which we use to obtain monotonicity properties. The following representation is as in [23]; see also [7]. If $G = (V, E)$ is a finite graph, let $G^\star$ be the graph formed by adding an edge from every $v \in V$ to a "ghost vertex" $v^\star$, that is, $G^\star = (V^\star, E^\star)$ where $V^\star = V \cup \{v^\star\}$ and $E^\star = E \cup \{(v, v^\star) : v \in V\}$. Writing $\underline{\sigma}$ for elements of $\mathscr{X}^{V^\star}$ and $\underline{\eta}$ for elements of $\{0,1\}^{E^\star}$ (*bond configurations*), consider the probability measure on pairs $(\underline{\sigma}, \underline{\eta})$ defined by

$$\begin{aligned}\varpi_G^{\underline{\beta},\underline{B}}(\underline{\sigma}, \underline{\eta}) \\ \cong \mathbf{1}\{\sigma_{v^\star} = 1\} \prod_{\eta_{ij}=1} \{(e^{\beta_{ij}\cdot \mathbf{1}\{\sigma_i=\sigma_j\}} - 1)\} \prod_{\eta_i=1} \{(e^{B_i \cdot \mathbf{1}\{\sigma_i = \sigma_{v^\star}\}} - 1)\}.\end{aligned} \tag{4.1}$$

The marginal on $\underline{\sigma}_V$ is the inhomogeneous Potts measure $\nu_G^{\underline{\beta},\underline{B}}$, while the marginal on $\underline{\eta}$ is the (*inhomogeneous*) *random-cluster measure*

$$\pi_G^{\underline{\beta},\underline{B}}(\underline{\eta}) \cong \prod_{e\in E^\star} p_e^{\eta_e}(1-p_e)^{1-\eta_e} \prod_{C\in\eta} \Theta(C), \tag{4.2}$$

where $p_{ij} \equiv 1 - e^{-\beta_{ij}}$ for $(i, j) \in E$ and $p_{iv^\star} \equiv 1 - e^{-B_i}$ for $i \in V$, and the last product is taken over connected components C of $\underline{\eta}$, with $\Theta(C) = q$ unless $v^\star \in C$ in which case $\Theta(C) = 1$. Given a configuration $\underline{\eta}$, a realization of the conditional law $\varpi_G^{\beta,B}(\underline{\sigma} = \cdot|\underline{\eta})$ is obtained by choosing a constant spin on each connected component C of $\underline{\eta}$ independently and uniformly over $[q]$, except for C containing $v^\star$ which is given spin 1.

For a detailed account the random-cluster model, see [24]; we will use only the following basic properties:

PROPOSITION 4.3. *The random-cluster measure $\pi_G^{\underline{\beta},\underline{B}}$ is FKG. It is also increasing, in the sense of stochastic domination, in $(\underline{\beta}, \underline{B})$.*

PROOF. The FKG property follows by a straightforward modification of the proof of [7], Theorem III.1(i). Monotonicity in $(\underline{\beta}, \underline{B})$ follows by modifying the proof of [24], Theorem 3.21. □

Recalling Definition 1.8, for U, a finite sub-graph of a graph G and $\ddagger \in \{\mathrm{f}\} \cup [q]$ (with f = free), let $\nu_{U,G}^{\ddagger,\beta,B}$ denote the Potts model on U with $\ddagger$ boundary conditions.

COROLLARY 4.4. *For the Potts model with parameters $\beta, B \geq 0$ on U a finite sub-graph of a graph G with boundary conditions $\ddagger \in \{\mathrm{f}, 1\}$, and for any vertices $v, w \in U$, the quantities*

$$\nu_{U,G}^{\ddagger,\beta,B}(\sigma_v = 1), \qquad \nu_{U,G}^{\ddagger,\beta,B}(\sigma_v = \sigma_w)$$

are nondecreasing in β and B, nonincreasing in U for $\ddagger = 1$ and nondecreasing in U for $\ddagger = \mathrm{f}$.

PROOF. Note that $\nu_{U,G}^{\mathrm{f},\beta,B}$ is the marginal on $\underline{\sigma}_U$ of the measure $\varpi_G^{\underline{\beta},\underline{B}}$ with

$$B_i = B \qquad \forall i \in V, \qquad \beta_e = \beta \cdot \mathbf{1}\{e \in E_U\}.$$

Similarly, $\nu_{U,G}^{1,\beta,B}$ is the marginal on $\underline{\sigma}_U$ of the measure $\varpi_G^{\underline{\beta}',\underline{B}'}$ with

$$\beta'_e = \beta \qquad \forall e \in E, \qquad B'_i = B \cdot \mathbf{1}\{i \in V_U\} + \infty \cdot \mathbf{1}\{i \notin V_U\}.$$

Clearly, $(\underline{\beta}, \underline{B})$ is nondecreasing in U while $(\underline{\beta}', \underline{B}')$ is nonincreasing, and both are nondecreasing in β, B. The result therefore follows from Proposition 4.3 by showing that for any $(\underline{\beta}, \underline{B})$, the conditional probabilities $\varpi_G^{\underline{\beta},\underline{B}}(\sigma_v = 1|\underline{\eta})$ and $\varpi_G^{\underline{\beta},\underline{B}}(\sigma_v = \sigma_w|\underline{\eta})$ are monotone functions of $\underline{\eta}$. Indeed, letting $\varpi \equiv \varpi_G^{\underline{\beta},\underline{B}}$ and writing $v \leftrightsquigarrow w$ to indicate that v, w belong to the same connected component of $\underline{\eta}$, we have

$$\varpi(\sigma_v = 1|\underline{\eta}) = \mathbf{1}\{v \leftrightsquigarrow v^\star\} + \frac{1 - \mathbf{1}\{v \leftrightsquigarrow v^\star\}}{q},$$
$$\varpi(\sigma_v = \sigma_v|\underline{\eta}) = \mathbf{1}\{v \leftrightsquigarrow w\} + \frac{1 - \mathbf{1}\{v \leftrightsquigarrow w\}}{q}.$$

These are increasing functions of $\underline{\eta}$ so the proof is complete. □

Under the measures with $\ddagger \in \{\mathrm{f}, 1\}$, any one-vertex marginal must be uniform on the spins $\neq 1$, and so is characterized by the probability given to spin 1. In particular, recall from Section 1.2.2 the definitions of $\bar{h}_T^{t,\ddagger}$ for $\ddagger \in \{\mathrm{f}, 1\}$; existence of the $t \to \infty$ limits $\bar{h}_T^{\ddagger}$ is now justified by Corollary 4.4, so we can define $h^\ddagger \in \mathcal{H}$ by $h^\ddagger_{x\to y} = \bar{h}^\ddagger_{T_{x\to y}}$. The following lemma gives the boundary values for the interpolation in (β, B) using $h^\ddagger$:

LEMMA 4.5. *For the Potts model on $G_n \to_{lwc} \mu$, let*

$$\widetilde{\Phi}_\mu(\beta, B) \equiv B + \beta\mathbb{E}_\mu[D_o]/2 + \mathbb{E}_\mu\big[\bar{\varphi}(|T|)\big],$$
$$\bar{\varphi}(n) \equiv \bar{\varphi}^B(n) \equiv n^{-1}\log\big(1 + (q-1)e^{-Bn}\big).$$

(a) *For all $B \in \mathbb{R}$ and any $h \in \mathcal{H}$, $\phi(0, B) = \log(e^B + q - 1) = \Phi_\mu(0, B, h)$.*
(b) *For $\beta \geq 0$ and $h \in \mathcal{H}^\star$,*

$$\lim_{B\to\infty} \limsup_{n\to\infty} \big|\phi_n(\beta, B) - \widetilde{\Phi}_\mu(\beta, B)\big| = 0 = \lim_{B\to\infty} z\big|\Phi_\mu(\beta, B, h) - \widetilde{\Phi}_\mu(\beta, B)\big|.$$

(c) *For $B \geq 0$, $\lim_{\beta\to\infty} \limsup_{n\to\infty} |\phi_n(\beta, B) - \widetilde{\Phi}_\mu(\beta, B)| = 0$.*
(d) *For $B > 0$ and $\ddagger \in \{\mathrm{f}, 1\}$, $\lim_{\beta\to\infty} |\Phi_\mu(\beta, B, h^\ddagger) - \widetilde{\Phi}_\mu(\beta, B)| = 0$.*

PROOF. (a) At $\beta = 0$, $\psi \equiv 1$ so the spins are independent. Thus, for all $n \geq 1$, $h \in \mathcal{H}$ and $T \in \mathcal{T}_\bullet$,

$$\phi_n(0, B) = \log\bigl(e^B + q - 1\bigr) = \Phi_T^{\mathrm{vx}}(0, B, h) = \Phi_T(0, B, h),$$

since $\Phi_T^{(oj)} \equiv 0$ for all $j \in \partial o$.

(b) The value of $Z_n(\beta, B)$ is bounded below by considering only the ground state $\underline{\sigma} \equiv 1$, and bounded above by decomposing $\mathscr{X}^V$ according to the subset of k vertices where the spin is not 1. For $\beta \geq 0$ this gives

$$1 \leq Z_n(\beta, B) e^{-Bn - \beta|E_n|} \leq \sum_{k=0}^{n} \binom{n}{k} (q-1)^k e^{-Bk} = \bigl(1 + (q-1)e^{-B}\bigr)^n,$$

so if we define $\bar{\phi}_n(\beta, B) \equiv \phi_n(\beta, B) - B - \beta \mathbb{E}_n[|E_n|]/n$, then $\lim_{B\to\infty} \limsup_{n\to\infty} |\bar{\phi}_n(\beta, B)| = 0$. Recalling (2.2), this proves the left identity in (b).

We next define

$$\bar{\Phi}_T^{\mathrm{vx}} \equiv \Phi_T^{\mathrm{vx}} - B - \beta D_o, \qquad \bar{\Phi}_T^{\mathrm{e}} \equiv \Phi_T^{\mathrm{e}} - \beta D_o/2, \qquad \bar{\Phi}_T \equiv \bar{\Phi}_T^{\mathrm{vx}} - \bar{\Phi}_T^{\mathrm{e}},$$
$$\bar{\Phi}_\mu \equiv \mathbb{E}_\mu \bar{\Phi}_T,$$

so that to prove the right identity in (b) it suffices to show $\lim_{B\to\infty} \bar{\Phi}_\mu(\beta, B, h) = 0$ for any $h \in \mathcal{H}^\star$. Indeed, (1.11) gives that μ-a.s., $\lim_{B\to\infty} h_{o\to j}^{\beta,B}(\sigma) = \mathbf{1}\{\sigma = 1\}$ for all $j \in \partial o$, hence also $\lim_{B\to\infty} h_{j\to o}^{\beta,B}(\sigma) = \mathbf{1}\{\sigma = 1\}$ for all $j \in \partial o$ by equivalence of $\mu^\uparrow$ and $\mu^\downarrow$. Thus

$$\lim_{B\to\infty} \bar{\Phi}_T^{\mathrm{vx}}(\beta, B, h) = 0 = \lim_{B\to\infty} \bar{\Phi}_T^{\mathrm{e}}(\beta, B, h), \qquad \mu\text{-a.s.}$$

It is easily verified that

$$-\beta D_o \leq \bar{\Phi}_T^{\mathrm{vx}}(\beta, B, h) \leq \log q, \qquad -\beta D_o/2 \leq \bar{\Phi}_T^{\mathrm{e}}(\beta, B, h) \leq 0, \tag{4.3}$$

so $\bar{\Phi}_\mu(\beta, B, h) \to 0$ by dominated convergence.

(c) Suppose first that G_n is connected. Then $Z_n(\beta, B)$ is bounded below by considering only the q constant-spin configurations, and bounded above by decomposing $\mathscr{X}^V$ according to the subset of ℓ edges across which the spins disagree. Since G_n is connected, removing ℓ edges leaves at most $\ell + 1$ connected components, of sizes $k_0, \ldots, k_\ell$ summing to n. Therefore, with $\varphi(n) \equiv \varphi^B(n) \equiv n\bar{\varphi}^B(n)$, we have

$$e^{\varphi(n)} \leq Z_n(\beta, B) e^{-Bn - \beta|E_n|} \leq \sum_{\ell=0}^{|E_n|} \binom{|E_n|}{\ell} e^{-\beta\ell} \max_{k_0,\ldots,k_\ell} \left\{\exp\left\{\sum_{r=0}^{\ell} \varphi(k_r)\right\}\right\},$$

where the maximum is taken over $k_0, \ldots, k_\ell \in \mathbb{Z}_{\geq 0}$ summing to n. By convexity of φ this maximum is achieved with $k_r = n$ for some r, so

$$\begin{aligned} \varphi(n) \leq n\bar{\phi}_n(\beta, B) &\leq \varphi(n) + \mathbb{E}_n\left[\log\left\{\sum_{\ell=0}^{|E_n|} \binom{|E_n|}{\ell} e^{-\beta\ell} q^\ell\right\}\right] \\ &= \varphi(n) + \mathbb{E}_n\bigl[|E_n|\bigr] \log\bigl(1 + qe^{-\beta}\bigr). \end{aligned} \tag{4.4}$$

If G_n has connected components $C^j = (V^j, E^j)$, $j \ge 1$, with $|V^j| = n^j$, then clearly $Z_n(\beta, B) = \prod_j Z_{C^j}(\beta, B)$, so

$$0 \le \bar{\phi}_n(\beta, B) - \frac{1}{n}\mathbb{E}_n\biggl[\sum_j \varphi(n^j)\biggr] \le \frac{1}{n}\mathbb{E}_n[|E_n|]\log(1 + qe^{-\beta}). \tag{4.5}$$

With $j(i)$ denoting the index of the connected component of G_n containing vertex i, we have $n^{-1}\mathbb{E}_n[\sum_j \varphi(n^j)] = \mathbb{E}_n[\bar{\varphi}(n^{j(I_n)})]$. Then, since $\bar{\varphi}'(n) \le 0$,

$$\mathbb{E}_n[\bar{\varphi}(|B_t(I_n)|) \cdot \mathbf{1}\{B_t(I_n) = C^{j(I_n)}\}] \le \mathbb{E}_n[\bar{\varphi}(n^{j(I_n)})] \le \mathbb{E}_n[\bar{\varphi}(|B_t(I_n)|)].$$

Since $G_n \to_{lwc} \mu$, letting $n \to \infty$ followed by $t \to \infty$ in the above inequalities gives $\mathbb{E}_n[\bar{\varphi}(n^{j(I_n)})] \to \mathbb{E}_\mu[\bar{\varphi}(|T|)]$, and so (c) follows from (4.5) by taking first $n \to \infty$ and then $\beta \to \infty$.

(d) Clearly $h_T^{\mathrm{f}} = h_T^1$ for any finite $T \in \mathcal{T}_\bullet$ (as $\partial T^t = \varnothing$ for large enough t). In the $\beta \to \infty$ limit only the constant-spin configurations contribute, so

$$\lim_{\beta\to\infty} h_T^{\ddagger,\beta,B}(\sigma) = e^{-\varphi(|T|) - B|T|(1-\mathbf{1}\{\sigma=1\})}, \qquad \ddagger \in \{\mathrm{f}, 1\}. \tag{4.6}$$

For T infinite, recall from Corollary 4.4 that $h_{T^t}^{\mathrm{f}}(1) \le h_T^{\mathrm{f}}(1) \le h_T^1(1)$, so if $B > 0$, then

$$1 = \lim_{t\to\infty}\lim_{\beta\to\infty} h^{t,\mathrm{f},\beta,B}(1) \le \lim_{\beta\to\infty} h^{\mathrm{f},\beta,B}(1) \le \lim_{\beta\to\infty} h^{1,\beta,B}(1),$$

so that (4.6) again holds for T infinite. We then compute

$$\lim_{\beta\to\infty} \bar{\Phi}_T^{\mathrm{vx}}(\beta, B, h^\ddagger) = -\sum_{j\in\partial o} \varphi(|T_{j\to o}|) + \varphi(|T|),$$

$$\lim_{\beta\to\infty} \bar{\Phi}_T^{\mathrm{e}}(\beta, B, h^\ddagger) = -\frac{1}{2}\sum_{j\in\partial o} \varphi(|T_{j\to o}|) - \frac{1}{2}\sum_{j\in\partial o}\varphi(|T_{o\to j}|) + \frac{D_o}{2}\varphi(|T|),$$

μ-a.s., where the first identity uses $|T| = 1 + \sum_{j\in\partial o}|T_{j\to o}|$ and the second uses $|T| = |T_{o\to j}| + |T_{j\to o}|$. Convergence also holds in μ-expectation, using the upper bounds in (4.3) together with

$$\bar{\Phi}_T^{\mathrm{vx}}(\beta, B, h^\ddagger) \ge \sum_{j\in\partial o} \log h_{j\to o}^{\mathrm{f},\beta,B}(1),$$

$$\bar{\Phi}_T^{\mathrm{e}}(\beta, B, h^\ddagger) \ge \frac{1}{2}\sum_{j\in\partial o} \log h_{j\to o}^{\mathrm{f},\beta,B}(1) + \frac{1}{2}\sum_{j\in\partial o}\log h_{o\to j}^{\mathrm{f},\beta,B}(1),$$

and the fact that $h_{x\to y}^{\mathrm{f},\beta,B}(1) \ge 1/q$ for $\beta, B \ge 0$ (by Corollary 4.4). Thus, using unimodularity of μ, we have

$$\lim_{\beta\to\infty} \bar{\Phi}_\mu(\beta, B, h^\ddagger) = \mathbb{E}_\mu[(1 - D_o/2)\varphi(|T|)],$$

and we conclude by showing that this coincides with $\mathbb{E}_\mu[\bar{\varphi}(|T|)]$. The case $|T| = \infty$ is trivial; otherwise, another application of unimodularity gives

$$\frac{1}{2}\mathbb{E}_\mu[D_o\varphi(|T|)] = \frac{1}{2}\mathbb{E}_\mu\Big[D_o\sum_{x\in T}\bar{\varphi}(|T|)\Big] = \frac{1}{2}\mathbb{E}_\mu\Big[\sum_{x\in T}D_x\bar{\varphi}(|T|)\Big]$$
$$= \mathbb{E}_\mu[\bar{\varphi}(|T|)|E_T|] = \mathbb{E}_\mu[\varphi(|T|)] - \mathbb{E}_\mu[\bar{\varphi}(|T|)].$$

Therefore, $\lim_{\beta\to\infty}\bar{\Phi}_\mu(\beta, B, h) = \mathbb{E}_\mu[\bar{\varphi}(|T|)]$ which concludes the proof. □

PROOF OF THEOREM 1.10. The Potts model (1.3) is of form (1.1) with $\mathscr{X} = [q]$, $\xi(\sigma,\sigma') = \beta\cdot\mathbf{1}\{\sigma = \sigma'\}$, and $\bar{\xi}(\sigma) = B\cdot\mathbf{1}\{\sigma = 1\}$, so (H1) and (H2) are clearly satisfied. It follows from the recursive structure of the tree that $h^\ddagger \in \mathcal{H}^\star$ for $\ddagger \in \{\mathrm{f}, 1\}$. For part (a), along any interpolation path contained in $\mathcal{R}_\mu$, both (H3$^\beta$) and (H3^B) are satisfied by Corollary 4.4 and the same argument used in the proof of Theorem 1.12. For part (b), (H3$^\beta$) and (H3^B) are satisfied by the additional hypothesis of continuity.

The inequalities in part (b) then follow from Theorem 1.15 once we verify [cf. (1.19), (1.20)]

$$a^{\mathrm{vx}}(\beta, B, h^{\mathrm{f}}) \le \liminf_{n\to\infty} a_n^{\mathrm{vx}}(\beta, B) \le \limsup_{n\to\infty} a_n^{\mathrm{vx}}(\beta, B) \le a^{\mathrm{vx}}(\beta, B, h^1),$$
$$a^{\mathrm{e}}(\beta, B, h^{\mathrm{f}}) \le \liminf_{n\to\infty} a_n^{\mathrm{e}}(\beta, B) \le \limsup_{n\to\infty} a_n^{\mathrm{e}}(\beta, B) \le a^{\mathrm{e}}(\beta, B, h^1),$$

where $a_n^{\mathrm{vx}}(\beta, B) = \mathbb{E}_n[\langle\mathbf{1}\{\sigma_{I_n} = 1\}\rangle_n^{\beta,B}]$ and $a_n^{\mathrm{e}}(\beta, B) = \frac{1}{2}\mathbb{E}_n[\sum_{j\in\partial I_n}\langle\mathbf{1}\{\sigma_{I_n} = \sigma_j\}\rangle_n^{\beta,B}]$. Indeed, by Vitali's convergence theorem, the assumption $G_n \to_{lwc} \mu$ and Corollary 4.4 [with $U = B_t(I_n) \subseteq G_n$], we have

$$a^{\mathrm{e}}(\beta, B, h^{\mathrm{f}}) = \liminf_{t\to\infty}\frac{1}{2}\mathbb{E}_\mu\Big[\sum_{j\in\partial o}\langle\mathbf{1}\{\sigma_o = \sigma_j\}\rangle_{T^t,T}^{\mathrm{f},\beta,B}\Big] \le \liminf_{n\to\infty} a_n^{\mathrm{e}}(\beta, B),$$

and the other inequalities are proved similarly. Together these inequalities imply that

$$\lim_{n\to\infty}[\phi_n(\beta', B') - \phi_n(\beta, B)] = \Phi(\beta', B', h^\ddagger) - \Phi(\beta, B, h^\ddagger)$$

for any (β, B) and (β', B') joined by an interpolation path contained in $\mathcal{R}_\mu$. The result of part (a) then follows by letting (β', B') approach $\mathcal{R}_\infty$ and applying Lemma 4.5. □

4.3. *Potts model with d-regular limiting tree.* In this section we prove Theorem 1.11, which amounts to determining the shape of $\mathcal{R}_{\neq}$ and establishing continuity of h^{f} and h^1 in certain regimes.

Since the limiting measure is supported on T_d, only $h \equiv h_{(\mathsf{T}_d, x \to y)}$ is of relevance. Further, $h^\ddagger$ is symmetric among the spins $\neq 1$ for $\ddagger \in \{\mathrm{f}, 1\}$, so determination of $h^\ddagger$ reduces to solving a univariate recursion for $h^\ddagger(1)$,

$$h \mapsto \frac{e^B[e^\beta h + (1-h)]^{d-1}}{e^B[e^\beta h + (1-h)]^{d-1} + (q-1)[h + ((1-h)/(q-1))(e^\beta + q - 2)]^{d-1}}.$$

Our result follows from analysis of the fixed points of this mapping; similar computations have appeared, for example, in [43, 47] so some overlap among the analyses may occur.

A convenient parametrization is given by the log likelihood ratio $r \equiv \log h - \log[(1-h)/(q-1)]$, in terms of which the recursion becomes

$$r \mapsto f(r) \equiv f(r; \beta, B) = B + (d-1)\log\left(\frac{e^{\beta+r} + q - 1}{e^r + e^\beta + q - 2}\right).$$

With $f^{(t)}$ the t-fold iteration of f, let r^{f} denote the increasing limit of $f^{(t)}(0)$ and r^1 the decreasing limit of $f^{(t)}(\infty)$, as $t \to \infty$. The region $\mathcal{R}_{\neq}$ corresponds to those $\beta, B \geq 0$ for which $r^{\mathrm{f}} \neq r^1$.

LEMMA 4.6. *There exists $\beta_- > 0$ such that for $\beta \leq \beta_-$ the map f has exactly one fixed point for any $B \in \mathbb{R}$. For $\beta > \beta_-$ there exist real-valued $B_-(\beta) < B_+(\beta)$ (smooth in β) such that f has one, two or three fixed points depending on whether B is in $[B_-, B_+]^c$, $\{B_-, B_+\}$ or (B_-, B_+). The curves extend continuously to $B_-(\beta_-) = B_+(\beta_-)$.*

PROOF. We have

$$f'(r) = \frac{(d-1)e^r(e^\beta - 1)(q + e^\beta - 1)}{(q + e^r + e^\beta - 2)(q + e^{r+\beta} - 1)} \tag{4.7}$$

so f is increasing in r with $f'(r) \to 0$ as $r \to \pm\infty$. Since $f(r; \beta, B) = f(\beta; r, B)$, it easily follows from (4.7) that $\partial_\beta f(r)$ has the same sign as r while $\partial_\beta[f'(r)] > 0$. Further

$$f''(r) = -\frac{(d-1)e^{r+\beta}(e^\beta - 1)(q + e^\beta - 1)(e^{2r} - \alpha)}{(q + e^r + e^\beta - 2)^2(q + e^{r+\beta} - 1)^2},$$
$$\alpha \equiv (q-1)\bigl(1 + (q-2)e^{-\beta}\bigr),$$

with $\alpha > 0$ since $q > 1$. Notice that $f''(r) > 0$ for r sufficiently negative and $f''(r) < 0$ for r sufficiently positive, with a single sign change occurring at $(\log \alpha)/2$ which is zero for $q = 2$ and strictly positive for $q > 2$. This proves that f has between one and three fixed points. When $B = 0$, one fixed point is always given by $r^{\mathrm{f}}(\beta, 0) = 0$. Further $f(r; 0, 0) \equiv 0$, so (by monotonicity of f' in β) there exists $\infty \geq \beta_- \geq 0$ such that $f' \leq 1$ everywhere for $\beta \leq \beta_-$, and f' exceeds 1 somewhere for $\beta > \beta_-$.

Solving the equation $f'(r) = 1$ in terms of $t \equiv e^r$ yields solutions

$$t_\pm(\beta) = -\gamma \pm \sqrt{\gamma^2 - \alpha}, \qquad \gamma \equiv e^\beta + q - 2 - \frac{d}{2}(1 - e^{-\beta})(e^\beta + q - 1).$$

Since $\alpha > 0$, $t_\pm(\beta)$ are not positive if $\gamma > -\sqrt{\alpha}$, equal to $\sqrt{\alpha} > 0$ if $\gamma = -\sqrt{\alpha}$, and positive but not equal if $\gamma < -\sqrt{\alpha}$. If $d \geq 2$, it is easy to check that both α and γ decrease smoothly in β, starting at $\gamma|_{\beta=0} = q - 1$ and $\alpha|_{\beta=0} = (q-1)^2$, so there is a unique value $\beta = \beta_- > 0$ at which $\gamma = -\sqrt{\alpha}$: if $d = 2$, then $\beta_- = \infty$, and if $d > 2$, then β_- is the logarithm of the unique finite positive root b_- of

$$(d-2)^2 b^2 + (d-2)^2(q-2)b - d^2(q-1) = 0. \tag{4.8}$$

Hence, the equation $f'(r) = 1$ has no solutions for $\beta < \beta_-$, and it has solutions $\rho_\pm(\beta) \equiv \log t_\pm(\beta)$ for $\beta \geq \beta_-$, with $\rho_-(\beta_-) = \rho_+(\beta_-)$ and $\rho_-(\beta) < \rho_+(\beta)$ for $\beta > \beta_-$. The values of $B_-(\beta)$, $B_+(\beta)$ are then given explicitly by

$$B_\pm(\beta) = \rho_\mp(\beta) - f\bigl(\rho_\mp(\beta); \beta, 0\bigr), \tag{4.9}$$

which clearly meet at $\beta = \beta_-$ and are smooth for $\beta > \beta_-$. $\square$

Considering hereafter only $d > 2$ (so that $\beta_- < \infty$), suppose $\beta > \beta_-$, so that the functions $\rho_\pm$ are defined. Since $\partial_\beta[f'(r)] > 0$, ρ_- and ρ_+ must be, respectively, decreasing and increasing in β. Further, since f has a unique inflection point at $(\log\alpha)/2$, we must have $\rho_-(\beta) \leq (\log\alpha)/2 \leq \rho_+(\beta)$, with strict inequalities unless $\rho_-(\beta) = \rho_+(\beta)$. For $q = 2$ (Ising), this implies $\rho_- \leq 0 \leq \rho_+$ from which it is easy to see that whenever $B > 0$ we have $r^{\mathrm{f}}(\beta, B) = r^1(\beta, B)$, which is then continuous in (β, B) by the same argument as in the proof of Theorem 1.12. When $B = 0$, $r^{\mathrm{f}}(\beta, 0)$ is zero for all β, while $r^1(\beta, B)$ is zero for $\beta \leq \beta_-$ and strictly positive for $\beta > \beta_-$.

For $q > 2$ (Potts), this implies that $\rho_+(\beta, B) > 0$ while $\rho_-(\beta, B) \geq 0$ if and only if $f'(0; \beta, B) \leq 1$. From the calculations above, $f'(0)$ is zero at $\beta = 0$ and increases in β. We therefore define

$$\begin{aligned}
&\beta_{\mathrm{f}} \equiv \inf\{\beta \geq 0 : f(r; \beta, 0) = r \text{ for some } r > 0\},\\
&\beta_+ \equiv \inf\{\beta \geq 0 : \rho_-(\beta) \leq 0\} = \inf\{\beta \geq 0 : f'(0; \beta, 0) \geq 1\}\\
&\qquad = \log\left(1 + \frac{q}{d-2}\right)
\end{aligned} \tag{4.10}$$

[where the formula for β_+ comes from (4.7)]. Clearly $\beta_- \leq \beta_{\mathrm{f}} \leq \beta_+$, and in fact these inequalities are strict: at β_{f}, f' must exceed one between zero and the positive fixed point, so $\beta_- < \beta_{\mathrm{f}}$.[5] Likewise, if $f'(0) \geq 1$ at $\beta = \beta_{\mathrm{f}}$, the concavity of $f(r)$ at $r = 0$ would imply the existence of a positive fixed point at some β below β_{f}

[5]Note that $r^1(\beta_{\mathrm{f}}, 0) > 0$, that is, the 1-biased fixed point "arises discontinuously."

which is a contradiction, so $\beta_{\mathrm{f}} < \beta_+$. We refer again to Figure 2 which shows the maps $f(r;\beta,B)$ for the Ising and Potts models at several values of β while holding $B=0$. Figure 3(b) shows the regime of (β,B) values delineated by the curves $B_\pm(\beta)$.

PROOF OF THEOREM 1.11. (a) We found above that $\mathcal{R}_{\neq}=\varnothing$ for $d=2$ and $\mathcal{R}_{\neq}=(\beta_-,\infty)$ for $q=2$, so suppose $d,q>2$. If $B>0$, $r^{\mathrm{f}}=r^1$ holds for all $\beta\ge 0$ with $\beta\notin(\beta_-,\beta_+)$. For $\beta\in(\beta_-,\beta_+)$ there is a closed interval $[B_-(\beta)\vee 0, B_+(\beta)]$ of B values for which $r^{\mathrm{f}}<r^1$: this interval is strictly positive for $\beta<\beta_{\mathrm{f}}$ and includes zero for $\beta\ge\beta_{\mathrm{f}}$. If $B=0$, $r^{\mathrm{f}}=r^1$ for $0\le\beta<\beta_{\mathrm{f}}$ and $r^{\mathrm{f}}<r^1$ for $\beta\ge\beta_{\mathrm{f}}$. Recalling (4.9),

$$\begin{aligned}\partial_\beta B_\pm(\beta)&=\partial_\beta\bigl[\rho_\mp(\beta)-f\bigl(\rho_\mp(\beta)\bigr)\bigr]=\bigl[1-f'\bigl(\rho_\mp(\beta)\bigr)\bigr]\partial_\beta\rho_\mp(\beta)-(\partial_\beta f)\bigl(\rho_\mp(\beta)\bigr)\\&=-(\partial_\beta f)\bigl(\rho_\mp(\beta)\bigr).\end{aligned}$$

This has the same sign as $-\rho_\mp(\beta)$, which are both negative for $0\le\beta<\beta_+$, so the curves $B_\pm(\beta)$ are decreasing. Inverting them gives the curves $\beta_{\mathrm{f}}(B),\beta_+(B)$ which delineate the region $\mathcal{R}_{\neq}$ as described in the theorem statement, with $\beta_{\mathrm{f}}(0)=\beta_{\mathrm{f}}$ and $\beta_+(0)=\beta_+$.

(b) Away from the boundary of $\mathcal{R}_{\neq}$, h^{f} and h^1 correspond to isolated zeros of a smooth function, and so are continuous by the implicit function theorem. From part (a), any point of $\mathcal{R}$ is connected to $\mathcal{R}_\infty$ by an interpolation path contained in $\mathcal{R}$, so applying Theorem 1.10(a) verifies the Bethe prediction for $(\beta,B)\notin\mathcal{R}_{\neq}$.

Since changing B only translates $f(r;\beta,B)$, it is not difficult to see that when $\beta\in(\beta_-,\beta_+)$, the function $h^{\mathrm{f}}(\beta,B)$ is continuous in B for $B\in[0,B_+(\beta)]$ while $h^1(\beta,B)$ is continuous for $B\in[B_-(\beta)\vee 0,\infty)$. It follows by Lemma 2.1 that for $(\beta,B)\in\partial\mathcal{R}_{\neq}$ with $\beta=\beta_{\mathrm{f}}(B)$, $\phi(\beta,B)=\Phi(\beta,B,h^{\mathrm{f}})$, while for $(\beta,B)\in\partial\mathcal{R}_{\neq}$ with $\beta\ge\beta_+(B)$, $\phi(\beta,B)=\Phi(\beta,B,h^1)$.

Recall our convention that $\beta_0\le\beta_1$, $B_0\le B_1$: by Theorem 1.10(b) we may interpolate in B from $(\beta,B_0)\in\mathcal{R}_{\neq}^\circ$ to $(\beta,B_1)\in\mathcal{R}$ using the message h^1, yielding $\liminf_{n\to\infty}\phi_n(\beta,B)\ge\Phi(\beta,B,h^1)$ for $(\beta,B)\in\mathcal{R}_{\neq}^\circ$. Likewise, we may interpolate in B from $(\beta,B_0)\in\mathcal{R}$ to $(\beta,B_1)\in\mathcal{R}_{\neq}$ using h^{f} (and once inside $\mathcal{R}_{\neq}$ we may also interpolate in β using h^{f}), which gives $\liminf_{n\to\infty}\phi_n(\beta,B)\ge\Phi(\beta,B,h^{\mathrm{f}})$ for $(\beta,B)\in\mathcal{R}_{\neq}^\circ$.

Next, since $h^{\mathrm{f}}(\beta,B)$ and $h^1(\beta,B)$ are lower and upper semi-continuous, respectively, in β, and both are nondecreasing in β, for $0<B<B_+$ we have that $h^{\mathrm{f}}(\beta,B)\uparrow h^{\mathrm{f}}(\beta_+(B),B)$ as $\beta\uparrow\beta_+(B)$ and $h^1(\beta,B)\downarrow h^1(\beta_{\mathrm{f}}(B),B)$ as $\beta\downarrow\beta_{\mathrm{f}}(B)$. Again by Theorem 1.10(b), we may interpolate in β from $(\beta_0,B)=(\beta_{\mathrm{f}}(B),B)\in\partial\mathcal{R}_{\neq}$ to $(\beta_1,B)\in\mathcal{R}_{\neq}^\circ$ using h^1, and from $(\beta_0,B)\in\mathcal{R}_{\neq}^\circ$ to $(\beta_1,B)=(\beta_+(B),B)\in\partial\mathcal{R}_{\neq}$ using h^{f}, giving

$$\limsup_{n\to\infty}\phi_n(\beta,B)\le\min\bigl\{\widetilde{\Phi}^{\mathrm{f}}(\beta,B),\widetilde{\Phi}^1(\beta,B)\bigr\},\qquad(\beta,B)\in\mathcal{R}_{\neq}^\circ,$$

which completes the proof. $\square$

Acknowledgments. We thank Allan Sly and Ofer Zeitouni for many helpful conversations. A. Dembo and N. Sun thank the Microsoft Research Theory Group for supporting a visit during which part of this work was completed.

REFERENCES

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

[2] 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

[3] BANDYOPADHYAY, A. and GAMARNIK, D. (2006). Counting without sampling. New algorithms for enumeration problems using statistical physics. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms* 890–899. ACM, New York. MR2373815

[4] BANDYOPADHYAY, A. and GAMARNIK, D. (2008). Counting without sampling: Asymptotics of the log-partition function for certain statistical physics models. *Random Structures Algorithms* **33** 452–479. MR2462251

[5] BAYATI, M., GAMARNIK, D. and TETALI, P. (2010). Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. In *STOC*'10—*Proceedings of the* 2010 *ACM International Symposium on Theory of Computing* 105–114. ACM, New York. MR2743259

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

[7] BISKUP, M., BORGS, C., CHAYES, J. T. and KOTECKÝ, R. (2000). Gibbs states of graphical representations of the Potts model with external fields. *J. Math. Phys.* **41** 1170–1210. MR1757955

[8] COOK, S. A. (1971). The complexity of theorem-proving procedures. In *STOC* '71 *Proceedings of the Third Annual ACM Symposium on Theory of Computing* 151–158. ACM, New York.

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

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

[11] DEMBO, A., MONTANARI, A., SLY, A. and SUN, N. (2012). The replica symmetric solution for Potts models on d-regular graphs. Preprint. Available at http://arxiv.org/abs/1207.5500.

[12] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. Springer, New York. MR1619036

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

[14] DYSON, F. J. (1969). Existence of a phase-transition in a one-dimensional Ising ferromagnet. *Comm. Math. Phys.* **12** 91–107. MR0436850

[15] EVANS, W., KENYON, C., PERES, Y. and SCHULMAN, L. J. (2000). Broadcasting on trees and the Ising model. *Ann. Appl. Probab.* **10** 410–433. MR1768240

[16] FORTUIN, C. M. and KASTELEYN, P. W. (1972). On the random-cluster model. I. Introduction and relation to other models. *Physica* **57** 536–564. MR0359655

[17] FRANZ, S. and LEONE, M. (2003). Replica bounds for optimization problems and diluted spin systems. *J. Stat. Phys.* **111** 535–564. MR1972121

[18] FRANZ, S., LEONE, M. and TONINELLI, F. L. (2003). Replica bounds for diluted non-Poissonian spin systems. *J. Phys. A* **36** 10967–10985. MR2025238

[19] GALANIS, A., ŠTEFANKOVIČ, D. and VIGODA, E. (2012). Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. Preprint. Available at http://arxiv.org/abs/1203.2226v2.

[20] GALANIS, A., GE, Q., ŠTEFANKOVIČ, D., VIGODA, E. and YANG, L. (2011). Improved inapproximability results for counting independent sets in the hard-core model. In *Approximation, Randomization, and Combinatorial Optimization* (L. Goldberg, K. Jansen, R. Ravi and J. Rolim, eds.). *Lecture Notes in Computer Science* **6845** 567–578. Springer, Heidelberg. MR2863291

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

[22] GREENHILL, C. (2000). The complexity of counting colourings and independent sets in sparse graphs and hypergraphs. *Comput. Complexity* **9** 52–72. MR1791090

[23] GRIMMETT, G. (2009). Correlation inequalities of GKS type for the Potts model. Preprint. Available at arXiv:0901.1625.

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

[25] GUERRA, F. and TONINELLI, F. L. (2002). The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.* **230** 71–79. MR1930572

[26] HAMMERSLEY, J. M. (1962). Generalization of the fundamental theorem on sub-additive functions. *Proc. Cambridge Philos. Soc.* **58** 235–238. MR0137800

[27] HIGUCHI, Y. (1977/78). Remarks on the limiting Gibbs states on a $(d + 1)$-tree. *Publ. Res. Inst. Math. Sci.* **13** 335–348. MR0676482

[28] KARP, R. M. (1972). Reducibility among combinatorial problems. In *Complexity of Computer Computations* (*Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, NY*, 1972) 85–103. Plenum, New York. MR0378476

[29] KELLY, F. P. (1985). Stochastic models of computer communication systems (with discussion). *J. R. Stat. Soc. Ser. B Stat. Methodol.* **47** 379–395, 415–428. MR0844469

[30] LAWLER, E. L., LENSTRA, J. K. and RINNOOY KAN, A. H. G. (1980). Generating all maximal independent sets: NP-hardness and polynomial-time algorithms. *SIAM J. Comput.* **9** 558–565. MR0584512

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

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

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

[34] 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

[35] MOSSEL, E., WEITZ, D. and WORMALD, N. (2009). On the hardness of sampling independent sets beyond the tree threshold. *Probab. Theory Related Fields* **143** 401–439. MR2475668

[36] PANCHENKO, D. and TALAGRAND, M. (2004). Bounds for diluted mean-fields spin glass models. *Probab. Theory Related Fields* **130** 319–336. MR2095932

[37] RIVA, V. and CARDY, J. (2006). Holomorphic parafermions in the Potts model and stochastic Loewner evolution. *J. Stat. Mech. Theory Exp.* **12** P12001, 19 pp. (electronic). MR2280251

[38] RUELLE, D. (1969). *Statistical Mechanics: Rigorous Results*. W. A. Benjamin, New York-Amsterdam. MR0289084

[39] SIMON, B. (1979). A remark on Dobrushin's uniqueness theorem. *Comm. Math. Phys.* **68** 183–185. MR0543198

[40] SLY, A. (2010). Computational transition at the uniqueness threshold. *2010 IEEE 51st Annual Symposium on Foundations of Computer Science* 287–296. Las Vegas, USA.

[41] SLY, A. and SUN, N. (2012). The computational hardness of counting in two-spin models on d-regular graphs. Preprint. Available at http://arxiv.org/abs/1203.2602.

[42] SOKAL, A. D. (2005). The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In *Surveys in Combinatorics* 2005. *London Mathematical Society Lecture Note Series* **327** 173–226. Cambridge Univ. Press, Cambridge. MR2187739

[43] SPITZER, F. (1975). Markov random fields on an infinite tree. *Ann. Probab.* **3** 387–398. MR0378152

[44] WEITZ, D. (2006). Counting independent sets up to the tree threshold. In *STOC'*06: *Proceedings of the* 38*th Annual ACM Symposium on Theory of Computing* 140–149. ACM, New York. MR2277139

[45] WU, F. Y. (1982). The Potts model. *Rev. Modern Phys.* **54** 235–268. MR0641370

[46] YEDIDIA, J. S., FREEMAN, W. T. and WEISS, Y. (2005). Constructing free-energy approximations and generalized belief propagation algorithms. *IEEE Trans. Inform. Theory* **51** 2282–2312. MR2246363

[47] ZACHARY, S. (1983). Countable state space Markov random fields and Markov chains on trees. *Ann. Probab.* **11** 894–903. MR0714953

[48] ZINN-JUSTIN, J. (1993). *Quantum Field Theory and Critical Phenomena*, 2nd ed. *International Series of Monographs on Physics* **85**. The Clarendon Press Oxford Univ. Press, New York. MR1227790

[49] ZUCKERMAN, D. (2006). Linear degree extractors and the inapproximability of max clique and chromatic number. In *STOC'*06: *Proceedings of the* 38*th Annual ACM Symposium on Theory of Computing* 681–690. ACM, New York. MR2277193

A. DEMBO
DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
BUILDING 380, SLOAN HALL
STANFORD, CALIFORNIA 94305
USA
AND
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
SEQUOIA HALL, 390 SERRA MALL
STANFORD, CALIFORNIA 94305
USA

A. MONTANARI
DEPARTMENT OF ELECTRICAL ENGINEERING
STANFORD UNIVERSITY
350 SERRA MALL
STANFORD, CALIFORNIA 94305
USA
AND
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
SEQUOIA HALL, 390 SERRA MALL
STANFORD, CALIFORNIA 94305
USA

N. SUN
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
SEQUOIA HALL, 390 SERRA MALL
STANFORD, CALIFORNIA 94305
USA