COUNTING IN TWO-SPIN MODELS ON d-REGULAR GRAPHS

BY ALLAN SLY[1] AND NIKE SUN[2]

*University of California, Berkeley and Australian National University,
and Stanford University*

We establish that the normalized log-partition function of any two-spin system on bipartite locally tree-like graphs converges to a limiting "free energy density" which coincides with the (nonrigorous) Bethe prediction of statistical physics. Using this result, we characterize the local structure of two-spin systems on locally tree-like bipartite expander graphs without the use of the second moment method employed in previous works on these questions. As a consequence, we show that for both the hard-core model and the anti-ferromagnetic Ising model with arbitrary external field, it is NP-hard to approximate the partition function or approximately sample from the model on d-regular graphs when the model has nonuniqueness on the d-regular tree. Together with results of Jerrum–Sinclair, Weitz, and Sinclair–Srivastava–Thurley, this gives an almost complete classification of the computational complexity of homogeneous two-spin systems on bounded-degree graphs.

1. Introduction. Spin systems are stochastic models defined by local interactions on networks. While playing a central role in statistical physics, they are also closely associated with a range of combinatorial counting problems. In this paper, we give a detailed analysis of two-spin models at all temperatures on locally tree-like bipartite d-regular graphs and prove that the free energy—the limiting log-partition function normalized by the number of vertices in the graph—is given by the Bethe prediction. We use this result to obtain detailed information about the local properties of such measures, allowing us to essentially complete a long-standing program of classifying the computational complexity of approximating the partition function for all homogeneous two-spin systems on bounded-degree graphs.

The study of locally tree-like graphs in statistical physics was initiated by Bethe [5] as a way to investigate mean-field phenomena expected in high-dimensional systems. In theoretical computer science and combinatorics, locally tree-like graphs play an important role in the study of randomized constraint satisfaction problems such as random k-SAT or proper graph colorings. The under-

standing of the partition functions of such systems is closely linked to the study of Gibbs measures on trees.

It is natural to ask if the free energy of an ensemble of locally tree-like graphs depends on the choice of ensemble. Dembo–Montanari [7] showed that the free energy density for the ferromagnetic Ising model on a class of locally tree-like graphs is given by the Bethe prediction, defined in terms of a distributional fixed point of a certain recursion on the limiting tree (and hence not depending on the particular graph sequence); this result was subject to a second moment condition on the degree distribution which was relaxed to a $(1 + \varepsilon)$-moment condition by Dommers et al. [10]. The Bethe prediction was proved for a wide class of models and graph sequences in regimes of Gibbs uniqueness [9], and it was recently verified in all regimes for the ferromagnetic Potts model on d-regular graphs with d even [8]. In contrast, for anti-ferromagnetic models at low temperatures, the graph ensemble plays an important role in the asymptotic partition function: for the hard-core model at high fugacity, the free energy density for random d-regular graphs can easily be shown to differ from that for random bipartite d-regular graphs. Indeed the statistical physics theory of replica symmetry breaking was developed to deal with the frustration and complicated long-range dependencies induced by the anti-ferromagnetic nature of such models.

In the case of two-spin systems on d-regular bipartite graphs, we show that at all temperatures the limiting free energy density is independent of the ensemble chosen and is again given by the Bethe prediction.

THEOREM 1. *For any nondegenerate homogeneous two-spin model on bipartite d-regular locally tree-like graphs, the log-partition function normalized by the number of vertices has an asymptotic value which coincides with the Bethe free energy prediction.*

The Bethe prediction is defined precisely in Section 2.2; for the precise statement, see Theorem 4. Establishing lower bounds on partition functions is challenging in general, and for random graph ensembles this is often done using the second moment method. This approach typically leads to difficult optimization problems which become increasingly challenging in systems with more parameters (see, e.g., [1–3]).

In this paper, we follow a different approach which is more conceptual and completely circumvents second moment method calculations, and further yields results for more general graph ensembles. The idea is to bound the derivative of the log-partition function with respect to the inverse temperature or fugacity of the model, similarly to what was done in [7, 9, 10] for the ferromagnetic Ising and Potts models. However, unlike those models, the systems we consider not have an FK-representation or the FKG property.

We show that the derivative is bounded above by a maximization over Gibbs measures on the d-regular tree which is attained by the extremal semi-translation-invariant Gibbs measures (see Definition 1.7). At low temperatures, these measures

favor a higher density of +-spins on one side of the bipartition. Underpinning our arguments is that on bipartite graphs (in contrast to general graphs) the limiting free energy can be calculated at zero temperature ($\beta \uparrow \infty$ or $\lambda \uparrow \infty$) as well as at infinite temperature ($\beta = 0$ or $\lambda \downarrow 0$). Interpolating our sharp upper bound on the derivative between these endpoint values then gives matching upper and lower bounds for the free energy.

Our proof further yields additional information about the local structure of the model on the finite graphs. Adapting methods developed in [19], we show that the local weak limit is a mixture of the two semi-translation-invariant Gibbs measures. More can be said under additional assumptions that the graph is an expander and symmetric with respect to the two sides of the bipartition: the spins have long-range correlations due to the one-sided bias ("phase") of a typical configuration, but conditioned on the phase the spins of a collection of vertices are essentially independent with marginals depending on the side of the graph chosen. A full description of the local behavior is developed in Theorem 5 and Proposition 4.1.

An extensive literature in theoretical computer science has analyzed the question of when it is possible to approximate the partition function of Gibbs measures on sparse graphs. Using our detailed analysis of the local structure, we are able to essentially complete the classification of the complexity of approximating the partition function for two-spin systems on *d*-regular graphs. Jerrum and Sinclair [16] gave a fully polynomial-time randomized approximation scheme (a randomized algorithm whose running time is polynomial in both in the number of vertices and the required accuracy, hereafter abbreviated FPRAS) for approximating the partition function of the ferromagnetic Ising model, which covers all ferromagnetic two-spin systems.

For anti-ferromagnetic systems such as the hard-core and anti-ferromagnetic Ising models, the complexity of approximating the partition function depends on the model parameters, and is known to be NP-hard when the interactions are sufficiently strong. In this paper, we establish that the computational transition for such models on *d*-regular graphs is located precisely at the uniqueness threshold (see Definition 1.6) for the corresponding model on the *d*-regular tree.

THEOREM 2. *For $d \geq 3$ and $\lambda > \lambda_c(d) = \frac{(d-1)^{d-1}}{(d-2)^d}$, unless* RP = NP *there exists no* PRAS *for the partition function of the hard-core model with fugacity λ on d-regular graphs.*

The famous conjecture P $\neq$ NP states that deterministic polynomial-time algorithms cannot solve all NP-problems (roughly, problems whose solutions are polynomial-time verifiable). Our hardness results assume the conjecture RP $\neq$ NP which states that polynomial-time algorithms using randomness cannot solve all NP-problems; this assumption is standard in computational complexity theory. For information on the RP complexity class, see, for example, [21].

The uniqueness threshold $\lambda_c(d)$ marks the point above which distant boundary conditions have a nonvanishing influence on the spin at the root. In a seminal paper [24], Weitz used computational tree methods to provide a (deterministic) fully polynomial-time approximation scheme (FPTAS) for the partition function of the hard-core model on graphs of maximum degree d at any $\lambda < \lambda_c(d)$. Together with Weitz's result, Theorem 2 completes the classification of the complexity of the hard-core model except at the threshold λ_c.

Previously, it was shown that there is no FPRAS for the hard-core model at $\lambda d \geq 10,000$ [18]. In the case of $\lambda = 1$, this was improved to $d \geq 25$ [11, 12], using random regular bipartite graphs as basic gadgets in a hardness reduction. Mossel et al. [20] showed that local MCMC algorithms are exponentially slow for $\lambda > \lambda_c(d)$, and conjectured that λ_c is in fact the threshold for existence of an FPRAS.

The first rigorous result establishing a computational transition at the uniqueness threshold appeared in [23], where hardness was shown for $\lambda_c(d) < \lambda < \lambda_c(d) + \varepsilon(d)$ for some $\varepsilon(d) > 0$. The proof relies on a detailed analysis of the hard-core model on random bipartite graphs, which are then used in a randomized reduction to MAX-CUT. More precisely, the result of [23] gives hardness subject to a technical condition which was an artifact of a difficult second moment calculation from [20], and which could only be verified for $\lambda < \lambda_c(d) + \varepsilon(d)$. Hardness was subsequently shown by Galanis et al. [13] for all $\lambda > \lambda_c(d)$ when $d \neq 4, 5$ by verifying the technical condition of [23].

As mentioned above, our method avoids the difficult second moment calculations. Moreover, essentially the same method of proof gives the analogous result for anti-ferromagnetic Ising models with arbitrary external field:

THEOREM 3. *For* $d \geq 3$, $B \in \mathbb{R}$ *and* $\beta < \beta_{c,af}(d, B) < 0$, *unless* RP = NP *there exists no* PRAS *for the partition function of the anti-ferromagnetic Ising model with inverse temperature β and external field B on d-regular graphs.*

Here, $\beta_{c,af}(d, B)$ denotes the uniqueness threshold for the anti-ferromagnetic Ising model with external field B on the d-regular tree. Extending the methods of Weitz [24], Sinclair et al. [22] (see also [17]) gave an FPTAS for the anti-ferromagnetic Ising model on d-regular graphs at inverse temperature $\beta > \beta_{c,af}(d, B)$, so together with Theorem 3 this again establishes that the computational transition coincides with the tree uniqueness threshold.

We emphasize that while Theorem 1 applies to a class of bipartite tree-like d-regular graphs, Theorems 2 and 3 concern the problem of computing the partition function on the class of (all) d-regular graphs. The hard-core and anti-ferromagnetic Ising models together encompass all (nondegenerate) homogeneous two-spin systems on d-regular graphs (see Section 2.1). Thus, the results of [16, 22, 24] combined with Theorems 2 and 3 give a full classification of the computational complexity of approximating the partition function for (homogeneous)

two-spin systems on d-regular graphs, except at the uniqueness thresholds $\lambda_c(d)$ and $\beta_{c,af}(d, B)$.

In fact, we will show inapproximability in nonuniqueness regimes in a strong sense: not only does there not exist a PRAS, but for any fixed choice of model parameters and d there exists $c > 0$ such that it is NP-hard even to approximate the partition function within a factor of e^{cn} on the class of d-regular graphs.

Independent results of Galanis–Štefankovič–Vigoda. In a simultaneous and independent work, Galanis, Štefankovič and Vigoda [14] established the result of Theorem 2, and Theorem 3 in the case of zero external field ($B = 0$). Their methods differ from ours: they analyze the second moment of the partition function on random bipartite d-regular graphs, and establish the condition necessary to apply the approach of [23]. Their proof analyzes a difficult optimization of a real function in several variables by relating the problem to certain tree recursions.

1.1. *Computational reduction via Gibbs measures on bipartite graphs.* The computational results of Theorems 2 and 3 are proved by a variation on the construction of [23], using the bipartite graphs in a randomized reduction to approximate MAX-CUT on 3-regular graphs, which is known to be NP-hard [4]. First, we use Theorem 5 to construct a symmetric bipartite d-regular locally tree-like graph $\mathcal{G}$ of large constant size such that in the hard-core or anti-ferromagnetic Ising model, conditioned on the phase of the global configuration, spins at distant vertices are asymptotically independent with known marginals depending only on the side of the graph (Proposition 4.1).

Given a 3-regular graph H on which we wish to approximate MAX-CUT, first we take a disjoint copy $\mathcal{G}_v$ of $\mathcal{G}$ for each vertex $v \in H$ which we call gadgets. After removing $6k$ edges from each $\mathcal{G}_v$, for each edge $(u, v) \in H$, we add $2k$ edges joining each side of $\mathcal{G}_u$ to the corresponding side of $\mathcal{G}_v$ in such a way that the resulting graph $H^{\mathcal{G}}$ is d-regular.

The connections between gadgets do not substantially change the spin distributions inside them, and in particular the $\pm$ phases remain. The anti-ferromagnetic nature of the interaction, however, results in neighboring copies of $\mathcal{G}$ in $H^{\mathcal{G}}$ preferring to be in opposing phases. Using the asymptotic conditional independence result Proposition 4.1, we can estimate the partition function for the model on $H^{\mathcal{G}}$ restricted to configurations of given phase on each copy of $\mathcal{G}$ within a factor of $e^{\varepsilon|H|}$ (Lemma 4.2). We find that the distribution is concentrated on configurations where the vector of phases gives a good cut of H, and the effect is strengthened as k is increased. Thus, for any $\varepsilon > 0$, by taking k (hence $\mathcal{G}$) to be sufficiently large a $(1 + \varepsilon)$-approximation of MAX-CUT(H) can be determined from the partition function of the model on $H^{\mathcal{G}}$, thereby completing the reduction.

In the remainder of this introductory section, we formally introduce the models which we consider. We then define the notion of local (weak) convergence of graphs and give precise statements of our results on the partition function (Theorem 4) and local structure (Theorem 5) of these models on bipartite graphs.

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

$$(1.1) \qquad \nu_G^{\underline{\psi}}(\underline{\sigma}) = \frac{1}{Z_G(\underline{\psi})} \prod_{(ij) \in E} \psi(\sigma_i, \sigma_j) \prod_{i \in V} \bar{\psi}(\sigma_i),$$

where ψ is a symmetric function $\mathscr{X}^2 \to \mathbb{R}_{\geq 0}$, $\bar{\psi}$ is a positive function $\mathscr{X} \to \mathbb{R}_{\geq 0}$, and $Z_G(\underline{\psi})$ is the normalizing constant, called the *partition function*. The pair $\underline{\psi} \equiv (\psi, \bar{\psi})$ is called a *specification* for the spin system (1.1).

In this paper, we consider spin systems with an alphabet of size two; without loss $\mathscr{X} \equiv \{\pm 1\}$. The *Ising* model on G at inverse temperature β and external field B is given by

$$(1.2) \qquad \nu_G^{\beta, B}(\underline{\sigma}) = \frac{1}{Z_G(\beta, B)} \prod_{(ij) \in E} e^{\beta \sigma_i \sigma_j} \prod_{i \in V} e^{B \sigma_i}.$$

The model is said to be *ferromagnetic* (favoring the alignment of neighboring spins) when $\beta \geq 0$, and *anti-ferromagnetic* (neighboring spins repel) when $\beta < 0$. The *hard-core* (or *independent set*) model on G at activity or fugacity λ is given by

$$(1.3) \qquad \nu_G^{\lambda}(\underline{\sigma}) = \frac{1}{Z_G(\lambda)} \prod_{(ij) \in E} \mathbf{1}\{\bar{\sigma}_i \bar{\sigma}_j \neq 1\} \prod_{i \in V} \lambda^{\bar{\sigma}_i},$$

where $\bar{\sigma} \equiv \mathbf{1}\{\sigma = +1\} = \frac{1}{2}(1 + \sigma)$. The edge interaction has no temperature parameter and is a hard constraint forbidding neighboring occupied sites, so the hard-core model is considered *anti-ferromagnetic* for all $\lambda > 0$. Our definition (1.3) is trivially equivalent to the standard definition of the hard-core model which has spin 0 in place of -1, but we take $\mathscr{X} = \{\pm 1\}$ throughout to unify the notation.

1.3. *Local convergence and the Bethe prediction.* If G is any graph and v a vertex in G, write $B_t(v)$ for the subgraph induced by the vertices of G at graph distance at most t from v, and $\partial v \equiv B_1(v) \setminus \{v\}$ for the neighbors of v. We let $T \equiv (T, o)$ denote a general tree with root o, with $T^t \equiv B_t(o) \subseteq T$ the subtree of depth t. We also fix d throughout and write $\mathbb{T} \equiv (\mathbb{T}, o)$ for the rooted d-regular tree. (Every vertex of $\mathbb{T}$ has d neighbors, in contrast with the $(d-1)$-ary tree where the root has $d-1$ neighbors.)

DEFINITION 1.1. Let $G_n = (V_n = [n], E_n)$ be a sequence of (random) finite undirected graphs, and let $I_n \in V_n$ denote a uniformly random vertex. The sequence G_n is said to *converge locally* to the d-regular tree $\mathbb{T}$ if for all $t \geq 0$, $B_t(I_n)$ converges to $\mathbb{T}^t$ in distribution with respect to the joint law $\mathbb{P}_n$ of (G_n, I_n): that is, $\lim_{n \to \infty} \mathbb{P}_n(B_t(I_n) \cong \mathbb{T}^t) = 1$ (where $\cong$ denotes graph isomorphism).

We write $\mathbb{E}_n$ for expectation with respect to $\mathbb{P}_n$ and impose the following integrability condition on the degree of I_n.

DEFINITION 1.2.   The sequence G_n is *uniformly sparse* if the random variables $|\partial I_n|$ are uniformly integrable, that is, if

$$\lim_{L\to\infty} \limsup_{n\to\infty} \mathbb{E}_n\big[|\partial I_n|\mathbf{1}\{|\partial I_n|\geq L\}\big]=0.$$

We assume throughout that G_n ($n\geq 1$) is a uniformly sparse graph sequence converging locally to the d-regular tree $\mathbb{T}$; this setting is hereafter denoted $G_n \to_{\mathrm{loc}} \mathbb{T}$. The *free energy density* for a specification $\underline{\psi}$ on G_n is defined by

$$(1.4) \qquad \phi \equiv \lim_{n\to\infty} \phi_n \equiv \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_n[\log Z_n], \qquad Z_n \equiv Z_{G_n}(\underline{\psi}),$$

provided the limit exists. For ferromagnetic spin systems on a broad class of locally tree-like graphs, heuristic methods from statistical physics yield an explicit (conjectural) formula for the value of ϕ, the so-called "Bethe prediction" Φ whose definition we recall in Section 2.2. For anti-ferromagnetic two-spin models, the Bethe prediction is well defined only on graph sequences G_n which are near-bipartite, in the following sense: let $\mathbb{T}_+$ denote the d-regular tree $\mathbb{T}$ with vertices colored $+1$ (black) or -1 (white) according to whether they are at even or odd distance from the root o; let $\mathbb{T}_-$ be $\mathbb{T}_+$ with the colors reversed. Let $_b\mathbb{T}$ be the random tree which equals $\mathbb{T}_+$ or $\mathbb{T}_-$ with equal probability; write $_b\mathbb{P}$ for the law of $_b\mathbb{T}$ and $_b\mathbb{E}$ for expectation with respect to $_b\mathbb{P}$.

DEFINITION 1.3.   For $G_n \to_{\mathrm{loc}} \mathbb{T}$, we say the G_n are *near-bipartite*, and write $G_n \to_{\mathrm{loc}} {}_b\mathbb{T}$ (equivalently $G_n \to_{\mathrm{loc}} {}_b\mathbb{P}$), if there exists a black–white coloring of G_n such that for all $t\geq 0$, $B_t(I_n) \to {}_b\mathbb{T}^t$ in distribution.

In Definition 1.3, the coloring on G_n can be random, and is not required to be proper, though its limit in distribution must be proper for the graph sequence to be near-bipartite. The canonical example of a uniformly sparse graph sequence converging to the d-regular tree $\mathbb{T}$ is the random d-regular multigraph $\mathcal{G}_{n,d}^{\mathrm{cm}}$, sampled according to the "configuration model" as described in Lemma 4.3 (see, e.g., [6] for the proof that $\mathcal{G}_{n,d}^{\mathrm{cm}} \to_{\mathrm{loc}} \mathbb{T}$). However, $\mathcal{G}_{n,d}^{\mathrm{cm}}$ is not near-bipartite: an easy first-moment calculation (see Lemma 4.3) shows that for any small constant $\delta > 0$ there exists constant $\lambda_\delta > 0$ such that with high probability as $n \to \infty$, every subset of vertices S in $\mathcal{G}_{n,d}^{\mathrm{cm}}$ of size $n(\frac{1}{2}-\delta) \leq |S| \leq n(\frac{1}{2}+\delta)$ has at least $|S|\lambda_\delta$ internal edges. From this, it is clear that for any black–white coloring of $\mathcal{G}_{n,d}^{\mathrm{cm}}$, either the proportions of black and white vertices are asymptotically unequal, or there will be an asymptotically positive density of vertices with like-colored neighbors, meaning the local limit cannot be $_b\mathbb{T}$. On the other hand, the gadgets used in our reduction are constructed (Section 4.1) from the bipartite double cover of $\mathcal{G}_{n,d}^{\mathrm{cm}}$ which indeed converges to $_b\mathbb{T}$. The precise statement of Theorem 1 is as follows.

THEOREM 4. *For the computation of the free energy density ϕ, all nondegenerate homogeneous two-spin systems on graph sequences $G_n \to_{\text{loc}} \mathbb{T}$ reduce to either the Ising model* (1.2) *or the hard-core model* (1.3).

(a) *In the ferromagnetic Ising model, the free energy density ϕ exists for any $G_n \to_{\text{loc}} \mathbb{T}$ and equals $_{\text{nb}}\Phi$ as defined by* (2.2) [*and given more explicitly by* (2.5)].

(b) *In the hard-core or anti-ferromagnetic Ising models, the free energy density ϕ exists for any $G_n \to_{\text{loc}} {}_b\mathbf{T}$ and equals $_b\Phi$ as defined in* (2.2) [*and given more explicitly by* (2.4)].

The reduction to Ising or hard-core is shown in Section 2.1. While part (a) of Theorem 4 applies to general d-regular tree-like graphs, part (b) is false without the assumption of bipartiteness. For example, for the hard-core model on the random d-regular multigraph $\mathcal{G}_{n,d}^{\text{cm}} \to_{\text{loc}} \mathbb{T}$, one can directly calculate $\mathbb{E}Z_n$ to see that (applying Jensen's inequality) $\limsup_{n\to\infty} n^{-1}\mathbb{E}[\log Z_n] \leq \limsup_{n\to\infty} n^{-1}\log\mathbb{E}Z_n < {}_b\Phi$.

REMARK 1.4. Hereafter, we treat $G_n \to_{\text{loc}} \mathbb{T}$ and $G_n \to_{\text{loc}} {}_b\mathbf{T}$ in a unified manner when possible by writing $G_n \to_{\text{loc}} \mathbb{P}_{\mathcal{T}}$ for $\mathbb{P}_{\mathcal{T}}$ the uniform measure on $\mathcal{T}$, which always denotes either $\{\mathbb{T}\}$ or $\{\mathbb{T}_\pm\}$. We write $\mathbb{E}_{\mathcal{T}}$ for expectation with respect to $\mathbb{P}_{\mathcal{T}}$.

1.4. *Local structure of measures.* Under some additional assumptions on G_n, Theorem 4, together with the arguments of [19], characterizes the asymptotic local structure of the spin systems $\nu_n \equiv \nu_{G_n}$. For $G_n \to_{\text{loc}} {}_b\mathbf{T}$, let $\tau : V_n \to \mathcal{X} \equiv \{\pm\}$ denote the given black–white coloring of the vertices of G_n (hereafter writing $\pm$ as shorthand for ± 1). We say that G_n is *symmetric* if it is isomorphism-invariant to reversing the black–white coloring. For a spin configuration $\underline{\sigma}$ on G_n we define the *phase* of $\underline{\sigma}$ to be

$$Y(\underline{\sigma}) \equiv \text{sgn} \sum_{i \in V_n} \tau_i \sigma_i \qquad \text{where } \text{sgn } x \equiv \mathbf{1}\{x \geq 0\} - \mathbf{1}\{x < 0\}.$$

For $s \in \mathcal{X}$, let ν_n^s denote the measure ν_n conditioned on the s-phase configurations: that is,

$$(1.5) \qquad \nu_n^s(\underline{\sigma}) \equiv \frac{1}{Z_n^s} \mathbf{1}\{Y(\underline{\sigma}) = s\} \prod_{(ij) \in E_n} \psi(\sigma_i, \sigma_j) \prod_{i \in V_n} \bar{\psi}(\sigma_i),$$

where Z_n^s is the partition function restricted to the s-phase configurations. We will characterize the local structure of the measures ν_n^s on graph sequences satisfying an edge-expansion assumption, as follows.

DEFINITION 1.5. A graph $G = (V, E)$ is a $(\delta, \gamma, \lambda)$-*edge expander* if, for any set of vertices $S \subseteq V$ with $\delta|V| \leq |S| \leq \gamma|V|$, there are at least $\lambda|S|$ edges joining S to $V \setminus S$.

The measures ν_n^s will be related to Gibbs measures on the infinite tree. In particular, recall the definition of (Gibbs) uniqueness:

DEFINITION 1.6. For a rooted tree T, let $\mathscr{G}_T$ denote the set of Gibbs measures for the specification $\underline{\psi}$ on T. The specification is said to have (*Gibbs*) *uniqueness* (on T) if $|\mathscr{G}_T| = 1$.

For $s \in \mathscr{X}$, let μ_s be the element of $\mathscr{G}_{\mathbb{T}}$ defined by conditioning on all spins identically equal to s on the vertices at depth $2t$, and taking the weak limit as $t \to \infty$. In the ferromagnetic Ising model, the measures μ_s are extremal and translation-invariant, with $\mu_- \preccurlyeq \mu \preccurlyeq \mu_+$ for all $\mu \in \mathscr{G}_{\mathbb{T}}$ where $\preccurlyeq$ denotes stochastic domination with respect to the coordinate-wise partial ordering $\leq$ on $\mathscr{X}^{\mathbb{T}}$: thus the model has uniqueness if and only if $\mu_+ = \mu_-$. In the hard-core or anti-ferromagnetic Ising model, the measures μ_s are extremal and *semi*-translation-invariant. For $\underline{\sigma}, \underline{\acute{\sigma}} \in \mathscr{X}^{\mathbb{T}}$, write $\underline{\sigma} \leq_b \underline{\acute{\sigma}}$ if $\sigma_v \tau_v \leq \acute{\sigma}_v \tau_v$ for all $v \in \mathbb{T}$ where τ_v is $+1$ if the distance from the root o to v is even, and -1 if odd. Then $\mu_- \preccurlyeq_b \mu \preccurlyeq_b \mu_+$ for all $\mu \in \mathscr{G}_{\mathbb{T}}$ where $\preccurlyeq_b$ denotes stochastic domination with respect to $\leq_b$, so again the model has uniqueness if and only if $\mu_+ = \mu_-$.

Recalling Remark 1.4, let $\mathscr{G}_{\mathcal{T}}$ denote the space of mappings $\nu : T \mapsto \nu(T)$ where $\nu(T) \in \mathscr{G}_T$. When $\mathcal{T} = \{\mathbb{T}_\pm\}$, we denote $\mathscr{G}_{\mathcal{T}}$ by $_b\mathscr{G}$, and write ν_s as shorthand for $\nu(\mathbb{T}_s)$; note $\mathscr{G}_{\mathbb{T}}$ may be naturally regarded as the subset of mappings $\nu \in {_b}\mathscr{G}$ which satisfy $\nu_+ = \nu_-$.

DEFINITION 1.7. An element $\nu \in \mathscr{G}_{\mathcal{T}}$ is *translation-invariant* if for $(T, o) \in \mathcal{T}$ and any vertex $x \in T$, $\nu(T, x)$ coincides with the law on spin configurations of (T, x) induced by $\nu(T, o)$. When $\mathcal{T} = \{\mathbb{T}\}$, this coincides with the usual definition of translation-invariance. In contrast, if $\nu \in {_b}\mathscr{G}$ then the image measures $\nu_+ \equiv \nu(\mathbb{T}_+) \in \mathscr{G}_{\mathbb{T}_+}$ and $\nu_- \equiv \nu(\mathbb{T}_-) \in \mathscr{G}_{\mathbb{T}_-}$—if regarded as Gibbs measures on the uncolored tree $\mathbb{T}$—need only be *semi*-translation-invariant, since $\nu(T, x)$ is allowed to depend on the coloring of x.

With μ_+, μ_-, the extremal semi-translation-invariant elements of $\mathscr{G}_{\mathbb{T}}$ define

$$(1.6) \qquad \nu^r \in {_b}\mathscr{G}, \qquad \nu^r : \mathbb{T}_s \mapsto \mu_{sr}.$$

In our abbreviated notation $\nu_s \equiv \nu(\mathbb{T}_s)$, the above reads $\nu_s^r = \mu_{sr}$. The ν^r are then translation-invariant in the sense of Definition 1.7—that is, invariant under isomorphisms of the tree which preserve the black–white coloring.

DEFINITION 1.8. For $G_n \sim \mathbb{P}_n$ a random graph sequence and ν_n any law on spin configurations $\underline{\sigma}_n$ of G_n, we say that $\mathbb{P}_n \otimes \nu_n$ *converges locally* (*weakly*) to $\mathbb{P}_{\mathcal{T}} \otimes \nu$ (for $\nu \in \mathscr{G}_{\mathcal{T}}$), and write $\mathbb{P}_n \otimes \nu_n \to_{\mathrm{loc}} \mathbb{P}_{\mathcal{T}} \otimes \nu$, if it holds for all $t \geq 0$ that $(B_t(I_n), \underline{\sigma}_{B_t(I_n)})$ converges in distribution to $(T^t, \underline{\sigma}_t)$ where $T \sim \mathbb{P}_{\mathcal{T}}$ and $\underline{\sigma}_t$ is the restriction to T^t of $\underline{\sigma} \sim \nu(T)$. In particular, ν must be translation-invariant in the sense of Definition 1.7.

THEOREM 5. *For any anti-ferromagnetic two-spin system on $G_n \to_{\mathrm{loc\ b}} \mathbf{T}$, the following hold*:

(a) *If the G_n are symmetric, then $\mathbb{P}_n \otimes \nu_n \to_{\mathrm{loc\ b}} \mathbf{P} \otimes (\frac{1}{2}\nu^+ + \frac{1}{2}\nu^-)$.*

(b) *If for all $\delta > 0$ the G_n are $(\delta, \frac{1}{2}, \lambda_\delta)$-edge expanders for some $\lambda_\delta > 0$, then*

$$(1.7) \qquad \mathbb{P}_n \otimes \nu_n^{\mathrm{r}} \to_{\mathrm{loc\ b}} \mathbf{P} \otimes \nu^{\mathrm{r}} \qquad \text{for } \mathrm{r} \in \mathscr{X}.$$

Further, with $\langle \rangle_\mu$ denoting expectation with respect to the Gibbs measure μ, the spatial average $\frac{1}{n}\sum_{i \in V} \tau_i \sigma_i$ is nearly constant in $\underline{\sigma}$:

$$\nu_n^{\mathrm{r}}\left(\left|Y(\underline{\sigma})\frac{1}{n}\sum_{i \in V} \tau_i \sigma_i - \frac{1}{2}\big[\langle \sigma_o \rangle_{\mu_+} - \langle \sigma_o \rangle_{\mu_-}\big]\right| \geq \varepsilon\right) \to 0$$

(1.8)
$$\text{in probability.}$$

1.5. *Outline of the paper.* In Section 2, we review the Bethe prediction in the d-regular setting and prove Theorem 1 (in its form Theorem 4). In Section 3, we show how to deduce Theorem 5 from Theorem 4 by the methods of [19]. In Section 4, we prove the approximate conditional independence statement (Proposition 4.1) and demonstrate the randomized reduction to MAX-CUT to prove Theorems 2 and 3.

2. Partition function for two-spin models. In this section, we prove Theorem 4, establishing the free energy density ϕ (and verifying the Bethe prediction) for two-spin models on graph sequences $G_n \to_{\mathrm{loc\ b}} \mathbf{T}$. In Section 2.1, we show that for purposes of computing ϕ on d-regular locally tree-like graph sequences, all nondegenerate two-spin systems reduce to Ising or hard-core. In Section 2.2, we review the Bethe prediction in the d-regular setting; for more general background and references, we refer to [6, 9]. In Section 2.3, we review an interpolation scheme for the Bethe free energy described in [9], and in Section 2.4 we apply the scheme to compute the free energy density for the hard-core and Ising models, thereby proving Theorem 4.

2.1. *Reduction to Ising and hard-core on d-regular graphs.* We first show that for the computation of the free energy density, all (nondegenerate) homogeneous two-spin models on graph sequences $G_n \to_{\mathrm{loc}} \mathbb{T}$ reduce to either the Ising or hard-core model. Indeed, let $\underline{\psi} \equiv (\psi, \bar{\psi})$ be a specification for a two-spin system with alphabet $\mathscr{X} = \{\pm\}$. If we define $\underline{\psi}'$ by $\psi'(\sigma, \sigma') \equiv \psi(\sigma, \sigma')\bar{\psi}(\sigma)^{1/d}\bar{\psi}(\sigma')^{1/d}$ and $\bar{\psi}'(\sigma) \equiv 1$, then

$$\frac{1}{n}\log Z_G(\underline{\psi}) - \frac{1}{n}\log Z_G(\underline{\psi}') = O\big(\mathbb{E}_n\big[|\partial I_n|\mathbf{1}\{|\partial I_n| \neq d\}\big]\big),$$

which for $G_n \to_{\mathrm{loc}} \mathbb{T}$ tends to zero as $n \to \infty$ by uniform sparsity. Therefore, we assume without loss $\bar{\psi} \equiv 1$, and consider the possibilities for ψ:

1. If $\psi > 0$, then $\psi(\sigma, \sigma') = e^{B_0} e^{\beta \sigma \sigma'} e^{B\sigma/d} e^{B\sigma'/d}$ for β, B, B_0 defined by

$$\frac{\psi(+, +)}{\psi(-, -)} = e^{4B/d},$$

$$\frac{\psi(+, +)\psi(-, -)}{\psi(+, -)^2} = e^{4\beta},$$

$$\psi(+, +)\psi(+, -)^2\psi(-, -) = e^{4B_0},$$

so $\phi_n - B_0 \frac{d}{2}$ is asymptotically equal to the free energy density for the Ising model on G_n with parameters (β, B).

2. If $\psi(+, -) = \psi(-, +) > 0$ and $\psi(-, -) > \psi(+, +) = 0$, then, recalling $\bar{\sigma} \equiv \mathbf{1}\{\sigma = +\}$, we have $\psi(\sigma, \sigma') = e^{B_0} \mathbf{1}\{\bar{\sigma} \bar{\sigma}' \neq 1\} \lambda^{\bar{\sigma}/d} \lambda^{\bar{\sigma}'/d}$ for B_0, λ defined by

$$\psi(-, -) \equiv e^{B_0}, \qquad \psi(+, -)/\psi(-, -) \equiv \lambda^{1/d}.$$

Therefore, $\phi_n - B_0 \frac{d}{2}$ is asymptotically equal to the free energy density for the independent set model on G_n at fugacity λ.

The remaining two-spin models are degenerate, with free energy density which is easy to calculate:

3. Suppose $\psi(+, -) = \psi(-, +) = 0$, so that $\psi(\sigma, \sigma')$ may be written as $\mathbf{1}\{\sigma = \sigma'\} e^{B_0} e^{B\sigma/d} e^{B\sigma'/d}$. Then

$$\phi_n = B_0 \frac{1}{n} \mathbb{E}_n[|E_n|] + B + \frac{1}{n} \mathbb{E}_n \left[\sum_{j=1}^{k(G_n)} \log\left(1 + e^{-2B|C_j|}\right) \right],$$

where the sum is taken over the connected components $C_1, \ldots, C_{k(G_n)}$ of G_n. We claim $\phi_n \to \phi = B_0 \frac{d}{2} + B$: we have $\liminf_{n\to\infty}(\phi_n - \phi) \geq 0$ (using uniform sparsity) and

$$\limsup_{n\to\infty}(\phi_n - \phi) \leq \limsup_{n\to\infty} \frac{1}{n} \mathbb{E}_n[k(G_n)] \log 2,$$

so it suffices to show $\frac{1}{n} \mathbb{E}_n[k(G_n)] \to 0$. Indeed, if this fails then there exists $\varepsilon > 0$ such that for infinitely many n, the event $\{k(G_n) \geq n\varepsilon\}$ occurs with $\mathbb{P}_n$-probability at least ε. On this event, G_n has at least $\frac{1}{2}n\varepsilon$ components of size $\leq 2/\varepsilon$, so for $t > \log_k(2/\varepsilon)$, $\limsup_{n\to\infty} \mathbb{P}_n(B_t(I_n) \not\cong \mathbb{T}^t) \geq \frac{1}{2}\varepsilon^2 > 0$, in contradiction of $G_n \to_{\text{loc}} \mathbb{T}$.

4. Suppose instead $\psi(+, +) = \psi(-, -) = 0$ while $\psi(+, -) = \psi(-, +) > 0$. If the G_n are not exactly bipartite then $\phi_n = -\infty$. If they are exactly bipartite, then

$$\phi_n = \frac{1}{n} \mathbb{E}_n[|E_n|] \log \psi(+, -) + \frac{1}{n} \mathbb{E}_n[k(G_n)] \log 2$$

and by the observation of (3) this converges to $\phi = \frac{d}{2} \log \psi(+, -)$.

2.2. *The Bethe prediction.* Recalling the notation of Remark 1.4, we now review the Bethe prediction for a generic spin model $\underline{\psi}$ with finite spin alphabet $\mathscr{X}$ defined on a graph sequence $G_n \to_{\text{loc}} \mathbb{P}_{\mathcal{T}}$.[3] The prediction is based on a heuristic connection between the free energy density ϕ on a graph sequence $G_n \to_{\text{loc}} \mathbb{P}_{\mathcal{T}}$ and a special class of Gibbs measures on $\mathcal{T}$. These measures are translation-invariant in the sense of Definition 1.7, and are characterized by fixed points $h^\star$ of a certain "Bethe recursion." The Bethe prediction is a nonrigorous formula for ϕ, expressed as an explicit function $\Phi(h)$ evaluated with h equal to a Bethe fixed-point $h^\star$. (It can arise that there are multiple fixed-points $h^\star$, in which case one typically selects the fixed point which maximizes Φ.)

DEFINITION 2.1. Let $\mathcal{T}_e \equiv \mathcal{T}_e(\mathcal{T})$ denote the set of $(T, x \to y)$ (tree T rooted at oriented edge $x \to y$) such that $(T, x) \in \mathcal{T}$. A *message on* $\mathcal{T}$ is a mapping h from $\mathcal{T}_e$ to the $(|\mathscr{X}| - 1)$-dimensional simplex $\boldsymbol{\Delta}$ of probability measures on $\mathscr{X}$.

For $T \in \mathcal{T}$ and h a message on $\mathcal{T}$, we write $h_{x \to y}$ for the image of $(T, x \to y)$ under h. Elements of $\mathcal{T}$ and $\mathcal{T}_e$ are regarded modulo isomorphism, where if the tree is equipped with a black–white coloring then the isomorphism must preserve the coloring: thus, if $\mathcal{T} = \{\mathbb{T}\}$ then $\mathcal{T}_e = \{(\mathbb{T}, o \to j)\}$, whereas if $\mathcal{T} = \{\mathbb{T}_\pm\}$ then $\mathcal{T}_e = \{(\mathbb{T}_\pm, o \to j)\}$. With $\mathcal{H}(\mathcal{T})$ denoting the space of messages on $\mathcal{T}$, we abbreviate $_{\text{nb}}\mathcal{H} \equiv \mathcal{H}(\{\mathbb{T}\})$ and $_b\mathcal{H} \equiv \mathcal{H}(\{\mathbb{T}_\pm\})$.

DEFINITION 2.2. The *Bethe* or *belief propagation recursion* is the map $\text{BP}_{\mathcal{T}} : \mathcal{H}(\mathcal{T}) \to \mathcal{H}(\mathcal{T})$ on the space of messages on $\mathcal{T}$ defined by

$$(\text{BP}h)_{x \to y}(\sigma) \equiv \mathbf{f}_{d-1}\big[(h_{v \to x})_{v \in \partial x \setminus y}\big],$$

(2.1)

$$\big[\mathbf{f}_{d-1}(\underline{h})\big](\sigma) \cong \bar{\psi}(\sigma) \prod_{j=1}^{d-1} \left\{ \sum_{\sigma_j} \psi(\sigma, \sigma_j) h_j(\sigma_j) \right\},$$

where $\underline{h} \equiv (h_1, \ldots, h_{d-1}) \in \boldsymbol{\Delta}^{d-1}$, and $\cong$ denotes equivalence up to a positive normalizing factor which is uniquely determined by the requirement that $\mathbf{f}_{d-1}$ is a mapping $\boldsymbol{\Delta}^{d-1} \to \boldsymbol{\Delta}$.

We write $\mathcal{H}_\star(\mathcal{T}) \subseteq \mathcal{H}(\mathcal{T})$ for the set of all fixed points of $\text{BP}_{\mathcal{T}}$. Since elements of $\mathcal{T}$ are regarded modulo isomorphism, all the incoming messages $h_{v \to x}$ in (2.1) must be the same; and if $\mathcal{T} = \{\mathbb{T}\}$ then they must also coincide with the output message $(\text{BP}h)_{x \to y}$. Thus, $_{\text{nb}}\mathcal{H}_\star \equiv \mathcal{H}_\star(\{\mathbb{T}\})$ corresponds simply to the fixed points of the mapping $\mathbf{f} : \boldsymbol{\Delta} \to \boldsymbol{\Delta}$, $h \mapsto \mathbf{f}_{d-1}(h, \ldots, h)$. If $\mathcal{T} = \{\mathbb{T}_\pm\}$, then the mappings $h \in \mathcal{H}(\mathcal{T})$ are allowed to take two different values $h_s \equiv h(\mathbb{T}_s, o \to j)$ in $\boldsymbol{\Delta}$,

---

[3]The discussion in this subsection is fairly general; in particular $\mathscr{X}$ is not required here to be binary.

which must then satisfy $h_{\mathrm{s}} = \mathbf{f}(h_{-\mathrm{s}})$: thus ${}_{\mathrm{b}}\mathcal{H}_\star \equiv \mathcal{H}_\star(\{\mathbb{T}_\pm\})$ corresponds to the fixed points of the *double* recursion $\mathbf{g} \equiv \mathbf{f}^{(2)} \equiv \mathbf{f} \circ \mathbf{f}$.

The *Bethe free energy functional* on $\mathcal{H}(\mathcal{T})$ is defined by

$$\Phi_\mathcal{T}(h) \equiv \mathbb{E}_\mathcal{T}\big[\Phi_T(h)\big],$$

where Φ_T is a difference of "vertex" and "edge" terms, $\Phi_T \equiv \Phi_T^{\mathrm{vx}} - \Phi_T^{\mathrm{e}}$ with

$$\Phi_T^{\mathrm{vx}}(h) \equiv \log\bigg\{ \sum_{\sigma_o} \bar\psi(\sigma_o) \prod_{j \in \partial o} \bigg(\sum_{\sigma_j} \psi(\sigma_o, \sigma_j) h_{j \to o}(\sigma_j) \bigg) \bigg\},$$

$$\Phi_T^{\mathrm{e}}(h) \equiv \frac{1}{2} \sum_{j \in \partial o} \log\bigg\{ \sum_{\sigma_o, \sigma_j} \psi(\sigma_o, \sigma_j) h_{o \to j}(\sigma_o) h_{j \to o}(\sigma_j) \bigg\}.$$

DEFINITION 2.3.   For any homogeneous spin system on $G_n \to_{\mathrm{loc}} \mathbb{P}_\mathcal{T}$, the *Bethe prediction* is that the free energy density ϕ of (1.4) exists and equals

(2.2) $\Phi_\mathcal{T} \equiv \sup\{\Phi_\mathcal{T}(h) : h \in \mathcal{H}_\star(\mathcal{T})\}.$

We hereafter write ${}_{\mathrm{nb}}\Phi \equiv \Phi_\mathbb{T}$ and ${}_{\mathrm{b}}\Phi \equiv \Phi_{\mathrm{b}T} \equiv \frac{1}{2}[\Phi_+ + \Phi_-]$ with $\Phi_{\mathrm{s}} \equiv \Phi_{\mathbb{T}_{\mathrm{s}}}$.

We emphasize that the Bethe prediction depends on the limiting measure $\mathbb{P}_\mathcal{T}$: fixed points of $\mathbf{g} \equiv \mathbf{f}^{(2)}$ need not be fixed points of $\mathbf{f}$, so in general ${}_{\mathrm{nb}}\mathcal{H}_\star$ is a subset of ${}_{\mathrm{b}}\mathcal{H}_\star$, meaning that the Bethe prediction ${}_{\mathrm{b}}\Phi$ for bipartite graph sequences can be strictly larger than the prediction ${}_{\mathrm{nb}}\Phi$ for nonbipartite sequences. Figure 1 illustrates the hard-core Bethe recursion, which can be expressed as the univariate mapping $\mathbf{f}(q) = [1 + \lambda q^{d-1}]^{-1}$ with $q \equiv h(-)$ for $h \in \boldsymbol{\Delta}$. The function $\mathbf{f}$ is decreasing on $0 \le q \le 1$, so ${}_{\mathrm{nb}}\mathcal{H}_\star$ is always a singleton corresponding to the unique solution $q_\star$ of $\mathbf{f}$. However, the double recursion $\mathbf{g}$ is an increasing function of y,
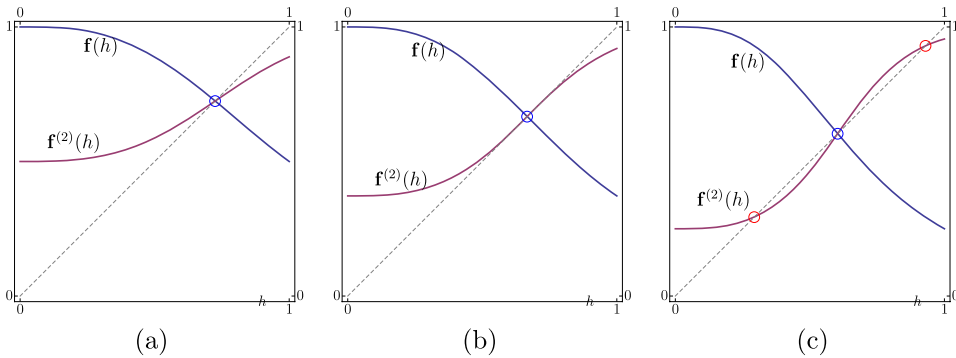

FIG. 1.   *Hard-core Bethe recursion $\mathbf{f} : h(-) \mapsto [1 + \lambda h(-)^{d-1}]^{-1}$ (blue curve, decreasing) shown together with double recursion $\mathbf{f}^{(2)}$ (red curve, increasing) for $d = 4$, with fixed points circled. For $\lambda \le \lambda_{\mathrm{c}}$ ${}_{\mathrm{nb}}\mathcal{H}_\star = {}_{\mathrm{b}}\mathcal{H}_\star$, while for $\lambda > \lambda_c$ ${}_{\mathrm{nb}}\mathcal{H}_\star \subsetneq {}_{\mathrm{b}}\mathcal{H}_\star$. (a) $\lambda = 1$, (b) $\lambda = \lambda_c(4) = \frac{27}{16}$, (c) $\lambda = 3$.*
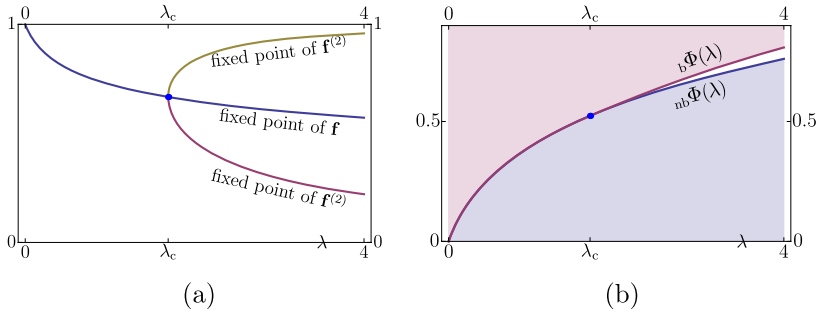
FIG. 2.   *Bethe fixed points $h \in {}_b\mathcal{H}_\star$ and the corresponding evaluations of the Bethe functional ${}_b\Phi(h)$, shown as functions of λ in hard-core model with $d = 4$. Above the uniqueness threshold λ_c, ${}_{nb}\mathcal{H}_\star$ corresponds to the blue curve, and the Bethe prediction ${}_b\Phi$ for $G_n \to_{loc} {}_b\mathbf{T}$ (boundary of shaded red region) is larger than the prediction ${}_{nb}\Phi$ for $G_n \to_{loc} \mathbb{T}$ (boundary of shaded blue region). (a) $h(-)$ evaluated for $h \in {}_b\mathcal{H}_\star$, (b) $\Phi(h)$ evaluated for $h \in {}_b\mathcal{H}_\star$.*

and for $\lambda > \lambda_c$ it turns out that $\mathbf{g}$ has two additional fixed points $q_- < q_\star < q_+$ [shown in Figure 1(c) for a fixed value of λ, and in Figure 2(a) as λ varies], so that ${}_b\mathcal{H}_\star$ contains two additional solutions not in ${}_{nb}\mathcal{H}_\star$. The explicit evaluation of the hard-core Bethe functional on ${}_{nb}\mathcal{H}_\star$ is

$$
{}_{nb}\Phi(h) = \log(\lambda q^d + 1) - \frac{d}{2}\log[1 - (1-q)^2] = -\log q - \frac{d}{2}\log[1 - (1-q)^2]
$$

$$
\text{for } h \in {}_{nb}\mathcal{H}_\star, q = h(-).
$$

For $h \in {}_b\mathcal{H}_\star$, recalling the notation $h_s \equiv h(\mathbb{T}_s, o \to j)$, the Bethe functional evaluates to

$$
{}_b\Phi(h) = \frac{1}{2}[\Phi_+(h) + \Phi_-(h)]
$$

(2.3)
$$
\text{where } \Phi_s(h) = -\log y_{-s} - \frac{d}{2}\log[1 - (1-y_s)(1-y_{-s})]
$$

$$
\text{for } h \in {}_b\mathcal{H}_\star, y_s \equiv h_s(-).
$$

If $h \in {}_{nb}\mathcal{H}_\star$, then ${}_{nb}\Phi(h) = {}_b\Phi(h)$. Figure 2(b) shows that ${}_b\Phi$ strictly exceeds ${}_{nb}\Phi$ for $\lambda > \lambda_c$.

In the course of proving Theorem 4, we will identify the fixed points attaining the supremum in (2.2). Let $g_s \in \Delta$ denote the fixed point of $\mathbf{f}^{(2)}$ giving maximal probability to spin s, and note $\mathbf{f}(g_s) = g_{-s}$. In the anti-ferromagnetic case, we will see that

$$
{}_b\Phi = {}_b\Phi(h^+) = {}_b\Phi(h^-)
$$

(2.4)
$$
\text{for messages } h^r \in {}_b\mathcal{H}_\star \qquad \text{defined by } h^r : (\mathbb{T}_s, o \to j) \mapsto g_{sr}.
$$

The ferromagnetic case reduces to the Ising model, and here we will see that

(2.5)
$$
{}_{nb}\Phi = {}_{nb}\Phi\Big(\arg\max_{h \in {}_{nb}\mathcal{H}_\star} h_{o \to j}(\operatorname{sgn} B)\Big).
$$

2.3. *Bethe Gibbs measures and interpolation scheme.* We now describe the connection between Bethe fixed points and Gibbs measures; for a discussion in a more general setting and further references, see [9], Remark 2.6. Recall that for $T \in \mathcal{T}$, $\mathscr{G}_T$ denotes the set of Gibbs measures for the specification $\underline{\psi}$ on T. An element $\mu \in \mathscr{G}_T$ is a *Markov chain* or *splitting Gibbs measure* (see [25]) if there exists a collection $h^\mu \equiv (h^\mu_{x \to y})$ of elements of $\mathbf{\Delta}$ indexed by the oriented edges of T such that for any finite connected induced subgraph $U = (V_U, E_U)$ of T, the marginal law of $\underline{\sigma}_U$ under μ is given by

$$
\mu(\underline{\sigma}_U) = z^{-1} \prod_{i \in V_U} \bar{\psi}(\sigma_i) \prod_{(ij) \in E_U} \psi(\sigma_i, \sigma_j)
$$

(2.6)

$$
\times \prod_{j \in \partial U} \left\{ \sum_{\sigma_j} \psi(\sigma_{p(j)}, \sigma_j) h^\mu_{j \to p(j)}(\sigma_j) \right\},
$$

where $p(j)$ denotes the unique neighbor of j inside U for j belonging to the external boundary ∂U of U. We emphasize that in contrast to the messages of Definition 2.1, the measures $h^\mu_{x \to y}$ need not be isomorphism-invariant. However, the finite-dimensional marginals of μ given by (2.6) must be consistent with one another, and this imposes relations on the measures $h^\mu_{x \to y}$ which closely resemble the Bethe fixed-point equation (2.1). Extremal Gibbs measures are Markov chains, but the converse is false. Also, for any distinct Markov chains μ, μ', any mixture $p\mu + (1 - p)\mu'$ with $0 < p < 1$ is not a Markov chain ([25], Theorem 4.4).

Next, recall that $\mathscr{G}_\mathcal{T}$ denotes the space of mappings $\nu : T \mapsto \nu(T) \in \mathscr{G}_T$ with $T \in \mathcal{T}$. We say that an element $\nu \in \mathscr{G}_\mathcal{T}$ is *Markovian* if $\nu(T)$ is a Markov chain for each $T \in \mathcal{T}$. The associated collection $h^\nu \equiv (h^{\nu(T)})_{T \in \mathcal{T}}$ is called an *entrance law*; the correspondence between Markovian $\nu \in \mathscr{G}_\mathcal{T}$ and entrance laws h^ν is bijective.

If ν is also *translation-invariant* in the sense of Definition 1.7, then each $h^{\nu(T)}_{x \to y}$ can depend only on the isomorphism class of $(T, x \to y)$ in $\mathcal{T}_\mathrm{e}$, so h is a message on $\mathcal{T}$ (Definition 2.1). It then follows from the consistency of the finite-dimensional marginals (2.6) that h satisfies the Bethe recursions (2.1), that is, $h \in \mathcal{H}_\star(\mathcal{T})$. Thus, there is a bijective correspondence

(2.7)
$$
\text{Bethe fixed points } h \in \mathcal{H}_\star(\mathcal{T})
$$
$$
\longleftrightarrow \quad \text{translation-invariant Markovian } \nu^h \in \mathscr{G}_\mathcal{T}.
$$

In particular, for two-spin models, the $\nu^\mathrm{r} \in {}_\mathrm{b}\mathscr{G}$ of (1.6) and the $h^\mathrm{r} \in {}_\mathrm{b}\mathcal{H}_\star$ of (2.4) are related by this correspondence and so may be regarded as essentially equivalent.

We now evaluate the hard-core and Ising free energy densities by interpolating in the model parameters. Write $\xi \equiv \log \psi$, $\bar{\xi} \equiv \log \bar{\psi}$, and for the hard-core model take $B \equiv \log \lambda$. Let $\langle \rangle^{\beta, B}_n$ denote expectation with respect to the measure $\nu^{\beta, B}_n \equiv \nu^{\beta, B}_{G_n}$ as defined by (1.2) or (1.3). Recalling Definition 1.1 that I_n denotes a

uniformly random vertex in G_n, define

$$\bar{\mathbf{a}}_n(\beta, B) \equiv \partial_B \phi_n(\beta, B) = \mathbb{E}_n \big[\langle \partial_B \bar{\xi}(\sigma_{I_n}) \rangle_n^{\beta, B} \big],$$

$$\mathbf{a}_n^{\mathrm{e}}(\beta, B) \equiv \partial_\beta \phi_n(\beta, B) = \frac{1}{2} \mathbb{E}_n \bigg[\sum_{j \in \partial I_n} \langle \partial_\beta \xi(\sigma_{I_n}, \sigma_j) \rangle_n^{\beta, B} \bigg].$$

In the hard-core model, $\bar{\mathbf{a}}_n(\beta, B)$ is the average occupation $\mathbb{E}_n[\langle \bar{\sigma}_{I_n} \rangle_n^B]$, and we define $\partial_\beta \xi \equiv 0$ so that $\mathbf{a}_n^{\mathrm{e}} \equiv 0$. In the Ising model, $\bar{\mathbf{a}}_n(\beta, B)$ is the average magnetization $\mathbb{E}_n[\langle \sigma_{I_n} \rangle_n^{\beta, B}]$ while $\bar{\mathbf{a}}_n(\beta, B)$ is the average edge correlation $\frac{1}{2} \mathbb{E}_n[\sum_{j \in \partial I_n} \langle \sigma_{I_n} \sigma_j \rangle_n^{\beta, B}]$.

We also define analogous quantities on the limiting tree $T \sim \mathbb{P}_{\mathcal{T}}$. By standard compactness arguments (see, e.g., [19], Section 3.1), $\mathbb{P}_n \otimes \nu_n$ has subsequential limits. If $\mathbb{P}_n \otimes \nu_n \to_{\mathrm{loc}} \mathbb{P}_{\mathcal{T}} \otimes \nu$ along any subsequence, then ν must be a translation-invariant element of $\mathscr{G}_{\mathcal{T}}$ (see Definition 1.8). Writing $\langle\,\rangle_\nu$ for expectation with respect to ν, the tree analogues of $\bar{\mathbf{a}}_n, \mathbf{a}_n^{\mathrm{e}}$ are given by

$$\bar{\mathbf{a}}_{\mathcal{T}}(\beta, B, \nu) \equiv \mathbb{E}_{\mathcal{T}} \big[\langle \partial_B \bar{\xi}(\sigma_o) \rangle_\nu^{\beta, B} \big],$$

$$\mathbf{a}_{\mathcal{T}}^{\mathrm{e}}(\beta, B, \nu) \equiv \frac{1}{2} \mathbb{E}_{\mathcal{T}} \bigg[\sum_{j \in \partial o} \langle \partial_\beta \xi(\sigma_o, \sigma_j) \rangle_\nu^{\beta, B} \bigg].$$

If $\nu = \nu^h$ corresponds to the Bethe fixed point $h \in \mathcal{H}_\star^{\beta, B}(\mathcal{T})$ via (2.7), then we write $\bar{\mathbf{a}}_{\mathcal{T}}(\beta, B, h) \equiv \bar{\mathbf{a}}_{\mathcal{T}}(\beta, B, \nu^h)$ and $\mathbf{a}_{\mathcal{T}}^{\mathrm{e}}(\beta, B, h) \equiv \mathbf{a}_{\mathcal{T}}^{\mathrm{e}}(\beta, B, \nu^h)$.

The following lemma, describing our interpolation scheme, may be verified directly by calculus, or obtained as a consequence of [9], Proposition 2.4. We always interpolate in one parameter at a time, keeping the other fixed and suppressing it from the notation.

LEMMA 2.4. *If for $B \in [B_0, B_1]$, we have $h \equiv h(B) \in \mathcal{H}_\star^B$ which is continuous and of bounded total variation in B, then*

$$\Phi_{\mathcal{T}}(B_1, h) - \Phi_{\mathcal{T}}(B_0, h) = \int_{B_0}^{B_1} \bar{\mathbf{a}}_{\mathcal{T}}(B, h)\, dB \qquad \text{where } \Phi(B, h) \equiv \Phi\big(B, h(B)\big).$$

The same holds for $B, \bar{\mathbf{a}}_{\mathcal{T}}$ replaced with $\beta, \mathbf{a}_{\mathcal{T}}^{\mathrm{e}}$.

The main implication of Lemma 2.4 is the following (which may also be obtained as a special case of [9], Theorem 1.13): if for $B \in [B_0, B_1]$ we have $h \equiv h(B) \in \mathcal{H}_\star^B$ which is continuous and of bounded total variation in B, then

$$\limsup_{n \to \infty} \bar{\mathbf{a}}_n(B) \le \bar{\mathbf{a}}_{\mathcal{T}}(B, h)$$

(2.8)

$$\text{implies } \limsup_{n \to \infty} [\phi_n(B_1) - \phi_n(B_0)] \le \Phi_{\mathcal{T}}(B_1, h) - \Phi_{\mathcal{T}}(B_0, h).$$

That is, asymptotic bounds on $\phi_n(B_1) - \phi_n(B_0)$ can be proved by relating $\bar{\mathbf{a}}_n$, the expectation of the local observable $\partial_B \bar{\xi}(\sigma_{I_n})$ under the measure ν_n on the finite graph, to $\bar{\mathbf{a}}_{\mathcal{T}}$, the expectation of the observable $\partial_B \bar{\xi}(\sigma_o)$ under translation-invariant Markov chains on the limiting tree. The analogous statements hold with $B, \bar{\mathbf{a}}$ replaced by $\beta, \mathbf{a}^e$. In the following, we write $_{\mathrm{nb}}\bar{\mathbf{a}} \equiv \bar{\mathbf{a}}_{\mathbb{T}}$, $_{\mathrm{b}}\bar{\mathbf{a}} \equiv \bar{\mathbf{a}}_{\mathrm{b}\mathbf{T}}$, and similarly $_{\mathrm{nb}}\mathbf{a}^e$, $_{\mathrm{b}}\mathbf{a}^e$.

The difficulty in proving relations of the type $\limsup_{n \to \infty} \bar{\mathbf{a}}_n \leq \sup_{h \in \mathcal{H}_\star(\mathcal{T})} \bar{\mathbf{a}}_{\mathcal{T}}$ is that translation-invariant Gibbs measures may not have a decomposition in terms of translation-invariant Markov chains. As noted above, if $\mathbb{P}_{\mathcal{T}} \otimes \nu$ is any subsequential local weak limit of $\mathbb{P}_n \otimes \nu_n$, then ν must be a translation-invariant element of $\mathscr{G}_{\mathcal{T}}$, but it need not be a Markov chain. By extremal decomposition, ν is a convex combination of Markov chains, but the measures appearing in the decomposition need not be translation-invariant. Thus, there is no a priori connection between the model on the finite graph model and the Bethe prediction on the infinite tree. Our proof of Theorem 4(b), which occupies the majority of Section 2.4, is based on two observations:

1. In the hard-core model with $B \equiv \log \lambda$, the maximum of $_{\mathrm{b}}\bar{\mathbf{a}}(B, \nu)$ over *all* translation-invariant $\nu \in _{\mathrm{b}}\mathscr{G}$ is achieved by the translation-invariant Markov chains ν^{r} of (1.6), so (2.8) can be used to deduce the upper bound $\limsup_{n \to \infty}[\phi_n(\lambda_1) - \phi_n(\lambda_0)] \leq _{\mathrm{b}}\Phi(\lambda_1) - _{\mathrm{b}}\Phi(\lambda_0)$ for any $0 < \lambda_0 \leq \lambda_1$. Since $\phi = _{\mathrm{b}}\Phi$ is already known for $\lambda \leq \lambda_{\mathrm{c}}$, we obtain $\limsup_{n \to \infty} \phi_n \leq _{\mathrm{b}}\Phi$ for all $\lambda > 0$.

In the anti-ferromagnetic Ising model, the minimum of $_{\mathrm{b}}\mathbf{a}^e(\beta, \nu)$ over *all* translation-invariant $\nu \in _{\mathrm{b}}\mathscr{G}$ is achieved by the ν^{r}, so $\liminf_{n \to \infty}[\phi_n(\beta_1) - \phi_n(\beta_0)] \geq _{\mathrm{b}}\Phi(\beta_1) - _{\mathrm{b}}\Phi(\beta_0)$ for any $\beta_0 \leq \beta_1 \leq 0$. Since $\phi = _{\mathrm{b}}\Phi$ holds trivially for $\beta = 0$, we obtain $\limsup_{n \to \infty} \phi_n \leq _{\mathrm{b}}\Phi$ for all $\beta \leq 0$.

2. For $G_n \to_{\mathrm{loc}} {}_{\mathrm{b}}\mathbf{T}$, the near-bipartite structure of the graphs G_n gives easy lower bounds on the partition function which can be used to see that the limit of $[\liminf_n \phi_n - _{\mathrm{b}}\Phi]$ as $B \to \infty$ (hard-core) or $\beta \to -\infty$ (Ising) is nonnegative, implying that the preceding upper bound must be tight.

2.4. *Verification of hard-core and Ising Bethe predictions.* In the remainder of this paper, we restrict all consideration to $\mathscr{X} = \{\pm\}$.

2.4.1. *Interpolation for hard-core.*

LEMMA 2.5. *For the hard-core model at fugacity λ, the supremum of $\langle \bar{\sigma}_o + d^{-1} \sum_{j \in \partial o} \bar{\sigma}_j \rangle_\mu$ over $\mu \in \mathscr{G}_{\mathbb{T}}$ is achieved precisely by the extremal semi-translation-invariant measures μ_+, μ_-.[4] Consequently, $_{\mathrm{b}}\mathbf{E}[\langle \bar{\sigma}_o \rangle_\nu]$ is strictly maximized over all translation-invariant $\nu \in _{\mathrm{b}}\mathscr{G}$ by the elements ν^+, ν^- defined in (1.6).*

---

[4]Note μ need not be translation-invariant. The lemma is nontrivial only when $\mu_- \neq \mu_+$.

PROOF.   By extremal decomposition, assume without loss that μ is itself an extremal Gibbs measure, hence also a splitting Gibbs measure, with finite-dimensional marginals given by (2.6) with entrance law $h^\mu \equiv (h^\mu_{x \to y})$. In particular, the marginal of μ on the depth-1 subtree $\mathbb{T}^1$ is given by

$$\mu(\bar{\sigma}_o = 1, \underline{\bar{\sigma}}_{\partial o} \equiv 0) = z^{-1} \lambda \prod_{j \in \partial o} h^\mu_{j \to o}(-)$$

and

$$\mu(\bar{\sigma}_o = 0, \underline{\bar{\sigma}}_{\partial o}) = z^{-1} \prod_{j \in \partial o} h^\mu_{j \to o}(\sigma_j)$$

with all other configurations receiving zero measure as they violate the hard-core constraint. Abbreviating $q_j \equiv h^\mu_{j \to o}(-)$, we find $z = 1 + \lambda \prod_{j \in \partial o} q_j$, and so

$$\left\langle \bar{\sigma}_o + d^{-1} \sum_{j \in \partial o} \bar{\sigma}_j \right\rangle_\mu = \frac{\lambda \prod_{j \in \partial o} q_j + d^{-1} \sum_{j \in \partial o}(1 - q_j)}{1 + \lambda \prod_{j \in \partial o} q_j} = 1 - \frac{d^{-1} \sum_{j \in \partial o} q_j}{1 + \lambda \prod_{j \in \partial o} q_j}.$$

For fixed $q \equiv (\prod_{j \in \partial o} q_j)^{1/d}$, it follows from Jensen's inequality that the above is (strictly) maximized by taking all $q_j = q$, therefore,

$$\left\langle \bar{\sigma}_o + d^{-1} \sum_{j \in \partial o} \bar{\sigma}_j \right\rangle_\mu \le 1 - \left[\max_{q_- \le q \le q_+} (q^{-1} + \lambda q^{d-1}) \right]^{-1},$$

where q_- and q_+ are the minimal and maximal achievable values for q: by the discussion following Definition 1.6, they correspond to μ_+ and μ_-, respectively, and thus are the fixed points of $\mathbf{g} \equiv \mathbf{f}^{(2)}$ where $\mathbf{f}(q) = [1 + \lambda q^{d-1}]^{-1}$. Since $q^{-1} + \lambda q^{d-1}$ is convex in q, its maximum can only be attained at the interval endpoints, and by the relation $\mathbf{f}(q_s) = q_{-s}$ the maximum is attained at both endpoints with value $1/q_- + 1/q_+ - 1$. This proves the first statement of the lemma. To conclude, recall that for $\nu \in {}_b\mathscr{G}$,

$$_b\mathbf{E}[\langle \bar{\sigma}_o \rangle_\nu] \equiv \tfrac{1}{2}[\langle \bar{\sigma}_o \rangle_{\nu_+} + \langle \bar{\sigma}_o \rangle_{\nu_-}] \qquad \text{where } \nu_s \equiv \nu(\mathbb{T}_s).$$

If further ν is translation-invariant (Definition 1.7), then $\langle \bar{\sigma}_o \rangle_{\nu_s} = \langle \bar{\sigma}_j \rangle_{\nu_{-s}}$ for $j \in \partial o$, so we find $_b\mathbf{E}[\langle \bar{\sigma}_o \rangle_\nu] = \tfrac{1}{2}\langle \bar{\sigma}_o + d^{-1} \sum_{j \in \partial o} \bar{\sigma}_j \rangle_{\nu_+} = \tfrac{1}{2}\langle \bar{\sigma}_o + d^{-1} \sum_{j \in \partial o} \bar{\sigma}_j \rangle_{\nu_-}$, and the second statement of the lemma follows directly from the first.   □

LEMMA 2.6.   *For the hard-core model, $_{nb}\mathcal{H}^\lambda_\star$ always consists of a single message $h^\star \equiv h^\star(\lambda)$. $_b\mathcal{H}^\lambda_\star$ consists of the messages $h^\star, h^+, h^-$ which coincide for $\lambda \le \lambda_c$ and are distinct for $\lambda > \lambda_c$. The messages are continuous in λ, smooth except possibly at $\lambda = \lambda_c$.*

The supremum in (2.2) is achieved by $h^\star$ for $_{nb}\Phi$, and by h^+, h^- for $_b\Phi$, with $_{nb}\Phi = {}_b\Phi$ for $\lambda \le \lambda_c$ and $_{nb}\Phi < {}_b\Phi$ for $\lambda > \lambda_c$.

PROOF.   In the hard-core model, measures $h \in \Delta$ are naturally parameterized in terms of $q = h(-)$, and so $_{\mathrm{nb}}\mathcal{H}_\star$ corresponds to fixed points of $\mathbf{f}(q) = (1 + \lambda q^{d-1})^{-1}$ while $_{\mathrm{b}}\mathcal{H}_\star$ corresponds to fixed points of $\mathbf{g} \equiv \mathbf{f}^{(2)}$. As q increases from 0 to 1, $\mathbf{f}$ decreases from 1 to $(1 + \lambda)^{-1}$, so $\mathbf{f}$ has a unique fixed point $q_\star$ which is smoothly decreasing in λ for all $\lambda > 0$.

Next we determine the messages in $_{\mathrm{b}}\mathcal{H}_\star$, corresponding to fixed points of the double recursion $\mathbf{g} \equiv \mathbf{f}^{(2)}$. We calculate

$$\mathbf{f}'(q) = -\frac{d-1}{q}\mathbf{f}[1 - \mathbf{f}], \qquad \mathbf{f}''(q) = -\left(\frac{d-1}{q}\right)^2 \mathbf{f}(1 - \mathbf{f})\left(2\mathbf{f} - \frac{d}{d-1}\right)$$

and use these to simplify the expression for $\mathbf{g}''(q) = (\mathbf{f}' \circ \mathbf{f})\mathbf{f}'' + (\mathbf{f}'' \circ \mathbf{f})(\mathbf{f}')^2$:

$$\mathbf{g}''(q) = -2\left(\frac{d-1}{q}\right)^4 [\mathbf{f}(1 - \mathbf{f})]^2 \mathbf{g}(1 - \mathbf{g})(\mathbf{g} - \mathbf{g}_0),$$

where

$$\mathbf{g}_0 \equiv -\frac{d}{2(d-1)^2}q\left[\mathbf{f}^{-1} - \frac{d-2}{d}(1 - \mathbf{f})^{-1} - (d-1)q^{-1}\right].$$

Since $\mathbf{f}$ is decreasing in q, $\mathbf{g}$ is increasing in q while $\mathbf{g}_0$ is decreasing. Thus, $\mathbf{g}$ can have at most one inflection point, hence at most three fixed points. If $\mathbf{g}$ has any fixed point other than $q_\star$, then necessarily it has exactly two additional fixed points $q_- < q_\star < q_+$ with $\mathbf{f}(q_{\mathrm{s}}) = q_{-\mathrm{s}}$. In this case, the function $\mathbf{g}$ is convex (concave) to the left (right) of its unique inflection point, so $\mathbf{g}'(q_\star) > 1$. But $\mathbf{g}'(q_\star) = \mathbf{f}'(q_\star)^2 = (d-1)^2(1 - q_\star)^2$ is smoothly increasing in λ with $\mathbf{g}'(q_\star) = 1$ precisely at $\lambda = \lambda_{\mathrm{c}}(d)$, so we see $\mathbf{g}$ has a unique fixed point $q_- = q_\star = q_+$ when $\lambda \le \lambda_{\mathrm{c}}$, and when $\lambda > \lambda_{\mathrm{c}}$ it has three fixed points $q_- < q_\star < q_+$, which are smooth on the open interval $\lambda_{\mathrm{c}} < \lambda < \infty$ (being isolated zeroes of a polynomial equation).

We now verify that $\lim_{\lambda \downarrow \lambda_{\mathrm{c}}}(q_{\mathrm{s}} - q_\star) = 0$. Suppose otherwise, so that $q_+ = q_\star + 2\varepsilon_+$ with $\liminf_{\lambda \downarrow \lambda_{\mathrm{c}}} \varepsilon_+ \ge \varepsilon > 0$. It is possible to take a sequence $\lambda \downarrow \lambda_{\mathrm{c}}$ along which the inflection point of $\mathbf{g}$ always lies on the same side of $q_\star$: assume it is $\le q_\star$ (the argument for the $\ge q_\star$ case is symmetric), so that $\mathbf{g}'$ is decreasing on $q \ge q_\star$. Then

$$\mathbf{g}'(q_\star + \varepsilon_+) \ge \frac{1}{\varepsilon}\left[\mathbf{g}(q_\star + 2\varepsilon_+) - \mathbf{g}(q_\star + \varepsilon_+)\right]$$

$$= \frac{1}{\varepsilon}\left[q_\star + 2\varepsilon_+ - q_\star - \varepsilon_+\mathbf{g}'(q_\star)\right]$$

$$= 2 - \mathbf{g}'(q_\star),$$

so we find $\mathbf{g}'(q_\star) \ge \mathbf{g}'(q) \ge 2 - \mathbf{g}'(q_\star)$ for all $q_\star \le q \le q_\star + \varepsilon_+$. At $\lambda = \lambda_{\mathrm{c}}$, this implies $\mathbf{g}'(q) = 1$ for all $q_\star \le q \le q_\star + \varepsilon_+$—contradicting our above observation that $\mathbf{g}''$ can have at most one zero on the interval $0 \le q \le 1$ and proving our claim that $\lim_{\lambda \downarrow \lambda_{\mathrm{c}}}(q_{\mathrm{s}} - q_\star) = 0$.

It remains to identify the Bethe fixed points achieving the supremum in (2.2). It is clear from the above that $_{\mathrm{nb}}\Phi = {}_{\mathrm{nb}}\Phi(h^\star) \leq {}_{\mathrm{b}}\Phi$ with equality for $\lambda \leq \lambda_{\mathrm{c}}$ since $_{\mathrm{nb}}\mathcal{H}_\star = {}_{\mathrm{b}}\mathcal{H}_\star = \{h^\star\}$ in this regime. For $\lambda > \lambda_{\mathrm{c}}$, we have $_{\mathrm{b}}\bar{\mathbf{a}}(B, h^\star) < {}_{\mathrm{b}}\bar{\mathbf{a}}(B, h^+) = {}_{\mathrm{b}}\bar{\mathbf{a}}(B, h^-)$ as a special case of Lemma 2.5. From the above calculations, the regularity conditions of Lemma 2.4 are clearly satisfied by $h^\star, h^+, h^-$, so we may integrate to see that $_{\mathrm{nb}}\Phi < {}_{\mathrm{b}}\Phi = {}_{\mathrm{b}}\Phi(h^+) = {}_{\mathrm{b}}\Phi(h^-)$ for $\lambda > \lambda_{\mathrm{c}}$, concluding the proof. $\square$

PROPOSITION 2.7. *For the hard-core model,*

(a) *If $G_n \to_{\mathrm{loc}} \mathbb{T}$ then $\phi = {}_{\mathrm{nb}}\Phi = {}_{\mathrm{b}}\Phi$ for $\lambda \leq \lambda_{\mathrm{c}}$ and $\limsup_{n\to\infty} \phi_n \leq {}_{\mathrm{b}}\Phi$ for $\lambda > \lambda_{\mathrm{c}}$.*

(b) *If $G_n \to_{\mathrm{loc}} {}_{\mathrm{b}}\mathbf{T}$ then $\phi = {}_{\mathrm{b}}\Phi$ for all $\lambda > 0$.*

PROOF. (a) Since any subsequential local weak limit $\mathbb{P}_n \otimes \nu_n \to_{\mathrm{loc}} {}_{\mathrm{b}}\mathbf{P} \otimes \nu$ must have ν a translation-invariant element of $_{\mathrm{b}}\mathscr{G}$, the second part of Lemma 2.5 implies

$$\limsup_{n\to\infty} \bar{\mathbf{a}}_n(B) = \limsup_{n\to\infty} \mathbb{E}_n\big[\langle \bar{\sigma}_{I_n} \rangle_n^B\big] \leq {}_{\mathrm{b}}\mathbf{E}\big[\langle \bar{\sigma}_o \rangle_{\nu^{\mathrm{s}}}\big] \equiv {}_{\mathrm{b}}\bar{\mathbf{a}}(B, h^{\mathrm{s}}),$$

where the right-hand side is the same for $\mathrm{s} = +, -$. (We comment that this inequality can be alternatively obtained by expressing $\bar{\mathbf{a}}_n(B)$ as $o(1) + \frac{1}{2}\mathbb{E}_n[\langle \bar{\sigma}_{I_n} + d^{-1}\sum_{j\in\partial I_n} \bar{\sigma}_j \rangle_n^B]$, then directly applying the first part of Lemma 2.5 to the local weak limit $\mathbb{P}_n \otimes \nu_n \to_{\mathrm{loc}} \mathbb{P} \otimes \mu$, where μ belongs to $\mathscr{G}_{\mathbb{T}}$ rather than $_{\mathrm{b}}\mathscr{G}$.)

Recall from Lemma 2.6 that the regularity conditions of Lemma 2.4 are satisfied by h^+, h^-, so the above implies [cf. (2.8)] that for $\lambda > \lambda_{\mathrm{c}}$, $\mathrm{s} \in \mathscr{X}$,

$$\limsup_{n\to\infty}\big[\phi_n(\lambda) - \phi_n(\lambda_{\mathrm{c}})\big] \leq {}_{\mathrm{b}}\Phi\big(\lambda, h^{\mathrm{s}}\big) - {}_{\mathrm{b}}\Phi\big(\lambda_{\mathrm{c}}, h^{\mathrm{s}}\big),$$

where the right-hand side is the same for $\mathrm{s} = +, -$. It was shown in [9], Theorem 1.11, that $\phi = {}_{\mathrm{nb}}\Phi = {}_{\mathrm{b}}\Phi(h^+) = {}_{\mathrm{b}}\Phi(h^-)$ for $\lambda \leq \lambda_{\mathrm{c}}$ so the claim follows.[5]

(b) It suffices to show that for $G_n \to_{\mathrm{loc}} {}_{\mathrm{b}}\mathbf{T}$,

$$(2.9) \qquad \liminf_{\lambda\to\infty}\Big[\liminf_{n\to\infty} \phi_n - {}_{\mathrm{b}}\Phi\Big] \geq 0,$$

implying that the upper bound obtained in (a) must be tight. If A_n denotes the maximum size of an independent set on G_n, then accounting for all subsets of the maximum independent set shows that the hard-core partition function on G_n is at least $(\lambda + 1)^{A_n}$, so $\phi_n \geq \frac{1}{n}\mathbb{E}_n[A_n]\log(\lambda + 1)$. But A_n is at least the number of

---

[5]Note that continuity of h^+, h^- at λ_{c} (Lemma 2.6) was necessary to conclude continuity of $_{\mathrm{b}}\Phi$ at λ_{c}.

black vertices with no black neighbors, so $G_n \to_{\text{loc}} {}_\flat\mathbf{T}$ implies $\liminf_{n\to\infty} \phi_n \geq \frac{1}{2}\log(\lambda + 1)$. From (2.3),

$$
\begin{aligned}
{}_\flat\Phi &- \frac{1}{2}\log(\lambda + 1) \\
&= \frac{1}{2}\log(q_+^d + \lambda^{-1}) - \frac{1}{2}\log(1 + \lambda^{-1}) - \frac{1}{2}\log q_+ \\
&\quad - \frac{d}{2}\log[1 - (1 - q_+)(1 - q_-)].
\end{aligned}
$$

Since $\lim_{\lambda\to\infty} q_+ = 1 = \lim_{\lambda\to\infty}(1 - q_-)$, the above tends to zero as $\lambda \to \infty$. This implies (2.9), concluding the proof. $\square$

We remark that in the entire hard-core interpolation argument the near-bipartite structure plays a role only in the proof of (2.9).

2.4.2. *Interpolation for Ising.*

LEMMA 2.8.   *For the Ising model with parameters $\beta < 0$ and $B \in \mathbb{R}$, the infimum of $d^{-1}\sum_{j\in\partial o}\langle\sigma_o\sigma_j\rangle_\mu$ over $\mu \in \mathscr{G}_{\mathbb{T}}$ is achieved uniquely by the extremal semi-translation-invariant measures μ_+, μ_-. Consequently, ${}_\flat\mathbf{E}[d^{-1}\sum_{j\in\partial o}\langle\sigma_o\sigma_j\rangle_\nu]$ is strictly minimized over all translation-invariant $\nu \in {}_\flat\mathscr{G}$ by the elements ν^+, ν^- defined in (1.6).*

PROOF.   We argue as in the proof of Lemma 2.5: assume μ is extremal with entrance law h^μ, and write $x_j \equiv h_{j\to o}(+)/h_{j\to o}(-)$. Then

$$
\langle\sigma_o\sigma_j\rangle_\mu = \frac{(aR_j - c)e^{2B} + (a/R_j - c)\prod_{k\in\partial o} R_k}{e^{2B} + \prod_{k\in\partial o} R_k},
$$

where

$$
R_j \equiv R(x_j) \equiv \frac{e^{-\beta}x_j + e^\beta}{e^\beta x_j + e^{-\beta}}, \qquad a \equiv \frac{2}{e^{-2\beta} - e^{2\beta}}, \qquad c \equiv \frac{e^{-2\beta} + e^{2\beta}}{e^{-2\beta} - e^{2\beta}}.
$$

Summing over $j \in \partial o$ and simplifying gives

$$
d^{-1}\sum_{j\in\partial o}\langle\sigma_o\sigma_j\rangle_\mu = -c + \frac{ad^{-1}\sum_{j\in\partial o}(R_j e^{2B} + 1/R_j \prod_{k\in\partial o} R_k)}{e^{2B} + \prod_{k\in\partial o} R_k}.
$$

For $\beta < 0$, note that a and c are positive, and $R(x)$ increases between $e^{2\beta}$ and $e^{-2\beta}$ for $x \geq 0$. For fixed $r \equiv (\prod_{j\in\partial o} R_j)^{1/d}$, it follows from Jensen's inequality that the above is (strictly) minimized by taking all $R_j = r$, therefore,

$$
d^{-1}\sum_{j\in\partial o}\langle\sigma_o\sigma_j\rangle_\mu \geq -c + a\min_{R(x_-)\leq r\leq R(x_+)}\gamma(r), \qquad \gamma(r) \equiv \frac{e^{2B}r + r^{d-1}}{e^{2B} + r^d},
$$

where x_-, x_+ are the minimal, maximal achievable values for x. We compute

$$(1 + e^{-2B}r^d)^2\gamma'(r) = -S[S - S^{-1} + (d - 1)(r - r^{-1})]|_{S=e^{-2B}r^{d-1}}.$$

The right-hand side is strictly decreasing in $r > 0$, so γ' can have at most one zero on the positive real line. Using the Bethe fixed-point equation, we evaluate

$$\gamma[R(x_+)] = \frac{e^\beta(x_+ + x_-) + e^{-\beta}(1 + x_+x_-)}{e^{-\beta}(x_+ + x_-) + e^\beta(1 + x_+x_-)} = \gamma[R(x_+)],$$

so on the interval $R(x_-) \le r \le R(x_+)$ the function $\gamma(r)$ is strictly minimized at the endpoints, from which the lemma follows. $\quad\square$

LEMMA 2.9. *For the anti-ferromagnetic Ising model, $_{\mathrm{nb}}\mathcal{H}_\star^{\beta,B}$ always consists of a single message $h^\star \equiv h^\star(\beta, B)$. $_{\mathrm{b}}\mathcal{H}_\star^{\beta,B}$ consists of the messages $h^\star, h^+, h^-$ which coincide for $\beta_{c,af} \le \beta \le 0$ and are distinct for $\beta < \beta_{c,af}$. The messages are continuous in β, B, smooth except possibly at $\beta = \beta_{c,af}$.*

The supremum in (2.2) is achieved by $h^\star$ for $_{\mathrm{nb}}\Phi$, and by h^+, h^- for $_{\mathrm{b}}\Phi$, with $_{\mathrm{nb}}\Phi = {_{\mathrm{b}}\Phi}$ for $\beta_{c,af} \le \beta \le 0$ and $_{\mathrm{nb}}\Phi < {_{\mathrm{b}}\Phi}$ for $\beta < \beta_{c,af}$.

PROOF. In the hard-core model, measures $h \in \mathbf{\Delta}$ are naturally parameterized in terms of $t \equiv \log x \equiv \log[h(+)/h(-)]$, and so $_{\mathrm{nb}}\mathcal{H}_\star$ corresponds to fixed points of

$$(2.10) \qquad \mathbf{f}(t) = 2B + (d - 1)\log[(e^t + \theta)/(\theta e^t + 1)], \qquad \theta \equiv e^{-2\beta}.$$

Observe that $\mathbf{f} - 2B$ is an odd function of $t \in \mathbb{R}$, identically zero when $\beta = 0$ and strictly monotone otherwise, going from $(d - 1)\log\theta$ to $-(d - 1)\log\theta$ as t increases from $-\infty$ to ∞, and with $\partial_\theta\mathbf{f}(t)$ taking the opposite sign as t. If $\beta < 0$, then $\mathbf{f}$ has a unique fixed point $t_\star$ of the same sign as B, smoothly increasing in B, and with absolute value smoothly decreasing in θ.

Next, we determine the messages in $_{\mathrm{b}}\mathcal{H}_\star$, corresponding to fixed points of the double recursion $\mathbf{g} \equiv \mathbf{f}^{(2)}$. For $\beta \ne 0$, we calculate

$$\mathbf{f}'(t) = -\frac{(d - 1)(\theta^2 - 1)}{(\theta + e^t)(\theta + e^{-t})}$$

and

$$\mathbf{f}''(t) = \frac{-\theta(e^t - e^{-t})\mathbf{f}'(t)}{(\theta + e^t)(\theta + e^{-t})} = \frac{(e^t - e^{-t})\mathbf{f}'(t)^2}{(d - 1)(\theta - \theta^{-1})},$$

so

$$\frac{\mathbf{g}''(t)}{-\mathbf{f}'(t)^2\mathbf{f}'(\mathbf{f}(t))} = \frac{\theta(e^{\mathbf{f}} - e^{-\mathbf{f}})}{(\theta + e^{\mathbf{f}})(\theta + e^{-\mathbf{f}})} - \frac{e^t - e^{-t}}{(d - 1)(\theta - \theta^{-1})}.$$

For any $\beta \neq 0$, $\mathbf{g}''(t)$ is strictly decreasing in t, so $\mathbf{g}$ can have at most one inflection point. For $\beta < 0$, we calculate the total derivative of $\mathbf{f}'(t_\star)$ with respect to θ to be

$$\partial_\theta\big[\mathbf{f}'(t_\star)\big] = \big[\partial_\theta(\mathbf{f}')\big](t_\star) + \mathbf{f}''(t_\star)(\partial_\theta t_\star) < 0,$$

implying that $\mathbf{g}'(t_\star) = \mathbf{f}'(t_\star)^2$ is smoothly increasing in θ for $\theta > 1$. Consequently, there is a unique threshold $\beta_{\mathrm{c,af}} \equiv \beta_{\mathrm{c,af}}(d, B) < 0$ such that $\mathbf{g}$ has a single fixed point $t_- = t_\star = t_-$ for $\beta_{\mathrm{c,af}} \leq \beta \leq 0$, and has three fixed points $t_- < t_\star < t_+$ for $\beta < \beta_{\mathrm{c,af}}$. The lemma then follows by repeating the argument of Lemma 2.6. $\quad\square$

PROPOSITION 2.10. *For the anti-ferromagnetic Ising model with external field $B \in \mathbb{R}$:*

(a) *If $G_n \to_{\mathrm{loc}} \mathbb{T}$ then $\phi = {}_{\mathrm{nb}}\Phi = {}_\mathrm{b}\Phi$ for $\beta_{\mathrm{c,af}} \leq \beta \leq 0$ and $\limsup_{n\to\infty} \phi_n \leq {}_\mathrm{b}\Phi$ for $\beta < \beta_{\mathrm{c,af}}$.*
(b) *If $G_n \to_{\mathrm{loc}} {}_\mathrm{b}\mathbf{T}$ then $\phi = {}_\mathrm{b}\Phi$ for all $\beta \leq 0$.*

PROOF. (a) Since any subsequential local weak limit $\mathbb{P}_n \otimes \nu_n \to_{\mathrm{loc}} {}_\mathrm{b}\mathbf{P} \otimes \nu$ must have ν a translation-invariant element of ${}_\mathrm{b}\mathscr{G}$, Lemma 2.8 gives

$$\liminf_{n\to\infty} \mathbf{a}_n^\mathrm{e}(\beta) = \frac{1}{2} \liminf_{n\to\infty} \mathbb{E}_n\bigg[\sum_{j\in\partial I_n} \langle \sigma_{I_n}\sigma_j\rangle_n^\beta \bigg] \geq \frac{1}{2}{}_\mathrm{b}\mathbf{E}\bigg[\sum_{j\in\partial o} \langle \sigma_o\sigma_j\rangle_{\nu^\mathrm{s}} \bigg] = {}_\mathrm{b}\mathbf{a}^\mathrm{e}(\beta, h^\mathrm{s}),$$

where the right-hand side is the same for $\mathrm{s} = +, -$. Recall from Lemma 2.9 that the regularity conditions of Lemma 2.4 are satisfied by h^+, h^-, so the above implies [cf. (2.8)] that for $\beta_0 \leq \beta_1 \leq 0$ and $\mathrm{s} \in \mathscr{X}$,

$$\liminf_{n\to\infty}\big[\phi_n(\beta_1) - \phi_n(\beta_0)\big] \geq {}_\mathrm{b}\Phi(\beta_1, h^\mathrm{s}) - {}_\mathrm{b}\Phi(\beta_0, h^\mathrm{s}),$$

where the right-hand side is the same for $\mathrm{s} = +, -$. If $\beta = 0$, then trivially $\phi = \log(e^B + e^{-B}) = {}_{\mathrm{nb}}\Phi = {}_\mathrm{b}\Phi$, so we conclude $\limsup_{n\to\infty} \phi_n(\beta) \leq {}_\mathrm{b}\Phi(\beta, h^\mathrm{s})$ for all $\beta \leq 0$. For $\beta_{\mathrm{c,af}} \leq \beta \leq 0$, Gibbs uniqueness implies $\lim_{n\to\infty} \mathbf{a}_n^\mathrm{e}(\beta) = {}_{\mathrm{nb}}\mathbf{a}^\mathrm{e}(\beta, h^\star) = {}_\mathrm{b}\mathbf{a}^\mathrm{e}(\beta, h^\star)$ where ${}_{\mathrm{nb}}\mathcal{H}_\star = {}_\mathrm{b}\mathcal{H}_\star = \{h^\star\}$, therefore, $\phi = {}_{\mathrm{nb}}\Phi = {}_\mathrm{b}\Phi$.

(b) By considering the configuration $\underline{\sigma}$ on G_n which gives spin $+$ to all black vertices and spin $-$ to all white vertices, we see clearly that $G_n \to_{\mathrm{loc}} {}_\mathrm{b}\mathbf{T}$ implies $\liminf_{n\to\infty} \phi_n \geq -\frac{d}{2}\beta$. Writing $y_\mathrm{s} \equiv (1 + e^{-t_\mathrm{s}})^{-1}$ for t_-, t_+ the minimal, maximal fixed points of the function $\mathbf{g}$ analyzed in Lemma 2.9, we also calculate

$$\begin{aligned}
2{}_\mathrm{b}\Phi(\beta) + d\beta = {}&{-}d\log\big[1 - (1 - e^{2\beta})[y_+y_- + (1-y_+)(1-y_-)]\big] \\
&+ \log\big[e^B(e^{2\beta}y_+ + (1-y_+))^d + e^{-B}(y_+ + e^{2\beta}(1-y_+))^d\big] \\
&+ \log\big[e^B(e^{2\beta}y_- + (1-y_-))^d + e^{-B}(y_- + e^{2\beta}(1-y_-))^d\big].
\end{aligned}$$

It is straightforward to check that $\lim_{\beta\to-\infty} y_+ = \lim_{\beta\to-\infty}(1 - y_-) = 1$, so in the above the first term tends to 0, the second to $-B$, and the third to $+B$, so we conclude $\lim_{\beta\to-\infty}[2{}_\mathrm{b}\Phi(\beta) + d\beta] = 0$. Therefore, $\lim_{\beta\to-\infty}[\liminf_{n\to\infty} \phi_n -$

$_{\rm b}\Phi(\beta)] \geq 0$, which implies that the upper bound obtained in (a) must be tight and concludes the proof. □

For completeness, we review what is known for the ferromagnetic Ising model.

PROPOSITION 2.11. *For the ferromagnetic Ising model on $G_n \to_{\rm loc} \mathbb{T}$, ϕ exists and equals $_{\rm nb}\Phi$ as defined by (2.2) [and given more explicitly by (2.5)].*

PROOF. It follows from [7], Theorem 2.4 (see also [9], Theorem 1.8) that for $G_n \to_{\rm loc} \mathbb{T}$, ϕ exists and equals Φ as defined by (2.5). Therefore, it remains to verify that the supremum in (2.2) is indeed achieved by the Bethe fixed point $h \in {}_{\rm nb}\mathcal{H}_\star$ maximizing $h_{o \to j}(\operatorname{sgn} B)$.

The messages $h \in {}_{\rm nb}\mathcal{H}_\star$ correspond simply to fixed points of a single iteration of the mapping $\mathbf{f}$ of (2.10). For $\beta > 0$, $\mathbf{f}$ is strictly increasing in t, with a single inflection point at $t = 0$. At $B = 0$, a fixed point is always given by $0 = t_{\rm o} = t_- = t_+$, and it is unique provided $\mathbf{f}'(0) = (d - 1) \tanh \beta < 1$. For $\beta > \beta_{\rm c,f} = \beta_{\rm c,f}(d, B)|_{B=0} = \operatorname{atanh}(\frac{1}{d-1})$, $\mathbf{f}$ has three fixed points $t_- < 0 = t_{\rm o} < t_+$, with $t_+ = -t_- \downarrow 0 = t_{\rm o}$ as $\beta \downarrow \beta_{\rm c,f}$.

Since adding the external field B simply shifts the map $\mathbf{f}$ by the constant $2B$, it is easy to deduce the behavior for general $\beta \geq 0$, $B \in \mathbb{R}$: in the uniqueness regime, $\mathbf{f}$ has a single fixed point $t_- = t_{\rm o} = t_+$, and in the nonuniqueness regime it has three fixed points $t_- < t_{\rm o} < t_+$ which all converge to the same point as (β, B) approaches the boundary of the uniqueness regime. Thus, in the ferromagnetic Ising model $_{\rm nb}\mathcal{H}_\star$ consists of the messages h°, h^+, h^- defined by $h^\circ_{o \to j}(+) = y_\circ \equiv (1 + e^{-t_\circ})^{-1}$ and $h^{\rm s}_{o \to j}(+) = y_{\rm s} \equiv (1 + e^{-t_{\rm s}})^{-1}$ (in the uniqueness regime the messages coincide). The messages are continuous in the parameters (β, B), smooth except possibly at the uniqueness threshold. At $\beta = 0$, clearly $\phi = {}_{\rm nb}\Phi = \log(e^B + e^{-B}) = {}_{\rm nb}\Phi(h)$ for all $h \in {}_{\rm nb}\mathcal{H}_\star$. We now claim that

$$(2.11) \qquad {}_{\rm nb}\Phi(h^+) \geq {}_{\rm nb}\Phi(h^\circ) \vee {}_{\rm nb}\Phi(h^-) \qquad \text{for all } B \geq 0;$$

the proposition follows by symmetry in B. For $h \in {}_{\rm nb}\mathcal{H}_\star$, we have

$${}_{\rm nb}\mathbf{a}^{\rm e}(\beta, B, h) = \frac{1}{2} \sum_{j \in \partial o} \langle \sigma_o \sigma_j \rangle = \frac{(e^\beta + e^{-\beta})[y^2 + (1 - y)^2] - e^{-\beta}}{(e^\beta - e^{-\beta})[y^2 + (1 - y)^2] + e^{-\beta}}$$

$$\text{with } y \equiv h_{o \to j}(+),$$

which is an even function of $2y - 1$, decreasing for $2y \leq 1$ and increasing for $2y \geq 1$. In the nonuniqueness regime, if $B \geq 0$ then $2y_- < 1 < 2y_+$ with $1 - 2y_- \leq 2y_+ - 1$ (with equality if and only if $B = 0$). Therefore, $_{\rm nb}\mathbf{a}^{\rm e}(\beta, B, h^+) \geq {}_{\rm nb}\mathbf{a}^{\rm e}(\beta, B, h^\circ) \vee {}_{\rm nb}\mathbf{a}^{\rm e}(\beta, B, h^-)$ for all $B \geq 0$, so (2.11) follows by interpolation in $\beta \geq 0$ using Lemma 2.4. □

PROOF OF THEOREMS 1 AND 4. Follows by combining the reduction of Section 2.1 with the results of Propositions 2.7, 2.10 and 2.11. □

3. Local structure of measures.   In this section, we show how Theorem 4 can be used to deduce Theorem 5 by straightforward modifications of the arguments of [19].

PROOF OF THEOREM 5(a).   In the hard-core model, observe that $\partial_B^2 \phi_n = n^{-1}\mathbb{E}_n[\langle S^2 \rangle_n^B - (\langle S \rangle_n^B)^2]$ with $S = \sum_{i \in V_n} \bar{\sigma}_i$, implying that the ϕ_n are convex in B, and hence so is the limiting free energy density ϕ (which equals $_b\Phi$ by Theorem 4). Convex functions are absolutely continuous, so it holds for a.e. B that ϕ_n, ϕ are differentiable in B with $\partial_B \phi_n \to \partial_B \phi = \partial_B(_b\Phi)$. From the identification in Lemma 2.6 of the Bethe fixed points achieving the supremum in (2.2), and recalling Lemma 2.4, $\partial_B(_b\Phi) = {_b\bar{\mathbf{a}}}(B, h^s) = {_b\mathbb{E}}[\langle \sigma_o \rangle_{\nu^s}]$ for s $= +, -$. On the other hand, as we have noted before, $\mathbb{P}_n \otimes \nu_n$ must have subsequential local weak limits $_b\mathbf{P} \otimes \nu$ with ν a translation-invariant element of $_b\mathscr{G}$. Along any such subsequence $\partial_B \phi_n = \bar{\mathbf{a}}_n(B)$ must converge to $_b\mathbb{E}[\langle \bar{\sigma}_o \rangle_\nu]$, therefore, $_b\mathbb{E}[\langle \bar{\sigma}_o \rangle_\nu] = {_b\mathbb{E}}[\langle \bar{\sigma}_o \rangle_{\nu^s}]$. Since we saw in Lemma 2.5 that ν^+ and ν^- are the only translation-invariant maximizers for $_b\mathbb{E}[\langle \bar{\sigma}_o \rangle_{\nu^s}]$, we conclude ν must be a convex combination of ν^+, ν^-; and since the G_n are symmetric, necessarily $\nu = \frac{1}{2}(\nu^+ + \nu^-)$. The argument for the anti-ferromagnetic Ising model is very similar, with β in place of B and applying Lemma 2.8 in place of Lemma 2.5.   □

We now analyze the conditional measures ν_n^s of (1.5), beginning with an easy observation.

LEMMA 3.1.   *For anti-ferromagnetic two-spin models on* $G_n \to_{\mathrm{loc}} {_b\mathbf{T}}$,

$$\lim_{n \to \infty} \mathbb{E}_n\left[\nu_n\left(\sum_{i \in V_n} \tau_i \sigma_i = 0\right)\right] = 0.$$

PROOF.   For the Ising model, see [19], Lemma 4.1. For the hard-core model, let A_n denote the set of vertices $i \in V_n$ with $B_2(i)$ isomorphic to $\mathbb{T}_+^2$, the depth-two subtree of $\mathbb{T}_+$; then A_n is necessarily an independent set of black vertices. The law of $X \equiv \sum_{i \in A_n} \bar{\sigma}_i$ under $\nu_n(\cdot|\underline{\sigma}_{V \setminus A_n})$ is $\mathrm{Bin}(N, \lambda/(1+\lambda))$ where $N \equiv \{i \in A_n : \underline{\sigma}_{\partial i} \equiv 0\}$. If $N \geq n\varepsilon$ for ε a small positive constant, then the local CLT or the Berry–Esséen theorem implies $\sup_{j \in \mathbb{Z}} \mathbb{P}(X = j) = O(1/\sqrt{n\varepsilon})$, so we conclude $\nu_n(\sum_{i \in V_n} \tau_i \sigma_i = 0 | N \geq n\varepsilon) = O(1/\sqrt{n\varepsilon})$. If $N \leq n\varepsilon$, decompose

$$\frac{1}{n}\sum_{i \in V_n} \tau_i \sigma_i = \frac{2}{n}\sum_{i \in A_n} \tau_i \bar{\sigma}_i + \frac{2}{n}\sum_{i \in \partial A_n} \tau_i \bar{\sigma}_i + \frac{2}{n}\sum_{i \notin A_n \cup \partial A_n} \tau_i \bar{\sigma}_i - \frac{1}{n}\sum_{i \in V_n} \tau_i$$

$$= \frac{2}{n}\sum_{i \in A_n} \bar{\sigma}_i - \frac{2}{n}\sum_{i \in \partial A_n} \bar{\sigma}_i + \frac{2}{n}\sum_{i \notin A_n \cup \partial A_n} \tau_i \bar{\sigma}_i - \frac{1}{n}\sum_{i \in V_n} \tau_i.$$

On the right-hand side, the first term is $\leq 2\varepsilon$. The second term, recalling the definition of A_n, is $\leq -2(|A_n| - n\varepsilon)/(nd)$, which tends in probability to $-(1 - 2\varepsilon)/d$

since $G_n \to_{\mathrm{loc}} {}_\mathrm{b}\mathbf{T}$. The third and fourth terms tend to zero in probability, so for small ε the overall sum is negative with high probability, implying in particular $\lim_{n\to\infty} \nu_n(\sum_{i\in V_n} \tau_i\sigma_i = 0 | N \le n\varepsilon) = 0$. Combining these observations concludes the proof for the hard-core model. $\square$

In view of Lemma 3.1, we may without loss restrict attention to the measures ν_n^+. Define the local observables [cf. [19], equation (3.9)]

$$F_i^t \equiv F_i^t(\delta, \underline{\sigma}) \equiv \mathbf{1}\left\{ \sum_{j\in B_t(i)} \tau_j\sigma_j \le -\delta |B_t(i)| \right\};$$

roughly speaking F_i^t indicates the vertices of G_n which are locally not in the $+$ phase.

PROOF OF THEOREM 5(b).   We outline the steps of the proof of (1.7) following [19], describing minor modifications where needed.

1. Let $_\mathrm{b}\mathbf{P} \otimes \nu^*$ denote any subsequential local weak limit of the $\mathbb{P}_n \otimes \nu_n^+$. Then $\nu^* \in \mathscr{G}_{\mathcal{T}}$ (see [19], Lemma 3.4). By Lemma 3.1, ν_n^+ has free energy density converging to ϕ, so the proof of Theorem 5(a) implies that $\nu^* = (1-q)\nu^+ + q\nu^-$ for some $q \in [0, 1]$.

2. By local weak convergence, $\lim_{n\to\infty} \mathbb{E}_n[\langle F_{I_n}^t \rangle_n] = {}_\mathrm{b}\mathbf{E}[\langle F_o^t \rangle_{\nu^*}]$; further, if J_n denotes a uniformly random neighbor of I_n, then

$$\lim_{n\to\infty} \mathbb{E}_n[\langle \mathbf{1}\{F_{I_n}^t \ne F_{J_n}^t\} \rangle_n] = {}_\mathrm{b}\mathbf{E}[\langle \mathbf{1}\{F_o^t \ne F_j^t\} \rangle_{\nu^*}], \qquad j \in \partial o$$

(cf. [19], Lemma 3.7).

3. For the hard-core or anti-ferromagnetic Ising model in nonuniqueness regimes, there exists $\delta > 0$ such that

$$\lim_{t\to\infty} \langle F_o^t \rangle_{\nu^+} = 0 = 1 - \lim_{t\to\infty} \langle F_o^t \rangle_{\nu^-},$$

$$\lim_{t\to\infty} \langle F_o^t \ne F_j^t \rangle_{\nu^+} = 0 = \lim_{t\to\infty} \langle F_o^t \ne F_j^t \rangle_{\nu^-}$$

(cf. [19], Lemma 3.8). It follows that for sufficiently large t

$$\lim_{n\to\infty} \mathbb{E}_n[\langle F_{I_n}^t \rangle_n] \ge q - \varepsilon, \qquad \lim_{n\to\infty} \mathbb{E}_n[\langle \mathbf{1}\{F_o^t \ne F_j^t\} \rangle_n] \le \varepsilon.$$

The argument of [19], Proposition 3.9 (using the edge-expansion hypothesis) now gives a contradiction unless $q = 0$, establishing (1.7). The proof of (1.8) then follows from applying the proof of [19], Theorem 2.5, to the bipartite case. $\square$

4. Computational hardness.   In this final section, we prove the hardness results Theorems 2 and 3 by a randomized reduction, via certain bipartite expander gadgets, to the approximate MAX-CUT problem.

4.1. *Bipartite gadgets and randomized reduction.* In this subsection, we construct the bipartite gadgets $\mathscr{G}$ and explain how they can be used to encode an approximate MAX-CUT on a given input graph. The gadgets will be constructed by *deterministic search* over symmetric bipartite d-regular graphs of nondecreasing size. To prove that a gadget with the desired properties can be constructed deterministically in finite time, we shall consider the uniform probability measure on (a large subset of) symmetric bipartite d-regular graphs on $2n$ vertices, and show that the desired properties are satisfied with positive probability in the limit $n \to \infty$. Explicitly, for any fixed positive integer k, G_{2n}^k will be a random bipartite graph on $2n$ vertices, as follows:

1. Take H_n to be the uniformly random simple[6] d-regular graph on vertex set $[n]$.

2. Take G_{2n} to be the bipartite double cover of H_n: the vertex set has bipartition $(i_+)_{i=1}^n$ and $(i_-)_{i=1}^n$, with two edges (i_+, j_-) and (i_-, j_+) corresponding to each edge (i, j) of H_n. This is a simple symmetric bipartite d-regular graph.

3. Distinguish k vertices $(i^\ell)_{\ell=1}^k$ uniformly at random from H_n, and for each ℓ choose a uniformly random neighbor $j^\ell \in \partial i^\ell$. Remove the edges $(i_+^\ell, j_-^\ell)_{\ell=1}^k$ and $(i_-^\ell, j_+^\ell)_{\ell=1}^k$ from G_{2n} to form the simple symmetric bipartite graph G_{2n}^k. All vertices in G_{2n}^k have degree d except the distinguished vertices $W \equiv W^+ \cup W^-$ where $W^{\mathrm{s}} \equiv \{i_{\mathrm{s}}^\ell\}_{\ell=1}^k \cup \{j_{\mathrm{s}}^\ell\}_{\ell=1}^k$.

We abbreviate $\nu_{2n} \equiv \nu_{G_{2n}}$ and $\nu_{2n,k} \equiv \nu_{G_{2n}^k}$. For $\mathrm{r} \in \mathscr{X}$ let ν_{2n}^{r} and $\nu_{2n,k}^{\mathrm{r}}$ denote the corresponding r-phase measures, as in (1.5). For $h, \acute{h} \in \mathbf{\Delta}$ define $h \otimes_\psi \acute{h} \in \mathbf{\Delta}_{\mathscr{X}^2}$ by

$$(4.1) \qquad (h \otimes_\psi \acute{h})(\sigma, \sigma') \equiv z_\psi(h, \acute{h})^{-1}[h(\sigma)\psi(\sigma, \sigma')\acute{h}(\sigma')]$$

with $z_\psi(h, \acute{h})$ the normalizing constant. The following proposition is our main result concerning the local structure of measures on the graphs G_{2n}^k.

PROPOSITION 4.1. *For k fixed and $\mathrm{r} \in \mathscr{X}$, the measures $\nu_{2n,k}^{\mathrm{r}}(\underline{\sigma}_W = \cdot)$ converge in the limit $n \to \infty$ to the product measure*

$$Q^{\mathrm{r}}(\underline{\sigma}) \equiv \prod_{w \in W^+} g_{\mathrm{r}}(\sigma_w) \prod_{w \in W^-} g_{-\mathrm{r}}(\sigma_w) \qquad \text{with } g_+, g_- \in \mathbf{\Delta} \text{ as in (2.4).}$$

We defer the proof to Section 4.2, and now demonstrate how to use Proposition 4.1 to establish a randomized reduction from approximating the partition function to the problem of approximate MAX-CUT on 3-regular graphs, which is NP-hard [4].

---

[6]No self-loops or multiedges.

By Lemma 3.1 and Proposition 4.1, for any $\varepsilon > 0$ there exists $n(\varepsilon)$ large enough such that for all $n \geq n(\varepsilon)$, there is a positive number of graphs G_{2n}^{3k} which arise from the above construction and satisfy the following properties:

(I) G_{2n}^{3k} was formed by removing $6k$ distinct edges from G_{2n}, leaving a set $W \equiv W^+ \cup W^-$ of $12k$ distinct vertices of degree $d - 1$;

(II) $\frac{1}{2}(1 - \varepsilon) \leq v_{2n,3k}(Y(\underline{\sigma}) = \mathrm{s}) \leq \frac{1}{2}(1 + \varepsilon)$ for both $\mathrm{s} \in \mathscr{X}$;

(III) $1 - \varepsilon \leq v_{2n,3k}^{\mathrm{s}}(\underline{\sigma}_W)/Q^{\mathrm{s}}(\underline{\sigma}_W) \leq 1 + \varepsilon$ for both $\mathrm{s} \in \mathscr{X}$, for all $\underline{\sigma}_W \in \mathscr{X}^W$.

Consequently, for given ε we may find G_{2n}^{3k} satisfying properties (I)–(III) within finite time by deterministic search, and define the gadget $\mathscr{G}$ to be the first such graph which is found.

Given a 3-regular input graph $H \equiv (V_H, E_H)$ on m vertices, we construct from H and $\mathscr{G}$ a new graph $H^{\mathscr{G}}$ as follows. First, take $\widehat{H}^{\mathscr{G}}$ be the disjoint union of copies $\mathscr{G}_x$ of $\mathscr{G}$ as x runs over V_H. For $\mathrm{s} \in \mathscr{X}$, let W_x^{s} denote the vertices of $\mathscr{G}_x$ corresponding to W^{s} in $\mathscr{G}$. For each edge $(x, y) \in E_H$, add $2k$ edges between W_x^{s} and W_y^{s} for both $\mathrm{s} \in \mathscr{X}$. From property (I), $|W^+| = |W^-| = 6k$ so this can be done deterministically in such a way that the resulting graph $H^{\mathscr{G}}$ is exactly d-regular.

Write $\mathscr{W}$ for the union of the W_x, and $\mathscr{E}$ for the between-gadget edges added in going from $\widehat{H}^{\mathscr{G}}$ to $H^{\mathscr{G}}$. We write a spin configuration on $\widehat{H}^{\mathscr{G}}$ or $H^{\mathscr{G}}$ as $\underline{\sigma} \equiv (\underline{\sigma}_x)_{x \in H}$ with $\underline{\sigma}_x$ the restriction of $\underline{\sigma}$ to $\mathscr{G}_x$. We write $Y_x \equiv Y(\underline{\sigma}_x)$ for the phase of each $\underline{\sigma}_x$, and $\mathbf{Y}(\underline{\sigma})$ for the vector of phases $(Y(\underline{\sigma}_x))_{x \in H} \in \mathscr{X}^H$. Write $Z_{H^{\mathscr{G}}}(\mathbf{Y})$ for the partition function for the two-spin model on $H^{\mathscr{G}}$ restricted to configurations with phase vector $\mathbf{Y}$, and define likewise $Z_{\widehat{H}^{\mathscr{G}}}(\mathbf{Y})$. Recalling (4.1), let $\Gamma \equiv z_\psi(g_+, g_+) z_\psi(g_-, g_-)$ and $\Theta \equiv z_\psi(g_+, g_-)^2$, and note that for antiferromagnetic two-spin models in nonuniqueness regimes, $\Theta > \Gamma$.

LEMMA 4.2.   *For $\mathscr{G}$ satisfying properties* (I)–(III),

$$\frac{\log((Z_{H^{\mathscr{G}}}/Z_{\widehat{H}^{\mathscr{G}}})/(\Gamma^{2k|E_H|}(1 + \varepsilon)^m))}{2k \log(\Theta/\Gamma)}$$

$$\leq \mathrm{MAX\text{-}CUT}(H)$$

$$\leq \frac{\log((Z_{H^{\mathscr{G}}}/Z_{\widehat{H}^{\mathscr{G}}})/(\Gamma^{2k|E_H|}[(1/2)(1 - \varepsilon)^2]^m))}{2k \log(\Theta/\Gamma)}.$$

PROOF.   From the construction of $H^{\mathscr{G}}$,

$$\frac{Z_{H^{\mathscr{G}}}(\mathbf{Y})}{Z_{\widehat{H}^{\mathscr{G}}}(\mathbf{Y})} = \sum_{\underline{\sigma}_{\mathscr{W}}} \left(\prod_{x \in V_H} v_{\mathscr{G}_x}^{Y_x}(\underline{\sigma}_{W_x}) \right) \left(\prod_{(ij) \in \mathscr{E}} \psi(\sigma_i, \sigma_j) \right)$$

and by property (III) this is within a $(1 \pm \varepsilon)^m$ factor of

$$\sum_{\underline{\sigma}_{\mathscr{W}}} \left(\prod_{x \in V_H} Q^{Y_x}(\underline{\sigma}_{W_x}) \right) \left(\prod_{(ij) \in \mathscr{E}} \psi(\sigma_i, \sigma_j) \right).$$

By a simple calculation, the last expression equals $\Gamma^{2k|E_H|}(\Theta/\Gamma)^{2k\,\mathrm{cut}(\mathbf{Y})}$ where $\mathrm{cut}(\mathbf{Y})$ denotes the number of edges crossing the cut of H induced by $\mathbf{Y}$. Then

$$Z_{H^{\mathscr{G}}} = \sum_{\mathbf{Y}} \frac{Z_{H^{\mathscr{G}}}(\mathbf{Y})}{Z_{\widehat{H}^{\mathscr{G}}}(\mathbf{Y})}$$

$$\times Z_{\widehat{H}^{\mathscr{G}}}(\mathbf{Y}) \begin{cases} \leq \Gamma^{2k|E_H|}(\Theta/\Gamma)^{2k\mathrm{MAX\text{-}CUT}(H)} Z_{\widehat{H}^{\mathscr{G}}}(1+\varepsilon)^m, \\ \geq \Gamma^{2k|E_H|}(\Theta/\Gamma)^{2k\mathrm{MAX\text{-}CUT}(H)} Z_{\widehat{H}^{\mathscr{G}}}\left[\frac{1}{2}(1-\varepsilon)^2\right]^m, \end{cases}$$

where the lower bound uses the fact that $\min_{\mathbf{Y}} Z_{\widehat{H}^{\mathscr{G}}}(\mathbf{Y}) \geq [\frac{1}{2}(1-\varepsilon)]^m Z_{\widehat{H}^{\mathscr{G}}}$ which follows from property (II). Rearranging gives the stated result. $\quad\square$

Using this lemma, we now complete the reduction to approximate MAX-CUT.

PROOF OF THEOREMS 2 AND 3. Let H be a 3-regular graph on m vertices. The upper and lower bounds in Lemma 4.2 differ by $O(m/k)$. Since $\widehat{H}^{\mathscr{G}}$ is a disjoint collection of constant-size graphs, its partition function can be computed in polynomial time. If for any fixed $c > 0$, the partition function $Z_{H^{\mathscr{G}}}$ could be approximated within a factor of $\exp\{c|H^{\mathscr{G}}|\}$ in polynomial time, then both the upper and lower bounds in Lemma 4.2 can be computed up to additive error $O(mc|\mathscr{G}|/k)$ in polynomial time, giving a computation of MAX-CUT(H) up to additive error $O(m(c|\mathscr{G}|+1)/k)$. Then note that MAX-CUT(H) is at least linear in m, since it must exceed the expected value of a random balanced cut which is $3m/[8(1-m^{-1})] \asymp m$—therefore this computes MAX-CUT(H) up to a multiplicative factor $1 + O((c|\mathscr{G}|+1)/k)$.

The error term can be made arbitrarily small: first take k large; the size $|\mathscr{G}|$ can depend on k but then we may choose c to be small. This completes the reduction to a PRAS for MAX-CUT on 3-regular graphs, in contradiction of the result of [4]. $\quad\square$

4.2. *Local structure on bipartite gadgets.* We conclude with the proof of Proposition 4.1 which refines the local structure result Theorem 5 for the bipartite gadgets. The following lemma shows that the G_{2n}^k have with high probability the expansion properties required in Theorem 5(b). The proof is quite standard but is included here for completeness.

LEMMA 4.3. *Let k be fixed. For all $\delta > 0$ there exists $\lambda_\delta > 0$ such that the graphs G_{2n}^k are $(\delta, \frac{1}{2}, \lambda_\delta)$-edge expanders (Definition 1.5) with high probability.*

PROOF. Let us first show that H_n is a $(\delta, 1-\delta, 2\lambda_\delta)$-edge expander with high probability. To this end, let $\mathcal{G}_{n,d}^{\mathrm{cm}}$ be the uniformly random d-regular *multigraph* on vertex set $V = [n]$, generated by the configuration model—that is, the edge

set is given by a perfect matching on the set $[nd]$ of labelled half-edges, where half-edge i is incident to vertex $\lceil i/d \rceil$ (self-loops and multiedges allowed). Then H_n has the law of $\mathcal{G}_{n,d}^{\mathrm{cm}}$ conditioned on the event that $\mathcal{G}_{n,d}^{\mathrm{cm}}$ is a simple graph—and since this event occurs with asymptotically nonnegligible probability (see, e.g., [15], Chapter 9), the claim is proved showing that $\mathcal{G}_{n,d}^{\mathrm{cm}}$ is a $(\delta, 1-\delta, 2\lambda_\delta)$-edge expander with high probability.

Let $\mathscr{S}_n(x, z)$ denote the number of subsets of vertices S in $\mathcal{G}_{n,d}^{\mathrm{cm}}$ of size $|S| = m \equiv nx$ such that there are exactly $j \equiv n\, dz$ edges between S and its complement. With $\mathbb{E}'_n$ denoting expectation with respect to the configuration model law of $\mathcal{G}_{n,d}^{\mathrm{cm}}$, we have

$$\mathbb{E}'_n[\mathscr{S}_n(x, z)] = \binom{n}{m} \frac{1}{(nd-1)!!} \binom{md}{j}\binom{(n-m)\,d}{j}$$
$$\times j!(md-j-1)!!((n-m)\,d-j-1)!!,$$

where the double factorial $(k-1)!! \equiv k!/[(k/2)!2^{k/2}]$ counts the number of matchings on $[k]$. By Stirling's formula, $\mathbb{E}'_n[\mathscr{S}_n(x, z)] = n^{O(1)}\exp\{n[H(x)+dF(x, z)]\}$ where

$$F(x, z) \equiv xH\left(\frac{z}{x}\right) + (1-x)H\left(\frac{z}{1-x}\right) + z\log z + \frac{1}{2}(x-z)\log(x-z)$$
$$+ \frac{1}{2}(1-x-z)\log(1-x-z).$$

Since $\lim_{z\downarrow 0} F(x, z) = -\frac{1}{2}H(x)$ and $d \geq 3$, we see that for any $\delta > 0$ there exists $\lambda_\delta > 0$ such that $\mathbb{E}'_n[\mathscr{S}_n(x, z)]$ is exponentially small in n for all $\delta \leq x \leq 1-\delta$, $0 \leq z \leq 2\lambda_\delta$—implying that with high probability every subset of vertices S in $\mathcal{G}_{n,d}^{\mathrm{cm}}$ with $n\delta \leq |S| \leq n(1-\delta)$ has edge-expansion at least $2\lambda_\delta$. Similarly, note that $\lim_{x\uparrow 1/2}\lim_{z\uparrow x} F(x, z) = -\frac{1}{2}\log 2$, so (adjusting λ_δ as needed) with high probability every subset S in $\mathcal{G}_{n,d}^{\mathrm{cm}}$ with $n(\frac{1}{2}-\delta) \leq |S| \leq n(\frac{1}{2}+\delta)$ has edge-expansion at most $d - 2\lambda_\delta$, that is, has at least $|S|\lambda_\delta$ internal edges.

We now show expansion for G_{2n}^k: since k does not change with n, it suffices to show expansion for the bipartite double cover G_{2n} of $\mathcal{G}_{n,d}^{\mathrm{cm}}$. Let $S_\pm$ be subsets of the $\pm$ sides of G_{2n} such that $S \equiv S_+ \cup S_-$ has size $2n\delta \leq |S| \leq n$, and let π denote the projection from G_{2n} to H_n, so that $|\pi S| \geq \frac{1}{2}|S| \geq n\delta$. The calculation above implies that (with high probability) all such subsets S with $|\pi S| \leq n(1-\delta)$ have edge-expansion at least λ_δ. Suppose instead $|\pi S| \geq n(1-\delta)$: without loss $|S_+| \geq |S_-|$, so $|\pi S_+ \setminus \pi S_-| = |\pi S| - |\pi S_-| \geq n(\frac{1}{2}-\delta)$. By the preceding calculation, $\pi S_+ \setminus \pi S_-$ must have at least $n(\frac{1}{2}-\delta)\lambda_\delta$ internal edges, implying that S must have edge-expansion at least $(\frac{1}{2}-\delta)\lambda_\delta$, concluding the proof. $\square$

PROOF OF PROPOSITION 4.1. Without loss, fix $r = +$. Let B_t denote the union of the balls $B_t(w) \subseteq G_{2n}$ over $w \in \{i_+^\ell\}_{\ell=1}^k \cup \{i_-^\ell\}_{\ell=1}^k$; assume that B_t is

a disjoint union of graphs isomorphic to $\mathbb{T}^t$ with internal boundary $S_t \equiv B_t \setminus B_{t-1}$, which is the case with high probability. By construction, the G_{2n} are symmetric bipartite graphs with law $\mathbb{P}_{2n} \to_{\text{loc}} {}_{\text{b}}\mathbf{P}$. Lemma 4.3 implies that they satisfy with high probability the expansion condition of Theorem 5(b), so we conclude $\mathbb{P}_{2n} \otimes \nu_{2n}^+ \to_{\text{loc}} {}_{\text{b}}\mathbf{P} \otimes \nu^+$. Let $Y_t(\underline{\sigma}) \equiv \text{sgn}(\sum_{i \in V \setminus B_t} \tau_i \sigma_i)$: in the nonuniqueness regime, the concentration result (1.8) implies that $\nu_{2n}(Y(\underline{\sigma}) \ne Y_t(\underline{\sigma})) \to 0$ in probability, so that we also have the convergence $\mathbb{P}_{2n} \otimes \nu_{2n}^{+,t} \to_{\text{loc}} {}_{\text{b}}\mathbf{P} \otimes \nu^+$ where $\nu_{2n}^{+,t}(\cdot) \equiv \nu_{2n}(\cdot | Y_t(\underline{\sigma}) = +)$. In particular, with $\mathbb{E}_{2n}$ denoting expectation over the law $\mathbb{P}_{2n}$ of G_{2n}, we must have

$$(4.2) \qquad \limsup_{t \to \infty} \limsup_{n \to \infty} \mathbb{E}_{2n}\left[\left\| \nu_{2n}^{+,t}((\sigma_{i_s^\ell}, \sigma_{j_{-s}^\ell}) = \cdot) - g_+ \otimes_\psi g_- \right\|_{\text{TV}} \right] = 0.$$

For $s \in \mathscr{X}$ and $\underline{\eta} \in \mathscr{X}^{S_t}$, let $\xi_{t,\underline{\eta}}^{\ell,s}$ and $\zeta_{t,\underline{\eta}}^{\ell,s}$ denote the marginals with respect to $\nu_{2n,k}$ of the spins at i_s^ℓ and j_s^ℓ, respectively, conditioned on configuration $\underline{\eta}$ on S_t:

$$\xi_{t,\underline{\eta}}^{\ell,s} \equiv \nu_{2n,k}(\sigma_{i_s^\ell} = \cdot | \underline{\sigma}_{S_t} = \underline{\eta}),$$
$$\zeta_{t,\ell,\underline{\eta}}^s(\cdot) \equiv \nu_{2n,k}(\sigma_{j_s^\ell} = \cdot | \underline{\sigma}_{S_t} = \underline{\eta}).$$

Notice that ξ, ζ are defined with respect to the measure on G_{2n}^k (i.e., with no direct edge interaction between i_s^ℓ and j_{-s}^ℓ), whereas B_t is defined with respect to the graph structure of G_{2n}. The reason for conditioning on Y_t rather than Y is that we now have the decomposition

$$(4.3) \qquad \nu_{2n}^{+,t}((\sigma_{i_s^\ell}, \sigma_{j_{-s}^\ell}) = \cdot) = \sum_{\underline{\eta}} \nu_{2n}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta}) \left(\xi_{t,\underline{\eta}}^{\ell,s} \otimes_\psi \zeta_{t,\underline{\eta}}^{\ell,-s} \right).$$

Since the local neighborhoods of i_s^ℓ, j_s^ℓ in G_{2n}^k converge to $(d-1)$-ary trees, the marginals $\xi_{t,\underline{\eta}}^{\ell,s}$ and $\zeta_{t,\underline{\eta}}^{\ell,s}$ are asymptotically sandwiched (in the limit $n \to \infty$ followed by $t \to \infty$, uniformly over all $\underline{\eta}$) between g_- and g_+. On the other hand, it is easily seen that the maximum of $(h \otimes_\psi \acute{h})(+, -)$ over all $g_- \preccurlyeq h, \acute{h} \preccurlyeq g_+$ is attained uniquely by $(h, \acute{h}) = (g_+, g_-)$, so the only way for the average (4.3) to be close to $h \otimes_\psi \acute{h}$ is for most of the $\xi_{t,\underline{\eta}}^{\ell,+}$ and $\zeta_{t,\underline{\eta}}^{\ell,-}$ to be close to g_+ and g_-, respectively,

$$\limsup_{t \to \infty} \limsup_{n \to \infty} \mathbb{E}_{2n}\left[\sum_{\underline{\eta}} \nu_{2n}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta}) \right.$$

$$(4.4)$$

$$\left. \times \left(\| \xi_{t,\underline{\eta}}^{\ell,+} - g_+ \|_{\text{TV}} + \| \zeta_{t,\underline{\eta}}^{\ell,-} - g_- \|_{\text{TV}} \right) \right] = 0.$$

We now claim (4.4) continues to hold if $\underline{\eta}$ is averaged with respect to $\nu_{2n,k}^{+,t}$ in place of $\nu_{2n}^{+,t}$; since the marginals $\xi_{t,\underline{\eta}}^{s,\ell}$ and $\zeta_{t,\underline{\eta}}^{s,\ell}$ are $\underline{\eta}$-measurable, it suffices to show

that $v_{2n}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta}) \asymp v_{2n,k}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta})$ uniformly over $\underline{\eta}$. Indeed,

$$(4.5) \qquad \frac{v_{2n}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta})}{v_{2n,k}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta})} = \frac{Z_{\text{out}}^{+,t}(\underline{\eta}) Z_{\text{in}}(\underline{\eta})}{Z_{\text{out}}^{+,t}(\underline{\eta}) Z_{\text{in}}^k(\underline{\eta})} \cdot \frac{\sum_{\underline{\eta}'} Z_{\text{out}}^{+,t}(\underline{\eta}') \bar{\psi}(\underline{\eta}')^{-1} Z_{\text{in}}^k(\underline{\eta}')}{\sum_{\underline{\eta}'} Z_{\text{out}}^{+,t}(\underline{\eta}') \bar{\psi}(\underline{\eta}')^{-1} Z_{\text{in}}(\underline{\eta}')},$$

where $\bar{\psi}(\underline{\eta}) \equiv \prod_{v \in S_t} \bar{\psi}(\eta_v)$ and

$$Z_{\text{out}}^{+,t}(\underline{\eta}) \equiv Z_{G_{2n} \setminus B_{t-1}}\big[\{\underline{\sigma}_{G_{2n} \setminus B_{t-1}} : Y_t(\underline{\sigma}) = + \text{ and } \underline{\sigma}_{S_t} = \underline{\eta}\}\big],$$

$$Z_{\text{in}}(\underline{\eta}) \equiv Z_{B_t}\big[\{\underline{\sigma}_{B_t} : \underline{\sigma}_{S_t} = \underline{\eta}\}\big],$$

$$Z_{\text{in}}^k(\underline{\eta}) \equiv Z_{B_t \cap G_{2n}^k}\big[\{\underline{\sigma}_{B_t} : \underline{\sigma}_{S_t} = \underline{\eta}\}\big].$$

Now note that with k fixed, in the limit $n \to \infty$ followed by $t \to \infty$ we have $Z_{\text{in}}(\underline{\eta}) \asymp Z_{\text{in}}^k(\underline{\eta})$ uniformly over $\underline{\eta}$: for the hard-core model

$$\frac{Z_{\text{in}}(\underline{\eta})}{Z_{\text{in}}^k(\underline{\eta})} = \prod_{\ell=1}^{k} \big([1 - \xi_{t,\ell,\underline{\eta}}^+(1) \zeta_{t,\ell,\underline{\eta}}^-(1)][1 - \xi_{t,\ell,\underline{\eta}}^-(1) \zeta_{t,\ell,\underline{\eta}}^+(1)]\big),$$

which is $\asymp 1$ uniformly over $\underline{\eta}$; a similar argument applies for Ising. Thus, (4.4) continues to hold with $v_{2n,k}^{+,t}$ in place of $v_{2n}^{+,t}$. Since the spins $(\sigma_w)_{w \in W}$ are exactly independent under $v_{2n,k}^{+,t}(\cdot | \underline{\sigma}_{S_t} = \underline{\eta})$ for any $\underline{\eta}$, this further implies

$$(4.6) \qquad \limsup_{t \to \infty} \limsup_{n \to \infty} \mathbb{E}_{2n}\big[\| v_{2n,k}^{+,t}(\underline{\sigma}_W = \cdot) - Q_W^+ \|_{\text{TV}}\big] = 0.$$

Finally, $v_{2n}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta}) \asymp v_{2n,k}^{+,t}(\underline{\sigma}_{S_t} = \underline{\eta})$ implies $v_{2n,k}(Y(\underline{\sigma}) \neq Y_t(\underline{\sigma})) \to 0$ in probability, so (4.6) holds with $v_{2n,k}^+$ in place of $v_{2n,k}^{+,t}$ which proves the result. $\square$

REFERENCES

[1] ACHLIOPTAS, D. and NAOR, A. (2005). The two possible values of the chromatic number of a random graph. *Ann. of Math.* (2) **162** 1335–1351. MR2179732

[2] ACHLIOPTAS, D., NAOR, A. and PERES, Y. (2005). Rigorous location of phase transitions in hard optimization problems. *Nature* **435** 759–764.

[3] ACHLIOPTAS, D., NAOR, A. and PERES, Y. (2007). On the maximum satisfiability of random formulas. *J. ACM* **54** Art. 10, 21 pp. (electronic). MR2295994

[4] ALIMONTI, P. and KANN, V. (1997). Hardness of approximating problems on cubic graphs. In *Algorithms and Complexity* (*Rome*, 1997). *Lecture Notes in Computer Science* **1203** 288–298. Springer, Berlin. MR1488983

[5] BETHE, H. A. (1935). Statistical theory of superlattices. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **150** 552–575.

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

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

[8] DEMBO, A., MONTANARI, A., SLY, A. and SUN, N. (2014). The replica symmetric solution for Potts models on d-regular graphs. *Comm. Math. Phys.* **327** 551–575. MR3183409

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

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

[11] DYER, M., FRIEZE, A. and JERRUM, M. (1999). On counting independent sets in sparse graphs. In 40*th Annual Symposium on Foundations of Computer Science* (*New York*, 1999) 210–217. IEEE Comput. Soc., Los Alamitos, CA. MR1917561

[12] DYER, M., FRIEZE, A. and JERRUM, M. (2002). On counting independent sets in sparse graphs. *SIAM J. Comput.* **31** 1527–1541. MR1936657

[13] 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

[14] GALANIS, A., ŠTEFANKOVIČ, D. and VIGODA, E. (2012). Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. Available at arXiv:1203.2226v2.

[15] JANSON, S., ŁUCZAK, T. and RUCINSKI, A. (2000). *Random Graphs*. Wiley, New York. MR1782847

[16] JERRUM, M. and SINCLAIR, A. (1993). Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.* **22** 1087–1116. MR1237164

[17] LI, L., LU, P. and YIN, Y. (2013). Correlation decay up to uniqueness in spin systems. In *Proceedings of the* 24*th Annual ACM-SIAM Symposium on Discrete Algorithms* (*SODA*) 47–66. Available at arXiv:1111.7064v2.

[18] LUBY, M. and VIGODA, E. (1999). Fast convergence of the Glauber dynamics for sampling independent sets. *Random Structures Algorithms* **15** 229–241. MR1716763

[19] 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

[20] 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

[21] MOTWANI, R. and RAGHAVAN, P. (1995). *Randomized Algorithms*. Cambridge Univ. Press, Cambridge. MR1344451

[22] SINCLAIR, A., SRIVASTAVA, P. and THURLEY, M. (2012). Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms* 941–953. SIAM, Philadelphia, PA.

[23] SLY, A. (2010). Computational transition at the uniqueness threshold. In 2010 *IEEE* 51*st Annual Symposium on Foundations of Computer Science FOCS* 2010 287–296. IEEE Comput. Soc., Los Alamitos, CA. MR3025202

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

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

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY
367 EVANS HALL
BERKELEY, CALIFORNIA 94720
USA
AND
MATHEMATICAL SCIENCES INSTITUTE
AUSTRALIAN NATIONAL UNIVERSITY
JOHN DEDMAN BUILDING 27, UNION LANE
CANBERRA, ACT 0200
AUSTRALIA

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
SEQUOIA HALL, 390 SERRA MALL
STANFORD, CALIFORNIA 94305
USA
E-MAIL: nikesun@stanford.edu