THE REPLICA SYMMETRIC SOLUTION FOR POTTS MODELS ON d-REGULAR GRAPHS

AMIR DEMBO*, ANDREA MONTANARI†, ALLAN SLY‡, AND NIKE SUN§

Abstract. We establish an explicit formula for the limiting free energy density (log-partition function divided by the number of vertices) for ferromagnetic Potts models on uniformly sparse graph sequences converging locally to the d-regular tree for d even, covering *all* temperature regimes. This formula coincides with the Bethe free energy functional evaluated at a suitable fixed point of the belief propagation recursion on the d-regular tree, the so-called replica symmetric solution. For uniformly random d-regular graphs we further show that the replica symmetric Bethe formula is an upper bound for the asymptotic free energy for *any* model with permissive interactions.

1. Introduction

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

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

where ψ is a non-negative symmetric function on $\mathscr{X}^2$, $\bar{\psi}$ is a positive function on $\mathscr{X}$, and $Z_G(\underline{\psi}) \equiv Z_G$ is the normalizing constant, called the *partition function* (with its logarithm called the *free energy*). The pair $\underline{\psi} \equiv (\psi, \bar{\psi})$ is called a *specification* for the factor model (1), and we assume it to be *permissive*, meaning there exists $\sigma^{\mathrm{p}} \in \mathscr{X}$ with $\min_\sigma \psi(\sigma, \sigma^{\mathrm{p}}) > 0$.

A primary example we consider in this paper is the *q-state Potts model* on G with inverse temperature β and magnetic field B, given by specification

$$\psi(\sigma, \sigma') = e^{\beta \mathbf{1}\{\sigma=\sigma'\}}, \quad \bar{\psi}(\sigma) = e^{B\mathbf{1}\{\sigma=1\}}, \quad \mathscr{X} = [q] \equiv \{1, \dots, q\}. \tag{2}$$

We write $\nu_G^{\beta,B}$ for the corresponding measure on $[q]^V$. The model is said to be *ferromagnetic* if $\beta \geqslant 0$, and *anti-ferromagnetic* otherwise. The case $q = 2$ corresponds to the *Ising* model.

In this paper we study the asymptotics of the free energy for factor models (1) on graph sequences $G_n = (V_n, E_n)$ converging locally to the d-regular tree $\boldsymbol{T}_d$ ($d \geqslant 3$) in the sense of Benjamini–Schramm [BS01] (see Defn. 1.1). This class includes in particular any sequence of d-regular graphs with girth (minimal cycle length) diverging to infinity.

Date: December 10, 2013.

2010 *Mathematics Subject Classification.* 82B20, 82B23, 05C80, 60K35.

Key words and phrases. Free energy density, replica symmetry, Gibbs measures, Bethe measures, Potts model, factor models, random graphs, local weak convergence.

Research partially supported by NSF grants *†§DMS-1106627 and †CCF-0743978, ‡Alfred P. Sloan Research Fellowship, and §Department of Defense NDSEG Fellowship.

The study of statistical mechanics on regular trees has a long history going back to Bethe [Bet35]. While tree graphs do not capture the finite-dimensional structure of actual physical systems, models on trees are often amenable to exact analysis. Further, it is often argued that they are a good approximation to models on the lattice $\mathbb{Z}^d$ for large d or for long interaction range [Wei48, ATA73, CLR79, Tho86]. According to this expectation, models on trees provide a flexible and well-defined approach for investigating *mean-field theory* (i.e. the behavior of statistical mechanics models in high dimensions).

While this expectation proves to be correct in a number of examples, it has recently become clear that, in many cases, models on trees fail to capture the "correct" mean-field behavior. Spin glasses provide an important example of this phenomenon: a fairly natural class of spin glasses on trees was introduced by Thouless [Tho86] and further characterized by Chayes et al. [CCST86]. However, the thermodynamic behavior observed there is very different from the widely accepted mean-field theory of spin glasses, as obtained from analysis of the Sherrington–Kirkpatrick (SK) model [MPV87, Tal11]. In particular, the low-temperature phase of the tree models defined in [Tho86] does not exhibit replica symmetry breaking (in contrast with SK). A similar discrepancy was observed in the case of Anderson localization by Aizenman–Warzel [AW06].

In the case of spin glasses, Mézard–Parisi [MP01] argued that this difference arises because of a particular feature of tree graphs: in the subgraph induced by the first ℓ levels of the regular tree, the leaves constitute a non-vanishing fraction of the vertices as $\ell \to \infty$. They suggested that mean-field theory ought instead to be defined by considering graphs that are not themselves trees, but "look like regular trees" in the neighborhood of a typical vertex (which fails for the depth-ℓ subtree of the regular tree) — the canonical example being the (uniformly) random d-regular graph ensemble. This approach allows to reconcile discrepancies in several known cases. In particular, spin glasses on random regular graphs are expected to exhibit replica symmetry breaking with features analogous to the SK model (see [MP01] and [MM09, Ch. 17]).

Let us also mention that the study of statistical mechanics models on locally tree-like graphs has attracted renewed interest because of the connection with random combinatorial problems, such as k-SAT and graph coloring. Statistical physicists were indeed able to compute threshold locations for these models by analyzing suitable Gibbs measures on locally tree-like structures [MPZ02, KMR$^+$07, MM09]. Rigorous verification of these predictions is an outstanding mathematical challenge.

In this paper we consider the existence and value of the *free energy density* (asymptotic free energy per spin)

$$\phi \equiv \lim_{n\to\infty} \phi_n \equiv \lim_{n\to\infty} n^{-1}\mathbb{E}_n[\log Z_n], \quad Z_n \equiv Z_{G_n}(\underline{\psi}), \tag{3}$$

for G_n a (possibly random) graph sequence converging locally to the regular tree and $\mathbb{E}_n$ expectation over the law of G_n. For Ising (specification (2) with $q = 2$) models in the ferromagnetic regime, for *any* graph sequence with uniformly integrable average degree converging locally to a (possibly random) tree, the free energy density (3) exists and depends only on the local limiting tree [DM10b, DGvdH10, DMS13]. The computation of ϕ allows to compute various limits of interest with respect to the ν_{G_n}, as done for example in [MMS12, DGvdH10]. Proving existence of the free energy density for $q \geqslant 2$ and general specification $\underline{\psi}$ poses several challenges:[1]

[1]Existence of (3) for general $\underline{\psi}$ is equivalent to right convergence of G_n in the language of [BCKL13].

1. There are examples in which the free energy density (3) depends not only on the local limiting tree but also on the particular graph sequence. For example, in the anti-ferromagnetic Ising model at sufficiently low temperature (sufficiently negative β), it is not difficult to show that the free energy per spin on random d-regular graphs is asymptotically lower than on random *bipartite* d-regular graphs. As a consequence local weak convergence is not in full generality a sufficient condition for existence of the limit (3).

2. Statistical physicists have put forth a number of conjectures (corresponding to different models or regimes) on the free energy density (3) (see e.g. [MPV87, MM09]). This analysis generally imposes a probability distribution on the graph G_n which is suitable for calculations, typically the Erdős-Renyi or configuration models. Ensuing rigorous work has also focused on the same random graph ensembles (see e.g. [Tal11]) rather than understanding which graph sequences in general have a limit (3). In this paper we focus instead on individual graph sequences.

Characterizing the limit for ensembles of uniformly random graphs is already beyond current techniques for many factor models (1). Achieving the same goal for general locally tree-like graph sequences is all the more difficult, and requires to go beyond what is known from physics methods. A simple example is provided again by the anti-ferromagnetic Ising model: existence of the limit can be proved by a combinatorial interpolation [BGT10], but even a heuristic prediction of the value is unavailable.

In contrast, as mentioned above the free energy density for the *ferromagnetic* Ising model on locally tree-like graphs exists and can be explicitly computed. Its value is given by the (*replica symmetric*) *Bethe free energy prediction* $\mathbf{\Phi}^\star$, which is expressed in terms of solutions of a certain fixed-point equation ((5) and (6) give the prediction in the d-regular setting; for the general case see [DM10a, DMS13]). The Ising Bethe prediction was proved (for all $\beta \geqslant 0$, $B \in \mathbb{R}$) in the case of graphs with Galton–Watson local limiting trees via an interpolation scheme [DM10b]. In subsequent work [DMS13] a generalized scheme was developed which extended this result to graphs with general local limiting trees. This method was applied also to show $\liminf_n \phi_n \geqslant \mathbf{\Phi}^\star$ in the ferromagnetic q-Potts model ($\beta \geqslant 0$, $B \geqslant 0$) for all $q > 2$, but could only pin down $\phi = \mathbf{\Phi}^\star$ in limited regimes of (β, B).

The difficulty of the Potts model ($q > 2$) may be understood as follows: by a monotonicity argument, the local weak limit of Potts measures on G_n is sandwiched between the free and maximally 1-biased automorphism-invariant Gibbs measures on the local limiting tree. For $q = 2$ these measures coincide for all $\beta \geqslant 0, B > 0$.[2] By contrast, for $q > 2$ these measures disagree in certain regimes of (β, B). This corresponds to the appearance of "multiple stable fixed points" in the recursion (5) as soon as $q > 2$, as we demonstrate below in the d-regular setting (§1.3, Figs. 1 and 2).

In this paper we prove the matching upper bound $\limsup_n \phi_n \leqslant \mathbf{\Phi}^\star$, thereby explicitly determining the free energy density (3), for the Potts model for all $q > 2$, $\beta \geqslant 0$, and $B \geqslant 0$, on graphs converging to regular trees of *even* degree. Let us mention that where multiple fixed points arise, the statistical physics folklore prescribes that the fixed point with the *highest Bethe free energy density* should be selected. However, in the physics literature this is justified only via analogy with other models, without providing arguments which apply to locally tree-like graphs. Our result is the first rigorous verification of this variational principle in a non-trivial example for locally tree-like graphs.

[2] In the Ising model, the case $B < 0$ is handled by symmetry, and the result follows for $B = 0$ by continuity.

Additionally, we supply a rigorous probabilistic interpretation for this variational principle in the setting of the *uniformly* random d-regular graph ensemble. In this ensemble we show that $n^{-1} \log \mathbf{E}_n[Z_n]$ (which clearly upper bounds $\phi_n = n^{-1}\mathbf{E}_n[\log Z_n]$) converges exactly to $\mathbf{\Phi}^\star$. In fact, the computation of $\mathbf{E}_n[Z_n]$ can be understood to correspond in a very precise manner to the folklore prescription of maximizing Bethe free energy over all fixed points. Validity of the Bethe prediction $\phi = \mathbf{\Phi}^\star$ for the uniformly random d-regular ensemble therefore requires *concentration* of Z_n over the space of d-regular graphs (since $\mathbb{E}[\log X] \approx \log \mathbb{E}X$ indicates concentration of the positive random variable X). This is consistent with the physics picture that the regimes where the Bethe prediction fails are characterized by non-concentration of Z_n (*replica symmetry breaking*; see [KMR+07]).

A different variational principle was proved in [Gue03, ASS03] for mean-field spin glass models, but in that case the free energy density needs to be *minimized*. This difference is typically attributed by physicists to the difference between ferromagnetic and spin glass models; it remains an outstanding challenge to understand these two variational principles within a common framework. In the context of models on sparse graphs, Contucci et al. [CDGS13] recently proved that the variational principle of [Gue03, ASS03] provides a bound on the free energy of *anti-ferromagnetic* Potts models, which was proved to be tight at high temperature.

The rest of the paper is organized as follows: in the remainder of this introductory section we review the definitions of local convergence and the Bethe prediction and formally state our results, which we divide into two categories: in §2 we study the Bethe prediction on the *uniformly* random d-regular graph ensemble and prove the variational principle. In §3-4 we prove results in the more general setting of graphs converging locally to the d-regular tree. We conclude in §5 with some supplementary results on *local* maximizers in the Potts variational problem.

1.1. **Local convergence.** Throughout this paper, graphs are allowed to have multi-edges and self-loops unless otherwise stated. If G is any graph and U any subgraph, we write ∂U for the external boundary of U in G (the set of vertices in G adjacent to but not contained in U). For any vertex v of G, write $B_t(v)$ for the subgraph induced by the vertices of G at graph distance at most t from v.

Fix d throughout and let $\boldsymbol{T}_d \equiv (\boldsymbol{T}_d, o)$ denote the d-regular tree rooted at o, with $\boldsymbol{T}_{d,t} \equiv B_t(o)$ the subtree of depth t. For $G = (V, E)$ a finite undirected graph, let

$$\zeta_t(G) \equiv |V|^{-1}|\{v \in V : B_t(v) \ncong \boldsymbol{T}_{d,t}\}| \tag{4}$$

where $\cong$ denotes graph isomorphism.

Throughout this paper we consider a sequence of (random) graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$. Write $\mathbb{P}_n$ for the law of G_n, and $\mathbb{E}_n$ for expectation over $\mathbb{P}_n$. We consider graphs which are locally like the d-regular tree in the following sense:

Definition 1.1. We say that $G_n = (V_n = [n], E_n)$ is a *uniformly sparse graph sequence converging locally to the d-regular tree* $\boldsymbol{T}_d$ if the following hold:

(a) *(uniform sparsity)* $\lim_{L\to\infty} |V_n|^{-1} \sum_{v\in V_n} \mathbb{E}_n[|\partial v|\mathbf{1}\{|\partial v| \geqslant L\}] = 0$; and
(b) *(local weak convergence)* for each $t \geqslant 0$, the random variable $\zeta_t(G_n)$ as defined in (4) converges in probability to zero in the limit $n \to \infty$.

This setting, which we assume throughout, is denoted $G_n \to_{loc} \boldsymbol{T}_d$.
With no loss of generality, we assume hereafter $V_n = [n] \equiv \{1, \ldots, n\}$.

Proposition 1.2. *Consider the factor models* (1) *defined by a permissive specification* $\underline{\psi}$, *and suppose* $G_n = (V_n = [n], E_n)$ *is a graph sequence for which the free energy density* (3) *exists,* $\phi = \lim_{n\to\infty} n^{-1}\mathbb{E}_n[\log Z_n]$. *(Here we do not assume* $G_n \to_{loc} \boldsymbol{T}_d$*.)*

(a) If the G_n *have uniformly bounded degree then* $n^{-1}\log Z_n \to \phi$ *almost surely as* $n \to \infty$.
(b) If the G_n *are uniformly sparse (Defn. 1.1a) then* $n^{-1}\log Z_n \to \phi$ *in probability as* $n \to \infty$.

1.2. **Bethe prediction in the regular setting.** We now describe the Bethe free energy prediction in the special setting of a graph sequence converging locally to the d-regular tree; for a more general description see [DM10a, DMS13].

Definition 1.3. Let $\mathscr{H}$ denote the $(|\mathscr{X}|-1)$-dimensional simplex of probability measures on $\mathscr{X}$. The *belief propagation* or *Bethe* recursion is the mapping $\mathsf{BP}:\mathscr{H}\to\mathscr{H}$ defined by

$$(\mathsf{BP}h)(\sigma) \equiv \frac{1}{z_h}\,\bar\psi(\sigma)\Big(\sum_{\sigma'}\psi(\sigma,\sigma')h_{\sigma'}\Big)^{d-1}, \quad \sigma\in\mathscr{X}, \tag{5}$$

with z_h the normalizing constant. Denote by $\mathscr{H}^\star \subseteq \mathscr{H}$ the set of BP fixed points.

Remark 1.4. Note that for permissive $\underline{\psi}$, any fixed point $h\in\mathscr{H}^\star$ must be strictly positive. Let $\acute{\boldsymbol{T}}_d$ denote the $(d-1)$-ary tree (in which every vertex has $d-1$ children), and consider the factor model (1) on the depth-t subtree $\acute{\boldsymbol{T}}_{d,t}$ multiplied with "entrance law" or "boundary law" $h\in\mathscr{H}$ on each spin at depth t. Then $h\in\mathscr{H}^\star$ if and only if the resulting marginal law at the root of $\acute{\boldsymbol{T}}_d$ is again h. Fixed points $h\in\mathscr{H}^\star$ correspond to "Markov chain Gibbs measures" [Spi75, Zac83] (also "splitting Gibbs measures" or "Bethe Gibbs measures") — natural candidates for the local weak limit of ν_{G_n} (see e.g. [DMS13, Rmk. 1.12]).

Definition 1.5. For a factor model (1) on a graph sequence $G_n \to_{loc} \boldsymbol{T}_d$, the *Bethe functional* is the mapping $\Phi:\mathscr{H}\to\mathbb{R}$ defined by

$$\Phi(h) = \underbrace{\log\Big\{\sum_\sigma \bar\psi(\sigma)\Big(\sum_{\sigma'}\psi(\sigma,\sigma')h_{\sigma'}\Big)^d\Big\}}_{\text{“vertex term” } \Phi^{\mathrm{vx}}(h)} - \underbrace{(d/2)\log\Big\{\sum_{\sigma,\sigma'}\psi(\sigma,\sigma')h_\sigma h_{\sigma'}\Big\}}_{\text{“edge term” } \Phi^{\mathrm{e}}(h)}. \tag{6}$$

(See §1.3 for explanation of the $\Phi^{\mathrm{vx}}, \Phi^{\mathrm{e}}$ notation.) The *Bethe free energy prediction* is that

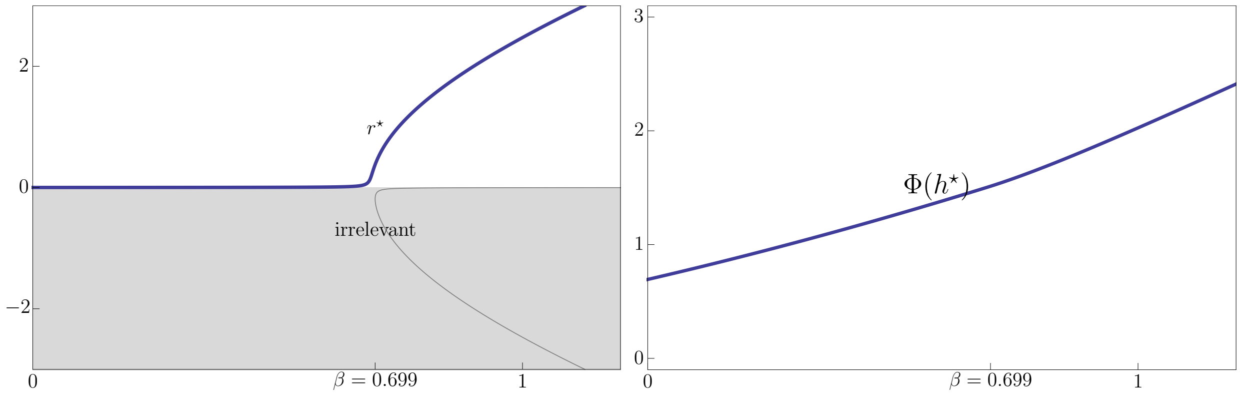
$$\text{the limit } \phi \text{ of (3) exists and equals } \boldsymbol{\Phi}^\star \equiv \sup\{\Phi(h) : h\in\mathscr{H}^\star\}. \tag{7}$$

1.3. **Results for general d-regular graphs.** The following is our main result, establishing the validity of the replica symmetric Bethe solution for ferromagnetic Potts models (2) on *general* graph sequences $G_n \to_{loc} \boldsymbol{T}_d$:

Theorem 1. *For the Potts model* (2) *on* $G_n \to_{loc} \boldsymbol{T}_d$ *with* d *even,* $\phi = \boldsymbol{\Phi}^\star$ *for all* $\beta, B \geqslant 0$.

In [DMS13] we established the lower bound $\liminf_n \phi_n \geqslant \boldsymbol{\Phi}^\star$ for all d. In this paper we prove the matching upper bound $\limsup_n \phi_n \leqslant \boldsymbol{\Phi}^\star$ for d even.

The result for $q=2$ (the Ising model) is a special case of the results of [DM10a], and is depicted in Fig. 1 (the Ising result also holds for d odd). Let us comment on some key differences between the Ising model and the Potts models with $q>2$. For the Ising model, the space $\mathscr{H}$ is one-dimensional, so the Ising belief propagation is a univariate recursion, which is straightforward to analyze. For the Potts model with $q>2$, $\mathscr{H}$ has dimension larger than one: the belief propagation mapping becomes more difficult to analyze; and indeed in §5 we exhibit a multiplicity of fixed points $h\in\mathscr{H}^\star$.



(A) Bethe fixed points (B) Bethe functional evaluation

FIGURE 1. *Ising model ($q = 2$) with $d = 4$, $B = 1/1000$, β on horizontal axis.* The left panel shows the Bethe fixed points $h \in \mathscr{H}^\star$, parametrized by $r \equiv \log(h_1/h_2)$. Solutions $r < 0$ (gray curve) are shown to be irrelevant since the model favors spin 1. The right panel shows the corresponding evaluation of the Bethe function, $\boldsymbol{\Phi}^\star = \Phi(h^\star)$.

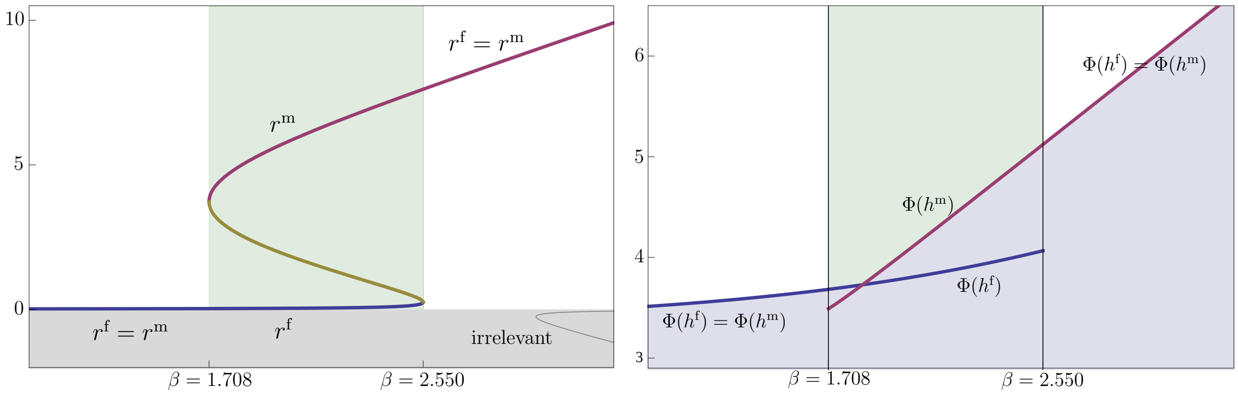
Most of the fixed points which we describe in §5 violate the physical "replica symmetric" intuition that, since the model (2) favors spin $\sigma = 1$ while making no distinction among the remaining spins, the solution $h \in \mathscr{H}^\star$ which attains the optimum in (7) should satisfy $h_1 \geqslant h_2 = \ldots = h_q$. The belief propagation *restricted* to this subspace is again a univariate recursion, naturally parametrized by $r \equiv \log(h_1/h_2)$.

In our view, the central difficulty of the Potts ($q > 3$) models over the Ising model is the presence of regimes of (β, B) with *multiple stable fixed points*, even under this restriction. The situation is illustrated in Fig. 2. Let $h^{\mathtt{f}}$ denote the limit of successive iterations of BP starting from the uniform probability measure on $[q]$, and let $h^{\mathtt{m}}$ denote the limit of successive iterations of BP starting from the probability measure on $[q]$ supported on spin 1. Then $h^{\mathtt{f}}, h^{\mathtt{m}}$ are elements of $\mathscr{H}^\star$, and there are regimes where they differ and give different evaluations of the Bethe functional Φ.

1.4. **Bounds by graph decomposition.** Our proof of Thm. 1 illustrates a more general principle which we now describe in the abstract factor model setting (1). We restrict the discussion below to *d-regular* graph sequences $G_n \to_{loc} \boldsymbol{T}_d$, since in §3 we will show that, for the purposes of computing the free energy, general sequences $G_n \to_{loc} \boldsymbol{T}_d$ can be reduced to the d-regular case using the uniform sparsity hypothesis.

For a finite graph G let I_G denote a uniformly random vertex in G, and write $I_n \equiv I_{G_n}$. From now on let $\mathbb{P}_n$ denote the joint law of (G_n, I_n), and $\mathbb{E}_n$ the expectation over $\mathbb{P}_n$. An equivalent definition of local convergence of G_n to $\boldsymbol{T}_d$ is that $\mathbb{P}_n(B_t(I_n) \not\cong \boldsymbol{T}_{d,t}) = \mathbb{E}_n[\zeta_t(G_n)]$ converges to zero as $n \to \infty$ for all $t \geqslant 0$. Uniform sparsity of G_n is equivalent to uniform integrability for the random degrees $|\partial I_n|$.

For d even, a d-regular graph G can be reduced by removing a random vertex I — leaving the "cavity graph" $G^\partial \equiv G \backslash I$ with d unmatched half-edges incident to the neighbors ∂I of I in G — then placing a matching $\mathfrak{m}$ on these half-edges to form a d-regular graph $G^{\mathfrak{m}}$ with one less vertex. Thm. 1 is proved by showing that over successive iterations of this reduction, the corresponding increments in the log-partition function are upper bounded by $\boldsymbol{\Phi}^\star$ until the graph is reduced to almost nothing.



(A) Bethe fixed points (B) Bethe functional evaluation

FIGURE 2. *Potts model ($q = 30$) with $d = 4$, $B = 1/50$, β on horizontal axis.* The left panel shows the Bethe fixed points $h \in \mathscr{H}^\star$ satisfying $h_2 = \ldots = h_q$, parametrized by $r \equiv \log(h_1/h_2)$. Solutions $r < 0$ are deemed irrelevant since the model favors spin 1. The minimal and maximal non-negative fixed points are $r^{\mathtt{f}}, r^{\mathtt{m}}$ (corresponding to $h^{\mathtt{f}}, h^{\mathtt{m}}$), and there is a regime of β (shaded between vertical lines) where they fail to coincide. We chose here a large value of q to exaggerate this effect, but precisely the same qualitative phenomenon occurs for $B > 0$ sufficiently small *for all* $d \geqslant 3$, *all* $q \geqslant 3$. The right panel shows the corresponding Bethe functional evaluations, $\Phi(h^{\mathtt{f}})$ and $\Phi(h^{\mathtt{m}})$. In [DMS13] we established the lower bound $\liminf_n \phi_n \geqslant \Phi(h^{\mathtt{f}}) \vee \Phi(h^{\mathtt{m}})$, together with an upper bound which was not sharp in the $r^{\mathtt{f}} < r^{\mathtt{m}}$ regime. In this paper we prove the matching upper bound for *all* values $\beta \geqslant 0$, provided d is even (Thm. 1). We also prove that $\Phi(h^{\mathtt{f}}) \vee \Phi(h^{\mathtt{m}})$ is in fact the maximum $\mathbf{\Phi}^\star$ of Φ over *all* of $\mathscr{H}^\star$ (Thm. 4).

The first observation is that the ratios Z_G/Z_{G^∂} and $Z_{G^{\mathfrak{m}}}/Z_{G^\partial}$ can be expressed as averages over the cavity measure $\nu_\partial \equiv \nu_{G^\partial}$:

$$\frac{Z_G}{Z_{G^\partial}} = \sum_{\underline{\sigma}_{\partial I}} \nu_\partial(\underline{\sigma}_{\partial I}) \sum_{\sigma} \bar{\psi}(\sigma) \prod_{j \in \partial I} \psi(\sigma, \sigma_j) \equiv \Psi^{\mathrm{vx}}(\nu_\partial),$$

$$\frac{Z_{G^{\mathfrak{m}}}}{Z_{G^\partial}} = \sum_{\underline{\sigma}_{\partial I}} \nu_\partial(\underline{\sigma}_{\partial I}) \prod_{(ij) \in \mathfrak{m}} \psi(\sigma_i, \sigma_j) \equiv \Psi^{\mathrm{e},\mathfrak{m}}(\nu_\partial).$$

Write $\Psi^{\mathfrak{m}}(\nu_\partial) \equiv \Psi^{\mathrm{vx}}(\nu_\partial)/\Psi^{\mathrm{e},\mathfrak{m}}(\nu_\partial)$. If ν_∂ is a d-fold product measure $h^{\otimes d}$ ($h \in \mathscr{H}$), then $\log \Psi^{\mathrm{vx}}(\nu_\partial)$ and $\log \Psi^{\mathrm{e}}(\nu_\partial)$ reduce to the functions $\Phi^{\mathrm{vx}}(h)$ and $\Phi^{\mathrm{e},\mathfrak{m}}(h)$ of (6), with difference $\log \Psi(\nu_\partial) = \Phi(h)$. Averaging over matchings $\mathfrak{m}$ we define the symmetrizations

$$\Psi^{\mathrm{e,sym}}(\nu_\partial) \equiv \frac{1}{(d-1)!!} \sum_{\mathfrak{m}} \Psi^{\mathrm{e},\mathfrak{m}}(\nu_\partial), \quad \Psi^{\mathrm{sym}}(\nu_\partial) \equiv \frac{\Psi^{\mathrm{vx}}(\nu_\partial)}{\Psi^{\mathrm{e,sym}}(\nu_\partial)}. \tag{8}$$

Definition 1.6. Let $\mathscr{H}^d$ denote the space of d-fold product measures $\underline{h} \equiv h_1 \otimes \cdots \otimes h_d$ ($h_i \in \mathscr{H}$), and let $\mathscr{M}_d(\mathscr{H})$ denote the space of all mixtures over $\mathscr{H}^d$: a measure ν on $\mathscr{X}^d$ belongs to $\mathscr{M}_d(\mathscr{H})$ if it can be expressed as

$$\nu(\sigma_1, \ldots, \sigma_d) = \int_{\mathscr{H}^d} h_1(\sigma_1) \cdots h_d(\sigma_d)\, d\rho(\underline{h}), \quad \text{mixing measure } \rho.$$

Theorem 2. *Consider the factor models (1) defined by a permissive specification $\underline{\psi}$. Suppose there is a subspace $\mathscr{H}^{\mathrm{g}} \subseteq \mathscr{H}$ for which the following hold:*

(i) *There is a uniformly bounded function* $\mathsf{err}(t,x)$ *satisfying* $\lim_{t\to\infty}\limsup_{x\downarrow 0}\mathsf{err}(t,x)=0$ *such that the minimal total variation distance between the cavity measure* $\nu_\partial \equiv \nu_{G\setminus I}$ *and* $\mathscr{M}_d(\mathscr{H}^{\mathrm{g}})$ *is upper bounded by* $\mathsf{err}(t,\zeta_t(G))$ *over all* d*-regular graphs* G*; and*

(ii) *The function* Ψ^{sym} *defined in* (8) *satisfies* $\log\Psi^{\mathrm{sym}}(\underline{h}) \leqslant \mathbf{\Phi}^\star$ *for all* $\underline{h}\in(\mathscr{H}^{\mathrm{g}})^d$.

Then $\limsup_n \phi_n \leqslant \mathbf{\Phi}^\star$ *for the factor model on* G_n *specified by* $\underline{\psi}$.

In §4 we show that the conditions of the preceding theorem are satisfied in the Potts model (implying $\limsup_n \phi_n \leqslant \mathbf{\Phi}^\star$) for any $\beta \geqslant 0$ and $B>0$. The bound extends to $B=0$ by continuity. Thm. 1 is then proved as the matching lower bound $\liminf_n \phi_n \geqslant \mathbf{\Phi}^\star$ was shown in [DMS13, Thm. 1.10]. Of several natural modifications of the graph reduction procedure which we considered for the case of d odd, all fail condition (ii).

1.5. **Bethe variational principle and uniformly random regular graphs.** The Bethe prediction (7) has the following variational characterization. Let $\mathbf{\Delta}$ denote the set of symmetric probability measures $\boldsymbol{h}$ on $\mathscr{X}^2$, with one-point marginals denoted by $\bar{h}$. Define the *Bethe rate function*

$$\begin{aligned}\mathbf{\Phi}(\boldsymbol{h}) &\equiv \langle \log\bar{\psi}\rangle_{\bar{h}} - (d-1)H(\bar{h}) + (d/2)[\langle\log\psi\rangle_{\boldsymbol{h}} + H(\boldsymbol{h})\\ &= -H(\bar{h}\,|\,\bar{\psi}) - (d/2)H(\boldsymbol{h}\,|\,\bar{h}\otimes_\psi \bar{h}).\end{aligned} \tag{9}$$

Above and hereafter, for p,q finite non-negative measures on a finite space, $H(p)$ denotes the Shannon entropy $-\sum_x p_x\log p_x$, and $H(q\,|\,p)$ denotes the relative entropy $\sum_x q_x\log(q_x/p_x)$ between q and p. We take the usual conventions $\log 0=-\infty$, $0\log 0=0$ and $0\log(0/0)=0$.

Proposition 1.7 (Bethe variational principle)**.** *Any interior stationary point* $\boldsymbol{h}$ *of* $\mathbf{\Phi}$ *corresponds to* $h\in\mathscr{H}^\star$ *by the bijective relation*

$$\boldsymbol{h}(\sigma,\sigma') = (h\otimes_\psi h)_{\sigma\sigma'}/\boldsymbol{z}_h \equiv \psi(\sigma,\sigma')h_\sigma h_{\sigma'}/\boldsymbol{z}_h \tag{10}$$

(with $\boldsymbol{z}_h$ *the normalizing constant). Any local maximizer* $\boldsymbol{h}$ *of* $\mathbf{\Phi}$ *is an interior point of* $\mathbf{\Delta}$*, so the Bethe free energy of* (7) *is given alternatively by*

$$\mathbf{\Phi}^\star = \sup_{\boldsymbol{h}\in\mathbf{\Delta}}\mathbf{\Phi}(\boldsymbol{h}),\quad \mathbf{\Phi}\text{ the Bethe rate function (9).} \tag{11}$$

Proof. Follows from [DMS13, Thm. 1.16] (using compactness of $\mathbf{\Delta}$). □

We supply the following simple interpretation for this variational characterization. For nd even, let $\boldsymbol{G}_n\equiv\boldsymbol{G}_{n,d}$ denote the uniformly random d-regular graph on n vertices, sampled according to the usual *configuration model* — that is, start with n isolated vertices each equipped with d half-edges, and form the graph by taking a uniformly random matching on the nd half-edges. Let $\mathbf{E}_n\equiv\mathbf{E}_{n,d}$ denote expectation with respect to the law of $\boldsymbol{G}_{n,d}$.

Conditioned on the event that $\boldsymbol{G}_n$ is free of self-loops and multi-edges, it has the law of the uniformly random *simple* d-regular graph. This event occurs (for fixed d) with uniformly positive probability in the limit $n\to\infty$ [JŁR00].

Theorem 3. *Consider the factor models* (1) *defined by a permissive specification* $\underline{\psi}$*. With* $\mathbf{E}_n$ *denoting expectation with respect to the configuration model for* d*-regular graphs on* $[n]$,

$$\limsup_{n\to\infty} n^{-1}\mathbf{E}_n[\log Z_n] \leqslant \lim_{n\to\infty} n^{-1}\log\mathbf{E}_n[Z_n] = \mathbf{\Phi}^\star \tag{12}$$

The following proposition serves as a counterpoint to (12):

Proposition 1.8. *In the setting of Thm. 3, if the inequality in* (12) *is strict, then there exists a sequence* G_n *of* d*-regular graphs for which* $\liminf_{n\to\infty} n^{-1}\log Z_n > \mathbf{\Phi}^\star$.

We highlight this proposition here because it demonstrates that if the uniformly random ensemble has free energy ϕ_n strictly below the replica symmetric solution $\mathbf{\Phi}^\star$ — as is expected to happen in replica symmetry breaking regimes — then there is breaking of homogeneity in the d-regular graph space as well, with a large subclass of graphs achieving free energy strictly above $\mathbf{\Phi}^\star$. An interesting open question is whether the maximal asymptotic free energy is achieved by random bipartite graphs, as is known to be the case in two-spin models [SS13].

1.6. **Explicit Potts Bethe prediction.** Surprisingly, another consequence of Thm. 3 is the following solution to the optimization problems (7) and (9) for the ferromagnetic Potts model. Let $h^{\mathtt{f}}$ denote the limit of successive iterations of BP starting from the uniform probability measure on $[q]$, and let $h^{\mathtt{m}}$ denote the limit of successive iterations of BP starting from the probability measure on $[q]$ supported on spin 1.

Theorem 4. *For the Potts model (2) with $\beta, B \geqslant 0$, $\mathbf{\Phi}^\star = \Phi(h^{\mathtt{f}}) \vee \Phi(h^{\mathtt{m}})$, and if $B > 0$ then this is strictly greater than $\Phi(h)$ for any $h \in \mathscr{H}^\star \backslash \{h^{\mathtt{f}}, h^{\mathtt{m}}\}$.*

In §5 we supplement Thm. 4 by a classification of stationary points of the Potts Bethe rate function $\mathbf{\Phi}$ (equivalently, via (10), solutions of the Potts Bethe recursion), as well as a study of which stationary points can be local maximizers. The detailed statements are given in Propns. 5.3 and 5.4. The motivation for considering *local* maximizers of $\mathbf{\Phi}$ — which after all are irrelevant to the Bethe prediction (11) if they are not global maximizers — is that we expect these are precisely the fixed points which can be seen in local weak limits of *conditioned* factor models on graph sequences $G_n \to_{loc} \boldsymbol{T}_d$, in the spirit of [MMS12]. That is, when $\boldsymbol{h}$ is a local maximizer of $\mathbf{\Phi}$, the factor model restricted to configurations of edge empirical measure close to $\boldsymbol{h}$ should converge locally weakly to the (Bethe) Gibbs measure corresponding to $\boldsymbol{h}$. However, §5 is independent of the rest of the paper.

2. Uniformly random d-regular graphs

2.1. **Expectation of the partition function.** Given spin configuration $\underline{\sigma}$ on the vertices of graph G, define the edge empirical measure

$$\boldsymbol{h}(\sigma,\sigma') \equiv \frac{1}{2|E|}\sum_{(ij)\in E}\Big(\mathbf{1}\{(\sigma_i,\sigma_j)=(\sigma,\sigma')\}+\mathbf{1}\{(\sigma_i,\sigma_j)=(\sigma',\sigma)\}\Big),$$

and write $\bar{h}$ for its one-point marginal (the vertex empirical measure). Recall the following strong form of Stirling's approximation [Rob55] $1 \leqslant n!/[\sqrt{2\pi n}(n/e)^n] \leqslant e$. Recall also that for n even, the number of matchings on $[n]$ is the double factorial $(n-1)!! = n!/[(n/2)!2^{n/2}]$.

Proof of Thm. 3. Let $Z \equiv Z_n$ denote the partition function for the factor model (1) on the random graph $\boldsymbol{G} \equiv \boldsymbol{G}_{n,d}$. We decompose $Z = \sum_{\boldsymbol{h}} Z(\boldsymbol{h})$ with $Z(\boldsymbol{h})$ the contribution from configurations $\underline{\sigma}$ with edge empirical measure $\boldsymbol{h}$. Recall that $\bar{h}$ denotes the one-point marginal of $\boldsymbol{h}$. The expected number of $\boldsymbol{h}$-configurations on $\boldsymbol{G}$ is given (with the obvious multi-index notation) by

$$\overbrace{\left[\binom{n}{n\bar{h}}\prod_\sigma\binom{nd\bar{h}(\sigma)}{nd\boldsymbol{h}(\sigma,\cdot)}\right]}^{\#\text{ assignments on half-edges}}\times\overbrace{\left[\frac{1}{(nd-1)!!}\prod_\sigma(nd\boldsymbol{h}(\sigma,\sigma)-1)!!\prod_{\sigma\neq\sigma'}\sqrt{(nd\boldsymbol{h}(\sigma,\sigma'))!}\right]}^{\text{probability of matching consistent with given spin assignment}}$$
$$= n^{O(1)}\exp\{n[(d/2)H(\boldsymbol{h})-(d-1)H(\bar{h})]\} \qquad \text{by Stirling's formula.}$$

Each $\boldsymbol{h}$-configuration receives the same weight $\bar{\psi}^{n\bar{h}}\psi^{(nd/2)\boldsymbol{h}} = \exp\{n\langle\log\bar{\psi}\rangle_{\bar{h}}+(nd/2)\langle\log\psi\rangle_{\boldsymbol{h}}\}$, so

$$\mathbf{E}Z(\boldsymbol{h}) = n^{O(1)}\exp\{n[\langle\log\bar{\psi}\rangle_{\bar{h}} - H(\bar{h}) + (d/2)[\langle\log\psi\rangle_{\boldsymbol{h}} + H(\boldsymbol{h})]]\} = n^{O(1)}\exp\{n\boldsymbol{\Phi}(\boldsymbol{h})\}.$$

By Propn. 1.7, $\boldsymbol{\Phi}$ attains its global maximum at an interior point $\boldsymbol{h}^\star \in \boldsymbol{\Delta}$, which must lie within distance $O(n^{-1})$ of one of the empirical measures $\boldsymbol{h}$ realizable on $\boldsymbol{G}$. On the other hand there are only polynomially many such measures, so we conclude $\mathbf{E}Z = n^{O(1)}\exp\{n\boldsymbol{\Phi}(\boldsymbol{h}^\star)\}$, implying the theorem. □

2.2. **Concentration of the log-partition function.** We now prove Propn. 1.2, which we again emphasize applies to general uniformly sparse graph sequences G_n (with no assumption on the local limiting tree).

Proof of Propn. 1.2. If $G = (V, E)$ is any finite graph with an isolated vertex i, and $G' = (V', E')$ is formed by adding d edges between i and V, then

$$\frac{Z_{G'}}{Z_G} = \sum_{\sigma_i}\sum_{\underline{\sigma}_{\partial i}}\nu_G(\sigma_i,\underline{\sigma}_{\partial i})\prod_{(ij)\in E'\setminus E}\psi(\sigma_i,\sigma_j) \leqslant (\psi_{\max})^d$$

with $\psi_{\max}$ the maximum over ψ. On the other hand, considering only the $\sigma_i = \sigma^{\mathrm{p}}$ term in the above sum shows $Z_{G'}/Z_G \geqslant (\psi_{\min})^{d+1}$ for $\psi_{\min} \equiv \min_\sigma[\psi(\sigma^{\mathrm{p}},\sigma)]$.

(a) On any random graph G_n we may consider the Doob martingale of $\log Z_n$ with respect to the vertex-revealing filtration. If all the G_n have maximum degree $\leqslant M$, then the above implies that the Doob martingale has uniformly bounded increments — therefore, by the Azuma–Hoeffding bound,

$$\mathbb{P}_n[|n^{-1}\log Z_n - \phi_n| \geqslant \epsilon] \leqslant e^{-nc(\epsilon/M)^2} \quad \text{for a constant } c \equiv c(\underline{\psi}) > 0. \tag{13}$$

This proves $n^{-1}\log Z_n - \phi_n \to 0$ almost surely, consequently $n^{-1}\log Z_n \to \phi$ almost surely.

(b) For any $M > 0$, define the truncated graphs $G_n[M]$ by removing all the edges incident to any vertex $i \in V_n$ with $|\partial i| \geqslant M$. Let $Z_n[M]$ denote the associated partition function, and $\phi_n[M] \equiv n^{-1}\mathbb{E}_n[\log Z_n[M]]$: then

$$|\phi_n - \phi_n[M]| \leqslant n^{-1}\mathbb{E}_n[|\log Z_n - \log Z_n[M]|] \leqslant c\,\mathbb{E}_n[|\partial I_n|\mathbf{1}\{|\partial I_n| \geqslant M\}]$$

where the first inequality is Jensen's while the second follows from the above, for a constant $c \equiv c(\underline{\psi}) > 0$. By the triangle inequality, $\mathbb{P}(|n^{-1}\log Z_n - \phi| \geqslant 4\epsilon) \leqslant \mathrm{A} + \mathrm{B} + \mathrm{C} + \mathrm{D}$ where

$$\begin{aligned}
\mathrm{A} &\equiv \mathbb{P}(n^{-1}|\log Z_n - \log Z_n[M]| \geqslant \epsilon) && \leqslant (c/\epsilon)\,\mathbb{E}_n[|\partial I_n|\mathbf{1}\{|\partial I_n| \geqslant M\}] \leqslant \delta \\
& && \qquad \text{for } M \geqslant M(c,\epsilon,\delta),\ n \geqslant n(M,c,\epsilon,\delta); \\
\mathrm{B} &\equiv \mathbb{P}(|n^{-1}\log Z_n[M] - \phi_n[M]| \geqslant \epsilon) && \leqslant \exp\{-nc(\epsilon/M)^2\} \leqslant \delta \text{ for } n \geqslant n(M,c,\epsilon,\delta); \\
\mathrm{C} &\equiv \mathbb{P}(|\phi_n[M] - \phi_n| \geqslant \epsilon) && \text{zero for } M \geqslant M(c,\epsilon),\ n \geqslant n(M,c,\epsilon); \\
\mathrm{D} &\equiv \mathbb{P}(|\phi_n - \phi| \geqslant \epsilon) && \text{zero for } n \geqslant n(\epsilon),
\end{aligned}$$

Therefore $\mathbb{P}(|n^{-1}\log Z_n - \phi| \geqslant \epsilon) \to 0$ as $n \to \infty$ for all $\epsilon > 0$, concluding the proof. □

2.3. **Replica symmetry breaking.** We turn now to Propn. 1.8 which gives an indication of the non-concentration of the partition function over the space of all d-regular graphs. The proof is based on the observation that if the inequality in (12) is strict, then the concentration bound (13) will force some graphs to have free energy $\geqslant \mathbf{\Phi}^\star + x$ for some $x > 0$. These graphs will constitute an exponentially small fraction of all d-regular graphs on $[n]$, but even this is enough to extract a sequence $G_n \to_{loc} \boldsymbol{T}_d$, due to the following

Lemma 2.1. *For any $\epsilon, t > 0$ there exists $\alpha \equiv \alpha(d,t,\epsilon) > 0$ with $\mathbf{P}(\zeta_t(\boldsymbol{G}) \geqslant \epsilon) \leqslant e^{-\alpha\, n \log n}$.*

Proof. Recall the following classical inequality (see e.g. [McD98, Lem. 3.11]): if $(I_k)_{k=1}^m$ is a sequence of indicator random variables adapted to filtration $(\mathscr{F}_k)_{k=1}^m$, and $a_k \equiv \mathbb{E}[I_k \,|\, \mathscr{F}_{k-1}]$, then $\mathbb{P}(\sum_{k=1}^m I_k \geqslant mx) \leqslant \exp\{-mH(x\,|\,a)\}$ with a the average of the a_k's, and $H(x\,|\,a)$ the binary relative entropy between x and a.

Consider the process of revealing the graph $\boldsymbol{G}$ one edge at a time, with $(\mathscr{F}_k)_{k=1}^{nd/2}$ the corresponding filtration, and let I_k ($1 \leqslant k \leqslant nd/2$) be the indicator that the k-th edge forms a cycle of length $\leqslant 2t$ within the graph revealed so far. Each such cycle can create only a bounded number of vertices with depth-t neighborhood non-isomorphic to the depth-t subtree of the d-regular tree, that is to say, we must have $\sum_k I_k \geqslant n\zeta_t(\boldsymbol{G}) \cdot 2\alpha_0$ for a positive constant $\alpha_0 \equiv \alpha_0(d,t)$. For $k \leqslant n(d/2 - \alpha_0\epsilon)$, since the configuration model matches half-edges uniformly at random, we must have $a_k \equiv \mathbb{E}[I_k \,|\, \mathscr{F}_{k-1}] \leqslant \alpha_1/n$ for a positive constant $\alpha_1 \equiv \alpha_1(d,t,\alpha_0\epsilon)$. Therefore

$$\mathbf{P}(\zeta_t(\boldsymbol{G}) \geqslant \epsilon) \leqslant \mathbf{P}\Big(\sum_{k=1}^{n(d/2-\alpha_0\epsilon)} I_k \geqslant n\epsilon\alpha_0\Big) \leqslant \exp\{-n(d/2 - \alpha_0\epsilon)H(\epsilon\alpha_0 \,|\, a)\}$$

with a the average of the $(a_k)_{k \leqslant n(d/2-\alpha_0\epsilon)}$. Since $a \leqslant \alpha_1/n$ the result readily follows. □

Proof of Propn. 1.8. With $\mathbf{P}_n$ the d-regular configuration model law as before, let us define $\mathbf{p}_n(y) \equiv \mathbf{P}_n(n^{-1}\log Z_n \geqslant \mathbf{\Phi}^\star + y)$. Since clearly $n^{-1}\log Z$ is uniformly bounded over *all* d-regular graphs by a finite constant $C > \mathbf{\Phi}^\star$, we may bound (for $0 < x < C - \mathbf{\Phi}^\star$)

$$\mathbb{E}Z_n \leqslant e^{n(\mathbf{\Phi}^\star - \delta)} + [\mathbf{p}_n(-\delta) - \mathbf{p}_n(x)]e^{n(\mathbf{\Phi}^\star + x)} + \mathbf{p}_n(x)e^{nC}.$$

Recalling Thm. 3 the left-hand side is $n^{O(1)}e^{n\mathbf{\Phi}^\star}$, and rearranging gives

$$\frac{n^{O(1)} - e^{-n\delta} - \mathbf{p}_n(-\delta)e^{nx}}{e^{n(C-\mathbf{\Phi}^\star)} - e^{nx}} \leqslant \mathbf{p}_n(x).$$

Now suppose $\phi^{\mathrm{cm}} \equiv \limsup_n n^{-1}\mathbf{E}_n[\log Z_n]$ is strictly below $\mathbf{\Phi}^\star$: taking $0 < \delta < \mathbf{\Phi}^\star - \phi^{\mathrm{cm}}$, the concentration bound (13) implies that $\mathbf{p}_n(-\delta)$ will be exponentially small in n. It is then clear that one may choose x sufficiently small such that $\mathbf{p}_n(x) \geqslant e^{-n(C-\mathbf{\Phi}^\star)/2}$. Meanwhile, from Lem. 2.1, we can take $t(n) \uparrow \infty$ and $\epsilon(n) \downarrow 0$ slowly enough such that

$$\mathbf{P}_n(\zeta_{t(n)}(G_n) \geqslant \epsilon(n)) \leqslant e^{-1}\mathbf{p}_n(x)$$

Therefore the set of d-regular graphs G on $[n]$ with $n^{-1}\log Z_G \geqslant \mathbf{\Phi}^\star + x$ and $\zeta_{t(n)}(G) \geqslant \epsilon(n)$ has cardinality at least

$$(nd-1)!!\, \mathbf{p}_n(x)[1 - e^{-1}] = \exp\{(nd/2)\log n + O(n)\} \gg 1,$$

so clearly we may extract the desired sequence $G_n \to_{loc} \boldsymbol{T}_d$, $\liminf_n n^{-1}\log Z_n \geqslant \mathbf{\Phi}^\star + x$. □

2.4. **Solution of the Potts variational problem.** We now apply the calculation of Thm. 3 to prove Thm. 4, giving an essentially explicit solution to the Bethe variational problem for the ferromagnetic Potts model. We do not know of a proof which does not go through the probabilistic interpretation for the Bethe variational principle which is described in Thm. 3.

The Potts Bethe recursion for $\beta, B \geqslant 0$ preserves the subspace $\mathscr{H}^{\mathrm{bal}}$ of measures on $\boldsymbol{\Delta}$ which are biased towards $\sigma = 1$ while giving equal weight to all spins $\sigma \neq 1$:

$$\mathscr{H}^{\mathrm{bal}} \equiv \{h \in \mathscr{H} : h_1 = [1 + (1-q)b]/q \text{ and } h_2 = \cdots = h_q = (1-b)/q \text{ with } b \geqslant 0\}. \quad (14)$$

The map BP restricted to this subset is simply a univariate recursion: in terms of the log-likelihood ratio $r \equiv \log(h_1/h_2) \geqslant 0$, it has the particularly simple form

$$\widetilde{\mathrm{BP}} : r \mapsto B + (d-1)\log\frac{e^{\beta+r} + q - 1}{e^r + e^\beta + q - 2}, \quad (15)$$

which is straightforward to analyze; see e.g. [DMS13, Lem. 4.6]. The Bethe recursion fixed points $h^{\mathtt{f}}, h^{\mathtt{m}} \in \mathscr{H}^\star \cap \mathscr{H}^{\mathrm{bal}}$ with minimal and maximal bias b are given by the limits of repeated iterations of this recursion started from $r = 0$ and $r = \infty$ respectively; moreover these are the only fixed points in $\mathscr{H}^{\mathrm{bal}}$.

Let us now review the well-known Fortuin–Kasteleyn (random-cluster) representation of the Potts model. On a finite graph $G \equiv (V, E)$, denote a spin configuration $\underline{\sigma} \in [q]^V$ as before, and denote a *bond configuration* $\underline{\eta} \in \{0,1\}^E$. The *Edwards–Sokal* (ES) *measure* on a finite graph $G \equiv (V, E)$ is the probability measure on pairs $(\underline{\sigma}, \underline{\eta})$ given by

$$\mathrm{ES}_G(\underline{\sigma}, \underline{\eta}) = \frac{1}{\boldsymbol{Z}_G} \prod_{i \in V} e^{B\mathbf{1}\{\sigma_i = 1\}} \prod_{e=(ij)\in E} [(1-p)(1-\eta_e) + p\eta_e \mathbf{1}\{\sigma_i = \sigma_j\}], \quad (16)$$

with $\boldsymbol{Z}_G$ the normalizing constant. Taking $p \equiv 1 - e^{-\beta}$, the marginal of ES_G on the spin configurations $\underline{\sigma}$ is the q-Potts measure with parameters (β, B), and we see that the Potts partition function Z_G is simply $e^{\beta|E|}$ times the ES partition function $\boldsymbol{Z}_G$. The marginal of ES_G on the bond configurations $\underline{\eta}$ is the *Fortuin–Kasteleyn* (FK) measure

$$\mathrm{FK}_G(\underline{\eta}) = \frac{e^{B|V|}}{\boldsymbol{Z}_G} \prod_{i \in E} p^{\eta_e}(1-p)^{1-\eta_e} \prod_{C \in \mathscr{C}(\underline{\eta})} [1 + (q-1)e^{-B|C|}],$$

where the second product runs over the collection $\mathscr{C}(\underline{\eta})$ of connected components C of $\underline{\eta}$ (with $|C|$ denoting the number of vertices in component C).

The ES coupling between Potts and FK has the following simple description: conditioned on the spins $\underline{\sigma}$, $\underline{\eta}$ is defined by a p-percolation on the bonds joining like spins, while all bonds joining unlike spins are left unoccupied. In the other direction, conditioned on bond configuration $\underline{\eta}$ with $\mathscr{C}(\underline{\eta}) = (C_1, \ldots, C_k)$, the spin configuration $\underline{\sigma}$ is given by assigning the same spin to all the vertices of each component, independently over the different components. Component C receives spin $\sigma(C) = \sigma$ with probability

$$e^{B|C|\mathbf{1}\{\sigma=1\}}/[e^{B|C|} + (q-1)]. \quad (17)$$

Proof of Thm. 4. We assume without loss that $B > 0$, with the result for $B = 0$ following by continuity. We first show that the rate function $\boldsymbol{\Phi}$ cannot attain its global maximum outside the subspace

$$\boldsymbol{\Delta}^{\mathrm{bal}} \equiv \{\boldsymbol{h} \in \boldsymbol{\Delta} : \bar{h}(1) \geqslant 1/q \text{ and } \bar{h}(2) = \cdots = \bar{h}(q)\}.$$

To this end, consider conditioning on the FK configuration $\underline{\eta}$ in the ES coupling. Given $C \in \mathscr{C}(\underline{\eta})$, define the random variable $Y_C(\sigma) \equiv |C|\mathbf{1}\{\sigma(C) = \sigma\}$. For all $\sigma \neq 1$ this has the same cumulant generating function

$$\kappa_C(t) \equiv \log \mathbb{E}[\exp\{tY_C(\sigma)\} \,|\, \underline{\eta}] = \log \frac{e^{t|C|} + e^{B|C|} + q - 2}{e^{B|C|} + q - 1}.$$

For $-\infty < t \leqslant B/2$ we calculate

$$\kappa''_C(t) = \frac{|C|^2 e^{t|C|}(e^{B|C|} + q - 2)}{(e^{t|C|} + e^{B|C|} + q - 2)^2} \leqslant \frac{|C|^2 e^{t|C|}}{e^{t|C|} + e^{B|C|} + q - 2} \leqslant \frac{|C|^2}{e^{B|C|/2}} \leqslant (eB/4)^{-2} \equiv c_B,$$

uniformly over all $|C| \geqslant 0$. Noting $|\mathscr{C}(\underline{\eta})| \leqslant n$, for all $\sigma \neq 1$ we have

$$\varpi_{G_n}\Big(\Big|\sum_{C \in \mathscr{C}(\underline{\eta})} (Y_C(\sigma) - \mathbb{E}[Y_C(\sigma)) \,|\, \underline{\eta}]\Big| \geqslant n\epsilon\Big) \leqslant 2 \inf_{0 \leqslant t \leqslant B/2} \frac{e^{nc_B t^2/2}}{e^{nt\epsilon}} \leqslant \frac{2}{e^{n\epsilon^2/(2c_B)}},$$

where the last inequality holds provided $\epsilon/c_B \leqslant B/2$. Clearly, for all spins $\sigma \neq 1$ the mean $\mathbb{E}[Y_C(\sigma) \,|\, \underline{\eta}]$ takes the same value, which is less than $|C|/q$. Let $Z_n[\epsilon]$ denote the contribution to the Potts partition function on G_n from the space $\boldsymbol{\Delta}[\epsilon]$ of measures $\boldsymbol{h} \in \boldsymbol{\Delta}$ with $\operatorname{dist}(\boldsymbol{h}, \boldsymbol{\Delta}^{\mathrm{bal}}) \geqslant \epsilon$: the above implies there is a constant $\alpha \equiv \alpha(B, q) > 0$ such that

$$Z_n[\epsilon] \leqslant e^{-n\alpha\epsilon^2} Z_n \quad \text{for } \epsilon \leqslant c_B(B/2). \tag{18}$$

The estimate (18) holds for any graph G_n on $[n]$. To prove the result, take $G_n = \boldsymbol{G}_n$, the random d-regular graph drawn from the configuration model. Comining (18) with the calculation of Thm. 3 gives

$$n^{O(1)} \exp\{n \sup\{\boldsymbol{\Phi}(\boldsymbol{h}) : \boldsymbol{h} \in \boldsymbol{\Delta}[\epsilon]\}\} = \mathbf{E}_n[Z_n[\epsilon]] \leqslant \frac{\mathbf{E}_n[Z_n]}{e^{n\epsilon^2\alpha_B}} = \frac{n^{O(1)} \exp\{n\boldsymbol{\Phi}^\star\}}{e^{n\epsilon^2\alpha_B}}$$

for all $\epsilon \leqslant c_B(B/2)$, proving the claim that $\boldsymbol{\Phi}$ can only attain its global maximum on $\boldsymbol{\Delta}^{\mathrm{bal}}$.

To conclude the proof, we apply the variational principle Propn. 1.7. Let $\boldsymbol{h}$ be any global maximizer for $\boldsymbol{\Phi}$, so by the above $\boldsymbol{h} \in \boldsymbol{\Delta}^{\mathrm{bal}}$. Let $h \in \mathscr{H}^\star$ correspond via (10) to $\boldsymbol{h}$: summing (10) over one of the spins σ' gives $\boldsymbol{z}_h \bar{h}_\sigma = h_\sigma[(e^\beta - 1)h_\sigma + 1]$, which implies (since the right-hand side is increasing in h_σ for $h_\sigma > 0$) that $h \in \mathscr{H}^\star \cap \mathscr{H}^{\mathrm{bal}}$. As noted above, the only two possibilities for h are $h^{\mathtt{f}}$ and $h^{\mathtt{m}}$, concluding the proof. □

3. Recursive graph decomposition

In this section we prove Thm. 2 for graph sequences $G_n \to_{loc} \boldsymbol{T}_d$ (Defn. 1.1), d even. The following lemma, whose proof we defer to the end of the section, reduces the free energy computation to the case of d-regular graphs.

Lemma 3.1. *If $G_n \to_{loc} \boldsymbol{T}_d$ with d even, then there exists a d-regular graph sequence $G'_n \to_{loc} \boldsymbol{T}_d$ with free energy ϕ'_n such that* $\lim_{n\to\infty}(\phi_n - \phi'_n) = 0$.

The lemma is a refinement of the result of Propn. 1.2b. Assuming the lemma, we now prove Thm. 2 by analysis of the following

> **Reduction operation R** on d-regular graphs (d even).
> Given graph G, remove a uniformly random vertex I, leaving the cavity graph $G^\partial \equiv G \backslash I$ with d unmatched half-edges incident to the neighbors ∂I of I in G. Let $G^{\mathfrak{m}}$ be the d-regular graph formed by placing matching $\mathfrak{m}$ on

these half-edges, and set $\mathbf{R}G = G^{\mathfrak{m}}$ for some choice of $\mathfrak{m}$ which minimizes $\log Z_G - \log Z_{G^{\mathfrak{m}}}$.

We prove Thm. 2 via the following propositions concerning operation $\mathbf{R}$ on d-regular graphs.

Proposition 3.2. *Let G be any finite d-regular graph, and write $\mathbb{E}_G$ for expectation over the choice of a uniformly random vertex I in G. Under the conditions of Thm. 2,*

$$\mathbb{E}_G[\log Z_G - \log Z_{\mathbf{R}G}] \leqslant \mathbf{\Phi}^\star + \mathsf{err}(t, \zeta_t(G))$$

where $\mathsf{err}(t,x)$ is a uniformly bounded function with $\lim_{t\to\infty} \limsup_{x\downarrow 0} \mathsf{err}(t,x) = 0$.[3]

Proof. By definition of operation $\mathbf{R}$ and of the cavity measure $\nu_\partial \equiv \nu_{G\setminus I}$, and recalling the manipulations leading to the definition (8), we have

$$\mathbb{E}_G\Big[\log \frac{Z_G}{Z_{\mathbf{R}G}}\Big] = \mathbb{E}_G\Big[\min_{\mathfrak{m}} \log \frac{\Psi^{\mathrm{vx}}(\nu_\partial)}{\Psi^{\mathrm{e},\mathfrak{m}}(\nu_\partial)}\Big] \leqslant \mathbb{E}_G\Big[\log \Psi^{\mathrm{sym}}(\nu_\partial)\Big]. \tag{19}$$

Observe that for any permissive specification $\underline{\psi}$, $\Psi^{\mathrm{vx}}(\nu_\partial)$ and $\Psi^{\mathrm{e}}(\nu_\partial)$ are uniformly bounded and uniformly positive deterministically over all graphs G with maximum degree d — this is easily seen by calculations very similar to the estimates of the ratio $Z_{G'}/Z_G$ appearing in the proof of Propn. 1.2, together with the observation that the marginal of ν_∂ on any vertex must give uniformly positive measure to spin σ^{p}. It follows from condition (i) of Thm. 2 that for some mixing measure ρ over $(\mathscr{H}^{\mathrm{g}})^d$, the right-hand side of (19) is within $\mathsf{err}(t, \zeta_t(G))$ of

$$\mathbb{E}_G\Big[\log \frac{\int \Psi^{\mathrm{vx}}(\underline{h})\, d\rho(\underline{h})}{\int \Psi^{\mathrm{e,sym}}(\underline{h})\, d\rho(\underline{h})}\Big] \leqslant \sup_{\underline{h}\in(\mathscr{H}^{\mathrm{g}})^d} \log \Psi^{\mathrm{sym}}(\underline{h}).$$

Condition (ii) then gives that this is $\leqslant \mathbf{\Phi}^\star$, concluding the proof. □

Proposition 3.3. *On any d-regular graph sequence $G_n \to_{loc} \boldsymbol{T}_d$ it holds that*

$$\lim_{n\to\infty} \mathbb{E}_n\Big[\max_{0\leqslant j\leqslant (1-\epsilon_0)n} \zeta_t(\mathbf{R}^j G_n)\Big] = 0 \quad \textit{for all } \epsilon_0 > 0,\ t \geqslant 0.$$

Proof. Let G be any d-regular graph on n vertices. For each vertex v in G consider the event

$$\mathbf{\Omega}_{t,\epsilon}(v) \equiv \{B_t(v) \cong \boldsymbol{T}_{d,t} \text{ in } G, \text{ but } B_{t/2}(v) \ncong \boldsymbol{T}_{d,t/2} \text{ in } \mathbf{R}^j G \text{ for some } j \leqslant n\epsilon\}.$$

This implies that during the first $n\epsilon$ applications of $\mathbf{R}$, at least $t/2$ vertices were deleted along some length-t geodesic path started from v. For each such geodesic, the same inequality cited in the proof of Propn. 2.1 implies that the probability of $\geqslant t/2$ deletions along the geodesic is (for $\epsilon > 0$ small)

$$\leqslant \exp\{-n\epsilon H(t/(2n\epsilon)\,|\,t/(2n))\} = \exp\{-(t/2)\log(1/\epsilon) + O(t)\},$$

which can clearly be made $\leqslant d^{-3t}$ by taking $\epsilon \leqslant \epsilon(d)$. Taking a union bound over d^t geodesics shows that $\mathbf{\Omega}_{t,\epsilon}(v)$ has probability $\leqslant d^{-2t}$, therefore

$$\mathbb{P}\Big[\max_{0\leqslant j\leqslant n\epsilon} (n-j)\zeta_{t/2}(\mathbf{R}^j(G)) \geqslant n\zeta_t(G) + nd^{-t}\Big] \leqslant \mathbb{P}\Big[\sum_{v\in G} \mathbf{\Omega}_{t,\epsilon}(v) \geqslant nd^{-t}\Big] \leqslant d^{-t}$$

where the last step is by Markov's inequality.

Now iterate the above over $L \equiv L(\epsilon_0, \epsilon(d))$ passes, removing an $\epsilon(d)$-fraction of vertices on each pass until only an ϵ_0-fraction of the original vertex set remains. (L is roughly

[3] The function $\mathsf{err}(t,x)$ may need to be adjusted going from Thm. 2 to Propn. 3.2, but can be chosen to depend only on d and $\underline{\psi}$.

$\log_{1-\epsilon(d)}\epsilon_0$, though not precisely so because each pass must remove an integer number of vertices.) Taking a crude upper bound on the accumulation of errors gives

$$\mathbb{P}\Big[\max_{0\leqslant j\leqslant n(1-\epsilon_0)}(n-j)\zeta_{t/2^L}(\mathbf{R}^j(G))\geqslant n\zeta_t(G)+nLd^{-t/2^L}\Big]\leqslant Ld^{-t/2^L}.$$

Thus, for the graph sequence G_n,

$$\mathbb{E}_n\Big[\max_{0\leqslant j\leqslant n(1-\epsilon_0)}\zeta_{t/2^L}(\mathbf{R}^j(G))\Big]\leqslant\frac{n\mathbb{E}_n[\zeta_t(G_n)]+nLd^{-t/2^L}}{n\epsilon_0}+Ld^{-t/2^L}$$

The right-hand side tends to zero in the limit $n\to\infty$ followed by $t\to\infty$, but the left-hand side is non-decreasing in t so it must in fact tend to zero as $n\to\infty$ for all t, as claimed. □

Proof of Thm. 2. By Lem. 3.1 we reduce to the setting of a d-regular graph sequence $G_n\to_{loc}\boldsymbol{T}_d$. Take $n_0\equiv n(1-\epsilon)$, and express the free energy of the factor model on G_n as the telescoping sum

$$\phi_n=n^{-1}\sum_{j=0}^{n_0-1}\mathbb{E}_n[\log Z_{\mathbf{R}^jG_n}-\log Z_{\mathbf{R}^{j+1}G_n}]+n^{-1}\mathbb{E}_n[\log Z_{\mathbf{R}^{n_0}G_n}].$$

Since $\mathbf{R}^{n_0}G_n$ is a d-regular graph on $n\epsilon$ vertices, the last term is bounded in absolute value by $c\epsilon$ for some constant $c\equiv c(d,\underline{\psi})$. Next, Propn. 3.2 gives

$$\max_{0\leqslant j<n_0}\mathbb{E}_n[\log Z_{\mathbf{R}^jG_n}-\log Z_{\mathbf{R}^{j+1}G_n}]\leqslant\mathbf{\Phi}^\star+\mathbb{E}_n\Big[\max_{0\leqslant\ell<n_0}\mathsf{err}\big(t,\zeta_t(\mathbf{R}^\ell G_n)\big)\Big].$$

By Propn. 3.3, the last term tends to zero in the limit $n\to\infty$ followed by $t\to\infty$, so

$$\limsup_{n\to\infty}\phi_n\leqslant(1-\epsilon)\mathbf{\Phi}^\star+c\epsilon,$$

and the result follows by taking $\epsilon\downarrow 0$. □

Proof of Lem. 3.1. Writing $G\equiv G_n\equiv(V,E)$, we form the d-regular graph $G'\equiv G'_n$ in two steps: (1) delete edges in G incident to vertices of degree larger than d until none remain to form the graph $G''\equiv(V,E'')$, then (2) add $d-|\partial v|$ half-edges to every $v\in V$, and take a matching on these half-edges to form the d-regular graph $G'=(V,E')$.

Let U denote the set of all vertices incident to any edge in $E\backslash E''$, and observe that $(\psi_{\max})^{|E\backslash E''|}\geqslant Z_G/Z_{G''}\geqslant(\psi_{\min})^{|E\backslash E''|}\nu_{G''}(\underline{\sigma}_U\equiv\sigma^{\mathrm{p}})$. Since G'' has maximum degree $\leqslant d$, there exists a positive constant $\tilde{c}\equiv\tilde{c}(d,\underline{\psi})>0$ for which $\nu_{G''}(\underline{\sigma}_U\equiv\sigma^{\mathrm{p}})\geqslant\tilde{c}^{|U|}$. Since clearly $|U|\leqslant 2|E\backslash E''|$, we conclude there exists a positive constant $c\equiv c(d,\underline{\psi})$ such that $|\log(Z_G/Z_{G''})|\leqslant c|E\backslash E''|$. By similar considerations (and adjusting c as needed) we conclude $|\log(Z_{G'}/Z_{G''})|\leqslant c\zeta_2(G)$, therefore

$$|\phi_n-\phi'_n|\leqslant c\,\mathbb{E}_n[\zeta_2(G_n)]+c\,\mathbb{E}_n[|\partial I_n|\mathbf{1}\{|\partial I_n|\neq d\}].$$

The first term tends to zero in the limit $n\to\infty$, while the second term is upper bounded by

$$cL\,\mathbb{E}_n[\zeta_1(G_n)]+c\,\mathbb{E}_n[|\partial I_n|\mathbf{1}\{|\partial I_n|\geqslant L\}].$$

By the uniform sparsity hypothesis this tends to zero in the limit $n\to\infty$ followed by $L\to\infty$, concluding the proof. □

4. The Potts free energy density

In this section we prove our main result Thm. 1 giving the free energy density of the q-Potts model on graphs converging locally to the d-regular tree with d even. Let $\mathscr{H}^{\rm FK} \subsetneq \mathscr{H}^{\rm bal}$ denote the set of all convex combinations of the measures $h^{\mathtt{f}}, h^{\mathtt{m}}$ defined in §1.6.

4.1. Product decomposition of Potts cavity measure.

We begin by showing that for $\beta \geqslant 0$ and $B > 0$, the Potts model satisfies condition (i) of Thm. 2 with $\mathscr{H}^{\rm g} = \mathscr{H}^{\rm FK}$:

Proposition 4.1. *For the Potts model with $\beta \geqslant 0$ and $B > 0$, there is a uniformly bounded function $\mathsf{err}(t,x)$ with $\lim_{t\to\infty} \limsup_{x\downarrow 0} \mathsf{err}(t,x) = 0$ such that the minimal total variation distance between the cavity measure ν_∂ and $\mathscr{M}_d(\mathscr{H}^{\rm FK})$ is upper bounded by $\mathsf{err}(t,\zeta_t(G))$ over all d-regular graphs G.*

Proof. On the graph $G^\partial \equiv G\backslash I$ we will consider the Edwards–Sokal measure $\textsc{es}_\partial$ (see (16)), with spin marginal ν_∂ (the Potts measure) and bond marginal $\textsc{fk}_\partial$. Condition on the event that $B_t(I) \cong \boldsymbol{T}_{d,t}$, which occurs with probability $1-\zeta_t(G)$. On this event, $\boldsymbol{B}_t \equiv B_t(I) \cap G^\partial$ is simply a collection of d disjoint trees, each isomorphic to $\acute{\boldsymbol{T}}_{d,t-1}$ (see Rmk. 1.4).

Fixing $s < t$, for $\underline{\eta}$ any bond configuration on G^∂ we shall decompose $\underline{\eta} \equiv (\underline{\eta}_{\rm O}, \underline{\eta}_{\rm A}, \underline{\eta}_{\rm I})$ with $\underline{\eta}_{\rm O}$ the bond configuration on $G^\partial\backslash\boldsymbol{B}_t$, $\underline{\eta}_{\rm I}$ the bond configuration on $\boldsymbol{B}_s$, and $\underline{\eta}_{\rm A}$ the bond configuration on $\boldsymbol{B}_t\backslash\boldsymbol{B}_s$ (O = "outer", I = "inner", A = "annuli"). Then

$$\nu_\partial(\underline{\sigma}_{\partial I}) = \sum_{\underline{\eta}_{\rm O,A}} \textsc{fk}_\partial(\underline{\eta}_{\rm O,A}) \sum_{\underline{\eta}_{\rm I}} \textsc{fk}_\partial(\underline{\eta}_{\rm I} \,|\, \underline{\eta}_{\rm O,A}) \textsc{es}_\partial(\underline{\sigma}_{\partial I} \,|\, \underline{\eta}).$$

We claim that the inner sum remains essentially unaffected if we set $\underline{\eta}_{\rm O}$ to the identically-0 configuration $\underline{\zeta}_{\rm O}$:

1. For any $\underline{\eta}_{\rm O}, \underline{\eta}_{\rm A}$ we calculate the ratio
$$\frac{\textsc{fk}_\partial(\underline{\eta}_{\rm I} \,|\, \underline{\eta}_{\rm O}, \underline{\eta}_{\rm A})}{\textsc{fk}_\partial(\underline{\eta}_{\rm I} \,|\, \underline{\zeta}_{\rm O}, \underline{\eta}_{\rm A})} = \boldsymbol{c}(\underline{\eta}_{\rm O,A}) \frac{\prod_{C\in\mathscr{D}(\underline{\eta}_{\rm O}, \underline{\eta}_{\rm A,I})} [1 + e^{-B|C|}(q-1)]}{\prod_{C'\in\mathscr{D}(\underline{\zeta}_{\rm O}, \underline{\eta}_{\rm A,I})} [1 + e^{-B|C'|}(q-1)]}$$
where $\boldsymbol{c}(\underline{\eta}_{\rm O,A})$ is a proportionality constant not involving $\underline{\eta}_{\rm I}$, and $\mathscr{D}(\underline{\eta})$ denotes the connected components of $\underline{\eta}$ which cross between $\boldsymbol{B}_s$ and $G\backslash\boldsymbol{B}_t$. Any such component must contain at least $t-s$ vertices, and the total number of such components in any $\underline{\eta}$ is at most d^s. Since $B > 0$, if we take say $s = \log t$, then the above ratio tends to 1 in the limit $t \to \infty$, uniformly over all $\underline{\eta}_{\rm O}$.
2. Similarly, given any $\underline{\eta}$, let $\partial I(t) \subseteq \partial I$ denote the vertices in ∂I which are joined to $G\backslash\boldsymbol{B}_t$ via the occupied bonds in $\underline{\eta}_{\rm I,A}$. Then $\textsc{es}_\partial(\underline{\sigma}_{\partial I} \,|\, \underline{\eta}_{\rm O}, \underline{\eta}_{\rm A,I}) - \textsc{es}_\partial(\underline{\sigma}_{\partial I} \,|\, \underline{\zeta}_{\rm O}, \underline{\eta}_{\rm A,I})$ factorizes as
$$[\textsc{es}_\partial(\underline{\sigma}_{\partial I(t)} \,|\, \underline{\eta}_{\rm O}, \underline{\eta}_{\rm A,I}) - \textsc{es}_\partial(\underline{\sigma}_{\partial I(t)} \,|\, \underline{\zeta}_{\rm O}, \underline{\eta}_{\rm A,I})] \times \prod_{v\in\partial I\backslash\partial I(t)} \textsc{es}_\partial(\sigma_v \,|\, \underline{\zeta}_{\rm O}, \underline{\eta}_{\rm I,A}).$$
Since $B > 0$, in the limit $t \to \infty$, $\textsc{es}_\partial(\underline{\sigma}_{\partial I(t)} \,|\, \underline{\eta}_{\rm O}, \underline{\eta}_{\rm A,I})$ converges to 1 if $\underline{\sigma}_{\partial I(t)}$ is identically 1, and converges to 0 otherwise, uniformly over $\underline{\eta}$. Thus the total variation distance between $\textsc{es}_\partial(\underline{\sigma}_{\partial I} = \cdot \,|\, \underline{\eta}_{\rm O}, \underline{\eta}_{\rm A,I})$ and $\textsc{es}_\partial(\underline{\sigma}_{\partial I} = \cdot \,|\, \underline{\zeta}_{\rm O}, \underline{\eta}_{\rm A,I})$ tends to zero as $t \to \infty$, uniformly over $\underline{\eta}$.

Combining the above estimates shows that on the event $B_t(I) \cong \boldsymbol{T}_{d,t}$, $\nu_\partial(\underline{\sigma}_{\partial I})$ is well approximated (in total variation distance, for t large) by the measure

$$\begin{aligned}\widetilde{\nu}_\partial(\underline{\sigma}_{\partial I}) &= \sum_{\underline{\eta}_{\mathrm{O,A}}} \mathrm{FK}_\partial(\underline{\eta}_{\mathrm{O,A}}) \sum_{\underline{\eta}_{\mathrm{I}}} \mathrm{FK}_\partial(\underline{\eta}_{\mathrm{I}} \,|\, \underline{\zeta}_{\mathrm{O}}, \underline{\eta}_{\mathrm{A}}) \mathrm{ES}_\partial(\underline{\sigma}_{\partial I} \,|\, \underline{\zeta}_{\mathrm{O}}, \underline{\eta}_{\mathrm{A,I}}) \\ &= \sum_{\underline{\eta}_{\mathrm{A}}} \mathrm{FK}_\partial(\underline{\eta}_{\mathrm{A}}) \mathrm{ES}_\partial(\underline{\sigma}_{\partial I} \,|\, \underline{\zeta}_{\mathrm{O}}, \underline{\eta}_{\mathrm{A}}) = \sum_{\underline{\eta}_{\mathrm{A}}} \mathrm{FK}_\partial(\underline{\eta}_{\mathrm{A}}) \prod_{v \in \partial I} \mathrm{ES}_\partial(\sigma_v \,|\, \underline{\zeta}_{\mathrm{O}}, \underline{\eta}_{\mathrm{A}}).\end{aligned}$$

Each $\mathrm{ES}_\partial(\sigma_v \,|\, \underline{\zeta}_{\mathrm{O}}, \underline{\eta}_{\mathrm{A}})$ is simply the root marginal of the Potts measure on $\acute{\boldsymbol{T}}_{d,s}$ with some boundary conditions, which from (17) must be biased towards 1. It follows (cf. Rmk. 1.4) that in the limit $s \equiv \log t \to \infty$, the distance between $\mathrm{ES}_\partial(\sigma_v \,|\, \underline{\zeta}_{\mathrm{O}}, \underline{\eta}_{\mathrm{A}})$ and $\mathscr{H}^{\mathrm{FK}}$ will tend to zero (uniformly over $\underline{\eta}_{\mathrm{A}}$). We therefore conclude

$$\min\{\|\nu_\partial - \nu'\|_{\mathrm{TV}} : \nu' \in \mathscr{M}_d(\mathscr{H}^{\mathrm{FK}})\} \leqslant \zeta_t(G) + \epsilon_t$$

with $\epsilon_t \to 0$ as $t \to \infty$, implying the result. □

4.2. **Optimization over product measures.** We now verify condition (ii) of Thm. 2:

Proposition 4.2. *For the Potts model with $\beta, B \geqslant 0$,* $\sup\{\log \Psi^{\mathrm{sym}}(\underline{h}) : \underline{h} \in (\mathscr{H}^{\mathrm{FK}})^d\} = \boldsymbol{\Phi}^\star$.

Recall the function $\Psi^{\mathfrak{m}} \equiv \Psi^{\mathrm{vx}}/\Psi^{\mathrm{e},\mathfrak{m}}$ defined in §1.3. The following lemma is proved by manipulating the BP identity (5), and applies to general factor models.

Lemma 4.3. *If $h^\star \in \mathscr{H}^\star$ then $\Psi^{\mathfrak{m}}(h, h^\star, \ldots, h^\star)$ is constant over $h \in \mathscr{H}$ for each $\mathfrak{m}$.*

Proof. Note $\Psi^{\mathrm{vx}}(\underline{h})$ equals $Z_S(\underline{h})$, the partition function on the star graph $S \equiv \boldsymbol{T}_{d,1}$ with boundary conditions $\sigma_j \sim h_j$ independently over $j \in \partial o$. Let S' be S with the edge $(o,1)$ disconnected: from (5), $Z_{S'}(h, h^\star, \ldots, h^\star) = z_{h^\star}$ regardless of h, while

$$Z_S(h, h^\star, \ldots, h^\star) = z_{h^\star} \sum_{\sigma,\sigma'} \psi(\sigma,\sigma') h^\star_\sigma h_{\sigma'}.$$

Similarly, for any $\mathfrak{m}$, $\Psi^{\mathrm{e},\mathfrak{m}}(h, h^\star, \ldots, h^\star)$ equals $Z_R(\underline{h})$, the partition function on the graph R consisting of the $d/2$ disjoint edges with boundary condition $h^\star$ on all but one vertex which instead receives boundary condition h. Let R' be R with the edge incident to that one vertex disconnected: then $Z_{R'}(h, h^\star, \ldots, h^\star)$ does not depend on h, and

$$\frac{Z_R(h, h^\star, \ldots, h^\star)}{Z_{R'}(h, h^\star, \ldots, h^\star)} = \sum_{\sigma,\sigma'} \psi(\sigma,\sigma') h^\star_\sigma h_{\sigma'} = \frac{Z_S(h, h^\star, \ldots, h^\star)}{Z_{S'}(h, h^\star, \ldots, h^\star)}.$$

Thus it holds for any $\mathfrak{m}$ that

$$\frac{\Psi^{\mathfrak{m}}(h, h^\star, \ldots, h^\star)}{\Psi^{\mathfrak{m}}(h^\star, h^\star, \ldots, h^\star)} = \frac{(Z_S/Z_{S'})(h, h^\star, \ldots, h^\star)}{(Z_R/Z_{R'})(h, h^\star, \ldots, h^\star)} \frac{(Z_R/Z_{R'})(h^\star, h^\star, \ldots, h^\star)}{(Z_S/Z_{S'})(h^\star, h^\star, \ldots, h^\star)} = 1,$$

proving the claim. □

Proof of Propn. 4.2. Let us reparametrize $\mathscr{H}$ in terms of the bias b, defined as in (14). Assume $b^{\mathfrak{f}} < b^{\mathfrak{m}}$, otherwise $(\mathscr{H}^{\mathrm{FK}})^d$ is a single point and there is nothing to prove. With an abuse of notation we write $\Psi(\underline{b})$ for the evaluation of Ψ at the product measure $\underline{h} \in (\mathscr{H}^{\mathrm{FK}})^d$ corresponding to the vector of biases $\underline{b} \equiv (b_1, \ldots, b_d)$. Then

$$\frac{\Psi^{\mathrm{vx}}(\underline{b})}{C^d} = e^B \prod_{j=1}^d (1 + \gamma b_j) + (q-1) \prod_{j=1}^d \Big(1 - \frac{\gamma b_j}{q-1}\Big), \quad \frac{\Psi^{\mathrm{e},\mathfrak{m}}(\underline{b})}{C^{d/2}} = \prod_{(ij) \in \mathfrak{m}} (1 + \gamma b_i b_j)$$

where $C \equiv (e^\beta + q - 1)/q$, $\gamma \equiv (q-1)(e^\beta - 1)/(e^\beta + q - 1) > 0$. Both $\Psi^{\rm vx}$ and $\Psi^{\rm e,sym}$ are affine in each b_j (with $(b_k)_{k\neq j}$ fixed), so their ratio is maximized at one (or both) of the endpoints $b^{\tt f}, b^{\tt m}$. We therefore conclude that $\Psi^{\rm sym}$ must be maximized over $(\mathscr{H}^{\rm FK})^d$ at a corner $\underline{b} \in \{b^{\tt f}, b^{\tt m}\}^d$.

Let $\underline{b}_\ell$ denote the vector which is $b^{\tt m}$ in the first ℓ coordinates, $b^{\tt f}$ in the remaining $d-\ell$ coordinates. Recalling the assumption $b^{\tt f} < b^{\tt m}$, we can express

$$\mathbf{f}^{\rm vx}(\ell) \equiv \log \Psi^{\rm vx}(\underline{b}_\ell) = \log(A_0 e^{a_0 \ell} + A_1 e^{-a_1 \ell}), \quad A_j, a_j > 0.$$

By Jensen's inequality,

$$\log \Psi^{\rm e,sym}(\underline{b}_\ell) \geqslant \frac{1}{(d-1)!!} \sum_{\tt m} \log \Psi^{\rm e,\tt m}(\underline{b}_\ell) \equiv \mathbf{f}^{\rm e}(\ell)$$

where, writing $C_{\tt ss'} \equiv \log(1 + \gamma b^{\tt s} b^{\tt s'})$ for ${\tt s}, {\tt s'} \in \{{\tt f}, {\tt m}\}$, we calculate

$$\mathbf{f}^{\rm e}(\ell) = \frac{\ell(\ell-1) C_{\tt mm} + 2\ell(d-\ell) C_{\tt mf} + (d-\ell)(d-\ell-1) C_{\tt ff}}{[C^{d/2}(d/2)]^{-1} \cdot d(d-1)} = a_4 \ell^2 + a_3 \ell + a_2$$

where $a_4 > 0$ (for $b^{\tt f} < b^{\tt m}$) using the arithmetic-geometric mean inequality.

If we now consider $\mathbf{f} \equiv \mathbf{f}^{\rm vx} - \mathbf{f}^{\rm e}$ as a function of $\ell \in \mathbb{R}$, then

$$\mathbf{f}'(\ell) = \frac{A_0 a_0 e^{a_0 \ell} - A_1 a_1 e^{-a_1 \ell}}{A_0 e^{a_0 \ell} + A_1 e^{-a_1 \ell}} - 2a_4 \ell - a_3$$

tends to $\mp\infty$ as $\ell \to \pm\infty$. Moreover

$$\mathbf{f}'''(\ell) = -\frac{A_0 A_1 (a_0 + a_1)^3 e^{(a_0 - a_1)\ell}}{(A_0 e^{a_0 \ell} + A_1 e^{-a_1 \ell})^3} (A_0 e^{a_0 \ell} - A_1 e^{-a_1 \ell})$$

which can have at most one real zero. Thus $\mathbf{f}'$ has at most one inflection point, hence at most three real zeroes; further, if there are three zeroes then the middle one corresponds to a local minimum of $\mathbf{f}$. But Lem. 4.3 implies $\mathbf{f}(0) = \mathbf{f}(1)$ and $\mathbf{f}(d-1) = \mathbf{f}(d)$, meaning $\mathbf{f}'$ has zeroes inside the intervals $(0,1)$ and $(d-1,d)$. This shows that $\mathbf{f}$ cannot have a local maximum in $[1, d-1]$, so it is maximized over $\{0, 1, \ldots, d\}$ at 0 or d: thus

$$\sup\{\log \Psi(\underline{h}) : \underline{h} \in (\mathscr{H}^{\rm FK})^d\} = \log[\Psi(h^{\tt f}, \ldots, h^{\tt f}) \vee \Psi(h^{\tt m}, \ldots, h^{\tt m})],$$

which by Thm. 4 is precisely $\mathbf{\Phi}^\star$, concluding the proof. □

Proof of Thm. 1. The lower bound $\liminf_n \phi_n \geqslant \mathbf{\Phi}^\star$ for $\beta, B \geqslant 0$ was proved in [DMS13]. For $B > 0$ the matching upper bound $\limsup_n \phi_n \leqslant \mathbf{\Phi}^\star$ follows by combining Thm. 2, Propn. 4.1, and Propn. 4.2. The result for $B = 0$ follows by continuity. □

5. Local maximizers

In view of the proof of Thm. 3, we expect local maximizers of $\boldsymbol{h}$ of the Bethe rate function $\mathbf{\Phi}$ to have the following probabilistic interpretation, which is in the spirit of results of [MMS12]: if $G_n \to_{loc} \boldsymbol{T}_d$, the factor model on G_n conditioned to the subspace of configurations with edge empirical measure close to $\boldsymbol{h}$ should converge locally weakly to the (Bethe) Gibbs measure corresponding to $\boldsymbol{h}$. With this motivation in mind, in this section we study stationary points and local maximizers of the Bethe rate function for the ferromagnetic Potts model. The results obtained here are independent of the rest of the paper.

5.1. **Local maximizers of the Bethe rate function.** We begin with a characterization of local maximizers of the Bethe rate function in the general (d-regular) setting.

Proposition 5.1. *Let $\mathbf{\Phi}$ be the Bethe rate function (9) for a permissive specification $\underline{\psi}$. An interior stationary point $\boldsymbol{h}$ of $\mathbf{\Phi}$ is a local maximizer if and only if, for (X,Y) having (exchangeable) law $\boldsymbol{h}$,*

$$\rho_{XY} \equiv \sup\left\{\frac{\operatorname{Var}\mathbb{E}[\boldsymbol{\varphi}(X,Y)\,|\,X]}{\operatorname{Var}\boldsymbol{\varphi}(X,Y)} : \boldsymbol{\varphi}\not\equiv 0, \boldsymbol{\varphi}(\sigma,\sigma') = \boldsymbol{\varphi}(\sigma',\sigma)\right\} \leqslant \frac{d}{2(d-1)}. \tag{20}$$

Proof. Given specification $\underline{\psi}$, let $\mathbf{\Delta}^{\pm}$ denote the set of symmetric functions $\boldsymbol{\delta} : \mathscr{X}^2 \to \mathbb{R}$ with $\operatorname{supp}\boldsymbol{\delta} \subseteq \operatorname{supp}\psi$, $\langle\boldsymbol{\delta}, 1\rangle = 0$, and $\|\boldsymbol{\delta}\| = 1$. An interior stationary point $\boldsymbol{h}$ of $\mathbf{\Phi}$ is a local maximizer if and only if

$$(\partial_\eta)^2\mathbf{\Phi}(\boldsymbol{h}+\eta\boldsymbol{\delta})|_{\eta=0} = (d-1)\langle(\bar{\delta}/\bar{h})^2\rangle_{\bar{h}} - (d/2)\langle(\boldsymbol{\delta}/\boldsymbol{h})^2\rangle_{\boldsymbol{h}} \leqslant 0 \quad \text{for all } \boldsymbol{\delta}\in\mathbf{\Delta}^{\pm}. \tag{21}$$

Taking $\boldsymbol{\varphi} = (\boldsymbol{h}+\boldsymbol{\delta})/\boldsymbol{h}$ and rearranging gives condition (20). ☐

Remark. The "symmetric correlation coefficient" ρ_{XY} measures dependence within the exchangeable pair (X,Y). (Note ρ_{XY} is not the classical correlation coefficient between $\sigma(X), \sigma(Y)$; see e.g. [DKS01] and references therein.) By the standard variance decomposition $\operatorname{Var}\boldsymbol{\varphi} = \operatorname{Var}(\mathbb{E}[\boldsymbol{\varphi}|X]) + \mathbb{E}[\operatorname{Var}(\boldsymbol{\varphi}|X)]$, we have $0 \leqslant \rho_{XY} \leqslant 1$, with $\rho_{XY} = 1$ if and only if $Y = f(X)$ for some deterministic function f (which by exchangeability must be involutive). For X and Y *independent*, Hoeffding's decomposition

$$\boldsymbol{\varphi} = \tilde{\boldsymbol{\varphi}} + \mathbb{E}[\boldsymbol{\varphi}\,|\,X] + \mathbb{E}[\boldsymbol{\varphi}\,|\,Y] - \mathbb{E}\boldsymbol{\varphi}$$

(with $\mathbb{E}[\tilde{\boldsymbol{\varphi}}\,|\,X] = 0 = \mathbb{E}[\tilde{\boldsymbol{\varphi}}\,|\,Y]$) gives

$$\begin{aligned}&\mathbb{E}[\operatorname{Var}(\boldsymbol{\varphi}\,|\,X)] = \mathbb{E}[(\boldsymbol{\varphi}-\mathbb{E}(\boldsymbol{\varphi}\,|\,X))^2]\\ &= \mathbb{E}[\tilde{\boldsymbol{\varphi}}^2] + 2\,\mathbb{E}[\tilde{\boldsymbol{\varphi}}\cdot(\mathbb{E}[\boldsymbol{\varphi}\,|\,Y]-\mathbb{E}\boldsymbol{\varphi})] + \mathbb{E}[(\mathbb{E}[\boldsymbol{\varphi}\,|\,Y]-\mathbb{E}\boldsymbol{\varphi})^2] = \mathbb{E}[\tilde{\boldsymbol{\varphi}}^2] + \operatorname{Var}(\mathbb{E}[\boldsymbol{\varphi}\,|\,Y]),\end{aligned}$$

where the cross term vanishes using $\mathbb{E}[\tilde{\boldsymbol{\varphi}}\,|\,Y] = 0$. This implies $\rho_{XY} \leqslant 1/2$; and in fact $\rho_{XY} = 1/2$ with supremum achieved by $\boldsymbol{\varphi}(\sigma,\sigma')$ of form $\varphi_\sigma + \varphi_{\sigma'}$. We do not know of an argument to show $\rho_{XY} \geqslant 1/2$ for general exchangeable pairs (X,Y).

5.2. **Fixed points of the Potts Bethe recursion.** Rearranging (5), we see that an interior point h of $\mathscr{H}$ belongs to $\mathscr{H}^\star$ if and only if

$$\frac{\bar{\psi}(\sigma)}{h_\sigma}\Big(\sum_{\sigma'}\psi(\sigma,\sigma')h_{\sigma'}\Big)^{d-1} \quad \text{takes the same value } z_h \text{ for all } \sigma\in\mathscr{X}. \tag{22}$$

We now classify the fixed points for the ferromagnetic Potts model. If $\beta = 0$ the unique solution is given by taking $h_1/h_2 = \cdots = h_2/h_q = e^B$, so assume from now on that $\beta > 0$. We then reparametrize $m \equiv e^B \geqslant 1$, $\theta \equiv 1/(e^\beta - 1) > 0$, so that (22) simplifies to

$$m\boldsymbol{F}(h_1) = \boldsymbol{F}(h_2) = \ldots = \boldsymbol{F}(h_q) = z_h, \quad \boldsymbol{F}(x) \equiv x^{-1}(x/\theta+1)^{d-1}.$$

With $h^\circ \equiv \theta/(d-2)$, we observe that the restriction $\boldsymbol{F}_- \equiv \boldsymbol{F}|_{(0,h^\circ]}$ is monotone decreasing while the restriction $\boldsymbol{F}_+ \equiv \boldsymbol{F}|_{[h^\circ,1]}$ is monotone increasing, so clearly $|\{h_2,\ldots,h_q\}| \leqslant 2$. More precisely, we have the following classification:

Definition 5.2. Say that $h\in\mathscr{H}^\star$ is an ℓ-type solution if among the entries $h_2,\ldots,h_q$ it has $\ell-1$ coordinates equal to $p_+ \equiv (\boldsymbol{F}_+)^{-1}(z_h)$, and $q-\ell$ coordinates equal to $p_- \equiv (\boldsymbol{F}_-)^{-1}(z_h)$. We further subdivide the ℓ-type solutions into ℓ_+-type or ℓ_--type according to whether h_1 equals $Q_+ \equiv (\boldsymbol{F}_+)^{-1}(z_h/m)$ or $Q_- \equiv (\boldsymbol{F}_-)^{-1}(z_h/m)$.

This terminology degenerates in some cases, in particular when $m = 1$: in this case $Q_\pm = p_\pm$, so the ℓ_+-type solutions coincide (up to permuting the labels) with $(\ell+1)_-$-type solutions for $1 \leqslant \ell < q$, and the only 1_-- or q_+-type solution is the uniform distribution on $[q]$.

Proposition 5.3. *For the Potts model with $\theta \equiv 1/(e^\beta - 1) > 0$, $m \equiv e^B \geqslant 1$ (both finite), every fixed point $h \in \mathscr{H}^\star$ is an ℓ-type solution for some $1 \leqslant \ell \leqslant q$. Further*

(a) If $h^\circ \equiv \theta/(d-2) \geqslant 1$ then all solutions are 1_--type.
For $\ell \geqslant 2$, if $m\boldsymbol{F}(h^\circ) \geqslant \boldsymbol{F}[1/(\ell-1)]$ then there are no solutions of type $\geqslant \ell$.

(b) If $qh^\circ < 1$, there are no 1_--type solutions; but for each $\ell \geqslant 2$ there exists $m_\ell(\theta) > 1$ such that both ℓ_+- and ℓ_--type solutions exist for all $1 \leqslant m \leqslant m_\ell(\theta)$.

Proof. It is clear from the preceding discussion that every $h \in \mathscr{H}^\star$ is an ℓ-type solution for some $1 \leqslant \ell \leqslant q$. Fixed points of (15) correspond to 1- or q-type solutions.

(a) If $h^\circ \geqslant 1$ then $\boldsymbol{F} = \boldsymbol{F}_-$ is injective on $(0, 1]$ so necessarily $\ell = 1$ and solutions are 1_--type. If $h \in \mathscr{H}^\star$ has $\ell \geqslant 2$ then, since $\min_\sigma h_\sigma > 0$ while $\max_\sigma h_\sigma = p_+$, necessarily $p_+ < 1/(\ell-1)$. Thus for $\boldsymbol{F}(Q) = \boldsymbol{F}(p_+)/m$ to have any solution we must have $m\boldsymbol{F}(h^\circ) < \boldsymbol{F}[1/(\ell-1)]$.

(b) A 1_--type solution must have $\max_\sigma h_\sigma \leqslant h^\circ$, so if $qh^\circ < 1$ then no such solution exists. For $\ell \geqslant 2$ and $\mathsf{s} \in \{\pm\}$, the function

$$\mathbf{g}^{\ell\mathsf{s}}(p; m) \equiv (\boldsymbol{F}_\mathsf{s})^{-1}[\boldsymbol{F}(p)/m] + (\ell-1)(\boldsymbol{F}_+)^{-1}[\boldsymbol{F}(p)] + (q-\ell)(\boldsymbol{F}_-)^{-1}[\boldsymbol{F}(p)]$$

is well-defined for $p_0(m) \leqslant p \leqslant 1$ where $p_0(m) \equiv (\boldsymbol{F}_+)^{-1}[m\boldsymbol{F}(h^\circ)]$ is less than 1 for small enough $m \geqslant 1$, due to the assumption $qh^\circ < 1$. Note

$$\mathbf{g}^{\ell\mathsf{s}}(p; m) > \ell - 1 \geqslant 1 \quad \text{while} \quad \lim_{m \downarrow 1} \mathbf{g}^{\ell\mathsf{s}}(p_0(m); m) = qh^\circ < 1.$$

It follows by continuity that if $\ell \geqslant 2$ and $qh^\circ < 1$ then $\mathbf{g}^{\ell\mathsf{s}}(p; m) = 1$ for some $p_0(m) \leqslant p \leqslant 1$, giving an ℓ_s-type solution as claimed. □

5.3. **Local maximizers for the Potts Bethe rate function.** We next study which of the stationary points classified in Propn. 5.3 correspond to local maximizers for $\boldsymbol{\Phi}$.

Proposition 5.4. *In the setting of Propn. 5.3,*

(a) Solutions of type $\ell > 2$ are never local maximizers.

(b) For $m \geqslant 1, \theta > 0$ both sufficiently small, there exist both 1_+-type and 2_--type solutions which are strict local maximizers, with strictly negative-definite Hessians.

Proof. (a) Let $\boldsymbol{h} \in \boldsymbol{\Delta}$ be the stationary point of $\boldsymbol{\Phi}$ corresponding to $h \in \mathscr{H}^\star$ via (10): then

$$\boldsymbol{h}(\sigma' \,|\, \sigma) \equiv \frac{\boldsymbol{h}(\sigma, \sigma')}{\bar{h}_\sigma} = \frac{h_{\sigma'}(\theta + \mathbf{1}\{\sigma = \sigma'\})}{\theta + h_\sigma}.$$

We will apply the correlation criterion (20) with $\boldsymbol{\varphi}(\sigma, \sigma') \equiv \varphi_\sigma + \varphi_{\sigma'}$. If $\langle \varphi \rangle_h = 0$, then

$$\mathbb{E}[\varphi_Y \,|\, X = \sigma] = \sum_{\sigma'} \varphi_{\sigma'} \boldsymbol{h}(\sigma' \,|\, \sigma) = \frac{\theta}{\theta + h_\sigma} \langle \varphi \rangle_h + \frac{h_\sigma}{\theta + h_\sigma} \varphi_\sigma = \gamma_\sigma \varphi_\sigma, \quad \gamma_\sigma \equiv \frac{h_\sigma}{\theta + h_\sigma}.$$

Thus $\mathbb{E}[\boldsymbol{\varphi}(X, Y) \,|\, X] = (1 + \gamma_X)\varphi_X$, and (20) becomes

$$\begin{aligned} 2(\mathbb{E}\varphi_X)^2 &\geqslant \mathbb{E}\Big[(1 + \gamma_X)(\varphi_X)^2 \frac{(d-1)\gamma_X - 1}{d-2}\Big] \\ &= \mathbb{E}\Big[(1 + \gamma_X)(\varphi_X)^2 \frac{h_X - h^\circ}{\theta + h_X}\Big] = \frac{\sum_\sigma h_\sigma (1 + \gamma_\sigma)(h_\sigma - h^\circ)(\varphi_\sigma)^2}{\theta + \|h\|^2} \end{aligned} \tag{23}$$

(using $\bar{h}_\sigma = h_\sigma(\theta + h_\sigma)/(\theta + \|h\|^2)$ for the last identity). Now suppose without loss that $h_2 \geqslant \ldots \geqslant h_q$: if h is an ℓ-type solution with $\ell > 2$, then $\varphi_\sigma \equiv \mathbf{1}\{\sigma = 2\} - \mathbf{1}\{\sigma = 3\}$ clearly violates (23), so $\boldsymbol{h}$ cannot be a local maximizer of $\boldsymbol{\Phi}$.

(b) Let $m = 1$ and $\theta > 0$ sufficiently small so that a 1_+-type (equivalently, up to permuting the labels, 2_--type) solution $h = (Q_+, p_-, \ldots, p_-)$ exists, given by taking the log-likelihood ratio $\mathbf{r} \equiv \log(Q_+/p_-)$ to be the maximal fixed point of the mapping $\widetilde{\mathrm{BP}}$ of (15). For $d \geqslant 3$ and $0 < \epsilon \leqslant 1$ we calculate that

$$\widetilde{\mathrm{BP}}[(d-1-\epsilon)\beta] \geqslant (d-1)(\beta - \log q) > (d-1-\epsilon)\beta \quad \text{for } \epsilon\beta > (d-1)\log q,$$

so crudely we have $\mathbf{r} \geqslant (3/2)\beta$ for all $\beta \geqslant 2(d-1)\log q$. Let $\boldsymbol{h} \in \boldsymbol{\Delta}$ be the stationary point corresponding to this fixed point: recalling (21), for $\boldsymbol{\delta} \in \boldsymbol{\Delta}^{\pm}$ we calculate

$$\frac{\langle(\bar\delta/\bar h)^2\rangle_{\bar h}}{\boldsymbol{z}_h} = \frac{\bar\delta(1)^2}{Q_+(e^\beta Q_+ + (q-1)p_-)} + \frac{\sum_{\sigma\neq 1}\bar\delta(\sigma)^2}{p_-(Q_+ + (e^\beta + q - 2)p_-)} \leqslant \frac{\bar\delta(1)^2}{e^\beta Q_+^2} + \frac{\sum_{\sigma\neq 1}\bar\delta(\sigma)^2}{Q_+ p_-},$$

$$\frac{\langle(\boldsymbol{\delta}/\boldsymbol{h})^2\rangle_{\boldsymbol{h}}}{\boldsymbol{z}_h} \geqslant \frac{\boldsymbol{\delta}(1,1)^2}{Q_+^2 e^\beta} + \frac{2\sum_{\sigma\neq 1}\boldsymbol{\delta}(1,\sigma)^2}{Q_+ p_-} + \frac{\sum_{\sigma,\sigma'\neq 1}\boldsymbol{\delta}(\sigma,\sigma')^2}{e^\beta p_-^2}.$$

Since $p_- \leqslant e^{-(3/2)\beta} Q_+$ for sufficiently large β,

$$\lim_{\beta\to\infty} \frac{e^\beta p_-^2}{\boldsymbol{z}_h}\langle(\bar\delta/\bar h)^2\rangle_{\bar h}^2 = 0, \quad \liminf_{\beta\to\infty} \frac{e^\beta p_-^2}{\boldsymbol{z}_h}\langle(\boldsymbol{\delta}/\boldsymbol{h})^2\rangle_{\boldsymbol{h}} \geqslant \sum_{\sigma,\sigma'\neq 1}\boldsymbol{\delta}(\sigma,\sigma')^2,$$

so for any fixed $\epsilon > 0$ we have $(\partial_\eta)^2\boldsymbol{\Phi}(\boldsymbol{h} + \eta\boldsymbol{\delta})|_{\eta=0} < 0$ uniformly over all $\boldsymbol{\delta} \in \boldsymbol{\Delta}^{\pm}$ satisfying $(q-1)^2\sum_{\sigma,\sigma'\neq 1}\boldsymbol{\delta}(\sigma,\sigma')^2 \geqslant \epsilon^2$ once β is sufficiently large (depending on ϵ). Suppose instead $(q-1)^2\sum_{\sigma,\sigma'\neq 1}\boldsymbol{\delta}(\sigma,\sigma')^2 \leqslant \epsilon^2$: by Cauchy–Schwarz $\sum_{\sigma,\sigma'\neq 1}|\boldsymbol{\delta}(\sigma,\sigma')| \leqslant \epsilon$, so

$$\begin{aligned}
&[\beta \to \infty]\limsup \frac{Q_+ p_-}{\boldsymbol{z}_h}\Big[2(d-1)\langle(\bar\delta/\bar h)^2\rangle_{\bar h} - d\langle(\boldsymbol{\delta}/\boldsymbol{h})^2\rangle_{\boldsymbol{h}}\Big] \\
&\leqslant 2(d-1)\sum_{\sigma\neq 1}\bar\delta(\sigma)^2 - 2d\sum_{\sigma\neq 1}\boldsymbol{\delta}(1,\sigma)^2 \leqslant 2(d-1)\sum_{\sigma\neq 1}[|\boldsymbol{\delta}(1,\sigma)| + \epsilon]^2 - 2d\sum_{\sigma\neq 1}\boldsymbol{\delta}(1,\sigma)^2 \\
&\leqslant -2\sum_{\sigma\neq 1}\boldsymbol{\delta}(1,\sigma)^2 + 2(d-1)\Big[2\epsilon\sum_{\sigma\neq 1}|\boldsymbol{\delta}(1,\sigma)| + (q-1)\epsilon^2\Big].
\end{aligned}$$

On the other hand, $\boldsymbol{\delta} \in \boldsymbol{\Delta}^{\pm}$ implies

$$2\Big|\sum_{\sigma\neq 1}\boldsymbol{\delta}(1,\sigma)\Big| = |\boldsymbol{\delta}(1,1) + \epsilon| \geqslant |\boldsymbol{\delta}(1,1)| - \epsilon \geqslant \Big[1 - 2\sum_{\sigma\neq 1}\boldsymbol{\delta}(1,\sigma)^2 - \frac{\epsilon^2}{(q-1)^2}\Big]^{1/2} - \epsilon$$

so by choosing $\epsilon > 0$ sufficiently small we can guarantee that for β large enough we will have $(\partial_\eta)^2\boldsymbol{\Phi}(\boldsymbol{h}+\eta\boldsymbol{\delta})|_{\eta=0} < 0$ uniformly over all $\boldsymbol{\delta} \in \boldsymbol{\Delta}^{\pm}$, implying that $\boldsymbol{h}$ is a strict local maximizer of $\boldsymbol{\Phi}$ with strictly negative-definite Hessian.

This concludes the proof for $m = 1$, and the conclusion for $m > 1$ sufficiently small follows by a perturbative argument: arguing similarly as in the proof of Propn. 5.3b, for $1 \leqslant m < m_0$ the equations

$$\begin{aligned}
\mathbf{g}^{1+}(p) &\equiv (\boldsymbol{F}_+)^{-1}[\boldsymbol{F}(p)/m] + (q-1)p = 1, \\
\mathbf{g}^{2-}(p) &\equiv (\boldsymbol{F}_-)^{-1}[\boldsymbol{F}(p)/m] + (\boldsymbol{F}_+)^{-1}[\boldsymbol{F}(p)] + (q-2)p = 1
\end{aligned}$$

have solutions $p_-^{1+}(m), p_-^{2-}(m)$, corresponding to 1_+-type and 2_--type solutions respectively, which are continuous in m with initial values $p^{1+}(1) = p^{2-}(1) = p_-$ corresponding to the solution considered above at $m = 1$. For sufficiently small m, it follows by continuity that the Hessians at the stationary points $\boldsymbol{h}^{1+}(m), \boldsymbol{h}^{2-}(m)$ corresponding to $p_-^{1+}(m), p_-^{2-}(m)$ will be strictly negative-definite, implying strict local maximizers as claimed. □

Remark 5.5. Related to the study of local maxima is the question of the local stability of the Bethe recursion. For the Potts specification (2), the linear (differential) mapping $D_h \equiv D\mathsf{BP}(h)$ defined on the space $\{\delta : \sum_\sigma \delta_\sigma = 0\}$ by

$$D_h\delta \equiv \lim_{\eta\to 0} \frac{\mathsf{BP}(h+\eta\delta) - \mathsf{BP}(h)}{\eta}$$

can be explicitly diagonalized when $h \in \mathscr{H}^\star$ and shown to have all eigenvalues positive, with maximal eigenvalue greater than 1 at ℓ-type solutions with $\ell > 2$ and at 2_+-type solutions. At a 2_--type solution (assuming $m > 1$, so it is not also a 1_+-type solution) the maximal eigenvalue is less than 1 if and only if

$$\frac{p_+^2}{p_+ - h^\circ} > \frac{d-2}{d-1} + \frac{Q_-^2}{h^\circ - Q_-} + (q-2)\frac{p_-^2}{h^\circ - p_-}. \tag{24}$$

However, if h is not the uniform measure on $[q]$ then D_h is *not* symmetric and so does not have orthonormal eigenbasis, so having all eigenvalues less than 1 need not imply contractivity of D_h. It is not clear how to relate (24) to the local stability of the non-linear map BP.

Acknowledgements. We thank the anonymous referee for helpful comments.

References

[ASS03] M. Aizenman, R. Sims, and S. L. Starr. Extended variational principle for the Sherrington–Kirkpatrick spin-glass model. *Phys. Rev. B*, 68:214403, Dec 2003.

[ATA73] R. Abou-Chacra, D. Thouless, and P. Anderson. A selfconsistent theory of localization. *J. Phys. C*, 6(10):1734, 1973.

[AW06] M. Aizenman and S. Warzel. The canopy graph and level statistics for random operators on trees. *Math. Phys. Anal. Geom.*, 9(4):291–333 (2007), 2006.

[BCKL13] C. Borgs, J. Chayes, J. Kahn, and L. Lovász. Left and right convergence of graphs with bounded degree. *Rand. Struct. Alg.*, 42(1):1–28, 2013.

[Bet35] H. A. Bethe. Statistical theory of superlattices. *P. Roy. Soc. Lond. A*, 150(871):552–575, 1935.

[BGT10] M. Bayati, D. Gamarnik, and P. Tetali. Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. In *Proc. ACM Symp. (STOC)*, pages 105–114. ACM, New York, 2010.

[BS01] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6(23):13 pp. (electronic), 2001.

[CCST86] J. T. Chayes, L. Chayes, J. P. Sethna, and D. J. Thouless. A mean field spin glass with short-range interactions. *Comm. Math. Phys.*, 106(1):41–89, 1986.

[CDGS13] P. Contucci, S. Dommers, C. Giardinà, and S. Starr. Antiferromagnetic Potts model on the Erdős-Rényi random graph. *Comm. Math. Phys.*, 323(2):517–554, 2013.

[CLR79] J. Chalupa, P. Leath, and G. Reich. Bootstrap percolation on a Bethe lattice. *J. Phys. C*, 12(1):L31, 1979.

[DGvdH10] S. Dommers, C. Giardinà, and R. van der Hofstad. Ising models on power-law random graphs. *J. Stat. Phys.*, 141(4):638–660, 2010.

[DKS01] A. Dembo, A. Kagan, and L. A. Shepp. Remarks on the maximum correlation coefficient. *Bernoulli*, 7(2):343–350, 2001.

[DM10a] A. Dembo and A. Montanari. Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.*, 24(2):137–211, 2010.

[DM10b] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, 20(2):565–592, 2010.

[DMS13] A. Dembo, A. Montanari, and N. Sun. Factor models on locally tree-like graphs. *Ann. Probab.*, 41(6):4162–4213, 2013.

[Gue03] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.

[JŁR00] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[KMR+07] F. Krząkała, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proc. Natl. Acad. Sci. USA*, 104(25):10318–10323 (electronic), 2007.

[McD98] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.

[MM09] M. Mézard and A. Montanari. *Information, physics, and computation.* Oxford Graduate Texts. Oxford University Press, Oxford, 2009.

[MMS12] A. Montanari, E. Mossel, and A. Sly. The weak limit of Ising models on locally tree-like graphs. *Probab. Theory Related Fields*, 152(1-2):31–51, 2012.

[MP01] M. Mézard and G. Parisi. The Bethe lattice spin glass revisited. *Eur. Phys. J. B Condens. Matter Phys.*, 20(2):217–233, 2001.

[MPV87] M. Mézard, G. Parisi, and M. A. Virasoro. *Spin glass theory and beyond*, volume 9 of *World Scientific Lecture Notes in Physics.* World Scientific Publishing Co. Inc., Teaneck, NJ, 1987.

[MPZ02] M. Mézard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297(5582):812–815, 2002.

[Rob55] H. Robbins. A remark on Stirling's formula. *Amer. Math. Monthly*, 62:26–29, 1955.

[Spi75] F. Spitzer. Markov random fields on an infinite tree. *Ann. Probab.*, 3(3):387–398, 1975.

[SS13] A. Sly and N. Sun. Counting in two-spin models on d-regular graphs. *Ann Probab., to appear*, 2013.

[Tal11] M. Talagrand. *Mean field models for spin glasses. Volume I*, volume 54 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge.* Springer-Verlag, Berlin, 2011. Basic examples.

[Tho86] D. Thouless. Spin-glass on a Bethe lattice. *Phys. Rev. Lett.*, 56(10):1082, 1986.

[Wei48] P. R. Weiss. The application of the Bethe-Peierls method to ferromagnetism. *Phys. Rev.*, 74(10):1493, 1948.

[Zac83] S. Zachary. Countable state space Markov random fields and Markov chains on trees. *Ann. Probab.*, 11(4):894–903, 1983.

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

†Department of Electrical Engineering, Stanford University
350 Serra Mall, Stanford, California 94305

‡Department of Statistics, University of California, Berkeley
Evans Hall, Berkeley, California 94720; and
Mathematical Science Institute, Australian National University
John Dedman Building, Union Lane, Canberra, ACT 0200

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