Reconstruction for models on random graphs

Antoine Gerschenfeld
Ecole Normale Supérieure
45, rue d'Ulm 75005 Paris, France
gerschen@clipper.ens.fr

Andrea Montanari
Departments of Electrical Engineering and Statistics
Stanford University, Stanford CA-9305 USA
montanari@stanford.edu

Abstract

Consider a collection of random variables attached to the vertices of a graph. The reconstruction problem requires to estimate one of them given 'far away' observations. Several theoretical results (and simple algorithms) are available when their joint probability distribution is Markov with respect to a tree. In this paper we consider the case of sequences of random graphs that converge locally to trees. In particular, we develop a sufficient condition for the tree and graph reconstruction problem to coincide. We apply such condition to colorings of random graphs.

Further, we characterize the behavior of Ising models on such graphs, both with attractive and random interactions (respectively, 'ferromagnetic' and 'spin glass').

1. Introduction and outline

Let $G = (V, E)$ be a graph, and $X = \{X_i : i \in V\}$ a proper coloring of its vertices sampled uniformly at random. The *reconstruction problem* amounts to estimating the color of a distinguished (root) vertex $r \in V$, when the colors $\{X_j = x_j : j \in U\}$ of subset of vertices are revealed. In particular, we want to understand whether the revealed values induce a substantial bias on the distribution of X_i.

We shall consider the more general setting of *graphical models*. Such a model is defined by a graph $G = (V, E)$, and a set of weights $\psi = \{\psi_{ij} : (ij) \in E\}$, $\psi_{ij} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$. Given a graph-weights pair (G, ψ), we let

$$\mathbb{P}\big\{X = x\big|(G,\psi)\big\} \equiv \frac{1}{Z} \prod_{(ij)\in E} \psi_{ij}(x_i, x_j)\,, \qquad (1)$$

where we assume $\psi_{ij}(x, y) = \psi_{ij}(y, x)$. The example of proper colorings is recovered by letting $\mathcal{X} = \{1, \dots, q\}$ (q being the number of colors) and $\psi_{ij}(x, y) = 1$ if $x \neq y$ and $= 0$ otherwise. Ising models from statistical mechanics provide another interesting class, whereby $\mathcal{X} = \{+1, -1\}$. In the 'ferromagnetic' case the weights are $\psi_{ij}(+,+) = \psi_{ij}(-,-) = 1 - \epsilon$ and $\psi_{ij}(+,-) = \psi_{ij}(-,+) = \epsilon$ for some $\epsilon \in [0, 1/2]$.

For economy of notation, we shall often write $\mathbb{P}\{\,\cdot\,|G\}$ as a shorthand for $\mathbb{P}\{\,\cdot\,|(G,\psi)\}$, and 'the graph G' for 'the graph-weights pair (G, ψ).' It is understood that, whenever G is given, the weights ψ are given as well. Further, for $U \subseteq V_N$, we let $X_U = \{X_j : j \in U\}$ and $\mathbb{P}_U\{x_U|G\} = \mathbb{P}\{X_U = x_U|G\}$ be its marginal distribution that can be obtained by marginalizing Eq. (1).

For $i, j \in V$, let $d(i, j)$ be their graph theoretic distance. Further for any $t \geq 0$, we let $\overline{\mathsf{B}}(i, t)$ be the set of vertices j such that $d(i, j) \geq t$, (and, by abuse of notation, the induced subgraph). The reconstructibility question asks whether the 'far away' variables $X_{\overline{\mathsf{B}}(r,t)}$ provide significant information about X_r. This is captured by the following definition (recall that, given two distributions p, q on the same space $\mathcal{S}$, their total variation distance is $||p - q||_{\text{TV}} \equiv (1/2)\sum_{x\in\mathcal{S}} |p_x - q_x|$).

Definition 1.1. *The reconstruction problem is* (t, ε)-solvable *(*reconstructible*) for the graphical model* (G, ψ) *rooted at* $r \in V$ *if*

$$\|\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|G\} - \mathbb{P}_r\{\,\cdot\,|G\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|G\}\|_{\text{TV}} \geq \varepsilon\,.$$

In the following we will consider graphs G that are themselves random. By this we mean that we will specify a joint distribution of the graph $G_N = (V_N = [N], E_N)$, of the weights $\{\psi_{ij}\}$, and of the root vertex r whose variable we are interested in reconstructing. Equation (1) then specifies the conditional distribution of X, given the random structure (G_N, ψ) (again, we'll drop reference to ψ).

Definition 1.2. *The reconstruction problem is* solvable *(*reconstructible*) for the sequence of random graphical models* $\{G_N\}$ *if there exists* $\varepsilon > 0$ *such that, for all* $t \geq 0$ *it is* (t, ε)*-solvable with positive probability, i.e. if*

$$\|\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|G_N\}\|_{\text{TV}} \geq \varepsilon\,. \qquad (2)$$

with positive probability[1].

[1] Here and below, we say that the sequence of events $\{A_N\}$ holds with

To be specific, we shall assume G_N to be a sparse random graph. In this case, any finite neighborhood of r converges in distribution to a tree [1]. Further, imagine to mark the boundary vertices of such a neighborhood, and then take the neighborhood out of G_N (thus obtaining the subgraph denoted above as $\overline{\mathsf{B}}(r,t)$). The marked vertices will be (with high probability) 'far apart' from each other in $\overline{\mathsf{B}}(r,t)$. This suggests that the corresponding random variables $\{X_j\}$ will be approximately independent when the tree-like neighborhood is taken out. Hence, approximating G_N by its local tree structure might be a good way to determine correlations between X_r and the boundary variables $\{X_j : d(r,j) = t\}$. In other words, one would expect reconstructibility on G_N to be determined by reconstructibility on the associated random tree.

Of course the above conclusion does not hold in general, as it is based on a circular argument. We assumed that 'far apart' variables (with respect to the residual graph $\overline{\mathsf{B}}(r,t)$) are weakly correlated, to understand whether 'far apart' variable (in G_N) are. In fact, we will prove that tree and graph reconstruction do not coincide in the simplest example one can think of, namely the Ising ferromagnet (binary variables with attractive interactions).

On the positive side, we prove a general sufficient condition for the tree and graph behaviors to coincide. The condition has a suggestive geometrical interpretation, as it requires two independent random configurations $X^{(1)}$ and $X^{(2)}$ to be, with high probability, at an approximately fixed 'distance' from each other. In the example of coloring, we require two uniformly random independent colorings of the same graph to take the same value on about $1/q$ of the vertices. The set of 'typical configurations' looks like a sphere when regarded from any typical configuration. Under such a condition, the above argument can be put on firmer basis. We show that, once the the neighborhood of the root r is taken out, boundary variables become roughly independent. This in turns implies that graph and tree reconstruction do coincide.

We apply this sufficient condition to the Ising spin glass (where the condition can be shown to hold as a consequence of a recent result by Guerra and Toninelli [2]), and to antiferromagnetic colorings of random graphs (building on the work of Achlioptas and Naor [3]). In both cases we will introduce a family of graphical models parametrized by their average degree. It is natural to expect reconstructibility to hold at large degrees (as the graph is 'more connected') and not to hold at small average degrees (since the graph 'falls' apart into disconnected components). In the spin glass case we are indeed able to estabilish a threshold behavior (i.e. a critical degree value above which reconstruction is solvable). While we didn't achieve the same for colorings, we essentially reduced the problem to establishing a threshold for the tree model.

positive probability (wpp) if there exists $\delta > 0$ and an infinite sequence $\mathcal{N} \subseteq \mathbb{N}$, such that $\mathbb{P}\{A_N\} \geq \delta$ for any $N \in \mathcal{N}$. Notice that, in a random graph, r might be in a small connected component. Therefore Eq. (2) cannot be required to hold with high probability.

1.1. Applications and related work

Let us discuss a selection of related problems that are relevant to our work.

Markov Chain Monte Carlo (MCMC) algorithms provide a well established way of approximating marginals of the distribution (1). If the chain is reversible and has local updates, the mixing time is known to be related to the correlation decay properties of the stationary distribution $\mathbb{P}\{\,\cdot\,|G_N\}$ [4, 5]. In this context, correlations between X_r and $X_{\overline{\mathsf{B}}(r,t)}$ are usually characterized by measures of the type $\Delta(t) \equiv \sup_x \|\mathbb{P}_{r|\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|x_{\overline{\mathsf{B}}(r,t)}, G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}\|_{\text{TV}}$. The 'uniqueness ' condition requires $\Delta(t)$ to decay at large t, and is easily shown to imply non-reconstructibility. On graphs with sub-exponential growth, a fast enough decay is a necessary and sufficient condition for fast mixing. On the other hand, in more general cases this is too strong a criterion, and one might want to replace it with the non-reconstructibility one.

In [6] it was proved that non-reconstructibility is equivalent to polynomial spectral gap for a class of models on trees. The equivalence was sharpened in [7], showing that non-reconstructibility is equivalent to fast mixing in the same models. Further, [6] proved that non-reconstructibility is a necessary condition for fast mixing on general graphs. While a converse does not hold in general, non-reconstructibility is sufficient for rapid decay of the variance of local functions (which is often regarded as the criterion for fast dynamics in physics) [8].

Random constraint satisfaction problems. Given an instance of a constraint satisfaction problem (CSP), consider the uniform distribution over its solutions. This takes the form (1), where ψ_{ij} is the indicator function over the constraint involving variables x_i, x_j being satisfied (Eq. (1) is trivially generalized to k-variables constraints). For instance, in coloring it is the indicator function on $x_i \neq x_j$.

Computing the marginal $\mathbb{P}_r\{\,\cdot\,|G_N\}$ can be useful both for finding and for counting the solutions of such a CSP. Assume to be able to generate *one* uniformly random solution X. In general, this is not sufficient to approximate the marginal of X_i in any meaningful way. However one can try the following: fix all the variables 'far from r' to take the same value as in the sampled configuration, namely $X_{\overline{\mathsf{B}}(r,t)}$, and compute the conditional distribution at the root. If the graph is locally tree-like, the conditional distribution of X_r can be computed through an efficient dynamic programming procedure. The result of this computation needs not to be near the actual marginal. However, non-

reconstructibility implies the result to be with high probability within ε (in total variation distance) from the marginal.

As a consequence, a single sample (a single random solution x) is sufficient to approximate the marginal $\mathbb{P}_r\{\,\cdot\,|G_N\}$. The situation is even simpler under the sufficient condition in our main theorem (Theorem 1.4). In fact this implies that the boundary condition $x_{\overline{\mathsf{B}}(r,t)}$ can be replaced by an iid uniform boundary.

For random CSP's, G_N becomes a sparse random graph. Statistical mechanics studies [9] suggest that, for typical instances the set of solutions decomposes into 'clusters' at sufficiently large constraint density [10, 11]. This leads to the speculation that sampling from the uniform measure $\mathbb{P}\{\,\cdot\,|G_N\}$ becomes harder in this regime.

The decomposition in clusters is related to reconstructibility, as per the following heuristic argument. Assume the set of solutions to be splitted into clusters, and that two solutions whose Hamming distance is smaller than $N\varepsilon$ belong to the same cluster. Then knowing the far away variables $x_{\overline{\mathsf{B}}(r,t)}$ (i.e. all but a bounded number of variables) does determine the cluster. This in turns provides some information on X_r.

In fact, it was conjectured in [12] that tree and graph reconstruction thresholds should coincide for 'frustrated' models on random graphs. Both should coincide with the clustering phase transition in the set of solutions [13].

Statistical inference and message passing. Graphical models of the form (1) are used in a number of contexts, from image processing to artificial intelligence, etc. Statistical inference requires to compute marginals of such a distribution and message passing algorithms (in particular, belief propagation, BP) are the methods of choice for accomplishing this task.

The (unproven) assumption in such algorithms is that, if a tree neighborhood of vertex i is cut away from G_N, then the variables $\{X_j\}$ on the boundary of this tree are approximately independent. Assuming the marginals of the boundary variables to be known, the marginal of X_i can be computed through dynamic programming. Of course the marginals to start from are unknown. However, the dynamic programming procedure defines an mapping on the marginals themselves. In BP this mapping is iterated recursively over all the nodes, without convergence guarantees.

Lemma 3.2 shows that, under the stated conditions, the required independence condition does indeed hold. As stressed above, this is instrumental in proving equivalence of graph and tree reconstructibility in Theorem 1.4.

The connection with message passing algorithm is further explored in [14]. Roughly speaking that paper proves that, if the reconstruction problem is unsolvable, than BP admits an approximate fixed point that allows to compute the correct marginals.

Reconstruction problems also emerge in a variety of other contexts: (i) Phylogeny [15] (given some evolved genomes, one aims at reconstructing the genome of their common ancestor); (ii) Network tomography [16] (given end-to-end delays in a computer network, infer the link delays in its interior); (iii) Statistical mechanics [17, 18] (reconstruction being related to the extremality of Gibbs measures).

1.2. Previous results

If the graph G_N is a tree, the reconstruction problem is relatively well understood [19]. The fundamental reason is that the distribution $\mathbb{P}\{X = x|G_N\}$ admits a simple description. First sample the root variable X_r from its marginal $\mathbb{P}\{X_r = x_r|G_N\}$. Then recursively for each node j, sample its children $\{X_l\}$ independently conditional on their parent value.

Because of this Markov structure, one can prove a recursive distributional equation for the conditional marginal at the root $\mathbb{P}_{r|\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|X_{\overline{\mathsf{B}}(r,t)}, G_N\} \equiv \eta_t(\,\cdot\,)$ given the variable values at generation t. Notice that this is a random quantity even for a deterministic graph G_N, because $X_{\overline{\mathsf{B}}(r,t)}$ is itself drawn randomly from the distribution $\mathbb{P}\{\,\cdot\,|G_N\}$. Further, it contains all the information (it is a 'sufficient statistic') in the boundary about the root variable X_r. In fact asymptotic behavior of $\eta_t(\,\cdot\,)$ as $t \to \infty$ then determines the solvability of the reconstruction problem. Studying the asymptotic behavior of the sequence $\eta_t(\,\cdot\,)$ (which satisfies a recursive distributional equation) is the standard approach to tree reconstruction.

Among the other results, reconstructibility has been thoroughly characterized for Ising models on generic trees [18, 20, 21]. For an infinite tree T the reconstruction problem is solvable if and only if $\mathrm{br}(\mathsf{T})(1-2\epsilon)^2 > 1$, whereby (for the cases treated below) $\mathrm{br}(\mathsf{T})$ coincides with the mean descendant number of any vertex. This result establishes a sharp threshold in the tree average degree (or in the parameter ϵ), that we shall generalize to random graphs below. However, as we will see, the behavior is richer than in the tree case.

Reconstruction on general graphs poses new challenges, since the above recursive description of the measure $\mathbb{P}\{\,\cdot\,|G_N\}$ is lacking. The result of [6] allows to deduce non-reconstructibility from fast mixing of reversible MCMC with local updates. However, proving fast mixing is far from an easy task. Further, the converse does not usually hold (one can have slow mixing and non-reconstructibility).

An exception is provided by the recent paper by Mossel, Weitz and Wormald [22] that establishes a threshold for fast mixing for weighted independent sets on random bipartite graphs (the threshold being in the weight parameter λ). Arguing as in Section 5, it can be shown that this is also the graph reconstruction threshold. This result is analogous

to ours for the ferromagnetic Ising model: it provides an example in which the graph reconstruction threshold does not coincide with the tree reconstruction threshold. In both cases the graph reconstruction threshold coincides instead with the tree 'uniqueness threshold' (i.e. the critical parameter for the uniqueness condition mentioned above to hold).

1.3. Basic definitions

We consider two families of random graphical models: *regular* and *Poisson* models. In both cases the root $r \in V$ is uniformly random and independent of G_N. A *regular ensemble* is specified by assigning an alphabet $\mathcal{X}$ (the variable range), a degree $(k+1)$ and and edge weight $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$. For any $N > 0$, a random model is defined by letting G_N be a uniformly random regular graph of degree $(k+1)$ over vertex set $V = [N]$. The joint distribution of $(X_1, \dots, X_N)$ is given by Eq. (1), with $\psi_{ij}(\,\cdot\,,\,\cdot\,) = \psi(\,\cdot\,,\,\cdot\,)$. A variation of this ensemble is obtained by letting G be a random regular *multi-graph* according to the configuration model [23] (notice that our definitions make sense for multigraphs as well). Indeed in the following we assume this model when working with regular graphs.

As an example, the *random regular Ising ferromagnet* is obtained by letting $\mathcal{X} = \{+1, -1\}$ and, for some $\epsilon \le 1/2$, $\psi(x_1, x_2) = 1 - \epsilon$ if $x_1 = x_2$ and $\psi(x_1, x_2) = \epsilon$ otherwise.

Specifying a *Poisson ensemble* requires an alphabet $\mathcal{X}$, a *density* $\gamma \in \mathbb{R}_+$, a finite collection of weights $\{\psi_a(\,\cdot\,,\,\cdot\,) : a \in \mathcal{C}\}$, and a probability distribution $\{p(a) : a \in \mathcal{C}\}$ over the weights. In this case G is a multigraph where the number edges among any pair of vertices i and j is an independent Poisson random variable of parameter $2\gamma/n$. Each loop (i, i) is present with multiplicity which is Poisson of mean[2] γ/n. Finally, for each edge in the multi-graph, we draw an independent random variable a with distribution $p(\,\cdot\,)$ and set $\psi_{ij}(\,\cdot\,,\,\cdot\,) = \psi_a(\,\cdot\,,\,\cdot\,)$.

Two examples of Poisson ensembles to be treated below are the *Ising spin glass*, and *antiferromagnetic colorings* (aka 'antiferromagnetic Potts model'). In the first case $\mathcal{X} = \{+1, -1\}$ and two type of weights appear with equal probability (i.e. $\mathcal{C} = \{+, -\}$ and $p(+) = p(-) = 1/2$): $\psi_+(x_1, x_2) = 1 - \epsilon$ for $x_1 = x_2$, $\psi_+(x_1, x_2) = \epsilon$ for $x_1 \neq x_2$, while $\psi_-(x_1, x_2) = \epsilon$ for $x_1 = x_2$, $\psi_-(x_1, x_2) = 1 - \epsilon$ for $x_1 \neq x_2$. For proper colorings $\mathcal{X} = \{1, \dots, q\}$, and $|\mathcal{C}| = 1$ with $\psi(x_1, x_2) = 1$ if $x_1 \neq x_2$, and $\psi(x_1, x_2) = \epsilon < 1$ otherwise (for $\epsilon = 0$ one recovers the uniform measure over proper colorings of G).

Both graphical model ensembles defined above converge locally to trees. In the case of regular models, the corresponding tree model is an infinite rooted tree of uniform degree $(k+1)$, each edge being associated the same weight $\psi(\,\cdot\,,\,\cdot\,)$. For Poisson models, the relevant tree is a rooted Galton-Watson tree with Poisson distributed degrees of mean 2γ. Each edge carries the weight $\psi_a(\,\cdot\,,\,\cdot\,)$ independently with probability $p(a)$.

Given such infinite weighted trees, let T_ℓ, $\ell \ge 0$ be the weighted subgraph obtained by truncating it at depth ℓ. One can introduce random variables $X = \{X_j : j \in \mathsf{T}_\ell\}$, by defining $\mathbb{P}\{X = x|\mathsf{T}_\ell\}$ as in Eq. (1) (with G replaced by T_ℓ). With an abuse of notation we shall call r the root of T_ℓ. It is natural to ask whether reconstruction on the original graphical models and on the corresponding trees are related.

Definition 1.3. *Consider a sequence of random graphical models $\{G_N\}$ (distributed according either to the regular or to the Poisson ensemble), and let $\{\mathsf{T}_\ell\}$ be the corresponding sequence of tree graphical models. We say that the reconstruction problem is* tree-solvable *for the sequence $\{G_N\}$ if there exists $\varepsilon > 0$ such that, for any $t \ge 0$*

$$\|\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|\mathsf{T}_\ell\} - \mathbb{P}_r\{\,\cdot\,|\mathsf{T}_\ell\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|\mathsf{T}_\ell\}\|_{\text{TV}} > \varepsilon\,, \quad (3)$$

with positive probability (as $\ell \to \infty$)

Notice that tree-reconstruction is actually a question on the sequence of tree graphical models $\{\mathsf{T}_\ell\}$ indexed by ℓ. The only role of the original random graphs sequence $\{G_N\}$ is to determine the distribution of T_ℓ.

Despite the similarity of Eqs. (3) and (2), passing from the original graph to the tree is a huge simplification because $\mathbb{P}\{\,\cdot\,|\mathsf{T}_\ell\}$ has a simple description as mentioned above. For instance, in the case of a ferromagnetic Ising model, one can sample the variables X_j on the tree through a 'broadcast' process. First, generate the root value X_r uniformly at random in $\{+1, -1\}$. Then recursively, for each node j, generate the values of its children $\{l\}$ conditional on $X_j = x_j$ by letting $X_l = x_j$ independently with probability $1 - \epsilon$, and $X_l = -x_j$ otherwise. Analogous descriptions exist for the spin-glass and colorings models.

1.4. Main results

Our first result is a sufficient condition for graph-reconstruction to be equivalent to tree reconstruction. In order to phrase it, we need to define the 'two-replicas type.' Consider a graphical model G_N and two two iid assignments of the variables $X^{(1)}$, $X^{(2)}$ with common distribution $\mathbb{P}\{\,\cdot\,|G_N\}$ (we will call them *replicas* following the spin glass terminology). The *two replica type* is a matrix $\{\nu(x, y) : x, y \in \mathcal{X}\}$ where $\nu(x, y)$ counts the fraction of vertices j such that $X_j^{(1)} = x$ and $X_j^{(2)} = y$. (Conversely, the set of distributions ν on $\mathcal{X} \times \mathcal{X}$ such that $N\nu(x, y) \in \mathbb{N}$ will be called the set of *valid two-replicas types* $\mathcal{R}_N$. When we drop the constraint $N\nu(x, y) \in \mathbb{N}$, we shall use $\mathcal{R}$.)

[2] Notice that in a typical realization there will be only a few loops and non-simple edges.

The matrix ν is random. If $\mathbb{P}\{\,\cdot\,|G_N\}$ were the uniform distribution, then ν would concentrate around $\overline{\nu}(x,y) \equiv 1/|\mathcal{X}|^2$. Our sufficient condition requires this to be approximately true.

Theorem 1.4. *Consider a sequence of random Poisson graphical models $\{G_N\}$, and let $\nu(\,\cdot\,,\,\cdot\,)$ be the type of two iid replicas $X^{(1)}$, $X^{(2)}$, and $\Delta\nu(x,y) \equiv \nu(x,y) - \overline{\nu}(x,y)$. Assume that, for any $x \in \mathcal{X}$,*

$$\mathbb{E}\left\{\left[\Delta\nu(x,x) - 2|\mathcal{X}|^{-1}\sum_{x'}\Delta\nu(x,x')\right]^2\right\} \stackrel{N}{\to} 0\,. \quad (4)$$

Then the reconstruction problem for $\{G_N\}$ is solvable if and only if it is tree-solvable.

Remark 1: Notice that the expectation in Eq. (4) is both over the two replicas $X^{(1)}$, $X^{(2)}$ (which the type $\nu(\,\cdot\,,\,\cdot\,)$ is a function of) conditional on G_N, and over G_N. Explicitly $\mathbb{E}\{\cdots\} = \mathbb{E}\{\mathbb{E}[\cdots|G_N]\}$. *Remark 2:* In fact, as is hinted by the proof, the condition (4) can be weakened, e.g. $\overline{\nu}(\,\cdot\,\cdot\,)$ can be chosen more generally than the uniform matrix. This will be treated in a longer publication.

The condition (4) emerges naturally in a variety of contexts, a notable one being second moment method applied to random constraint satisfaction problems [24]. As an example, consider proper colorings of random graphs. In bounding on the colorability threshold, one computes the second moment of the number of colorings, and, as an intermediate step, an upper bound on the large deviations of the type ν. Oversimplifying, one might interpret Theorem 1.4 by saying that, when second moment method works, then tree and graph reconstruction are equivalent. Building on [3] we can thus establish the following.

Theorem 1.5. *Consider antiferromagnetic q-colorings of a Poisson random graph and let $\gamma_q \equiv (q-1)\log(q-1)$. Then there exists a set Γ of zero (Lebesgue) measure such that the following is true. If $\gamma \in [0,\gamma_q) \setminus \Gamma$ and $\epsilon \in (0,1]$, then the reconstruction problem is solvable if and only if it is tree solvable.*

We expect the result to hold down to $\epsilon = 0$ (proper colorings), with $\Gamma = \emptyset$, but did not prove it because of some technical difficulties (indeed we need a sharper control of ν that guaranteed by [3], and our proof technique, cf. Lemma 4.3, relied on an average over γ).

The above theorems might suggests that graph and tree reconstruction do generally coincide. This expectation is falsified by the simplest possible example: the Ising model. This has been studied in depth for trees [18, 20, 21]. If the tree is regular with degree $(k+1)$, the problem is solvable if and only if $k(1-2\epsilon)^2 > 1$. The situation changes dramatically for graphs.

Theorem 1.6. *Reconstruction is solvable for random regular Ising ferromagnets if and only if $k(1-2\epsilon) > 1$.*

This result possibly generalizes to Ising ferromagnets on other graphs that converge locally to trees. The proof of reconstructibility for $k(1-2\epsilon) > 1$ essentially amounts to finding a bottleneck in Glauber dynamics. As a consequence it immediately implies that the mixing time is exponential in this regime. We expect this to be a tight estimate of the threshold for fast mixing.

On the other hand, for an Ising spin-glass, the tree and graph thresholds do coincide. In fact, for an Ising model on a Galton-Watson tree with Poisson(2γ) offspring distribution, reconstruction is solvable if and only if $2\gamma(1-2\epsilon)^2 > 1$ [20]. The corresponding graph result is established below.

Theorem 1.7. *Reconstruction is solvable for Poisson Ising spin-glasses if $2\gamma(1-2\epsilon)^2 > 1$, and it is unsolvable if $2\gamma(1-2\epsilon)^2 < 1$.*

2. Random graph preliminaries

Let us start with a few more notations. Given $i \in V$, and $t \in \mathbb{N}$, $\mathsf{B}(i,t)$ is the set of vertices j such that $d(i,j) \le t$ (as well as the subgraph formed by those vertices and by edges that are not in $\overline{\mathsf{B}}(i,t)$). Further we introduce the set of vertices $\mathsf{D}(i,t) \equiv \mathsf{B}(i,t) \cap \overline{\mathsf{B}}(i,t)$.

The proof of Theorem 1.4 relies on two remarkable properties of Poisson graphical models: the local convergence of $\mathsf{B}(r,t)$ to the corresponding Galton-Watson tree of depth t (whose straightforward proof we omit), and a form of independence of $\overline{\mathsf{B}}(r,t)$ relatively to $\mathsf{B}(r,t)$. Notice that, because of the symmetry of the graph distribution under permutation of the vertices, we can fix r to be a deterministic vertex (say, $r = 1$).

Proposition 2.1. *Let $\mathsf{B}(r,t)$ be the depth-t neighborhood of the root in a Poisson random graph G_N, and T_t a Galton-Watson tree of depth t and offspring distribution* Poisson(2γ). *Given any (labeled) tree T_*, we write $\mathsf{B}(r,t) \simeq \mathsf{T}_*$ if T_* is obtained by the depth-first relabeling of $\mathsf{B}(r,t)$ following a pre-established convention*[3]. *Then $\mathbb{P}\{\mathsf{B}(r,t) \simeq \mathsf{T}_*\}$ converges to $\mathbb{P}\{\mathsf{T}_t \simeq \mathsf{T}_*\}$ as $N \to \infty$.*

Proposition 2.2. *Let $\mathsf{B}(r,t)$ be the depth-t neighborhood of the root in a Poisson random graph G_N. Then, for any $\lambda > 0$ there exists $C(\lambda,t)$ such that, for any N, $M \ge 0$*

$$\mathbb{P}\{|\mathsf{B}(r,t)| \ge M\} \le C(\lambda,t)\,\lambda^{-M}\,. \quad (5)$$

Proof. Imagine to explore $\mathsf{B}(r,t)$ in breadth-first fashion. For each t, $|\mathsf{B}(r,t+1)| - |\mathsf{B}(r,t)|$ is upper bounded by the sum of $|\mathsf{D}(r,t)|$ iid binomial random variables with parameters $N - |\mathsf{B}(r,t)|$ and $1 - e^{-2\gamma/N} \le 2\gamma/N$ (the number

[3] For instance one might agree to preserve the original lexicographic order among siblings

of neighbors of each node in $\mathsf{D}(r,t)$). Therefore $|\mathsf{B}(r,t)|$ is stochastically dominated by $\sum_{s=0}^{t} Z_N(s)$, where $\{Z_N(t)\}$ is a Galton-Watson process with offspring distribution $\text{Binom}(N, 2\gamma/N)$. By Markov inequality $\mathbb{P}\{|\mathsf{B}(r,t)| \geq M\} \leq \mathbb{E}\{\lambda^{\sum_{s=0}^{t} Z_N(s)}\}\lambda^{-M}$. By elementary branching processes theory $g_t^N(\lambda) \equiv \mathbb{E}\{\lambda^{\sum_{s=0}^{t} Z_N(s)}\}$ satisfies the recursion $g_{t+1}^N(\lambda) = \lambda\xi_N(g_t^N(\lambda))$, $g_0^N(\lambda) = \lambda$, with $\xi_N(\lambda) = \lambda(1 + 2\gamma(\lambda-1)/N)^N$. The thesis follows by $g_t^N(\lambda) \leq g_t(\lambda)$, the latter being obtained by replacing $\xi_N(\lambda)$ with $\xi(\lambda) = e^{2\gamma(\lambda-1)} \geq \xi_N(\lambda)$. □

Proposition 2.3. *Let G_N be a Poisson random graph on vertex set $[N]$ and edge probability $p = 2\gamma/N$. Then, conditional on $\mathsf{B}(r,t)$, $\overline{\mathsf{B}}(r,t)$ is a Poisson random graph on vertex set $[N] \setminus \mathsf{B}(r,t-1)$ and same edge probability.*

Proof. Condition on $\mathsf{B}(r,t) = \mathsf{G}(t)$, and let $\mathsf{G}(t-1) = \mathsf{B}(r,t-1)$ (notice that this is uniquely determined from $\mathsf{G}(t)$). This is equivalent to conditioning on a given edge realization for any two vertices k, l such that $k \in \mathsf{G}(t-1)$ and $l \in \mathsf{G}(t)$ (or viceversa).

On the other hand, $\overline{\mathsf{B}}(r,t)$ is the graph with vertices set $[N]\setminus\mathsf{G}(t)$ and edge set $(k,l) \in G_N$ such that $k, l \notin \mathsf{G}(t-1)$. Since this set of vertices couples is disjoint from the one we are conditioning upon, and by independence of edges in G_N, the claim follows. □

3. Proof of Theorem 1.4

We start from a simple technical result.

Lemma 3.1. *Let p, q be probability distribution over a finite set $\mathcal{S}$, and denote by $q_0(x) = 1/|\mathcal{S}|$ the uniform distribution over the same set. Define $\widehat{p}(x) \equiv p(x)q(x)/z$, where $z \equiv \sum_x p(x)q(x)$. Then $||\widehat{p} - p||_{\text{TV}} \leq 3|\mathcal{S}|^2\,||q - q_0||_{\text{TV}}$.*

Proof. Since $||\widehat{p} - p||_{\text{TV}} \leq 1$ we can assume, without loss of generality, that $||q - q_0||_{\text{TV}} \leq (2|\mathcal{S}|)^{-1}$. If we write $p(x) = p(x)q_0(x)/z_0$, with $z_0 = 1/|\mathcal{S}|$, then $|z - z_0| \leq |\sum_x [p(x)q(x) - p(x)q_0(x)]| \leq ||q - q_0||_{\text{TV}}$ and in particular $z \geq z_0/2$. From triangular inequality we have on the other hand

$$||\widehat{p} - p||_{\text{TV}} \leq \frac{1}{2}\left|z^{-1} - z_0^{-1}\right| + \frac{1}{2z_0}\sum_x p(x)|q(x) - q_0(x)|\,.$$

Using $|x^{-1} - y^{-1}| \leq |x-y|/\min(x,y)^2$, the first term is bounded by $2|z - z_0|/z_0^2 \leq 2|\mathcal{S}|^2||q - q_0||_{\text{TV}}$. The second is at most $|\mathcal{S}|\,||q - q_0||_{\text{TV}}$ which proves the thesis. □

In order to prove Theorem 1.4 we first establish that, under the condition (4), any (fixed) subset of the variables $\{X_1, \dots, X_N\}$ is (approximately) uniformly distributed.

Proposition 3.2. *Let $i(1), \dots, i(k) \subseteq [N]$ be any (fixed) set of vertices, and $\xi_1, \dots, \xi_k \in \mathcal{X}$. Then, under the hypotheses of Theorem 1.4, for any $\varepsilon > 0$*

$$\left|\mathbb{P}_{i(1),\dots,i(k)}\{\xi_1, \dots, \xi_k | G_N\} - \frac{1}{|\mathcal{X}|^k}\right| \leq \varepsilon\,, \tag{6}$$

with high probability.

Proof. Given two replicas $X^{(1)}$, $X^{(2)}$, define, for $\xi \in \mathcal{X}$ (with $\mathbb{I}_{\dots}$ the indicator function)

$$\mathsf{Q}(\xi) = \frac{1}{N}\sum_{i=1}^{N}\left\{\mathbb{I}_{X_i^{(1)}=\xi} - \frac{1}{|\mathcal{X}|}\right\}\left\{\mathbb{I}_{X_i^{(2)}=\xi} - \frac{1}{|\mathcal{X}|}\right\}.$$

Notice that $\mathsf{Q}(\xi) = \Delta\nu(\xi,\xi) - (2/|\mathcal{X}|)\sum_x \Delta\nu(\xi,x)$ is the quantity in Eq. (4). Therefore, under the hypothesis of Theorem 1.4, $\mathbb{E}\{\mathsf{Q}(\xi)^2\} \stackrel{N}{\to} 0$. Further, since $|Q(\xi)| \leq 1$ and using Cauchy-Schwarz, for any $\xi_1, \dots, \xi_k \in \mathcal{X}$

$$|\mathbb{E}\left\{\mathsf{Q}(\xi_1)\cdots\mathsf{Q}(\xi_k)\right\}| \leq \mathbb{E}|\mathsf{Q}(\xi_1)| \stackrel{N}{\to} 0\,.$$

If we denote by $\mathsf{Q}_i(\xi)$ the quantity on the right hand side of the sum in Eq. (7) then $\mathsf{Q}(\xi)$ is the uniform average of $\mathsf{Q}_i(\xi)$ over a uniformly random $i \in [N]$. By symmetry of the graph distribution with respect to permutation of the vertices in $[N]$, and since $|\mathsf{Q}(\xi)| \leq 1$ we get

$$\begin{aligned}\mathbb{E}\left\{\mathsf{Q}(\xi_1)\cdots\mathsf{Q}(\xi_k)\right\} &= \mathbb{E}\left\{\mathsf{Q}_{i(1)}(\xi_1)\cdots\mathsf{Q}_{i(k)}(\xi_k)\right\} + \varepsilon_{k,N}\\ &= \mathbb{E}\{\mathbb{E}\{\prod_{a=1}^{k}(\mathbb{I}_{X_{i(a)}=\xi_a} - |\mathcal{X}|^{-1})|G_N\}^2\} + \varepsilon_{k,N}\,,\end{aligned}$$

where $|\varepsilon_{k,N}|$ is upper bounded by the probability that k random variable uniform in $[N]$ are not distinct (which is $O(1/N)$). Therefore the expectation on right hand side vanishes as $N \to \infty$ as well, which implies (since the quantity below is, again, bounded by 1)

$$\left|\mathbb{E}\left\{\prod_{a=1}^{k}(\mathbb{I}_{X_{i(a)}=\xi_a} - |\mathcal{X}|^{-1})\middle| G_N\right\}\right| \leq \varepsilon \tag{7}$$

with high probability for any $\varepsilon > 0$. The proof is completed by noting that the left hand side of Eq. (6) can be written as

$$\left|\sum_{\emptyset\neq U\subseteq[k]}\mathbb{E}\left\{\prod_{a\in U}(\mathbb{I}_{X_{i(a)}=\xi_{i(a)}} - |\mathcal{X}|^{-1})\middle| G_N\right\}\right| \leq 2^k\varepsilon\,,$$

where the last bound holds whp thanks to Eq. (7) and ε can eventually be rescaled. □

In order to write the proof Theorem 1.4 we need to introduce a few shorthands. Given a graphical model G_N, and $U \subseteq [N]$, we let $\mu_U(x_U) \equiv \mathbb{P}\{X_U = x_U | G_N\}$ (omitting subscripts if $U = V$). If r is its root, $\ell \in \mathbb{N}$ and

$U \subseteq \mathsf{B}(r,\ell)$, we define $\mu^{<}_U(x_U) \equiv \mathbb{P}\{X_U = x_U | \mathsf{B}(r,\ell)\}$ (i.e. $\mu^{<}$ is the distribution obtained by restricting the product in (1) to edges $(i,j) \in \mathsf{B}(r,\ell)$). Analogously $\mu^{>}_U(x_U) \equiv \mathbb{P}\left\{X_U = x_U | \overline{\mathsf{B}}(r,\ell)\right\}$. Finally for $U \subseteq [N]$, we let $\rho_U(x_U) = 1/|\mathcal{X}|^{|U|}$ be the uniform distribution on $\mathcal{X}^U$.

Lemma 3.3. *Let G_N be a graphical model rooted at r, and $\ell \in \mathbb{N}$. Then for any $t \le \ell$,*

$$\left| ||\mu_{r,\overline{\mathsf{B}}(r,t)} - \mu_r \mu_{\overline{\mathsf{B}}(r,t)}||_{\text{TV}} - ||\mu^{<}_{r,\overline{\mathsf{B}}(r,t)} - \mu^{<}_r \mu^{\le}_{\overline{\mathsf{B}}(r,t)}||_{\text{TV}} \right| \le$$
$$\le 9|\mathcal{X}|^{2|\mathsf{B}(r,\ell)|}\, ||\mu^{>}_{\mathsf{D}(r,\ell)} - \rho_{\mathsf{D}(r,\ell)}||_{\text{TV}}\,. \qquad (8)$$

Proof. First notice that, by elementary properties of the total variation distance, $||\mu_U - \mu^{<}_U||_{\text{TV}} \le ||\mu_{\overline{\mathsf{B}}(r,\ell)} - \mu^{<}_{\mathsf{B}(r,\ell)}||_{\text{TV}}$ for any $U \subseteq \mathsf{B}(r,\ell)$. Applying this remark and triangular inequality, the left hand side of Eq. (8) can be upper bounded by $3\,||\mu_{\overline{\mathsf{B}}(r,\ell)} - \mu^{<}_{\mathsf{B}(r,\ell)}||_{\text{TV}}$. Next notice that, as a consequence of Eq. (1) and of the fact that $\mathsf{B}(r,\ell)$ and $\overline{\mathsf{B}}(r,\ell)$ are edge disjoint (and using the shorthands $\mathsf{B}(\ell)$ and $\mathsf{D}(\ell)$ for $\mathsf{B}(r,\ell)$ and $\mathsf{D}(r,\ell)$)

$$\mu_{\mathsf{B}(\ell)}(x_{\mathsf{B}(\ell)}) = \frac{\mu^{<}_{\mathsf{B}(\ell)}(x_{\mathsf{B}(\ell)})\mu^{>}_{\mathsf{D}(\ell)}(x_{\mathsf{D}(\ell)})}{\sum_{x'_{\mathsf{B}(\ell)}} \mu^{<}_{\mathsf{B}(\ell)}(x'_{\mathsf{B}(\ell)})\mu^{>}_{\mathsf{D}(\ell)}(x'_{\mathsf{D}(\ell)})}\,.$$

We can therefore apply Lemma 3.1 whereby p is $\mu^{<}_{\mathsf{B}(\ell)}$, $\widehat{p}$ is $\mu_{\mathsf{B}(\ell)}$, q is $\mu^{>}_{\mathsf{D}(\ell)}$, and $\mathcal{S} = \mathcal{X}^{\mathsf{B}(\ell)}$. This proves the thesis. □

Proof of Theorem 1.4. Let Δ_N denote the left hand side of Eq. (8). We claim that its expectation (with respect to a random graph G_N) vanishes as $N \to \infty$. Since the probability that $\mathsf{B}(r,\ell) \ge M$ can be made arbitrarily small by letting M large enough, cf. Lemma 2.2, and using the fact that the left hand side of Eq. (8) is bounded by 1, it is sufficient to prove that

$$\sum_{|\mathsf{G}|\le M} \mathbb{P}\{\mathsf{B}(r,\ell) = \mathsf{G}\}\mathbb{E}\{\Delta_N | \mathsf{B}(r,\ell) = \mathsf{G}\} \le$$
$$\le K^{M+1} \sum_{|\mathsf{G}|\le M} \mathbb{E}\{||\mu^{>}_{\mathsf{D}(r,\ell)} - \rho_{\mathsf{D}(r,\ell)}||_{\text{TV}} | \mathsf{B}(r,\ell) = \mathsf{G}\}\,,$$

vanishes as $N \to \infty$. Each term in the sum is the expectation, with respect to a random graph over $N - |\mathsf{G}| \ge N - M$ vertices of the total variation distance between the joint distribution of a fixed set of vertices, and the uniform distribution (for $\mathsf{D} = \mathsf{D}(r,\ell)$):

$$||\mu^{>}_{\mathsf{D}} - \rho_{\mathsf{D}}||_{\text{TV}} = \frac{1}{2}\sum_{x_{\mathsf{D}}} \left| \mathbb{P}_{\mathsf{D}}\{x_{\mathsf{D}} | \overline{\mathsf{B}}(i,\ell)\} - |\mathcal{X}|^{-|\mathsf{D}|} \right|\,.$$

This vanishes by Lemma 3.2, thus proving the above claim.

This implies that there exists $\varepsilon > 0$ such that $||\mu_{r,\overline{\mathsf{B}}(r,t)} - \mu_r \mu_{\overline{\mathsf{B}}(r,t)}||_{\text{TV}} \ge \varepsilon$ with positive probability, if and only if there exists $\varepsilon' > 0$ such that $||\mu^{<}_{r,\overline{\mathsf{B}}(r,t)} - \mu^{<}_r \mu^{\le}_{\overline{\mathsf{B}}(r,t)}||_{\text{TV}} \ge \varepsilon'$ with positive probability. In other words, since $\mu(\,\cdot\,) \equiv \mathbb{P}\{\cdots | G_N\}$, reconstruction is solvable if and only if $||\mu^{<}_{r,\overline{\mathsf{B}}(r,t)} - \mu^{<}_r \mu^{\le}_{\overline{\mathsf{B}}(r,t)}||_{\text{TV}} \ge \varepsilon'$ with positive probability.

Finally, recall that $\mu^{<}(\,\cdot\,) \equiv \mathbb{P}\{\,\cdot\,|\mathsf{B}(r,\ell)\}$ and that $\mathsf{B}(r,\ell)$ converges in distribution to $\mathsf{T}(\ell)$, by Lemma 2.1. Since $||\mu^{<}_{r,\overline{\mathsf{B}}(i,t)} - \mu^{<}_r \mu^{\le}_{\overline{\mathsf{B}}(r,t)}||_{\text{TV}}$ is a bounded function of $\overline{\mathsf{B}}(r,t)$ (and using as above Lemma 2.2 to reduce to a finite set of graphs), it converges in distribution to $\|\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|\mathsf{T}_\ell\} - \mathbb{P}_r\{\,\cdot\,|\mathsf{T}_\ell\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|\mathsf{T}_\ell\}\|_{\text{TV}}$. We conclude that $||\mu^{<}_{r,\overline{\mathsf{B}}(i,t)} - \mu^{<}_r \mu^{\le}_{\overline{\mathsf{B}}(r,t)}||_{\text{TV}} \ge \varepsilon'$ with positive probability if and only if reconstruction is tree solvable, thus proving the thesis. □

4. Two successful applications

4.1. The Ising spin glass

Proof of Theorem 1.7. The behavior of spin-glasses on Poisson random graphs has been studied extensively in [2]. In particular, the *two-replica overlap* $q_{12} = N^{-1}\sum_i X^{(1)}_i X^{(2)}_i$ satisfies $\mathbb{E}\{q^2_{12}\} = O\left(N^{-1}\right)$ for $2\gamma(1-2\epsilon)^2 < 1$ ("high-temperature" region). It is easy to check that the quantity in the expectation in Eq. (4) equals $(q_{12}/4)^2$ both for $x = +1$, and -1. Hence, if $2\gamma(1-2\epsilon)^2 < 1$, Theorem 1.4 applies. Since tree reconstruction is unsolvable in that case [20] (notice that on trees, reconstruction for the spin glass model and the ferromagnet are equivalent), we obtain that graph reconstruction is unsolvable as well.

Conversely, suppose that $2\gamma(1-2\epsilon)^2 > 1$. First assume $\mathbb{E}\{q^2_{12}\} \stackrel{N}{\to} 0$. Since tree reconstruction is solvable in this case, Theorem 1.4 would imply that graph reconstruction is solvable as well. It is thus sufficient to prove that graph reconstruction is solvable if $\mathbb{E}\{q^2_{12}\} \not\to 0$. Equivalently, if for any $\varepsilon > 0$, there exists $t = t(\varepsilon)$ such that $||\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|G_N\}||_{\text{TV}} \le \varepsilon$ with high probability, then $\mathbb{E}\{q^2_{12}\} \stackrel{N}{\to} 0$.

Since q_{12} is the average of $X^{(1)}_i X^{(2)}_i$ over a uniformly random $i \in [N]$, then $\mathbb{E}\{q^2_{12}\}$ is the average of $\mathbb{E}\{X^{(1)}_r X^{(2)}_r X^{(1)}_j X^{(2)}_j\} = \mathbb{E}\{\mathbb{E}\{X_r X_j | G_N\}^2\}$ over r, $j \in [N]$ uniform and independent. Fixing $t = t(\varepsilon)$ as above, we can neglect the probability that $j \in \mathsf{B}(r,t)$, since this is N^{-1} times the expected size of $\mathsf{B}(r,t)$, that is bounded by Lemma 2.2. Therefore $\mathbb{E}\{q^2_{12}\} = \mathbb{E}\left\{\mathbb{E}[X_r X_j | G_N]^2\, \mathbb{I}_{j\notin\mathsf{B}(r,t)}\right] + o(1)$. On the other hand, if $j \notin \mathsf{B}(r,t)$

$$\left|\mathbb{E}[X_r X_j | G_N]\right| \le$$
$$\le ||\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|G_N\}||_{\text{TV}}\,.$$

We deduce that $\left|\mathbb{E}[X_r X_j|G_N]\right| \le \varepsilon$ with high probability ant hence $\lim_N \mathbb{E}\{q_{12}\} \le \varepsilon^2$ The thesis follows by recalling that ε is an arbitrary positive number. □

4.2. q-colorings of Poisson random graphs

The application Theorem 1.4 to this case require some technical groundwork. For space reasons we limit to quoting the results deferring the proofs to a complete publication. We denote by $U(x)$ be the number of monochromatic edges under coloring x, by $Z = \sum_x \epsilon^{U(x)}$ the partition function, by $Z_{\rm b}$ the modified partition function where the sum is restricted to balanced colorings (such that each color occupies N/q vertices), and by $Z_2(\nu) = \sum_{x^{(1)},x^{(2)}}^{\nu} \epsilon^{U(x^{(1)})+U(x^{(2)})}$, where the sum is restricted to couples of colorings with two-replica type $\nu = \{\nu(x,y)\}_{x,y\in[q]}$. As above we denote by $\overline{\nu}(x,y) = 1/q^2$ for any $x, y \in [q]$ the uniform matrix. Finally we introduce the following function over two-replica types ν (i.e. over $q\times q$ matrices with non-negative entries normalized to one):

$$\phi(\nu) = -\sum_{xy}\nu(x,y)\log\nu(x,y) + \\ + \bar\gamma \log\Big\{1 - \bar\epsilon F(\nu) + \bar\epsilon^2 \sum_{x,y}\nu(x,y)^2\Big\}\,,$$

where $\bar\epsilon = 1-\epsilon$ and $F(\nu) \equiv 2\sum_x(\sum_y \nu(x,y))^2$.

The first two preliminary remarks are a combinatorial calculation that straightforwardly generalizes the result of [3] for proper colorings, and a good estimate on the balanced partition function.

Lemma 4.1. *Let M be the number of edges in a Poisson graph. Then $\mathbb{E}[Z_2(\nu)|M = \bar\gamma N] \le K\, e^{n\phi(\nu)}$.*

Lemma 4.2. *Let $\gamma < \gamma_q = (q-1)\log(q-1)$. Then, for any $\xi > 0$, $Z_{\rm b} \ge e^{N[\phi(\overline{\nu})-\xi]/2}$ with high probability. Further, $\mathbb{E}\log Z_{\rm b} \ge N[\log q + \gamma\log(1-\bar\epsilon/q)] + O(N^{1/2})$.*

Our last remark is that, for $\gamma < \gamma_q$, balanced colorings dominate the measure μ.

Lemma 4.3. *Let Γ and γ_q be as in the statement of Theorem 1.5, and $\nu(x,y)$ be as in 1.4. Then, for any $\gamma \in [0,\gamma_q)\setminus\Gamma$, any $x \in \mathcal{X}$, and any $\delta > 0$, $|\sum_y \nu(x,y) - q^{-1}| \le \delta$ with high probability.*

Proof. Recall that, if X is a Poisson random variable of mean λ, and $f(\lambda) \equiv \mathbb{E}\,F(X)$, then $f'(\lambda) = \mathbb{E}[F(X+1) - F(X)]$. Further notice that, if $Z(G)$ is the partition function for the graph G, then

$$Z(G\cup(ij)) = Z(G)\left[1 - \bar\epsilon\mathbb{P}\{X_i = X_j|G\}\right]. \qquad (9)$$

Applying these identities to $\alpha(\gamma) \equiv N^{-1}\mathbb{E}\log Z(G_N)$, we get

$$\frac{{\rm d}\alpha(\gamma)}{{\rm d}\gamma} = \frac{1}{N^2}\sum_{i,j}\mathbb{E}\log\{1 - \bar\epsilon\mathbb{P}\{X_i = X_j|G_N\}\}\,. \quad (10)$$

Since $\alpha(0) = \log q$, by using Jensen inequality we get

$$\alpha(\gamma) \le \log q + \int_0^\gamma \log\{1 - \bar\epsilon\,\mathbb{E}\,[A(\gamma')]\}{\rm d}\gamma'\,, \qquad (11)$$

where (the dependence on γ being through the distribution of G_N and hence of X)

$$A(\gamma) \equiv \frac{1}{N^2}\sum_{ij}\mathbb{I}\{X_i = X_j\} = \sum_x\Big(\sum_y \nu(x,y)\Big)^2.$$

On the other hand, for $\gamma < \gamma_q$

$$\alpha(\gamma) \ge \log q + \gamma\log(1-\bar\epsilon/q) + O(N^{-1/2})\,. \qquad (12)$$

This follows from $Z \ge Z_{\rm b}$ together with Lemma 4.2.

From Eq. (11) and (12), and since $A(\gamma) \ge 1/q$ by definition, we get $A(\gamma) \le (1+\varepsilon)/q$ with high probability for any $\gamma \in [0,\gamma_q)\setminus\Gamma$, where Γ has zero Lebesgue measure. Finally by Cauchy-Schwarz

$$\sum_x\Big|\sum_y\nu(x,y) - \frac{1}{q}\Big| \le q\sum_x\Big[\sum_y\nu(x,y) - \frac{1}{q}\Big]^2 \le \\ \le q\,A(\gamma) - 1 \le \varepsilon\,,$$

where the last inequality holds with high probability by the above. □

Lemma 4.4. *Let $\mathcal{A} \subset \mathcal{R}$ be any subset of the set of two replicas types, and assume $\sup_{\nu\in\mathcal{A}}\phi(\nu) < \phi(\overline{\nu})$. Then $\nu \notin \mathcal{A}$ with high probability.*

Proof. Fix $\xi > 0$ and denote by Typ the set of graphs such that $Z \ge Z_{\rm b} > e^{N[\phi(\overline{\nu})-\xi]/2}$, and that the number of edges M satisfies $|M - N\gamma| \le o(N)$. For any $G_N \in$ Typ, we have (denoting, with an abuse of notation, the two replica type of $x^{(1)}$ and $x^{(2)}$ by ν as well)

$$\mathbb{P}\{\nu\in\mathcal{A}|G_N\} = \frac{1}{Z^2}\sum_{x^{(1)},x^{(2)}}\epsilon^{U(x^{(1)})+U(x^{(2)})}\,\mathbb{I}(\nu\in\mathcal{A}) \\ \le e^{-N[\phi(\overline{\nu})-\xi]}\sum_{x^{(1)},x^{(2)}}\epsilon^{U(x^{(1)})+U(x^{(2)})}\,\mathbb{I}(\nu\in\mathcal{A})\,.$$

On the other hand, it follows from Lemma 4.2 that $\mathbb{P}\{\nu\in\mathcal{A}\} \le \mathbb{E}\{\mathbb{P}\{\nu\in\mathcal{A}|G_N\}\,\mathbb{I}_{G_N\in\mathsf{Typ}}\} + o_N(1)$. Using Lemma 4.1, this implies

$$\mathbb{P}\{\nu\in\mathcal{A}\} \le Ke^{N\xi}\sum_{\nu\in\mathcal{R}_N\cap\mathcal{A}}e^{-N[\phi(\overline{\nu})-\phi(\nu)+o_N(1)]} + o_N(1)\,,$$

where the $o(1)$ term in the exponent accounts for the fact that $M = N[\gamma + o(1)]$ (as $\phi(\nu)$ is continuous in $\overline{\gamma}$). The thesis follows by choosing $\xi = \inf_{\nu\in\mathcal{A}}[\phi(\overline{\nu}) - \phi(\nu)]/2 > 0$ and noting that the number of terms in the sum is at most $|\mathcal{R}_N| = O(N^q)$. □

Proof of Theorem 1.5. The quantity appearing in the expectation in Eq. (4) is upper bounded by $3\max_{x,y}|\nu(x,y) - \overline{\nu}(x,y)|$. We will prove that, for $\gamma \in [0,\gamma_q)\setminus\Gamma$, for any $\delta > 0$, and any $x, y \in \mathcal{X}$, $|\nu(x,y) - \overline{\nu}(x,y)| \le \delta$ with high probability, which implies the sufficient condition in Theorem 1.4.

Notice that $F(\nu) \ge 2/q$ with $F(\nu) = 2/q$ if and only if $\sum_y \nu(x,y) = 1/q$. Because of Lemma 4.3, $F(\nu) \le q^{-1} + \delta'$ with high probability for any $\delta' > 0$ (to be fixed below). The thesis follows by applying Lemma 4.4 to the event $\mathcal{A} = \{|\nu(x,y) - \overline{\nu}(x,y)| > \delta;\, F(\nu) \le q^{-1} + \delta'\}$, thus showing that $\nu \notin \mathcal{A}$ with high probability and hence $|\nu(x,y) - \overline{\nu}(x,y)| \le \delta$.

We are left with the task of checking the hypothesis of Lemma 4.4, namely $\sup_{\nu\in\mathcal{A}}\phi(\nu) < \phi(\overline{\nu})$. Achlioptas and Naor proved that, if $F(\nu) = q^{-1}$ (i.e. the column and row sums of ν are all equal), and $\gamma < \gamma_q$, then $\phi(\nu) \le \phi(\overline{\nu}) - A'||\nu - \overline{\nu}||_{\text{TV}}^2$ for some $A' > 0$ (see [3], Theorem 7 and discussion below). Under the condition $||\nu - \overline{\nu}||_{\text{TV}} \ge |\nu(x,y) - \overline{\nu}(x,y)| \ge \delta$ $\phi(\nu) < \phi(\overline{\nu})$, always subject to $F(\nu) = q^{-1}$. But by continuity of $F(\nu)$ and $\psi(\nu)$, we can chose δ' small enough such that $\phi(\nu) < \phi(\overline{\nu})$ for $F(\nu) \le q^{-1} + \delta'$ as well. □

5. The case of the Ising ferromagnet

We now set out to demonstrate a counter-example to the graph-tree reconstruction equivalence encountered above: the reconstruction threshold for the random $(k+1)$-regular Ising ferromagnet is $k(1-2\epsilon) = 1$ (Theorem 1.6). It is convenient to use a symmetric notation by letting $\theta \equiv 1 - 2\epsilon > 0$, and to generalize the model introducing a second parameter $\lambda \ge 0$ (corresponding to a 'magnetic field' in the physics terminology). We then let $\psi(+,+) = (1+\lambda)(1+\theta)$, $\psi(-,-) = (1-\lambda)(1+\theta)$, and $\psi(+,-) = \psi(-,+) = (1-\theta)$. The original problem is recovered by letting $\lambda = 0$. In terms of these parameters the distribution of X reads

$$\mathbb{P}\{X = x|G_N\} = \frac{1}{Z_{\theta,\lambda}}\theta_+^{e_=(x)}\theta_-^{e_\neq(x)}\lambda_+^{n_+(x)}\lambda_-^{n_-(x)}\,, \quad (13)$$

whereby $\theta_\pm \equiv 1 \pm \theta$, $\lambda_\pm \equiv 1 \pm \lambda$, $n_\pm(x)$ is the number of vertices with $x_i = \pm$, and $e_=(x)$ (respectively $e_\neq(x)$) denotes the number of edges (ij) with $x_i = x_j$ (respectively $x_i \neq x_j$).

A crucial role is played by of the partition function $Z_{\theta,\lambda}$ (defined by the normalization condition of $\mathbb{P}\{\,\cdot\,|G_N\}$) as well as the constrained partition functions

$$\widehat{Z}_{\theta,M} = \sum_{n_+(x)-n_-(x)=NM}\theta_+^{e_=(x)}\theta_-^{e_\neq(x)}\,. \quad (14)$$

The rationale for introducing $\widehat{Z}_{\theta,M}$ is that it allows to estimate the distribution of the number of $+$'s (or $-$'s) through the identity (valid for $\lambda = 0$) $\mathbb{P}\{n_+(X) - n_-(X) = NM|G_N\} = \widehat{Z}_{\theta,M}/Z_{\theta,0}$.

The first technical tool is a well known tree calculation.

Lemma 5.1. *Assume* T *to be a regular tree with branching* k *and depth* t*, rooted at* r*, let* L *be its leaves, and let* $\mathbb{P}\{X = x|\mathsf{T}\}$ *be defined as in Eq. (13) whereby* $n_\pm(x)$ *does not count variables in* x_L*. For* $h_0 \in [-1,+1]$*, let* $F_{h_0}(x_L)$ *be the law of* $|L|$ *iid Bernoulli variables of parameter* $(1+h_0)/2$*, and define* $\mathbb{P}_{h_0}\{X = x|\mathsf{T}\} \equiv \mathbb{P}\{X = x|\mathsf{T}\}\,F_{h_0}(x_L)/C$ *(with* C *a normalization constant).*

Then $\mathbb{P}_{h_0}\{X_i = \pm|\mathsf{T}\} = (1 \pm h_t)/2$*, where* $h_t \equiv f_{\theta,\lambda}^{\circ t}(h_0)$ *(*$f_{\theta,\lambda}^{\circ t}$ *being the* t*-fold composition of* $f_{\theta,\lambda}$*) and*

$$f_{\theta,\lambda}(h) \equiv \frac{(1+\lambda)(1+\theta h)^k - (1-\lambda)(1-\theta h)^k}{(1+\lambda)(1+\theta h)^k + (1-\lambda)(1-\theta h)^k}\,. \quad (15)$$

Lemma 5.2. *For any* $\lambda, \varepsilon > 0$*,* $|f_{\theta,\lambda}^{\circ t}(+1) - f_{\theta,\lambda}^{\circ t}(0)| \le \varepsilon$ *for* t *large enough. Further, for* $k\theta \le 1$ *and any* $h \in [-1,+1]$ $|f_{\theta,0}^{\circ t}(h)| \le \varepsilon$ *for* t *large enough.*

The following estimate of $\widehat{Z}_{\theta,M}$ is a standard exercise in combinatorics, whose proof we omit.

Lemma 5.3. *There exist* $C, D > 0$ *independent of* N *such that* $\mathbb{E}\,\widehat{Z}_{\theta,M} \le C\,2^N\,e^{DN\,M^2}$.

Lemma 5.4. *For any* $\lambda \ge 0$*, let* h_* *the unique non-negative solution of* $h_* = f_{\theta,\lambda}(h_*)$ *and define* $e_{\theta,\lambda}(h) = (\theta + h^2)/(1+\theta h^2)$*. Then, for any* $\varepsilon > 0$*, and a uniformly random edge* $(i,j) \in E$*,* $|\mathbb{E}\{X_iX_j|G_N\} - e_{\theta,\lambda}(h_*)| \le \varepsilon$ *whp.*

Proof. Let $h_{t,+} = f_{\theta,\lambda}^{\circ t}(+1)$ and $h_{t,0} = f_{\theta,\lambda}^{\circ t}(0)$. Since $e_{\theta,\lambda}(h)$ is continuous in h, and because of Lemma 5.2, we can fix $t = t(\varepsilon)$ in such a way that $|e_{\theta,\lambda}(h_{t,+}) - e_{\theta,\lambda}(h_{t,0})| \le \varepsilon$. We will show that, whp, $e_{\theta,\lambda}(h_{t,0}) \le \mathbb{E}\{X_iX_j|G_N\} \le e_{\theta,\lambda}(h_{t,+})$, thus proving the thesis, since (by monotonicity of $f_{\theta,\lambda}(\,\cdot\,)$ and $e_{\theta,\lambda}(\,\cdot\,)$) $e_{\theta,\lambda}(h_{t,0}) \le e_{\theta,\lambda}(h_*) \le e_{\theta,\lambda}(h_{t,+})$ as well.

In order to prove our claim, notice that $\mathsf{B} = \mathsf{B}(i,t) \cup \mathsf{B}(j,t)$ is whp a tree (obtained by joining through their roots two regular trees with branching k and depth t), and denote by D its leaves. Griffiths inequalities imply [17] that $\mathbb{E}\{X_iX_j|G_N\}$ can be lower bounded by replacing G_N with any subgraph, and upper bounded conditioning on $X_k = +1$ for any set of vertices k. In particular we have

$$\mathbb{E}\{X_iX_j|\mathsf{B}\} \le \mathbb{E}\{X_iX_j|G_N\} \le \mathbb{E}\{X_iX_j|X_{\mathsf{D}} = \underline{+1}, \mathsf{B}\}\,,$$

where (in the upper bound) we emphasized that, by the Markov property of $\mathbb{P}\{\,\cdot\,|G_N\}$, X_{B} is conditionally independent of $G_N \setminus \mathsf{B}$, given X_{D}.

The proof is finished by evaluating the upper and lower bound under the assumption, mentioned above, that B is a tree. This can be done through a dynamic programming-type calculation, which we omit from this abstract. The final result is $\mathbb{E}\{X_iX_j|\mathsf{B}\} = e_{\theta,\lambda}(f^{\circ t}_{\theta,\lambda}(\lambda)) \geq e_{\theta,\lambda}(h_{t,0})$ and $\mathbb{E}\{X_iX_j|X_{\mathsf{D}} = \underline{+1}, \mathsf{B}\} = e_{\theta,\lambda}(h_{t,+})$, which finishes the proof. □

Lemma 5.5. *For any $\lambda \geq 0$, let h_* be the unique non-negative solution of $h_* = f_{k,\lambda}(h_*)$ and $\varphi(\theta,\lambda) \equiv \phi(\theta,\lambda,h_*)$, where*

$$\phi(\theta,\lambda,h) \equiv -\frac{k+1}{2}\log(1+\theta h^2)+ \\ +\log[(1+\lambda)(1+\theta h)^{k+1}+(1-\lambda)(1-\theta h)^{k+1}]\,.$$

Then $(Z_{\theta,\lambda}/e^{N\varphi}) \in [e^{-N\varepsilon}, e^{N\varepsilon}]$ whp for any $\varepsilon > 0$.

Proof. Let $\varphi_N(\theta,\lambda) \equiv N^{-1}\log Z_{\theta,\lambda}$. The proof consists in showing that $\mathbb{E}|\varphi_N(\theta,\lambda) - \varphi(\theta,\lambda)| \stackrel{N}{\to} 0$, whence the thesis follows by Markov inequality applied to the event $(Z_{\theta,\lambda}/e^{N\varphi}) \notin [e^{-N\varepsilon}, e^{N\varepsilon}]$.

It can be proved that $\varphi_N(\theta,\lambda)$ is uniformly (in N) continuous with respect to λ. We can therefore restrict, without loss of generality to $\lambda > 0$.

Next we notice that the above claim is true for $\theta = 0$ by elementary algebra: $Z_{0,\lambda} = 2^N = e^{N\varphi(0,\lambda)}$. In order to prove it for $\theta > 0$, we write (omitting the dependence on λ that is fixed throughout)

$$\mathbb{E}|\varphi_N(\theta) - \varphi(\theta)| \leq \int_0^\theta \mathbb{E}\,|\partial_\theta\varphi_N(\theta') - \partial_\theta\varphi(\theta')|\,\mathrm{d}\theta'\,.$$

We will then show that $|\partial_\theta\varphi_N(\theta,\lambda) - \partial_\theta\varphi(\theta,\lambda)|$ is bounded by $(k+1)/(1-\theta^2)$, and is smaller than ε whp for any $\varepsilon > 0$. This implies the thesis by applying dominated convergence theorem to the above integral.

An elementary calculation omitted from this abstract leads to $(1-\theta^2)\partial_\theta\varphi(\theta,\lambda) = \frac{k+1}{2}\frac{\theta+h_*^2}{1+\theta h_*^2}$. Analogously, simple calculus yields $(1-\theta^2)\partial_\theta \log Z_{\theta,\lambda} = \sum_{(k,l)\in E}\mathbb{E}\{X_kX_l|G_N\}$, and therefore $(1-\theta^2)\partial_\theta\varphi_N(\theta,\lambda) = \frac{k+1}{2}\mathbb{E}\{X_kX_l|G_N\}$ averaged over a uniformly random edge $(k,l) \in E$. As a consequence we have $|\partial_\theta\varphi|, |\partial_\theta\varphi_N| \leq (k+1)/2(1-\theta^2)$ and, because of Lemma 5.4 $|\partial_\theta\varphi - \partial_\theta\varphi_N| \leq \varepsilon$ whp. This proves our claim. □

Proof of Theorem 1.6. Throughout the proof, we set $\lambda = 0$. Let us first prove that $k\theta \leq 1$ reconstruction is unsolvable, i.e. for any $\varepsilon > 0$ there exists t such that $||\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,|G_N\}||_{\text{TV}} \leq \varepsilon$ whp. By the Markov property of $\mathbb{P}\{\,\cdot\,|G_N\}$, (and using shorthands $\overline{\mathsf{B}}$ for $\overline{\mathsf{B}}(r,t)$ and D for $\mathsf{D}(r,t)$)

$$||\mathbb{P}_{r,\overline{\mathsf{B}}}\{\,\cdot\,,\,\cdot\,|G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}\mathbb{P}_{\overline{\mathsf{B}}}\{\,\cdot\,|G_N\}||_{\text{TV}} \leq \\ \leq \sup_{x_{\mathsf{D}}} ||\mathbb{P}_{i|\mathsf{D}}\{\,\cdot\,|X_{\mathsf{D}} = x_{\mathsf{D}}, G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}||_{\text{TV}}\,. \tag{16}$$

By symmetry of $\mathbb{P}\{\,\cdot\,|G_N\}$ under exchange of $+1$ and -1 at $\lambda = 0$, cf. Eq. (13), $\mathbb{P}_r\{+1|G_N\} = 1/2$. On the other hand, by Griffiths inequalities, the probability for $X_r = +1$ is a monotone function of the values other spins are conditioned to. Therefore the right hand side of Eq. (16) equals $|\mathbb{P}_{r|\mathsf{D}}\{+1|X_{\mathsf{D}} = \underline{+1}, \mathsf{B}\} - 1/2|$ (we emphasized that, conditional on X_{D}, X_r depends on G_N only through B).

Finally, we recall that $\mathsf{B}(r,t)$ is with high probability a $k+1$ regular tree of depth t rooted at r. Assuming this to be the case, the conditional distribution of the root variable can be computed through a recursive dynamic-programming procedure, that we omit. The result is (for $h_{+,s} \equiv f^{\circ s}_{\theta,0}(+1)$)

$$\mathbb{P}_{i|\mathsf{D}}\{+1|X_{\mathsf{D}} = \underline{+1}, \mathsf{B}\} = \frac{1}{2}\Big\{1 + g_{\theta,0}(h_{+,t-1})\Big\}\,,$$

where $g_{\theta,\lambda}(h)$ is defined as $f_{\theta,\lambda}(h)$, cf. Eq. (15), with k replaced by $k+1$. The thesis follows from Lemma 5.2.

We shall now prove that non-reconstructibility implies $k\theta \leq 1$. If $\overline{X} \equiv N^{-1}\sum_i X_i$ is the 'magnetization,' we claim that non-reconstructibility implies $|\overline{X}| \leq \delta$ whp for any $\delta > 0$. In fact, because of non-reconstructibility, we can fix t in such a way that $||\mathbb{P}_{r,\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,\,\cdot\,|G_N\} - \mathbb{P}_r\{\,\cdot\,|G_N\}\mathbb{P}_{\overline{\mathsf{B}}(r,t)}\{\,\cdot\,,|G_N\}||_{\text{TV}} \leq \delta$ whp. Further notice that $\mathbb{E}\{\overline{X}^2\} = \mathbb{E}\{X_rX_j\}$ for two uniformly random vertex $r,j \in [N]$. Since $r \notin \mathsf{B}(i,t)$ whp and $\mathbb{E}\{X_r|G_N\} = 0$ by symmetry, we then have $\mathbb{E}\{X_rX_j|G_N\} \leq \delta$ whp, and, as a consequence, $\mathbb{E}\{\overline{X}^2\} \leq 2\delta$ for all N large enough. The claim follows from this result together with $|\overline{X}| \leq 1$

The thesis is proved by contradiction showing that for $k\theta > 1$, there is some $\delta > 0$ such that $|\overline{X}| > \delta$ whp. Denote by Typ the set of graphs G_N such that $Z_{\theta,0} \geq e^{N[\varphi(\theta,0)-\xi]}$ for some ξ to be fixed below. Then, for any $G_N \in \mathsf{Typ}$, by Eq. (14) and discussion below,

$$\mathbb{P}\{|\overline{X}| \leq \delta|G_N\} \leq e^{-N[\varphi(\theta,0)-\xi]}\sum_{|M|\leq\delta}\widehat{Z}_{\theta,M}\,.$$

Since by Lemma 5.5, $\mathbb{P}\{|\overline{X}| \leq \delta\} \leq \mathbb{E}\{\mathbb{P}\{|\overline{X}| \leq \delta|G_N\}\,\mathbb{I}_{G_N\in\mathsf{Typ}}\} + o_N(1)$, we then have (estimating $\mathbb{E}\widehat{Z}_{\theta,M}$ with Lemma 5.3,

$$\mathbb{P}\{|\overline{X}| \leq \delta\} \leq N\,e^{-N[\varphi(\theta,0)-\xi]} \times C2^N\,e^{ND\delta^2} + o_N(1)\,.$$

The proof follows by showing that $\varphi(\theta,0) > \log 2$, and taking δ and ξ small enough to make the first term above exponentially small as $N \to \infty$.

To show that $\varphi(\theta,0) > \log 2$ for $k\theta > 1$, observe that $\phi(\theta,0,0) = \log 2$ and (after some calculus)

$$\left.\frac{\partial\phi}{\partial h}\right|_{\lambda=0} = -\frac{(k+1)\theta h}{1+\theta h^2} + \frac{(k+1)\theta f_{\theta,0}(h)}{1+\theta h f_{\theta,0}(h)}\,.$$

Since $f_{\theta,0}(0) = k\theta h + O(h^2)$, $h = 0$ is a local minimum (at θ fixed) of $\phi(\theta,0,h)$. Further $\partial_h\phi = 0$ if and only if $h = f_{\theta,\lambda}(h)$. Recall [7] that, for, $k\theta > 1$, $f_{\theta,\lambda}$ has 3 fixed points: $h = 0$ and $h = \pm h_*$, for $h_* > 0$. As a consequence h_* must be a local maximum and hence $\varphi(\theta,0) = \phi(\theta,0,h_*) > \phi(\theta,0,0) = \log 2$. □

Acknowledgments

This work was largely motivated by a discussion of one of the authors (A.M.) with Elchanan Mossel, and by the desire to substantiate the claims made in that discussion. It is a pleasure to acknowledge the debt we have with him.

References

[1] D. Aldous, J. M. Steele, 'The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence,' in *Probability on discrete structures*, H. Kesten (ed.), New York, 2004

[2] F. Guerra and F. L. Toninelli, 'The high temperature region of the Viana-Bray diluted spin glass model,' *J. Stat. Phys.* **115** (2004), 531-555.

[3] D. Achlioptas and A. Naor, 'The two possible values of the chromatic number of a random graph,' Proc. of ACM STOC 2004

[4] A. Guionnet and B. Zegarlinski, 'Lectures on Logarithmic Sobolev Inequalities,' Séminaire de Probabilites de Strasbourg, **36** (2002) 1–134

[5] M. Dyer, A. Sinclair, E. Vigoda and D. Weitz, 'Mixing in time and space for lattice spin systems: A combinatorial view,' Rand. Struc. and Alg. 24, 461-479 (2004)

[6] N. Berger, C. Kenyon, E. Mossel and Y. Peres, 'Glauber dynamics on trees and hyperbolic graphs,' *Prob. Theory Rel. Fields* **131**, 311 (2005)

[7] F. Martinelli, A. Sinclair, and D. Weitz, 'The Ising model on trees: boundary conditions and mixing time,' Proc. of IEEE FOCS 2003

[8] A. Montanari and G. Semerjian, 'Rigorous Inequalities Between Length and Time Scales in Glassy Systems,' *J. Stat. Phys.* 125 (2006), 23-54

[9] M. Mézard, G. Parisi and R. Zecchina 'Analytic and Algorithmic Solution of Random Satisfiability Problems,' *Science* **297** (2002), 812-815.

[10] M. Mézard, T. Mora and R. Zecchina, 'Clustering of Solutions in the Random Satisfiability Problem,' *Phys. Rev. Lett.* **94**, 197205.

[11] D. Achlioptas and F. Ricci-Tersenghi, 'On the solution-space geometry of random constraint satisfaction problems,' Proc. of ACM STOC 2006

[12] M. Mézard and A. Montanari 'Reconstruction on Trees and Spin Glass Transition,' *J. Stat. Phys.* **124** (2006), 1317-1350.

[13] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian and L. Zdeborova, 'Gibbs States and the Set of Solutions of Random Constraint Satisfaction Problems,' Proc. Natl. Acad. Sci. 1004, 10318 (2007)

[14] A. Dembo and A. Montanari, in preparation

[15] C. Daskalakis, E. Mossel, and S. Roch, 'Optimal Phylogenetic Reconstruction,' Proc. of ACM STOC 2006.

[16] S. Bhamidi, R. Rajagopal and S. Roch 'Network Delay Inference from Additive Metrics,' preprint, 2007

[17] H.-O. Georgii. *Gibbs Measures and Phase Transition*, de Gruyter, Berlin, 1988.

[18] P. M. Bleher, J. Ruiz, and V. A. Zagrebnov, 'On the purity of limiting Gibbs state for the Ising model on the Bethe lattice,' *J. Stat. Phys.*, **79** (1995), 473–482.

[19] E. Mossel and Y. Peres, 'Information Flow on Trees,' *Ann. Appl. Probab.* **13** (2003), 817-844.

[20] W. Evans, C. Kenyon, Y. Peres, and L. J. Schulman. 'Broadcasting on trees and the Ising model,' *Ann. Appl. Probab.*, **10** (2000) 410–433.

[21] C. Borgs, J. Chayes, E. Mossel and S. Roch, 'The Kesten-Stigum Reconstruction Bound Is Tight for Roughly Symmetric Binary Channels,' Proc. of IEEE FOCS 2006

[22] E. Mossel, D. Weitz, and N. Wormald, 'On the hardness of sampling independent sets beyond the tree threshold,' `arXiv:math/0701471`

[23] B. Bóllobas, 'A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs,' Eur. J. Combin. **1** (1980), 311.

[24] D. Achlioptas, A. Naor, and Y. Peres, 'Rigorous location of phase transitions in hard optimization problems,' Nature **435** (2005), 759