Phase transitions of random constraint satisfaction problems

Allan Sly

Abstract. Random constraint satisfaction problems, such as the random K-SAT model and colourings of random graphs naturally emerge in the study of combinatorics and theoretical computer science. Ideas from statistical physics describe a series of phase transitions these models undergo as the density of constraints increases. Throughout we will focus on the example of the maximal independent set of a random regular graph as a simple example of this phenomena and then conclude by describing the additional technical challenges needed to establish the Satisfiability Conjecture for large K.

Contents

1. Introduction

In this paper we will describe the theory behind different phase transitions that random constraint satisfaction problems are predicted to undergo and the progress that has been made in proving some of these predictions.

In particular we will discuss the one-step replica symmetry breaking (1RSB) universality class of models which each are believed to undergo the same sequence of transitions as the density of constraints grows. Questions of interest include how many constraints a random constraint satisfaction problem can have while still being satisfiable (the *satisfiability threshold*), how many satisfying assignments of the variables there are, the geometry of how the solutions are arranged and the algorithmic question of finding solutions efficiently. These problems have been extensively studied in probabilistic combinatorics and theoretical computer science [2,9,27,33,34,51] but ideas originating from the statistical mechanics of disordered systems have dramatically advanced the subject. Beginning with the study of spin glasses, particularly the work of Parisi on the Sherrington-Kirkpatrick model [37,44,45], a mean field model, these ideas have been adapted to the sparse setting where they give a comprehensive description of the different phase transitions [32,35,38,41].

A *constraint satisfaction problem* or CSP consists of a set of variables subject to a collection of constraints. A *random constraint satisfaction problem* is a CSP where the collection of constraints are chosen randomly from some distribution. We will focus on four random CSPs: colourings, independent sets, K-SAT and NAESAT. The first two are simple examples from combinatorics. In a proper q-colouring of a graph, the variables are the colours of the vertices and the edges are the constraints, that is that neighbouring vertices must have different colours. So a random colouring of a random graph, such as on an Erdos Renyi random graph $G(n, d/n)$ or a random d-regular graph, are random CSPs.

Another variant of CSPs is to find the satisfying assignment that maximizes some objective function. A natural example of this is the maximal independent set question. An *independent set* of a graph $G = (V, E)$ is a subset $I \subset V$ such that no two vertices of I share an edge. Independent sets of large density on random d-regular graphs (when d is large) are believed to undergo the same series of phase transitions.

The final two random CSPs are models of random Boolean formulas. The random K-SAT model gives a random Boolean formula over n Boolean variables $\underline{x} = (x_1 \ldots, x_n) \in \{T, F\}^n$. Its clauses or constraints are formed by taking the OR of K of the variables or their negations chosen uniformly at random e.g. $(x_1 \vee \neg x_7 \vee x_9)$. The formula is then formed by taking the AND of αn clauses chosen independently. A solution, or *satisfying assignment* is a vector $\underline{x}$ such that the formula evaluates to true or equivalently that all the constraints evaluate to true. We say the formula is satisfiable when there exists at least one satisfying assignment.

The final model is random K-NAESAT meaning "not all equal" SAT. This is simply a modification of a random K-SAT formula where we ask that both $\underline{x}$ and its negation satisfy the formula. Equivalently in each clause at least one term must evaluate to true and at least one to false, hence the name not-all-equal. This shares similar properties to K-SAT but has additional

symmetries which have made it more tractable. In many instances, it is the first 1RSB model for which various prediction are established.

It will be very useful to consider the behaviour of a random solution to a CSP, the distribution of which can be described as a spin system on the underlying graph. We will begin with some preliminaries on spin systems, particularly spin systems on trees as they play a crucial role in the theory. We will then focus on computing the free energy of spin systems on random graphs, in particular on the relationship between fixed points of the belief propagation equations and computing the moments of the partition function. With these computations we will see how clustering naturally emerges as an obstacle to computing thresholds and then describe the way in which clusters of random CSP solutions can themselves be recast as a spin system. In Section 8 we explain the series of phase transitions the 1RSB class of models are predicted to undergo. We demonstrate such a threshold in Section 9 where we describe the proof of the size of the maximal independent set for a random d-regular graph. Finally in Section 10 we give, at a high level, the main ideas of the proof of the satisfiability threshold for random K-SAT when K is large.

2. Preliminaries

In this theory it is often useful to consider a random solution of the CSP which leads us to the study of *spin systems*. These models, also called *graphical models* or *Markov random fields* in some areas, are a broad class of stochastic processes on networks giving a probability distribution on $\mathcal{X}^V$ for some (usually) discrete set $\mathcal{X}$, satisfying a local Markov property. Let us take the example of a random uniformly chosen independent set which we consider as an element $\sigma \in \{0,1\}^V$ where σ_u is the indicator that u is in the independent set. We can write this as

$$\mathbb{P}[\sigma] = \frac{1}{Z} \prod_{u \sim v} I(\sigma_u \sigma_v = 0)$$

where the normalizing constant Z, called the partition function, is in this case simply the number of independent sets. Of course we will be interested in larger independent sets so it is useful to give larger weight to larger independent sets. The *Hardcore Model* is a distribution over independent sets given by

$$\mathbb{P}[\sigma] = \frac{1}{Z_\lambda} \lambda^{\sum_u \sigma_u} \prod_{u \sim v} I(\sigma_u \sigma_v = 0)$$

This is a special case of a more general collection of models called *spin systems*. These are distributions over $\mathcal{X}^V$ of the form

$$\mu(\sigma) = \frac{1}{Z} \prod_{u \in V} \psi_u(\sigma_u) \prod_{(u,v) \in E} \psi_{u,v}(\sigma_u, \sigma_v) \tag{2.1}$$

where $\{\psi_u\}_{u\in V}$ and $\{\psi_{u,v}\}_{(u,v)\in E}$ are non-negative functions on $\mathcal{X}$ and $\mathcal{X}^2$ respectively. In most cases we will consider homogeneous spin systems where the ψ_u (resp. $\psi_{u,v}$) do not depend on the vertex u (resp. edge (u,v)). In each case the partition function Z is the normalizing constant to make the distribution a probability measure. The value σ_u at the vertex u is called the *spin* at u.

2.1. Markov random field property. Spin systems share a spatial Markov property. If $A \subset V$ we write the exterior boundary of A as

$$\partial A := \{v \in V \setminus A : \exists u \in A, (u,v) \in E\}.$$

The Markov Random Field Property says that to understand the conditional distribution of σ on A given the rest of the configuration it is enough to know σ on ∂A. That is, for any $x \in \mathcal{X}^V$,

$$\mathbb{P}[\sigma_A = x_A \mid \sigma_{A^c} = x_{A^c}] = \mathbb{P}[\sigma_A = x_A \mid \sigma_{\partial A} = x_{\partial A}]. \tag{2.2}$$

The distribution of σ_A given $\sigma_{\partial A} = x_{\partial A}$ is simply the spin system on $A \cup \partial A$ with the spins on ∂A fixed to $x_{\partial A}$. It may be the case, for example for colourings, that not every boundary condition gives rise to configurations with positive probability. We call a spin system *permissive* if for every boundary condition on ∂A, there is a configuration of A with positive probability. The hardcore model is permissive because the empty independent set always has positive probability.

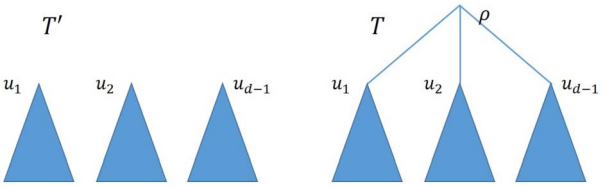
2.2. Infinite graphs. So far we have only defined the model on finite graphs and in general (2.1) only makes sense on finite graphs but taking local weak limits of spin systems on random graphs leads to models on infinite trees. The definition is extended to infinite graphs via the Markov Random Field Property from equation (2.2). On an infinite graph V with a set of weights $\{\psi_u\}, \{\psi_{u,v}\}$ a measure μ on $\mathcal{X}^V$ is an infinite *Gibbs measure* for the spin system if for any finite $A \subset V$,

$$\mu[\sigma_A = x_A \mid \sigma_{A^c} = x_{A^c}] = \mu[\sigma_A = x_A \mid \sigma_{\partial A} = x_{\partial A}]$$

holds and the conditional distribution is given by (2.1) on the finite graph $A \cup \partial A$. This is called the DLR property for Dobrushin–Lanford–Ruelle. We will particularly be interested in the Gibbs measures on infinite trees, particularly the infinite d-regular tree which we denote by T^d.

3. Spin systems on trees

As trees are the local weak limits of random graphs, spin systems on trees play a major role in understanding spin systems on random graphs. The Markov property is particularly useful here as conditional on the spin at one vertex, it's neighbours are conditionally independent. This conditional independence makes the models amenable to recursive calculations of marginals.



Let T be a finite tree rooted at ρ with children $u_1, \ldots, u_d$ and let T' be the graph with the edges at ρ removed. We will let $\mathbb{P}_G$ denote the Gibbs measure on a graph G. We will relate the marginal of ρ in T in terms of the marginal of the u_i in the forest T' which will give an efficient method for recursively calculating marginal distributions.

Let T_i be the subtree rooted at u_i and let $m_{u_i \to \rho}(x_i) = \mathbb{P}_{T_i}[\sigma_{u_i} = x_i]$ denote the marginal of u_i in T_i. In the graph T' subtrees T_i are disconnected and so their spins are independent. Then summing over all the spin configurations on ρ and the u_i,

$$\begin{aligned} Z_T &= \sum_{x \in \mathcal{X}} \psi(x) \sum_{\{x_i\} \in \mathcal{X}^d} \prod_{i=1}^{d} \psi(x, x_i) m_{u_i \to \rho}(x_i) Z_{T_i} \\ &= \sum_{x \in \mathcal{X}} \psi(x) \prod_{i=1}^{d} \left(\sum_{x_i \in \mathcal{X}} \psi(x, x_i) m_{u_i \to \rho}(x_i) Z_{T_i} \right) \end{aligned}$$

and

$$\mathbb{P}_T[\sigma_\rho = x_\rho] = \frac{\psi(x_\rho) \prod_{i=1}^{d} \left(\sum_{x_i \in \mathcal{X}} \psi(x_\rho, x_i) m_{u_i \to \rho}(x_i) \right)}{\sum_{x' \in \mathcal{X}} \psi(x') \prod_{i=1}^{d} \left(\sum_{x_i \in \mathcal{X}} \psi(x', x_i) m_{u_i \to \rho}(x_i) \right)} \tag{3.1}$$

Building up from the leaves this gives an algorithm to determine the marginal at the root. We can formalize this in terms of what are called message passing algorithms. Let $\overrightarrow{E}$ be the set of directed edges in the graph. For $(u, v) \in \overrightarrow{E}$ we write

$$m_{u \to v}(x_u) := \mathbb{P}_{T \setminus (u,v)}[\sigma_u = x_u]$$

where $T \setminus (u, v)$ is the tree T with the edge (u, v) removed. This is interpreted as the message from u telling v what it believes its marginal is removing the effect of v. When u is a leaf of the tree,

$$m_{u \to v}(x_u) = \frac{\psi(x_u)}{\sum_{x' \in \mathcal{X}} \psi(x')}.$$

The beliefs on internal edges can be calculated by applying the same reasoning as in equation (3.1),

$$\begin{aligned} m_{u \to v}(x_u) &= \frac{\psi(x_u) \prod_{w \in \partial u \setminus v} \left(\sum_{x_w \in \mathcal{X}} \psi(x_u, x_w) m_{w \to u}(x_i) \right)}{\sum_{x' \in \mathcal{X}} \psi(x') \prod_{w \in \partial u \setminus v} \left(\sum_{x_w \in \mathcal{X}} \psi(x', x_w) m_{w \to u}(x_i) \right)} \\ &=: BP[\{m_{w \to u}\}_{w \in \partial u \setminus v}](x_u). \end{aligned}$$

The function BP is called the *belief propagation* function. For a more in depth discussion of belief propagation see [36]. The marginal distribution at a vertex is simply

$$\mathbb{P}_T[\sigma_v = x_v] = BP[\{m_{u\to v}\}_{u\in\partial v}](x_u)$$

The partition function of the spin system on the tree can also be readily calculated from the belief propagation messages. Defining the vertex and edge terms as follows

$$\Phi_v = \log \sum_{x\in\mathcal{X}} \psi(x_v) \prod_{i=1}^{d} \left(\sum_{x_{u_i}\in\mathcal{X}} \psi(x_v, x_{u_i}) m(x_{u_i\to v}) \right)$$

$$\Phi_{(v,u)} = \log \sum_{x,x'\in\mathcal{X}} \psi(x,x') m_{u\to v}(x) m_{v\to u}(x'),$$

which correspond to the change in $\log Z$ of removing a vertex $v \in V$ or and edge $(v,u) \in E$ respectively. Then the log partition can be written as [36],

$$\log Z_{G_n} = \sum_{v\in V} \Phi_v - \sum_{(u,v)\in E} \Phi_{(u,v)}. \tag{3.2}$$

While this formula is valid only for trees, an important heuristic from physics is the assumption that it holds in many cases on locally treelike graphs.

A set of messages $\{r_{u\to v}\}_{(u,v)\in\overrightarrow{E}}$ is a BP-fixed point if

$$r_{u\to v} = BP[\{r_{w\to u}\}_{w\in\partial u\setminus v}] \tag{3.3}$$

for all $(u,v) \in \overrightarrow{E}$. Our construction show that on a tree this is unique. However, on an infinite tree or a graph with cycles there may be multiple BP-fixed points

On an infinite graph each set of BP messages $\{r_{u\to v}\}$ defines a Gibbs measure. The distribution of σ_A for a finite connected set $A \subset V$ is given by

$$\mu_r[\sigma_A = x_A] = \frac{1}{Z} \sum_{x_{\partial A}\in\mathcal{X}^{\partial A}} \prod_{\substack{u\sim v \\ u\in\partial A, v\in A}} r_{u\to v}(x_u)\psi(x_u,x_v) \prod_{v\in A} \psi(x_v) \prod_{\substack{v\sim v' \\ v,v'\in A}} \psi(x_v,x_{v'})$$

which can be verified to satisfy the properties of a Gibbs measure. The effect of the measure outside of $A \cup \partial A$ is to change ψ_u with $r_{u\to v}(x_u)$ on ∂A. We will let μ_r denote the Gibbs measure generated from the BP-fixed point r.

A special case of a BP-fixed point in a d-regular tree T^d is when all the messages in the tree are constant and

$$m = BP[m^{\times d-1}] \tag{3.4}$$

where $m^{\times d-1} = m, \ldots, m$ denotes $d-1$ copies of m. We will call this a *translation invariant BP-fixed point* or TIFP. For any homogeneous spin system there is always a TIFP by the Brouwer fixed-point theorem and gives

rise to a translation invariant Gibbs measure. In the case of the hardcore model, the BP fixed point equation is

$$BP[m](1) = \frac{\lambda(1-m(1))^d}{1+\lambda(1-m(1))^d}. \tag{3.5}$$

Since this is a decreasing function of $m(1)$ there is a unique TIFP.

The Gibbs measure correspond to a TIFP has a simple description as a *broadcast model* on T^d. A *broadcast model* is a random spin configuration on the tree given by the following Markov model. Let M be a $\mathcal{X} \times \mathcal{X}$-Markov transition matrix M, reversible with respect to ν. Then the broadcast model on T given by M chooses the state of the root σ_ρ according to ν and then the state of children is assigned as $\mathbb{P}[\sigma_u = y \mid \sigma_{u^+} = x] = M_{xy}$ where u^+ is the parent of u. Thus along each path in the tree, the states are given by the Markov chain with transition M in stationarity. Reversibility means that the distribution does not depend on the location of the root.

For a TIFP with BP fixed point m the law of the root is given by

$$\nu(x) = BP[m^{\times d}](x)$$

and the transition matrix is

$$M_{xy} = \frac{m(y)\psi(x,y)}{\sum_{y' \in \mathcal{X}} m(y')\psi(x,y')}.$$

4. Reconstruction threshold

The *reconstruction problem* asks whether distant spins provide information about the state at the root or in its neighbourhood under some Gibbs measure μ. We let B_ℓ be the set of vertices at distance at most ℓ from the root ρ and $S_\ell = \{u : d(u,\rho) = \ell\}$ denote the vertices at distance exactly ℓ. For a finite set $A \subset V$, we write

$$n_{A,\ell}(x_A) = \mu(\sigma_A = x_A \mid \sigma_{B_\ell^c}).$$

Since $\sigma_{B_\ell^c}$ is a random variable, so is $n_{A,\ell}$. Note that for large enough ℓ we have that $A \subset B_\ell$ and so by the Markov Random Field Property $n_{A,\ell}(x_A) = \mu(\sigma_A = x_A \mid \sigma_{S_\ell})$. Furthermore, $n_{A,\ell}$ is a bounded backwards martingale so converges almost surely,

$$n_{A,\ell}(x_A) \overset{a.s}{\to} n_A(x_A).$$

Taking expected values,

$$\mathbb{E}n_A(x_A) = \mathbb{E}n_{A,\ell}(x_A) = \lim_\ell n_{A,\ell}(x_A) = \mu(\sigma_A = x_A).$$

If $n_A(x_A)$ were always identically equal to $\mu(\sigma_A = x_A)$ then distant spins provide no information asymptotically about σ_A. If this holds for all A then we say that the reconstruction problem is *non-solvable* for μ. It is also equivalent to tail-triviality of the Gibbs measure which we describe as follows. Conversely if for some x_A,

$$\mathbb{P}[n_A(x_A) \neq \mu(\sigma_A = x_A)] > 0$$

then we say that the reconstruction problem is *solvable*. At the same time we may construct a random BP-fixed point derived from σ as follows. We set

$$m^{(0)}_{u\to v}(x) = I(\sigma_u = x)$$

and

$$m^{(t)}_{u\to v}(x) = BP[\{m^{(t-1)}_{u'\to u}(x)\}_{u'\in\partial u\setminus v}],$$

that is $m^{(t)}_{u\to v}$ is the conditional distribution of u in the tree $T\setminus(u,v)$ with a boundary condition on at depth t from u given by σ. The distribution of the spins on an edge given the spins at distance t is

$$\frac{1}{Z} m^{(t)}_{u\to v}(x_u) m^{(t)}_{v\to u}(x_v)\psi(x_u,x_v)$$

Since this converges to $n_{(u,v)}(x_u,x_v)$ we must have that $m^{(t)}_{u\to v}$ and $m^{(t)}_{v\to u}$ converge as $t\to\infty$. We denote this limit $m_{u\to v} = m^{\sigma}_{u\to v}$ which by construction must be a BP-fixed point. The probability n_A can be constructed as

$$n_A(x_A) = \frac{1}{Z}\sum_{x_{\partial A}}\prod_{u\in A}\psi(x_u)\prod_{\substack{u\sim v\\ u,v\in A}}\psi(x_u,x_v)\prod_{\substack{u\sim v\\ u\in\partial A, v\in A}} m_{u\to v}(x_u)\psi(x_u,x_v).$$

An equivalent formulation of solvability of the reconstruction problem is extremality of Gibbs measures [42]. Recall that any convex combination of Gibbs measures is also a Gibbs measures. We say a Gibbs measure is *extremal* or *pure* if it cannot be written as a non-trivial convex combination of other Gibes measures. This decomposition will be important as we consider the clustering of solutions of random constraint satisfaction problems.

4.1. Reconstruction for the colouring model. In the colouring model it may be the case that the boundary condition exactly determines the value at the root which is called *freezing*. When the boundary is at depth one this corresponds to all the other colours appearing among the children of the root, leaving the colour of the root as the only possibility. We write p_ℓ for the depth ℓ freezing probability,

$$p_\ell := \mathbb{P}[m^{(\ell)}_{\rho\to\rho^+}(x) = 1 \mid \sigma_\rho = x],$$

which by symmetry does not depend on x. Given a boundary condition at depth ℓ, $\mathrm{Bin}(d,p_{\ell-1})$ of the children of ρ can be determined exactly.

Having all colours appear among the children is an instance of the coupon collector problem. Let $f(n,m)$ denote the probability that after m IID samples drawn uniformly from n possible 'coupons' that all n have appeared at least once. The well known theorem for the coupon collector problem is that for any $\delta>0$,

$$f(n,(1+\delta)n\log n)\to 1, \qquad f(n,(1-\delta)n\log n)\to 0$$

as $n \to \infty$. If m of the children of the root are known exactly, then the state of the root is known with probability $f(q-1,m)$. Thus

$$p_n = \mathbb{E}[f\,(q-1, \mathrm{Bin}(d, p_{n-1}))]$$

If $d \geq (1+\epsilon)q\log q$ then for large enough q and $\delta > 0$ sufficiently small,

$$\mathbb{P}[\mathrm{Bin}(d, 1-\delta) \geq (1+\epsilon/2)q\log q] \geq 1 - \delta/2,$$

and

$$f\,(q-1, (1+\epsilon/2)q\log q) \geq \frac{1-\delta}{1-\delta/2}.$$

Thus if $p_{n-1} \geq 1-\delta$ then

$$p_n \geq \mathbb{E}[f\,(q-1, \mathrm{Bin}(d, 1-\delta))] \geq 1-\delta,$$

and $\inf p_n \geq 1-\delta$ which implies the reconstruction problem is solvable [42] since the root is frozen with probability bounded away from 0. Conversely if $d \leq q\log q$ it can be shown that there is non-reconstruction for large q so $q\log q$ gives the correct asymptotics [47].

5. Free energy and the Ising model on random graphs

A key quantity in the analysis of spin systems and random CSPs is the *free energy*, the normalized log-partition function

$$\Phi = \lim_{n\to\infty} \frac{1}{n} \log Z_n$$

assuming the limit exists. We will discuss two related methods for evaluating the free energy, the cavity method and moments. We will begin with the cavity method which is a non-rigourous method from statistical mechanics which can be made rigorous in some cases and gives correct predictions in many others.

A sequence of graphs G_n is *locally treelike* if for all ℓ,

$$\frac{1}{n}\#\{v : B_\ell(v) \text{ is a tree}\} \to 1.$$

We will consider methods for evaluating the free energy of a d-regular locally treelike sequence of graphs which have the infinite d-regular tree as the local weak limit. The classic example of such a sequence is random d-regular graphs. Consider the following operation for removing two vertices of the graph. Start with the graph $G = G_n$, pick two vertices ρ, ρ' uniformly at random and remove them to form the graph G^- with $n-2$ vertices, $2d$ of which have degree $d-1$ (note we are removing two vertices in case d is odd). Choosing a uniform perfect matching of these vertices we add d edges to form G_{n-2}, a d-regular graph on $n-2$ vertices. If G_n is chosen from the configuration model, then so is G_{n-2}. If we only assume that G_n is locally treelike then the following lemma (a small variation of Proposition 3.3 of [20]) shows that after many iterations of this operation, it remains locally treelike.

LEMMA 1. *For all $\delta, \ell > 0$ and $0 < \alpha < 1$, there exists $\delta', \ell' > 0$ such that the following holds. If G_n is a graph satisfying*

$$\frac{1}{n}\#\{v \in G_n : B_{\ell'}(v) \text{ is a tree in } G_n\} \geq 1 - \delta',$$

then with high probability for each $k \in [0, \frac{1-\alpha}{2}n]$,

$$\frac{1}{|G_{n-2k}|}\#\{v \in G_{n-2k} : B_{\ell}(v) \text{ is a tree in } G_{n-2k}\} \geq 1 - \delta.$$

Now suppose that the local weak limit of the graph G_n and a configuration σ from the spin system ψ is a Gibbs measure μ on the infinite d-regular tree given by an extremal TIFP with message m. If $u_1, \ldots, u_d$ are the neighbours of ρ then for large ℓ, extremality means that the distribution of $\sigma_{\rho \cup \partial\rho}$ is almost independent of $\sigma_{S_\ell(\rho)}$ and

$$\mathbb{P}[\sigma_{\rho\cup\partial\rho} = x_{\rho\cup\partial\rho} \mid S_\ell(\rho)] \approx \frac{\psi(x_\rho)\prod_{i=1}^d (\psi(x_\rho, x_{u_i})m(x_{u_i}))}{\sum_{x' \in \mathcal{X}} \psi(x') \prod_{i=1}^d \left(\sum_{x'_{u_i} \in \mathcal{X}} \psi(x', x_{u_i})m(x_{u_i})\right)}.$$

Then provided $d(\rho, \rho') > 2\ell + 2$,

$$\begin{aligned}
&\mathbb{P}[\sigma_{\rho\cup\partial\rho} = x_{\rho\cup\partial\rho}, \sigma_{\rho'\cup\partial\rho'} = x'_{\rho'\cup\partial\rho'}] \\
&\quad = \mathbb{E}\mathbb{P}[\sigma_{\rho\cup\partial\rho} = x_{\rho\cup\partial\rho}, \sigma_{\rho'\cup\partial\rho'} = x'_{\rho'\cup\partial\rho'} \mid S_\ell(\rho), S_\ell(\rho')] \\
&\quad = \mathbb{E}\mathbb{P}[\sigma_{\rho\cup\partial\rho} = x_{\rho\cup\partial\rho} \mid S_\ell(\rho)]\mathbb{P}[\sigma_{\rho'\cup\partial\rho'} = x'_{\rho'\cup\partial\rho'} \mid S_\ell(\rho')] \\
&\quad \approx \mathbb{P}[\sigma_{\rho\cup\partial\rho} = x_{\rho\cup\partial\rho}]\mathbb{P}[\sigma_{\rho'\cup\partial\rho'} = x'_{\rho'\cup\partial\rho'}]
\end{aligned}$$

where the second equality used the Markov Random Field property. Thus the distribution in the local negibourhoods of ρ and ρ' are asymptotically independent and we can calculate the effect of the change in the partition function from G_n to G_{n-2}. First note that since

$$\mathbb{P}_{G_n}[\sigma_{\rho\cup\partial\rho} = x_{\rho\cup\partial\rho}] \propto \psi(x_\rho)\prod_{i=1}^d (\psi(x_\rho, x_{u_i})m(x_{u_i}))$$

we have that

$$\mathbb{P}_{G^-}[\sigma_{\partial\rho} = x_{\partial\rho}] = \prod_{i=1}^d m(x_{u_i}).$$

Then we define the effect on the partition function of removing a vertex from a tree with messages m as

$$\Phi_{vertex}(m) := \log \sum_{x_\rho \in \mathcal{X}} \psi(x_\rho) \prod_{i=1}^d \left(\sum_{x_{u_i} \in \mathcal{X}} \psi(x_\rho, x_{u_i})m(x_{u_i})\right)$$

Applying this to the neighbourhood of ρ and ρ' and the fact that they are approximately independent we have that

$$\log \frac{Z_{G_n}}{Z^-} = 2\Phi_{vertex}(m) + o(1)$$

Similarly the effect of removing an edge is

$$\Phi_{edge}(m) := \log \sum_{x,x' \in \mathcal{X}} \psi(x,x') m(x) m(x'),$$

and so removing d edges from G_{n-2} to form G^- we have that

$$\log \frac{Z_{G_{n-2}}}{Z^-} = d\Phi_{edge}(m) + o(1).$$

Together we get that

$$\log Z_{G_n} - \log Z_{G_{n-2}} = 2\Phi_{vertex} - d\Phi_{edge} + o(1).$$

Writing this as a telescoping sum, and using the fact that G_{n-2k} remain locally treelike,

$$\begin{aligned} \log Z_{G_n} &= \log Z_{\alpha n} + \sum_{k=1}^{\frac{1-\alpha}{2}n} \log Z_{G_{n-2(k-1)}} - \log Z_{G_{n-2k}} \\ &= \log Z_{\alpha n} + \frac{1-\alpha}{2} n \left(2\Phi_{vertex} - d\Phi_{edge}\right) + o(n). \end{aligned}$$

The log partition function $\log Z_{\alpha n}$ is $O(\alpha n)$ as it is a sum of $|\mathcal{X}|^{\alpha n}$ terms each of which has magnitude $e^{O(\alpha n)}$. Thus taking $\alpha \to 1$ the free energy is given by

$$\Phi = \lim_n \frac{1}{n} \log Z_{G_n} = \Phi_{vertex}(m) - \frac{d}{2}\Phi_{edge}(m), \tag{5.1}$$

which is the analogue of the formula for the tree (3.2). When there is a unique Gibbs measure this formula determines the free energy. But what if there are multiple Gibbs measures, which one is the right one? In general there is no recipe to decide but the case of the Ising model is instructive [18, 29]. The BP message for the Ising model letting $y = m_{u\to v}(+) - \frac{1}{2}$ gives

$$\begin{aligned} y &= BP[m,\dots,m] - \frac{1}{2} \\ &= \frac{((\frac{1}{2}+y)e^{\beta} + (\frac{1}{2}-y)e^{-\beta})^d}{((\frac{1}{2}+y)e^{\beta} + (\frac{1}{2}-y)e^{-\beta})^d + ((\frac{1}{2}+y)e^{\beta} + (\frac{1}{2}-y)e^{-\beta})^d} - \frac{1}{2} \\ &= \frac{\frac{1}{2}(1+2y\tanh\beta)^d - \frac{1}{2}(1-2y\tanh\beta)^d}{(1+2y\tanh\beta)^d + (1-2y\tanh\beta)^d} =: f(y). \end{aligned}$$

The fixed point $y = f(y)$ has a unique fixed point when $0 \le \beta \le \tanh^{-1}\frac{1}{d-1}$. The unique fixed point is $y = 0$ so there is a unique TIFP and in fact a unique Gibbs measure. When $\beta > \tanh^{-1}\frac{1}{d-1}$ there are three fixed points, $y_-, 0$ and y_+ which each correspond to TIFPs, $\mu_- \preceq \mu_f \preceq \mu_+$. Using monotonicity argument specific to the ferromagnetic Ising model it is known [40] that the local weak limit in this case is $\frac{1}{2}\mu_- + \frac{1}{2}\mu_+$. While not extremal, the

calculations of (5.1) still hold and with $m_+ = y_+ + \frac{1}{2}$,

$$\begin{aligned}\Phi_{\beta,0} &= \lim_{h\downarrow 0} \Phi_{\beta,h} \\ &= \log\left((m_+e^{\beta} + (1-m_+)e^{-\beta})^d + (m_+e^{-\beta} + (1-m_+)e^{\beta})^d\right) \\ &\quad - \frac{d}{2}\log\left((m_+^2 + (1-m_+)^2)e^{\beta} + 2m_+(1-m_+)e^{-\beta}\right).\end{aligned}$$

For the antiferromagnetic Ising model, the picture is much less complete. When $-\beta \leq \tanh^{-1}\frac{1}{d-1}$ there is a unique Gibbs measure and equation (5.1) gives the free energy. However, at low enough temperatures the value of the free energy remains open.

In establishing the free energy for the Ferromagnetic Ising Model, the argument crucially used monotonicty in the FK-model. This does not hold for the antiferromagnetic Ising model, as it does not have a representation in the Fortuin Kasteleyn model. In fact establishing an expression for the free energy remains an open problem. Why is it different? Consider the "ground state" the state with the highest probability. In the ferromagnetic model this is simply the all plus and all minus states. In the limit at $\beta \to \infty$ most of the mass of the distribution goes to states close to the ground states. The ground states for the antiferromagnetic Ising model correspond to the partition of the vertices with the greatest number of edges between the groups. This corresponds to the random CSP called the MAX-CUT problem, which on a general graph is NP-hard to find. In the case of the random regular graph the empirical distribution of the ground state remains open.

Now consider the anti-ferromagnetic Ising model. Recall there is a unique translation invariant measure with $m_{u\to v}(+) = \frac{1}{2}$. Then

$$\mathbb{P}[x_u = x_v] = \frac{\frac{1}{2}e^{\beta}}{\frac{1}{2}e^{\beta} + \frac{1}{2}e^{-\beta}} = \frac{1}{2} + \frac{1}{2}\tanh\beta. \tag{5.2}$$

As $\beta \to -\infty$, we have $\mathbb{P}[x_u = x_v] \to 0$ so this local weak limit would correspond to an almost equal split of the vertices with very few edges between the two halves. In fact large d the asymptotics of the MAXCUT on a random regular graph were determined in [19] as

$$\lim_n \frac{1}{n}\mathrm{MAXCUT}(G) = \frac{d}{4} + c_\star d^{1/2} + o(d^{1/2})$$

for some known constant $c_\star$ derived from the Sherrington-Kirkpatrick model. For large d the maximal cut is just over half the number of edges so this fact, together with equation (5.2) establish that for large β the translation invariant Gibbs measure is not the local weak limit of the anti-ferromagnetic Ising model for large d (less precise asymptotics imply that it is not the limit for small d as well).

6. Moments

One of the most powerful methods for understanding spin systems on random graphs is calculating moments of their expected partition function. For a configuration $\underline{x} \in \mathcal{X}^V$ we will write the edge empirical density as

$$h_{\underline{x}}(x,x') = \frac{1}{dn}\sum_{u\in V}\sum_{u'\in\partial u} I(x_u = x, x_{u'} = x')$$

and the vertex empirical density

$$\overline{h}_{\underline{x}}(x) = \sum_{x'\in\mathcal{X}} h_{\underline{x}}(x,x') = \frac{1}{n}\sum_{u\in V} I(x_u = x).$$

Clear $h_{\underline{x}}$ is non-negative, symmetric and sums to 1. Let $\mathcal{H} \subset [0,1]^{\mathcal{X}^2}$ be the set of such functions. Let $\mathcal{H}_n$ be the set of such functions taking values $\frac{a}{n}$ for $a \in \mathbb{Z}$ which is the set of empirical distributions that can be achieved on a graph. Taking into account the degrees of freedom by the symmetry conditions, $\mathcal{H}_n \asymp n^{\binom{|\mathcal{X}|+1}{2}-1}$.

Let $A_h \subset \mathcal{X}^V$ be the set of configurations with edge empirical density h. Next we will calculate the expected size of A_h under the d-regular configuration model. The number of ways to assign states to the n vertices is

$$\frac{n!}{\prod_{x\in\mathcal{X}}(\overline{h}(x)n)!}.$$

Next, for each half edge we assign it a state indicating the states of the opposite vertex which can be done in

$$\prod_{x\in\mathcal{X}}\frac{(d\overline{h}(x)n)!}{\prod_{x'\in\mathcal{X}}(dh(x,x')n)!}$$

ways. When matching edges according to the configuration model, each half edge of type (x,x') must be matched with one of type (x',x). The number of such matchings is

$$\prod_{x\in\mathcal{X}}(dh(x,x)n-1)!!\prod_{x'\neq x}\sqrt{(dh(x,x')n)!},$$

where the square root comes from double counting (x,x') and (x',x). The total number of matchings is $(dn-1)!!$ so overall

$$\mathbb{E}|A_h| = \frac{1}{(dn-1)!!}\frac{n!}{\prod_{x\in\mathcal{X}}(\overline{h}(x)n)!}\prod_{x\in\mathcal{X}}\frac{(d\overline{h}(x)n)!}{\prod_{x'\in\mathcal{X}}(dh(x,x')n)!}$$
$$\cdot\prod_{x\in\mathcal{X}}(dh(x,x)n-1)!!\prod_{x'\neq x}\sqrt{(dh(x,x')n)!}. \tag{6.1}$$

By Stirling's Approximation $m! \approx \sqrt{2\pi m}e^{-m}m^m$ and

$$(2m-1)!! = \frac{(2m)!}{2^m m!} \approx \sqrt{2}e^{-m}(2m)^m.$$

Substituting this into (6.1), after counting the number of factor of $\sqrt{n}$, all the exponentials cancel out and we have

$$\begin{aligned}\mathbb{E}|A_h| &\asymp n^{-\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)}\frac{1}{(dn)^{dn/2}}\frac{n^n}{\prod_{x\in\mathcal{X}}(\overline{h}(x)n)^{\overline{h}(x)n}}\\ &\quad\cdot\prod_{x\in\mathcal{X}}\frac{(d\overline{h}(x)n)^{d\overline{h}(x)n}}{\prod_{x'\in\mathcal{X}}(dh(x,x')n)^{dh(x,x')n}}\\ &\quad\cdot\prod_{x\in\mathcal{X}}(dh(x,x)n)^{dh(x,x)n/2}\prod_{x'\neq x}(dh(x,x')n)^{dh(x,x')n/2}\\ &= n^{-\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)}\prod_{x\in\mathcal{X}}\overline{h}(x)^{(d-1)\overline{h}(x)n}\prod_{(x,x')\in\mathcal{X}^2}h(x,x')^{\frac{dn}{2}h(x,x')}\\ &= n^{-\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)}\exp\left(n\left[\frac{d}{2}H(h)-(d-1)H(\overline{h})\right]\right) \qquad (6.2)\end{aligned}$$

where H is the Shannon Entropy of a probability distribution where we view $\overline{h}$ and h as probability vectors.

For a spin system with weights $(\overline{\psi},\psi)$ the weight of a configuration $\underline{x}\in A_h$ is

$$\prod_{u\in V}\overline{\psi}(x_u)\prod_{(u,u')\in E}\psi(u,u')=\exp(n[<\overline{h},\log\overline{\psi}>+\frac{d}{2}<h,\log\psi>]).$$

We will write

$$Z_h=Z_{h,n,\overline{\psi},\psi}:=\sum_{\underline{x}\in A_h}\prod_{u\in V}\overline{\psi}(x_u)\prod_{(u,u')\in E}\psi(u,u')$$

and so

$$\mathbb{E}Z_h\asymp n^{-\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)}\exp\left(n\Phi_h\right)$$

where

$$\Phi_h=\Phi_{h,\overline{\psi},\psi}:=\frac{d}{2}H(h)-(d-1)H(\overline{h})+<\overline{h},\log\overline{\psi}>+\frac{d}{2}<h,\log\psi>$$

We call Φ_h the *empirical rate function* and will omit $\overline{\psi},\psi$ when the spin system is clear.

THEOREM 2 ([20]). *The limit of the normalized log expected partition function is given by*

$$\lim_n\frac{1}{n}\log\mathbb{E}Z_n=\sup_{h\in\mathcal{H}}\Phi_h.$$

PROOF. The function Φ_h is continuous on $\mathcal{H}$ and so achieves its supremum. Let h_* be such an argmax and let h_n be the element of $\mathcal{H}_n$ closest to h_*. Then by continuity, $|\Phi_{h_*}-\Phi_{h_n}|=o(1)$ and hence

$$\lim_n\frac{1}{n}\log\mathbb{E}Z_n\geq\lim_n\frac{1}{n}\log\mathbb{E}Z_{h_*,n}\geq\Phi_{h_*}$$

and

$$\lim_n \frac{1}{n} \log \mathbb{E} Z_n = \lim_n \frac{1}{n} \log \sum_{h \in \mathcal{H}_n} \mathbb{E} Z_{h,n}$$
$$\leq \lim_n \frac{1}{n} \log n^{\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)} \exp(n\Phi_h) = \Phi_h,$$

since $\mathcal{H}_n \asymp n^{\binom{|\mathcal{X}|+1}{2}-1}$. ■

Thus computing the expected moments is effectively reduced to finding the maxima of Φ_h. In the following lemma we see that under mild assumptions the moment is determined up to a multiplicative constant.

LEMMA 3. *If Φ_h has a unique global maxima h_* in the interior of $\mathcal{H}$ and its Hessian $D^2\Phi_h$ is negative definite at h_* then*

$$\mathbb{E} Z_n \asymp \exp(n\Phi_{h_*}).$$

PROOF. Since the Hessian is negative definite at the maxima and $\mathcal{H}$ is compact, there exist constants $c_1, c_2 > 0$ such that for all $h \in \mathcal{H}$,

$$c_1|h - h_*|^2 \leq \Phi_{h_*} - \Phi_h \leq c_2|h - h_*|^2.$$

Then since $\mathcal{H}$ is a $\binom{|\mathcal{X}|+1}{2} - 1$ dimensional subspace,

$$\begin{aligned}
\mathbb{E} Z_n &\geq \sum_{\substack{h \in \mathcal{H}_n \\ |h-h_*| \leq n^{-\frac{1}{2}}}} \mathbb{E} Z_{h,n} \\
&\geq C \sum_{\substack{h \in \mathcal{H}_n \\ |h-h_*| \leq n^{-\frac{1}{2}}}} n^{-\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)} \exp(n\Phi_h) \\
&\geq C n^{-\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)} \exp(n\Phi_{h_*} - c_2) \left| \{h \in \mathcal{H}_n : |h - h_*| \leq n^{-\frac{1}{2}}\} \right| \\
&\geq C' \exp(n\Phi_{h_*}).
\end{aligned}$$

For the upper bound

$$\begin{aligned}
\mathbb{E} Z_n &= \sum_{k \geq 0} \sum_{\substack{h \in \mathcal{H}_n \\ kn^{-\frac{1}{2}} \leq |h-h_*| < (k+1)n^{-\frac{1}{2}}}} \mathbb{E} Z_{h,n} \\
&\leq \sum_{k \geq 0} C n^{-\frac{1}{2}(\binom{|\mathcal{X}|+1}{2}-1)} \exp\left(n\Phi_{h_*} - c_1(k-1)^2\right)) \\
&\quad \times \left| \{h \in \mathcal{H}_n : |h - h_*| \leq n^{-\frac{1}{2}}\} \right| \\
&\leq O(\exp(n\Phi_{h_*})) \sum_{k \geq 0} k^{\binom{|\mathcal{X}|+1}{2}-1} \exp\left(-c_1(k-1)^2\right) \\
&= O(\exp(n\Phi_{h_*})).
\end{aligned}$$

■

6.1. Maxima and BP fixed points. Next we show that the maxima, and indeed all stationary points, of Φ_h correspond to TIFPs.

THEOREM 4. *If h is a stationary point of Φ_h then there exists a BP fixed point $m = BP[m]$ such that*

$$h(x,x') = \frac{1}{z_h} m(x)m(x')\psi(x,x') \tag{6.3}$$

where $z_h = \sum_{x,x'} m(x)m(x')\psi(x,x')$ is a normalizing constant. Conversely, any BP fixed point m corresponds to a stationary point of Φ_h.

We will give two proofs of this Theorem. The first from [21] is essentially calculus while the second is more probabilistic but will yield additional information.

PROOF. Let

$$T\mathcal{H} := \{\delta \in \mathbb{R}^{\mathcal{X}^2} : \delta(x,x') = \delta(x',x), \sum_{x,x'} \delta(x,x') = 0\},$$

be the set of vectors tangent to $\mathcal{H}$. Then h is a stationary point if and only if for all $\delta \in T\mathcal{H}$,

$$\frac{d}{dt}\Phi_{h+\delta t} \mid_{t=0} = 0.$$

The derivative is given by

$$\frac{d}{dt}\Phi_{h+\delta t} \mid_{t=0} = \frac{d}{2} < \delta, \log\psi - \log h > + < \overline{\delta}, \log\overline{\psi} + (d-1)\log\overline{h} >$$

where $\overline{\delta}(x) = \sum_{x'} \delta(x,x')$. Set $\lambda_a = \frac{1}{2}\log\frac{h(a,a)}{\psi(a,a)}$ so $h(a,a) = \psi(a,a)\exp(\lambda_a + \lambda_a)$. For $a, b \in \mathcal{X}$ let

$$\delta'(x,x') = \begin{cases} 1, & \text{if } x = x' \in \{a,b\} \\ -1, & \text{if } x = a, x' = b \text{ or } x = b, x' = a \\ 0, & \text{otherwise.} \end{cases}$$

Since $\overline{\delta}' = 0$ we have that

$$\begin{aligned} 0 = \frac{d}{dt}\Phi_{h+\delta' t} \mid_{t=0} &= \frac{d}{2} < \delta', \log\psi - \log h > \\ &= \frac{d}{2}\left(\log\frac{\psi(a,a)}{h(a,a)} + \log\frac{\psi(b,b)}{h(b,b)} - 2\log\frac{\psi(a,b)}{h(a,b)}\right) \\ &= \frac{d}{2}\left(-2\lambda_a - 2\lambda_b - 2\log\frac{\psi(a,b)}{h(a,b)}\right) \end{aligned}$$

and hence,

$$\log\frac{h(a,b)}{\psi(a,b)} = \lambda_a + \lambda_b$$

and so

$$h(a,b) = \psi(a,b)\exp(\lambda_a + \lambda_b).$$

Then

$$\log \psi(x, x') - \log h(x, x') = -\lambda_x - \lambda_{x'}$$

and hence by the symmetry of δ,

$$< \delta, \log \psi - \log h >= 2 \sum_x \sum_{x'} \delta(x, x')(-\lambda_x) = -2 < \overline{\delta}, \lambda > .$$

Substituting this into the derivative of Φ_h we have that

$$0 = \frac{d}{dt} \Phi_{h+\delta t} \mid_{t=0} = < \overline{\delta}, -d\lambda + \log \overline{\psi} + (d-1) \log \overline{h} >$$

Now let $m(x)$ be a probability vector with $m(x) = e^{\lambda_x}/z$ where $z = \sum_{x'} e^{\lambda_{x'}}$. We have that

$$\log \overline{h}(x) = \log \sum_{x'} \psi(x, x') \exp(\lambda_x + \lambda_{x'}) = \lambda_x + \log \sum_{x'} \psi(x, x') m(x')/z$$

and so

$$\begin{aligned} 0 &= \frac{d}{dt} \Phi_{h+\delta t} \mid_{t=0} \\ &= \sum_x \overline{\delta}(x) \left(-\lambda_x + \log \overline{\psi}(x) + (d-1) \log \sum_{x'} \psi(x, x') m(x')/z \right). \end{aligned}$$

This must hold for all vectors $\overline{\delta}$ which sum to 0 and hence there exists a constant C such that for all $x \in \mathcal{X}$,

$$-\lambda_x + \log \overline{\psi}(x) + (d-1) \log \sum_{x'} \psi(x, x') m(x')/z = C.$$

Taking an exponential this is equivalent to

$$m(x) \propto \overline{\psi}(x) \left(\sum_{x'} \psi(x, x') m(x') \right)^{d-1} \propto BP[m]$$

and hence $m(x)$ is a fixed point of the BP equations and

$$h(x, x') \propto \psi(x, x') m(x) m(x').$$

The converse follows similarly. ■

We now given an alternative version of the theorem using a resampling argument which additionally yields the desired behaviour of the Hessian under conditions on the BP fixed point. This resampling scheme and variants of it is at the heart of our approach to making moment calculations tractable. It is a generalization of the scheme used in [48].

First let us introduce the *annealed model* where we take the random pair (G, σ) where G is a d-regular graph and σ is a configuration with the weight

$$\nu(G, \sigma) = \frac{1}{Z^*} \prod_{u \in V} \overline{\psi}(x_u) \prod_{(u,u') \in E} \psi(u, u').$$

Then since graphs are weighted uniformly, $Z_* = (dn-1)!!\mathbb{E}Z_G$. Note that the distribution of G under $\mathbb{P}_\nu$ is the size biased distribution on d-regular random graphs.

PROPOSITION 5. *If $\hbar$ is a global maximizer of Φ_h then there exists a BP fixed point $m = BP[m]$ such that*

$$\hbar(x,x') = \frac{1}{z} m(x)m(x')\psi(x,x') \tag{6.4}$$

where $z = \sum_{x,x'} m(x)m(x')\psi(x,x')$ is a normalizing constant. If there exists $\kappa, \delta > 0$ such that

$$|BP[m'] - m'| \geq \kappa |m' - m| \tag{6.5}$$

for all m' such that $|m' - m| \leq \delta$ then $D^2\Phi$ is negative definite in a neighbourhood of $\hbar$.

PROOF. Suppose that $\hbar$ is a global maximizer of Φ_h and let $\overline{\hbar}$ be the marginal of $\hbar$. We begin by choosing (G,σ) under $\mathbb{P}_\nu[\cdot \mid h_\sigma = \hbar]$ the conditional distribution that the edge empirical distribution is $\hbar$.

Fix $\gamma > 0$ a small constant. We place some arbitrary ordering on the vertices. Given (G,σ) let $W^+ = \{v_1, \dots, v_{2\gamma n}\}$ be the set of the first $2\gamma n$ vertices in V and let W be the set of vertices in W^+ not adjacent to any other vertices in W^+. If γ is small most vertices in W^+ will be isolated and so

$$\mathbb{P}[|W| \geq \gamma n \mid h_\sigma = \hbar] \geq 1 - e^{-c\gamma n}. \tag{6.6}$$

Next, let us write

$$\xi(x,y) = \frac{1}{d\gamma n} \#\{(w \in W, v \in \partial W : \sigma_w = x, \sigma_v = y\}$$

and its marginals

$$\overline{\xi}_1(x) = \sum_y \xi(x,y), \quad \overline{\xi}_2(y) = \sum_x \xi(x,y)$$

Since we sampled W uniformly from the vertices of G, we must have that $\xi = \hbar + o(1)$ with exponentially high probability and in particular

$$\begin{aligned} \mathbb{P}[|\xi - \hbar| > \epsilon/2] &\leq \exp(-cn\gamma\epsilon^2) \\ \mathbb{P}[|\overline{\xi}_2 - \overline{\hbar}| > \epsilon/2] &\leq \exp(-cn\gamma\epsilon^2) \end{aligned} \tag{6.7}$$

We are now going to resample the pair (G,σ) as follows. First we delete the edges adjacent to W and the spins σ_W. We retain the graph $G \setminus W$ and the spins σ_{W^c} and resample the spins and edges of W according to $\mathbb{P}_\nu$ to form a new pair (G',σ'). We will define ξ' analogously for the new model. Figure 1 gives the example of a 3-colouring displaying G, $G \setminus W$ and G' respectively.

We will couple it with a new model on a graph consisting of $|W|$ disconnected d-stars, that is a set of $|W|$ vertices which we will call U, each with

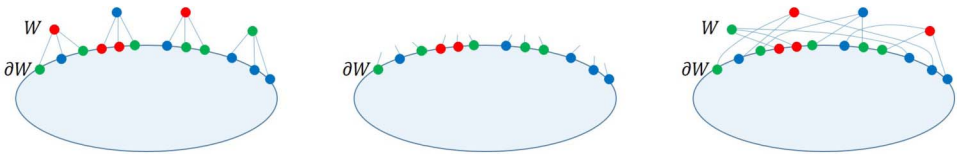


FIGURE 1. Resampling, G, $G \setminus W$ and G' respectively.

d neighbours. Define the new spin system on $U \cup \partial U$ with no external field on ∂U as

$$\tilde{\mathbb{P}}[\sigma^*_{U\cup\partial U} = \underline{x}] = \prod_{i\in U} \overline{\psi}(x_i) \prod_{j\in\partial i} \psi(x_i, x_j).$$

If the empirical distribution of $\sigma^*_{\partial U}$ is equal to $\overline{\xi}_2$ then we can construct (G', σ') in the following way. Let $f : U \to W$ be a one to one mapping from U to W. Abusing notation we will extend f to also give a bijection from the edges of $U \cup \partial U$ into the half edges of W. Assign σ'_W by setting $\sigma'_{f(u)} = \sigma^*_u$. Similarly assign states to half-edges of W according to the correspond state of σ^* on ∂U.

Having assigned colours to W, we now construct the edges connecting W to ∂W. For each $x \in \mathcal{X}$ choose a uniform matching of the half-edges of U in state x and the vertices in ∂W of state x (note that our assumption on the empirical distribution ensures these are of the same size). Connect vertices of W and ∂W according to the matchings to construct the edges of G'. Note that the number of matchings depends only on $\overline{\xi}_2$ and not the particular configuration of σ^*. The weight of σ^* under $\tilde{\mathbb{P}}$ is chosen such that this construction gives the correct conditional law for (G', σ'). This implies that

$$\tilde{\mathbb{P}}[\sigma^*_U = \underline{x} \mid \overline{h}_{\sigma^*_{\partial U}} = \tau] = \tilde{\mathbb{P}}[\sigma'_W = \underline{x} \mid \overline{\xi}_2 = \tau].$$

For most choices of τ, this will be a large deviations event which suggests tilting the measure according to make the conditional distribution likely. With this in mind for a probability measure m on $\mathcal{X}$ we define

$$\tilde{\mathbb{P}}_m[\sigma^*_{U\cup\partial U} = \underline{x}] = \prod_{i\in U} \overline{\psi}(x_i) \prod_{j\in\partial i} \psi(x_i, x_j) m(x_j).$$

Since we can write the new weight $\prod m(x_j) = \exp(dn\overline{h}_{\sigma^*_{\partial U}} \cdot (\log m))$ in terms of the empirical distribution $\overline{h}_{\sigma^*_{\partial U}}$, after conditioning on $\overline{h}_{\sigma^*_{\partial U}}$ the new weight is uniform and so

$$\tilde{\mathbb{P}}[\sigma^*_{U\cup\partial U} = \underline{x} \mid \overline{h}_{\sigma^*_{\partial U}} = \tau] = \tilde{\mathbb{P}}_m[\sigma^*_{U\cup\partial U} = \underline{x} \mid \overline{h}_{\sigma^*_{\partial U}} = \tau].$$

Next we use the fact that σ^* is a product measure over the $|W|$ stars. If $v \in U$, standard methods from large deviations theory imply that we can find an m_τ such that

$$\tilde{\mathbb{E}}_{m_\tau}[\overline{h}_{\sigma^*_{\partial v}}] = \frac{\tilde{\mathbb{E}}[\overline{h}_{\sigma^*_{\partial v}} \exp(d\overline{h}_{\sigma^*_{\partial v}} \cdot (\log m_\tau))]}{\tilde{\mathbb{E}}[\exp(d\overline{h}_{\sigma^*_{\partial v}} \cdot (\log m_\tau))]} = \tau.$$

Since the spins on the stars are identically distributed,

$$\tilde{\mathbb{E}}_{m_\tau}[\overline{h}_{\sigma^*_{\partial U}}] = \tau$$

and by the Local Central Limit Theorem,

$$\tilde{\mathbb{P}}_{m_\tau}[\overline{h}_{\sigma^*_{\partial U}} = \tau] \geq cn^{-C}$$

for some constants c, C. Now define the probability measure

$$g_m(x,y) = \frac{1}{z_m} BP[m](x)m(y)\psi(x,y)$$

where z_m is a normalizing constant. For an edge (u,v) in $U \cup \partial U$ with $u \in U, v \in U^*$, under $\mathbb{P}_m$, the distribution of σ^* is given by the belief propagation equations,

$$\mathbb{P}_m[\sigma^*_u = x, \sigma^*_v = y] = g_m(x,y).$$

For any $\epsilon > 0$, by a large deviations bound, the edge empirical distribution satisfies

$$\tilde{\mathbb{P}}_{m_\tau}[|h_{\sigma^*_{U\cup\partial U}} - g_{m_\tau}| > \epsilon] \leq \exp(-c|W|\epsilon^2).$$

Combining the above estimates,

$$\begin{aligned}
&\mathbb{P}[|\xi' - g_{m_{\overline{\xi}_2}}| > \epsilon \mid \overline{\xi}_2, |W| \geq \gamma n] \\
&\quad = \mathbb{P}_{m_{\overline{\xi}_2}}[|h_{\sigma^*_{U\cup\partial U}} - g_{m_{\overline{\xi}_2}}| > \epsilon \mid \overline{h}_{\sigma^*_{\partial U}} = \overline{\xi}_2, |W| \geq \gamma n] \\
&\quad \leq \frac{\mathbb{P}_{m_{\overline{\xi}_2}}[|h_{\sigma^*_{U\cup\partial U}} - g_{m_{\overline{\xi}_2}}| > \epsilon \mid |W| \geq \gamma n]}{\mathbb{P}_{m_{\overline{\xi}_2}}[\overline{h}_{\sigma^*_{\partial U}} = \overline{\xi}_2 \mid |W| \geq \gamma n]} \leq \exp(-c\gamma\epsilon^2 n)n^C.
\end{aligned}$$

Suppose that $\hbar = g_{m_{\overline{\hbar}}}$. Since $\hbar(x,y) = \hbar(y,x)$ this must also hold for $g_{m_{\overline{\hbar}}}$ which implies that $m_{\overline{\hbar}} = BP[m_{\overline{\hbar}}]$ and hence that $\hbar$ is of the form (6.4). Otherwise we may assume that $\|\hbar - g_{m_{\overline{\hbar}}}\| > \epsilon$. We may pick $\delta \in (0, \epsilon/3)$ small enough so that for any τ with $\|\tau - \overline{\hbar}\| \leq \delta$ we have that $\|g_{m_\tau} - g_{m_{\overline{\hbar}}}\| \leq \epsilon/3$.

Now we use our bounds on resampling. We will write $G = (G_{W^c}, G_W)$ where G_{W^c} is the graph with W removed and G_W denotes how W is connected to W^c. Then with

$$\mathcal{A}_\delta = \{|W| \geq \gamma n, |\overline{\xi}_2 - \overline{\hbar}| \leq \delta\}, \quad \mathcal{B}_\delta = \mathcal{A}_\epsilon \cap \{|\xi - \hbar| \leq \delta\}$$

by equations (6.6) and (6.7) for any fixed $\epsilon > 0$,

$$\sum_{(G,\sigma):h_\sigma = \hbar} \nu(G,\sigma) I(\mathcal{B}_\delta) \geq \frac{1}{2} \sum_{(G,\sigma):h_\sigma=\hbar} \nu(G,\sigma). \tag{6.8}$$

By definition of $\mathcal{B}_\delta$ and the triangle inequality

$$\|\xi - g_{m_{\overline{\xi}_2}}\| \geq \|\hbar - g_{m_{\overline{\hbar}}}\| - \|g_{m_{\overline{\xi}_2}} - g_{m_{\overline{\hbar}}}\| - \|\xi - \hbar\| \geq \epsilon/3$$

For fixed G_{W^c}, σ_{W^c} we have that

$$\sum_{G_W,\sigma_W} \nu((G_{W^c},G_W),(\sigma_{W^c},\sigma_W))I(\mathcal{B}_\delta)$$
$$\leq \sum_{G_W,\sigma_W} \nu((G_{W^c},G_W),(\sigma_{W^c},\sigma_W))I(\mathcal{A}_\delta)\mathbb{P}[|\xi' - g_{m_{\overline{\xi}_2}}| \geq \epsilon/3 \mid \overline{\xi}_2, |W| \geq \gamma n]$$
$$\leq \exp(-\frac{1}{9}c\gamma\epsilon^2 n)n^C \sum_{G_W,\sigma_W} \nu((G_{W^c},G_W),(\sigma_{W^c},\sigma_W))$$

Summing over all G_{W^c}, σ_{W^c} and applying (6.8) we have that for some c' and large enough n,

$$\sum_{(G,\sigma):h_\sigma=\hbar} \nu(G,\sigma) \leq \exp(-c'\epsilon^2 n) \sum_{(G,\sigma)} \nu(G,\sigma)$$

and hence

$$\Phi_\hbar - \max_h \Phi_h \leq -c'\gamma\epsilon^2. \tag{6.9}$$

Hence any global maximizer must be of the form (6.4).

Now suppose that $\hbar$ is a global maximizer which satisfies equations (6.4) and (6.5) for some belief propagation fixed point m. Suppose that h is in a neighbourhood of $\hbar$. Both m and g are smooth so there exists a constant $C_1, C_2 > 0$ such that

$$|g_{m_{\overline{h}}} - g_{m_{\overline{\hbar}}}| \leq C_1\|m_{\overline{h}} - m_{\overline{\hbar}}\|, \quad |m_{\overline{h}} - m_{\overline{\hbar}}| \leq C_2\|h - \hbar\|.$$

Since $\hbar = g_{m_{\overline{\hbar}}}$ if $\|m_{\overline{h}} - m_{\overline{\hbar}}\| \leq \frac{1}{2C_1}\|h - \hbar\|$ then

$$\|h - g_{m_{\overline{h}}}\| \geq \|h - \hbar\| - |g_{m_{\overline{h}}} - g_{m_{\overline{\hbar}}}| \geq \frac{1}{2}\|h - \hbar\|. \tag{6.10}$$

Otherwise if $\|m_{\overline{h}} - m_{\overline{\hbar}}\| \geq \frac{1}{2C_1}\|h - \hbar\|$ then by (6.5)

$$\|m_{\overline{h}} - BP[m_{\overline{h}}]\| \geq \frac{\kappa}{2C_1}\|h - \hbar\|.$$

Now we use the fact that $g_{m_{\overline{h}}}(x,y)$ is not symmetric since $m_{\overline{h}} \neq BP[m_{\overline{h}}]$ while by contrast h is symmetric. Quantifying this for some $C_3 > 0$,

$$\|h - g_{m_{\overline{h}}}\| \geq \frac{1}{2}\sum_{x,y} |g_{m_{\overline{h}}}(x,y) - g_{m_{\overline{h}}}(y,x)| \geq C_3\|m_{\overline{h}} - BP[m_{\overline{h}}]\| \geq \frac{C_3\kappa}{2C_1}\|h - \hbar\|.$$

Thus combined with (6.10) for $\theta = \frac{1}{2} \wedge \frac{C_3\kappa}{2C_1} > 0$ we have that

$$\|h - g_{m_{\overline{h}}}\| \geq \theta\|h - \hbar\|.$$

Applying the argument that lead to (6.9) we have that

$$\Phi_h - \Phi_\hbar \leq -c'\gamma\|h - g_{m_{\overline{h}}}\|^2 \leq -c'\gamma\theta^2\|h - \hbar\|^2$$

which gives the required quadratic decay of the Hessian around $\hbar$. ■

The combination of Lemma 3 and Proposition 5 suggests the following approach to determining the moments of a spin system. We will often begin by restricting

(1) Directly establish that the maximizer must lie in the interior of some set $H \subset \mathcal{H}$.
(2) Find the (hopefully) finite set of BP fixed points $\{m_1, \dots, m_k\}$ which correspond to pair empirical distributions in H.
(3) Find the m_i that maximizes Φ_h.
(4) Show that this maximizing belief propagation fixed point satisfies (6.5).

Let us apply this to the problem of counting the number of independent sets in a random d-regular graph. This is simply Z the partition function of the hardcore model with fugacity 1,

$$\mathbb{P}[\sigma] = \frac{1}{Z} \prod_{u \sim v} I(\sigma_u \sigma_v = 0).$$

Setting $q = m(1)$, the belief propagation equation satisfies

$$q = f(q) = \frac{(1-q)^{d-1}}{1 + (1-q)^{d-1}}.$$

This has a unique solution q_* and so

$$h_*(x, y) \propto m(x) m(y) I(xy = 0),$$

which has density of,

$$\alpha = \frac{(1-q_*)^{d-1}}{1 + (1-q_*)^{d-1}}$$

that is the fraction of 1's. The maximizing empirical distribution gives,

$$\Phi_{h_*} = \log(1 + (1-q)^d) - \frac{d}{2} \log(1 - q^2).$$

Since $f(q)$ is decreasing, $|f(q) - q_*| \geq |q - q_*|$ and so Φ_h is negative definite at h_*. Hence we have that

$$\mathbb{E} Z \asymp \exp(n \Phi_{h_*}).$$

This gives the exponential growth rate of the expected number of independent sets. But what if we are interested in the largest independent set and so want to calculate the number of independent sets of a particular size. We want to restrict the optimization of Φ according to Φ_h according to the independent set density $\overline{h}(1)$.

6.2. Restricted maximizers. In this section we consider optimizing Φ_h over the intersection of $\mathcal{H}$ with a hyperplane $w^T \overline{h} = \alpha$ or $w^T h = \alpha$ for some vector w. For now we will focus on the constraints of the marginal $\overline{h}$. Constrained optimization naturally suggests Lagrange Multipliers, that is

looking for stationary points of $\Phi_h + \lambda w^T \overline{h}$ on the hyperplane. We can write this as

$$\begin{aligned}\Phi_h &+ \lambda w^T \overline{h} \\ &= \frac{d}{2} H(h) - (d-1) H(\overline{h}) + < \overline{h}, \log(\overline{\psi}) + \lambda w > + \frac{d}{2} < h, \log \psi > \\ &= \Phi_{h, \overline{\psi} e^{\lambda w}, \psi}\end{aligned}$$

The addition of a Lagrange multiplier simply changed the external field of the model. So to find the constrained optimizer, we look for pairs (λ, m) such that m is a fixed point of the corresponding belief propagation equation, $BP_\lambda[m] = m$ and so that the corresponding pair empirical distribution satisfies $w^T \overline{h} = \alpha$. In practice we will look for fixed points for each λ and then find which ones satisfy $w^T \overline{h} = \alpha$.

In the case of independent sets constrained to have $h(1) = \alpha$, that is density α of independent sets, the addition of a Lagrange multiplier λ corresponds to the hardcore model with fugacity e^λ and the belief propagation equation is

$$BP_\lambda[q] = \frac{e^\lambda (1-q)^{d-1}}{1 + e^\lambda (1-q)^{d-1}}, \tag{6.11}$$

which has a unique solution q_λ. The density of the independent set is

$$\alpha_\lambda = \frac{e^\lambda (1-q_\lambda)^d}{1 + e^\lambda (1-q_\lambda)^d},$$

which is a smooth monotone function of λ and for any $\alpha \in (0, \frac{1}{2})$ there is a unique λ with $\alpha_\lambda = \alpha$. The restricted maximizer h_α corresponds to

$$\Phi_{h_\alpha} = \log(1 + e^\lambda (1-q)^d) - \frac{d}{2} \log(1-q^2) - \alpha\lambda \tag{6.12}$$

where the last term corresponds to subtracting the Lagrange multiplier. Viewing Φ_{h_α} as a function of α, it is of course maximized at α_0 as this corresponds to counting unweighted independent sets. It is a decreasing function as α increases and there is $\alpha_\square$ such that $\Phi_{h_{\alpha_\square}} = 0$. This corresponds to the density of independent sets such that the expected number is $O(1)$. Certainly this is an upper bound on the size of the maximal independent set. However, it is an overestimate. The probability under the corresponding TIFP that a vertex and all of its neighbours will be 0 is

$$\frac{1}{1 + e^\lambda (1-q)^d} > 0.$$

This means that when we sample (G, σ) with $h_\sigma = h_{\alpha_\square}$ then we expect there to be ϵn vertices which can be added to the independent set and hence there is an independent set of density $(\alpha_\square + \epsilon)n$, something we have already ruled out. With suitable concentration arguments this can be made precise. What causes the first moment to fail to predict the maximal independent set is that large independent sets come in clusters. In fact, whenever you have one

you have a cluster of exponentially many. We will come back to this point in Section 7 when we discuss clustering of solutions. But first we will move to achieving lower bounds using the second moment method.

6.3. Second moments. If the first moment of the partition function is a natural way to give upper bounds on the free energy, the second moment method is one of the simplest ways to prove a lower bound using the Paley–Zygmund inequality,

$$\mathbb{P}[Z \geq \theta \mathbb{E}Z] \geq \frac{(1-\theta)^2(\mathbb{E}Z)^2}{\mathbb{E}[Z^2]}.$$

We can write the second moment of Z as

$$\mathbb{E}Z^2 = \mathbb{E}\sum_{\underline{x},\underline{x}'} \prod_{u \in V} \psi_u(x_u)\psi_u(x'_u) \prod_{(u,v)\in E} \psi_{u,v}(x_u, x_v)\psi_{u,v}(x'_u, x'_v).$$

The second moment can be interpreted as the partition function of a spin system on G taking values in $\mathcal{X}^2$. We can think of this as two copies of the spin system and the weight of the configuration is simply the product of weights of the two configurations. To bound the second moment we can study the this spin system called the *two copy model*. Of course we could also study systems with higher moments but typically the second moment suffices.

We will write $\Phi_h^{\otimes 2}$ for the empirical rate function of the two copy model. If h is corresponds to a BP fixed point m and $h \otimes h(x, x', y, y') := h(x, y)h(x', y')$ then $h \otimes h$ is corresponds to a BP fixed point as the equation factorises and

$$\Phi_{h\otimes h}^{\otimes 2} = 2\Phi_h.$$

The following proposition follows immediately from Lemma 3 and the second moment method.

PROPOSITION 6. *If h and $h \otimes h$ are the global maxima of Φ and $\Phi^{\otimes 2}$ respectively and both have negative definite Hessian at the maxima then*

$$\mathbb{E}[Z^2] \asymp (\mathbb{E}[Z])^2 \asymp \exp(2\Phi_h n)$$

and $\liminf_n \mathbb{P}[Z > \frac{1}{2}\mathbb{E}Z] > \delta > 0$.

Even when the conditions of Proposition 6 hold in general we won't have $\frac{\mathbb{E}[Z^2]}{(\mathbb{E}[Z])^2} \to 1$ so this does not immediately imply convergence of the normalized log partition function. The reason is an extra source of fluctuations from the small cycles of the graph. In a random graph, the number of small cycles of a constant length is Poisson with constant mean. However, each such cycle will have a constant multiplicative effect on the partition function. If the spin system is at finite temperature or is permissive then by revealing the edges one at a time and applying the Azuma–Hoeffding inequality to the Doob martingale of $\log Z$ we have that

$$\mathbb{P}[|\log Z - \mathbb{E}\log Z| > t\sqrt{n}] \leq 2\exp(-ct^2), \tag{6.13}$$

and so $\mathbb{E}\log Z = \log \mathbb{E}Z + O(\sqrt{n})$. This implies that $\frac{1}{n}\log Z$ is concentrated around Φ_h which is therefore the free energy. For general constraint satisfaction problems we cannot apply this method since Z may be 0 so we should not write $\mathbb{E}\log Z$. There are several approaches to this issue. The main one is the use of the small graph conditioning method introducing in [46]. This is a way of accounting for the variations of Z in terms of the small cycles of the graph. This approach involves comparing the ratio of the first and second moments to the effect of different cycle lengths which can be a non-trivial calculation (see [28] for one approach to this). A similar approach is to directly control the Doob martingale of $\log(Z + \epsilon Z)$. This approach was taken in [24] and establishes concentration of $\log Z$ subject to certain decay of correlation conditions on the limiting Gibbs measure on the infinite tree. A final technique is to use the method of Friedgut's Theorem [26] on sharp thresholds and its extensions, a topic described in Section 10.4.

6.4. Local weak convergence. The second moment is sufficient to control the distribution in a local neighbourhood of the origin. We will show that under the conditions of Proposition 6 that the local measure of the configuration at almost all vertices is a translation invariant Gibbs measure given by the TIFP.

To prove such a convergence we will prove it for each fixed depth ℓ. The local weak limit of the random d-rgular graphs is T^d the infinite d-regular tree. Let T^d_ℓ denote the tree truncated to depth ℓ. For π, a measure on T^d, let π_ℓ denote its restriction to T^d_ℓ. For a vertex $v \in V$ let $\zeta_v = \zeta_{v,G,\ell}$ denote the random measure μ on $\mathcal{X}^{T_{d,\ell}}$ which we interpret as $\mathbb{P}[\sigma_{B_\ell(v)} \in \cdot]$, where $B_\ell(v)$ is the ball of radius ℓ around v (on the $O(1)$ set of vertices whose depth ℓ neighbourhood is not isomorphic to T^d_ℓ we can define ζ_v arbitrarily). We say that the measure of the spin system on the random d-regular graph G_n converges in probability locally to a measures π if

$$\frac{1}{n}\sum_{v\in V} d_{TV}(\zeta_{v,G_n,\ell}, \pi_\ell) \to 0$$

in probability as $n \to \infty$ for all depths ℓ. The following lemma describes the empirical local distribution in terms of the edge empirical distribution h and established that it is concentrated.

LEMMA 7. *For $h' \in \mathcal{H}$ let $\pi^{(h')}$ be the broadcast model with Markov transition matrix*

$$M_{x,y} = \frac{h'(x,y)}{\sum_{y'\in\mathcal{X}} h'(x,y')},$$

and root distribution $\overline{h}'$. If we select (G,σ) under the annealed measure with edge empirical distribution h then for any $\epsilon > 0$ there exists $c > 0$ such that

(6.14)

$$\mathbb{P}_\nu\Big[\sum_{x\in\mathcal{X}^{T_{d,\ell}}} \Big|\pi^{(h')}_\ell(x) - \frac{1}{n}\sum_v I(\sigma_{B_\ell(v)} = x)\Big| > \epsilon \mid h_\sigma = h'\Big] \le \exp(-cn).$$

PROOF. Given the edge empirical distribution h' the random pair (G, σ) can be constructed according to the recipe given in establishing equation (6.2),

- First assign the vertices pairs of spins according to $\overline{h}'$.
- For each half edge, assign the pair of spin of the other end according to the law h'.
- Choose a perfect matching given the assignment.

Consider the exploration from a random vertex v. Its spin σ_v is state $x \in \mathcal{X}$ with probability $\overline{h}'(x)$. We explore the spin of the first neighbour of v is which conditionally has state y with probability $M_{x,y} = \frac{h'(x,y)}{\overline{h}'(x)}$. If we explore the next neighbour of v the probability of the states will have shifted slightly since the pool of unexplored vertices has decreased. However, the probability of the next state changes by only $O(1/n)$. Exploring for the first ℓ levels, asymptotically the law of $\sigma_{B_\ell(v)}$ is given by the broadcast model $\pi_\ell^{(h')}$ on T_ℓ^d.

Now suppose we reveal the local neighbourhoods of a random sequence of vertices $v_1, \dots, v_{\kappa n}$ and let $\mathcal{F}_j$ be the filtration generated by the first j neighbourhoods. If we have revealed at most j spins so far, the distribution of the spins of random half-edge will have change from h' by at most $O(j/n)$. Moreover, the probability of selecting a half-edge in an already revealed vertex will be at most $O(j/n)$. Thus at step j the distribution of $\sigma_{B_\ell(v_j)}$ will still be conditionally close to $\pi_\ell^{(h')}$ and

$$d_{TV}\left(\mathbb{P}_\nu(\sigma_{B_\ell(v_j)} = x \mid h_\sigma = h', \mathcal{F}_{j-1}), \pi_\ell^{(h')}(x)\right) \le \frac{C_\ell j}{n}.$$

Since the conditional law of $\sigma_{B_\ell(v_j)}$ is close in total variation distance to $\pi_\ell^{(h')}$ by the Azuma-Hoeffding,

$$\mathbb{P}_\nu\Big[\sum_{x \in \mathcal{X}^{T_{d,\ell}}} \Big|\pi_\ell^{(h')}(x) - \frac{1}{\kappa n}\sum_{i=1}^{\kappa n} I(\sigma_{B_\ell(v_i)} = x)\Big| > C_\ell(\kappa + \delta) \mid h_\sigma = h'\Big]$$

$$\le \exp(-c\delta^2 \kappa n). \tag{6.15}$$

Furthermore since the vertices were sampled uniformly at random by Azuma-Hoeffding for standard sampling without replacement,

$$\mathbb{P}_\nu\Big[\sum_{x \in \mathcal{X}^{T_{d,\ell}}} \Big|\frac{1}{n}\sum_{v \in V} I(\sigma_{B_\ell(v)} = x) - \frac{1}{\kappa n}\sum_{i=1}^{\kappa n} I(\sigma_{B_\ell(v_i)} = x)\Big| > \delta \mid h_\sigma = h'\Big]$$

$$\le \exp(-c\delta^2 \kappa n). \tag{6.16}$$

Thus combining equations (6.15) and (6.16) and taking κ, δ to be sufficiently small we have establish (6.14) which completes the proof of the lemma. ∎

We now prove that when the second moment method succeeds we also have local weak convergence. If there is a unique global maxima $\hbar$ of Φ_h then

by Proposition 5 it corresponds to a BP fixed point m and the broadcast model $\pi^{(\hbar)}$ is the corresponding translation invariant Gibbs measure.

THEOREM 8. *For a permissive spin system ψ on a random d-regular graph, if $\hbar$ and $\hbar\otimes\hbar$ are the unique global maxima of Φ and $\Phi^{\otimes 2}$ respectively and both have negative definite Hessian at the maxima then (G,σ) converges in probability locally to the translation invariant Gibbs measure $\pi^{(\hbar)}$.*

PROOF. By Proposition 6 and equation (6.13) we have that for any $\frac{1}{2} < r < 1$,

$$\mathbb{P}[Z_G > e^{n\Phi_\hbar - n^r}] \to 1.$$

Recall that for two probability measures η, η' on a discrete set A that the L^2 and total variation distances satisfy

$$d_{TV}(\eta,\eta') \le \|\eta - \eta'\|_{L^2(\eta)}$$

where

$$\|\eta - \eta'\|^2_{L^2(\eta)} = \sum_{x\in A} |\frac{\eta'(x)}{\eta(x)} - 1|^2 \eta(x) = \sum_{x\in A} \frac{(\eta'(x))^2}{\eta(x)} - 1$$

Define $\mathcal{A}_\delta$ to be the set of graphs such that

$$\mathcal{A}_\delta = \{G : \delta \le \frac{1}{n}\sum_{v\in V} I(d_{TV}(\zeta_{v,G,\ell}, \pi^{(\hbar)}_\ell) \ge \delta)\} \cap \{G : Z_G > e^{n\Phi_\hbar - n^r}\}$$

To establish the desired convergence it is sufficient to prove that $\mathbb{P}[G \in \mathcal{A}_\delta] \to 0$ for all $\delta > 0$. If σ, σ' are two independent samples from the spin system on G then define

$$\psi(G,\sigma,\sigma') = \frac{1}{n}\sum_{v\in V}\sum_{x\in \mathcal{X}^{T^d_\ell}} \frac{I(\sigma_{B_\ell(v)} = \sigma'_{B_\ell(v)} = x)}{\pi^{(\hbar)}_\ell(x)} - 1$$

Then

$$\mathbb{E}_G[\psi(G,\sigma,\sigma')] = \frac{1}{n}\sum_{v\in V}\sum_{x\in \mathcal{X}^{T^d_\ell}} \frac{(\zeta_{v,G,\ell}(x))^2}{\pi^{(\hbar)}_\ell(x)} - 1 = \frac{1}{n}\sum_{v\in V} \|\zeta_{v,G,\ell} - \pi^{(\hbar)}_\ell\|^2_{L^2(\pi^{(\hbar)}_\ell)}$$

Now if $G \in \mathcal{A}_\delta$ then

$$\begin{aligned}\mathbb{E}_G[\psi(G,\sigma,\sigma')] &= \frac{1}{n}\sum_{v\in V} \|\eta_{v,g,\ell} - \pi^{(\hbar)}_\ell\|^2_{L^2(\pi^{(\hbar)}_\ell)} \\ &\ge \frac{1}{n}\sum_{v\in V} (d_{TV}(\eta_{v,G,\ell}, \pi^{(\hbar)}_\ell))^2 \ge \delta^3\,,\end{aligned}$$

and so

$$\mathbb{P}_G[\psi(G,\sigma,\sigma') \ge \frac{1}{2}\delta^3] \ge \frac{1}{2}\delta^3.$$

So it will be enough to prove that pairs (σ,σ') with ψ at least $\frac{1}{2}\delta^3$ are very rare. In order to show this we will apply Lemma 7 to the two copy model. Edge empirical distributions h' that are far from $\hbar\otimes\hbar$ such that

$|h' - \hbar \otimes \hbar| \geq \epsilon'$ are rare and indeed we have that $\Phi_{h'}^{\otimes 2} < 2\Phi_\hbar - c(\epsilon')^2$ and hence

$$\mathbb{P}[|h_{\sigma,\sigma'} - \hbar \otimes \hbar| \geq \epsilon', Z_G > e^{n\Phi_\hbar - n^r}] \leq \frac{n^{O(1)} e^{n(2\Phi_\hbar - c(\epsilon')^2)}}{(e^{n\Phi_\hbar - n^r})^2} \to 0\,,$$

for any fixed $\epsilon' > 0$. Hence we may focus on empirical distributions close to $\hbar \otimes \hbar$. Write $\pi^{\otimes 2,(h')}$ for the broadcast model on $\mathcal{X}^2$ with edge empirical distribution h'. If the depth ℓ empirical distribution of (σ, σ') is $\pi_\ell^{\otimes 2,(\hbar\otimes\hbar)} = \pi_\ell^{(\hbar)} \otimes \pi_\ell^{(\hbar)}$ then

$$\begin{aligned}\psi(G,\sigma,\sigma') &= \frac{1}{n}\sum_{v\in V}\sum_{x\in\mathcal{X}^{T_\ell^d}} \frac{I(\sigma_{B_\ell(v)} = \sigma'_{B_\ell(v)} = x)}{\pi_\ell^{(\hbar)}(x)} - 1 \\ &= \sum_{x\in\mathcal{X}^{T_\ell^d}} \frac{(\pi_\ell^{(\hbar)}(x))^2}{\pi_\ell^{(\hbar)}(x)} - 1 = 0\,.\end{aligned}$$

We may choose $\epsilon' > 0$ small enough such that if $|h' - \hbar \otimes \hbar| < \epsilon'$ then

$$\sum_{x\in\mathcal{X}^{T_\ell^d}} \left| \frac{\pi_\ell^{\otimes 2(h')}(x,x)}{\pi_\ell^{(\hbar)}(x)} - \frac{(\pi_\ell^{(\hbar)}(x))^2}{\pi_\ell^{(\hbar)}(x)} \right| \leq \frac{1}{10}\delta^3.$$

Thus equation (6.14) implies that for some $c > 0$ and for all h' such that $|h' - \hbar \otimes \hbar| < \epsilon'$ we have that

$$\begin{aligned}\mathbb{P}_\nu\Big[\sum_{x\in\mathcal{X}^{T_{d,\ell}}} \Big| \frac{(\pi_\ell^{(\hbar)}(x))^2}{\pi_\ell^{(\hbar)}(x)} - \frac{1}{n}\sum_v \frac{I((\sigma_{B_\ell(v)}, \sigma'_{B_\ell(v)}) = (x,x))}{\pi_\ell^{(\hbar)}(x)} \Big| > \frac{1}{4}\delta^3 \mid h_{\sigma,\sigma'} = h' \Big] \\ \leq \exp(-cn)\,,\end{aligned}$$

and hence

$$\mathbb{P}_\nu\Big[\psi(G,\sigma,\sigma') > \frac{1}{4}\delta^3 \mid h_\sigma = h'\Big] \leq \exp(-cn)\,.$$

If we write

$$\begin{aligned}Y_G = \sum_{\sigma,\sigma'}\prod_{u\in V} \psi_u(\sigma_u)\psi_u(\sigma'_u) \prod_{(u,v)\in E} \psi_{u,v}(\sigma_u,\sigma_v)\psi_{u,v}(\sigma'_u,\sigma'_v) \\ \cdot I(\psi(G,\sigma,\sigma') \geq \frac{1}{2}\delta^3)\end{aligned}$$

then we have that

$$\mathbb{E}Y_G \leq e^{-cn} e^{2n\Phi_\hbar}.$$

Now if $G \in \mathcal{A}_\delta$ then

$$Y_G > \frac{1}{2}\delta^3 Z_G^2 \geq \frac{1}{2}\delta^3 e^{2n\Phi_\hbar - 2n^r}$$

and by Markov's inequality

$$\mathbb{P}[Y_G > \frac{1}{2}\delta^3 e^{2n\Phi_\hbar - 2n^r}] \leq \frac{e^{-cn}e^{2n\Phi_\hbar}}{\frac{1}{2}\delta^3 e^{2n\Phi_\hbar - 2n^r}} \to 0$$

and so $\mathbb{P}[G \in \mathcal{A}_\delta] \to 0$ which completes the proof. ∎

6.5. Independent sets. Let us move back to the question of understanding the large independent sets of random d-regular graphs. It has long been known [9, 27, 34, 51] that for large d the size of the largest independent set is approximately $\frac{(2+o_d(1))\log d}{d}$ where the $o_d(1)$ tends to 0 as $d \to \infty$. To see this suppose we take $e^\lambda = d^\beta$ then the BP fixed point is

$$q = BP_\lambda[q] = \frac{d^\beta(1-q)^d}{d}$$

in which case $q_\lambda = \frac{(1+\beta+o_d(1))\log d}{d}$ and $\alpha = (1-q)q = \frac{(1+\beta+o_d(1))\log d}{d}$. Applying equation (6.12) we have that

$$\begin{aligned}\Phi_{h_\alpha} &= \log(1+e^\lambda(1-q)^d) - \frac{d}{2}\log(1-q^2) - \alpha\lambda \\ &= (1-\beta^2+o_d(1))\frac{\log^2 d}{2d}.\end{aligned}$$

This is of course maximized when $\beta = 0$ which says that the typical independent set is of order $\frac{(1+o_d(1))\log d}{d}$ and Φ_{h_α} is positive provided $|\beta| \le 1$ so the largest α that gives a nonnegative empirical rate function is $\alpha_\square = \frac{(2+o_d(1))\log d}{d}$ which provides an upper bound on the size of the maximal independent set.

To find a lower bound we need to bound the second moment. The BP equation is

$$BP[q](x,x') = \frac{A_{xx'}}{A_{00}+A10+A_{01}+A_{11}}$$

$$A_{00}=1, A_{10}=e^{\lambda_1}(q_{01}+q_{00})^{d-1}, A_{01}=e^{\lambda_2}(q_{10}+q_{00})^{d-1}, A_{11}=e^{\lambda_1+\lambda_2}q_{00}^{d-1}.$$

The two Lagrange multiplies λ_1, λ_2 are chosen to fix the density of each independent set to be α. It is easy to check that for any such fixed point there is a symmetry between the two copies so that $\lambda_1 = \lambda_2$ and $q_{01}+q_{10}$. For $\beta > 0$, large d and $\alpha = \frac{(1+\beta)\log d}{d}$ there are two fixed points. One is simply the tensor product of the one copy fixed point with Lagrange multipliers $e^{\lambda_i} \approx d^\beta$ the same as the one copy multiplier. We call this the *independent copies* fixed point and have that

$$\Phi^{\otimes 2}_{h_\alpha \otimes h_\alpha} = 2\Phi_{h_\alpha} = 2(1-\beta^2+o_d(1))\frac{\log^2 d}{2d}$$

The second fixed point $h^{=}_\alpha$ corresponds to two almost identical independent sets, with Lagrange multiplier $e^{\lambda_i} \approx d^{\beta/2}$. This is, roughly speaking, one independent set of size αn together with a small perturbation of it in the second copy. Furthermore,

$$\Phi^{\otimes 2}_{h^{=}_\alpha} = (1+o_d(1))\Phi_{h_\alpha} = (1-\beta^2+o_d(1))\frac{\log^2 d}{2d}$$

so for fixed $\beta < 1$ and large enough d the independent copies fixed point gives the global maxima. Thus we can apply the second moment method to count the number of independent sets of size $\frac{(1+\beta)\log d}{d}$. The Azuma–Hoeffding

Inequality implies that the maximal independent set is concentrated around its expected value so with high probability it is at least $\frac{(1+\beta)\log d}{d}$ giving the asymptotics of $\frac{(2+o_d(1))\log d}{d}$.

6.6. Non-regular models. In this section we discuss moments of non-regular spin systems, for example on Erdos Renyi random graphs, and the challenges they pose. First we should distinguish between two models, the standard Erdos Renyi model $G(n, d/n)$ where all edges are independent and an alternative $G(n, m)$ where m edges are chosen uniformly at random. The same average degree is achieved if we take $m = \frac{d}{2}n$. When considering moments of CSPs the latter is the right formulation to take because it removes the effects of fluctuations of the number of edges. In most CSPs, each edge will decrease the partition function by a constant factor so the $\sqrt{n}$ fluctuations of the number of edges mean that Z cannot be concentrated around its mean. To illustrate this point consider the expected number of q-colourings under $G(n, d/n)$

$$\mathbb{E}Z = \sum_{\underline{p}} \frac{n!}{\prod_{i=1}^{q}(p_i n)!}(1-\frac{d}{n})^{\sum_i \binom{np_i}{2}}$$
$$= \exp\left(n(\max_{\underline{p}} H(\underline{p}) - \frac{d}{2}\sum p_i^2 + o(1))\right) = \exp\left(n(\log q - \frac{d}{2q} + o(1))\right)$$

where the p_i denote the empirical distribution of the configuration. Both terms in the maximization are maximized when $p_i \equiv \frac{1}{q}$. Suppose instead we did the same calculation with $G(n, m)$ with $m = \frac{dn}{2}$. Then

$$\mathbb{E}Z = \sum_{\underline{p}} \frac{n!}{\prod_{i=1}^{q}(p_i n)!} \frac{\binom{\binom{n}{2}-\sum_i \binom{np_i}{2}}{m}}{\binom{\binom{n}{2}}{m}}$$
$$= \exp\left(n(\log q + \frac{d}{2}\log(1-\frac{1}{q}) + o(1))\right)$$

Since $\log(1-\frac{1}{q}) < -\frac{1}{q}$ the exponent is strictly less when the number of edges is conditioned on. The reason is that in $G(n, d/n)$, while graphs with fewer than $\frac{dn}{2}$ edges are rare, they have many more colourings and make the dominant contribution to the expectation. Graphs with ϵn fewer edges will have, in expectation, a factor of $(\frac{q}{q-1})^{\epsilon n}$ more colourings while the probability of such graphs is $\exp(-c\epsilon^2 n)$. Because of the ϵ^2 term in the exponent, we see that for small ϵ graphs with fewer edges make a larger contribution to the expected value than typical graphs. Thus to understand a CSP on an Erdos-Renyi graph it is, therefore, better to condition on the number of edges.

This effect of fluctuations of the partition function due to the number of edges is the simplest example of *neighbourhood fluctuations*. In a graph we can look at the empirical distribution of the depth k neighbourhoods, of vertices, the simplest being just the degree distribution of the graph.

Neighbourhood fluctuations describes when rare graphs with atypical neighbourhood empirical distributions make the dominant contribution to the expected partition function. This can occur even after conditioning on the number of edges. We will see contrasting behaviour between the q-colouring model and the hardcore model on $G(n,m)$.

We have already calculated the first moment of the partition function of the q-state colouring model on $G(n,m)$. The second moment, which involves a lot more work, was studied in [2] where they showed that $\mathbb{E}Z^2 \le C(\mathbb{E}Z)^2$ for $d \le (2-o_q(1))q\log q$ which is quite close to the colouring threshold. This implies that choosing a sized biased random graph, weighted by the number of colourings does not have a large effect on the distribution over graphs and thus the model does not exhibit neighbourhood fluctuations.

The picture is different in the hardcore model. The partition function is

$$\begin{aligned}\mathbb{E}Z &= \sum_{\alpha} \binom{n}{\alpha n} \frac{\binom{\binom{n}{2}-\binom{\alpha n}{2}}{m}}{\binom{\binom{n}{2}}{m}} \\ &= \exp\left(n(\max_{\alpha} H(\alpha) + \frac{d}{2}\log(1-\alpha^2) + o(1))\right)\end{aligned}$$

With $\alpha_\star$ chosen to maximize the exponent, the graphs which make the dominant contribution to the expected partition function can be described as follows. Choose an independent set $I \subset V$ of $\alpha_\star n$ vertices and then throw m edges such that no edge joins vertices of I. The average degree of a vertex in I will be $d\frac{1-\alpha}{1-\alpha^2}$ while the average degree in I^c is $d\frac{1}{1-\alpha^2}$. Asymptotically, the degree distribution of such a graph will be a mixture of a Poisson distributions with means $d\frac{1-\alpha}{1-\alpha^2}$ and $d\frac{1}{1-\alpha^2}$ which is of course different from a Poisson of mean d which we expect to see in $G(n,m)$. Thus, unlike q-colourings, the hardcore model is affected by neighbourhood fluctuations.

To understand the difference suppose we write Z_i for the partition function after adding i edges so

$$Z = Z_m = |\mathcal{X}|^n \prod_{i=1}^{m} \frac{Z_i}{Z_{i-1}}.$$

The ratios $\frac{Z_i}{Z_{i-1}}$ are not IID but we should expect at least the last $o(n)$ to be approximately IID. A product of IID random variables is concentrated around its mean only when the variables are constant. Thus we want that for each edge the ratio $\frac{Z_i}{Z_{i-1}}$ to be essentially deterministic, that is adding each edge changes the number of solutions by the same factor. In the case of colourings this holds because the marginal distribution at each vertex is uniform and two randomly chosen vertices will be asymptotically independent. In the case of the hardcore model though the marginal at different vertices varies. For isolated vertices of degree 0, the marginal of 1 is $\frac{1}{2}$ while for vertices of high degree the marginal is much smaller. This inhomogeneity of marginals is what causes neighbourhood fluctuations.

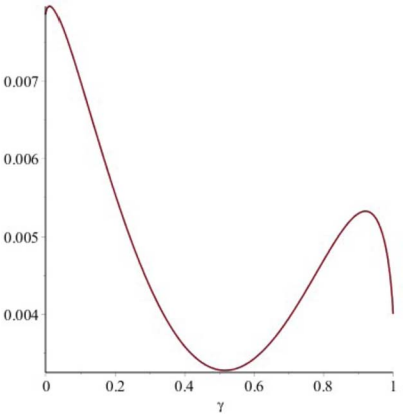


FIGURE 2. Rate function of pairs of independent sets of density $\alpha = \frac{1.5 \log d}{d}$ and intersection $\gamma\alpha$ when $d = 1000$.

In models that have neighbourhood fluctuations, a number of techniques have been employed including adding weights to the partition function [4] so that all marginals remain the same. Another approach is restricting the sum to a smaller set of spins configurations in such as way that the graph is not distorted [14]. In Section 10 we discuss how this is achieved in the random K-SAT model.

7. Clustering

A key part of the theory of random constraint satisfaction problems is understanding the way in which the space of solutions breaks into clusters. Models in the 1RSB universality class such as colourings, K-SAT and large independent sets split into clusters in a simple way. In particular there is no further structure, for example there are no clusters of clusters as is found for instance in the SK-model.

To see the existence of clusters, we can look at the expected number of pairs of independent sets of size $\alpha = \frac{1.5 \log d}{d}$ according to the density of their overlap. In Figure 2 we plot the empirical rate function for the two copy model with $d = 1000$ where the density of $(1, 1)$ vertices is $\gamma\alpha$. Note the two peaks correspond to independent copies and almost identical copies.

This observation, combined with the second moment method, was used by Achlioptas and Coja–Oghlan [1] to prove the existence of a clustering regime for colourings of random graphs, random k-sat and random NAE–SAT. The idea is to consider the *planted model*, that is where first a "planted" solution σ is chosen and then the graph or CSP is chosen conditionally on σ being a solution. This does not give a uniformly chosen graph, the graph is size biased according to the number of solutions. But when the second moment method succeeds, the expected number of

solutions and the typical number are of the same order so size biasing has only a mild effect on probabilities. It is then sufficient to prove that there are no solutions at an intermediate distance from the planted solution.

The existence of clustering is often seen as an explication of the apparent computational hardness of finding solutions. In the case of independent sets, clustering begins at density $\frac{(1+o_d(1))\log d}{d}$ while no algorithm [13,31] is known to find sets of size $\frac{(1+\epsilon)\log d}{d}$. Similar bottlenecks exists for colourings [3] and random k-SAT [17]. Intuitively, local search algorithms such as MCMC will perform poorly when the landscape is clustered and has many local maxima, many of which are far from the global maxima. However, showing actual computational hardness in this regime remains elusive.

7.1. Cluster as BP fixed points. The simplest way to formally define the meaning of a cluster is in terms of connected components of solutions where two solutions are adjacent if they differ in one (or perhaps a constant) number of variables. While the argument of [1,5] can prove clustering with this definition, it does not give a particularly tractable description of the cluster. A powerful alternate way is to view clusters in terms of fixed points of BP equations. Suppose we start from a solution σ of the CSP and set an initial message $m^{(0)}_{u\to v}(x) = I(\sigma_u = x)$. With this initial message we then iterate the BP equations letting

$$m^{(t+1)}_{u\to v} = BP[\{m^{(t)}_{u\to v}(x)\}].$$

If these converge to a limit $m_{u\to v} = \lim_t m^{(t)}_{u\to v}$ then this is a fixed point of the belief propagation equations on the graph. It turns out that such fixed points are essentially in one to one correspondence with the set of clusters of solutions. Studying such fixed points has been an essential part of the development of the theory from statistical physics as it is amenable to both numerical simulations and theoretical analysis.

Considered as a spin system the BP messages obey a certain rigidity property. When the graph is a tree, the BP messages outside of a finite set completely determine those inside of it. Indeed on an infinite tree, it is not possible to change a finite number of the messages and arrive at another BP fixed point. On a graph this corresponds to BP fixed points being far apart, generally at least a linear fraction of the messages must be changed to move from one BP fixed point to another. In effect, the BP messages are clusters of solutions mapped to a single point and these points are well separated from each other. For this reason the second moment method can succeed up to the satisfiability threshold when applied to BP messages.

When a cluster is described by a BP fixed point in this way we should expect that the messages $m_{u\to v}$ should describe the measure restricted to the cluster. Locally it should behave as an extremal Gibbs measure on a tree

where one expects decay of correlation. In such a setting we might expect that weight the configurations in cluster m is given by the Bethe Prediction of equation (3.2)

$$\log w(m) = \sum_{v \in V} \Phi_v(m) - \sum_{(u,v) \in E} \Phi_{(u,v)}(m) \tag{7.1}$$

where

$$\Phi_v(m) = \log \sum_{x \in \mathcal{X}} \psi(x) \prod_{i=1}^{d} \left(\sum_{x_{v_i} \in \mathcal{X}} \psi(x_\rho, x_{v_i}) m_{v_i \to v}(x_{v_i}) \right)$$

$$\Phi_{(v,u)}(m) = \log \sum_{x,x' \in \mathcal{X}} \psi(x,x') m_{u \to v}(x) m_{v \to u}(x').$$

We can also think of the BP fixed point itself as a solution to a random CSP. Now the variables are pairs of messages $(m_{u \to v}, m_{v \to u}) \in \mathcal{P}(\mathcal{X})^2$ on the edges of the graph. The vertices become the constraints, that at each vertex v, all the outgoing messages are given by the BP equations,

$$\forall u \in \partial v : m_{v \to u} = BP[\{m_{u' \to v}\}_{u' \in \partial v \setminus u}]$$

We will call this the *cluster model*. Viewing the clusters themselves as a random CSP means we can apply tools such as the moment method directly to counting clusters which will be a crucial step in understanding these models. Furthermore, when we come to consider the free energy of the model, it will be useful to weight these clusters according to their size and equation (7.1) gives a local systems of interactions whose sum is the predicted log weight of the cluster.

As the states of the cluster model are themselves probability measures, the belief propagation equation for the cluster model involves measures over these measures or $s \in \mathcal{P}(\mathcal{P}(\mathcal{X}))$. Let us for a moment ignore two issues, first that m lies in a continuous space and secondly the fact that each edge has two messages and consider only the messages going up the tree. Formally then the belief propagation on the cluster model with uniform weights should be

$$s(A) = \frac{1}{z} \int I(BP[\{m_i\}] \in A) I(\Phi_v(\{m_i\}) > 0) \prod_{i=1}^{d-1} s_i(dm_i)$$

where $A \subset \mathcal{P}(\mathcal{X})$ and the second indicator checks that there is a valid output for the BP equations for the input $\{m_i\}$. If instead we were considering the cluster model weighted according to the Gibbs weight $w(m)$ then the recursion would be

$$s(A) = \frac{1}{z} \int I(BP[\{m_i\}] \in A) \Phi_v(\{m_i\}) \prod_{i=1}^{d-1} s_i(dm_i).$$

Of course we have traded a CSP with spins taking a discrete set of values for one taking values as messages living in the simplex. We will see, however, that in many cases the messages can be restricted to a discrete set.

7.2. Freezing transition. We say a variable in a cluster is frozen if it always takes the same value and that the cluster is frozen if it has a linear number of the frozen variables. This, of course, can only occur in models at zero temperature, and occurs in the colouring, K-SAT and NAESAT models. In Section 4.1 we saw that in the q-colouring model, when the degree is high enough then most BP messages will be point masses. Indeed, on the tree most vertices will be frozen and just a small fraction of the vertices will be free, those that could take more than one colour. The spins of the free variables can interact with each other but they form finite, highly subcritical trees.

In the colouring models on a random graph, the picture is the same. When $d = \gamma q \log q$ for $\gamma \in (1, 2)$ and q is large, the space of solutions splits into clusters, in which most of the vertices are frozen and the free variables are all in small connected components. Molloy [39] showed that the location of this freezing transition exactly corresponds to the freezing threshold on the tree.

The analysis to determine the freezing threshold is simpler than the reconstruction threshold since we really only need to know if a message is a point mass on some state. We introduce a state $\star$ which will mean that the state message is not a point mass and write $\mathcal{X}_\star = \mathcal{X} \cup \{\star\}$. For a belief propagation message $m_{u\to v} \in \mathcal{P}(\mathcal{X})$ we will write the corresponding *Warning Propagation* message $r_{u\to v} \in \mathcal{X}_\star$ defined such that

$$r_{u\to v} = \begin{cases} x, & \text{if } m_{u\to v}(x) = 1 \\ \star, & \text{otherwise if } \max_{x'} m_{u\to v}(x') < 1. \end{cases}$$

For the CSPs such as colouring and K-SAT after the freezing threshold, the warning propagation messages satisfy a Warning Propagation equation

$$r_{v\to u} = WP[\{r_{u'\to v}\}_{u'\in\partial v\setminus u}],$$

and is undefined if there is no state at v compatible with the messages. This now provides us with a discrete spin system to analyse and apply the moment method.

Survey Propagation [10] is defined as belief propagation equations applied to the warning propagation messages. A message is then $s_{u\to v} \in \mathcal{P}(\mathcal{X}_\star)$ and satisfies

$$\begin{aligned} s_{u\to v}(r) &= SP[\{s_{u_i\to u}\}](r) \\ &= \frac{1}{z} \sum_{r_1,\dots,r_d} I(WP[\{r_i\}] = r) \prod_i s_{u_i\to u}(r_i) \end{aligned} \tag{7.2}$$

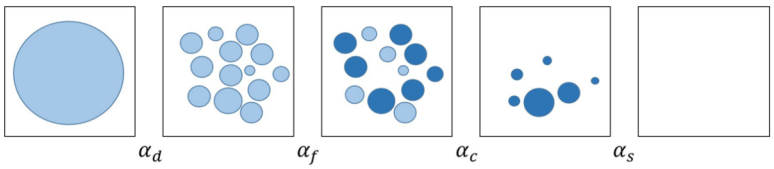


FIGURE 3. The 1RSB Phase transitions. Dark colour represents frozen clusters.

In the case of q-colouring this corresponds to for $r \in \mathcal{X}$,

$$s_{u\to v}(r) = \frac{1}{z} \sum_{r_1,\dots,r_d} I(r \notin \{r_i\}, |\{r_i\} \cap \mathcal{X}| = q-1) \prod_i s_{u_i\to u}(r_i),$$

$$s_{u\to v}(\star) = \frac{1}{z} \sum_{r_1,\dots,r_d} I(|\{r_i\} \cap \mathcal{X}| \leq q-2) \prod_i s_{u_i\to u}(r_i).$$

This fixed point can be substantially simplified by making the natural assumption that each of the colours is equally likely which reduces it to a recursion in a single real variable.

8. Predicted phase transitions

Having established a description of the clusters, the BP equations and the moment method we are now ready to describe and explain the series of phase transitions predicted by statistical mechanics [32]. In Figure 3 we see a progression of blobs which represent the geometry of the space of clusters as α the density of constraints increases in a 1RSB model. In a random graph, α can be thought of as the degree or average degree of the vertices. The transitions are reconstruction α_d, freezing α_f, condensation α_c and satisfiability α_s.

The ordering of the transitions presented here is expected for q-colourings, random k-SAT and k-NAESAT when q and k are large. For smaller values things may change for instance for K-SAT the freezing and condensation transitions are believed to interchange for $k < 8$.

8.1. Clustering and freezing. For small densities of constraints, it is expected that all, or almost all of the solutions lie inside a single cluster. At the clustering threshold, also sometimes the shattering threshold, the space of solutions splits into exponentially many clusters, each with an exponentially small fraction of the total solutions. The predicated location of the clustering transition is given simply by the reconstruction threshold of the Gibbs measure of the corresponding TIFP. When there is non-reconstruction, the BP messages simply converge to the constant fixed point and so do not indicate clustering. The freezing threshold α_f corresponds to the point at which there is freezing in the reconstrution problem, that is the root is reconstructed exactly with positive probability.

In the case of the q-colouring model on the tree these thresholds correspond to [47, 50]

$$\alpha_d \geq q(\log q + \log\log q + 1 - \log 2 + o_q(1))$$
$$\alpha_d \leq q(\log q + \log\log q + 1 - \epsilon + o_q(1))$$
$$\alpha_f = q(\log q + \log\log q + 1 + o_q(1))$$

for some $\epsilon > 0$ and where $o_q(1) \to 0$ as $q \to \infty$. In particular, at least for large q, the two thresholds are distinct [50]. The existence of clustering in the colouring model was proven for $\alpha > (1 + o_q(1))q \log q$ by [1] while later Molloy [39] established clustering and freezing down to the freezing threshold. It remains open in the colouring model, and the other models we consider here, whether the clustering threshold is indeed at the tree reconstruction threshold.

The d in α_d denotes a dynamical phase transition, roughly speaking that the Glauber dynamics or other Markov chains will stop mixing at this transition due to the clustering creating bottlenecks for the chain. This is in contrast to a phase transition involving the free energy which is expected to be analytic throughout the clustered and frozen regimes.

As random CSPs emerged from computer science, the algorithmic question of when solutions can be efficiently found has received a great deal of attention. For colouring, quite simple algorithms [3] are known find colourings only up to $\alpha \leq (1 + o_q(1))q \log q$ but no algorithms are known for $\alpha > (1+\epsilon)q \log q$. The onset of clustering clearly poses a barrier for MCMC and other local algorithms which may get stuck in local maxima with no nearby solutions. The freezing threshold poses even greater difficulties for algorithms, for instance in an algorithm that assigns the variables one by one, a single bad choice for one variable in the middle of the algorithm may cause the algorithm to fail. There is no strong evidence either way as to which, if either, of these thresholds is responsible for computational hardness but more experts favour the freezing threshold.

8.2. Condensation. A later predicted threshold, much closer to the satisfiability threshold is the condensation threshold α_c at which point the free energy has a non-analyticity. Between the condensation and satisfiability thresholds is the condensation regime in which solutions still appear in clusters but most of the solutions lie in a constant number of clusters. A more refined prediction is that the relative cluster sizes obey a Poisson–Dirichlet distribution. We'll explain the heuristics from which these predictions are arrived upon.

When α is small, the free energy is given by

$$\Phi_\alpha = \sup_{h \in \mathcal{H}} \Phi_{h,\alpha}.$$

This formula is expected to be analytic and to hold for α both before and after the clustering threshold, up to α_c. We can also heuristically compute the weight of a single cluster chosen according to the translation invariant

Gibbs measure with empirical distribution $h = \text{argmax}\Phi_{h,\alpha}$. Suppose that $\{r_{v\to v'}\}$ is the random BP fixed point constructed from the Gibbs measure. Then by equation (7.1) the predicted cluster size is $\exp(n\Psi + o(n))$ where

$$\Psi_\alpha = \mathbb{E}\log\sum_{x\in\mathcal{X}}\psi(x)\prod_{i=1}^{d}\left(\sum_{x_{v_i}\in\mathcal{X}}\psi(x_\rho,x_{v_i})r_{v_i\to v}(x_{v_i})\right) - \frac{\alpha}{2}\mathbb{E}\log\sum_{x,x'\in\mathcal{X}}\psi(x,x')r_{u\to v}(x)r_{v\to u}(x'). \tag{8.1}$$

For some large enough α we will have $0 < \sup_{h\in\mathcal{H}}\Phi_{h,\alpha} < \Psi_\alpha$ which would suggest that the free energy is positive but smaller than a typical cluster size and so the heuristics must break down. The condensation threshold corresponds to the point at which $\Phi_{h,\alpha} = \Psi_\alpha$. What is happening is that the biggest contribution to the expected value is coming from very large clusters of solutions that are very rare, exponentially unlikely to show up.

It is natural then to ask what the right prediction is for the condensation regime. The idea is to count the contribution of clusters of the right size, the largest that actually appear. We have already seen in the case of independent sets how introducing a Lagrange Multiplier tilts the measure to configurations subject to some restriction. Letting $w(m)$ denote the mass of the cluster with messages $m = \{m_{u\to v}\}$ we write

$$Z_\beta = \sum_m (w(m))^\beta$$

which is the partition function of the spin system with weights

$$\beta\sum_{v\in V}\Phi_v(m) - \beta\sum_{(u,v)\in E}\Phi_{(u,v)}(m).$$

We write $\Phi_{\alpha,\beta} = \lim_n \frac{1}{n}\log\mathbb{E}Z_\beta$. Since we are focusing on smaller cluster sizes we will take $\beta\in(0,1)$. The distribution of messages is then given by the fixed point equation,

$$s_{\alpha,\beta}(A) = \frac{1}{z}\int I(BP[\{m_i\}]\in A)\Big(\Phi_v(\{m_i\})\Big)^\beta\prod_{i=1}^{d-1}s_{\alpha,\beta}(dm_i).$$

The normalized log weight of a typical cluster $\Psi_{\alpha,\beta}$ under this measure can be computed similarly to (8.1). The exponent of the number of clusters of size $\Psi_{\alpha,\beta}$ is

$$\Phi_{\alpha,\beta} - \beta\Psi_{\alpha,\beta}$$

To calculate the free energy, the right tilting to take is $\beta_\star = \beta_\star(\alpha)$ given by

$$\beta_\star = \sup\{\beta\in[0,1] : \Phi_{\alpha,\beta} - \beta\Psi_{\alpha,\beta}\geq 0\}$$

With tilting $\beta_\star$, we have that $\Phi_{\alpha,\beta} - \beta\Psi_{\alpha,\beta} = 0$ which corresponds to order one number of clusters. These will be the largest contribution to the partition

function and so the 1RSB prediction for the free energy is

$$\Phi = \Psi_{\alpha,\beta_\star},$$

that is the cluster size given by the tilting.

Another way to interpret this is by assuming that there is a function $\Sigma(y) = \Sigma_\alpha(y)$, sometimes called the *complexity function*, such that the expected number of clusters of size of order e^{ny} is $e^{n\Sigma(y)+o(n)}$. Then

$$\Phi_{\alpha,\beta} = \max_y y\beta + \Sigma_\alpha(y)$$

which is the Legendre Transform of Σ_α. Assuming that Σ_α is concave, we can then invert $\Phi_{\alpha,\beta}$ to recover $\Sigma(y)$. The 1RSB prediction is then

$$\Phi = \sup\{y : \Sigma(y) \geq 0\}.$$

With this description, the condensation threshold corresponds to the point when the derivative at the largest root of $\Sigma_\alpha(y)$ is -1. If $\Sigma(y)$ is interpreted not just as the number of clusters but as determining a Poisson density then the conjecture that relative cluster sizes form a Poisson–Dirichlet distribution follows from properties of Poisson processes.

The non-analyticity of the free energy at the condensation threshold has been established in a number of models, most notably [6] in the q-colouring model on Erdos-Renyi random graphs for large q. These results established that the free energy corresponds to the first moment for all $\alpha < \alpha_c$ and then showed that the free energy was strictly less above α_c.

The above program to calculate the free energy throughout the condensation regime has only been rigorously carried out [49] for the regular k-NAESAT model for large k. In this case as the model is frozen so most messages are 0 and 1. The proof derived lower bounds on the free energy by considering only clusters where free variables occur in clusters of size bounded by L so that the set of allowable messages takes values in a finite set. As L tends to infinity the lower bounds converge to the predicted free energy. A matching upper bound is given by the interpolation method developed in [8, 25, 43].

Establishing the condensation threshold provides too coarse information to describe the relative cluster sizes. So far there are no models in which the set of solutions is known to be restricted to a finite number of clusters (although see [15]).

8.3. Satisfiability threshold. The final transition is the Satisfiability Threshold α_s after which there are no more solutions. Here it is sufficient to consider unweighed cluster, equivalently $\beta = 0$. The fixed point equation becomes

$$s_{\alpha,0}(A) = \frac{1}{z}\int I(BP[\{m_i\}] \in A) I\Big(\Phi_v(\{m_i\}) > 0\Big) \prod_{i=1}^{d-1} s_{\alpha,0}(dm_i).$$

Since $\Psi_{\alpha,0} = 0$ then it suffices to check if $\Phi_{\alpha,0} \geq 0$ and

$$\alpha_s = \sup\{\alpha : \Phi_{\alpha,0} \geq 0\}.$$

An equivalent formulation in terms of $\Sigma(y)$ the complexity function from the previous section is that

$$\alpha_s = \sup\{\alpha : \sup_y \Sigma_\alpha(y) \geq 0\}.$$

In fact, however, the satisfiability threshold can be determined more simply using warning propagation which simplifies things greatly as there are only a small number of states. Equation (7.2) gives a simple formula for the distribution of WP messages in $\mathcal{X}^\star$. Then

$$\Phi_\alpha^{(WP)} = \log \sum_{r,r_1,\ldots,r_d} I(WP[\{r_i\}] = r) \prod_{i=1}^{\alpha} s(r_i) - \frac{\alpha}{2} \log \left(1 - \sum_{r,r' \in \mathcal{X}} I(\psi(r,r') = 0) s(r) s(r') \right).$$

Then

$$\alpha_s = \sup\{\alpha : \Phi_\alpha^{(WP)} \geq 0\}.$$

Let's consider the example of the prediction for q-colourings on d-regular random graphs. We make the simplifying assumption that each colour is equally likely under the Survey Propagation equation $s(r)$ is a measure of the form

$$\xi_p(r) := \begin{cases} p/q, & \text{if } r \in \mathcal{X} \\ 1-p, & r = \star. \end{cases}$$

The fixed point equation for p is

$$p = \frac{\sum_{r_1\ldots,r_{d-1} \in \mathcal{X}_\star} I(|\{r_1 \ldots, r_{d-1}\} \cap \mathcal{X}| = q-1) \prod_i \xi_p(r_i)}{\sum_{r_1\ldots,r_{d-1} \in \mathcal{X}_\star} I(|\{r_1 \ldots, r_{d-1}\} \cap \mathcal{X}| \leq q-1) \prod_i \xi_p(r_i)}. \tag{8.2}$$

The numerator counts the probability that exactly $q-1$ other colours appear among the neighbours leaving one more fixed. The denominator includes the possibility that the root vertex can be coloured in two or more ways. The ratio is then p, the probability of a frozen state. We will write p_d for the largest fixed point of equation (8.2). Then the free energy for the warning propagation model is

$$\Phi_d^{(WP)} = \log \sum_{r_1\ldots,r_{d-1} \in \mathcal{X}_\star} I(|\{r_1 \ldots, r_{d-1}\} \cap \mathcal{X}| \leq q-1) \prod_i \xi_{p_d}(r_i) - \frac{d}{2} \log(1 - \frac{p_d^2}{q})$$

Then $d_{\max} = \max\{d : \Phi_d^{(WP)} \geq 0\}$ is the 1RSB prediction for the largest d for which a random d-regular graph is q-colourable. This has yet to be

established rigorously but the gap between upper and lower bounds is now quite small [2, 12, 16].

Satisfiability thresholds have now been proved in a number of models, first for K-regular NAESAT [24] and a regular version of K-SAT [14] both for large enough K. In the following section we will describe how the same methods were used to establish the size of the largest independent set for random d-regular graphs when d is large enough. Finally, the satisfiability threshold for random K-SAT [22] was established for large K using methods described in Section 10.

A key assumption in the proof is that the free (non-frozen) variables with the clusters can be assigned in a satisfiable way. Particularly when K is large for SAT or q is large for colourings, the free variables will be small highly subcritical trees and there will always be a way to assign the variables. In some CSPs though, for example MAX-CUT the free variables are supercritical and as a result the model is no longer 1RSB.

9. Maximal independent sets

In this section we consider the size of the maximal independent set of a random d-regular graph when d is large following the proof of [23]. Physicists predict [7] that the model is 1RSB provided $d \geq 20$. In this section we will focus on lower bounds for the size of the maximal independent set. Upper bounds can be derived using the interpolation method or in fact by an appropriate first moment estimate as was shown in [23].

To reach the maximal independent sets, one might think to consider the hardcore model with very high fugacity λ. This, however, leads to TIFPs which have density approaching $\frac{1}{2}$, much higher than the maximal density of $\frac{(2+o_d(1))\log d}{d}$.

The 1RSB formalism tells us to count clusters, so we can look at the BP fixed equations from (6.11) as we take $\lambda \to \infty$. In the limit it becomes

$$BP[m_1, \ldots, m_{d-1}] = \begin{cases} 0, & \text{if } \prod_i (1 - m_i) = 0 \\ 1, & \text{otherwise.} \end{cases}$$

This heuristic turns out to be the right object to count. The messages $m_{u \to v}$ lie in $\{0, 1\}$ and satify the constraints that an outgoing message is 1 if and only if all the incoming messages are 0.

How does this cluster configuration relate to maximal independent sets? The answer is that they encode clusters of *locally maximal independent sets* which are ones that cannot be increased without a large number of vertices being changed. This is more easily seen in an equivalent formulation called the *frozen model*. Each vertex is takes values in $\{1, 0, f\}$ where f denotes a free state. The configuration must follow the following rules:

- A vertex of type 1 must be adjacent to only vertices of type 0. These are the vertices with all 0 incoming messages.

- A vertex of type 0 must be adjacent to at least two 1's. These are vertices with two or more incoming 1 messages.
- Vertices of type f come in matched pairs and are not adjacent to 1's. They are vertices with a single incoming and a single outgoing 1 message on the same edge. Edges with a 1 message in both directions are the pairs of free variables.

The interpretation of the frozen model is as a cluster of locally maximal independent sets where 1 means always in the independent set as 0 means never in. The free variables are vertices which may or may not be in the independent set. At most one vertex from each pair of free variables is the independent set. When the density of free variables is small enough the clusters of free variables are very subcritical trees and it is always possible to find an independent set comprising of half of the free variables. This can be shown to hold when d is large enough.

Thus for the frozen model we expect the density of the maximal independent set to be $\overline{h}_1+\frac{1}{2}\overline{h}_f$. We want to count independent sets with density α we introduce a Lagrange multiplier and count clusters weighted by $e^{\theta n(\overline{h}_1+\frac{1}{2}\overline{h}_f)}$. We can then look for fixed points of the BP equation. However, the state space of spins is a pair of messages on each edge. The recursion would be much simpler if we instead considered it on a tree with messages only in one direction. In this case the BP fixed point is

$$BP_\theta[q] = \frac{e^\theta(1-q)^{d-1}}{1+(e^\theta-1)(1-q)^{d-1}}.$$

Note that when the root is a free variable there is no change in the weight of a cluster, one if its children must have previously been a 1 but becomes part of a free pair – the contributions cancel out. So the only change in weight is when the root is in state 1. For each θ there is a unique fixed point q_θ. The corresponding marginals are

$$\overline{h}_1 = \frac{e^\theta(1-q_\theta)^d}{1+(e^\theta-1)(1-q_\theta)^d}, \qquad \overline{h}_f = \frac{dq_\theta(1-q_\theta)^{d-1}}{1+(e^\theta-1)(1-q_\theta)^d}$$

and so the density of the independent set achieved is

$$\alpha_\theta = \frac{e^\theta(1-q_\theta)^d}{1+(e^\theta-1)(1-q_\theta)^d} + \frac{1}{2}\frac{dq_\theta(1-q_\theta)^{d-1}}{1+(e^\theta-1)(1-q_\theta)^d}$$

The free energy in this weighted model is then

$$\log(1+(e^\theta-1)(1-q_\theta)^d) - \frac{d}{2}\log(1-(1-e^{-\theta})q_\theta^2)$$

and since each cluster has weight $e^{\theta n(\overline{h}_1+\frac{1}{2}\overline{h}_f)}$ the rate function for the number of unweighted clusters is

$$\Phi_\theta = \log(1+(e^\theta-1)(1-q_\theta)^d) - \frac{d}{2}\log(1-(1-e^{-\theta})q_\theta^2) - \alpha_\theta\theta.$$

We choose $\theta_\star = \max\{\theta : \Phi_\theta \geq 0\}$ and then the density of the maximal independent set is estimated to be $\alpha_{\theta_\star}$. This was established in [23] for all large enough d.

The proof is via the second moment method applied to a density $\alpha' = \alpha_\star - \epsilon$. In the pair model the BP equation is

$$BP_\theta[q](x,x') = \frac{A_{xx'}}{A_{00}+A_{10}+A_{01}+A_{11}}$$
$$A_{00} = 1-(q_{01}+q_{00})^{d-1}-(q_{10}+q_{00})^{d-1}+q_{00}^{d-1}, A_{11} = e^{2\theta}q_{00}^{d-1},$$
$$A_{01} = e^{\theta}((q_{10}+q_{00})^{d-1}-q_{00}^{d-1}), A_{10} = e^{\theta}((q_{01}+q_{00})^{d-1}-q_{00}^{d-1}).$$

We need to find θ's with fixed points corresponding to a density α'. Let θ' be the weight such that $\alpha_{\theta'} = \alpha'$. An analysis of fixed points shows that there are two,

$$q^{\otimes 2}(x,x') = q_{\theta'}(x)q_{\theta'}(x'), \qquad q^{=}(x,x') = q_{\theta'}(x)I(x=x')$$

with corresponding weights $e^{\theta'}$ and $e^{\theta'/2}$ respectively. The first fixed point corresponds to two independent copies, that is with product empirical distribution while the second corresponds to a pair of configurations that are identical. The correspond rate functions are $2\Phi_{\theta'}$ and $\Phi_{\theta'}$. Since $\alpha' < \alpha_\star$, we have that $\Phi_{\theta'} > 0$ and so the global maxima corresponds to the independent copies maxima. It follows that the second moment method succeeds establishing the density of the maximal independent set.

Let us return to the earlier assumption that it was legitimate to consider the simpler one directional BP fixed point. If we consider a Gibbs measure on a $(d-1)$-ary tree all the spins are determined by the messages at the boundary, that is messages at the leaves plus one at the root. We can ask whether changing the message at the root propagates as a supercritical branching process or does it quickly die out as a subcritical branching process.

In [23, 24] it was shown that a bi-directional TIFP with the property that the messages from the root is subcritical is given by the simpler one directional fixed point. This can be seen in the argument of Proposition 5 where we establish that a maximizer of Φ_h is given by a BP fixed point via a resampling argument. In the cluster model this means a bi-directional fixed point. However, suppose that we perform the resampling in Proposition 5 but fix only the outgoing messages on ∂W, not the incoming messages. Not knowing the incoming message at $u \in \partial W$ means that we may not be able to determine the messages in a neighbourhood of u either. However, our assumption that the effect at the root does not propagate far means that these sets are typically small and disjoint. We can then perform the resampling of W given just the outgoing messages from ∂W. The messages from W are then the new incoming messages into ∂W from which the rest of the configuration can be constructed.

The approach of [24] to establish a bounded effect of the message at the root involved a priori bounds to restrict the possible range of empirical

distributions h that could be a global maximizer. An alternative approach which was used in [48] for the NAESAT model is to change the spin system in such a way that the message from the root must have a bounded effect. Roughly, this corresponds to restricting free components to have a bounded size. While this leads to a more complicated BP message, it is automatically sufficient to consider one directional fixed point.

10. Random K-SAT

In this final section we will discuss the satisfiability threshold for random K-SAT, a problem known as the Satisfiability Conjecture. Recall that random K-SAT gives a random Boolean formula which is the AND of $m = \alpha n$ clauses, each of which is the OR of K variables chosen uniformly from $\{+x_1, -x_1, \dots, +x_n, -x_n\}$. The case of $K = 2$ is substantially simpler and was established, independently, by multiple groups [11, 30] in 1992. It can be analysed using a branching process analysis which gives the location of the threshold. For $K \geq 3$ it is believed to be in the 1RSB universality class.

The graphical representation of random K-SAT will be important. An instance of K-SAT can be described as a factor graph, a bi-partite graph (V, F, E) with vertices $V \cup F$ and edges E. The vertices V represent the variables of the formula, while F represents the set of $m = \alpha n$ clauses. An edge from vertex v to clause a is present if the variable v appears in clause a in the formula. For each edge $e = av$ we will also have an additional variable $L_{av} \in \{\pm\}$ representing whether or not the variable v appears negated in clause a. While all clause vertices have degree K by construction, the variable vertices have a random degree which is asymptotically Poisson with mean $K\alpha$. For the most part so far we have focussed on models on regular graphs. The major additional challenge posed by the random K-SAT model is due to the effect of neighbourhood fluctuations described in Section 6.6.

The phenomena of neighbourhood fluctuations is due to inhomogeneity of the marginals of the variables. If $+x_i$ appears in many clauses but $-x_i$ appears in none of them then x_i is more likely to be $+$ in a random solution. When K is large, this effect tends to be small, most variables have marginals close to $\frac{1}{2}$ but any deviation is enough to cause the second moment method to fail. Achlioptas and Peres [4] weighted solutions by λ raised to the power of the number of satisfied literals in the formula. By choosing the right choice of $\lambda < 1$ this precisely balanced the effect on marginals so that they remained $\frac{1}{2}$. This resulted in a major improvement on the previous best lower bound, improving it to $\alpha_s \geq 2^K \log 2 - O(K)$. However, the effect of adding in this external field tilts the measure towards solutions with a biased distribution and does not address the clustering of solutions so it cannot be expected to reach the satisfiability threshold using this approach.

10.1. Clustering of solutions. As with order 1RSB models, there are simple combinatorial representations of the clusters of random K-SAT. In the frozen model variables can be assigned to $\{+, -, f\}$ where $+, -$ will mean

that the vertex is frozen and free means that it varies within the cluster. We will call a clause a forcing for a variable v if all other variables in the clause evaluate to false. The frozen model can be described by the following set of rules:

- Every clause with no free variable is satisfied.
- Each frozen variable is forced by at least one clause.
- Free variable are forced by no clauses.

An equivalent but particularly tractable representation by Coja-Oghlan and Panagiotou [14] assigns colours $\{b, g, r, y\}$ the edges av as follows:

- Green: If v is free.
- Red: If v is forced in clause a.
- Blue: If v evaluates to true in clause a but is not forced.
- Yellow: If v evaluates to false in clause a.

The colouring must satisfy the following rules. In every clause with a red edge the other $K-1$ edges are all yellow. In any clause without a red edge, at least two edges are in $\{b, g\}$. Each free vertex is all green. A vertex frozen to $+$ has all its edges with a negative literal coloured yellow and all edges with a positive literal coloured blue or red, at least one of which much be red. For a vertex with frozen to $-$ it is reversed.

The challenge in pushing a second moment method to the satisfiability threshold is that the distribution of solutions depends on the local neighbourhood but it is such fluctuations that lead to neighbourhood fluctuations. This suggests dividing the vertices into types according to their neighbourhood. The depth R type t_v of a variable v is the graph structure of $B_{2R}(v)$ of the radius $2R$ ball in the factor graph along with the literals $\{L_e\}_{e\in B_{2R}(v)}$. The *clause type* of a clause $a \in F$ is the graph structure of $B_{2R+1}(a)$ of the radius $2R+1$ ball in the factor graph along with the literals $\{L_e\}_{e\in B_{2R+1}(a)}$. With this definition the clause type contains the information of the variable types of the variables in the clause. In a frozen model configuration, let $\omega_{t,j}$ denote the empirical distribution colours on the j-th edge of clauses of type t.

An important step in this direction was carried out by [14]. They restricted their counting to what they called *judicious* colourings such that $\omega_{t,j} = \pi_{t,j}$ where π is a measure that may be specified. It is chosen such that $\pi_{t,j} = \pi_{t_{v_j}}$, that is it is a function not of the whole clause type but only of t_{v_j}, the variable type of the j-th variable in the clause. It is this specification that avoids the neighbourhood fluctuations as the marginal at a vertex becomes effectively independent of its wider neighbourhood. In [14] this was performed with depth 1 neighbourhoods and achieved what was then the best lower bound on the satisfiability threshold. To get a tight lower bound [22] took a sequence of lower bounds as $R \to \infty$ that converge to the true satisfiability threshold. The choice of π is derived from survey propagation.

10.2. BP fixed point and 1RSB prediction. As the factor graphs have both variable and clause vertices there will be two sorts of messages. Messages $m_{v\to a}$ from a variable v to clause a will take values $\{+,-,f\}$ with $+$ (respectively $-$) denoting that the variable is frozen and that value in the clause including possible negations evaluates to true (respectively false). Otherwise if it is not frozen then it is f. The clause to variable messages $m_{a\to v}$ take values $\{+,f\}$ with $+$ indicating that the edge av is forced by the clause and otherwise it is f. The BP equations are

$$m_{a\to v}=\begin{cases}+, & \text{if } m_{v'\to a}=- \text{ for all } v'\in\partial a\setminus v,\\ f, & \text{otherwise,}\end{cases}$$

and

$$m_{v\to a}=\begin{cases}+, & \text{if } M_+\geq 1, M_-=0\\ -, & \text{if } M_-\geq 1, M_+=0\\ f, & M_+=M_-=0,\end{cases}$$

where

$$M_\pm=\sum_{\substack{a'\in\partial v\setminus a\\ L_{a'v}=\pm L_{av}}} I(m_{a'\to v}=+).$$

Finally, each edge av much satisfy the compatibility constraint that $m_{v\to a}\neq -$ or $m_{a\to v}=f$ which ensures that there are no unsatisfied clauses.

The local weak limit of the graph is a branching process tree with alternating levels of variable nodes having Poisson with mean $K\alpha$ children and clause nodes with $K-1$ children and where each edge has an independent literal L_e. Let T be such a tree rooted as ρ and let $T^{(R)}$ denote the tree truncated at depth $2R$ (meaning R levels of variable vertices). We will let ρ^+ denote a parent clause of ρ. In $T^{(R)}$ we set a boundary condition on each leaf u setting them as frozen with probability $\frac{1}{2}$ uniformly on $\pm$ so

$$m^{(R)}_{u\to u^+}(+)=m^{(R)}_{u\to u^+}(-)=\frac{1}{2},\quad m^{(R)}_{u\to u^+}(f)=0. \tag{10.1}$$

Then given this boundary condition let $\eta_R=m^{(R)}_{\rho\to\rho^+}$ be the message from ρ to its parent. Then [22] shows that η_R converges almost surely as $R\to\infty$ to a measure η which represents the probability in the limit that the root is frozen to 1 in the random tree T. Let μ_α be the law of η. The interpretation of this is that for a randomly chosen variable in a K-SAT formula, the fraction of clusters in which it is frozen as $+$ is random and asymptotically has law μ_α. We can now define the 1RSB prediction for the random K-SAT

model. Define

$$\Phi(\alpha) = \sum_{\underline{d}} p_\alpha(\underline{d}) \int \ln\left(\Pi^+ + \Pi^- - \Pi^+\Pi^-\right) \prod_j d\mu_\alpha(\eta_j) \prod_{i,j} d\mu_\alpha(\eta_{ij}^\pm)$$

$$-\alpha(k-1) \int \ln\left(1 - \prod_{j=1}^{K} \eta_j\right) \prod_j d\mu_\alpha(\eta_j) \prod_{i,j} d\mu_\alpha(\eta_{ij}^\pm) \tag{10.2}$$

where $\underline{d} = (d_+, d_-)$ and

$$p_\alpha(\underline{d}) = \frac{e^{-K\alpha}(K\alpha/2)^{d^+ + d^-}}{(d^+)!(d^-)!}, \quad \Pi^\pm \equiv \Pi^\pm(\underline{d}, \underline{\eta}) \equiv \prod_{i=1}^{d^\pm}\left(1 - \prod_{j=1}^{K-1} \eta_{ij}^\pm\right).$$

In equation (10.2) $\Phi(\alpha)$ represents the predicted free energy of the cluster model. The first term is the effect of adding a new vertex with $d_\pm$ neighbours with literal $\pm$ while the second come from removing, on average, $(\alpha - 1)K$ clauses to maintain a constant average degree as the number of vertices increases from n to $n+1$. The 1RSB prediction for the satisfiability threshold is then

$$\alpha_s = \sup\{\alpha : \Phi(\alpha) \geq 0\},$$

that is the highest density of constraints such that the cluster model has a nonnegative free energy. The main theorem of [22] is that for large enough K this correctly predicts the satisfiability threshold.

THEOREM 9. *The Satisfiability Conjecture holds for all sufficiently large K at the 1RSB prediction α_s. That is, for $\alpha > \alpha_s$ (resp. $\alpha < \alpha_s$) the probability of the random K-SAT formula with constraint density α tends to 0 (resp. tends to 1) as $n \to \infty$.*

The proof gives a sequence of lower bounds via the second moment method with vertex types given by a radius R. First a preprocessing step removes $\epsilon_R n$ vertices to remove the most problematic vertex types in a process described in the following subsection. In order to achieve the satisfiability threshold it is essential that $\epsilon_R \to 0$ as $R \to \infty$, that is to achieve a tight bound the fraction of removed vertices must be vanishing. The proof counts clusters, encoded as colourings, that satisfy the judicious condition.

The empirical distribution $\pi_{t,j}$ is given as follows. In a tree with clause a having clause type $t_a = t$, let v_j be the j-th variable in the clause. Then root the tree at v_j and put a boundary condition at depth $2R$ according to (10.1). The distribution of the colouring on the edge av_j then gives $\pi_{t,j}$. Note that by construction this depends only on the variable type, not the full clause type.

When K is large, the $\pi_{t,j}$ will be very concentrated. Since the typical degree of a variable vertex is of order 2^K, the fluctutations are $O(2^{K/2})$

which will typically have a negligible affect on the marginal. Typically

$$\pi_{t,j}(x) = \begin{cases} \frac{1}{2} + O(2^{-K/10}, & \text{if } x \in \{b, y\} \\ 2^{-K} + O(2^{-11K/10}), & \text{if } x = r \\ 2^{-K-1} + O(2^{-11K/10}), & \text{if } x = g. \end{cases} \tag{10.3}$$

Also typically the number of times the variable appears in the formula with literal $\pm$ is $K2^{K-1}\log 2 + O(2^{2K/3})$. Such variable types are called *nice*. There will be some non-nice types, for example those outside the giant component of the graph but these are double exponentially rare in K.

10.3. Graph preprocessing. In this subsection we describe preprocessing of the graph. As the full description in [22] runs to 3 pages we will describe just the high level motivation. It is necessary to first remove a small fraction of the vertices for a couple of reasons. One is to restrict to a bounded number of variable and clause types by removing very large degree vertices. A second is to ensure that the empirical distribution imposed on a clause is in fact an achievable distribution. For example there are rare clause types such that $\sum_{j=1}^{k} \pi_{t,j}(r) > 1$ which is invalid as each clause may only have at most one red forcing edge. Clauses that satisfy properties to ensure that the distribution is achievable are called *coherent*. As $R \to \infty$ the fraction of coherent clauses rapidly tends to 1.

While, as noted above, most vertices are nice the fraction of vertices that are not nice does not vanish as $r \to \infty$. We cannot remove all the non-nice vertices without shifting the satisfiability threshold. Indeed, perhaps the most significant challenge in [22] is to construct a system that deals with the non-nice variables. Instead we retain non-nice vertices which are surrounded by sufficiently many layers of nice variables. In particular vertices are grouped into *enclosures* such that each enclosure A is a tree of diameter at most $R/10$ and such that for any $v \in \partial A$

$$\sum_{u \in \partial A, u \neq v} 2^{-\frac{101K}{100} J_{u,v} + \frac{K}{100} d(u,v)} \leq 2^{-K/2} \tag{10.4}$$

where $J_{u,v}$ is the number of nice vertices between u and v. Such enclosures are merged together and treated as a single *compound* vertex type. The vast majority are singletons of simple vertex type.

For each fixed R a small fraction of the variables removed in this process but the proportion will tend to 0. The preprocessing algorithm removes components of vertices that are trees or have at most a single cycle and any such configuration can be shown to have solutions. It will be assumed that the removed variables do not satisfy the clauses they appear in so in the remaining clauses which are of size $K - 1$. The number of colourings on the remaining graph has rate function $\Phi(\alpha, R) = \Phi(\alpha) - \epsilon(\alpha, R)$ such that $\epsilon(\alpha, R) \to 0$ as $R \to \infty$ so for $\alpha < \alpha_s$ and R large enough we have $\Phi(\alpha, R) > 0$. Thus it is enough to prove this is indeed the free energy on the preprocessed graph.

10.4. Optimization and sharp thresholds. The bulk of the proof involves the application of the second moment method. It begins with the preprocessed graph and conditions on the number of each clause type. The proof counts the number of pairs of judicious colouring configurations on the pre-processed graph and in particular shows that the dominant contribution comes from configurations with an independent empirical distribution. It is possible to show that most pairs of frozen model solutions are either almost identical or approximately independent via simple a priori bounds. By controlling the number of small perturbations of solutions it is possible to show that the almost identical solutions make a negligible contribution. The argument then shows that most pairs of colourings have empirical distribution close to a product measure on the singleton variable types.

To prove that the empirical distribution h that maximizes the rate function for the pair model we choose the clause type t_a and edge j such that the empirical distribution on $h_{t_a,j}$ is furthest from $\pi_{t_a,j} \otimes \pi_{t_a,j}$ (in an appropriate metric). Let t_v be the variable type of the j-th variable of clause type t_a. The proof resamples all the variables of type t_v and all attached clauses, similarly to Proportion 5. Since the local neighbourhood is a tree the distribution can be given by appropriately chosen external fields, in effect Lagrange multipliers. These are arrived at iteratively starting from the product measure and then adding external fields to achieve the right distribution on the neighbouring vertex types and to enforce the judicious condition on each of the copies. These adjustments can be bounded to show that the resampled distribution h' the distance from the product measure on edge type $h_{t_a,j}$ has decreased by half.

When all the variables in the tree are nice there is an effective decay of correlation of $2^{-(2-\epsilon)K}$ for large enough K which together with (10.4) implies the resampling bound required. Perhaps the most significant insight in [22] is that by relative entropy considerations not much is lost in terms of decay of correlation when there are non-nice variables whose local marginals $\pi_{t,j}$ can be essentially arbitrary. The fact that we do not have to quantitatively analyse the recursion through with these variables means that the proof can deal with all neighbourhood types, an essential ingredient in reaching the satisfiability threshold.

This implies that the maximizer is the product measure and that the negative definite condition holds. Thus, the second moment implies solutions exist with probability bounded away from 0. By taking large enough R this holds for any $\alpha < \alpha_s$.

The final element in establishing a lower bound is an appeal to the sharp threshold theorem of Friedgut [26]. This states that there is a sequence $\alpha(n)$ such that for any $\epsilon > 0$ the probability that the random formula is satisfiable with constraint density $\alpha(n) - \epsilon$ (resp. $\alpha(n) + \epsilon$) tends to 1 (resp. 0) as n tends to infinity. That is there is a sharp threshold in the probability of being satisfiable around $\alpha(n)$. What is missing in terms of the Satisfiability Conjecture is that $\alpha(n)$ may depend on n. The proof of Friedgut's theorem

lies in the abstract theory of threshold phenomena of Boolean functions and so does not yield any information on $\alpha(n)$. Combined with the second moment lower bound it implies that for any $\alpha < \alpha_s$ the probability of being satisfiable tends to 1. Together with the upper bound from the interpolation method, this implies the Satisfiability Conjecture for large enough K at the 1RSB prediction.

Acknowledgements

We would like to thank the referee for a very thorough reading of the paper.

References

[1] D. Achlioptas and A. Coja-Oghlan, *Algorithmic barriers from phase transitions*, 49th annual ieee symposium on foundations of computer science, 2008. focs'08, 2008, pp. 793–802.

[2] D. Achlioptas and A. Naor, *The two possible values of the chromatic number of a random graph*, Annals of Mathematics (2005), 1335–1351. MR2179732

[3] D. Achlioptas and M. Molloy, *The analysis of a list-coloring algorithm on a random graph*, Foundations of computer science, 1997. Proceedings., 38th annual symposium on, 1997, pp. 204–212.

[4] D. Achlioptas and Y. Peres, *The threshold for random k-sat is* $2^k \log 2 - o(k)$, Journal of the American Mathematical Society **17** (2004), no. 4, 947–973. MR2083472

[5] D. Achlioptas and F. Ricci-Tersenghi, *On the solution-space geometry of random constraint satisfaction problems*, Proceedings of the thirty-eighth annual acm symposium on theory of computing, 2006, pp. 130–139. MR2277138

[6] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Raßmann, and D. Vilenchik, *The condensation phase transition in random graph coloring*, Communications in Mathematical Physics **341** (2016), no. 2, 543–606. MR3440196

[7] J. Barbier, F. Krzakala, L. Zdeborová, and P. Zhang, *The hard-core model on random graphs revisited*, Journal of physics: Conference series **473** (2013), 012021.

[8] M. Bayati, D. Gamarnik, and P. Tetali, *Combinatorial approach to the interpolation method and scaling limits in sparse random graphs*, Proceedings of the forty-second acm symposium on theory of computing, 2010, pp. 105–114. MR2743259

[9] B. Bollobás, *The independence ratio of regular graphs*, Proceedings of the American Mathematical Society (1981), 433–436. MR0624948

[10] A. Braunstein, M. Mézard, and R. Zecchina, *Survey propagation: An algorithm for satisfiability*, Random Structures & Algorithms **27** (2005), no. 2, 201–226. MR2155706

[11] V. Chvátal and B. Reed, *Mick gets some (the odds are on his side)*, Proc. ieee symp. (focs), 1992, pp. 620–627.

[12] A. Coja-Oghlan, C. Efthymiou, and S. Hetterich, *On the chromatic number of random regular graphs*, Journal of Combinatorial Theory, Series B **116** (2016), 367–439. MR3425250

[13] A. Coja-Oghlan and C. Efthymiou, *On independent sets in random graphs*, Random Structures & Algorithms **47** (2015), no. 3, 436–486. MR3385742

[14] A. Coja-Oghlan and K. Panagiotou, *The asymptotic k-sat threshold*, Advances in Mathematics **288** (2016), 985–1068. MR3436404

[15] A. Coja-Oghlan and W. Perkins, *Spin systems on bethe lattices*, arXiv preprint arXiv:1808.03440 (2018). MR4032871

[16] A. Coja-Oghlan and D. Vilenchik, *Chasing the k-colorability threshold*, Foundations of computer science (focs), 2013 ieee 54th annual symposium on, 2013, pp. 380–389. MR3246240

[17] A. Coja-Oghlan, *A better algorithm for random k-sat*, SIAM Journal on Computing **39** (2010), no. 7, 2823–2864. MR2645890

[18] A. Dembo and A. Montanari, *Ising models on locally tree-like graphs*, The Annals of Applied Probability **20** (2010), 565–592. MR2650042

[19] A. Dembo, A. Montanari, S. Sen, et al., *Extremal cuts of sparse random graphs*, The Annals of Probability **45** (2017), no. 2, 1190–1217. MR3630296

[20] A. Dembo, A. Montanari, A. Sly, and N. Sun, *The replica symmetric solution for potts models on d-regular graphs*, Communications in Mathematical Physics **327** (2014), no. 2, 551–575. MR3183409

[21] A. Dembo, A. Montanari, N. Sun, et al., *Factor models on locally tree-like graphs*, The Annals of Probability **41** (2013), no. 6, 4162–4213. MR3161472

[22] J. Ding, A. Sly, and N. Sun, *Proof of the satisfiability conjecture for large k*, Proceedings of the forty-seventh annual acm on symposium on theory of computing, 2015, pp. 59–68. MR3388183

[23] J. Ding, A. Sly, and N. Sun, *Maximum independent sets on random regular graphs*, Acta Mathematica **217** (2016), no. 2, 263–340. MR3689942

[24] J. Ding, A. Sly, and N. Sun, *Satisfiability threshold for random regular NAE-SAT*, Communications in Mathematical Physics **341** (2016), no. 2, 435–489. MR3440193

[25] S. Franz and M. Leone, *Replica bounds for optimization problems and diluted spin systems*, Journal of Statistical Physics **111** (2003), no. 3-4, 535–564. MR1972121

[26] E. Friedgut, J. Bourgain, et al., *Sharp thresholds of graph properties, and the k-sat problem*, Journal of the American mathematical Society **12** (1999), no. 4, 1017–1054. MR1678031

[27] A. M Frieze and T Łuczak, *On the independence and chromatic numbers of random regular graphs*, Journal of Combinatorial Theory, Series B **54** (1992), no. 1, 123–132. MR1142268

[28] A. Galanis, D. Štefankovič, and E. Vigoda, *Inapproximability for antiferromagnetic spin systems in the tree non-uniqueness region*, Proceedings of the forty-sixth annual acm symposium on theory of computing, 2014, pp. 823–831. MR3239010

[29] A Gerschcnfeld and A Monianari, *Reconstruction for models on random graphs*, 48th annual ieee symposium on foundations of computer science (focs'07).

[30] A. Goerdt, *A threshold for unsatisfiability*, J. Comput. System Sci. **53** (1996), no. 3, 469–486. MR1423858 (98i:03012)

[31] G. R Grimmett and C. J. McDiarmid, *On colouring random graphs*, Mathematical proceedings of the cambridge philosophical society, 1975, pp. 313–324. MR0369129

[32] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborova, *Gibbs states and the set of solutions of random constraint satisfaction problems*, Proceedings of the National Academy of Sciences **104** (2007), no. 25, 10318. MR2317690

[33] T. Łuczak, *The chromatic number of random graphs*, Combinatorica **11** (1991), no. 1, 45–54. MR1112273

[34] B. McKay, *Independent sets in regular graphs of high girth*, Ars Combinatoria **23** (1987), 179–185. MR0890138

[35] S. Mertens, M. Mézard, and R. Zecchina, *Threshold values of random k-sat from the cavity method*, Random Structures & Algorithms **28** (2006), no. 3, 340–373. MR2213115

[36] M. Mezard and A. Montanari, *Information, physics, and computation*, Oxford University Press, USA, 2009. MR2518205

[37] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond*, World scientific, Singapore, 1987. MR1026102

[38] M. Mézard, G. Parisi, and R. Zecchina, *Analytic and algorithmic solution of random satisfiability problems*, Science **297** (2002), no. 5582, 812.
[39] M. Molloy, *The freezing threshold for k-colourings of a random graph*, Proceedings of the forty-fourth annual acm symposium on theory of computing, 2012, pp. 921–930. MR2961555
[40] A. Montanari, E. Mossel, and A. Sly, *The weak limit of Ising models on locally tree-like graphs*, To appear in Probability Theory and Related Fields (2010). MR2875752
[41] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian, *Clusters of solutions and replica symmetry breaking in random k-satisfiability*, Journal of Statistical Mechanics: Theory and Experiment **2008** (2008), no. 04, P04004.
[42] E. Mossel, Y. Peres, et al., *Information flow on trees*, The Annals of Applied Probability **13** (2003), no. 3, 817–844. MR1994038
[43] D. Panchenko and M. Talagrand, *Bounds for diluted mean-fields spin glass models*, Probability Theory and Related Fields **130** (2004), no. 3, 319–336. MR2095932
[44] G. Parisi, *Infinite number of order parameters for spin-glasses*, Physical Review Letters **43** (1979), no. 23, 1754. MR0702601
[45] G. Parisi, *The order parameter for spin glasses: a function on the interval 0-1*, Journal of Physics A: Mathematical and General **13** (1980), no. 3, 1101. MR0702601
[46] R. W. Robinson and N. C. Wormald, *Almost all regular graphs are hamiltonian*, Random Structures & Algorithms **5** (1994), no. 2, 363–374. MR1262985
[47] A. Sly, *Reconstruction of random colourings*, Communications in Mathematical Physics **288** (2009), no. 3, 943–961. MR2504861
[48] A. Sly, N. Sun, and Y. Zhang, *The number of solutions for random regular nae-sat*, Foundations of computer science (focs), 2016 ieee 57th annual symposium on, 2016, pp. 724–731. MR3631035
[49] A. Sly, N. Sun, and Y. Zhang, *The number of solutions for random regular NAE-SAT* (2016), 724–731. MR3631035
[50] A. Sly and Y. Zhang, *Reconstruction of colourings without freezing*, arXiv preprint arXiv:1610.02770 (2016).
[51] N. C Wormald et al., *Differential equations for random processes and random graphs*, The annals of applied probability **5** (1995), no. 4, 1217–1235. MR1384372

Department of Mathematics, Princeton University, Princeton, New Jersey, U.S.A.
E-mail address: asly@math.princeton.edu