PHASE TRANSITIONS IN RANDOM STRUCTURES

AMIN COJA-OGHLAN

ABSTRACT. These notes provides a perspective on phase transitions in random discrete structures informed by recent ideas and techniques from statistical physics. The lectures are based on the publications [11, 20, 22].

1. RANDOM GRAPHS

1.1. **Random graphs.** The theory of random graphs started in 1960 with the work of Paul Erdős and Alfred Rényi on the phase transition for the emergence of the 'giant component' [28]. Specifically, Erdős and Rényi investigated the size of the largest component of the random graph $\mathbb{G} = \mathbb{G}(n,p)$ with vertex set $V = V_n = \{x_1, \ldots, x_n\}$ in which each of the $\binom{n}{2}$ possible edges is present with probability $0 \le p \le 1$ independently. The most interesting regime of the edge probability for this problem is $p = d/n$ for a fixed number $d > 0$ and $n \gg 1$. Erdős and Rényi, whose seminal papers on the theory of random graphs have over 12,000 citations at this time, proved the following result.

Theorem 1.1 ([28])**.** *Let $\mathscr{L}(\mathbb{G})$ be the number of vertices in the largest component of $\mathbb{G}$. Moreover, let $\lambda(d) \in (0,1]$ be the smallest fixed point of the function $x \mapsto \exp(d(x-1))$. Then $\frac{1}{n}\mathscr{L}(\mathbb{G})$ converges to $1-\lambda(d)$ in probability.*

The function $d \mapsto 1-\lambda(d)$ is identically 0 for $d \le 1$ because the only fixed point of $\exp(d(x-1))$ is $x = 1$ in this case. But for $d > 1$ we have $1-\lambda(d) > 0$. Thus, $d \mapsto 1-\lambda(d)$ is non-analytic at the point $d = 1$. That is why we say that a *phase transition* occurs.

Since the days of Erdős and Rényi the theory of random graphs has gone from strength to strength. Not only have various models of random graphs been invented in order to study the evolution of complex networks [27, 35], an objective already mentioned by Erdős and Rényi. But also have random graphs become the protagonists of modern coding theory [54]. For instance, novel probabilistic constructions known as *low-density parity check codes*, based on random graphs, achieve channel capacity and allow for efficient encoding and decoding algorithms [41], thereby solving the fundamental problem of coding theory since its inception by Shannon. Another application of random graphs is remarkable constructions of integrality gap instances for convex optimisation problems [30]. Moreover, random graphs have been harnessed as gadgets in computer science to prove tight bounds on the computational complexity of counting and sampling problems [32, 50, 55] as well as to construct candidate one-way functions [24]. Furthermore, random graphs provide challenging benchmarks for the study of algorithms and inspired the discovery and design of new powerful algorithmic techniques, e.g., [3, 16, 15]. Moreover, in physics

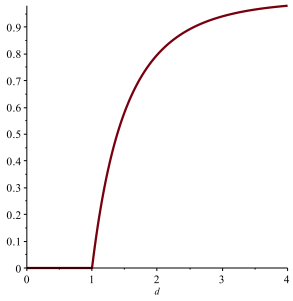

FIGURE 1. The function $1-\lambda(d)$.

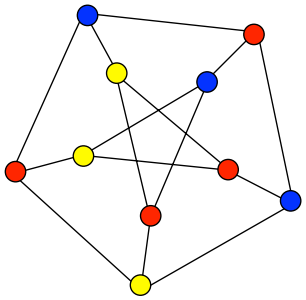

FIGURE 2. The graph colouring problem.

random graphs have been investigated as models of 'disordered systems' such as glasses [44]. In addition, in combinatorics random graphs have been instrumental to the proofs of theorems whose statements are entirely deterministic [9]. A recent example is the proof of the existence of combinatorial designs [38]. For a recent introduction to the theory of random graphs see [31].

The quest for phase transitions has remained the guiding theme of the theory of random graphs and of probabilistic combinatorics generally. There is by now a vast literature on phase transitions in discrete structures. A lot of this work deals with the component structure of random objects, a question that is known as the percolation problem. Generally speaking this type of question is very well understood by now, at least on models without an underlying finite-dimensional geometry. Many other phase transitions are not understood nearly as well, and it these challenging problems that these notes are about.

1.2. **Random graph colouring.** The single most prominent example is the *random graph colouring* problem. In their 1960 paper [28] the sages posed a number of questions that guided the development of the theory of random graphs for decades. Only a single one of these questions remains open to this day:

given $d > 0$, what is the chromatic number $\chi(\mathbb{G})$?

Recall that the chromatic number of a graph is the least number q of colours that suffice to colour the vertices in such a way that no edge joins two vertices of the same colour.

Based on computer experiments and heuristic arguments we expect that for any $q \geq 3$ there is a *sharp threshold* for q-colorability, i.e., a specific number $d_{\mathrm{col}}(q) > 0$ such that

$$\lim_{n\to\infty} \mathrm{P}\big[\chi(\mathbb{G}) \leq q\big] = \begin{cases} 1 & \text{if } d < d_{\mathrm{col}}(q), \\ 0 & \text{if } d > d_{\mathrm{col}}(q). \end{cases} \tag{1.1}$$

(The case $q = 2$ is special, and much easier—why?) A result that gets close to establishing the existence of such a $d_{\mathrm{col}}(q)$ for which (1.1) holds was proved by Achlioptas and Friedgut [4]. But what might be the value of $d_{\mathrm{col}}(q)$? Let $Z_q(\mathbb{G})$ be the number of q-colourings of $\mathbb{G}$. A simple calculation of the first moment

$$\mathrm{E}[Z_q(\mathbb{G}) \,|\, |E(\mathbb{G})| = m] = \Theta(q^n(1-1/q)^m) \tag{1.2}$$

shows that $d_{\mathrm{col}}(q) \leq (2q-1)\ln q + o_q(1)$. Moreover, Achlioptas and Naor [6] managed to also estimate the second moment $\mathrm{E}[Z_q(\mathbb{G})^2 \,|\, |E(\mathbb{G})| = m]$ to prove that

$$d_{\mathrm{col}}\left(q\right) \geq (2q-2)\ln q + o_q(1), \tag{1.3}$$

leaving a gap of about $\ln q$. Coja-Oghlan and Vilenchik [17, 23] refined these arguments to prove that

$$(2q-1)\ln q - 2\ln 2 + o_q(1) \leq d_{\mathrm{col}}\left(q\right) \geq (2q-2)\ln q - 1 + o_q(1), \tag{1.4}$$

reducing the gap to about $2\ln 2 - 1 \approx 0.39$. But both the upper bound and the lower bound proofs hit the wall at these slightly different values, i.e., they fully exploit the potential of the respective proof methods.

An explanation as to why improving these bounds might be difficult can be put forward on the basis of heuristic but compelling deliberations from statistical physics [42]. For the upper bound the reason is simply that the bound is expected to be tight, up to the $o_q(1)$ error term. Indeed, the physics arguments even allow for an (as yet unproven) prediction as to the exact value of $d_{\mathrm{col}}(q)$ for any $q \geq 3$. The reason why improving the lower bound is difficult is more subtle. Namely, the physics calculations predict that, up to the error term, the lower bound

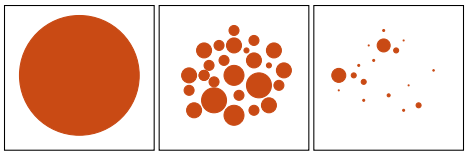

FIGURE 3. The evolution of the set $\mathscr{S}_q(\mathbb{G})$

in (1.4) marks a *second* phase transition in the random graph colouring problem. To be precise, the so-called *condensation phase transition* marks a point where the quantity $d \mapsto \frac{1}{n}\ln Z_q(\mathbb{G})$ is expected to converge to a non-analytic limit, not unlike in the case of Theorem 1.1. Specifically, the physics arguments predict that for a certain $d_{\mathrm{cond}}(q) = (2q-1)\ln q - 2\ln 2 + o_q(1)$ we have, in probability,

$$\lim_{n\to\infty}\frac{1}{n}\ln Z_q(\mathbb{G})\begin{cases}=\ln q+\frac{d}{2}\ln(1-1/q) & \text{if } d<d_{\mathrm{cond}}(q),\\ <\ln q+\frac{d}{2}\ln(1-1/q) & \text{if } d>d_{\mathrm{cond}}(q).\end{cases} \tag{1.5}$$

Thus, up to $d_{\mathrm{cond}}(q)$ the number $Z_q(\mathbb{G})$ takes the largest possible value permitted by the first moment bound (1.2). Mathematically it is not even currently known that the random variables on the left hand side converge to a deterministic limit in probability.

The conjecture (1.5) is based on a hypothetical cartoon of the evolution of the set $\mathscr{S}_q(\mathbb{G})$ of all q-colourings of the random graph. According to this prediction the set $\mathscr{S}_q(\mathbb{G})$ undergoes two fundamental changes. First, for $d < (1+o_q(1))q\ln q$ the set $\mathscr{S}_q(\mathbb{G})$ is well-connected and we may expect Markov chains to succeed in traversing this set rapidly. This regime roughly coincides with the degrees for which efficient algorithms are known to find a q-colouring of the random graph. But for $d > (1+o_q(1))q\ln q$ the set of all q-colourings shatters into an enormous number of well-separated tiny 'clusters', a phenomenon called *dynamic replica symmetry breaking* in physics. We will learn more about this in the subsequent lectures. Furthermore, at the condensation point $d_{\mathrm{cond}}(q)$ the set $\mathscr{S}_q(\mathbb{G})$, although still composed of an enormous number of clusters, is dominated by the few largest ones. The internal structure of the clusters is quite rigid and, e.g., the colours of most vertices are completely determined or 'frozen' throughout the entire cluster. In particular, the expected agreement

$$\alpha(\sigma,\tau) = \max\left\{\frac{q}{(q-1)n}\sum_{i=1}^{n}\mathbf{1}\{\sigma(x_i)=\kappa\circ\tau(x_i)\}:\kappa\in\mathfrak{S}_q\right\}-\frac{1}{q-1}$$

of two randomly chosen colourings within the same cluster is about $n(1-1/q+o_q(1))$ as $n\to\infty$. By contrast, two colourings in different clusters are essentially 'orthogonal', i.e., their expected agreement is $o(1)$.

One of the main results that we will prove in this lecture confirms the existence and precise location of the condensation phase transition as predicted by the statistical physics calculations. To state the theorem we introduce a bit of notation. For an integer $q\geq 1$ let $[q]=\{1,\ldots,q\}$. Moreover, for a finite set Ω let $\mathscr{P}(\Omega)$ be the set of all probability measures on Ω. We tacitly identify $\mathscr{P}(\Omega)$ with the standard simplex in $\mathbb{R}^\Omega$. Further, let $\mathscr{P}^2(\Omega)$ be the set of all probability distribution on $\mathscr{P}(\Omega)$. Additionally, define

$$\mathscr{P}^2_*(\Omega)=\left\{\pi\in\mathscr{P}^2(\Omega):\forall\omega\in\Omega:\int_{\mathscr{P}(\Omega)}\mu(\omega)\mathrm{d}\pi(\mu)=|\Omega|^{-1}\right\}.$$

Finally, let $\Lambda(x)=x\ln x$, with the convention that $\Lambda(0)=0$.

Theorem 1.2 ([20])**.** *Suppose that $q\geq 3$. For $\beta\in[0,\infty]$, $d>0$ and $\pi\in\mathscr{P}^2_*([q])$ let*

$$\mathscr{B}_{q,\beta,d}(\pi)=\mathrm{E}\left[\frac{\Lambda(\sum_{\sigma=1}^q\prod_{i=1}^{\boldsymbol{\gamma}}1-(1-e^{-\beta})\boldsymbol{\mu}_i^{(\pi)}(\sigma))}{q(1-(1-e^{-\beta})/q)^{\boldsymbol{\gamma}}}-\frac{d}{2}\frac{\Lambda(1-(1-e^{-\beta})\sum_{\sigma=1}^q\boldsymbol{\mu}_1^{(\pi)}(\sigma)\boldsymbol{\mu}_2^{(\pi)}(\sigma))}{1-(1-e^{-\beta})/q}\right],$$

where $\boldsymbol{\gamma}=\mathrm{Po}(d)$ and $\boldsymbol{\mu}_1^{(\pi)},\boldsymbol{\mu}_2^{(\pi)},\ldots$ are drawn from π, all mutually independent. Then with

$$d_{\mathrm{cond}}(q)=\inf\left\{d>0:\sup_{\pi\in\mathscr{P}^2_*([q])}\mathscr{B}_{q,\infty,d}(\pi)>\ln q+\frac{d}{2}\ln(1-1/q)\right\}$$

the following two statements are true.

(i) If $d<d_{\mathrm{cond}}(q)$, then $\frac{1}{n}\ln Z_q(\mathbb{G})$ converges to $\ln q+\frac{d}{2}\ln(1-1/q)$ in probability.

(ii) If $d > d_{\mathrm{cond}}(q)$, then $\limsup_{n\to\infty} \frac{1}{n}\mathrm{E}[\ln Z_q(\mathbb{G})] < \ln q + \frac{d}{2}\ln(1-1/q)$.

Theorem 1.2 opens the door to a detailed analysis of the graph colouring problem for $d < d_{\mathrm{cond}}(q)$. For instance, it is possible to derive a precise mathematical version of the 'dynamic replica symmetry breaking' picture and to investigate the agreement between random q-colourings [18]. Furthermore, $d_{\mathrm{cond}}(q)$ is the best current lower bound on the q-colorability threshold for all $q > 3$. Remarkably though, for $q = 3$ a strictly better lower bound can be obtained via the analysis of a colouring algorithm [5].

1.3. **The stochastic block model.** The following inference problem can be viewed as a "noisy" variation of the random graph colouring problem [36]. Suppose that $\boldsymbol{\sigma}^* : V = \{x_1, \ldots, x_n\} \to [q]$ is a random colouring assignment. Let $\mathbb{G}^* = \mathbb{G}^*(n, \beta, \boldsymbol{\sigma}^*)$ be the random graph obtained by connecting any two vertices x_i, x_j, $i \neq j$, with probability

$$p_{ij} = \frac{d}{n} \cdot \frac{q}{q-1+e^{-\beta}} \cdot \exp(-\beta \mathbf{1}\{\boldsymbol{\sigma}^*(x_i) = \boldsymbol{\sigma}^*(x_j)\}) \tag{1.6}$$

independently. Given $\mathbb{G}^*(n, \beta, \boldsymbol{\sigma}^*)$, can we recover $\boldsymbol{\sigma}^*$? If not perfectly, is it at least possible to infer a colouring $\tau_{\mathbb{G}^*}$ such that $\mathrm{E}[\alpha(\boldsymbol{\sigma}^*, \tau_{\mathbb{G}^*})] = \Omega(1)$? Even more modestly, given a graph can we tell whether it was created via the stochastic block model or whether it is a mere Erdős-Rényi graph?

What might we expect? In equation (1.6) we can think of $\exp(-\beta)$ as a noise parameter. Moreover, (1.6) ensures that the expected degree of every vertex is asymptotically equal to d. As d gets larger, more information about the 'ground truth' $\boldsymbol{\sigma}^*$ is revealed. Thus, we can think of d as the signal strength. The natural question therefore is:

for what signal-to-noise ratios is it possible to (approximately) recover the ground truth?

An intriguing conjecture asserts that this question is intimately related to the condensation phase transition [25]. More precisely, given a value of $\beta > 0$ it has been conjectured, once more on the basis of insightful but non-rigorous physics deliberations, that the evolution of the inference problem on $\mathbb{G}^*$ proceeds similarly as in the graph colouring problem. Of course, in this case the structure of interest is not the set of all q-colourings but the posterior distribution $\mathrm{P}[\boldsymbol{\sigma}^* = \sigma | \mathbb{G}^*]$ of the *ground truth* $\boldsymbol{\sigma}^*$ given $\mathbb{G}^*$. Clearly, in the absence of a clean combinatorial structure such as the discrete set $\mathscr{S}_q(\mathbb{G}) \subset \Omega^n$ the question arises how 'clusters' and 'condensation' might be defined mathematically. We will come to that. But the bottom line is that the inference problem is soluble if and only if d exceeds a certain critical value $d_{\mathrm{cond}}(q, \beta)$ corresponding to the condensation threshold. Formally, we have the following result, confirming the conjecture from [25].

Theorem 1.3 ([20])**.** *Suppose that $q \geq 2$, $\beta > 0$ and let*

$$d_{\mathrm{cond}}(q,\beta) = \inf\left\{ d > 0 : \sup_{\pi \in \mathscr{P}_*^2([q])} \mathscr{B}_{q,\beta,d}(\pi) > \ln q + \frac{d}{2}\ln(1-(1-e^{-\beta})/q) \right\}.$$

The following two statements are true.

(i) If $d < d_{\mathrm{cond}}(q)$, then there for any random variable $\mathbb{G}^* \mapsto \tau_{\mathbb{G}^*}$ we have $\mathrm{E}[\alpha(\boldsymbol{\sigma}^*, \tau_{\mathbb{G}^*})] = o(1)$.
(ii) If $d > d_{\mathrm{cond}}(q)$, then there is a random variable $\mathbb{G}^* \mapsto \tau_{\mathbb{G}^*}$ such that $\mathrm{E}[\alpha(\boldsymbol{\sigma}^*, \tau_{\mathbb{G}^*})] = \Omega(1)$.

The value $d_{\mathrm{cond}}(q, \beta)$ is called the *information-theoretic threshold* for the stochastic block model $\mathbb{G}^*$ because it marks the point from where at least approximative inference of the ground truth $\boldsymbol{\sigma}^*$ is possible. Specifically, the random variable from (ii) is to draw a sample $\tau_{\mathbb{G}^*}$ from the posterior distribution of $\boldsymbol{\sigma}^*$ given $\mathbb{G}^*$.

Apart from the information-theoretic threshold itself it might be interesting to find out how much information about $\boldsymbol{\sigma}^*$ is 'stored' in $\mathbb{G}^*$. The following theorem provides the exact answer for all $d > 0$.

Theorem 1.4 ([20])**.** *Suppose that $q \geq 2$, $\beta > 0$. Then the mutual information*

$$I(\mathbb{G}^*, \boldsymbol{\sigma}^*) = \sum_{G,\sigma} \mathrm{P}\left[(\mathbb{G}^*, \boldsymbol{\sigma}^*) = (G, \sigma)\right] \ln \frac{\mathrm{P}\left[(\mathbb{G}^*, \boldsymbol{\sigma}^*) = (G, \sigma)\right]}{\mathrm{P}\left[\mathbb{G}^* = G\right] \mathrm{P}\left[\boldsymbol{\sigma}^* = \sigma\right]}$$

satisfies

$$\lim_{n\to\infty} \frac{1}{n} I(\mathbb{G}^*, \boldsymbol{\sigma}^*) = \ln q - \frac{d}{2} \cdot \frac{\beta e^{-\beta}}{q-1+e^{-\beta}} - \sup_{\pi \in \mathscr{P}_*^2([q])} \mathscr{B}_{q,\beta,d}(\pi) \qquad \text{for all } d > 0.$$

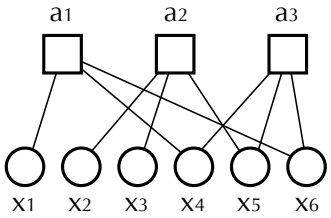

FIGURE 4. A factor graph.

A subtle but not very difficult argument shows that Theorem 1.3 actually follows from Theorem 1.4. We shall therefore concentrate on the proof of Theorem 1.4.

Of course, from a computer science viewpoint a further question arises. Namely, given $\mathbb{G}^*$ can we actually compute a colouring with a strictly positive agreement efficiently, i.e., in polynomial time? According to a further conjecture from [25] there is a second threshold from where such a polynomial time algorithm exists, known as the *Kesten-Stigum bound* [39]

$$d_{\mathrm{KS}}(q,\beta) = \left(\frac{q-1+e^{-\beta}}{1-e^{-\beta}}\right)^2.$$

The positive part of this conjecture has been confirmed rigorously [3], but unsurprisingly the negative bit is wide open. However, for $q=2$, historically the case studied first, it is known that $d_{\mathrm{KS}}(q,\beta) = d_{\mathrm{cond}}(q,\beta)$ [43, 48, 49]. For $q=3$ it is conjectured that $d_{\mathrm{KS}}(q,\beta) = d_{\mathrm{cond}}(q,\beta)$ as well, but for $q>4$ we know that $d_{\mathrm{KS}}(q,\beta) > d_{\mathrm{cond}}(q,\beta)$. For an overview of the enormous literature on the stochastic block model see [1, 47].

1.4. **Factor graphs.** The appropriate mathematical framework for much of the material in these lectures is a comprehensive class of graphical representations of probability distributions known as *factor graph models* [44]. They can be seen as a generalisation of Markov random fields or Bayesian networks. Let $\Omega \neq \emptyset$ be a finite set of 'spins' or 'colours'. An Ω-*factor graph* $G = (V, F, (\partial a)_{a\in F}, (\psi_a)_{a\in F})$ consists of a set V of *variable nodes*, a set F of *constraint nodes*, a subset $\partial a \subset V$ and a *weight function* $\psi_a : \Omega^{\partial a} \to [0,\infty)$ for each $a \in F$. Factor graphs can be represented as bipartite graphs such that the constraint node a is adjacent to the variable nodes ∂a. In particular, we will use standard graph-theoretic concepts such as the distance between vertices (viz. number of edges on a shortest path). But of course the weight functions provide additional information.

The graph colouring problem and the stochastic block model can be represented naturally as factor graph models. Indeed, given a graph $g = (V_g, E_g)$ and a number $q \geq 2$ of colours we can turn g into a factor graph G with variable nodes V_g and constraint nodes E_g where each $e = \{v,w\} \in E_g$ is connected with $\partial e = \{v,w\}$. The weight function $\psi_\infty(\sigma,\tau) = \mathbf{1}\{\sigma \neq \tau\}$ enforces the constraint that the two vertices must be coloured differently. A similar representation is natural in the case of the stochastic block model, except that in this case the weight function

$$\psi_\beta(\sigma,\tau) = \exp(-\beta\mathbf{1}\{\sigma \neq \tau\}) \tag{1.7}$$

seems more natural.

For a factor graph G and an *configuration* $\sigma : V \to \Omega$ it is natural to define the *total weight* as

$$\psi_G(\sigma) = \prod_{a\in F} \psi_a(\sigma|_{\partial a}).$$

Additionally, we define the *partition function*

$$Z(G) = \sum_{\sigma\in\Omega^V} \psi_G(\sigma).$$

Assuming that $Z(G) > 0$ we define a probability measure μ_G on Ω^V by letting

$$\mu_G(\sigma) = \psi_G(\sigma)/Z(G). \tag{1.8}$$

For instance, in the case of the graph colouring problem we have $\psi_G(\sigma) = 1$ iff σ is a proper q-colouring of G. Thus, $Z(G)$ is equal to the total number of q-colourings and μ_G is the uniform distribution on the set of q-colourings.

Random factor graph models naturally describe problems such as random graph colouring or the stochastic block model. Suppose that we are given a prior distribution P on a finite set Ψ of k-ary weight functions $\Omega^k \to$

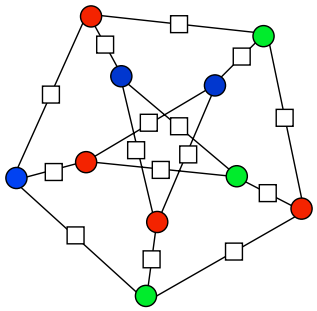

FIGURE 5. Graph colouring via factor graphs.

$[0,\infty)$. Moreover, assume that this prior distribution is invariant under permutations of the coordinates, i.e., for any $\kappa \in \mathfrak{S}_k$ and any $\psi \in \Psi$ we have $P(\psi^\kappa) = P(\psi)$. Then we define the random factor graph $\boldsymbol{G}(n,m) = \boldsymbol{G}(n,m,P)$ with variable nodes $V = \{x_1,\ldots,x_n\}$ and constraint nodes $F = \{a_1,\ldots,a_m\}$ by independently choosing for each variable node a neighbourhood ∂a_i consisting of k (distinct) variable nodes uniformly and a weight function ψ from P. Further, given a number $d > 0$ we let $\boldsymbol{m}$ be a Poisson variable with mean d and we introduce $\boldsymbol{G} = \boldsymbol{G}(n,\boldsymbol{m},P)$. This type of model clearly describes problems such as random graph colouring.

A variant of this construction can be used to describe inference problems. Indeed, given a colouring $\sigma : V \to \Omega$ we introduce the random factor graph $\boldsymbol{G}^*(n,m,P,\sigma)$ with variable nodes V and constraint nodes F with distribution

$$\mathrm{P}\left[\boldsymbol{G}^*(n,m,P,\sigma) = G\right] = \frac{\psi_G(\sigma)\mathrm{P}\left[\boldsymbol{G}(n,m) = G\right]}{\mathrm{E}[\psi_{\boldsymbol{G}(n,m)}(\sigma)]}. \tag{1.9}$$

Thus, we reweigh the prior distribution $\boldsymbol{G}$ according to the weight assigned to σ. The stochastic block model is then (essentially) equivalent to the random graph distribution obtained by choosing $\boldsymbol{\sigma}^* : V \to \Omega$ uniformly and then setting up $\boldsymbol{G}^*(n,\boldsymbol{m},P,\boldsymbol{\sigma}^*)$.

The mutual information from Theorem 1.4 is closely related to the partition function $\ln Z(\boldsymbol{G}^*)$. More precisely, in the case of the stochastic block model a few simple manipulations reveal the following.

Lemma 1.5. *In the case of the stochastic block model we have*

$$\frac{1}{n}I(\boldsymbol{G}^*,\boldsymbol{\sigma}^*) = \ln q - \frac{d}{2}\cdot\frac{\beta e^{-\beta}}{q-1+e^{-\beta}} - \frac{1}{n}\mathrm{E}[\ln Z(\boldsymbol{G}^*)] + o(1).$$

Thus, the proof of Theorem 1.4 boils down to calculating $\mathrm{E}[\ln Z(\boldsymbol{G}^*)]$.

1.5. **The Nishimori identity.** While we are at it we may as well emphasise an important connection between the distributions $\boldsymbol{G}^*$ and $\boldsymbol{G}$. Namely, in analogy to (1.9) we may also define a reweighed distribution on colourings by letting

$$\mathrm{P}\left[\hat{\boldsymbol{\sigma}}_{n,m} = \sigma\right] = \frac{\mathrm{E}[\psi_{\boldsymbol{G}(n,m)}(\sigma)]}{\mathrm{E}[Z(\boldsymbol{G}(n,m))]}.$$

Moreover, we define a further random factor graph model $\hat{\boldsymbol{G}}(n,m) = \hat{\boldsymbol{G}}(n,m,P)$ by letting

$$\mathrm{P}\left[\hat{\boldsymbol{G}}(n,m) = G\right] = \frac{\mathrm{P}\left[\boldsymbol{G}(n,m) = \boldsymbol{G}\right]Z(G)}{\mathrm{E}[Z(\boldsymbol{G}(n,m))]}. \tag{1.10}$$

Thus, we reweigh the graphs according to their partition function. As before we let $\hat{\boldsymbol{\sigma}} = \hat{\boldsymbol{\sigma}}_{n,\boldsymbol{m}}$ and $\hat{\boldsymbol{G}} = \hat{\boldsymbol{G}}(n,\boldsymbol{m})$. With this notation we obtain the following *Nishimori identity*.

Proposition 1.6 ([20])**.** *Suppose that $\psi(\sigma) > 0$ for all $\psi \in \Psi$, $\sigma \in \Omega^k$. Then for all n,m we have*

$$\mathrm{P}\left[\hat{\boldsymbol{G}}(n,m) = G\right]\mu_G(\sigma) = \mathrm{P}\left[\hat{\boldsymbol{\sigma}}_{n,m} = \sigma\right]\mathrm{P}\left[\boldsymbol{G}^*(n,m,\sigma) = G\right].$$

Proof. By construction,

$$\mathrm{P}\left[\hat{\boldsymbol{G}}(n,m)=G\right]\mu_G(\sigma)=\frac{\mathrm{P}\left[\boldsymbol{G}(n,m)=G\right]Z(G)}{\mathrm{E}[Z(\boldsymbol{G}(n,m))]}\cdot\frac{\psi_G(\sigma)}{Z(G)}=\frac{\mathrm{E}[\psi_{\boldsymbol{G}(n,m)}(\sigma)]}{\mathrm{E}[Z(\boldsymbol{G}(n,m))]}\cdot\frac{\mathrm{P}\left[\boldsymbol{G}(n,m)=G\right]\psi_G(\sigma)}{\mathrm{E}[\psi_{\boldsymbol{G}(n,m)}(\sigma)]}$$
$$=\mathrm{P}\left[\hat{\boldsymbol{\sigma}}_{n,m}=\sigma\right]\mathrm{P}\left[\boldsymbol{G}^*(n,m,\sigma)=G\right],$$

as desired. □

In the case of the stochastic block model a simple calculation reveals that $\hat{\boldsymbol{\sigma}}_{n,m}$ and the uniform distribution $\boldsymbol{\sigma}^*$ are *mutually contiguous*, i.e., for any sequence of events $(\mathscr{A}_n)_{n\geq 1}$ we have

$$\lim_{n\to\infty}\mathrm{P}\left[\boldsymbol{\sigma}^*\in\mathscr{A}_n\right]=0\text{ iff }\lim_{n\to\infty}\mathrm{P}\left[\hat{\boldsymbol{\sigma}}_{n,m}\in\mathscr{A}_n\right]=0.$$

Thus, for the purposes of the present lecture the stochastic block model is identical to the random factor graph model $\boldsymbol{G}^*(\hat{\boldsymbol{\sigma}})$.

What is the connection between the stochastic block model and the random graph colouring problem? Consider the random factor graph model $\boldsymbol{G}$ with the weight function (1.7). For $\beta=\infty$ the corresponding partition function $Z(\boldsymbol{G})$ is equal to the number of k-colourings. For $\beta<\infty$ the partition function $Z(\boldsymbol{G})$ is greater, but clearly

$$\lim_{\beta\to\infty}\ln Z(\boldsymbol{G})=\ln Z_q(\mathbb{G}). \tag{1.11}$$

Thus, we might attempt to prove Theorem 1.2 by way of studying $Z(\boldsymbol{G})$ for $\beta<\infty$. The following lemma shows that this comes down to studying $\mathbb{G}^*$.

Lemma 1.7. *For any $q\geq 3$, $\beta>0$ the following is true. Assume that $d<d_{\mathrm{cond}}(q,\beta)$. Then*

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{E}[\ln Z(\boldsymbol{G})]=\ln q+\frac{d}{2}\ln(1-(1-e^{-\beta})/q). \tag{1.12}$$

Conversely, (1.12) holds only if $d<d_{\mathrm{cond}}(q,\beta)$.

Proof. The proof is based on a *second moment* calculation. Specifically, fix a small $\varepsilon>0$ and define

$$\mathscr{Z}(G)=Z(G)\mathbf{1}\{n^{-1}\ln Z(G)\leq\ln q+d\ln(1-(1-e^{-\beta})/q)/2+\varepsilon\}.$$

Then (1.10) implies that

$$\mathrm{E}[\mathscr{Z}(\boldsymbol{G}(n,m))]=\mathrm{E}[Z(\boldsymbol{G}(n,m))]\cdot\mathrm{P}\left[n^{-1}\ln Z(\hat{\boldsymbol{G}}(n,m))\leq\ln q+d\ln(1-(1-e^{-\beta})/q)/2+\varepsilon\right].$$

Furthermore, the proof of Theorem 1.2 will reveal that for $\frac{2}{m}/n\sim d<d_{\mathrm{cond}}$ we have

$$\mathrm{P}\left[n^{-1}\ln Z(\hat{\boldsymbol{G}}(n,m))\leq\ln q+d\ln(1-(1-e^{-\beta})/q)/2+\varepsilon\right]\sim 1.$$

Since $n^{-1}\ln\mathrm{E}[Z(\boldsymbol{G}(n,m))]\sim\ln q+d\ln(1-(1-e^{-\beta})/q)/2$, we conclude that

$$\mathrm{E}[Z(\boldsymbol{G}(n,m))]=\exp(o(n))q^n(1-(1-e^{-\beta})/q)^m,\qquad\mathrm{E}[Z(\boldsymbol{G}(n,m))^2]=\exp(\varepsilon n+o(n))\mathrm{E}[Z(\boldsymbol{G}(n,m))]^2$$

if $d<d_{\mathrm{cond}}$. The first assertion therefore follows from the Paley-Zygmund inequality. The second assertion requires a subtle application of Azuma's inequality (details omitted). □

Many of the ideas and techniques from these lectures generalise to other random factor graph models [11, 22, 20]. This leads to important applications. For instance, error correcting codes such as low-density parity check codes or low-density generator matrix codes can be cast as random factor graph models [46].

2. The Boltzmann distribution

2.1. **A whiff of statistical physics.** Following [40], in this section we give a very brief (and non-rigorous) introduction to statistical mechanics. The aim of statistical physics is to understand how the macroscopic behaviour of a system $\mathscr{S}$ arises out of the interactions of its microscopic particles. Key concepts of statistical physics are temperature and entropy.

Suppose that we aim to study a large closed system with n elementary particles; for a physical system we would realistically have $n\approx 10^{23}$. Each of these particles can be in one of a finite number of elementary quantum states. At a given overall energy U there would thus be a (*very* large) number g of possible or accessible configurations describing the quantum states of all the particles. Think of $g\approx\exp(10^{23})$. The fundamental assumption that underpins statistical mechanics is that at equilibrium, *the system is equally likely to be in any of the g configurations.*

We refer to the quantity $S = \ln g$ as the *entropy*. Furthermore, we consider the uniform distribution $\mu = \mu_{\mathscr{S},U}$ on the set $\Gamma = \Gamma(n,U)$ of all g configurations and we write $\boldsymbol{\sigma} = \boldsymbol{\sigma}_\mu$ for a random configuration. For a random variable X we write

$$\langle X(\boldsymbol{\sigma})\rangle = \big\langle X(\boldsymbol{\sigma}_\mu)\big\rangle_\mu = \sum_{\sigma\in\Gamma} X(\sigma)\mu(\sigma).$$

Now suppose that we have systems $\mathscr{S}_1, \mathscr{S}_2$ of sizes n_1, n_2 with energies U_1, U_2. What happens if we bring these two systems into 'thermal contact'? That is, we enable the systems to exchange energy but not particles. Upon contact the two systems form a larger one $\mathscr{S} = \mathscr{S}_1 \oplus \mathscr{S}_2$. Since the total energy $U = U_1 + U_2$ of the two systems is preserved, the overall number of possible configurations comes to

$$g(n,U) = \sum_{U_1+U_2=U} g_1(n_1,U_1)g_2(n_2,U_2). \tag{2.1}$$

In a physical system the energy U will realistically be on the order of $n = n_1 + n_2$, while $g(n,U)$ will be exponential in n. Therefore, the sum (2.1) is dominated by the values $\hat{U}_1, \hat{U}_2$ that render the largest contribution, i.e.,

$$g(n,U) \approx \max_{U_1+U_2=U} g_1(n_1,U_1)g_2(n_2,U_2). \tag{2.2}$$

In fact, the energies $U_1(\boldsymbol{\sigma}), U_2(\boldsymbol{\sigma})$ of a random sample from $\mu_{\mathscr{S},U}$ will typically be very tightly concentrated (say, within $O(\sqrt{n})$). Thus, if we pretend that U is a continuous quantity, then the product rule gives

$$\mathrm{d}g = \frac{\partial g_1}{\partial U_1} g_2 \mathrm{d}U_1 + g_1 \frac{\partial g_2}{\partial U_2}\mathrm{d}U_2 = 0 \qquad \text{and} \qquad \mathrm{d}U_1 + \mathrm{d}U_2 = 0.$$

Hence, $\frac{\partial \ln g_1}{\partial U_1} = \frac{\partial \ln g_2}{\partial U_2}$, which we can rewrite in terms of the entropy as

$$\frac{\partial S_1}{\partial U_1} = \frac{\partial S_2}{\partial U_2}. \tag{2.3}$$

We therefore define the *temperature* T of a system by

$$\frac{1}{T} = \frac{\partial S}{\partial U} \tag{2.4}$$

Thus, while S is a pure number, the temperature comes in the units of energy. This is one variant of the *first law of thermodynamics*. In the situation of the two systems $\mathscr{S}_1, \mathscr{S}_2$ in thermal contact, (2.3) and (2.4) show that at the equilibrium we have $T_1 = T_2$, i.e., the temperatures of the two sub-systems equalise. Moreover, (2.2) expresses the fact that the entropy of the combined system is at least as large as the sum of the entropies of the sub-systems before they entered into thermal contact. This is a version of the *second law of thermodynamics*.

What if we bring a system $\mathscr{S}$ into contact with a *much* bigger system $\mathscr{R}$, called the *reservoir*? Suppose that σ_1, σ_2 are two configurations of $\mathscr{S}$ with energies $E(\sigma_1), E(\sigma_2)$. Then by our fundamental assumption, within the system $\mathscr{R}\oplus\mathscr{S}$ we find

$$\frac{\mu_{\mathscr{R}\oplus\mathscr{S},U}(\sigma_1)}{\mu_{\mathscr{R}\oplus\mathscr{S},U}(\sigma_2)} = \frac{g_{\mathscr{R}}(U-E(\sigma_1))}{g_{\mathscr{R}}(U-E(\sigma_2))} = \exp\left(S_{\mathscr{R}}(U-E(\sigma_1)) - S_{\mathscr{R}}(U-E(\sigma_2))\right). \tag{2.5}$$

Assuming that $\mathscr{R}$ is much bigger than $\mathscr{S}$, we can reasonably write the approximation

$$S_{\mathscr{R}}(U-E(\sigma_1)) - S_{\mathscr{R}}(U-E(\sigma_2)) \approx S_{\mathscr{R}}(U) - \frac{\partial S_{\mathscr{R}}}{\partial U}(E(\sigma_1)-E(\sigma_2)) \approx S_{\mathscr{R}}(U) - \frac{E(\sigma_1)-E(\sigma_2)}{T}. \tag{2.6}$$

Combining (2.5) and (2.6) , we find

$$\frac{\mu_{\mathscr{R}\oplus\mathscr{S},U}(\sigma_1)}{\mu_{\mathscr{R}\oplus\mathscr{S},U}(\sigma_2)} = \frac{\exp(-E(\sigma_1)/T)}{\exp(-E(\sigma_2)/T)}.$$

Any specifics of the reservoir cancelled out! Its only role is to maintain $\mathscr{S}$ at a given temperature T ("heat bath"). Hence, writing $\mu = \mu_{\mathscr{R}\oplus\mathscr{S}}$ and introducing the *partition function*

$$Z = Z_{\mathscr{S}}(T) = \sum_\sigma \exp(-E(\sigma)/T),$$

we find

$$\mu(\sigma) = \exp(-E(\sigma)/T)/Z. \tag{2.7}$$

Finally, we introduce the *Helmholtz free energy*

$$F = U - ST \tag{2.8}$$

and for notational convenience we introduce $\beta = 1/T$, the *inverse temperature*.

2.2. **Boltzmann distributions of factor graphs.** If we compare equations (1.8) and (2.7), then we see that the distribution induced by the factor graph is nothing but the Boltzmann distribution with

$$-E(\sigma)/T = \sum_{a\in F} \ln \psi_a(\sigma|_{\partial a}).$$

The temperature parameter T is obsolete, viz. can be incorporated into the definition of the ψ_a; we therefore just set $T = 1$. We hence call μ_G the *Boltzmann distribution* of G. The average with respect to μ_G is denoted by $\langle\cdot\rangle_G$. Furthermore, we define by analogy with (2.8),

$$U(G) = -\sum_{a\in F}\left\langle \ln\psi_a\right\rangle_G, \qquad S(G) = H(\mu_G) = -\left\langle \ln\mu_G(\boldsymbol{\sigma})\right\rangle_G, \qquad F(G) = U(G) - S(G).$$

Lemma 2.1. *We have* $F(G) = -\ln Z(G)$.

Proof. For any configuration $\sigma \in \Omega^V$ we have $\mu_G(\sigma) = \frac{1}{Z}\prod_{a\in F}\psi_a(\sigma|_{\partial a})$ and thus

$$\ln Z = -\ln\mu_G(\sigma) + \sum_{a\in F}\ln\psi_a(\sigma|_{\partial a}).$$

Averaging this equation over σ chosen from μ_G yields the assertion. □

Let us now glimpse at a simple example. In the *Curie-Weiss model* we have n variable nodes $V = \{x_1,\ldots,x_n\}$ that take values in the set $\Omega = \{\pm 1\}$. There is a constraint node a_{ij} associated with any pair $1 \le i < j \le n$ of variable nodes, and one factor node a_i associated with each variable x_i. Thus, there are $n(n+1)/2$ factor nodes in total. The weight functions satisfy

$$\ln\psi_{a_{ij}}(\sigma_i,\sigma_j) = \sigma_i\sigma_j/(Tn), \qquad \ln\psi_{a_i}(\sigma_i) = B\sigma_i/T,$$

where $T > 0$ and $B \ge 0$ are real parameters. Let

$$h(z) = -z\ln z - (1-z)\ln(1-z).$$

Theorem 2.2. *For* $\lambda \in [-1,1]$ *let*

$$\phi(\lambda,T,B) = -\frac{1-\lambda^2}{2T} + \frac{B\lambda}{T} + h\left(\frac{1+\lambda}{2}\right).$$

Then

$$\lim_{n\to\infty}\frac{1}{n}\ln Z = \phi(T,B) = \max_{\lambda\in[-1,1]}\phi(\lambda,T,B).$$

Proof. For a configuration $\sigma = (\sigma_1,\ldots,\sigma_n) \in \{\pm 1\}^n$ we define the *magnetisation*

$$\lambda(\sigma) = \frac{1}{n}\sum_{i=1}^n \sigma_i.$$

The key observation is that

$$T\cdot U(\sigma) = \frac{1}{2}n - \frac{1}{2}n\lambda(\sigma)^2 - nB\lambda(\sigma).$$

Therefore, summing over the (finite number of) possible empirical magnetisations λ, we find

$$Z = \exp(-n/(2T))\sum_{\lambda}\binom{n}{n(1+\lambda)/2}\exp\left[\frac{n/T}{2}\cdot\lambda^2 + nB\lambda/T\right]. \tag{2.9}$$

The binomial coefficient $\binom{n}{n(1+\lambda)/2}$ simply enumerates the number of σ with $\lambda(\sigma) = \lambda$. By Stirling's formula,

$$\frac{1}{n}\ln\binom{n}{n(1+\lambda)/2} \sim h((1+\lambda)/2).$$

Hence,

$$Z = \exp(o(n))\sum_{\lambda}\exp(n\phi(\lambda,T,B)). \tag{2.10}$$

Since there are $O(n)$ summands in (2.10) and because $\lambda \mapsto \phi(\lambda, T, B)$ is continuous, we see that

$$\frac{1}{n} \ln Z \sim \max_{\lambda \in [-1,1]} \phi(\lambda, T, B),$$

as claimed. □

Let

$$\mathscr{M}_n(T, B) = \frac{1}{n} \sum_{i=1}^{n} \langle \boldsymbol{\sigma}(x_i) \rangle \sim T \frac{\partial}{\partial B} \phi(T, B)$$

and

$$\mathscr{M}_+(T) = \lim_{B \searrow 0} \lim_{n \to \infty} \mathscr{M}_n(T, B) = \lim_{B \searrow 0} T \frac{\partial}{\partial B} \phi(T, B).$$

Corollary 2.3. *For $T > 1$ we have $\mathscr{M}_+(T) = 0$. By contrast, $\mathscr{M}_+(T) > 0$ for $T < 1$. In particular, there is a phase transition at $T = 1$.*

Proof. Suppose first that $B = 0$. Then the function $\lambda \mapsto \phi(\lambda, T, B)$ is symmetric in λ, and thus its differential at $\lambda = 0$ vanishes. But where the function attains its global maximum depends on T. More precisely, for $T > 1$, the global maximum is attained at $\lambda = 0$. Indeed, the differentials of $\phi(\lambda, \beta, B)$ are

$$\frac{\partial}{\partial \lambda} \phi(\lambda, T, 0) \quad = \quad \lambda / T - \frac{\ln(1+\lambda) - \ln(1-\lambda)}{2}, \qquad \frac{\partial^2}{\partial \lambda^2} \phi(\lambda, T, 0) = 1 - \frac{1}{1-\lambda^2}.$$

Thus, for $T > 1$ the function is concave, with its maximum at $\lambda = 0$. However, for $T < 1$ the global maximum is attained at another point $\lambda_+(T) > 0$ (and, by symmetry, at its mirror image $\lambda_-(T) < 0$), while there is a local minimum at $\lambda = 0$.

In effect, for $T > 1$ we have $\mathscr{M}_+(T) = 0$, because for sufficiently small $B > 0$, the maximum of $\phi(\lambda, T, B)$ is going to remain at $\lambda = 0$. By contrast, for $T > 1$ we have that $\mathscr{M}_+(T) > 0$ (because $B > 0$ rules out $\lambda_-(T)$ as a global maximum). □

Observe that the Curie-Weiss model with $B = 0$ is symmetric, i.e., $\mathscr{E}(\sigma) = \mathscr{E}(-\sigma)$ for every configuration σ. Thus, if $T < 1$ and we imagine many separate, isolated copies of the system, so-called *replicas*, then we should expect that in about half the replicas the magnetisation is negative and in the others positive. In other words, the replicas are not symmetric. This is reflected in the following observation. In each replica the average magnetisation satisfies $\langle \boldsymbol{\sigma}(x_i) \rangle = 0$ for every variable node x_i. Yet because $\mathscr{M}_+(T) > 0$ we have $\langle \boldsymbol{\sigma}(x_i) \boldsymbol{\sigma}(x_j) \rangle > 0$ for all i, j, i.e., pairs of spins are positively correlated. In particular, for any $T < 1$ there is $\varepsilon(T) > 0$ such that for large enough n,

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \left\| \mu_{x_i, x_j} - \mu_{x_i} \otimes \mu_{x_j} \right\|_{\mathrm{TV}} \geq \varepsilon(T),$$

where μ_{x_i, x_j} is the joint distribution of the spins of x_i, x_j and μ_{x_i}, μ_{x_j} are the marginal distributions. On the other hand, once we condition, say, on the average magnetisation $\lambda(\boldsymbol{\sigma})$ being positive, asymptotic pairwise independence is recovered:

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \left\| \mu_{x_i, x_j}[\,\cdot\,|\lambda(\boldsymbol{\sigma}) > 0] - \mu_{x_i}[\,\cdot\,|\lambda(\boldsymbol{\sigma}) > 0] \otimes \mu_{x_j}[\,\cdot\,|\lambda(\boldsymbol{\sigma}) > 0] \right\|_{\mathrm{TV}} = o(1).$$

Similarly,

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \left\| \mu_{x_i, x_j}[\,\cdot\,|\lambda(\boldsymbol{\sigma}) < 0] - \mu_{x_i}[\,\cdot\,|\lambda(\boldsymbol{\sigma}) < 0] \otimes \mu_{x_j}[\,\cdot\,|\lambda(\boldsymbol{\sigma}) < 0] \right\|_{\mathrm{TV}} = o(1).$$

We might thus call $\{\lambda(\boldsymbol{\sigma}) > 0\}$ and $\{\lambda(\boldsymbol{\sigma}) < 0\}$ two *pure states* of the system. By contrast, for $T > 1$ there is but one single pure state, namely the paramagnetic one where $\lambda(\boldsymbol{\sigma}) = o(n)$.

Remark 2.4. *There is a "spin glass" variant of the Curie-Weiss model. In the* Sherrington-Kirkpatrick model *we introduce couplings $\boldsymbol{J} = (J_{ij})_{1 \leq i < j \leq N}$ that are mutually independent standard Gaussians. This gives rise to the energy function*

$$\mathscr{E}(\sigma) = -\frac{1}{\sqrt{n}} \sum_{1 \leq i < j \leq n} J_{ij} \sigma_i \sigma_j - B \sum_{i=1}^{n} \sigma_i.$$

This innocuous change makes the problem much *more complicated. The free energy in this model was first calculated in a celebrated paper by Talagrand* [57].

3. Pure states

3.1. **Discrete probability distributions.** This section is based on [11]. The main result of this section shows that *any* probability measure on a discrete cube Ω^n can be partitioned into a bounded number of pure states. Let us introduce some terminology. Throughout we assume that $\Omega \neq \emptyset$ is a finite set. Let $\mu \in \mathscr{P}(\Omega^n)$ be a probability distribution on the discrete cube Ω^n. For a finite set $I \subset [n]$ we let μ_I denote the joint distribution of the coordinates $i \in I$. Thus, $\mu_I \in \mathscr{P}(\Omega^I)$ is defined by

$$\mu_I(\sigma) = \sum_{\tau \in \Omega^n} \mathbf{1}\{\forall i \in I : \tau_i = \sigma_i\}\mu(\tau).$$

An (ε, k)*-pure state* of μ is a set $S \subset \Omega^n$ such that $\mu(S) > 0$ and

$$\frac{1}{n^k} \sum_{x_1,\ldots,x_k \in [n]} \left\| \mu_{x_1,\ldots,x_k}[\cdot|S] - \mu_{x_1}[\cdot|S] \otimes \cdots \otimes \mu_{x_k}[\cdot|S] \right\|_{\mathrm{TV}} < \varepsilon.$$

Furthermore, we call μ (ε, k)*-symmetric* if $S = \Omega^n$ itself is an (ε, k)-state. When k is omitted it is understood that we mean $k = 2$. The main result of this section is

Theorem 3.1 ([11])**.** *For any $\varepsilon > 0$, $k \geq 2$ there exists $\eta = \eta(\varepsilon, k, \Omega) > 0$ such that for every $n > 1/\eta$ any measure $\mu \in \mathscr{P}(\Omega^n)$ has pairwise disjoint (ε, k)-pure states $S_1, \ldots, S_N$ such that $\mu(S_i) \geq \eta$ for all $i \in [N]$ and $\sum_{i=1}^N \mu(S_i) \geq 1 - \varepsilon$.*

The following consequence of Theorem 3.1 shows that the parameter k is relatively unimportant.

Corollary 3.2 ([11])**.** *For any $\varepsilon > 0$, $k \geq 3$ there exists $\delta > 0$ such that for all $n > 1/\delta$ and all $\mu \in \mathscr{P}(\Omega^n)$ the following is true. If μ is $(\delta, 2)$-symmetric, then μ is (ε, k)-symmetric.*

In the example of the Curie-Weiss model we saw that the balance between the two pure states for $T < 1$ is extremely delicate. Indeed, Corollary 2.3 demonstrates that we can tip the balance by exposing an arbitrarily weak external field. The following result, which is a generalisation of a result from [46], shows that this is a universal fact.

Theorem 3.3 ([20])**.** *For any $\varepsilon > 0$ there is $T = T(\varepsilon, \Omega) > 0$ such that for every $n > T$ and every probability measure $\mu \in \mathscr{P}(\Omega^n)$ the following is true. Obtain a random probability measure $\check{\boldsymbol{\mu}} \in \mathscr{P}(\Omega^n)$ as follows.*

> *Draw a sample $\check{\boldsymbol{\sigma}} \in \Omega^n$ from μ, independently choose a number $\boldsymbol{\theta} \in (0, T)$ uniformly at random, then obtain a random set $\boldsymbol{U} \subset [n]$ by including each $i \in [n]$ with probability $\boldsymbol{\theta}/n$ independently and let*

$$\check{\boldsymbol{\mu}}(\sigma) = \frac{\mu(\sigma)\mathbf{1}\{\forall i \in \boldsymbol{U} : \sigma_i = \check{\boldsymbol{\sigma}}_i\}}{\mu(\{\tau \in \Omega^n : \forall i \in \boldsymbol{U} : \tau_i = \check{\boldsymbol{\sigma}}_i\})} \qquad (\sigma \in \Omega^n).$$

Then $\check{\boldsymbol{\mu}}$ is ε-symmetric with probability at least $1 - \varepsilon$.

The rest of this section deals with the proof of Theorems 3.1. The proof of Corollary 3.2 can be found in [11] and that of Theorem 3.3 in [20].

3.2. **Homogeneity.** This section introduces a key concept upon which the proof of Theorem 3.1 is based. If $\boldsymbol{V} = (V_1, \ldots, V_k)$ is a partition of some set V, then we call $\#\boldsymbol{V} = k$ the *size* of $\boldsymbol{V}$. Furthermore, a partition $\boldsymbol{W} = (W_1, \ldots, W_l)$ *refines* another partition $\boldsymbol{V} = (V_1, \ldots, V_k)$ if for each $i \in [l]$ there is $j \in [k]$ such that $W_i \subset V_j$. Moreover, for a configuration $\sigma \in \Omega^n$ and a subset $U \subset [n]$ we define

$$\sigma[\omega|U] = |U|^{-1} \sum_{i \in U} \mathbf{1}\{\sigma_i = \omega\}.$$

Thus, $\sigma[\cdot|U]$ is the empirical distribution of the spins on U.

For $\varepsilon > 0$ we say that $\mu \in \mathscr{P}(\Omega^n)$ is ε*-regular* on a set $U \subset [n]$ if for every subset $W \subset U$ of size $|W| \geq \varepsilon|U|$ we have

$$\langle \|\boldsymbol{\sigma}[\cdot|W] - \boldsymbol{\sigma}[\cdot|U]\|_{\mathrm{TV}} \rangle_\mu < \varepsilon.$$

Further, μ is ε*-regular* with respect to a partition $\boldsymbol{V}$ if there is a set $J \subset [\#\boldsymbol{V}]$ such that $\sum_{i \in [\#\boldsymbol{V}] \setminus J} |V_i| < \varepsilon n$ and such that μ is ε-regular on V_i for all $i \in J$. Additionally, if $\boldsymbol{V}$ is a partition of $[n]$ and $\boldsymbol{S}$ is a partition of Ω^n, then we say that μ is ε*-homogeneous* with respect to $(\boldsymbol{V}, \boldsymbol{S})$ if there is a subset $I \subset [\#\boldsymbol{S}]$ such that the following is true.

- **HM1:** We have $\mu(S_i) > 0$ for all $i \in I$ and $\sum_{i \in [\#\boldsymbol{S}] \setminus I} \mu(S_i) < \varepsilon$.
- **HM2:** for all $i \in [\#\boldsymbol{S}]$ and $j \in [\#\boldsymbol{V}]$ we have $\max_{\sigma,\sigma' \in S_i} \left\| \sigma[\cdot|V_j] - \sigma'[\cdot|V_j] \right\|_{\mathrm{TV}} < \varepsilon$.
- **HM3:** for all $i \in I$ the measure $\mu[\cdot|S_i]$ is ε-regular with respect to $\boldsymbol{V}$.
- **HM4:** μ is ε-regular with respect to $\boldsymbol{V}$.

Theorem 3.4 ([11]). *For any $\varepsilon > 0$ there exists $N = N(\varepsilon, \Omega) > 0$ such that for every $n > N$, any measure $\mu \in \mathscr{P}(\Omega^n)$ and any partition $\boldsymbol{V}_0$ of $[n]$ of size $\#\boldsymbol{V}_0 \leq 1/\varepsilon$ the following is true. There exist a refinement $\boldsymbol{V}$ of $\boldsymbol{V}_0$ and a partition $\boldsymbol{S}$ of Ω^n such that $\#\boldsymbol{V} + \#\boldsymbol{S} \leq N$ and such that μ is ε-homogeneous with respect to $(\boldsymbol{V}, \boldsymbol{S})$.*

Informally speaking, Theorem 3.4 shows that any probability measure $\mu \in \mathscr{P}(\Omega^n)$ admits a partition $(\boldsymbol{V}, \boldsymbol{S})$ such that the following is true. Almost the entire probability mass of μ belongs to parts S_i such that the conditional measure $\mu[\cdot|S_i]$ is ε-regular w.r.t. $\boldsymbol{V}$. This means that almost every coordinate $x \in [n]$ belongs to a class V_j such that for every "large" $U \subset V_j$ for $\boldsymbol{\sigma}$ chosen from $\mu[\cdot|S_i]$ very likely the empirical distribution $\boldsymbol{\sigma}[\cdot|U]$ is close to the marginal distribution $\langle \boldsymbol{\sigma}[\cdot|V_j] \rangle_{\mu[\cdot|S_i]}$ of the entire class.

Theorem 3.4 and its proof are inspired by Szemerédi's regularity lemma, a result about the decomposition of graphs.

4. Belief Propagation

4.1. **Message passing on factor graphs.** Suppose that $G = (V, F, (\partial a)_{a \in F}, (\psi_a)_{a \in F})$ is an Ω-factor graph. Further, define the *message space* $\mathscr{M}(G)$ as the set of all families $\nu = (\nu_{x \to a}, \nu_{a \to x})_{x \in V, a \in F, x \in \partial a}$ such that $\nu_{x \to a}, \nu_{a \to x} \in \mathscr{P}(\Omega)$. The *Belief Propagation operator* $\mathrm{BP} = \mathrm{BP}_G : \mathscr{M}(G) \to \mathscr{M}(G)$ maps $\nu \in \mathscr{M}(G)$ to the point $\hat{\nu} \in \mathscr{M}(G)$ defined by

$$\hat{\nu}_{a \to x}(\sigma_x) = \frac{\sum_{\tau \in \Omega^{\partial a}} \mathbf{1}\{\tau_x = \sigma_x\} \psi_a(\tau) \prod_{y \in \partial a \setminus x} \nu_{y \to a}(\sigma_y)}{\sum_{\tau \in \Omega^{\partial a}} \psi_a(\tau) \prod_{y \in \partial a \setminus x} \nu_{y \to a}(\sigma_y)}, \qquad \hat{\nu}_{x \to a}(\sigma_x) = \frac{\prod_{b \in \partial x \setminus a} \hat{\nu}_{b \to x}(\sigma_x)}{\sum_{\tau \in \Omega} \prod_{b \in \partial x \setminus a} \hat{\nu}_{b \to x}(\tau)}.$$

Furthermore, for a point $\nu \in \mathscr{M}(G)$ and a variable node x and $\sigma \in \Omega$ we define

$$\nu_x(\sigma) = \frac{\prod_{b \in \partial x} \hat{\nu}_{b \to x}(\sigma_x)}{\sum_{\tau \in \Omega} \prod_{b \in \partial x} \hat{\nu}_{b \to x}(\tau)}.$$

Similarly, for a constraint node a we define $\nu_a \in \mathscr{P}(\Omega^{\partial a})$ by

$$\nu_a(\sigma) = \frac{\psi_a(\sigma) \prod_{y \in \partial a} \nu_{y \to a}(\sigma_y)}{\sum_{\tau \in \Omega^{\partial a}} \psi_a(\tau) \prod_{y \in \partial a} \nu_{y \to a}(\tau_y)}.$$

Additionally, we define a functional $\mathscr{B} : \mathscr{M}(G) \to \mathbb{R}$ called the *Bethe free entropy* by letting

$$\mathscr{B}(\nu) = \sum_{x \in V} H(\nu_x) + \sum_{a \in F} \left[D_{\mathrm{KL}}(\nu_a \| \otimes_{x \in \partial a} \nu_x) + \langle \ln \psi_a \rangle_{\nu_a} \right] \tag{4.1}$$

Of course, the formula can be applied to any family $(\nu_x, \nu_a)_{x,a}$ of marginal distributions.

Proposition 4.1 ([58]). *The stationary points $(\nu_x, \nu_a)_{x,a}$ of $\mathscr{B}$ that satisfy the following consistency condition:*

$$\forall x \in V, a \in \partial x, \sigma \in \Omega : \sum_{\tau \in \Omega^{\partial a}} \mathbf{1}\{\tau(x) = \sigma\} \nu_a(\tau) = \nu_x(\sigma) \tag{4.2}$$

and the Belief Propagation fixed points are in one-to-one correspondence.

4.2. **Acyclic graphs.** Throughout this section we assume that the factor graph G is a acyclic. Let $V = \{x_1, \ldots, x_n\}$. The ultimate goal is to devise a formalism for computing the $\frac{1}{n} \ln Z$.

For starters, we devise an approach to compute the marginals μ_{x_i}. Recall that the *diameter* of a graph is the maximum distance between any two vertices, where "distance" refers to the number of edges on the shortest path.

Theorem 4.2. *In a tree factor graph of diameter t^* the following is true.*

(1) There exists $\nu^* \in \mathscr{M}(G)$ such that $\mathrm{BP}(\nu^*) = \nu^*$.
(2) For any $\nu^0 \in \mathscr{M}(G)$ and any $t > t^*$ we have $\mathrm{BP}^t(\nu^0) = \nu^*$.
(3) For any $x \in V$ we have $\nu^*_x = \mu_x$.

Theorem 4.2 shows that on an acyclic factor graph Belief Propagation has exactly one fixed point, which we can determine from any starting point by iterating the BP operator more than t^* times.

To prove Theorem 4.2, we introduce a bit of notation. For a variable x and $a \in \partial x$ we denote by $G_{x \to a}$ the subgraph of the factor graph obtained by removing a. Let $V_{x \to a}$ be the set of variable nodes and let $F_{x \to a}$ be the set of constraint nodes in $G_{x \to a}$. Furthermore, let $\mu_{x \to a}(\cdot)$ denote the marginal of x in $G_{x \to a}$. Analogously, if a is a factor node and $x \in \partial a$, then we denote by $G_{a \to x}$ the subgraph of G obtained by removing all $b \in \partial x \setminus a$. Let $\mu_{a \to x}$ signify the Boltzmann marginal of x in $G_{a \to x}$.

Lemma 4.3. *In a tree factor graph $(\mu_{x \to a}, \mu_{a \to x})_{a \in \partial x}$ is a Belief Propagation fixed point.*

Proof. Consider a constraint node a and $x \in \partial a$. Then by the definition of $\mu_{x\to a}$,

$$\frac{\mu_{x\to a}(\sigma)}{\mu_{x\to a}(\sigma')} = \frac{\sum_{\tau\in\Omega^V}\mathbf{1}\{\tau(x)=\sigma\}\psi_{G_{x\to a}}(\tau)}{\sum_{\tau\in\Omega^V}\mathbf{1}\{\tau(x)=\sigma'\}\psi_{G_{x\to a}}(\tau)}.$$

Similarly, if $b \in \partial x \setminus a$, then

$$\frac{\mu_{b\to x}(\sigma)}{\mu_{b\to x}(\sigma')} = \frac{\sum_{\tau\in\Omega^V}\mathbf{1}\{\tau(x)=\sigma\}\psi_{G_{b\to x}}(\tau)}{\sum_{\tau\in\Omega^V}\mathbf{1}\{\tau(x)=\sigma'\}\psi_{G_{b\to x}}(\tau)}.$$

The last expression can be simplified as follows. Let $H_{b\to x}$ be the component of b in $G_{b\to x}$. Then we can write

$$\frac{\mu_{b\to x}(\sigma)}{\mu_{b\to x}(\sigma')} = \frac{\sum_{\tau\in\Omega^{V(H_{b\to x})}}\mathbf{1}\{\tau(x)=\sigma\}\psi_{H_{b\to x}}(\tau)}{\sum_{\tau\in\Omega^{V(H_{b\to x})}}\mathbf{1}\{\tau(x)=\sigma'\}\psi_{H_{b\to x}}(\tau)}.$$

Furthermore, because G is a tree the components $H_{b\to x}$ only have the node x in common. Therefore,

$$\frac{\mu_{x\to a}(\sigma)}{\mu_{x\to a}(\sigma')} = \frac{\prod_{b\in\partial a\setminus x}\sum_{\tau\in\Omega^{V(H_{b\to x})}}\mathbf{1}\{\tau(x)=\sigma\}\psi_{H_{b\to x}}(\tau)}{\prod_{b\in\partial a\setminus x}\sum_{\tau\in\Omega^{V(H_{b\to x})}}\mathbf{1}\{\tau(x)=\sigma'\}\psi_{H_{b\to x}}(\tau)} = \prod_{b\in\partial x\setminus a}\frac{\mu_{b\to x}(\sigma)}{\mu_{b\to x}(\sigma')}.$$

We apply a similar argument to the constraint-to-variable messages. Thus, for $y \in \partial a\setminus x$ let $H_{y\to a}$ be the component of y in $G_{y\to a}$. Then by construction,

$$\begin{aligned}\frac{\mu_{a\to x}(\sigma)}{\mu_{a\to x}(\sigma')} &= \frac{\sum_{\tau\in\Omega^V}\mathbf{1}\{\tau(x)=\sigma\}\psi_{G_{a\to x}}(\tau)}{\sum_{\tau\in\Omega^V}\mathbf{1}\{\tau(x)=\sigma'\}\psi_{G_{a\to x}}(\tau)}\\ &= \frac{\sum_{\tau\in\Omega^{\partial a}}\mathbf{1}\{\tau(x)=\sigma\}\psi_a(\tau)\prod_{y\in\partial a\setminus x}\mu_{y\to a}(\sigma_y)Z(H_{y\to a})}{\sum_{\tau\in\Omega^{\partial a}}\mathbf{1}\{\tau(x)=\sigma'\}\psi_a(\tau)\prod_{y\in\partial a\setminus x}\mu_{y\to a}(\sigma_y)Z(H_{y\to a})}\\ &= \frac{\sum_{\tau\in\Omega^{\partial a}}\mathbf{1}\{\tau(x)=\sigma\}\psi_a(\tau)\prod_{y\in\partial a\setminus x}\mu_{y\to a}(\sigma_y)}{\sum_{\tau\in\Omega^{\partial a}}\mathbf{1}\{\tau(x)=\sigma'\}\psi_a(\tau)\prod_{y\in\partial a\setminus x}\mu_{y\to a}(\sigma_y)},\end{aligned}$$

as desired. □

Proof of Theorem 4.2. Let x be a variable node and let $a \in \partial x$. The *depth* $t_{x\to a}$ is the maximum distance between x and a leaf of the component of x in $G_{x\to a}$. We may assume without loss that all leaves of the factor graph are variable nodes. Let $\nu^0 \in \mathscr{M}(G)$ be any starting point and let $\nu^t = \mathrm{BP}^t(\nu^0)$. We claim that for any $x \in V$, $a \in \partial x$ we have

$$\mu_{x\to a} = \nu^{(t)}_{x\to a} \quad \text{if } t > t_{i\to a}. \tag{4.3}$$

The proof is by induction on $t_{x\to a}$. If $t_{x\to a} = 0$, then x is a leaf of G. In particular, $\partial x = \{a\}$. Hence, by the construction of the Belief Propagation operator $\nu^t_{x\to a}$ is just the uniform distribution for any $t \geq 1$. So is $\mu_{x\to a}$. Thus, we have got (4.3) in the case $t_{x\to a} = 0$. The inductive step follows from Lemma 4.3 because both $\mu_{x\to a}$ and $\nu_{x\to a}$ are defined by the same recurrence. Of course, (4.3) implies that $(\mu_{x\to a}, \mu_{a\to x})_{a\in\partial x}$ is the unique fixed point of the Belief propagation operator. This establishes the first and the second assertion.

To obtain the third assertion, pick a variable node x. We modify the factor by adding another factor node a^* such that $\partial a^* = \{x\}$ and $\psi_{a^*} \equiv 1$. Then $\mu_{x\to a^*}$ coincides with the Boltzmann distribution on the original factor graph. Moreover, due to our choice of ψ_{a^*} the addition of the factor node a^* does not alter the Belief Propagation fixed point. Therefore, the last assertion follows from (4.3) applied to $\mu_{x\to a^*}$. □

In fact, we can express the *entire* distribution μ merely in terms of the marginal distributions μ_x, $\mu_{\partial a}$. The proof of the following fact is based on a simple induction, not unlike the proof of Theorem 4.2.

Corollary 4.4. *If G is a tree, then*

$$\mu(\sigma) = \prod_{a\in F}\mu_{\partial a}(\sigma|_{\partial a})\prod_{x\in V}\mu_x(\sigma(x))^{1-|\partial x|}.$$

Recall the Bethe functional from (4.1).

Corollary 4.5. *If G is acyclic, then* $\ln Z(G) = \mathscr{B}((\mu_x)_{x\in V}, (\mu_{\partial a})_{a\in F})$.

Proof. This is immediate from Corollaries 2.1 and 4.4. □

We make a note of the following alternative form of the Bethe formula.

Corollary 4.6. *Assume that G is a tree. Set*

$$\mathscr{B}_G = \sum_{a\in F}\mathscr{B}_a + \sum_{x\in V}\mathscr{B}_x - \sum_{x\in V, a\in\partial x}\mathscr{B}_{a,x}, \qquad \text{where}$$
$$\mathscr{B}_a = \ln\sum_{\tau\in\Omega^{\partial a}}\psi_a(\tau)\prod_{x\in\partial a}\mu_{x\to a}(\sigma(x)), \qquad \mathscr{B}_x = \ln\sum_{\sigma\in\Omega}\prod_{b\in\partial x}\mu_{b\to x}(\sigma), \qquad \mathscr{B}_{a,x} = \ln\sum_{\sigma\in\Omega}\mu_{x\to a}(\sigma)\mu_{a\to x}(\sigma).$$

Then $\ln Z = \mathscr{B}_G$.

4.3. **Belief Propagation and replica symmetry.** We recall the random factor graph model $\boldsymbol{G}$ from Section 1.4. Let us call this model *replica symmetric* if

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i,j=1}^{n}\mathrm{E}\left\|\mu_{\boldsymbol{G}_n,x_i,x_j} - \mu_{\boldsymbol{G}_n,x_i}\otimes\mu_{\boldsymbol{G}_n,x_j}\right\|_{\mathrm{TV}} = 0 \tag{4.4}$$

An important conjecture holds that (4.4) is sufficient for the fact that Belief Propagation and the Bethe formula can be used to calculate the free energy [42]. In this section we prove this conjecture. For a given factor graph G we call the messages $\mu_{G,\cdot\to\cdot} = (\mu_{G,x\to a},\mu_{G,a\to x})_{x\in V(G),a\in F(G),x\in\partial a}$ an *ε-Belief Propagation fixed point* on G if

$$\sum_{\substack{x\in V(G)\\ a\in\partial x\\ \sigma\in\Omega}}\left|\mu_{G,x\to a}(\sigma) - \frac{\prod_{b\in\partial x\setminus a}\mu_{G,b\to x}(\sigma)}{\sum_{\tau\in\Omega}\prod_{b\in\partial x\setminus a}\mu_{G,b\to x}(\tau)}\right| + \left|\mu_{G,a\to x}(\sigma) - \frac{\sum_{\tau\in\Omega^{\partial a}}\mathbf{1}\{\tau_x=\sigma\}\psi_a(\tau)\prod_{y\in\partial a\setminus x}\mu_{G,y\to a}(\tau_y)}{\sum_{\tau\in\Omega^{\partial a}}\psi_a(\tau)\prod_{y\in\partial a\setminus x}\mu_{G,y\to a}(\tau_y)}\right| < \varepsilon n.$$

Assume that all factor graphs have strictly positive weight functions from a fixed finite set Ψ.

Theorem 4.7 ([22])**.** *If (4.4) holds, then there is a sequence $(\varepsilon_n)_n\to 0$ such that $\mu_{\boldsymbol{G}_n,\cdot\to\cdot}$ is an ε_n-Belief Propagation fixed point w.h.p.*

We remember the formulas from Corollary 4.6.

Corollary 4.8 ([22])**.** *If (4.4) holds and $\frac{1}{n}\mathscr{B}_{\boldsymbol{G}_n}$ converges to a real number B in probability, then*

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{E}[\ln Z_{\boldsymbol{G}}] = B.$$

The rest of this section follows [22]. To prove Theorem 4.7 we need a few preparations. The proof is based on the following lemma, which follows from Theorem 3.1.

Lemma 4.9 ([22])**.** *For any integer $L>0$ and any $\alpha>0$ there exist $\varepsilon=\varepsilon(\alpha,L,\Psi)>0$, $n_0=n_0(\varepsilon,L)$ such that the following is true. Suppose that G is a factor graph with $n>n_0$ variable nodes. Moreover, assume that μ_G is $(\varepsilon,2)$-symmetric. If G^+ is obtained from G by adding L constraint nodes $b_1,\ldots,b_L$ with weight functions $\psi_{b_1},\ldots,\psi_{b_L}\in\Psi$ arbitrarily, then μ_{G^+} is $(\alpha,2)$-symmetric and*

$$\sum_{x\in V(G)}\left\|\mu_{G,x}-\mu_{G^+,x}\right\|_{\mathrm{TV}} < \alpha n. \tag{4.5}$$

Proof of Theorem 4.7. Fix $\varepsilon>0$, choose $L=L(\varepsilon)>0$ and $\gamma=\gamma(\varepsilon,L,\Psi)>\eta=\eta(\gamma)>\delta=\delta(\eta)>0$ small enough and assume that $n>n_0(\delta)$ is sufficiently large. Because the distribution of the random factor graph $\boldsymbol{G}_n$ is symmetric under permutations of the variable nodes, it suffices to prove that with probability at least $1-\varepsilon$ we have

$$\sum_{a\in\partial x_n,\sigma\in\Omega}\left|\mu_{\boldsymbol{G}_n,x_n\to a}(\sigma) - \frac{\prod_{b\in\partial x_n\setminus a}\mu_{\boldsymbol{G}_n,b\to x_n}(\sigma)}{\sum_{\tau\in\Omega}\prod_{b\in\partial x_n\setminus a}\mu_{\boldsymbol{G}_n,b\to x_n}(\tau)}\right| < \varepsilon \qquad \text{and} \tag{4.6}$$

$$\sum_{a\in\partial x_n,\sigma\in\Omega}\left|\mu_{\boldsymbol{G}_n,a\to x_n}(\sigma) - \frac{\sum_{\tau\in\Omega^{\partial a}}\mathbf{1}\{\tau(x_n)=\sigma\}\psi_a(\tau)\prod_{y\in\partial a\setminus x_n}\mu_{\boldsymbol{G}_n,y\to a}(\tau_y)}{\sum_{\tau\in\Omega^{\partial a}}\psi_a(\tau)\prod_{y\in\partial a\setminus x_n}\mu_{\boldsymbol{G}_n,y\to a}(\tau_y)}\right| < \varepsilon. \tag{4.7}$$

To prove (4.6)–(4.7) we use the following standard trick. Let $\boldsymbol{G}'$ be the random factor graph with variable nodes $x_1,\ldots,x_n$ comprising of $m'=\mathrm{Po}(dn(1-1/n)^k/k)$ random constraint nodes $a_1,\ldots,a_{m'}$ that do not contain x_n. Moreover, let $\Delta=\mathrm{Po}(dn(1-(1-1/n)^k)/k)$ be independent of m' and obtain $\boldsymbol{G}''$ from $\boldsymbol{G}'$ by adding independent random constraint nodes $b_1,\ldots,b_\Delta$ with $x_n\in\partial b_i$ for all $i\in[\Delta]$. Then the random factor graph $\boldsymbol{G}''$ has precisely the same distribution as $\boldsymbol{G}_n$. Therefore, it suffices to verify (4.6)–(4.7) with $\boldsymbol{G}_n$ replaced by $\boldsymbol{G}''$.

Since $dn(1-(1-1/n)^k)/k = d+o(1)$, we can choose $L=L(\varepsilon)$ so large that

$$\mathrm{P}[\Delta>L] < \varepsilon/3. \tag{4.8}$$

Furthermore, $\boldsymbol{G}'$ is distributed precisely as the random factor graph $\boldsymbol{G}_n$ given that $\partial x_n = \emptyset$. Therefore, Bayes' rule and our assumption (4.4) imply

$$\mathrm{P}\left[\boldsymbol{G}' \text{ fails to be } (\delta,2)\text{-symmetric}\right] < \delta, \tag{4.9}$$

provided that n_0 is chosen large enough. Combining (4.9) and Corollary 3.2, we see that

$$\mathrm{P}\left[\boldsymbol{G}' \text{ is } (\eta, 2+(k-1)L)\text{-symmetric} | \Delta \le L\right] > 1-\delta, \tag{4.10}$$

provided δ is sufficiently small.

Due to (4.8) and (4.10) and the symmetry amongst $b_1,\ldots,b_\Delta$ we just need to prove the following: given that $\boldsymbol{G}'$ is $(\eta, 2+(k-1)L)$-symmetric and $0<\Delta\le L$, with probability at least $1-\varepsilon/L$ we have

$$\sum_{\sigma\in\Omega}\left|\mu_{\boldsymbol{G}'',x_n\to b_1}(\sigma) - \frac{\prod_{i=2}^{\Delta}\mu_{\boldsymbol{G}'',b_i\to x_n}(\sigma)}{\sum_{\tau\in\Omega}\prod_{i=2}^{\Delta}\mu_{\boldsymbol{G}'',b_i\to x_n}(\tau)}\right| < \varepsilon/L \qquad \text{and} \tag{4.11}$$

$$\sum_{\sigma\in\Omega}\left|\mu_{\boldsymbol{G}'',b_1\to x_n}(\sigma) - \frac{\sum_{\tau\in\Omega^{\partial b_1}}\mathbf{1}\{\tau(x_n)=\sigma\}\psi_{b_1}(\tau)\prod_{y\in\partial b_1\setminus x_n}\mu_{\boldsymbol{G}_n,y\to b_1}(\tau(y))}{\sum_{\tau\in\Omega^{\partial b_1}}\psi_a(\tau)\prod_{y\in\partial b_1\setminus x_n}\mu_{\boldsymbol{G}_n,y\to b_1}(\tau(y))}\right| < \varepsilon/L. \tag{4.12}$$

To this end, let $U=\bigcup_{j\ge2}\partial b_j$ be the set of all variable nodes that occur in the constraint nodes $b_2,\ldots,b_\Delta$. Because $\mu_{\boldsymbol{G}'',x_n\to b_1}$ is the marginal of x_n in the factor graph $\boldsymbol{G}''-b_1$, the definition of the Boltzmann distribution entails that for any $\sigma\in\Omega$,

$$\mu_{\boldsymbol{G}'',x_n\to b_1}(\sigma) \propto \sum_{\tau\in\Omega^U}\mathbf{1}\{\tau(x_n)=\sigma\}\left\langle\mathbf{1}\{\forall y\in U\setminus\{x_n\}:\boldsymbol{\sigma}(y)=\tau(y)\right\rangle_{\mu_{\boldsymbol{G}'}}\prod_{j=2}^{\Delta}\psi_{b_j}(\tau|_{\partial b_j}). \tag{4.13}$$

Similarly, because $\mu_{\boldsymbol{G}'',b_i\to x_n}$ is the marginal of x_n in $\boldsymbol{G}'+b_i$, we have

$$\mu_{\boldsymbol{G}'',b_i\to x_n}(\sigma) \propto \sum_{\tau\in\Omega^{\partial b_i}}\mathbf{1}\{\tau(x_n)=\sigma\}\left\langle\mathbf{1}\{\forall y\in\partial b_i\setminus\{x_n\}:\boldsymbol{\sigma}(y)=\tau(y)\right\rangle_{\mu_{\boldsymbol{G}'}}\psi_{b_i}(\tau). \tag{4.14}$$

To prove (4.11), recall that the variable nodes $\partial b_j\setminus x_n$ are chosen uniformly and independently for each $j\ge2$. Therefore, if $\boldsymbol{G}'$ is $(\eta,(k-1)L)$-symmetric and $0<\Delta\le L$, then

$$\sum_{\tau\in\Omega^U}\mathrm{E}\left[\left|\left\langle\mathbf{1}\{\forall y\in U\setminus\{x_n\}:\boldsymbol{\sigma}(y)=\tau(y)\right\rangle_{\mu_{\boldsymbol{G}'}} - \prod_{y\in U}\mu_{\boldsymbol{G}',y}(\tau(y))\right|\middle|\boldsymbol{G}'\right] \le 2\eta. \tag{4.15}$$

Set

$$\nu_i(\sigma) = \sum_{\tau\in\Omega^{\partial b_i}}\mathbf{1}\{\tau(x_n)=\sigma\}\psi_{b_i}(\tau)\prod_{y\in\partial b_i\setminus x_n}\mu_{\boldsymbol{G}',y}(\tau(y)). \tag{4.16}$$

W.h.p. for any $1\le i<j\le\Delta$ we have $\partial b_i\cap\partial b_j=\{x_n\}$. Hence, assuming that $\eta=\eta(\gamma)$ is chosen small enough, we obtain from (4.13), (4.14), (4.15) and Markov's inequality that with probability at least $1-\gamma$,

$$\left|\mu_{\boldsymbol{G}'',x_n\to b_1}(\sigma) - \frac{\prod_{i=2}^{\Delta}\nu_i(\sigma_x)}{\sum_{\tau\in\Omega}\prod_{i=2}^{\Delta}\nu_i(\tau)}\right| < \gamma \qquad \text{and} \qquad \left|\mu_{\boldsymbol{G}'',b_i\to x_n}(\sigma) - \frac{\nu_i(\sigma)}{\sum_{\tau\in\Omega}\nu_i(\tau)}\right| < \gamma \qquad \text{for all } i\in[\Delta]. \tag{4.17}$$

Hence, (4.11) follows from (4.17), provided that γ is chosen small enough. A similar argument yields (4.12). □

The proof of Corollary 4.8 is based on the so-called *Aizenman-Sims-Starr scheme* [7], i.e., the observation that

$$\lim_{n\to\infty}\mathrm{E}\left[\ln\frac{Z_{\boldsymbol{G}_n}}{Z_{\boldsymbol{G}_{n-1}}}\right] = \lim_{n\to\infty}\frac{1}{n}\mathrm{E}[\ln Z(\boldsymbol{G}_n)]. \tag{4.18}$$

To calculate the quantity on the l.h.s. we set up a coupling of the factor graphs $\boldsymbol{G}_{n-1}$ and $\boldsymbol{G}_n$. Roughly speaking, the latter is obtained from the former by adding one variable node along with a few adjacent constraint node. (Strictly speaking, we also need to delete a few randomly chosen constraint nodes.) In order to control the effect of these modifications we use similar arguments as in the proof of Theorem 4.7.

5. The interpolation method

Recall the factor graph interpretation $\hat{\boldsymbol{G}}$ of the stochastic block model. In this section we will prove the following.

Proposition 5.1. *For* any $\pi \in \mathscr{P}$,

$$\frac{1}{n}\mathrm{E}[\ln Z(\hat{\boldsymbol{G}})] \geq \mathscr{B}(d,\pi) + o(1). \tag{5.1}$$

The proof of (5.1) is based on the *interpolation method*, originally invented by Guerra [34] in order to study spin glasses. The idea is that for a given $\pi \in \mathscr{P}_*^2(\Omega)$ we set up a family of random factor graph models parametrised by $t \in [0,1]$ such that the free energy of the $t = 0$ model is easily seen to be $-n\mathscr{B}(d,\pi) + o(n)$ and such that the $t = 1$ model is identical to $\hat{\boldsymbol{G}}$. Finally, we will show that the derivative of the free energy with respect to t is non-positive, whence (5.1) follows.

5.1. **The interpolation scheme.** To construct the intermediate models let $\gamma = (\gamma_v)_{v\in[n]}$ be a sequence of integers. Fix $\pi \in \mathscr{P}_*^2(\Omega)$. We define a random factor graph model $\boldsymbol{G} = \boldsymbol{G}(n,m,\gamma,\pi)$ as follows.

- **G1:** the variable nodes are $V = \{x_1,\ldots,x_n\}$.
- **G2:** there are binary constraint nodes $a_1,\ldots,a_m$; for each $i \in [m]$ independently choose $\partial a_i \in V^k$ uniformly and let $\psi_{a_i} = \psi$ be the weight function from (1.7).
- **G3:** for each $x \in V$ there are unary constraint nodes $b_{x,1},\ldots,b_{x,\gamma_x}$ adjacent to x whose weight functions are generated as follows: for each $j \in [\gamma_x]$ independently, pick $\boldsymbol{\mu}_{x,j}$ from π and let

$$\psi_{b_{x,j}} : \sigma \in \Omega \mapsto \sum_{\tau\in\Omega} \psi(\sigma,\tau)\boldsymbol{\mu}_{x,j}(\tau).$$

We recall these random factor graph models induce a few further distributions. First, the Boltzmann measure of G is

$$\mu_G(\sigma) = \frac{\psi_G(\sigma)}{Z(G)} \quad \text{with} \qquad \psi_G : \sigma \in \Omega^V \mapsto \prod_{i=1}^{m}\psi(\sigma)\prod_{x\in V}\prod_{j=1}^{\gamma_v}\psi_{b_{x,j}}(\sigma(v)), \qquad Z(G) = \sum_{\sigma\in\Omega^V}\psi_G(\sigma).$$

We also obtain a reweighed version $\hat{\boldsymbol{G}}(n,m,\gamma,\pi)$ by letting

$$\mathrm{P}\left[\hat{\boldsymbol{G}}(n,m,\gamma,\pi) \in \mathscr{A}\right] = \frac{\mathrm{E}[Z(\boldsymbol{G}(n,m,\gamma,\pi))\mathbf{1}\{\boldsymbol{G}(n,m,\gamma,\pi)\in\mathscr{A}\}]}{\mathrm{E}[Z(\boldsymbol{G}(n,m,\gamma,\pi))]} \qquad \text{for any event } \mathscr{A}.$$

Further, there is an induced distribution $\hat{\boldsymbol{\sigma}}_{n,m,\gamma,\pi}$ on assignments defined by

$$\mathrm{P}\left[\hat{\boldsymbol{\sigma}}_{n,m,\gamma,\pi} = \sigma\right] = \mathrm{E}[\psi_{\boldsymbol{G}(n,m,\gamma,\pi)}(\sigma)]/\mathrm{E}[Z(\boldsymbol{G}(n,m,\gamma,\pi))]. \tag{5.2}$$

Finally, each assignment σ induces a distribution $\boldsymbol{G}^*(n,m,\gamma,\pi,\sigma)$ on factor graphs by letting

$$\mathrm{P}\left[\boldsymbol{G}^*(n,m,\gamma,\pi,\sigma)\in\mathscr{A}\right] = \frac{\mathrm{E}\left[\psi_{\boldsymbol{G}(n,m,\gamma,\pi)}(\sigma)\mathbf{1}\{\boldsymbol{G}(n,m,\gamma,\pi)\in\mathscr{A}\}\right]}{\mathrm{E}[\psi_{\boldsymbol{G}(n,m,\gamma,\pi)}(\sigma)]} \qquad \text{for any event } \mathscr{A}.$$

We are ready to set up the interpolation scheme. Given $d > 0$, $t \in [0,1]$ we let $\boldsymbol{m}_t = \mathrm{Po}(tdn/k)$. Moreover, for each $x \in V$ independently we let $\boldsymbol{\gamma}_{t,x} = \mathrm{Po}((1-t)d)$. Let $\boldsymbol{\gamma}_t = (\boldsymbol{\gamma}_{t,x})_{x\in V}$. Finally, let

$$\hat{\boldsymbol{G}}_t = \hat{\boldsymbol{G}}(n,\boldsymbol{m}_t,\boldsymbol{\gamma}_t,\pi).$$

Then $\hat{\boldsymbol{G}}_1$ is identical to our original factor graph model. Moreover, all constraint nodes of $\hat{\boldsymbol{G}}_0$ are unary; in other words, each connected component of $\hat{\boldsymbol{G}}_0$ contains just a single variable node. The construction of $\hat{\boldsymbol{G}}_t$ is an adaptation of the interpolation schemes from [29, 52].

Finally, we need a correction term. Letting $c = 1 - e^{-\beta}$, we also introduce

$$\xi = q^{-2}\sum_{\tau\in\Omega^2}\psi(\tau) = q^{-2}(q^2 - cq) = 1 - c/q.$$

Let us observe that we can write

$$1 - \psi(\sigma,\tau) = c\mathbf{1}\{\sigma = \tau\}.$$

Further, let

$$\Gamma_t = \frac{td}{2\xi}\mathrm{E}\left[\Lambda\left(\sum_{\tau\in\Omega^2}\psi(\tau)\prod_{j=1}^{2}\boldsymbol{\mu}_j^{(\pi)}(\tau_j)\right)\right].$$

The following is the centrepiece of the interpolation argument.

Proposition 5.2. *For every $\varepsilon > 0$ for all large enough n the following is true. Let*

$$\phi : t \in [0,1] \mapsto (\mathrm{E}[\ln Z(\hat{\boldsymbol{G}}_t)] + \Gamma_t)/n.$$

Then $\phi'(t) > -\varepsilon$ for all $t \in [0,1]$.

We proceed to prove Proposition 5.2. We write $\langle\cdot\rangle_t$ for the expectation with respect to the Boltzmann measure of $\hat{\boldsymbol{G}}_t$ and denotes independent samples by $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \ldots$.

Proposition 5.3. *With $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_2$ chosen uniformly from the set of variable nodes and $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ chosen from π, all mutually independent and independent of $\hat{\boldsymbol{G}}_t$, let*

$$\Xi_{t,l} = \mathrm{E}\left[\left\langle \mathbf{1}\{\boldsymbol{\sigma}(\boldsymbol{y}_1) = \boldsymbol{\sigma}(\boldsymbol{y}_2)\}\right\rangle_t^l\right] - 2\mathrm{E}\left[\left\langle \boldsymbol{\mu}_1(\boldsymbol{\sigma}(\boldsymbol{y}_1))\right\rangle_t^l\right] + \mathrm{E}\left[\left(\sum_{\tau\in\Omega}\prod_{j=1}^{2}\boldsymbol{\mu}_j(\tau)\right)^l\right].$$

Then uniformly for all $t \in (0,1)$,

$$\frac{\partial}{\partial t}\phi(t) = o(1) + \frac{d}{2\xi}\sum_{l\geq 2}\frac{c^l\Xi_{t,l}}{l(l-1)}.$$

We proceed to prove Proposition 5.3. Let

$$\Delta_t = \mathrm{E}\left[\ln Z(\hat{\boldsymbol{G}}_t(\boldsymbol{m}_t+1, \boldsymbol{\gamma}_t))\right] - \mathrm{E}\left[\ln Z(\hat{\boldsymbol{G}}_t(\boldsymbol{m}_t, \boldsymbol{\gamma}_t))\right],$$
$$\Delta_t' = \frac{1}{n}\sum_{x\in V}\mathrm{E}\left[\ln Z(\hat{\boldsymbol{G}}_t(\boldsymbol{m}_t, \boldsymbol{\gamma}_t + \mathbf{1}_x))\right] - \mathrm{E}\left[\ln Z(\hat{\boldsymbol{G}}_t(\boldsymbol{m}_t, \boldsymbol{\gamma}_t))\right].$$

Lemma 5.4. *We have $\frac{1}{n}\frac{\partial}{\partial t}\mathrm{E}[\ln Z(\hat{\boldsymbol{G}}_t)] = \frac{d}{k}\Delta_t - d\Delta_t'$.*

Proof. We just need to compute the derivative of the generating function of the Poisson distribution. □

Lemma 5.5. *We have* $\Delta_t = o(1) - \dfrac{1-\xi}{\xi} + \dfrac{1}{n^2\xi}\displaystyle\sum_{y_1,y_2\in V}\sum_{l\geq 2}\frac{c^l}{l(l-1)}\mathrm{E}\left\langle\prod_{h=1}^{l}\mathbf{1}\{\boldsymbol{\sigma}_h(y_1), \boldsymbol{\sigma}_h(y_2)\}\right\rangle_{T,t}.$

Proof. A mildly delicate calculation shows that the two assignments $\hat{\boldsymbol{\sigma}}' = \hat{\boldsymbol{\sigma}}_{n,m,\boldsymbol{\gamma}_t,\boldsymbol{m}_t}, \hat{\boldsymbol{\sigma}}'' = \hat{\boldsymbol{\sigma}}_{n,m,\boldsymbol{\gamma}_t,\boldsymbol{m}_t+1}$ can be coupled so that

$$\mathrm{P}\left[\hat{\boldsymbol{\sigma}}' = \hat{\boldsymbol{\sigma}}''\right] = 1 - \tilde{O}(n^{-1}), \qquad \mathrm{P}\left[|\hat{\boldsymbol{\sigma}}' \triangle \hat{\boldsymbol{\sigma}}''| > \sqrt{n}\ln n\right] = O(n^{-2}). \tag{5.3}$$

Let us assume that indeed $\hat{\boldsymbol{\sigma}}' = \hat{\boldsymbol{\sigma}}''$; the other case merely contributed to the $o(1)$ error term. Then we can couple $\hat{\boldsymbol{G}}_{T,t}(n, \boldsymbol{m}_t, \boldsymbol{\gamma}_t, \pi)$, $\hat{\boldsymbol{G}}_{T,t}(n, \boldsymbol{m}_t+1, \boldsymbol{\gamma}_t, \pi)$ as follows. Define $\boldsymbol{G}'$ as the outcome of **INT1–INT4** with $\check{\boldsymbol{\sigma}} = \hat{\boldsymbol{\sigma}}' = \hat{\boldsymbol{\sigma}}''$. Then obtain $\boldsymbol{G}''$ by adding one single binary constraint node $\boldsymbol{a}$ such that $\partial\boldsymbol{a}$ has distribution

$$\mathrm{P}\left[\partial\boldsymbol{a} = (x_{i_1}, x_{i_2})\right] \propto \psi(\hat{\boldsymbol{\sigma}}'(x_{i_2}), \hat{\boldsymbol{\sigma}}'(x_{i_2})). \tag{5.4}$$

Hence,

$$\mathrm{E}\left[\ln Z(\hat{\boldsymbol{G}}_t(n, \boldsymbol{m}_t+1, \boldsymbol{\gamma}_t, \pi))\right] - \mathrm{E}\left[\ln Z(\hat{\boldsymbol{G}}_t(n, \boldsymbol{m}_t, \boldsymbol{\gamma}_t, \pi))\right] = \mathrm{E}\left[\ln\frac{Z(\boldsymbol{G}'')}{Z(\boldsymbol{G}')}\right] + o(1) = \mathrm{E}\left[\ln\left\langle\psi_{\boldsymbol{a}}(\boldsymbol{\sigma}_{\boldsymbol{G}'})\right\rangle_{\boldsymbol{G}'}\right] + o(1). \tag{5.5}$$

Writing $\boldsymbol{\sigma}, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \ldots$ for independent samples from $\mu_{\boldsymbol{G}'}$ and plugging in the definition (5.4) of $\boldsymbol{a}$, we find

$$\mathrm{E}\left[\ln\left\langle\psi_{\boldsymbol{a}}(\boldsymbol{\sigma}_{\boldsymbol{G}'})\right\rangle_{\boldsymbol{G}'}\right] = \frac{\sum_{y_1,y_2\in V}\mathrm{E}\left[\psi(\hat{\boldsymbol{\sigma}}'(y_1), \hat{\boldsymbol{\sigma}}'(y_2))\ln\left\langle\psi(\boldsymbol{\sigma}(y_1), \boldsymbol{\sigma}(y_2))\right\rangle_{\boldsymbol{G}'}\right]}{\sum_{y_1,\ldots,y_k\in V}\mathrm{E}\left[\psi(\hat{\boldsymbol{\sigma}}'(y_1), \hat{\boldsymbol{\sigma}}'(y_2))\right]}.$$

Since the empirical distribution of $\hat{\boldsymbol{\sigma}}'$ is asymptotically uniform with very high probability, the denominator in the above expression equals $n^2(\xi + o(1))$ with probability $1 - O(n^{-2})$. Thus,

$$\mathrm{E}\left[\ln\left\langle\psi_{\boldsymbol{a}}(\boldsymbol{\sigma}_{\boldsymbol{G}'})\right\rangle_{\boldsymbol{G}'}\right] = o(1) + \frac{1}{n^2\xi}\sum_{y_1,y_2\in V}\mathrm{E}\left[\psi(\hat{\boldsymbol{\sigma}}'(y_1), \hat{\boldsymbol{\sigma}}'(y_2))\ln\left\langle\psi(\boldsymbol{\sigma}(y_1), \boldsymbol{\sigma}(y_2))\right\rangle_{\boldsymbol{G}'}\right]. \tag{5.6}$$

Expanding the logarithm gives

$$\ln\left\langle\psi(\boldsymbol{\sigma}(y_1), \boldsymbol{\sigma}(y_2))\right\rangle_{\boldsymbol{G}'} = -\sum_{l\geq 1}\frac{1}{l}\left\langle 1 - \psi(\boldsymbol{\sigma}(y_1), \boldsymbol{\sigma}(y_2))\right\rangle_{\boldsymbol{G}'}^l = -\sum_{l\geq 1}\frac{1}{l}\left\langle\prod_{h=1}^{l} 1 - \psi(\boldsymbol{\sigma}_h(y_1), \boldsymbol{\sigma}_h(y_2))\right\rangle_{\boldsymbol{G}'};$$

the second equality sign holds because $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \ldots$ are mutually independent. Combining the last two equations, we obtain

$$\begin{aligned}\mathrm{E}\left[\ln\left\langle\psi_{\boldsymbol{a}}(\boldsymbol{\sigma}_{\boldsymbol{G}'})\right\rangle_{\boldsymbol{G}'}\right] &= o(1) - \sum_{l\geq 1}\sum_{y_1,y_2}\mathrm{E}\left[\frac{\psi(\hat{\boldsymbol{\sigma}}'(y_1),\hat{\boldsymbol{\sigma}}'(y_2))}{ln^2\xi}\left\langle\prod_{h=1}^{l}1-\psi(\boldsymbol{\sigma}_h(y_1),\boldsymbol{\sigma}_h(y_2))\right\rangle_{\boldsymbol{G}'}\right]\\ &= o(1) + \sum_{l\geq 1}\frac{1}{ln^2\xi}\sum_{y_1,y_2}\mathrm{E}\left[(1-\psi(\hat{\boldsymbol{\sigma}}'(y_1),\hat{\boldsymbol{\sigma}}'(y_2)))\left\langle\prod_{h=1}^{l}1-\psi(\boldsymbol{\sigma}_h(y_1),\boldsymbol{\sigma}_h(y_2))\right\rangle_{\boldsymbol{G}'}\right]\\ &\quad - \sum_{l\geq 1}\frac{1}{ln^2\xi}\sum_{y_1,y_2}\mathrm{E}\left\langle\prod_{h=1}^{l}1-\psi(\boldsymbol{\sigma}_h(y_1),\boldsymbol{\sigma}_h(y_2))\right\rangle_{\boldsymbol{G}'}. \end{aligned} \tag{5.7}$$

Since Proposition 1.6 implies that given $\boldsymbol{G}'$ the assignment $\hat{\boldsymbol{\sigma}}'$ is distributed as a sample from the Gibbs measure $\mu_{\boldsymbol{G}'}$, we obtain

$$\mathrm{E}\left[\frac{1-\psi(\hat{\boldsymbol{\sigma}}'(y_1),\hat{\boldsymbol{\sigma}}'(y_2))}{ln^2\xi}\left\langle\prod_{h=1}^{l}1-\psi(\boldsymbol{\sigma}_h(y_1),\boldsymbol{\sigma}_h(y_2))\right\rangle_{\boldsymbol{G}'}\right] = \mathrm{E}\left\langle\prod_{h=1}^{l+1}1-\psi(\boldsymbol{\sigma}_h(y_1),\boldsymbol{\sigma}_h(y_2))\right\rangle_{\boldsymbol{G}'}$$

for $l\geq 1$. Moreover,

$$\frac{1}{n^2}\sum_{y_1,y_2}\mathrm{E}\left\langle 1-\psi(\boldsymbol{\sigma}(y_1),\boldsymbol{\sigma}(y_2))\right\rangle_{\boldsymbol{G}'} = 1-\sum_{y_1,y_2}\frac{\mathrm{E}\left[\psi(\boldsymbol{\sigma}'(y_1),\boldsymbol{\sigma}'(y_2))\right]}{n^2} = 1-\xi+o(1).$$

Plugging these two into (5.7) and simplifying, we finally obtain

$$\mathrm{E}\left[\ln\left\langle\psi_{\boldsymbol{a}}(\boldsymbol{\sigma}_{\boldsymbol{G}'})\right\rangle_{\boldsymbol{G}'}\right] = o(1) - \frac{1-\xi}{\xi} + \sum_{l\geq 2}\sum_{y_1,y_2}\frac{1}{l(l-1)n^2\xi}\mathrm{E}\left\langle\prod_{h=1}^{l}1-\psi(\boldsymbol{\sigma}_h(y_1),\boldsymbol{\sigma}_h(y_2))\right\rangle_{\boldsymbol{G}'}$$

and the assertion follows from (5.5). □

The steps that we just followed from (5.6) onward to calculate $\mathrm{E}\ln\left\langle\psi_{\boldsymbol{a}}(\boldsymbol{\sigma}_{\boldsymbol{G}'})\right\rangle_{\boldsymbol{G}'}$ are similar to the manipulations from the interpolation argument of Abbe and Montanari [2]. Similar manipulations yield the other two terms in Proposition 5.3.

Proof of Proposition 5.2. The assertion follows from the fact that all the expressions $\Xi_{t,l}$ from Proposition 5.3 are non-negative. □

5.2. **Proof of Theorem 1.4.** By Lemma 1.5 our task comes down to calculating $\mathrm{E}[\ln Z(\hat{\boldsymbol{G}})]$. Proposition 1.6, Theorem 3.3 and Corollary 4.8 show that the Bethe free energy provides an upper bound, i.e., that

$$\limsup_{n\to\infty}\frac{1}{n}\mathrm{E}[\ln Z(\hat{\boldsymbol{G}})] \leq \sup_{\pi\in\mathscr{P}_*^2([q])}\mathscr{B}_{q,\beta,d}(\pi).$$

Proposition 5.1 yields the matching lower bound.

6. Outlook

Although the physics ideas are quite universal and applicable to a wide variety of problems almost mechanically [44, 45], from a rigorous point of view many important questions in this area of research remain open. For instance, the precise location of the q-colorability threshold (and its existence, actually) remain open. In the case of the k-SAT problem the corresponding problem has recently been solved under the assumption that k exceeds a large constant k_0 [26], but the proof is quite technical. Generally fairly the regime beyond the condensation threshold is currently not very well understood (with the exception of [56]). More is known in the case of densely connected models such as the Sherrington-Kirkpatrick model [51].

Also regarding inference problems many natural questions remain (rigorously) unsolved. For example, the *assortative* version of the stochastic block model has edge probabilities

$$p_{ij} = \frac{d}{n}\cdot\frac{\exp(\beta\mathbf{1}\{\boldsymbol{\sigma}^*(x_i)=\boldsymbol{\sigma}^*(x_j)\})}{q-1+e^{\beta}},$$

i.e., monochromatic edges are preferred. In this case the interpolation argument that we saw in Section 5 breaks down and, in effect, neither the information-theoretic threshold nor the mutual information is known at this time

for $q > 2$. The case $q = 2$ is special and the information-theoretic threshold can be identified by other methods [43, 48, 49].

REFERENCES

[1] E. Abbe: Community detection and stochastic block models: recent developments. arXiv:1703.10146 (2017).
[2] E. Abbe, A. Montanari: Conditional random fields, planted constraint satisfaction and entropy concentration. Theory of Computing **11** (2015) 413–443.
[3] E. Abbe, C. Sandon: Detection in the stochastic block model with multiple clusters: proof of the achievability conjectures, acyclic BP, and the information-computation gap. arXiv:1512.09080 (2015).
[4] D. Achlioptas, E. Friedgut: A sharp threshold for k-colorability. Random Struct. Algorithms **14** (1999) 63–70.
[5] D. Achlioptas, C. Moore: Almost all graphs of degree 4 are 3-colorable. Journal of Computer and System Sciences **67** (2003) 441–471.
[6] D. Achlioptas, A. Naor: The two possible values of the chromatic number of a random graph. Annals of Mathematics **162** (2005), 1333–1349.
[7] M. Aizenman, R. Sims, S. Starr: An extended variational principle for the SK spin-glass model. Phys. Rev. B **68** (2003) 214403.
[8] D. Aldous , J. Steele: The objective method: probabilistic combinatorial optimization and local weak convergence (2003).
[9] N. Alon, J. Spencer: The probabilistic method. Wiley.
[10] J. Banks, C. Moore, J. Neeman, P. Netrapalli: Information-theoretic thresholds for community detection in sparse networks. Proc. 29th COLT (2016) 383–416.
[11] V. Bapst, A. Coja-Oghlan: Harnessing the Bethe free energy. Random Structures and Algorithms **49** (2016) 694–741.
[12] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann, D. Vilenchik: The condensation phase transition in random graph coloring. Communications in Mathematical Physics **341** (2016) 543–606.
[13] B. Bollobás: Modern graph theory. Springer 1998.
[14] B. Bollobás: Random graphs. Cambridge University Press (2nd ed.).
[15] C. Bordenave, M. Lelarge, L. Massoulié: Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs. Proc. 56th FOCS (2015) 1347–1357.
[16] R. Boppana: Eigenvalues and graph bisection: an average-case analysis. Proc. 28th FOCS (1987) 280–285
[17] A. Coja-Oghlan: Upper-bounding the k-colorability threshold by counting covers. Electronic Journal of Combinatorics **20** (2013) P32.
[18] A. Coja-Oghlan, C. Efthymiou, N. Jaafari: Local convergence of random graph colorings. Combinatorica, in press.
[19] A. Coja-Oghlan, C. Efthymiou, N. Jaafari, M. Kang, T. Kapetanopoulos: Charting the replica symmetric phase. arXiv:1704.01043 (2017).
[20] A. Coja-Oghlan, F. Krzakala, W. Perkins, L. Zdeborová: Information-theoretic thresholds from the cavity method. arXiv:1611.00814 (2016).
[21] A. Coja-Oghlan, K. Panagiotou: The asymptotic k-SAT threshold. Advances in Mathematics **288** (2016) 985–1068.
[22] A. Coja-Oghlan, W. Perkins: Belief Propagation on replica symmetric random factor graph models. arXiv:1603.08191 (2016).
[23] A. Coja-Oghlan, D. Vilenchik: The chromatic number of random graphs for most average degrees. International Mathematics Research Notices 2016:5801-5859.
[24] J. Cook, O. Etesami, R. Miller, L. Trevisan: On the one-way function candidate proposed by Goldreich. ACM Transactions on Computation Theory **6** (2014) 14.
[25] A. Decelle, F. Krzakala, C. Moore, L. Zdeborová: Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. Phys. Rev. E **84** (2011) 066106.
[26] J. Ding, A. Sly, N. Sun: Proof of the satisfiability conjecture for large k. Proc. 47th STOC (2015) 59–68. Full version: arXiv:1411.0650.
[27] R. Durrett: Random graph dynamics. Cambridge.
[28] P. Erdős, A. Rényi: On the evolution of random graphs. Magayar Tud. Akad. Mat. Kutato Int. Kozl. **5** (1960) 17–61.
[29] S. Franz, M. Leone:Replica bounds for optimization problems and diluted spin systems. J. Stat. Phys. **111** (2003) 535–564.
[30] U. Feige, M. Langberg, G. Schechtman: Graphs with tiny vector chromatic numbers and huge chromatic numbers. SIAM Journal on Computing **33** (2004) 1338–1368
[31] A. Frieze, M. Karoński: Introduction to random graphs. Cambridge (2015).
[32] A. Galanis, D. Stefankovic, E. Vigoda: Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. J. ACM **62** (2015) 50
[33] R. Gallager: Information theory and reliable communication. Wiley 1968.
[34] F. Guerra: Broken replica symmetry bounds in the mean field spin glass model. Comm. Math. Phys. **233** (2003) 1–12.
[35] R. van der Hofstad: Random graphs and complex networks: volume 1. Cambridge (2016).
[36] P. Holland, K. Laskey, S. Leinhardt: Stochastic blockmodels: First steps. Social networks, **5** (1983) 109–137.
[37] S. Janson, T. Łuczak, A. Ruciński: Random Graphs, Wiley 2000.
[38] P. Keevash: The existence of designs. arXiv:1401.3665 (2014).
[39] H. Kesten, B. Stigum. Additional limit theorem for indecomposable multidimensional Galton-Watson processes. Ann. Math. Statist. **37** (1966) 1463 –1481.
[40] C. Kittel, H. Kroemer: Thermal physics. Freeman 1980.
[41] S. Kumar, A. Young, N. Macris, H. Pfister: Threshold saturation for spatially-coupled LDPC and LDGM Codes on BMS channels. IEEE Transactions on Information Theory **60** (2014) 7389–7415
[42] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborova: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. National Academy of Sciences **104** (2007) 10318–10323.
[43] L. Massoulié: Community detection thresholds and the weak Ramanujan property. Proc. 46th STOC (2014) 694–703.
[44] M. Mézard, A. Montanari: Information, physics and computation, Oxford 2009.
[45] M. Mézard, G. Parisi, R. Zecchina: Analytic and algorithmic solution of random satisfiability problems. Science **297** (2002) 812–815.

[46] A. Montanari: Estimating random variables from random sparse observations. European Transactions on Telecommunications **19** (2008) 385–403.
[47] C. Moore: The computer science and physics of community detection: landscapes, phase transitions, and hardness. arXiv:1702.00467 (2017).
[48] E. Mossel, J. Neeman, A. Sly: A proof of the block model threshold conjecture. Combinatorica, in press.
[49] E. Mossel, J. Neeman, A. Sly: Reconstruction and estimation in the planted partition model. Probability Theory and Related Fields (2014) 1–31.
[50] E. Mossel, D. Weitz, N. Wormald: On the hardness of sampling independent sets beyond the tree threshold. Probability Theory and Related Fields **143** (2009) 401–439.
[51] D. Panchenko: The Sherrington-Kirkpatrick model. Springer 2013.
[52] D. Panchenko, M. Talagrand: Bounds for diluted mean-fields spin glass models. Probab. Theory Relat. Fields **130** (2004) 319–336.
[53] J. Pearl: Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann 1988.
[54] T. Richardson, R. Urbanke: Modern coding theory. Cambridge University Press (2008).
[55] A. Sly: Computational transition at the uniqueness threshold. Proc. 51st FOCS (2010) 287–296.
[56] A. Sly, N. Sun, Y. Zhang: The number of solutions for random regular NAE-SAT. arXiv:1604.08546 (2016)
[57] M. Talagrand: The Parisi formula. Annals of Mathematics **163** (2006) 221–263.
[58] J. Yedidia, W. Freeman, Y. Weiss: Constructing free-energy approximations and generalized Belief Propagation algorithms. IEEE Transactions on Information Theory **51** (2005) 2282–2312.
[59] L. Zdeborová, F. Krzakala: Statistical physics of inference: thresholds and algorithms. Advances in Physics **65** (2016) 453–552.

AMIN COJA-OGHLAN, `acoghlan@math.uni-frankfurt.de`, GOETHE UNIVERSITY, MATHEMATICS INSTITUTE, 10 ROBERT MAYER ST, FRANKFURT 60325, GERMANY.