Charting the Replica Symmetric Phase*

Amin Coja-Oghlan†[1], Charilaos Efthymiou‡[2], Nor Jaafari[3], Mihyun Kang§[4], and Tobias Kapetanopoulos¶[5]

1 **Goethe University, Mathematics Institute, Frankfurt, Germany**
`acoghlan@math.uni-frankfurt.de`
2 **Goethe University, Mathematics Institute, Frankfurt, Germany**
`efthymiou@math.uni-frankfurt.de`
3 **Goethe University, Mathematics Institute, Frankfurt, Germany**
`jaafari@math.uni-frankfurt.de`
4 **Technische Universität Graz, Institute of Discrete Mathematics, Graz, Austria**
`kang@math.tugraz.at`
5 **Goethe University, Mathematics Institute, Frankfurt, Germany**
`kapetano@math.uni-frankfurt.de`

Abstract

Random graph models and associated inference problems such as the stochastic block model play an eminent role in computer science, discrete mathematics and statistics. Based on non-rigorous arguments physicists predicted the existence of a generic phase transition that separates a "replica symmetric phase" where statistical inference is impossible from a phase where the detection of the "ground truth" is information-theoretically possible. In this paper we prove a contiguity result that shows that detectability is indeed impossible within the replica-symmetric phase for a broad class of models. In particular, this implies the detectability conjecture for the disassortative stochastic block model from [Decelle et al.: Phys. Rev. E 2011]. Additionally, we investigate key features of the replica symmetric phase such as the nature of point-to-set correlations ('reconstruction').

1998 ACM Subject Classification G.2 Discrete Mathematics, G.3 Probability and Statistics

Keywords and phrases Random factor graph, bounds for condensation phase transition, Potts antiferromagnet, diluted k-spin model, stochastic block model.

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2017.40

1 Introduction

1.1 The cavity method

Models based on random graphs have come to play a role in combinatorics, probability, statistics and computer science that can hardly be overstated. For example, the random k-SAT model is of fundamental interest in computer science [4], the stochastic block model has gained prominence in statistics [1, 24, 36], low-density parity check codes have become a

---

* A full version of the paper is available at `https://arxiv.org/abs/1704.01043`.

† The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 278857–PTCC.
‡ Supported by DFG grant EF 103/1-1.
§ Supported by Austrian Science Fund (FWF): P26826.
¶ Supported by Stiftung Polytechnische Gesellschaft PhD grant.

pillar of modern coding theory [40] and problems such as random graph coloring have been the lodestars of probabilistic combinatorics since the days of Erdős and Rényi [4, 10, 39]. Additionally, very similar models have been studied in statistical physics as models of disordered systems [31] and over the past 20 years physicists developed an analytic but non-rigorous technique for the study of such models called the 'cavity method'. This non-rigorous approach has inspired numerous "predictions" with an impact on an astounding variety of problems (e.g., [15, 31, 33, 42]). Hence the task of putting the cavity method on a rigorous foundation has gained substantial importance. Despite recent successes (e.g., [13, 17, 22, 36, 8, 16, 28]) much remains to be done. In particular, while the cavity method can be applied almost mechanically to a wide variety of problems, most rigorous arguments still hinge on model-specific deliberations, a state of affairs that begs the questions of whether we can rigorise the physics calculations wholesale. This is the thrust of the present paper.

One of the most important predictions of the cavity method is that random graph models generically undergo a *condensation phase transition* [27] that separates a "replica symmetric phase" without extensive long-range correlations from a phase where long-range correlations prevail. The fact *that* a phase transition occurs at the location predicted by the cavity method was recently proved for a fairly broad family of models [13]. However, that result fell short of establishing the connection to the nature of correlations claimed by the physics work. We rigorise the entire "physics story" of how correlations evolve up to the condensation phase transition as predicted in [18, 27, 29], including the nature of long-range correlations and the onset of point-to-set correlations known as the "reconstruction threshold". Furthermore, verifying a prominent prediction from [15], we prove a contiguity statement that has an impact on statistical inference problems such as the stochastic block model.

The results of this paper cover a wide class of random graph models, even broader than the family of models for which the condensation threshold was previously derived in [13]. Before presenting the general results in Section 2, we illustrate their impact on three important examples: the Potts antiferromagnet on the Erdős-Rényi random graph, the stochastic block model and the diluted k-spin model.

1.2 The Potts antiferromagnet

Let $q \geq 2$ be an integer, let $\Omega = \{1, \ldots, q\}$ be a set of q "colors" and let $\beta > 0$. The *antiferromagnetic q-spin Potts model on a graph $G = (V, E)$ at inverse temperature β is the distribution on* Ω^V defined by

$$\mu_{G,q,\beta}(\sigma) = (Z_{q,\beta}(G))^{-1} \prod_{\{v,w\} \in E} \exp(-\beta \mathbf{1}\{\sigma(v) = \sigma(w)\}), \tag{1.1}$$

where $Z_{q,\beta}(G) = \sum_{\tau \in \Omega^V} \prod_{\{v,w\} \in E} \exp(-\beta \mathbf{1}\{\tau(v) = \tau(w)\})$.

The Potts model can be viewed as a version of the graph coloring problem where monochromatic edges are not strictly forbidden but merely incur a 'penality factor' of $\exp(-\beta)$. The model has received attention in the context of the complexity of counting (e.g., [20]).

The Potts model on the random graph $\mathbb{G} = \mathbb{G}(n, p)$ with vertex set $V_n = \{x_1, \ldots, x_n\}$ whose edge set $E(\mathbb{G})$ is obtained by including each of the possible edge with probability $p \in [0, 1]$ independently, has received considerable attention as well (e.g. [5, 12, 14]). The most challenging case turns out to be that $p = d/n$ for a fixed real $d > 0$. The key problem associated with the model is to determine the distribution of the variable $\ln Z_\beta(\mathbb{G}, q, \beta)$.

Recently Coja-Oghlan, Krzakala, Perkins and Zdeborová [13] determined the *condensation threshold* $d_{\text{cond}}(q, \beta)$. Specifically, this is defined as the smallest value of d where the function

$d \mapsto \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\ln Z_\beta(\mathbb{G}, q, \beta)]$ is non-analytic (the existence of the limit was proved by Bayati, Gamarnik and Tetali [9]). The precise formula for $d_{\mathrm{cond}}(q, \beta)$ is complicated and not important here, but we recall the explicit *Kesten-Stigum bound*

$$d_{\mathrm{cond}}(q, \beta) \leq d_{\mathrm{KS}}(q, \beta) = \left(\frac{q - 1 + \mathrm{e}^{-\beta}}{1 - \mathrm{e}^{-\beta}}\right)^2. \tag{1.2}$$

Moreover, Azuma's inequality shows that $\frac{1}{n}\ln Z_{q,\beta}(\mathbb{G})$ converges to $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\ln Z_{q,\beta}(\mathbb{G})]$ in probability, and thus $\ln Z_{q,\beta}(\mathbb{G})$ has fluctuations of order $o(n)$. On the other hand, given that, e.g., the size of the largest component of $\mathbb{G}$ exhibits fluctuations of order $\sqrt{n}$ even once we condition on the number $|E(\mathbb{G})|$ of edges, one might expect that so does $\ln Z_{q,\beta}(\mathbb{G})$. Yet remarkably, the following theorem shows that $\ln Z_{q,\beta}(\mathbb{G})$ merely has *bounded* fluctuations given $|E(\mathbb{G})|$. In fact, we can determine the precise limiting distribution.

▶ **Theorem 1.** *Let $q \geq 2$, $\beta > 0$ and $0 < d < d_{\mathrm{cond}}(q, \beta)$. With $(K_l)_{l\geq 3}$ a sequence of independent Poisson variables with mean $\mathbb{E}[K_l] = d^l/(2l)$, let*

$$\mathcal{K} = \sum_{l=3}^{\infty} K_l \ln(1 + \delta_l) - \frac{d^l \delta_l}{2l} \qquad \text{where} \quad \delta_l = (q-1)\left(\frac{\mathrm{e}^{-\beta} - 1}{q - 1 + \mathrm{e}^{-\beta}}\right)^l.$$

Then $\mathbb{E}|\mathcal{K}| < \infty$ and as $n \to \infty$ the random variable,

$$\ln Z_{q,\beta}(\mathbb{G}) - \left(n + \frac{1}{2}\right)\ln q - |E(\mathbb{G})| \ln\left(1 - \frac{1 - \mathrm{e}^{-\beta}}{q}\right)$$
$$+ \frac{q-1}{2}\ln\left(1 + \frac{d(1 - \mathrm{e}^{-\beta})}{q - 1 + \mathrm{e}^{-\beta}}\right) + \frac{d\delta_1}{2} + \frac{d^2\delta_2}{4}$$

converges in distribution to $\mathcal{K}$.

Arguably the key element of the physics narrative is that for $d < d_{\mathrm{cond}}(q, \beta)$ the measure $\mu_{\mathbb{G},q,\beta}$ is free from extensive long-range correlations, while such correlations emerge for $d > d_{\mathrm{cond}}(q, \beta)$. Our next result verifies this conjecture. Formally, we define the *overlap* of two colorings $\sigma, \tau : V_n \to \Omega$ as the probability distribution $\rho_{\sigma,\tau} = (\rho_{\sigma,\tau}(s,t))_{s,t\in\Omega}$ on $\Omega \times \Omega$ with $\rho_{\sigma,\tau}(s,t) = |\sigma^{-1}(s) \cap \tau^{-1}(t)|/n$ for $s, t \in \Omega$. Thus, $\rho_{\sigma,\tau}(s,t)$ is the probability that a random vertex v is colored s under σ and t under τ. Let $\bar{\rho}$ denote the uniform distribution on $\Omega \times \Omega$. We write $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$ for two independent samples from $\mu_{\mathbb{G},q,\beta}$, denote the expectation with respect to $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$ by $\langle\,\cdot\,\rangle_{\mathbb{G},q,\beta}$ and the expectation over the choice of $\mathbb{G}$ by $\mathbb{E}\,[\,\cdot\,]$.

▶ **Theorem 2.** *For all $q \geq 2, \beta > 0$ we have*

$$d_{\mathrm{cond}}(q, \beta) = \inf\left\{d > 0 : \limsup_{n\to\infty} \mathbb{E}\,\langle\|\rho_{\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2} - \bar{\rho}\|_{\mathrm{TV}}\rangle_{\mathbb{G}} > 0\right\}.$$

Theorem 2 implies the absence of extensive long-range correlations in the replica symmetric phase. Indeed, for two vertices $x, y \in V_n$ and $s, t \in \Omega$ let

$$\mu_{\mathbb{G},x,y}(s,t) = \langle \mathbf{1}\{\boldsymbol{\sigma}_1(x) = s, \boldsymbol{\sigma}_1(y) = t\}\rangle_{\mathbb{G}}$$

be the joint distribution of the spins assigned to x, y. It is known (e.g., [6, Section 2]) that

$$\lim_{n\to\infty} \mathbb{E}\,\langle\|\rho_{\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2} - \bar{\rho}\|_{\mathrm{TV}}\rangle_{\mathbb{G}} = 0 \qquad \text{iff} \qquad \lim_{n\to\infty} \frac{1}{n^2}\sum_{x,y\in V_n} \mathbb{E}\|\mu_{\mathbb{G},x,y} - \bar{\rho}\|_{\mathrm{TV}} = 0. \tag{1.3}$$

Hence, Theorem 2 implies that for $d < d_{\mathrm{cond}}(q,\beta)$, with probability tending to 1, the colors assigned to two random vertices x, y of $\mathbb{G}$ are asymptotically independent. By contrast, Theorem 2 and (1.3) also show that the same ceases to be true beyond $d_{\mathrm{cond}}(q,\beta)$.

The condensation transition is conjectured to be preceded by another threshold where certain "point-to-set correlations" emerge [27]. Intuitively, the *reconstruction threshold* is the point from where for a random vertex $\boldsymbol{y} \in V_n$ correlations between the color assigned to $\boldsymbol{y}$ and the colors assigned to *all* vertices at a large enough distance ℓ from $\boldsymbol{y}$ persist. Formally, with $\boldsymbol{\sigma}$ chosen from $\mu_{\mathbb{G}}$ let $\nabla_{\ell,}(\mathbb{G}, y)$ be the σ-algebra on Ω^{V_n} generated by the random variables $\boldsymbol{\sigma}(z)$ with z ranging over all vertices at distance at least ℓ from $\boldsymbol{y}$. Then

$$\mathrm{corr}(d) = \lim_{\ell\to\infty} \limsup_{n\to\infty} \frac{1}{n} \sum_{y\in V_n} \sum_{s\in\Omega} \mathbb{E}\left\langle \left| \left\langle \mathbf{1}\{\boldsymbol{\sigma}(y) = s\} \middle| \nabla_\ell(\mathbb{G}, y) \right\rangle_{\mathbb{G},} - 1/q \right| \right\rangle_{\mathbb{G}} \tag{1.4}$$

measures the extent of correlations between $\boldsymbol{y}$ and a random boundary condition in the limit $\ell, n \to \infty$ (the outer limit exists due to mononicity). Indeed, with the expectation $\mathbb{E}[\,\cdot\,]$ in (1.4) referring to the choice of $\mathbb{G}$, the outer $\langle\,\cdot\,\rangle_{\mathbb{G}}$ chooses a random coloring of the vertices at distance at least ℓ from y and the inner $\langle\,\cdot\,|\nabla_\ell(\mathbb{G}, y)\rangle_{\mathbb{G}}$ averages over the color of y given the boundary condition.

The *reconstruction threshold* is defined as $d_{\mathrm{rec}}(q,\beta) = \inf\{d > 0 : \mathrm{corr}_{q,\beta}(d) > 0\}$. A priori, calculating $d_{\mathrm{rec}}(q,\beta)$ appears to be quite challenging because we seem to have to control the joint distribution of all the colors at distance ℓ from y. However, according to physics predictions $d_{\mathrm{rec}}(q,\beta)$ is identical to the corresponding threshold on a random tree [27], conceptually a *much* simpler object. Formally, let $\mathbb{T}(d)$ be the Galton-Watson tree with offspring distribution $\mathrm{Po}(d)$. Let r be its root and for an integer $\ell \geq 1$ let $\mathbb{T}^\ell(d)$ be the finite tree obtained by deleting all vertices at distance greater than ℓ from r. Then

$$\mathrm{corr}^\star(d) = \lim_{\ell\to\infty} \sum_{s\in\Omega} \mathbb{E}\left\langle \left| \left\langle \mathbf{1}\{\boldsymbol{\sigma}(r) = s\} \middle| \nabla_\ell(\mathbb{T}^\ell(d), r) \right\rangle_{\mathbb{T}^\ell(d)} - 1/q \right| \right\rangle_{\mathbb{T}^\ell(d)}$$

measures the extent of correlations between the color of the root and the colors at the boundary of the tree. Accordingly, the *tree reconstruction threshold* is defined as $d^\star_{\mathrm{rec}}(q,\beta) = \inf\{d > 0 : \mathrm{corr}^\star(d) > 0\}$. Combining Theorem 2 with a result in [21], we obtain

▶ **Corollary 3.** *For every $q \geq 2$ and $\beta > 0$ we have $1 \leq d_{\mathrm{rec}}(q,\beta) = d^\star_{\mathrm{rec}}(q,\beta) \leq d_{\mathrm{cond}}(q,\beta)$.*

1.3 The stochastic block model

The disassortative *stochastic block model*, first introduced in [24], is defined as follows: First choose a random q-coloring $\boldsymbol{\sigma}^* : V_n \to \Omega$ of n vertices with $q \geq 2$. Then, setting

$$d_{\mathrm{in}} = \frac{dq\mathrm{e}^{-\beta}}{q - 1 + \mathrm{e}^{-\beta}} \quad \text{and} \quad d_{\mathrm{out}} = \frac{dq}{q - 1 + \mathrm{e}^{-\beta}} \tag{1.5}$$

we generate a random graph $\mathbb{G}^*$ by connecting any two vertices v, w of the same color with probability d_{in}/n and any two with distinct with probability d_{out}/n independently. Thus, the average degree of $\mathbb{G}^*$ converges to d in probability.

Two fundamental statistical problems arise [15]. First, given q, β, for what values of d is it possible to perform non-trivial inference, i.e., obtain a better approximation to $\boldsymbol{\sigma}^*$ given the random graph $\mathbb{G}^*$ that just a random guess (see [15] for a formal definition)? A second, more modest task is the *detection problem*, which merely asks whether the random graph $\mathbb{G}^*$ can be told apart from the natural "null model", i.e., the plain Erdős-Rényi graph $\mathbb{G}$.

Decelle, Krzakala, Moore and Zdeborová [15] predicted that for $d < d_{\mathrm{cond}}(q,\beta)$, i.e., below the Potts condensation threshold, it is information-theoretically impossible to solve either problem. On the other hand, they predicted that there exist *efficient* algorithms to solve either problem if $d > d_{\mathrm{KS}}(q,\beta)$ from (1.2). Both of these conjectures were proved in the case $q = 2$ by Mossel, Neeman and Sly [37, 38] and Massoulié [30]. The positive algorithmic conjecture was proved in full by Abbe and Sandon [2]. On the negative side, [13] shows that no algorithm can infer a non-trivial approximation to $\boldsymbol{\sigma}^*$ if $d < d_{\mathrm{cond}}(q,\beta)$ for any $q \geq 3$, $\beta > 0$. Further, Banks, Moore, Neeman, and Netrapalli [5] employed a second moment argument to determine an explicit range of d where it is impossible to discern $\mathbb{G}^*$ from $\mathbb{G}$. However, there remained an extensive gap between their explicit bound and the actual condensation threshold. Our next result closes this gap and thus settles the conjecture from [15].

$\mathbb{G}$ and $\mathbb{G}^*$ are *mutually contiguous* for $d > 0$ if for any sequence $(\mathcal{A}_n)_n$ of events we have

$$\lim_{n\to\infty} \mathbb{P}[\mathbb{G} \in \mathcal{A}_n] = 0 \quad \text{iff} \quad \lim_{n\to\infty} \mathbb{P}[\mathbb{G}^* \in \mathcal{A}_n] = 0.$$

If so, then clearly no algorithm (efficient or not) can discern with probability $1 - o(1)$ whether a given graph stems from the stochastic block model $\mathbb{G}^*$ or the "null model" $\mathbb{G}$.

▶ **Theorem 4.** *For all $q \geq 3$, $\beta > 0$, $d < d_{\mathrm{cond}}(q,\beta)$ the models $\mathbb{G}$ and $\mathbb{G}^*$ are mutually contiguous.*

This result is tight since [13, Theorem 2.6] implies that $\mathbb{G}, \mathbb{G}^*$ fail to be contiguous for $d > d_{\mathrm{cond}}(q,\beta)$.

▶ Remark. There is a similar conjecture regarding the assortative version of the stochastic block model, which can be seen as an inference version of the ferromagnetic Potts model. However, the assortative block model, and ferromagnetic models generally, are beyond the scope of the present work as such models violate one of the key technical assumptions that our proofs require (condition **POS** and **BAL** below).

1.4 The diluted k-spin model

Our third application deals with a model that is of fundamental interest in physics [23, 31, 34]. For integers $k \geq 2$, $n \geq 1$ and a real $p \in [0,1]$ let $\mathbb{H} = \mathbb{H}_k(n,p)$ be the random k-uniform hypergraph on $V_n = \{x_1, \ldots, x_n\}$ whose edge set $E(\mathbb{H})$ is obtained by including each of the $\binom{n}{k}$ possible k-subsets of V_n with probability p independently. Additionally, let $\boldsymbol{J} = (\boldsymbol{J}_e)_{e \in E(\mathbb{H})}$ be a family of independent standard Gaussians. The *k-spin model* on $\mathbb{H}$ at inverse temperature $\beta > 0$ is the distribution on the set $\{-1,1\}^{V_n}$ defined by

$$\mu_{\mathbb{H},\boldsymbol{J},\beta}(\sigma) = \frac{1}{Z_\beta(\mathbb{H},\boldsymbol{J})} \prod_{e \in E(\mathbb{H})} \exp\left(\beta \boldsymbol{J}_e \prod_{y \in e} \sigma(y)\right), \tag{1.6}$$

where $Z_\beta(\mathbb{H},\boldsymbol{J}) = \sum_{\tau \in \{\pm 1\}^{V_n}} \prod_{e \in E(\mathbb{H})} \exp\left(\beta \boldsymbol{J}_e \prod_{y \in e} \tau(y)\right)$.

The most interesting and at the same time most challenging scenario arises in the case of a sparse random hypergraph [32]. Specifically, set $p = d/\binom{n-1}{k-1}$ for a fixed $d > 0$.

Guerra and Toninelli [23] determined the condensation threshold in the special case where $k = 2$ but noticed that their argument does not extend to $k \geq 3$. Proving a conjecture from [19], the following theorem pinpoints the condensation thereshold for all $k \geq 3$.

Let us write $\mathcal{P}(\mathcal{X})$ for the set of all probability distributions on a finite set $\mathcal{X}$ and identify $\mathcal{P}(\mathcal{X})$ with the standard simplex in $\mathbb{R}^{\mathcal{X}}$. Moreover, let $\mathcal{P}^2(\mathcal{X})$ be the space of all probability measures on $\mathcal{P}(\mathcal{X})$ and let $\mathcal{P}^2_*(\mathcal{X})$ be the space of all $\pi \in \mathcal{P}^2(\mathcal{X})$ whose barycenter $\int_{\mathcal{P}(\mathcal{X})} \mu \mathrm{d}\pi(\mu)$ is the uniform distribution on $\mathcal{X}$. Finally, let $\Lambda(x) = x \ln x$.

▶ **Theorem 5.** *Suppose that $d > 0, \beta > 0$ and that $k \geq 3$. Let $\boldsymbol{\gamma}$ be a Poisson variable with mean d, let $\boldsymbol{I}_1, \boldsymbol{I}_2, \ldots$ be standard Gaussians and for $\pi \in \mathcal{P}^2_*(\{\pm 1\})$ let $\boldsymbol{\rho}^\pi_1, \boldsymbol{\rho}^\pi_2, \ldots \in \mathcal{P}(\{\pm 1\})$ be random variables with distribution π, all mutually independent. Define*

$$\begin{aligned}
&\mathcal{B}_{k-\mathrm{spin}}(d, \beta, \pi) \\
&= \frac{1}{2}\mathbb{E}\left[\Lambda\left(\sum_{\sigma_k \in \{\pm 1\}} \prod_{j=1}^{\boldsymbol{\gamma}} \sum_{\sigma_1, \ldots, \sigma_{k-1} \in \{\pm 1\}} (1 + \tanh(\beta \boldsymbol{I}_j \sigma_1 \cdots \sigma_k)) \prod_{h=1}^{k-1} \boldsymbol{\rho}^\pi_{kj+h}(\sigma_h)\right)\right] \\
&\quad - \frac{d}{k}\mathbb{E}\left[\Lambda\left(1 + \sum_{\sigma_1, \ldots, \sigma_k \{\pm 1\}} \tanh(\beta \boldsymbol{I}_1 \sigma_1 \cdots \sigma_k) \prod_{h=1}^{k} \boldsymbol{\rho}^\pi_h(\sigma_h)\right)\right].
\end{aligned}$$

and $d_{\mathrm{cond}}(k, \beta) = \inf\{d > 0 : \sup_{\pi \in \mathcal{P}^2_*(\{1,-1\})} \mathcal{B}_{k-\mathrm{spin}}(d, \beta, \pi) > \ln 2\}$. Then $0 < d_{\mathrm{cond}}(k, \beta) < \infty$ and

$$\lim_{n \to \infty} \frac{1}{n}\mathbb{E}[\ln Z_\beta(\mathbb{H}, \boldsymbol{J})] \begin{cases} = \ln 2 + \frac{d}{\sqrt{2\pi}k} \int_{-\infty}^{\infty} \ln(\cosh(z)) \exp(-z^2/2) \mathrm{d}z & \text{if } d \leq d_{\mathrm{cond}}(k, \beta), \\ < \ln 2 + \frac{d}{\sqrt{2\pi}k} \int_{-\infty}^{\infty} \ln(\cosh(z)) \exp(-z^2/2) \mathrm{d}z & \text{if } d > d_{\mathrm{cond}}(k, \beta). \end{cases}$$

As in the Potts model, the condensation threshold is conjectured to be related to the nature of correlations under $\mu_{\mathbb{H}, \boldsymbol{J}, \beta}$. The following theorem proves this conjecture for even values of k. We recall the overlap notation from Section 1.2.

▶ **Theorem 6.** *For all $\beta > 0$ and $k \geq 4$ even, it holds that*

$$d_{\mathrm{cond}}(k, \beta) = \inf\left\{d > 0 : \limsup_{n \to \infty} \mathbb{E}\left\langle \|\varrho_{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2} - \bar{\rho}\|_{\mathrm{TV}}\right\rangle_{\mathbb{H}, \beta} > 0\right\}.$$

The corresponding statement for $k = 2$ was proved by Guerra and Toninelli, but they point out that their argument does not extend to larger k [23]. Furthermore, arguing as for the Potts model, we get that $\mathbb{E}\langle \|\rho_{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2} - \bar{\rho}\|_{\mathrm{TV}}\rangle_{\mathbb{H}} = o(1)$ iff the spins of two randomly chosen vertices of $\mathbb{H}$ are asymptotically independent with probability tending to one.

2 Main results

2.1 Definitions and assumptions

Factor graphs have emerged as a unifying framework for a multitude of problems. The main results of this paper, which we present in this section, therefore deal with a general class of random factor graph models, subject merely to a few easy-to-check assumptions. Formally, let Ω be a finite set of *spins*, let $k \geq 2$ be an integer and let Ψ be a set of functions $\psi : \Omega^k \to (0, 2)$ that we call *weight functions*. A Ψ*-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*, an ordered k-tuple $\partial a = (\partial_1 a, \ldots, \partial_k a) \in V^k$ for each $a \in F$ and a weight function $\psi_a \in \Psi$ for each $a \in F$. We can picture G as a bipartite graph with variable nodes on one side and constraint nodes on the other in which each constraint node a is adjacent to $\partial_1 a, \ldots, \partial_k a$ and adorned with a weight function ψ_a. This allows us to speak of, e.g., the distance of two nodes. But we keep in mind that actually the neighborhood ∂a is an *ordered* tuple. The *Gibbs distribution* of G is the distribution on Ω^V defined by $\mu_G(\sigma) = \psi_G(\sigma)/Z(G)$ for $\sigma \in \Omega^V$, where

$$\psi_G(\sigma) = \prod_{a \in F} \psi_a(\sigma(\partial_1 a), \ldots, \sigma(\partial_k a)) \quad \text{and} \quad Z(G) = \textstyle\sum_{\tau \in \Omega^V} \psi_G(\tau).$$

For a weight function $\psi : \Omega^k \to (0, 2)$ and a permutation $\theta : [k] \to [k]$ we define $\psi^\theta : \Omega^k \to (0, 2)$, $(\sigma_1, \ldots, \sigma_k) \mapsto \psi(\sigma_{\theta(1)}, \ldots, \sigma_{\theta(k)})$. Throughout the paper we assume that Ψ is a measurable set of weight functions such that for all $\psi \in \Psi$ and all permutations θ we have $\psi^\theta \in \Psi$. Moreover, we fix a probability distribution P on Ψ. We always denote by $\boldsymbol{\psi}$ an element of Ψ chosen from P, and we set

$$q = |\Omega| \quad \text{and} \quad \xi = \xi(P) = q^{-k} \sum_{\sigma \in \Omega^k} \mathbb{E}[\boldsymbol{\psi}(\sigma)].$$

Furthermore, we always assume that P is such that the following three inequalities hold:

$$\begin{aligned} \mathbb{E}[\ln^8(1 - \max\{|1 - \boldsymbol{\psi}(\tau)| : \tau \in \Omega^k\})] &< \infty, \\ \mathbb{E}[\max\{\boldsymbol{\psi}(\tau)^{-4} : \tau \in \Omega^k\}] &< \infty, \\ \sum_{\tau \in \Omega^k} \mathbb{E}[(\boldsymbol{\psi}(\tau) - \xi)^2] &> 0. \end{aligned} \tag{2.1}$$

The first two bound the 'tails' of $\boldsymbol{\psi}(\tau)$ for $\tau \in \Omega^k$. The third one provides that $\boldsymbol{\psi}$ is non-constant.

We define the random Ψ-factor graph $\boldsymbol{G}(n, m, P)$ as follows. The set of variable nodes is $V_n = \{x_1, \ldots, x_n\}$, the set of constraint nodes is $F_m = \{a_1, \ldots, a_m\}$ and the neighborhoods $\partial a_i \in V_n^k$ are chosen uniformly and independently for $i = 1, \ldots, m$. Furthermore, the weight functions $\psi_{a_i} \in \Psi$ are chosen from the distribution P mutually independently and independently of $(\partial a_i)_{i=1,\ldots,m}$. Where P is apparent we just write $\boldsymbol{G}(n, m)$ rather than $\boldsymbol{G}(n, m, P)$. For a fixed $d > 0$, i.e. independent of n, let $\boldsymbol{m} = \boldsymbol{m}_d(n)$ have distribution $\mathrm{Po}(dn/k)$ and write $\boldsymbol{G} = \boldsymbol{G}(n, \boldsymbol{m}, P)$ for brevity. Then the expected degree of a variable node is equal to d.

Apart from the condition (2.1) the main results require (some of) the following four assumptions; crucially, they *only* refer to the distribution P on the set Ψ of weight functions.

SYM. For all $i \in \{1, \ldots, k\}$, $\omega \in \Omega$ and $\psi \in \Psi$ we have

$$\sum_{\tau \in \Omega^k} \mathbf{1}\{\tau_i = \omega\} \psi(\tau) = q^{k-1} \xi \tag{2.2}$$

and for every permutation θ and every measurable $\mathcal{A} \subset \Psi$ we have that $P(\mathcal{A}) = P(\{\psi^\theta : \psi \in \mathcal{A}\})$.

BAL. The function

$$\phi : \mu \in \mathcal{P}(\Omega) \mapsto \sum_{\tau \in \Omega^k} \mathbb{E}[\boldsymbol{\psi}(\tau)] \prod_{i=1}^k \mu(\tau_i)$$

is concave and attains its maximum at the uniform distribution on Ω.

MIN. Let $\mathcal{R}(\Omega)$ be the set of all probability distribution $\rho = (\rho(s, t))_{s,t \in \Omega}$ on $\Omega \times \Omega$ such that $\sum_{s \in \Omega} \rho(s, t) = \sum_{s \in \Omega} \rho(t, s) = q^{-1}$ for all $t \in \Omega$. The function

$$\rho \in \mathcal{R}(\Omega) \mapsto \sum_{\sigma, \tau \in \Omega^k} \mathbb{E}[\boldsymbol{\psi}(\sigma)\boldsymbol{\psi}(\tau)] \prod_{i=1}^k \rho(\sigma_i, \tau_i)$$

has the uniform distribution on $\Omega \times \Omega$ as its unique global minimizer.

POS. For all $\pi, \pi' \in \mathcal{P}_*^2(\Omega)$ the following is true. With $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots$ chosen from π, $\boldsymbol{\rho}'_1, \boldsymbol{\rho}'_2, \ldots$ chosen from π' and $\boldsymbol{\psi} \in \Psi$ chosen from P, all mutually independent, we have

$$\begin{aligned} 0 \leq\ & \mathbb{E}\left[\Lambda\left(\sum_{\tau \in \Omega^k} \boldsymbol{\psi}(\tau) \prod_{i \in [k]} \boldsymbol{\rho}_i(\tau_i)\right)\right] + (k-1)\mathbb{E}\left[\Lambda\left(\sum_{\tau \in \Omega^k} \boldsymbol{\psi}(\tau) \prod_{i \in [k]} \boldsymbol{\rho}'_i(\tau_i)\right)\right] \\ & - \mathbb{E}\left[k\Lambda\left(\sum_{\tau \in \Omega^k} \boldsymbol{\psi}(\tau) \boldsymbol{\rho}_1(\tau_1) \prod_{i \in [k] \setminus \{1\}} \boldsymbol{\rho}'_i(\tau_i)\right)\right]. \end{aligned} \tag{2.3}$$

Conditions similar to **SYM**, **BAL** and **POS** appeared in [13], too. The upshot is that all four conditions can be checked solely by inspecting the distribution P on weight functions, and this is not normally difficult. For a more detailed discussion of these conditions see the full version of this paper in [11].

It is not difficult to cast the Potts antiferromagnet and the k-spin model as factor graph models. For the Potts model we let $k = 2$ and we merely introduce a single weight function $\psi_{q,\beta}(\sigma, \tau) = \exp(-\beta \mathbf{1}\{\sigma = \tau\})$. The four conditions **SYM**, **BAL**, **POS** and **MIN** are easily verified. For the k-spin model we need infinitely many weight functions, one for each $J \in \mathbb{R}$, defined by $\psi_{J,\beta}(\sigma_1, \ldots, \sigma_k) = 1 + \tanh(J\beta)\sigma_1 \cdots \sigma_k$, and P is the distribution of $\psi_{\boldsymbol{J},\beta}$ with J a standard Gaussian. The conditions **SYM**, **BAL** and **POS** hold for this model for any k and **MIN** is satisfied for even k.

2.2 Results

We proceed with the results on the condensation phase transition, the limiting distribution of the free energy, the overlap, the reconstruction and the detection thresholds for general random factor graph models.

▶ **Theorem 7.** *Assume that P satisfies **SYM**, **BAL** and **POS** and let $d > 0$. With $\boldsymbol{\gamma}$ a $\mathrm{Po}(d)$-random variable, $\boldsymbol{\rho}_1^\pi, \boldsymbol{\rho}_2^\pi, \ldots$ chosen from $\pi \in \mathcal{P}_*^2(\Omega)$ and $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots \in \Psi$ chosen from P, all mutually independent, let*

$$\begin{aligned}\mathcal{B}(d, P, \pi) \quad &= \quad \mathbb{E}\left[\tfrac{1}{q\xi^{\boldsymbol{\gamma}}}\Lambda\left(\sum_{\sigma\in\Omega}\prod_{i\in[\boldsymbol{\gamma}]}\sum_{\tau\in\Omega^k}\mathbf{1}\{\tau_k = \sigma\}\boldsymbol{\psi}_i(\tau)\prod_{j\in[k-1]}\boldsymbol{\rho}_{ki+j}^\pi(\tau_j)\right)\right] \\ &\quad - \tfrac{d(k-1)}{k\xi}\mathbb{E}\left[\Lambda\left(\sum_{\tau\in\Omega^k}\boldsymbol{\psi}_1(\tau)\prod_{i\in[k]}\boldsymbol{\rho}_j^\pi(\tau_j)\right)\right]\end{aligned} \tag{2.4}$$

and let $d_{\mathrm{cond}} = \inf\left\{d > 0 \,:\, \sup_{\pi\in\mathcal{P}_*^2(\Omega)}\mathcal{B}(d, P, \pi) > \ln q + \frac{d}{k}\ln\xi\right\}$. *Then* $1/(k-1) \leq d_{\mathrm{cond}} < \infty$ *and*

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\ln Z(\boldsymbol{G})]\begin{cases} = \ln q + \frac{d}{k}\ln\xi & \text{if } d \leq d_{\mathrm{cond}}, \\ < \ln q + \frac{d}{k}\ln\xi & \text{if } d > d_{\mathrm{cond}}.\end{cases}$$

Theorem 7 generalizes [13, Theorem 2.7], which requires that the set Ψ of weight functions be finite (and thus does not cover the k-spin model).

Admittedly the formula for d_{cond} provided by Theorem 7 is neither very simple nor very explicit, but we are not aware of any reason why it ought to be. Yet there is a natural generalization of the Kesten-Stigum bound from (1.2) that provides an easy-to-compute upper bound on d_{cond} in terms of the spectrum of a certain linear operator. The operator is constructed as follows. For $\psi \in \Psi$ let $\Phi_\psi \in \mathbb{R}^{\Omega\times\Omega}$ be the matrix with entries

$$\Phi_\psi(\omega, \omega') = q^{1-k}\xi^{-1}\textstyle\sum_{\tau\in\Omega^k}\mathbf{1}\{\tau_1 = \omega, \tau_2 = \omega'\}\psi(\tau) \qquad (\omega, \omega' \in \Omega) \tag{2.5}$$

and let $\Xi = \Xi_P$ be the linear operator on the q^2-dimensional space $\mathbb{R}^\Omega \otimes \mathbb{R}^\Omega$ defined by

$$\Xi = \Xi_P = \mathbb{E}[\Phi_{\boldsymbol{\psi}} \otimes \Phi_{\boldsymbol{\psi}}]. \tag{2.6}$$

Furthermore, letting $\mathcal{E} = \{z \in \mathbb{R}^q \otimes \mathbb{R}^q : \forall y \in \mathbb{R}^q : \langle z, \mathbf{1} \otimes y\rangle = \langle z, y \otimes \mathbf{1}\rangle = 0\}$, with $\mathbf{1}$ denoting the vector with all entries equal to one, we introduce

$$d_{\mathrm{KS}} = \left((k-1)\max_{x\in\mathcal{E}:\|x\|=1}\langle \Xi x, x\rangle\right)^{-1}, \tag{2.7}$$

with the convention that $d_{\mathrm{KS}} = \infty$ if $\max_{x\in\mathcal{E}:\|x\|=1}\langle \Xi x, x\rangle = 0$.

▶ **Theorem 8.** *If P satisfies **SYM** and **BAL**, then $d_{\mathrm{cond}} \leq d_{\mathrm{KS}}$.*

We shall see in Section 3 that Ξ is related to the "broadcasting matrix" of a suitable Galton-Watson tree, which justifies referring to d_{KS} as a generalized version of the classical *Kesten-Stigum bound* from [26]. While this bound is not generally tight, it plays a major conceptual role, as will emerge in due course.

Theorem 7 easily implies that $n^{-1}\ln Z(\boldsymbol{G})$ converges to $\ln q + \frac{d}{k}\ln\xi$ in probability if $d < d_{\mathrm{cond}}$. Yet due to the scaling factor of $1/n$ this is but a rough first order approximation. The next theorem, arguably the principal achievement of the paper, yields the exact limiting distribution of the *unscaled* free energy $\ln Z(\boldsymbol{G})$ in the entire replica symmetric phase. Recalling (2.5), let the $\Omega \times \Omega$-matrix

$$\Phi = \Phi_P = \mathbb{E}[\Phi_{\boldsymbol{\psi}}]. \tag{2.8}$$

▶ **Theorem 9.** *Assume that P satisfies **SYM**, **BAL**, **POS** and **MIN** and that $0 < d < d_{\mathrm{cond}}$. Let $(K_l)_{l\geq 1}$ be a family of Poisson variables with means $\mathbb{E}[K_l] = \frac{1}{2l}(d(k-1))^l$ and let $(\boldsymbol{\psi}_{l,i,j})_{l,i,j\geq 1}$ be a sequence of samples from P, all mutually independent. Then the random variable*

$$\mathcal{K} = \sum_{l=1}^{\infty}\left[\frac{(d(k-1))^l}{2l}\left(1-\mathrm{tr}(\Phi^l)\right) + \sum_{i=1}^{K_l}\ln\mathrm{tr}\prod_{j=1}^{l}\Phi_{\boldsymbol{\psi}_{l,i,j}}\right] \tag{2.9}$$

satisfies $\mathbb{E}|\mathcal{K}| < \infty$ and we have the following convergence in distribution:

$$\ln Z(\boldsymbol{G}) - \left(n+\tfrac{1}{2}\right)\ln q - \boldsymbol{m}\ln(\xi) + \tfrac{1}{2}\sum_{\lambda\in\mathrm{Eig}(\Phi)\setminus\{1\}}\ln(1-d(k-1)\lambda) \quad \xrightarrow{n\to\infty} \quad \mathcal{K}. \tag{2.10}$$

Let $\bar{\rho}$ be the uniform distribution on $\Omega\times\Omega$, while for $\sigma,\tau\in\Omega^{V_n}$ we defined the *overlap* $\rho_{\sigma,\tau}$ such that $\rho_{\sigma,\tau}(\omega,\omega') = |\sigma^{-1}(\omega)\cap\tau^{-1}(\omega')|/n$. The following theorem confirms one of the core tenets of the physicists' cavity method, namely the absence of extensive long-range correlations for $d < d_{\mathrm{cond}}$.

▶ **Theorem 10.** *If **SYM**, **BAL**, **POS** , **MIN** hold, then it holds that*

$$d_{\mathrm{cond}} = \inf\left\{d > 0 : \limsup_{n\to\infty}\mathbb{E}\left\langle\|\rho_{\boldsymbol{\sigma},\boldsymbol{\tau}} - \bar{\rho}\|_{TV}\right\rangle_{\boldsymbol{G}} > 0\right\}.$$

The condensation phase transition is generally preceded by another threshold where certain point-to-set correlations emerge, the reconstruction threshold [27]. Indeed, the quantity $\mathrm{corr}(d)$ as defined in (1.4) generalises naturally to any random factor graph model. Further, we can easily construct a mulit-type Galton-Watson tree $\boldsymbol{T}(d,P)$ that mimics the local geometry of a random factor graph $\boldsymbol{G}$. Its types are variable and constraint nodes, each of the latter endowed with a weight function $\psi\in\Psi$. The root is a variable node r. The offspring of a variable node is a $\mathrm{Po}(d)$ number of constraint nodes whose weight functions are chosen from P independently. Moreover, the offspring of a constraint node is $k-1$ variable nodes. For an integer $\ell\geq 0$ we let $\boldsymbol{T}^{\ell}(d,P)$ denote the (finite) tree obtained from $\boldsymbol{T}(d,P)$ by deleting all variable nodes at distance greater than 2ℓ from r. We set

$$\mathrm{corr}^{\star}(d) = \lim_{\ell\to\infty}\sum_{s\in\Omega}\mathbb{E}\left\langle\left|\left\langle\mathbf{1}\{\boldsymbol{\sigma}(r)=s\}\middle|\nabla_{\ell}(\boldsymbol{T}^{\ell}(d,P),r)\right\rangle_{\boldsymbol{T}^{\ell}(d,P)} - 1/q\right|\right\rangle_{\boldsymbol{T}^{\ell}(d,P)}. \tag{2.11}$$

The *tree reconstruction threshold* is defined as $d_{\mathrm{rec}}^{\star} = \inf\{d>0 : \mathrm{corr}^{\star}(d) > 0\}$.

▶ **Theorem 11.** *If P satisfies **SYM**, **BAL**, **POS** and **MIN**, then $0 < d_{\mathrm{rec}} = d_{\mathrm{rec}}^{\star} \leq d_{\mathrm{cond}}$.*

Theorem 11 generalises results from [21, 35]. For further discussion see the full version [11].

Finally, there is a natural statistical inference version of the random factor graph model, the *teacher-student model* [42], a generalisation of the stochastic block model. The model is defined as follows.

TCH1 an assignment $\boldsymbol{\sigma}^* : V_n \to \Omega$, the *ground truth*, is chosen uniformly at random.

TCH2 independently of $\boldsymbol{\sigma}^*$, draw $\boldsymbol{m} = \boldsymbol{m}_d(n)$ from the Poisson distribution with mean dn/k.

TCH3 generate $\boldsymbol{G}^*$ with factor nodes $a_1, \ldots, a_{\boldsymbol{m}}$ by choosing the neighborhoods ∂a_j and the weight functions ψ_{a_j} from the distribution

$$\mathbb{P}\left[\partial a_j = (y_1, \ldots, y_k), \psi_{a_j} \in \mathcal{A}\right] \propto \mathbb{E}[\mathbf{1}\{\boldsymbol{\psi} \in \mathcal{A}\}\boldsymbol{\psi}(\sigma(y_1), \ldots, \sigma(y_k))], \tag{2.12}$$

independently for $i = 1, \ldots, \boldsymbol{m}$.

As in the case of the stochastic block model, the *detection problem* arises: given a factor graph G, for what d is it possible to discern whether G was chosen from the model $\boldsymbol{G}^*$ or from the "null model" $\boldsymbol{G}$? The following theorem shows that the detection threshold is always given by d_{cond}.

▶ **Theorem 12.** *If P satisfies **SYM**, **BAL**, **POS** and **MIN**, then $\boldsymbol{G}, \boldsymbol{G}^*$ are mutually contiguous for all $d < d_{\text{cond}}$, while $\boldsymbol{G}, \boldsymbol{G}^*$ fail to be mutually contiguous for $d > d_{\text{cond}}$.*

The disassortative stochastic block model and the teacher-student model $\boldsymbol{G}^*$ are known to be mutually contiguous [13] and thus Theorem 4 follows from Theorem 12.

3 Proof strategy

The apex of the present work is Theorem 9 about the limiting distribution of the free energy; all the other results follow from it almost immediately. For such a result the usual approach would be the second moment method, pioneered by Achlioptas and Moore [3], in combination with the small subgraph conditioning technique of Robinson and Wormald [25, 41]. However, this approach does not generally allow for tight results (in particular, it typically stops working well below the condensation threshold).

We craft a proof around the teacher-student model $\boldsymbol{G}^*$ instead. Specifically, the main achievement of the recent paper [13] was to verify the cavity formula for the leading order $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\ln Z(\boldsymbol{G}^*)]$ of the "free energy" $\ln Z(\boldsymbol{G}^*)$ (in the case that the set Ψ is finite). We will replace the second moment calculation by that free energy formula, generalized to infinite Ψ, and combine it with a suitably generalized small subgraph conditioning technique. The challenge is to integrate these two components seamlessly. We accomplish this by realizing that, remarkably, both arguments are inherently and rather elegantly tied together via the spectrum of the linear operator Ξ from (2.6). But to develop this novel approach we first need to recall the classical second moment argument and understand why it founders.

3.1 Two moments do not suffice

For any second moment calculation it is crucial to fix the number of constraint nodes as otherwise its fluctuations would boost the variance. Hence, we will work with an integer sequence $m = m(n) \geq 0$. We fix $d > 0$ and consider specific integer sequences $m = m(n) \geq 0$ such that $|m(n) - dn/k| \leq n^{3/5}$ for all n. Let $\mathcal{M}(d)$ be the set of all such sequences.

The second moment method rests on showing that $\mathbb{E}[Z(\boldsymbol{G}(n,m))^2] = O(\mathbb{E}[Z(\boldsymbol{G}(n,m))]^2)$. If this is the case, then from Azuma's inequality we get that $\lim_{n\to\infty} n^{-1}\mathbb{E}[\ln Z(\boldsymbol{G}(n,m))] = \lim_{n\to\infty} n^{-1}\ln\mathbb{E}[Z(\boldsymbol{G}(n,m))]$. The second limit is easy to compute because the expectation sits inside the logarithm, and thus we obtain the leading order of the "free energy" $\ln Z(\boldsymbol{G}(n,m))$. In fact, if we can calculate the second moment $\mathbb{E}[Z(\boldsymbol{G}(n,m))^2]$ sufficiently accurately, then it may be possible to determine the limiting distribution of $\ln Z(\boldsymbol{G}(n,m))$ precisely. Suppose that there is a sufficiently simple random variable $Q(\boldsymbol{G}(n,m))$ such that

$$\mathrm{Var}[Z(\boldsymbol{G}(n,m))] = (1+o(1))\mathrm{Var}[\mathbb{E}[Z(\boldsymbol{G}(n,m))|Q(\boldsymbol{G}(n,m))]]. \tag{3.1}$$

The formula

$$\mathrm{Var}[Z(\boldsymbol{G}(n,m))] = \mathrm{Var}[\mathbb{E}[Z(\boldsymbol{G}(n,m))|Q(\boldsymbol{G}(n,m))]] + \mathbb{E}[\mathrm{Var}[Z(\boldsymbol{G}(n,m))|Q(\boldsymbol{G}(n,m))]]$$

implies

$$\mathbb{E}[\mathrm{Var}[Z(\boldsymbol{G}(n,m))|Q(\boldsymbol{G}(n,m))]] = o(\mathbb{E}[Z(\boldsymbol{G}(n,m))]^2) \tag{3.2}$$

and it is not difficult to deduce from (3.2) that $\ln Z(\boldsymbol{G}(n,m)) - \ln\mathbb{E}[Z(\boldsymbol{G}(n,m))|Q(\boldsymbol{G}(n,m))]$ converges to 0 in probability. Hence, we get the limiting distribution of $\ln Z(\boldsymbol{G}(n,m))$ if $Q(\boldsymbol{G}(n,m))$ is simple enough so that the law of $\ln\mathbb{E}[Z(\boldsymbol{G}(n,m))|Q(\boldsymbol{G}(n,m))]$ is easy to express. The basic insight behind the small subgraph conditioning technique is that (3.1) sometimes holds with a variable Q that is determined by the statistics of bounded-length cycles in $\boldsymbol{G}(n,m)$ [25, 41].

Anyhow, the crux of the entire argument is to calculate $\mathbb{E}[Z(\boldsymbol{G}(n,m))^2]$. Stirling's formula yields the following approximation of $\mathbb{E}[Z(\boldsymbol{G}(n,m))^2]$ in terms of the overlaps:

$$\ln\mathbb{E}[Z(\boldsymbol{G}(n,m))^2] = \max_{\rho\in\mathcal{P}(\Omega^2)} n\mathcal{H}(\rho) + m\ln\left(\sum_{s,t\in\Omega^k}\mathbb{E}[\boldsymbol{\psi}(s)\boldsymbol{\psi}(t)]\prod_{i\in[k]}\rho(s_i,t_i)\right) + O(\ln n), \tag{3.3}$$

where $\mathcal{H}(\rho)$ denotes the entropy of ρ. Hence, computing the second moment comes down to identifying the overlap ρ that renders the dominant contribution to the second moment. Indeed, the second moment bound $\mathbb{E}[Z(\boldsymbol{G}(n,m))^2] = O(\mathbb{E}[Z(\boldsymbol{G}(n,m))]^2)$ holds if and only if the maximum (3.3) is attained at the uniform overlap $\bar{\rho}$. However, this is not generally true for d below but near the condensation threshold.

This problem was noticed and partly remedied in prior work by applying the second moment method to a suitably truncated random variable (e.g. [7, 12]). This method revealed, e.g., the condensation threshold in a few special cases such as the random graph q-coloring problem [7] and the random regular k-SAT model, albeit only for large q and k. Yet apart from introducing such extraneous conditions, arguments of this kind require a meticulous combinatorial study of the specific model.

3.2 The condensation phase transition and the overlap

The merit of the present approach is that we avoid combinatorial deliberations altogether. Instead we employ an asymptotic formula for $\mathbb{E}[\ln Z(\boldsymbol{G}^*)]$ for the teacher-student model $\boldsymbol{G}^*$.

▶ **Theorem 13.** *If P satisfies **SYM**, **BAL** and **POS** and $d>0$, then with $\mathcal{B}(d,P,\pi)$ from (2.4) we have* $\lim_{n\to\infty} n^{-1}\mathbb{E}[\ln Z(\boldsymbol{G}^*)] = \sup_{\pi\in\mathcal{P}_*^2(\Omega)} \mathcal{B}(d,P,\pi)$.

Theorem 13 was established in [13] for a set Ψ of weight functions that is finite and the proof of Theorem 13 is based on a limiting argument.

We deduce the following result from Theorem 13 by observing that $\frac{\partial}{\partial d} \ln Z(\boldsymbol{G}^*)$ can be expressed in terms of the overlap. Let $\boldsymbol{G}^*(n,m)$ be the teacher-student model with a fixed number m of constraint nodes.

▶ **Proposition 14.** *Assume that **BAL**, **SYM**, **POS** and **MIN** hold and that $d < d_{\text{cond}}$. There exists a sequence $\zeta = \zeta(n)$, $\zeta(n) = o(1)$ but $n^{1/6}\zeta(n) \to \infty$ as $n \to \infty$, such that for all $m \in \mathcal{M}(d)$ we have*

$$\mathbb{E}\left\langle \|\rho_{\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2} - \bar{\rho}\|_{\text{TV}} \right\rangle_{\boldsymbol{G}^*(n,m)} \leq \zeta^2. \tag{3.4}$$

Proposition 14 resolves our second moment troubles. Indeed, it enables a generic way of setting up a 'truncated' random variable: with ζ from Proposition 14 we define

$$\mathcal{Z}(G) = Z(G)\mathbf{1}\left\{\left\langle \|\rho_{\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2} - \bar{\rho}\|_{\text{TV}} \right\rangle_G \leq \zeta\right\}. \tag{3.5}$$

Hence, $\mathcal{Z}(G) = Z(G)$ if "most" pairs $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$ drawn from μ_G have overlap close to $\bar{\rho}$, and $\mathcal{Z}(G) = 0$ otherwise. Since up to contiguity the teacher-student model $\boldsymbol{G}^*(n,m)$ corresponds to a reweighted version of the random factor graph model $\boldsymbol{G}(n,m)$ where each graph G is weighted according to its partition function $Z(G)$, Proposition 14 shows immediately that this truncation does not diminish the first moment.

▶ **Corollary 15.** *If **BAL**, **SYM**, **POS** and **MIN** hold and $d < d_{\text{cond}}$, then $\mathbb{E}[\mathcal{Z}(\boldsymbol{G}(n,m))] \sim \mathbb{E}[Z(\boldsymbol{G}(n,m))]$ uniformly for all $m \in \mathcal{M}(d)$.*

The second moment calculation for $\mathcal{Z}$ is easy, too. Indeed, the very construction (3.5) of $\mathcal{Z}$ guarantees that the dominant contribution to the second moment of $\mathcal{Z}$ comes from pairs with an overlap close to $\bar{\rho}$. Hence, computing the second moment comes down to expanding the right hand side of (3.3) around $\bar{\rho}$ via the Laplace method. Yet in order to do so we need to verify that $\bar{\rho}$ is a local maximum of the function

$$\rho \in \mathcal{P}(\Omega^2) \mapsto \mathcal{H}(\rho) + \frac{d}{k} \ln \sum_{s,t \in \Omega^k} \mathbb{E}[\boldsymbol{\psi}(s)\boldsymbol{\psi}(t)] \prod_{i=1}^{k} \rho(s_i, t_i) \tag{3.6}$$

from (3.3). For the special case of the Potts antiferromagnet the overlap concentration (3.4) was established and the second moment argument for $\mathcal{Z}$ was carried out in [13]. While the generalization to random factor graph models is anything but straightforward, an even more important difference lies in the application of the Laplace method. But of course there ought to be a general, conceptual explanation. As we shall see momentarily, there is one indeed, namely the generalized Kesten-Stigum bound.

3.3 The Kesten-Stigum bound

To see the connection, we observe that the Hessian of (3.6) at the point $\bar{\rho}$ is equal to $q(\text{id} - d(k-1)\Xi)$, where Ξ us the matrix from (2.6). Hence, taking into account that the argument ρ is a probability distribution on $\Omega \times \Omega$, we find that $\bar{\rho}$ is a local maximum of (3.6) if and only if

$$\langle(\text{id} - d(k-1)\Xi)x, x\rangle > 0 \qquad \text{for all } x \in \mathbb{R}^q \otimes \mathbb{R}^q \text{ such that } x \perp \mathbf{1} \otimes \mathbf{1}. \tag{3.7}$$

In order to get a handle on the spectrum of the operator Ξ from (2.6) we begin with the following observation about the matrices Φ_ψ and Φ from (2.5) and (2.8).

▶ **Lemma 16.** *Let P satisfy* ***SYM****. Then the matrix Φ_ψ is stochastic and thus $\Phi_\psi \mathbf{1} = \mathbf{1}$ for every $\psi \in \Psi$. Moreover, Φ is symmetric and doubly-stochastic. If, additionally, P satisfies* ***BAL****, then $\max_{x \perp \mathbf{1}} \langle \Phi x, x \rangle \leq 0$.*

Proceeding to the operator Ξ, we recall the definition of $\mathcal{E}$ from (2.7) and we introduce

$$\mathcal{E}' = \{x \in \mathbb{R}^q \otimes \mathbb{R}^q : \langle x, \mathbf{1} \otimes \mathbf{1} \rangle = 0\} \supset \mathcal{E}. \tag{3.8}$$

▶ **Lemma 17.** *Assume that P satisfies* ***SYM***, ***BAL****. The operator Ξ is self-adjoint, $\Xi(\mathbf{1} \otimes \mathbf{1}) = \mathbf{1} \otimes \mathbf{1}$ and for every $x \in \mathbb{R}^q$ we have $\Xi(x \otimes \mathbf{1}) = (\Phi x) \otimes \mathbf{1}$, $\Xi(\mathbf{1} \otimes x) = \mathbf{1} \otimes (\Phi x)$ and*

$$\langle \Xi(x \otimes \mathbf{1}), x \otimes \mathbf{1} \rangle \leq 0, \qquad \langle \Xi(\mathbf{1} \otimes x), \mathbf{1} \otimes x \rangle \leq 0 \qquad \text{if } x \perp \mathbf{1}. \tag{3.9}$$

Furthermore, $\Xi\mathcal{E} \subset \mathcal{E}$ and $\Xi\mathcal{E}' \subset \mathcal{E}'$.

Lemma 17 shows that Ξ induces a self-adjoint operator on the space $\mathcal{E}$.

The following proposition yields a bound on the spectral radius of this operator. Let $\mathrm{Eig}^*(\Xi) = \{\lambda \in \mathbb{R} : \exists x \in \mathcal{E} \setminus \{0\} : \Xi x = \lambda x\}$.

▶ **Proposition 18.** *If P satisfies* ***SYM*** *and* ***BAL****, then $d_{\mathrm{cond}}(k-1) \max_{\lambda \in \mathrm{Eig}^*(\Xi)} |\lambda| \leq 1$.*

The proof of Proposition 18 is based on establishing an inherent connection between the spectrum of Ξ and the Bethe free energy functional $\mathcal{B}$ from (2.4). Specifically, we use the eigenvector of Ξ to construct a candidate maximum of the functional $\mathcal{B}$. Theorem 8 is immediate from Proposition 18.

Lemma 17 and Proposition 18 show that (3.7) is satisfied, and thus that $\bar{\rho}$ is a local maximum of (3.6), for all $d < d_{\mathrm{cond}}$. Indeed, it is immediate from (3.9) that $\langle (\mathrm{id} - d(k-1)\Xi)x, x \rangle > 0$ if x is of the form $\mathbf{1} \otimes y$ or $y \otimes \mathbf{1}$ for some $\mathbf{1} \perp y \in \mathbb{R}^q$, and Theorem 8 shows that $\langle (\mathrm{id} - d(k-1)\Xi)x, x \rangle > 0$ for all $x \in \mathcal{E}$. Hence, Proposition 18 links the free energy calculation for $\boldsymbol{G}^*$ with the second moment of $\mathcal{Z}$.

3.4 Second moment redux

Observe that by Lemma 16 the set $\mathrm{Eig}\,(\Phi)$ of eigenvalues of Φ contains precisely one non-negative element, namely 1. Therefore, the following formula makes sense.

▶ **Proposition 19.** *Suppose that P satisfies* ***SYM*** *and* ***BAL*** *and let $0 < d$. Then uniformly for all $m \in \mathcal{M}(d)$,*

$$\mathbb{E}[Z(\boldsymbol{G}(n,m))] \sim \frac{q^{n+\frac{1}{2}} \xi^m}{\prod_{\lambda \in \mathrm{Eig}(\Phi) \setminus \{1\}} \sqrt{1 - d(k-1)\lambda}}. \tag{3.10}$$

Proceeding to the second moment, we recall from Lemma 17 that Ξ induces an endomorphism on the subspace $\mathcal{E}'$ from (3.8) and for the spectrum of Ξ on $\mathcal{E}'$ we write

$$\mathrm{Eig}'(\Xi) = \{\lambda \in \mathbb{R} : \exists x \in \mathcal{E}' \setminus \{0\} : \Xi x = \lambda x\}.$$

Lemma 17 and Proposition 18 imply that $d_{\mathrm{cond}}(k-1)\lambda \leq 1$ for all $\lambda \in \mathrm{Eig}'(\Xi)$. Therefore, the following formula for the second moment makes sense, too.

▶ **Proposition 20.** *If P satisfies* ***SYM*** *and* ***BAL*** *and let $0 < d < d_{\mathrm{cond}}$. Then uniformly for all $m \in \mathcal{M}(d)$,*

$$\mathbb{E}[\mathcal{Z}(\boldsymbol{G}(n,m))^2] \leq \frac{(1+o(1)) q^{2n+1} \xi^{2m}}{\prod_{\lambda \in \mathrm{Eig}'(\Xi)} \sqrt{1 - d(k-1)\lambda}}. \tag{3.11}$$

Combining Corollary 15 with Propositions 19 and 20 and applying Lemma 17, we obtain for $m \in \mathcal{M}(d)$,

$$\frac{\mathbb{E}[\mathcal{Z}(\boldsymbol{G}(n,m))^2]}{\mathbb{E}[\mathcal{Z}(\boldsymbol{G}(n,m))]^2} \sim \frac{\prod_{\lambda \in \mathrm{Eig}(\Phi)\setminus\{1\}} 1 - d(k-1)\lambda}{\prod_{\lambda \in \mathrm{Eig}'(\Xi)} \sqrt{1 - d(k-1)\lambda}} = \prod_{\lambda \in \mathrm{Eig}^*(\Xi)} \frac{1}{\sqrt{1 - d(k-1)\lambda}} \quad \text{if } d < d_{\mathrm{cond}}. \tag{3.12}$$

In particular, the ratio of the second moment and the square of the first is bounded as $n \to \infty$.

3.5 Virtuous cycles

In order to determine the limiting distribution of $\ln Z(\boldsymbol{G}(n,m))$ we are going to "explain" the remaining variance of $\mathcal{Z}(\boldsymbol{G}(n,m))$ in terms of the statistics of the bounded-length cycles of $\boldsymbol{G}(n,m)$. However, by comparison to prior applications of the small subgraph conditioning technique, here it does not suffice to merely record how many cycles of a given length occur. We also need to take into account the specific weight functions along the cycle. Yet this approach is complicated substantially by the fact that there may be infinitely many different weight functions. To deal with this issue we are going to discretize the set of weight functions and perform a somewhat delicate limiting argument.

For integer $\ell > 0$, $E_1, \dots, E_\ell \subset \Psi$ and $s_1, t_1, \dots, s_\ell, t_\ell \in \{1, \dots, k\}$ a *signature of order ℓ* is a family

$$Y = (E_1, s_1, t_1, E_2, s_2, t_2, \dots, E_\ell, s_\ell, t_\ell)$$

such that $s_i \neq t_i$ for all $i \in \{1, \dots, \ell\}$ and $s_1 < t_1$ if $\ell = 1$. We let $\mathcal{Y}$ be the set of all signatures.

For a factor graph G we call a family $(x_{i_1}, a_{h_1}, \dots, x_{i_\ell}, a_{h_\ell})$ a *cycle of signature Y in G* if the following holds: All $i_1, \dots, i_\ell \in \{1, \dots, n\}$ are pairwise distinct, the same holds for $h_1, \dots, h_\ell \in \{1, \dots, m\}$. We impose an orientation on how we traverse the cycle, i.e. we start from x_{i_1} and we traverse towards the constraint node with the smaller index or $s_1 < t_1$ if $\ell = 1$. For this reason we require $i_1 = \min\{i_1, \dots, i_\ell\}$, while $h_1 < h_\ell$ if $\ell > 1$. The weight functions along the cycle belong to $E_1, \dots, E_\ell$, i.e. $\psi_{a_{h_j}} \in E_j$, for $j = 1, \dots, \ell$. Finally, we require that the cycle enters the jth constraint node in position s_j and leaves in position t_j.

Let $C_Y(G)$ denote the number of cycles of signature Y. Moreover, for an event $\mathcal{A} \subset \Psi$ with $P(\mathcal{A}) > 0$ and $h, h' \in \{1, \dots, k\}$ define the $q \times q$ matrix $\Phi_{\mathcal{A},h,h'}$ by letting

$$\Phi_{\mathcal{A},h,h'}(\omega, \omega') = q^{1-k} \xi^{-1} \sum_{\tau \in \Omega^k} \mathbf{1}\{\tau_h = \omega, \tau_{h'} = \omega'\} \mathbb{E}[\boldsymbol{\psi}(\tau) | \mathcal{A}] \qquad (\omega, \omega' \in \Omega). \tag{3.13}$$

In addition, for a signature $Y = (E_1, s_1, t_1, \dots, E_\ell, s_\ell, t_\ell)$ define

$$\kappa_Y = \frac{1}{2\ell} \left(\frac{d}{k}\right)^\ell \prod_{i=1}^{\ell} P(E_i), \qquad \Phi_Y = \prod_{i=1}^{\ell} \Phi_{E_i, s_i, t_i}, \qquad \hat{\kappa}_Y = \kappa_Y \operatorname{tr}(\Phi_Y). \tag{3.14}$$

A *cycle of order ℓ* is a family $(x_{i_1}, a_{h_1}, \dots, x_{i_\ell}, a_{h_\ell})$ of signature $(\Psi, s_1, t_1, \dots, \Psi, s_\ell, t_\ell)$ for some sequence $s_1, t_1, \dots, s_\ell, t_\ell$, and we let C_ℓ signify the number of such cycles. Finally, two signatures $Y = (E_1, s_1, t_1, \dots, E_\ell, s_\ell, t_\ell)$, $Y' = (E'_1, s'_1, t'_1, \dots, E'_{\ell'}, s'_{\ell'}, t'_{\ell'})$ are *disjoint* if either $\ell \neq \ell'$, or for some for some i we have $(s_i, t_i) \neq (s'_i, t'_i)$ or $E_i \cap E'_i = \emptyset$. We establish the following enhancement that takes the weight functions along the cycles into account.

▶ **Proposition 21.** *Suppose that P satisfies **SYM** and **BAL**. Let $Y_1, Y_2, \dots Y_l \in \mathcal{Y}$ be pairwise disjoint signatures and let $y_1, \dots, y_l$ be non-negative integers. Let $d > 0$. Then uniformly for all $m \in \mathcal{M}(d)$,*

$$\mathbb{P}\left[\forall t \le l: \ C_{Y_t}(\boldsymbol{G}(n,m)) = y_t\right] \sim \prod_{t=1}^{l} \mathbb{P}\left[\mathrm{Po}(\kappa_{Y_t}) = y_t\right],$$
$$\mathbb{P}\left[\forall t \le l: \ C_{Y_t}(\boldsymbol{G}^*(n,m)) = y_t\right] \sim \prod_{t=1}^{l} \mathbb{P}\left[\mathrm{Po}(\hat{\kappa}_{Y_t}) = y_t\right].$$

Thus, for disjoint $Y_1, \dots, Y_l$ the cycle counts C_{Y_t} are asymptotically independent Poisson.

Finally, we establish that $\mathcal{K}$ from Theorem 9 is well-defined. We view $\Psi \subset [0,2]^{\Omega^k}$ as a subset of a cube in Euclidean speace. For an integer $r \ge 1$ let $\mathfrak{C}_r$ be the partition of Ψ induced by slicing the cube into pairwise disjoint sub-cubes of side length $1/r$. Further, let $\mathcal{Y}_{\ell,r}$ denote the set of all signatures $(E_1, s_1, t_1, \dots, E_\ell, s_\ell, t_\ell)$ such that $E_1, \dots, E_\ell \in \mathfrak{C}_r$ and such that $P(E_i) > 0$ for all $i \le \ell$, and define $\mathcal{Y}_{\le \ell, r} = \bigcup_{l=1}^{\ell} \mathcal{Y}_{l,r}$. Furthermore, if $\psi \in \Psi$ belongs to a sub-cube $C \in \mathfrak{C}_r$, then we let

$$\psi^{(r)}(\tau) = \mathbb{E}[\boldsymbol{\psi}(\tau) | C] \qquad\qquad (\tau \in \Omega^k).$$

▶ **Proposition 22.** *Assume that P satisfies **SYM** and **BAL** and let $0 < d < d_{\mathrm{cond}}$. Let $(K_l)_{l \ge 1}$ be a family of independent Poisson variables with $\mathbb{E}[K_l] = (d(k-1))^l/(2l)$ and let $(\boldsymbol{\psi}_{l,i,j})_{l,i,j}$ be a family of independent samples from P. Furthermore, define*

$$\mathcal{K}_{\ell,r} = \sum_{l=1}^{\ell} \left[\frac{(d(k-1))^l}{2l} \left(1 - \mathrm{tr}(\Phi^l)\right) + \sum_{i=1}^{K_l} \ln \mathrm{tr} \prod_{j=1}^{l} \Phi_{\boldsymbol{\psi}^{(r)}_{l,i,j}} \right],$$
$$\mathcal{K}_{\ell} = \sum_{l=1}^{\ell} \left[\frac{(d(k-1))^l}{2l} \left(1 - \mathrm{tr}(\Phi^l)\right) + \sum_{i=1}^{K_l} \ln \mathrm{tr} \prod_{j=1}^{l} \Phi_{\boldsymbol{\psi}_{l,i,j}} \right]$$

and $\mathcal{K} = \sum_{\ell=1}^{\infty} \mathcal{K}_\ell$. Then all $\mathcal{K}_{\ell,r}$ are uniformly bounded in the L^1-norm, $\mathcal{K}_{\ell,r}$ is L^1-convergent to $\mathcal{K}_\ell$ as $r \to \infty$ and $\mathcal{K}_\ell$ is L^1-convergent to $\mathcal{K}$ as $\ell \to \infty$. Furthermore,

$$\lim_{\ell \to \infty} \lim_{r \to \infty} \exp \sum_{Y \in \mathcal{Y}_{\le \ell, r}} \frac{(\kappa_Y - \hat{\kappa}_Y)^2}{\kappa_Y} = \prod_{\lambda \in \mathrm{Eig}^*(\Xi)} \frac{1}{\sqrt{1 - d(k-1)\lambda}}.$$

Equipped with Propositions 19–22 we can determine the limiting distribution of $\ln Z(\boldsymbol{G})$ and thus prove Theorem 9 by applying Janson's version of the small subgraph conditioning theorem [25] if the set Ψ is finite. In the case of infinite Ψ additional steps are necessary, see in the full version of this paper in [11].

Acknowledgment. We thank Will Perkins, Guilhem Semerjian and Nick Wormald for helpful discussions.

References

1. Emmanuel Abbe. Community detection and stochastic block models: recent developments. *CoRR*, abs/1703.10146, 2017. URL: `http://arxiv.org/abs/1703.10146`.
2. Emmanuel Abbe and Colin Sandon. Detection in the stochastic block model with multiple clusters: proof of the achievability conjectures, acyclic bp, and the information-computation gap. *CoRR*, abs/1512.09080, 2015. URL: `http://arxiv.org/abs/1512.09080`.
3. Dimitris Achlioptas and Cristopher Moore. Random k-SAT: Two moments suffice to cross a sharp threshold. *SIAM J. Comput.*, 36(3):740–762, 2006.
4. Dimitris Achlioptas, Assaf Naor, and Yuval Peres. Rigorous location of phase transitions in hard optimization problems. *Nature*, 435(7043):759–764, 06 2005.

5 Jess Banks, Cristopher Moore, Joe Neeman, and Praneeth Netrapalli. Information-theoretic thresholds for community detection in sparse networks. In *Proceedings of the 29th Conference on Learning Theory, COLT 2016, New York, USA, June 23-26, 2016*, pages 383–416, 2016.

6 Victor Bapst and Amin Coja-Oghlan. Harnessing the bethe free energy. *Random Struct. Algorithms*, 49(4):694–741, 2016.

7 Victor Bapst, Amin Coja-Oghlan, Samuel Hetterich, Felicia Raßmann, and Dan Vilenchik. The condensation phase transition in random graph coloring. *Communications in Mathematical Physics*, 341(2):543–606, 2016.

8 Jean Barbier, Mohamad Dia, Nicolas Macris, Florent Krzakala, Thibault Lesieur, and Lenka Zdeborová. Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula. In *Advances in Neural Information Processing Systems 29*, pages 424–432. Curran Associates, Inc., 2016.

9 Mohsen Bayati, David Gamarnik, and Prasad Tetali. Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. *Ann. Probab.*, 41(6):4080–4115, 11 2013.

10 Amin Coja-Oghlan. Phase transitions in discrete structures. In *7th European Congress of Mathematics*, (In press) 2016.

11 Amin Coja-Oghlan, Charilaos Efthymiou, Nor Jaafari, Mihyun Kang, and Tobias Kapetanopoulos. Charting the replica symmetric phase. *CoRR*, abs/1704.01043, 2017. URL: `https://arxiv.org/abs/1704.01043`.

12 Amin Coja-Oghlan and Nor Jaafari. On the potts antiferromagnet on random graphs. *Electr. J. Comb.*, 23(4):P4.3, 2016.

13 Amin Coja-Oghlan, Florent Krzakala, Will Perkins, and Lenka Zdeborová. Information-theoretic thresholds from the cavity method. *CoRR*, abs/1611.00814, 2016. URL: `http://arxiv.org/abs/1611.00814`.

14 Pierluigi Contucci, Sander Dommers, Cristian Giardinà, and Shannon Starr. Antiferromagnetic potts model on the Erdős-Rényi random graph. *Communications in Mathematical Physics*, 323(2):517–554, 2013.

15 Aurelien Decelle, Florent Krzakala, Cristopher Moore, and Lenka Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Physical Review E*, 84(6):066106–, 12 2011.

16 Yash Deshpande, Emmanuel Abbe, and Andrea Montanari. Asymptotic mutual information for the binary stochastic block model. In *IEEE International Symposium on Information Theory, ISIT 2016, Barcelona, Spain, July 10-15, 2016*, pages 185–189, 2016.

17 Jian Ding, Allan Sly, and Nike Sun. Proof of the satisfiability conjecture for large k. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 59–68, 2015.

18 Ulisse Ferrari, Carlo Lucibello, Flaviano Morone, Giorgio Parisi, Federico Ricci-Tersenghi, and Tommaso Rizzo. Finite-size corrections to disordered systems on Erdős-Rényi random graphs. *Physical Review B*, 88(18):184201–, 11 2013.

19 Silvio Franz, Michele Leone, Federico Ricci-Tersenghi, and Riccardo Zecchina. Exact solutions for diluted spin glasses and optimization problems. *Physical Review Letters*, 87(12:127209), 08 2001.

20 Andreas Galanis, Daniel Stefankovic, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM*, 62(6):50:1–50:60, 2015.

21 Antoine Gerschenfeld and Andrea Montanari. Reconstruction for models on random graphs. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA, Proceedings*, pages 194–204, 2007.

22 Andrei Giurgiu, Nicolas Macris, and Rüdiger L. Urbanke. Spatial coupling as a proof technique and three applications. *IEEE Trans. Information Theory*, 62(10):5281–5295, 2016.

23 Francesco Guerra and Fabio Lucio Toninelli. The high temperature region of the Viana–Bray diluted spin glass model. *Journal of Statistical Physics*, 115(1):531–555, 2004.

24 Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137, 1983.

25 Svante Janson. Random regular graphs: Asymptotic distributions and contiguity. *Combinatorics, Probability & Computing*, 4:369–405, 1995.

26 Harry Kesten and Bernt P. Stigum. Additional limit theorems for indecomposable multidimensional galton-watson processes. *Ann. Math. Statist.*, 37(6):1463–1481, 1966. doi:10.1214/aoms/1177699139.

27 Florent Krzakała, Andrea Montanari, Federico Ricci-Tersenghi, Guilhem Semerjian, and Lenka Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proceedings of the National Academy of Sciences*, 104(25):10318–10323, 06 2007.

28 Marc Lelarge and Léo Miolane. Fundamental limits of symmetric low-rank matrix estimation. *CoRR*, abs/1611.03888, 2016. URL: https://arxiv.org/abs/1611.03888.

29 Carlo Lucibello, Flaviano Morone, Giorgio Parisi, Federico Ricci-Tersenghi, and Tommaso Rizzo. Finite-size corrections to disordered ising models on random regular graphs. *Physical Review E*, 90(1):012146–, 07 2014.

30 Laurent Massoulié. Community detection thresholds and the weak ramanujan property. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 – June 03, 2014*, pages 694–703, 2014.

31 Marc Mézard and Andrea Montanari. *Information, physics and computation.* Oxford University Press, 2009.

32 Marc Mézard and Giorgio Parisi. The bethe lattice spin glass revisited. *Eur. Phys. J. B*, 20(2):217–233, 3 2001.

33 Marc Mézard, Giorgio Parisi, and Ricardo Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297(5582):812, 08 2002.

34 Marc Mézard, Federico Ricci-Tersenghi, and Riccardo Zecchina. Two solutions to diluted p-spin models and XORSAT problems. *Journal of Statistical Physics*, 111(3):505–533, 2003.

35 Andrea Montanari, Ricardo Restrepo, and Prasad Tetali. Reconstruction and clustering in random constraint satisfaction problems. *SIAM J. Discrete Math.*, 25(2):771–808, 2011.

36 Cristopher Moore. The computer science and physics of community detection: Landscapes, phase transitions, and hardness. *CoRR*, abs/1702.00467, 2017. URL: http://arxiv.org/abs/1702.00467.

37 Elchanan Mossel, Joe Neeman, and Allan Sly. A proof of the block model threshold conjecture. *CoRR*, abs/1311.4115, 2013. URL: http://arxiv.org/abs/1311.4115.

38 Elchanan Mossel, Joe Neeman, and Allan Sly. Reconstruction and estimation in the planted partition model. *Probability Theory and Related Fields*, 162(3):431–461, 2015. doi:10.1007/s00440-014-0576-6.

39 Paul Erdős and Alfred Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 5:17–61, 1960.

40 Tom Richardson and Rüdiger Urbanke. *Modern coding theory.* Cambridge University Press, 2008.

41 Robert W. Robinson and Nicholas C. Wormald. Almost all cubic graphs are hamiltonian. *Random Struct. Algorithms*, 3(2):117–126, 1992.

42 Lenka Zdeborová and Florent Krzakala. Statistical physics of inference: thresholds and algorithms. *Advances in Physics*, 65(5):453–552, 2016.