The Condensation Phase Transition in the Regular k-SAT Model*

Victor Bapst[1] and Amin Coja-Oghlan[2]

1 **Mathematics Institute, Goethe University, Frankfurt, Germany**
bapst@math.uni-frankfurt.de
2 **Mathematics Institute, Goethe University, Frankfurt, Germany**
acoghlan@math.uni-frankfurt.de

Abstract

Much of the recent work on phase transitions in discrete structures has been inspired by ingenious but non-rigorous approaches from physics. The physics predictions typically come in the form of distributional fixed point problems that mimic Belief Propagation, a message passing algorithm. In this paper we show how the Belief Propagation calculation can be turned into a rigorous proof of such a prediction, namely the existence and location of a condensation phase transition in the regular k-SAT model.

1998 ACM Subject Classification G.2.1 Combinatorics, G.3 Probability and Statistics, F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases random k-SAT, phase transitions, Belief Propagation, condensation

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2016.22

1 Introduction

1.1 Background and motivation

Over the past three decades the study of random constraint satisfaction problems has been driven by ideas from statistical mechanics [3, 24, 25]. The physics ideas have since had a substantial impact on algorithms, coding theory and combinatorics [12, 15, 18, 19, 20, 21, 30]. The striking feature of the physics work is that it is based on one generic but non-rigorous technique called the *cavity method* that can be applied almost mechanically [23]. Its centerpiece is the *Belief Propagation* message-passing algorithm. By contrast, rigorous studies have largely been case-by-case.

This state of affairs begs the question of whether the Belief Propagation calculations can be put on a rigorous basis directly. This is precisely the thrust of the present paper. We show how the physics calculations can be turned into a proof in a highly non-trivial and somewhat representative case. We expect that this approach generalises to many other alike problems. Specifically, we determine the precise *condensation phase transition* in the random regular k-SAT model. The existence of such a phase transition in a wide variety of models is one of the key predictions of the cavity method [22] and its impact on algorithmic as well as information-theoretic question can hardly be overstated [29, 32]. For example, the condensation phenomenon has a bearing on the performance of message-passing

* 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.

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016).
Editors: Klaus Jansen, Claire Matthieu, José D. P. Rolim, and Chris Umans; Article No. 22; pp. 22:1–22:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

algorithm such as Belief Propagation guided decimation [29] as well as on statistical inference problems [5]. Moreover, the regular k-SAT problem shares many of the key properties of the better-known model where clauses are simple chosen uniformly and independently; in particular, a condensation phase transition is expected to occur in that model as well [22]. The proof builds upon on our abstract results [6] on the "regularity method" for discrete probability measures and the connection to spatial mixing properties.[1]

1.2 The regular k-SAT model

Consider variables $x_1, \ldots, x_n$ that may take the values 'true' or 'false', represented by $+1$ and -1. If $\Phi = \Phi_1 \wedge \cdots \wedge \Phi_m$ is a k-CNF formula, then we define a function $E_\Phi : \{\pm 1\}^n \to \{0, 1, \ldots, m\}$ on the set of truth assignments by letting $E_\Phi(\sigma)$ be the number of violated clauses. In physics jargon, E_Φ is called the *Hamiltonian*. Further, we define the *Gibbs measure* at "inverse temperature" $\beta \geq 0$ by letting

$$\sigma \in \{\pm 1\}^n \mapsto \exp(-\beta E_\Phi(\sigma))/Z_\Phi(\beta) \qquad \text{where} \qquad Z_\Phi(\beta) = \sum_{\sigma \in \{\pm 1\}^n} \exp(-\beta E_\Phi(\sigma)) \quad (1)$$

is called the *partition function*. Thus, the Gibbs measure is a probability measure on the cube $\{\pm 1\}^n$.

As β gets larger the mass of the Gibbs measure shifts to assignments that violated fewer clauses. Ultimately, if we let $\beta \to \infty$, then the Gibbs measure concentrates on the maximally satisfying assignments. Hence, by tuning β we can "scan" the landscape that the function E_Φ defines on the cube $\{\pm 1\}^n$. Among other things, grasping this landscape is key in order to study the performance of local search algorithms such as Simulated Annealing or the Metropolis process, which attempt to descend from a random starting point to a global minimum. For instance, if E_Φ is riddled with local minima, local search algorithms are bound to get trapped, while they might be efficient on a nice "convex" landscape [1, 9, 25].

It turns out that the key quantity upon which the study of the Hamiltonian hinges is the partition function. Therefore, we aim to calculate $Z_{\mathbf{\Phi}}(\beta)$ on a random k-CNF formula $\mathbf{\Phi}$. There are several natural probability distributions on k-SAT formulas. The one that we study here is perhaps the simplest non-trivial example, namely the *regular k-SAT model* [28]. It comes with two integer parameters $k \geq 3$ and $d > 1$, which is even. For n such that $2k$ divides dn we let $\mathbf{\Phi} = \mathbf{\Phi}_{d,k}(n)$ signify a uniformly random k-SAT formula with $m = dn/(2k)$ clauses of length k over $x_1, \ldots, x_n$ such that each variable x_i occurs precisely $d/2$ times as a positive literal x_i and precisely $d/2$ times as a negative literal $\neg x_i$.[2] For k exceeding a certain constant k_0 there is an explicitly known critical degree $d_{k-\mathrm{SAT}}$, the *satisfiability threshold*, where satisfying assignments cease to exist in a typical $\mathbf{\Phi}$ [12][3]. While the exact formula is cumbersome, asymptotically we have

$$d_{k-\mathrm{SAT}}/k = 2^k \ln 2 - k \ln 2/2 - (1 + \ln 2)/2 + o_k(1), \qquad (2)$$

where $o_k(1)$ hides a term that tends to 0 in the limit of large k. Since $Z_{\mathbf{\Phi}}(\beta)$ scales exponentially with n, we consider

$$\phi_{d,k} : \beta \in (0, \infty) \mapsto \lim_{n \to \infty} \frac{1}{n} \mathbb{E}[\ln Z_{\mathbf{\Phi}}(\beta)] \qquad (3)$$

[1] The present paper builds upon the arXiv version of [6] because the version that appeared in the proceedings of RANDOM 2015 contained a critical error.

[2] The regular k-SAT model shares many of the properties of the better known model where m clauses are chosen uniformly and independently but avoids the intricacies that result from degree fluctuations.

[3] We have $\liminf \mathbb{P}[\mathbf{\Phi} \text{ is satisfiable}] > 0$ if $d < d_{k-\mathrm{SAT}}$ and $\lim \mathbb{P}[\mathbf{\Phi} \text{ is satisfiable}] = 0$ if $d > d_{k-\mathrm{SAT}}$.

with the log *inside* the expectation and the expectation is over $\mathbf{\Phi}$. The existence of the limit follows from the interpolation method [10]. Moreover, Azuma's inequality implies that $\ln Z_{\mathbf{\Phi}}(\beta)$ concentrates about $\mathbb{E}[\ln Z_{\mathbf{\Phi}}(\beta)]$ for any d, k, β.

We call $\beta_0 \in (0,\infty)$ *smooth* if there exists $\varepsilon > 0$ such that the function

$$\beta \in (\beta_0 - \varepsilon, \beta_0 + \varepsilon) \mapsto \phi_{d,k}(\beta)$$

admits an expansion as an absolutely convergent power series around β_0. Otherwise a *phase transition* occurs at β_0.[4] Thus, with d fixed we aim to investigate the effect of tuning β.

Results. According to the "cavity method" for certain values of d close to the satisfiability threshold $d_{k-\mathrm{SAT}}$ there occurs a so-called *condensation phase transition* at a certain critical $\beta_{\mathrm{cond}}(d,k) > 0$ [22]. The main result of this paper proves this conjecture for k exceeding a certain constant k_0. Let us postpone the precise definition of $\beta_{\mathrm{cond}}(d,k)$ for just a moment.

▶ **Theorem 1.** *There exists $k_0 \geq 3$ such that for all $k \geq k_0$, $d \leq d_{k-\mathrm{SAT}}$ there is $\beta_{\mathrm{cond}}(d,k) \in (0,\infty]$ such that all $\beta \in (0, \beta_{\mathrm{cond}}(d,k))$ are smooth. If $\beta_{\mathrm{cond}}(d,k) < \infty$, then there occurs a phase transition at $\beta_{\mathrm{cond}}(d,k)$.*

We will see momentarily that $\beta_{\mathrm{cond}}(d,k) < \infty$ for d exceeding a specific critical degree $d_{\mathrm{cond}}(k) < d_{k-\mathrm{SAT}}$. Theorem 1 is the first rigorous result to identify the precise critical "inverse temperature" in a random constraint satisfaction problem, apart perhaps from the far simpler case of the stochastic block model [26].

Let us take a look at $\beta_{\mathrm{cond}}(d,k)$. As most predictions based on the cavity method, $\beta_{\mathrm{cond}}(d,k)$ results from a *distributional fixed point problem*, i.e., a fixed point problem on the space of probability measures on the open unit interval $(0,1)$. This fixed point problem derives mechanically from the physicists' "1RSB cavity equations" [23]. Specifically, writing $\mathcal{P}(0,1)$ for the set of probability measures on the unit interval, we define two maps

$$\mathcal{F}_{k,d,\beta} : \mathcal{P}(0,1) \to \mathcal{P}(0,1), \qquad \hat{\mathcal{F}}_{k,d,\beta} : \mathcal{P}(0,1) \to \mathcal{P}(0,1)$$

as follows. Given $\pi \in \mathcal{P}(0,1)$ let $\eta = (\eta_1, \ldots, \eta_{k-1}) \in (0,1)^{k-1}$ be a random $k-1$-tuple drawn from the distribution $(\hat{z}(\eta)/\hat{Z}(\pi))\,\mathrm{d}\bigotimes_{j=1}^{k-1} \pi(\eta_j)$, where

$$\hat{z}(\eta) = 2 - (1 - \exp(-\beta)) \prod_{j<k} \eta_j \qquad \text{and} \qquad \hat{Z}(\pi) = \int \hat{z}(\eta) \mathrm{d} \bigotimes_{j<k} \pi(\eta_j). \tag{4}$$

Then $\widehat{\mathcal{F}}_{k,d,\beta}(\pi)$ is the distribution of $(1 - (1 - \exp(-\beta)) \prod_{i=1}^{k-1} \eta_i)/\widehat{z}(\eta)$. Similarly, given $\hat{\pi} \in \mathcal{P}(0,1)$ draw $\hat{\eta} = (\hat{\eta}_1, \ldots, \hat{\eta}_{d-1})$ from $(z(\hat{\eta})/Z(\hat{\pi}))\mathrm{d}\bigotimes_{j=1}^{k-1} \hat{\pi}(\hat{\eta}_j)$, where

$$z(\widehat{\eta}) = \prod_{j<d/2} \widehat{\eta}_j \prod_{j\geq d/2} (1 - \widehat{\eta}_j) + \prod_{j<d/2} (1 - \widehat{\eta}_j) \prod_{j\geq d/2} \widehat{\eta}_j, \qquad Z(\hat{\pi}) = \int z(\widehat{\eta}) \mathrm{d} \bigotimes_{j<k} \hat{\pi}(\hat{\eta}_j). \tag{5}$$

Then $\mathcal{F}_{k,d,\beta}(\widehat{\pi})$ is the distribution of $(\prod_{j<d/2} \widehat{\eta}_j \prod_{j\geq d/2} (1 - \widehat{\eta}_j))/z(\widehat{\eta})$. Further, call a distribution $\pi \in \mathcal{P}(0,1)$ *skewed* if

$$\pi[(\exp(-k^{0.9}\beta), 1 - \exp(-k^{0.9}\beta))] < 2^{-0.9k}.$$

[4] This is the usual physics definition of a "phase transition". The motivation is that the non-analyticity of $\phi_{d,k}$ indicates a qualitative change. For illustration, observe that the fraction of vertices in the largest component of the Erdős-Rényi random graph is non-analytic at average degree one.

▶ **Proposition 2.** *Let $d_-(k) = d_{k-\mathrm{SAT}} - k^5$ and $\beta_-(k,d) = k\ln 2 - 10\ln k$. The map $\mathcal{G}_{k,d,\beta} = \mathcal{F}_{k,d,\beta} \circ \widehat{\mathcal{F}}_{k,d,\beta}$ has a unique skewed fixed point $\pi^\star_{k,d,\beta}$, provided that $k \geq k_0$, $d \in [d_-(k), d_{k-\mathrm{SAT}}]$ and $\beta > \beta_-(k,d)$.*

To extract $\beta_{\mathrm{cond}}(d,k)$, let $\nu_1,\ldots,\nu_k,\hat\nu_1,\ldots,\hat\nu_d$ be independent random variables such that the ν_i have distribution $\pi^\star_{k,d,\beta}$ and the $\hat\nu_i$ have distribution $\widehat{\mathcal{F}}_{k,d,\beta}(\pi^\star_{k,d,\beta})$. Setting

$$z_1 = \prod_{j\leq d/2}\hat\nu_j \prod_{j>d/2}(1-\hat\nu_j) + \prod_{j\leq d/2}(1-\hat\nu_j)\prod_{j>d/2}\hat\nu_j, \qquad z_2 = 1-(1-\exp(-\beta))\prod_{j\leq k}\nu_j$$

and $z_3 = \nu_1\hat\nu_1 + (1-\nu_1)(1-\hat\nu_1)$, we let

$$\mathcal{F}(k,d,\beta) = \ln\mathbb{E}\left[z_1\right] + \frac{d}{k}\ln\mathbb{E}\left[z_2\right] - d\ln\mathbb{E}[z_3], \tag{6}$$

$$\mathcal{B}(k,d,\beta) = \frac{\mathbb{E}\left[z_1\ln z_1\right]}{\mathbb{E}\left[z_1\right]} + \frac{d}{k}\frac{\mathbb{E}\left[z_2\ln z_2\right]}{\mathbb{E}\left[z_2\right]} - d\frac{\mathbb{E}[z_3\ln z_3]}{\mathbb{E}[z_3]}. \tag{7}$$

Finally, with the usual convention that $\inf\emptyset = \infty$ we let

$$\beta_{\mathrm{cond}}(k,d) = \begin{cases}\infty & \text{if } d < d_-(k),\\ \inf\{\beta > \beta_-(k,d) : \mathcal{F}(k,d,\beta) < \mathcal{B}(k,d,\beta)\} & \text{if } d \in [d_-(k), d_{k-\mathrm{SAT}}].\end{cases}$$

We proceed to highlight a few consequences of Theorem 1 and its proof. The following result shows that $\beta_{\mathrm{cond}}(d,k) < \infty$, i.e., that a condensation phase transition occurs, for degrees d strictly below the satisfiability threshold.

▶ **Corollary 3.** *If $k \geq k_0$, then $d_{\mathrm{cond}}(k) = \min\{d > 0 : \beta_{\mathrm{cond}}(d,k) < \infty\} < d_{k-\mathrm{SAT}} - \Omega(k)$.*

Furthermore, the following corollary shows that the so-called "replica symmetric solution" predicted by the cavity method yields the correct value of $\phi_{d,k}(\beta)$ for $\beta < \beta_{\mathrm{cond}}(d,k)$.

▶ **Corollary 4.** *If $k \geq k_0$, $d \leq d_{k-\mathrm{SAT}}$ and $\beta < \beta_{\mathrm{cond}}(d,k)$, then $\phi_{d,k}(\beta) = \mathcal{F}(k,d,\beta)$.*

Corollary 4 opens the door to studying the "shape" of the Hamiltonian $E_{\mathbf{\Phi}}$ for $\beta < \beta_{\mathrm{cond}}(d,k)$, a necessary step towards studying, e.g., the performance of local search algorithms. Specifically, Corollary 4 enables us to bring the "planting trick" from [1] to bear so that we can analyse typical properties of samples from the Gibbs measure.

Finally, complementing Corollary 4, the following result shows that $\mathcal{F}(k,d,\beta)$ overshoots $\phi_{d,k}(\beta)$ for $\beta > \beta_{\mathrm{cond}}(d,k)$.

▶ **Corollary 5.** *If $k \geq k_0$, $d \leq d_{k-\mathrm{SAT}}$ and $\beta > \beta_{\mathrm{cond}}(d,k)$, then there is $\beta_{\mathrm{cond}}(d,k) < \beta' < \beta$ such that $\phi_{d,k}(\beta') < \mathcal{F}(k,d,\beta')$.*

2 Techniques and related work

Admittedly, the definition of $\beta_{\mathrm{cond}}(k,d)$ is not exactly simple. For instance, even though the fixed point distribution from Proposition 2 stems from a discrete problem, it is a continuous distribution on $(0,1)$. Yet the analytic formula (6) is conceptually *far* simpler than the combinatorial definition of $\phi_{d,k}$. Indeed, we are going to see in Section 3 that the fixed point problem can be understood in terms of a branching process, i.e., a random infinite tree.

The proof of Theorem 1 builds upon an abstract result from [6] that, roughly speaking, breaks the study of the partition function down into two tasks. First, to prove that the Gibbs measure induced by a random formula $\hat{\mathbf{\Phi}}$ chosen from a reweighted probability distribution, the "planted model", enjoys the *non-reconstruction property*, a spatial mixing property. Second, to analyse Belief Propagation on $\hat{\mathbf{\Phi}}$. The technical contribution of the present work is to tackle these two problems in a fairly generic way. In fact, we expect that the proof strategy extends to other problems. A concrete example that springs to mind is the Potts antiferromagnet on a random graph, which is intimately related to the information-theoretic threshold in the "stochastic block model" with multiple classes [5]. While conceptually the proof strategy allows us to turn the Belief Propagation calculation into a rigorous theorem in a fairly direct way, the technical challenge of actually analysing the relevant Belief Propagation fixed point in a completely rigorous manner remains.

The overall proof strategy bears some resemblance to the work of Mossel, Neeman and Sly [26] on the "stochastic block model", but the details are quite different. Roughly speaking, the stochastic block model can be viewed as a planted version of the minimum bisection problem and the problem is to recover the labels that were used to generate the graph. The proof from [26] that this is not possible up to a certain point relies on non-reconstruction as well. Moreover, the contiguity estabished in [26] can be viewed as a condensation result, albeit with the much simpler interactions of the stochastic block model. In particular, the "condensation threshold" is merely given by a quadratic equation rather than a distributional fixed point equation.

The predictions of the "cavity method" typically come as distributional fixed points but there are only few proofs that establish such predictions rigorously. The one most closely related to the present work is [7] on condensation in random graph coloring. It determines the critical average degree d for which condensation starts to occur with respect to the number of proper k-colorings of the Erdös-Rényi random graph. This corresponds to taking $\beta \to \infty$ in (3). This simplifies the problem substantially because in the limit "frozen variables" emerge that are fixed deterministically to one specific value. Other previous results on condensation gave only approximate answers [8, 13, 14].

Interestingly, determining the satisfiability threshold on $\mathbf{\Phi}$ is conceptually easier than identifying the condensation threshold [12]. This is because the local structure of the regular random formula is essentially deterministic, namely a tree comprising of clauses and variables in which every variable appears $d/2$ times positively and $d/2$ times negatively. In effect, the satisfiability threshold is given by a fixed point problem on the unit interval, rather than on the space of probability measures on the unit interval. Similar simplifications occur in other regular models [17, 16]. By contrast, we will see in Section 3 that the condensation phase transition hinges on the reweighted distribution $\hat{\mathbf{\Phi}}$ with a genuinely random local structure.

Recent work on the k-SAT threshold in uniformly random formulas [12, 11] and in particular the breakthrough paper by Ding, Sly and Sun [18], also harnessed the Belief/Survey Propagation calculations and [18] verified the prediction in terms of the corresponding distributional fixed point problem.[5] In the uniformly random model a substantial technical complication is posed by variables of exceptionally high degree. While [12, 11, 18] apply the second moment method to a random variable whose construction is guided by Belief/Survey Propagation, here we employ Belief Propagation in the direct way enabled by [6].

Talagrand [31] and, by means of a different argument, Panchenko [27] studied the k-SAT model on uniformly random formulas in the "high-temperature" (i.e., small β) case. Specifically, with d the average degree of a variable, [27, 31] require that $\min\{4\beta, 1\}(k-1)d < 1$. This range of parameters is well below the conjectured condensation phase transition [22].

[5] Survey Propagation can be viewed as a Belief Propagation applied to a modified constraint satisfaction problem [23].

3 Proof outline

We assume that $k \geq k_0$ for a large enough constant k_0 and that $d < d_{k-\mathrm{SAT}}$.

3.1 Two moments do not suffice

The default approach to studying $\phi_{d,k}(\beta)$ would be the venerable "second moment method" [3]. Cast on a logarithmic scale, if

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)^2] \leq \lim_{n\to\infty} \frac{2}{n} \ln \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)], \qquad \text{then} \tag{8}$$

$$\phi_{d,k}(\beta) = \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)]. \tag{9}$$

Thus, if (8) holds, then we can "swap the log and the expecation". Unsurprisingly, calculating $\ln \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)]$ is fairly easy (see (11) below).

From a bird's eye view, both the physics intuition and the second moment are all about the geometry of the Gibbs measure of $\mathbf{\Phi}$ at a given $\beta \in (0, \infty)$. Indeed, according to the physics picture the condensation point $\beta_{\mathrm{cond}}(k)$ should be the supremum of all $\beta > 0$ such that w.h.p. for two random assignments $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \in \{\pm 1\}^n$ chosen from the Gibbs measure we have $|\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2| = o_n(n)$, i.e., $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$ are about orthogonal [22]. This is a necessary condition for the success of the second moment method as well [2, 4], which may instil hopes that (8) might hold for β right up to $\beta_{\mathrm{cond}}(d,k)$. In fact, (8) holds if either d or β is relatively small.

▶ **Lemma 6.** *For $d \leq d_{k-\mathrm{SAT}}$ and $\beta > 0$ let $q \in (0,1)$ be the unique solution to the equation*

$$1 - (1 - \exp(-\beta))q^k = 2(1-q). \tag{10}$$

Then

$$\frac{1}{n} \ln \mathbb{E}\left[Z_\beta(\mathbf{\Phi})\right] \sim \ln 2 + \frac{d}{k} \ln\left(1 - (1-\exp(-\beta))q^k\right) + \frac{d}{2} \ln(4q(1-q)). \tag{11}$$

Furthermore, if either $d \leq d_-(k)$ or $\beta \leq \beta_-(k,d)$ then (8) is true.

However, for d close to $d_{k-\mathrm{SAT}}$ and β near $\beta_{\mathrm{cond}}(d,k)$ the second moment method fails. Formally, if d is such that $\beta_{\mathrm{cond}}(d,k) < \infty$, then there exists $\beta' < \beta_{\mathrm{cond}}(d,k)$ such that (8) is violated for all $\beta \in (\beta', \beta_{\mathrm{cond}}(d,k))$.

3.2 Quenching the average

To understand what goes awry we turn the second moment into a first moment with respect to reweighted distribution. Specifically, the *planted model* is the random pair $(\hat{\mathbf{\Phi}}, \hat{\boldsymbol{\sigma}})$ chosen from the distribution

$$\mathbb{P}\left[(\hat{\mathbf{\Phi}}, \hat{\boldsymbol{\sigma}}) = (\hat{\Phi}, \hat{\sigma})\right] = \frac{\exp(-\beta E_{\hat{\Phi}}(\hat{\sigma}))}{(dn)! \cdot \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)]}. \tag{12}$$

Thus, the probability of that $(\hat{\Phi}, \hat{\sigma})$ comes up is proportional to $\exp(-\beta E_{\hat{\Phi}}(\hat{\sigma}))$. Further, the probability that a specific formula $\hat{\Phi}$ comes up equals $\mathbb{P}[\hat{\mathbf{\Phi}} = \hat{\Phi}] = Z_\beta(\hat{\Phi})/((dn)! \cdot \mathbb{E}[Z_\beta(\mathbf{\Phi})])$, proportional to the partition function. In effect,

$$\mathbb{E}[Z_{\mathbf{\Phi}}(\beta)^2] = \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)] \cdot \mathbb{E}[Z_{\hat{\mathbf{\Phi}}}(\beta)]. \tag{13}$$

Hence, in light of (11) computing the second moment is equivalent to calculating $\mathbb{E}[Z_{\hat{\mathbf{\Phi}}}(\beta)]$.

In fact, the second moment calculation from the proof of Lemma 6 reveals that $\mathbb{E}[Z_{\hat{\mathbf{\Phi}}}(\beta)]$ is dominated by two distinct contributions. First, assignments that are more or less orthogonal to $\hat{\boldsymbol{\sigma}}$ yield a term of order $\mathbb{E}[Z_{\mathbf{\Phi}}(\beta)]$. The second contribution is from $\boldsymbol{\sigma}$ close to $\hat{\boldsymbol{\sigma}}$; say, $\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\sigma}} \geq n(1-2^{-k/10})$. Geometrically, this reflects the fact that the "planted assignment" $\hat{\boldsymbol{\sigma}}$ sits in a "valley" of the Hamiltonian $E_{\hat{\mathbf{\Phi}}}$ w.h.p. The valleys are officially called *clusters* and we let

$$\mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta) = \sum_{\sigma \in \{\pm 1\}^n} \mathbf{1}\{\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\sigma}} > n(1-2^{-k/10})\} \exp(-\beta E_{\hat{\mathbf{\Phi}}}(\sigma)). \tag{14}$$

be the Gibbs-weighted *cluster size*. Hence, (13) shows that the second moment method functions iff $\mathbb{E}[\mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)] \leq \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)]$.

But for d close to $d_{k-\mathrm{SAT}}$ and $\beta > \beta_{\mathrm{cond}}(d,k)$ we have $\mathbb{E}[\mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)] \geq \exp(\Omega(n))\mathbb{E}[Z_{\mathbf{\Phi}}(\beta)]$. In other words, the expected cluster size blows up. At a second glance, this is unsurprising. For the cluster size scales exponentially with n and is therefore prone to large deviations effects. To suppress these we ought to work with $\mathbb{E}[\ln \mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)]$ instead of $\mathbb{E}[\mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)]$. A similar issue (that the expected cluster size explodes) occurred in earlier work on condensation [7, 8, 13, 14]. Indeed, borrowing the idea of a truncated second moment method from these papers, we can reduce the computation of $\phi_{d,k}(\beta)$ to the problem of determining $\mathbb{E}[\ln \mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)]$.

▶ **Lemma 7.** *Equation* (9) *holds iff*

$$\limsup_{n\to\infty} n^{-1}\mathbb{E}[\ln \mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)] \leq \lim_{n\to\infty} n^{-1} \ln \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)]. \tag{15}$$

Hence, we are left to calculate $\mathbb{E}[\ln \mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)]$, the "quenched average" in physics jargon. As we saw the log and the expectation do not commute. In such cases, computing the quenched average is notoriously difficult, certainly well beyond the reach of elementary methods. Tackling this problem is the main achievement of this paper; recall the expressions from (6)–(7).

▶ **Proposition 8.** *Assume that $d \in [d_-(k), d_{k-SAT}]$ and $\beta > \beta_-(k,d)$. Then*

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\ln \mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)] = \mathcal{B}(k,d,\beta), \qquad \textit{while} \qquad \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[Z_{\mathbf{\Phi}}(\beta)] = \mathcal{F}(k,d,\beta).$$

We observe that Theorem 1 is immediate from Lemma 6, Lemma 7 and Proposition 8.

3.3 Non-reconstruction

To calculate the quenched average we are going to have to understand the typical internal structure of the cluster in the planted model. According to the physicists "1-step replica symmetry breaking picture", the restriction of the Gibbs measure to the cluster should enjoy a spatial mixing property called *non-reconstruction*. In particular, the truth values assigned to variables that are "far apart" are predicted to be asymptotically independent.

If non-reconstruction holds, then a general result from [6] reduces the computation of the quenched average to determining the marginals of the restricted Gibbs distribution, which we are going to calculate via Belief Propagation.

Formally, by the restriction of the Gibbs measure to the cluster we mean the probability distribution on $\{\pm 1\}^n$ defined by

$$\sigma \in \{\pm 1\}^n \mapsto \mathbf{1}\{\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\sigma}} > n(1-2^{-k/10})\} \exp(-\beta E_{\hat{\mathbf{\Phi}}}(\sigma))/\mathcal{Z}_{\hat{\mathbf{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta). \tag{16}$$

For a random variable $X(\boldsymbol{\sigma})$ we denote the average with respect to (16) by

$$\langle X(\boldsymbol{\sigma})\rangle' = \langle X(\boldsymbol{\sigma})\rangle'_{\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\sigma}},\beta} = \frac{1}{\mathcal{Z}_{\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)} \sum_{\sigma\in\{\pm1\}^n} \mathbf{1}\{\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\sigma}} > n(1-2^{-k/10})\}\exp(-\beta E_{\hat{\boldsymbol{\Phi}}}(\sigma)).$$

Further, to define a metric we set up a bipartite graph whose vertices are the clauses and variable of $\hat{\boldsymbol{\Phi}}$. Each clause is adjacent to all the variables that it contains. Then the distance between two variables or clauses is, of course, the length of a shortest path in the graph.

We can now state the non-reconstruction condition. For a variable x, an integer $\ell \geq 0$ and $\tau \in \{\pm1\}^n$ let $\nabla(\hat{\Phi}, x, \ell, \tau)$ be the set of all $\chi \in \{\pm1\}^n$ such that $\chi(y) = \tau(y)$ for all y at distance at least 2ℓ from x in $\hat{\boldsymbol{\Phi}}$. Then

$$\left\langle \boldsymbol{\sigma}(x)|\nabla(\hat{\boldsymbol{\Phi}}, x, \ell, \tau)\right\rangle'$$

is the average of the truth value of x once we condition on the event that the truth values of all variables at distance at least 2ℓ from x are given by the "boundary condition" τ. Thus, we inspect the distribution of the truth value of x given the faraway variables.

The non-reconstruction condition requires that for most variables x, $\langle\boldsymbol{\sigma}(x)|\nabla(\hat{\boldsymbol{\Phi}}, x, \ell, \boldsymbol{\tau})\rangle'$ is close to $\langle\boldsymbol{\sigma}(x)\rangle'$ in expectation with respect to a boundary condition $\boldsymbol{\tau}$ that is itself chosen randomly from (16). Formally, $(\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{\sigma}})$ has the *non-reconstruction property w.h.p.* if for any $\varepsilon > 0$ there is $\ell > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n}\left\langle\left|\langle\boldsymbol{\sigma}(x_i)\rangle' - \left\langle\boldsymbol{\sigma}(x_i)|\nabla(\hat{\boldsymbol{\Phi}}, x_i, \ell, \boldsymbol{\tau})\right\rangle'\right|\right\rangle' < \varepsilon\right] = 1. \tag{17}$$

▶ **Proposition 9.** *Assume that $d \in [d_-(k), d_{k-SAT}]$ and $\beta > \beta_-(k,d)$. Then $(\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{\sigma}})$ has the non-reconstruction property w.h.p.*

Together with [6, Theorems 4.4–4.5] Proposition 9 reduces the computation of the quenched average to the problem of computing the marginals under the measure (16). Specifically, $\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\ln \mathcal{Z}_{\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)]$ is given by an expression called the *Bethe free energy* that is a function of the vector $(\langle\boldsymbol{\sigma}(x_i)\rangle')_{i=1,\ldots,n}$ of marginals only.[6] The Bethe free energy originally comes from the physicists cavity method [23, ch. 14].

3.4 A branching process

Hence, we are left to calculate the marginals of (16). Due to the correlation decay guaranteed by the non-reconstruction property, the marginals are governed by the local structure of the formula $\hat{\boldsymbol{\Phi}}$. To facilitate the marginal computation, we are going to condition on the event that the planted assignment $\hat{\boldsymbol{\sigma}} = \mathbf{1}$ is the all-ones vector; this is without loss of generality because under the planted model (12) $\hat{\boldsymbol{\sigma}}$ is uniformly distributed.

Of course, in $\hat{\boldsymbol{\Phi}}$ each variable occurs $d/2$ times positively and $d/2$ times negatively. But the distribution of the signs with which the variables occur in the clauses is non-trivial. We are going to describe it via a branching process with four types: variable nodes of type ±1 and clause nodes of type ±1. Starting from a single variable node r, the process is defined as follows; let $q \in (0,1)$ be the solution to (10).

[6] Stirctly speaking, Proposition 9 and [6, Theorems 4.4 and 4.5] merely imply that the Bethe free energy is an upper bound on $\lim \frac{1}{n}\mathbb{E}[\ln \mathcal{Z}_{\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)]$. To obtain the matching lower bound it is necessary to consider another version of the planted model, see the appendix for details.

BR1: For the root r let $b_{r,\uparrow} = 1$ with probability $1-q$ and $b_{r,\uparrow} = -1$ with probability q.

BR2: Suppose that x is a variable node of type $b_{x,\uparrow} = \pm 1$. Then x has $d-1$ children, which are clause nodes. Specifically, $\frac{d}{2}-1$ children a are clause nodes of type $b_{a,\uparrow} = b_{x,\uparrow}$, and the remaining $d/2$ children are clause ndoes of type $b_{a,\uparrow} = -b_{x,\uparrow}$.

BR3: Suppose that a is a clause node of type $b_{a,\uparrow} = 1$. Then a has $k-1$ children in total, which are variable nodes. Specifically, $X_a = \mathrm{Bin}(k-1, 1-q)$ children have type 1, and the remaining $k-1-X_a$ children have type -1.

BR4: Finally, suppose that a is a clause node of type $b_{a,\uparrow} = -1$. Then a has $k-1$ children, which are variable nodes and

- with probability $\exp(-\beta)q^{k-1}/(1-(1-\exp(-\beta))q^{k-1})$ all children have type -1,
- otherwise $Y_a = \mathrm{Bin}_{\geq 1}(k-1, 1-q)$ children have type 1 and the others have type -1.

Let us write $\boldsymbol{T}_\infty$ be the random infinite tree generated by this branching process (including the type assignment $b_{\cdot,\uparrow}$). Moreover, let $\mathcal{T}_\infty$ be the set of all possible outcomes.

We can think of $\boldsymbol{T}_\infty$ as an infinite k-SAT formula in which all variables other than r appear $d/2$ times positively and $d/2$ times negatively. Namely, for each clause node a we define a Boolean clause whose variables are the parent variable node of a and the $k-1$ children of a. The sign with which the parent x of a occurs in a is precisely $b_{a,\uparrow}$, the type of a. Thus, a contains the literal x if $b_{a,\uparrow} = 1$ and the literal $\neg x$ otherwise. Similarly, each child y of a occurs with sign $b_{y,\uparrow}$.

The root of $\boldsymbol{T}_\infty$ has degree $d-1$ rather than d. This will be useful to set up the Belief Propagation equations below, but to describe the local structure of $\hat{\boldsymbol{\Phi}}$ we actually need a tree in which the root has degree d. Thus, let $\boldsymbol{T}'_\infty$ be the infinite tree defined just as above except that the root has $d/2$ children of type $+1$ and $d/2$ children of type -1. Further, let $\mathcal{T}'_\infty$ be the set of all possible outcomes of this process.

The tree $\boldsymbol{T}'_\infty$ captures the local structure of $\hat{\boldsymbol{\Phi}}$. More precisely, for a formula Φ and a variable x let $\Delta^l_\Phi x$ be the sub-formula obtained from Φ by deleting all clauses and variables at distance at least l from x. Additionally, for a specific formula φ let

$$\rho_{\hat{\boldsymbol{\Phi}}}(\varphi) = \frac{1}{n}\left|\left\{x : \Delta^{2\ell+1}_{\hat{\boldsymbol{\Phi}}} x \cong \varphi\right\}\right|$$

be the fraction of variables x of $\hat{\boldsymbol{\Phi}}$ whose depth-2ℓ neighborhood is isomorphic to φ.

▶ **Lemma 10.** *For all ℓ, φ we have $\mathbb{E}\left|\rho_{\hat{\boldsymbol{\Phi}}}(\varphi) - \mathbb{P}\left[\Delta^{2\ell+1}\boldsymbol{T}'_\infty \cong \varphi\right]\right| = O(n^{-1/2}\ln n)$.*

In light of Lemma 10 we can study the marginals of (16) by way of the random tree $\boldsymbol{T}'_\infty$. Specifically, we are going construct a map $\mathcal{T}'_\infty \to \mathcal{P}(\{\pm 1\})$ that yields a probability measure on $\{\pm 1\}$ for each tree such that the marginal of a variable x is close to the conditional expectation of this map given the depth-2ℓ neighborhood $\Delta^{2\ell+1}_{\hat{\boldsymbol{\Phi}}} x$ for large enough ℓ. It will emerge that this map is intimately related to the fixed point problem from Proposition 2. To construct the map $\mathcal{T}'_\infty \to \mathcal{P}(\{\pm 1\})$ we employ Belief Propagation; for a detailed introduction to Belief Propagation and the physics intuition behind it see [23].

3.5 Belief Propagation

Fix some integer $\ell \geq 1$. Viewing the tree $T \in \mathcal{T}_\infty$ as a k-SAT formula as above, we let $V_{2\ell}$ be the set of all variable nodes at distance at most 2ℓ from the root of T and let $F_{2\ell}$ be the set of all clause nodes at distance at most 2ℓ from the root. Further, let $\partial V_{2\ell}$ be the set of all variable nodes of T at distance exactly 2ℓ from the root. Belief Propagation starts from a *boundary condition* $\partial\nu : \partial V_{2\ell} \to \mathcal{P}(\{\pm 1\})$, $x \mapsto \partial\nu_x$ that assigns each x a probability distribution on $\{\pm 1\}$. The *Belief Propagation messages* induced by the boundary condition $\partial\nu$ on T are the (unique) families

$$(\nu_{x,\uparrow}^{T,\partial\nu})_{x\in V_{2\ell}}, \qquad (\widehat{\nu}_{a,\uparrow}^{T,\partial\nu})_{a\in F_{2\ell}}$$

of probability measures on $\{\pm 1\}$ determined by the following three conditions. For a node u of T let $\partial_\downarrow u$ be the set of children.

BP1: For all $x \in \partial V_{2\ell}$ we have $\nu_{x,\uparrow}^{T,\partial\nu} = \partial\nu_x$.

BP2: For all $x \in V_{2\ell} \setminus \partial V_{2\ell}$ and $s \in \{-1,1\}$,

$$\nu_{x,\uparrow}^{T,\partial\nu}(s) = \frac{\prod_{a\in\partial_\downarrow x} \widehat{\nu}_{a,\uparrow}^{T,\partial\nu}(s)}{\sum_{s'\in\{-1,1\}} \prod_{a\in\partial_\downarrow x} \widehat{\nu}_{a,\uparrow}^{T,\partial\nu}(s')}. \tag{18}$$

BP3: For all $a \in F_{2\ell}$ and $s \in \{-1,1\}$,

$$\widehat{\nu}_{a,\uparrow}^{T,\partial\nu}(s) = \frac{\sum_{s_a\in\{-1,1\}^{\partial a}} \mathbf{1}\{s_x = s\}\, \psi_a(s_a) \prod_{y\in\partial_\downarrow a} \nu_{y,\uparrow}^{T,\partial\nu}(s_y)}{\sum_{s_a\in\{-1,1\}^{\partial a}} \psi_a(s_a) \prod_{y\in\partial_\downarrow a} \nu_{y,\uparrow}^{T,\partial\nu}(s_y)}. \tag{19}$$

Algorithmically, all messages can be calculated bottom-up from the boundary $V_{2\ell}$. The "result" of the Belief Propagation calculation on T given a certain boundary condition is the message emanating from the root:

$$\nu_T^{\partial\nu} = \nu_{r,\uparrow}^{T,\partial\nu}.$$

The fixed point distribution from Proposition 2 can be obtained organically by running Belief Propagation on $\boldsymbol{T}_\infty$. Indeed, define $\partial\nu^{(0)} : \partial V_{2\ell} \to \mathcal{P}(\{\pm 1\})$ by $\partial\nu_x^{(0)}(1) = 1$ for all $x \in \partial V_{2\ell}$ and let $\nu_T^{(2\ell)} = \nu_{r,\uparrow}^{T,\partial\nu^{(0)}}$.

▶ **Proposition 11.** *Assume that $d_-(k) < d \le d_{k-SAT}$ and $\beta > \beta_-(k,d)$. The sequence $(\nu_{\boldsymbol{T}_\infty}^{(2\ell)})_{\ell\ge 1}$ converges almost surely to a limit $\nu_{\boldsymbol{T}_\infty}^\star$. Moreover, $\pi_{k,d,\beta}^\star$ is the distribution of the random variable $\nu_{\boldsymbol{T}_\infty}^\star(-b_{r,\uparrow})$.*

We define the Belief Propagation messages $\nu_{T'}^{(2\ell)}$ for the trees $T' \in \mathcal{T}_\infty'$ in which the root r has degree d exactly as we did above. Of course, the calculation of the messages $\nu_{T'}^{(2\ell)}$ is closely related to that of the messages $\nu_T^{(2\ell)}$ for $T \in \mathcal{T}_\infty$; after all, the only difference occurs at the root. The proof of Proposition 11 shows that the Belief Propagation recurrence enjoys certain contraction properties. In combination with the non-reconstruction property we thus obtain an asymptotic formula for the marginals of the distribution (16).

▶ **Proposition 12.** *Assume that $d_-(k) < d \le d_{k-SAT}$ and $\beta > \beta_-(k,d)$. The sequence $(\nu_{\boldsymbol{T}_\infty'}^{(2\ell)})_{\ell\ge 1}$ converges almost surely to a limit $\nu_{\boldsymbol{T}_\infty'}^\star$. Moreover,*

$$\lim_{\ell\to\infty}\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left|\langle \mathbf{1}\{\boldsymbol{\sigma}(x_i)=1\}\rangle_{\hat{\boldsymbol{\Phi}},\mathbf{1},\beta}' - \mathbb{E}[\nu_{\boldsymbol{T}_\infty'}^\star(1)|\Delta^{2\ell+1}\boldsymbol{T}_\infty' \cong \Delta_{\hat{\boldsymbol{\Phi}}}^{2\ell+1}x]\right| = 0$$

Plugging the asymptotic marginals from Proposition 11 into the Bethe free energy formula, we obtain an expression for the quenched average $\lim \frac{1}{n}\mathbb{E}[\ln \mathcal{Z}_{\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\sigma}}}(\beta)]$. Due to the inherent connection between $\nu_{\boldsymbol{T}_\infty}^\star$ and $\nu_{\boldsymbol{T}_\infty'}^\star$, a (tedious) bit of calculus reveals that this formula can be expressed in terms of the fixed point distribution $\pi_{k,d,\beta}^\star$. The Bethe free energy formula then morphs into the expression $\mathcal{B}(k,d,\beta)$ from (7). In the course of this we also find that (6) matches the "annealed average" $\lim \frac{1}{n}\ln \mathbb{E}[Z_{\boldsymbol{\Phi}}(\beta)]$. Thus Proposition 8 follows.

In the following two sections we outline the two key parts of the proof in some more detail, namely the proof of the non-reconstruction property and the analysis of Belief Propagation on the random tree.

4 Non-reconstruction

We assume that $d \in [d_-(k), d_{k-SAT}]$ and that $\beta \geq \beta_-(k,d)$. Let $c_\beta = 1 - \exp(-\beta)$.

To prove Proposition 9 we exhibit six deterministic conditions that entail the non-reconstruction property. First, a formula Φ on variables $V = \{x_1, \ldots, x_n\}$ satisfies property **ℓ-Local Structure** if

ℓ-Local Structure: for all trees T of height $2\ell + 2$ we have

$$|\rho_\Phi(T) - \mathbb{P}\left[\Delta^{2\ell+3}\boldsymbol{T}'_\infty \cong T\right]| \leq n^{-0.49}.$$

In words, the empirical distribution of the depth-$2\ell + 2$ neighbourhoods is close to the distribution of the random tree $\boldsymbol{T}'$.

The second condition reads

Cycles: The formula Φ contains $o(\sqrt{n})$ cycles of length at most $\sqrt{\ln n}$.

To state the third condition we identify a large "well-behaved" bit of the random formula that we call the *core*. Similar constructions have been used extensively in prior work on random constraint satisfaction problems (e.g., [1, 7, 8]). Let $\partial_{\pm 1}x$ be the set of clauses where the variable x appears as a positive/negative literal and, conversely, let $\partial_{\pm 1}a$ be the set of variables that appear in clause a positively/negatively. Now, the λ-*core* of Φ (in symbols: $\mathrm{Core}_\lambda(\Phi)$) is the largest set W of variables such that all $x \in W$ satisfy the following conditions.

CR1: there are at least $\lambda^{-1}k^{0.99}$ clauses $a \in \partial_1 x$ such that $\partial_1 a = \{x\}$.
CR2: there are no more than 10λ clauses $a \in \partial x$ such that $|\partial_{-1}a| = k$.
CR3: for any $1 \leq l \leq k$ the number of $a \in \partial_{-1}x$ such that $|\partial_1 a| = l$ is bounded by $\lambda k^{l+3}/l!$.
CR4: there are no more than $\lambda k^{3/4}$ clauses $a \in \partial_1 x$ such that $|\partial_1 a| = 1$ but $\partial a \not\subset W$.
CR5: there are no more than $\lambda k^{3/4}$ clauses $a \in \partial_{-1}x$ such that $|\partial_{-1}a| < k$ and $|\partial_1 a \setminus W| \geq |\partial_1 a|/4$.

The λ-core is well-defined; for if W, W' satisfy the above conditions, then so does $W \cup W'$. Further, if $\lambda < \lambda'$, then $\mathrm{Core}_\lambda(\Phi) \subset \mathrm{Core}_{\lambda'}(\Phi)$. The formula Φ has the property **λ-Core** if

λ-Core: $|\mathrm{Core}_\lambda(\Phi)| \geq (1 - 2^{-0.95k})n$.

We are going to identify a large set $V_{\mathrm{good}} \subset V$ of variables that are very likely to be set to one under a typical assignment $\boldsymbol{\sigma}$ chosen from $\langle \cdot \rangle'_{\Phi,\beta}$. As a first attempt we might try $V_{\mathrm{good}} = \mathrm{Core}_{1/2}(\Phi)$. However, the conditions **CR1–CR5** are not quite strong enough to enable an estimate of the $\langle \cdot \rangle'_{\Phi,\beta}$-marginals. For instance, if the marginals of most of the neighbors of a given vertex $x \in \mathrm{Core}_{1/2}(\Phi)$ go astray, x will likely follow suit. Yet the variables x in the core such that $\langle \boldsymbol{\sigma}(x) \rangle'_{\Phi,\beta}$ is "small" must clump together. Formally, we say that a set $S \subset V$ is λ-*sticky* if for all $x \in S$ one of the following conditions holds.

ST1: There are at least $\lambda k^{3/4}$ clauses $a \in \partial_1 x$ such that $\partial_1 a = \{x\}$ and $\partial_{-1}a \cap S \neq \emptyset$.
ST2: Yhere are at least $\lambda k^{3/4}$ clauses $a \in \partial_{-1}x$ such that $|\partial_{-1}a| < k$ and $|\partial_1 a \cap S| \geq |\partial_1 a|/4$.

Further, Φ satisfies the property **λ-Sticky** if

λ-Sticky: Φ has no λ-sticky set of size between $2^{-0.95k}n$ and $2^{-k/20}n$.

The condition **Sticky** ensures that for $\lambda \in \{1/2, 1\}$ there is a unique maximal λ-sticky set $S_\lambda(\Phi) \subset \mathrm{Core}_\lambda(\Phi)$ of size $S_\lambda(\Phi) \leq 2^{-0.1k}n$. Indeed, if $S, S' \subset \mathrm{Core}_\lambda(\Phi)$ are two λ-sticky sets of size at most $2^{-0.1k}n$, then $S \cup S'$ is sticky as well. Consequently, **Sticky** guarantees that $|S \cup S'| \leq 2^{-0.95k}n$. In fact, this argument shows that $S_\lambda(\Phi) \leq 2^{-0.95k}n$.

Further, the next condition reads

Gap: $\left\langle \mathbf{1}\{\boldsymbol{\sigma}\cdot\mathbf{1} < (1-2^{-k/3})n\}\right\rangle' \leq \exp(-\Omega_n(n))$.
Hence, comparing the above with (14), we realise that **Gap** requires that assignments $\boldsymbol{\sigma}$ with $1-2^{-k/10} < \boldsymbol{\sigma}\cdot\mathbf{1}/n < 1-2^{-k/3}$ contribute little to the cluster size.

Finally, we come to the seventh and last condition. A variable $x \in V$ is $(\varepsilon, 2\ell)$-*cold* if the following two conditions are satisfied. Write $\partial^{2\ell}x$ for the set of all variables at distance exactly 2ℓ from x.
CD1: The sub-formula $T = \Delta^{2\ell+1}x$ is a tree.
CD2: If $\boldsymbol{\tau} : \partial^{2\ell}x \to \{\pm 1\}$ is a random assignment such that independently for all $y \in \partial^{2\ell}x$,

$$\boldsymbol{\tau}(y) = \begin{cases} -1 & \text{if } y \notin \mathrm{Core}_1(\Phi) \cup S_1(\Phi) \\ (-1)^{\mathrm{Be}(\exp(-k^{0.9}\beta))} & \text{otherwise} \end{cases},$$

then

$$\mathbb{E}\left[\max\left\{\left|\nu_T^{(2\ell)}(1) - \langle \mathbf{1}\{\boldsymbol{\sigma}(x)=1\}|\nabla(\Phi,x,\ell,\tau)\rangle'\right| : \tau \geq \boldsymbol{\tau}\right\}\right] \leq \varepsilon. \tag{20}$$

In words, suppose that we choose a random "boundary condition" $\boldsymbol{\tau}$ such that all y at distance 2ℓ from x that do not belong to the core are set to -1 and all y in the core are set to -1 with probability $\exp(-k^{0.9}\beta)$ independently. Then an adversary comes along and obtains τ from $\boldsymbol{\tau}$ maliciously by setting $\tau(y) = 1$ for a few y such that $\boldsymbol{\tau}(y) = -1$. (The adversary is not allowed to make changes in the opposite direction.) Then (20) requires that the spin $\boldsymbol{\sigma}(x)$ given the boundary condition τ be close to the Belief Propagation marginal $\nu_T^{(2\ell)}(1)$. Of course, the expectation in (20) is over $\boldsymbol{\tau}$ only.
$(\varepsilon, 2\ell)$-Cold: All but εn variables are $(\varepsilon, 2\ell)$-cold.

A formula Φ is $(\varepsilon, \ell, \lambda)$-*quasirandom* if the properties **ℓ-Local Structure**, **Cycles**, **λ-Core**, **λ-Sticky**, **Gap** and **$(\varepsilon, 2\ell)$-Cold** hold.

▶ **Proposition 13.** *For any $\varepsilon > 0$ there is $\ell > 0$ such that w.h.p. $\hat{\boldsymbol{\Phi}}$ is $(\varepsilon, \ell, 1)$-quasirandom.*

The proof that $\hat{\boldsymbol{\Phi}}$ has the first five properties w.h.p. is based on standard arguments. But the proof of the $(\varepsilon, 2\ell)$**-Cold** property is novel. The argument is intertwined with the study of the Belief Propagation recurrence on the random tree. In particular, that analysis, which we sketch in Section 5, is via a contraction argument that enables a comparison between the result $\nu_T^{\partial\nu}$ for a given bounardy condition and the result for the all-ones boundardy condition (Lemma 20 below). In order to transfer this result from the random tree to the random formula, in which the boundary condition depends on the core, we use a switching argument. The details can be found in the appendix.

Proposition 9 is immediate from Proposition 13 and the following statement.

▶ **Proposition 14.** *For any $\delta > 0$ there exists $\varepsilon > 0$ and $\ell_0(\varepsilon) > 0$ such that for any $\ell > \ell_0(\varepsilon)$ there exists $n_0(\varepsilon, \ell)$ such that for all $n > n_0$ the following is true. If Φ is (ε, ℓ)-quasirandom, then*

$$\frac{1}{n}\sum_{i=1}^{n}\left\langle\left|\nu^{(2\ell)}_{\Delta_\Phi^{2\ell}x_i}(1) - \langle \mathbf{1}\{\boldsymbol{\sigma}(x_i)=1\}|\nabla(\Phi,x_i,\ell,\boldsymbol{\tau})\rangle'\right|\right\rangle'_{\Phi,\beta} < \delta.$$

Proof of Proposition 14

Assume that Φ is (ε, ℓ)-quasirandom and that $n > n_0$ for some large $n_0 = n_0(\varepsilon, \ell)$. In particular, $|\mathrm{Core}_1(\Phi)| \geq (1 - 2^{0.95k})n$. Let $\sigma : V \to \{\pm 1\}$ be an assignment. A set $T \subset \mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$ is σ-*closed* if for any $x \in T$ and all $a \in \partial x$ we have

$$\{y \in \partial a \cap \mathrm{Core}_1(\Phi) \setminus S_1(\Phi) : \sigma(y) = -1\} \subset T. \tag{21}$$

Hence, if $y \in \mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$ is set to -1 and there is a clause a connecting y to some $x \in T$, then y itself must be in T. Moreover, for a clause b we say $T \subset \mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$ is (σ, b)-*closed* if (21) holds for all $x \in T$ and all $a \in \partial x \setminus b$. Additionally, let

$$\partial_{\pm 1,l} x = \{a \in \partial_{\pm 1} x : |\partial_1 a| = l\}$$

be the set of clauses a with a total number of l positive literals where x appears positively/negatively.

▶ **Lemma 15.** *Suppose that Φ is (ε, ℓ)-quasirandom. Then for any σ such that $\mathbf{1} \cdot \sigma \geq (1 - 2^{-k/9})n$ and for any (σ, b)-closed set $T \subset \mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$ the following is true. Let $\tilde{\sigma}(x) = (-1)^{\mathbf{1}\{x \in T\}} \sigma(x)$. Then*

$$E_\Phi(\tilde{\sigma}) \leq E_\Phi(\sigma) - k^{0.98}|T|. \tag{22}$$

Proof. Consider the following process:
- Let $\sigma_0 = \sigma$, $V_0 = T$ and $U_0 = \sigma^{-1}(-1) \setminus V_0$.
- While there is $i_t \in V_t$ such that $E_\Phi((-1)^{\mathbf{1}\{\,\cdot\, = i_t\}} \sigma_t(\,\cdot\,)) \leq E_\Phi(\sigma_t) - k^{0.98}$, pick one such i_t arbitrarily and let $\sigma_{t+1}(\,\cdot\,) = (-1)^{\mathbf{1}\{\,\cdot\, = i_t\}} \sigma_t(\,\cdot\,)$ and $V_{t+1} = V_t \setminus \{i_t\}$.

Clearly,

$$E_\Phi(\sigma_t) \leq E_\Phi(\sigma) - k^{0.98} t. \tag{23}$$

Let τ be the stopping time of this process and assume that $\tau < |T|$, or, in other words, that $V_\tau \neq \emptyset$. We claim that V_τ is a 1-sticky set. Indeed, because T is σ-closed for $i \in V_\tau$ we have

$$\begin{aligned} -k^{0.98} &\leq E_\Phi((-1)^{\mathbf{1}\{\,\cdot\, = i\}} \sigma_t(\,\cdot\,)) - E_\Phi(\sigma_\tau) \\ &\leq \mathbf{1}\{b \in \partial i\} - |\partial_{1,0}(i)| + |\{a \in \partial_{1,0} i, \partial_{-1} a \cap (V_\tau \cup U_0) \neq \emptyset\}| \\ &\qquad + |\partial_{-1,0} i| + |\cup_{1 \leq l \leq k} \{a \in \partial_{-1,l} i, \partial_1 a \subset V_\tau \cup U_0)\}|. \end{aligned}$$

Because $i \in \mathrm{Core}_1(\Phi)$ we have $|\partial_{1,0} i| \geq k^{0.99}$, $|\partial_{-1,0} i| \leq 10$, $|\{a \in \partial_{1,0} i, \partial_{-1} a \cap U_0 \neq \emptyset\}| \leq k^{3/4}$ and $|\{a \in \partial_{1,0} i, |\partial_{-1} a \cap U_0| \geq |\partial_{-1} a|/4\}| \leq k^{3/4}$. Therefore, one of the following must hold.

(a) $|\{a \in \partial_{1,0}, \partial_{-1} a \cap V_\tau \neq \emptyset\}| \geq k^{3/4}$,
(b) $|\{a \in \partial_{1,0} i, |\partial_{-1} a \cap V_\tau| \geq |\partial_{-1} a|/4\}| \geq k^{3/4}$.

It follows that the set $V_\tau \subset T \subset \mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$ is 1-sticky.

However, $\mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$ cannot contain a 1-sticky set of size $|V_\tau| \leq |T| \leq 2^{-k/10}$ as this would contradict the maximality of $S_1(\Phi)$. It follows that $\tau = |T|$, and therefore $\sigma_\tau = \tilde{\sigma}$, whence (22) follows using (23). ◀

Lemma 15 is going to be our principal tool to establish Proposition 14. To put it to work, we need the following simple observation that follows from the fact that Φ is d-regular.

▶ **Fact 16.** *For any variable x the following is true. Let $\gamma(x, L)$ be the number of trees with $L \geq 1$ vertices rooted at x that are contained in the factor graph of Φ. Then $\gamma(x, L) \leq L(100dk)^L$.*

Write $T(x, \sigma)$ for the smallest σ-closed set that contains x. If $\sigma(x) = 1$ we let $T(x, \sigma) = \emptyset$. The following lemma shows that $T(x, \sigma)$ is unlikely to be non-empty and very unlikely to be large.

▶ **Lemma 17.** *For all $x \in \mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$ we have*

$$\langle \boldsymbol{\sigma}(x) \rangle' \geq 1 - \exp(-\beta k^{0.97}) \quad \textit{and} \quad \langle \mathbf{1}\{|T(x, \boldsymbol{\sigma})| > \ln \ln n\} \rangle' \leq 1/\ln n.$$

Proof. Let $N = 2^{-k/4}n$. Due to **Gap** we have

$$\langle \mathbf{1}\{\mathbf{1} \cdot \boldsymbol{\sigma} < n - N/2\}\rangle' \leq \exp(-\Omega(n)).$$

Therefore,

$$\langle \mathbf{1}\{|T(x, \boldsymbol{\sigma})| > N\}\rangle' \leq \exp(-\Omega(n)). \tag{24}$$

Hence, let $t \leq N$ and let θ be a tree of order t with root x that is contained in the factor graph of Φ and whose vertices lie in $\mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$. If σ is such that $T(x, \sigma) = \theta$, then Lemma 15 implies that $\tilde{\sigma}(x) = (-1)^{\mathbf{1}\{x \in T(x,\sigma)\}}\sigma(x)$ satisfies $E_\Phi(\tilde{\sigma}) \leq E_\Phi(\sigma) - k^{0.98}t$. Consequently,

$$\frac{\langle \mathbf{1}\{\boldsymbol{\sigma} = \sigma\}\rangle'}{\langle \mathbf{1}\{\boldsymbol{\sigma} = \tilde{\sigma}\}\rangle'} \leq \exp(-\beta k^{0.98} t).$$

Hence, by Fact 16, the union bound and our assumptions on β and d,

$$\frac{\langle \mathbf{1}\{|T(x, \boldsymbol{\sigma})| = t\}\rangle'}{\langle \mathbf{1}\{\boldsymbol{\sigma}(x) = 1\}\rangle'} \leq t(100dk)^t \exp(-\beta k^{3/4} t) \leq \exp(-0.99\beta k^{0.98} t). \tag{25}$$

This bound readily implies the second assertion. To obtain the first assertion, we remember (24) and sum (25) over $1 \leq t \leq N$. ◀

Let r be a variable with depth-2ℓ neighborhood T. Guided by (20), we call $\tilde{\tau} : V \to \{\pm 1\}$ a *good boundary condition* for r if

$$\max\left\{\left|\nu_T^{(2\ell)}(1) - \langle \mathbf{1}\{\boldsymbol{\sigma}(x) = 1\}|\nabla(\Phi, x, \ell, \tau)\rangle'\right| : \tau \geq \tilde{\tau}\right\} \leq \varepsilon. \tag{26}$$

▶ **Lemma 18.** *Let r be a variable for which the following conditions hold.*
1. *r is $(\varepsilon, 2\ell)$-cold.*
2. *r has distance at least $\ln^{1/3} n$ from any cycle of length at most $\sqrt{\ln n}$.*

Let Γ_r be the event that $\boldsymbol{\sigma}$ is a good boundary condition for r. Then $\langle \mathbf{1}\{\boldsymbol{\sigma} \notin \Gamma_r\}\rangle' \leq 2\varepsilon$.

Proof. Let $X = (\partial^{2\ell} r) \cap \mathrm{Core}_1(\Phi) \setminus S_1(\Phi)$. Moreover, let $\mathcal{A}$ be the event that

$$\max_{x \in X} |T(x, \boldsymbol{\sigma})| \leq \ln\ln n \quad \text{and} \quad \boldsymbol{\sigma} \cdot \mathbf{1} \geq (1 - 2^{-k/4})n.$$

Because $|X| \leq (dk)^\ell < \ln\ln n$ by our assumption that $n > n_0(\varepsilon, \ell)$, Lemma 17 and the fact that Φ is quasirandom imply $\langle \mathbf{1}\{\boldsymbol{\sigma} \in \mathcal{A}\}\rangle' \sim 1$. Furthermore, if $\mathcal{A}$ occurs, then assumption (2) ensures that the subgraph of the factor graph induced on $Y = (\Delta^{2\ell+1} r) \cup \bigcup_{x \in X} T(x, \boldsymbol{\sigma})$ is acyclic.

Now, fix a variable $x \in X$ and $\sigma \in \mathcal{A}$ such that $\sigma(x) = -1$. Let a be the clause that is adjacent to x on its shortest path to r and let $T(x, a, \sigma)$ be the smallest (σ, a)-closed set that contains x. Further, define

$$\tilde{\sigma}(y) = (-1)^{\mathbf{1}\{y \in T(x,a,\sigma)\}}\sigma(y).$$

Then Lemma 15 shows that $E_\Phi(\tilde{\sigma}) \leq k^{0.98}|T(x, a, \sigma)|$. Moreover, because the subgraph induced on Y is acyclic we have $\tilde{\sigma}(x') = \sigma(x')$ for all $x' \in X \setminus \{x\}$. Consequently, by Fact 16 and the union bound,

$$\begin{aligned} \frac{\left\langle \mathbf{1}\{\boldsymbol{\sigma}(x) = -1\} \prod_{y \in X \setminus \{x\}} \mathbf{1}\{\boldsymbol{\sigma}(y) = \sigma(y)\} \mathbf{1}\{\boldsymbol{\sigma} \in \mathcal{A}\}\right\rangle'}{\left\langle \mathbf{1}\{\boldsymbol{\sigma}(x) = 1\} \prod_{y \in X \setminus \{x\}} \mathbf{1}\{\boldsymbol{\sigma}(y) = \sigma(y)\}\right\rangle'} &\leq \sum_{t \leq \ln\ln n} \frac{t(100dk)^t}{\exp(\beta k^{0.98} t)} \\ &\leq \exp(-\beta k^{0.98}/2). \end{aligned} \tag{27}$$

Since $\langle \mathbf{1}\{\boldsymbol{\sigma} \in \mathcal{A}\}\rangle' = 1 - o_n(1)$ and because for all $\tau : X \to \{\pm 1\}$ we have

$$\left\langle \prod_{y \in X \setminus \{x\}} \mathbf{1}\{\boldsymbol{\sigma}(y) = \tau(x)\} \right\rangle' \geq \exp(-dk\beta|X|) = \Omega_n(1),$$

(27) implies that for any $\tau : X \to \{\pm 1\}$,

$$\frac{\left\langle \mathbf{1}\{\boldsymbol{\sigma}(x) = -1\} \prod_{y \in X \setminus \{x\}} \mathbf{1}\{\boldsymbol{\sigma}(y) = \tau(y)\} \right\rangle'}{\left\langle \mathbf{1}\{\boldsymbol{\sigma}(x) = 1\} \prod_{y \in X \setminus \{x\}} \mathbf{1}\{\boldsymbol{\sigma}(y) = \tau(y)\} \right\rangle'} \leq \exp(-\beta k^{0.98}/3). \tag{28}$$

Finally, the assertion follows from (28) and the assumption that r is $(\varepsilon, 2\ell)$-cold. ◀

Proof of Proposition 14. The condition **Cycles** ensures that there are at most $o_n(n)$ variables r for which condition (2) from Lemma 18 is violated. Furthermore, due to $(\varepsilon, 2\ell)$-**Cold** all but εn variables r satisfy assumption (1). Therefore, the assertion follows from Lemma 18. ◀

5 Belief Propagation on the infinite tree

Assume that $d \in [d_-(k), d_{k-SAT}]$ and that $\beta \geq \beta_-(k,d)$. Let $c_\beta = 1 - \exp(-\beta)$.

We sketch the analysis of the Belief Propagation messages on the random tree $\boldsymbol{T}_\infty$ to prove Propositions 2 and 11. The key step is the proof of the following statement.

▶ **Lemma 19.** *There exists a number $\ell_0 = \ell_0(d,k,\beta)$ such that for all $\ell \geq \ell_0$ the following is true. Suppose that $\boldsymbol{\partial\nu} : \partial V_{2\ell} \to \{\pm 1\}$ is a random boundary condition, independent of $\boldsymbol{T}_\infty$, such that*
H: *for any $x \in \partial V_{2\ell}$, $\mathbb{P}\left[(\boldsymbol{\partial\nu})_x(1) \leq 1 - \exp(-k^{0.9}\beta)\,|(\boldsymbol{\partial\nu})_{y \neq x}\right] \leq 2^{-0.9k}$.*
Then

$$\mathbb{P}\left[\|\nu_{\boldsymbol{T}}^{\boldsymbol{\partial\nu}} - \nu_{\boldsymbol{T}}^{(2\ell)}\|_{\mathrm{TV}} \geq 2\ell^{-1}\right] \leq \ell^{-1}.$$

Thus, the condition **H** provides that for any x on the boundary the message $\boldsymbol{\partial\nu}_x(1)$ is likely close to one, even given $\boldsymbol{T}_\infty$ and all the other boundary messages $(\boldsymbol{\partial\nu}_y)_{y \neq x}$. Further, Lemma 19 states that the message at the root given that the boundary condition satisfies **H** is likely within $O_\ell(\ell^{-1})$ of the message obtained from the "all-ones" boundary condition.

We will need a version of Lemma 19 for the random tree $\boldsymbol{T}'_\infty$. Condition **H** becomes **H'**: for any $x \in \partial V'_{2\ell}$, $\mathbb{P}\left[(\boldsymbol{\partial\nu})_x(1) \leq 1 - \exp(-k^{0.9}\beta)\,|(\boldsymbol{\partial\nu})_{y \neq x}\right] \leq 2^{-0.9k}$.

▶ **Lemma 20.** *There is $\ell_0 = \ell_0(d,k,\beta) > 0$ such that for all $\ell \geq \ell_0$ the following is true. Assume that the random boundary condition $\boldsymbol{\partial\nu''}$, independent of $\boldsymbol{T}'$, satisfies $\boldsymbol{H'}$. Moreover, assume that $\boldsymbol{\partial\nu'}$ is a random boundary condition that may depend on $\boldsymbol{T}', \boldsymbol{\partial\nu''}$ such that $\boldsymbol{\partial\nu''}_x(1) \geq \boldsymbol{\partial\nu}'_x(1)$ for all $x \in \partial V'_{2\ell}$. Then for $\ell \geq \ell_0$ we have*

$$\mathbb{P}\left[\|\nu_{\boldsymbol{T}'}^{\boldsymbol{\partial\nu'}} - \nu_{\boldsymbol{T}'}^{(2\ell)}\|_\infty \geq 2\exp(dk\beta)\ell^{-1}\right] \leq 2dk\ell^{-1}.$$

To grasp the assumptions of Proposition 20, we may think of $\boldsymbol{\partial\nu}'$ as obtained from $\boldsymbol{\partial\nu}''$ by allowing an "adversary" to switch some of the -1s of $\boldsymbol{\partial\nu}''$ to $+1$s. The adversary knows both $\boldsymbol{T}'$ and $\boldsymbol{\partial\nu}''$. In the following we tacitly assume that $\ell \geq \ell_0$ for a large constant ℓ_0.

To prove Lemma 19–20 we are going to exhibit a deterministic condition on $(T, \partial\nu)$ that ensures that $\nu_T^{\partial\nu}$ is close to $\nu_T^{(2\ell)}$. For a variable node x we let $\partial_1 x$ be the set of all clauses in which x appears as a positive literal. Similarly, $\partial_{-1}x$ is the set of clauses containing the literal $\neg x$. Conversely, for a clause a we let $\partial_{\pm 1}a$ be the set of all variables that appear in a positively/negatively. Further, we define the *trunk* of T under the boundary condition $\partial\nu$, $\mathrm{Trunk}(T, \partial\nu)$, as the largest subset W of $V_{2\ell}$ such that for any $x \in W$ either

TR0: $x \in \partial V_{2\ell}$ and $\partial\nu_x(1) \geq 1 - \exp(-k^{0.9}\beta)$

or all of the five following conditions hold

TR1: there are at least $\lfloor 2k^{0.9} \rfloor$ clauses $a \in \partial_{\downarrow}x$ such that $\partial_1 a = \{x\}$.

TR2: there are no more than $\lceil \ln k \rceil$ clauses $a \in \partial x$ such that $|\partial_{-1}a| = k$.

TR3: for any $1 \leq l \leq k$ the number of $a \in \partial_{-1}x$ such that $|\partial_1 a| = l$ is bounded by $k^{l+3}/l!$.

TR4: there are no more than $k^{3/4}$ clauses $a \in \partial_1 x$ such that $|\partial_1 a| = 1$ but $\partial a \not\subset W$.

TR5: there are no more than $k^{3/4}$ clauses $a \in \partial_{-1}x$ such that $|\partial_{-1}a| < k$ and $|\partial_1 a \setminus W| \geq |\partial_1 a|/4$.

The trunk is well-defined; for if W, W' are sets that satisfy the above conditions, then so is $W \cup W'$. Somewhat unbotanically, the trunk is non-empty only if it contains some of the leaves. In fact, the construction is monotonous with respect to the boundary condition:

$$\text{if } \partial\nu_x(1) \leq \partial\nu'_x(1) \text{ for all } x \in \partial V_{2\ell}, \text{ then } \mathrm{Trunk}(T, \nu) \subset \mathrm{Trunk}(T, \nu'). \tag{29}$$

For $T \in \mathcal{T}_{2\ell}$ with root r and $x \in \partial V_{2\ell}$, we denote by $[x \to r]$ the shortest path from x to r in T. Moreover,

1. a variable node $x \in V_{2\ell}$ is *cold* if $x \in \mathrm{Trunk}(T, \partial\nu)$,
2. a clause node $a \in F_{2\ell}$ is *cold* if $\partial_1 a \cap \mathrm{Trunk}(T, \partial\nu) \neq \emptyset$,
3. the pair (x, a) with $x \in \partial_{\downarrow}a$ is *cold* if x is cold or a is cold,
4. a path $[x \to r]$ from $x \in \partial V_{2\ell}$ to r is *cold* if it contains at least $\lfloor 0.4\ell \rfloor$ cold pairs (x, a),
5. the pair $(T, \partial\nu) \in \mathcal{T}_{2\ell} \times \mathcal{P}(\{-1, 1\})^{\partial V_{2\ell}}$ is *cold* if all the paths $[x \to r]$ with $x \in \partial V_{2\ell}$ are cold.

The following estimate shows the use of these concepts.

▶ **Lemma 21.** *Assume that $\ell \geq \ell_0(d, k, \beta)$ is sufficiently large. If*

$$(T, \partial\nu) \in \mathcal{T}_{2\ell} \times \mathcal{P}(\{-1, 1\})^{\partial V_{2\ell}}$$

is cold, then

$$\|\nu_{T,\uparrow}^{\partial\nu} - \nu_{T,\uparrow}^{(2\ell)}\|_\infty \leq \ell^{-1}.$$

The proof of Lemma 21 is based on a (technically delicate) contraction argument. More precisely, the basic insight is that the Belief Propagation operation is a contraction along cold paths. The proof is based on estimating the derivatives of the formulas (18) and (19). Once contraction is established, convergence of the messages to a unique limit follows just as in the Banach fixed point theorem. Furthermore, a “classical” analysis of the random tree based on Chernoff bounds etc. yields

▶ **Lemma 22.** *Assume that $\ell \geq \ell_0(d, k, \beta)$ is sufficiently large and that $\boldsymbol{\partial\nu}$ satisfies **H**. Then*

$$\mathbb{P}\left[(\boldsymbol{T}, \boldsymbol{\partial\nu}) \text{ is cold}\right] \geq 1 - \ell^{-1}.$$

Proof of Lemma 19. The assertion follows from the combination of Lemmas 21 and 22 directly. ◀

Proof of Lemma 20. For $h = 1, \dots, (k-1)d$ consider the sub-trees $\boldsymbol{T}^{(h)}$ of $\boldsymbol{T}'$ pending on the variables at distance exactly two from the root r of $\boldsymbol{T}'$. Then with $\boldsymbol{\partial\nu}^{(h)}$ denoting the boundary condition on $\boldsymbol{T}^{(h)}$ induced by $\boldsymbol{\partial\nu}''$, (29) and Lemma 22 yield

$$\mathbb{P}\left[(\boldsymbol{T}^{(h)}, \boldsymbol{\partial\nu}^{(h)}) \text{ is cold }\right] \geq 1 - (\ell - 1)^{-1}.$$

Hence, Lemma 21 yields

$$\mathbb{P}\left[\|\nu_{\boldsymbol{T}^{(h)}}^{\boldsymbol{\partial\nu}^{(h)}} - \nu_{\boldsymbol{T}^{(h)}}^{(2\ell+1)}\|_{\mathrm{TV}} \leq \ell^{-1}\right] \geq 1 - (\ell - 1)^{-1}.$$

Therefore, the assertion follows from a coupling argument. ◀

The convergence of the sequence $(\nu_{\boldsymbol{T}_\infty}^{(2\ell)})_{\ell \geq 1}$ follows from Lemma 19 rather directly by induction on ℓ. Similarly, the existence and uniqueness of the distributional fixed point from Proposition 2 follows from Lemma 19, albeit with a bit of work. In fact, it is straightforward to verify that the law of $\nu_{\boldsymbol{T}_\infty}^{\star}(b_{r,\uparrow})$ is a fixed point for Proposition 2. Conversely, any fixed point distribution for Proposition 2 can be "unravelled" to obtain a $\mathcal{P}(\{\pm 1\})$-valued random variable $\lambda_{\boldsymbol{T}_\infty}^{\star}$ such that the original distribution is the law of $\lambda_{\boldsymbol{T}_\infty}^{\star}(b_{r,\uparrow})$ that satisfies condition **H**. The contraction property from Lemma 19 therefore implies that $\lambda_{\boldsymbol{T}_\infty}^{\star}$ coincides with $\nu_{\boldsymbol{T}_\infty}^{\star}$ almost surely. Hence the uniqueness of the distributional fixed point. Similarly, the proof of the $(\varepsilon, 2\ell)$**-Cold** property required in the non-reconstruction argument is based on Lemma 20.

References

1. D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proc. 49th FOCS*, pages 793–802, 2008.
2. D. Achlioptas and C. Moore. Random k-sat: two moments suffice to cross a sharp threshold. *SIAM Journal on Computing*, 36:740–762, 2006.
3. D. Achlioptas, A. Naor, and Y. Peres. Rigorous location of phase transitions in hard optimization problems. *Nature*, 435:759–764, 2005.
4. D. Achlioptas and Y. Peres. The threshold for random k-sat is $2^k \ln 2 - o(k)$. *Journal of the AMS*, 17:947–973, 2004.
5. J. Banks and C. Moore. Information-theoretic thresholds for community detection in sparse networks. *Information-theoretic thresholds for community detection in sparse networks*, arXiv:1601.02658, 2016.
6. V. Bapst and A. Coja-Oghlan. Harnessing the bethe free energy. *Harnessing the Bethe free energy*, arXiv:1504.03975, 2015.
7. V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Raßmann, and D. Vilenchik. The condensation phase transition in random graph coloring. *Communications in Mathematical Physics*, 341:543–606, 2016.
8. V. Bapst, A. Coja-Oghlan, and F. Raßmann. A positive temperature phase transition in random hypergraph 2-coloring. *Annals of Applied Probability*, 26:1362?1406, 2016.
9. B. Barak. Structure vs. combinatorics in computational complexity, 2013.
10. M. Bayati, D. Gamarnik, and P. Tetali. Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. *Annals of Probability*, 41:4080–4115, 2013.
11. A. Coja-Oghlan and K. Panagiotou. Going after the k-sat threshold. In *Proc. 45th STOC*, pages 705–714, 2013.
12. A. Coja-Oghlan and K. Panagiotou. The asymptotic k-sat threshold. *Advances in Mathematics*, 288:985–1068, 2016.

13 A. Coja-Oghlan and L. Zdeborová. The condensation transition in random hypergraph 2-coloring. In *Proc. 23rd SODA*, pages 241–250, 2012.

14 P. Contucci, S. Dommers, C. Giardina, and S. Starr. Antiferromagnetic potts model on the Erdős-Rényi random graph. *Communications in Mathematical Physics*, 323:517–554, 2013.

15 A. Dembo, A. Montanari, A. Sly, and N. Sun. The replica symmetric solution for potts models on d-regular graphs. *Comm. Math. Phys.*, 327:551–575, 2014.

16 J. Ding, A. Sly, and N. Sun. Maximum independent sets on random regular graphs. *Maximum independent sets on random regular graphs*, arXiv:1310.4787, 2013.

17 J. Ding, A. Sly, and N. Sun. Satisfiability threshold for random regular nae-sat. In *Proc. 46th STOC*, pages 814–822, 2014.

18 J. Ding, A. Sly, and N. Sun. Proof of the satisfiability conjecture for large k. In *Proc. 47th STOC*, pages 59–68, 2015.

19 D. Gamarnik, T. Nowicki, and G. Swirszcz. Maximum weight independent sets and matchings in sparse random graphs: Exact results using the local weak convergence method. *Random Structures and Algorithms*, 28:76–106, 2006.

20 D. Gamarnik and M. Sudan. Limits of local algorithms over sparse random graphs. In *Proc. 5th ITCS*, pages 369–376, 2014.

21 D. Gamarnik and M. Sudan. Performance of the survey propagation-guided decimation algorithm for the random nae-k-sat problem. *Performance of the Survey Propagation-guided decimation algorithm for the random NAE-K-SAT problem*, arXiv:1402.0052, 2014.

22 F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborova. Gibbs states and and the set of solutions of random constraint satisfaction problems. *Proc. National Academy of Sciences*, 104:10318–10323, 2007.

23 M. Mézard and A. Montanari. *Information, physics and computation.* Oxford University Press, 2009.

24 M. Mézard, G. Parisi, and M. Virasoro. *Spin Glass Theory and Beyond.* World Scientific, 1987.

25 M. Mézard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297:812–815, 2002.

26 E. Mossel, J. Neeman, and A. Sly. Reconstruction and estimation in the planted partition model. *Probability Theory and Related Fields*, pages 1–31, 2014.

27 D. Panchenko. On the replica symmetric solution of the k-sat model. *Electron. J. Probab.*, 19:1–17, 2014.

28 V. Rathi, E. Aurell, L. K. Rasmussen, and M. Skoglund. Bounds on threshold of regular random k-sat. In *Proc. 12th SAT*, pages 264–277, 2010.

29 F. Ricci-Tersenghi and G. Semerjian. On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms. *J. Stat. Mech.*, 9001, 2009.

30 T. Richardson and R. Urbanke. *Modern coding theory.* Cambridge University Press, 2008.

31 M. Talagrand. The high temperature case for the random k-sat problem. *Probab. Theory Related Fields*, 119:187–212, 2001.

32 L. Zdeborová and F. Krzakala. Statistical physics of inference: Thresholds and algorithms. *Statistical physics of inference: thresholds and algorithms*, arXiv:1511.02476, 2015.