ABSTRACT. Much of the recent work on random constraint satisfaction problems 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 are intended to mimic Belief Propagation, a message passing algorithm, applied to the random CSP. In this paper we propose a novel method for harnessing Belief Propagation directly to obtain a rigorous proof of such a prediction, namely the existence and location of a condensation phase transition in the random regular $k$-SAT model.

Mathematics Subject Classification: 60C05, 82B44

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 physics [22, 23]. This work has had a substantial impact on the theory of computing (e.g., proofs that certain benchmark instances are difficult for certain algorithms), coding theory (“low density parity check codes”) and combinatorics (random graphs, hypergraphs and formulas); e.g., [9, 13, 16, 17, 18, 19, 28]. All of these disciplines deal with a common setup. There are a large number of variables that interact through a similarly large number of constraints. Each variable ranges over a finite domain (such as the Boolean values ‘true’ and ‘false’) and every constraint binds a small number of variables, either encouraging or discouraging certain value combinations.

The striking feature of the physics work is that it is based on a non-rigorous but generic approach called the cavity method, centered around the Belief Propagation message-passing algorithm, that can be applied almost mechanically [21]. Hence the impact of a single technique on such a wide range of problems. By comparison, the rigorous study of random constraint satisfaction problems has largely been case-by-case. This 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 rigorous proof in a highly non-trivial and somewhat representative case. Specifically, we determine the “condensation phase transition” in the random regular $k$-SAT model. The proof is based on a novel approach that demonstrates how our recent general results on the structure of Gibbs measures from [5] can be put to work. The centrepiece of the proof is a direct analysis of Belief Propagation on the random $k$-SAT instance. Thus, we prove a structural result in an indirect and perhaps surprising way by analysing an algorithm. The arguments are rather generic and we expect them to extend to other problems.

The random regular $k$-SAT model is defined as follows [26]. There are Boolean variables $x_1, \ldots, x_n$ and $m$ constraints, namely propositional clauses of length $k$. Each variable occurs precisely $d/2$ times as a positive and precisely $d/2$ times as a negative literal. Hence, $m = dn/(2k)$; we assume tacitly that $d$ is even and that $k$ divides $dn$. Let $\Phi = \Phi_{d,k}(n)$ signify a uniformly random such $k$-SAT formula

For $k$
Theorem 1.1. There exists a threshold where \( \Phi \) ceases to be satisfiable is known \(^9\). While the exact formula is cumbersome, asymptotically \( d_{k,SAT}/k = 2^k \ln 2 - k \ln 2/2 + O(1) \) for large \( k \).

Of course, to tackle questions such as the performance of algorithms knowing the satisfiability threshold is not enough. Much more precise information is encoded in the function \( \sigma \rightarrow E_\Phi(\sigma) \) that maps each truth assignment \( \sigma \) to the number of clauses that it violates. Thus, we take a weighted sum over all \( 2^n \) assignment where each violated clause induces a “penalty factor” of \( \exp(-\beta) \). In particular, the larger the inverse temperature \( \beta \), the smaller the relative contribution of “good” assignments that violate few clauses. In effect, by tuning \( \beta \) we can zoom in on a specific cross-section of \( E_\Phi \), i.e., on the set of assignments violating a given number of clauses. Further, the maximum number of clauses that can be satisfied simultaneously equals \( m + \lim_{\beta \to \infty} \frac{2}{\beta} \ln Z_\Phi(\beta) \). Of course, we are interested in the asymptotics as \( n \to \infty \). Since \( Z_\Phi(\beta) \) scales exponentially with \( n \), we consider

\[
\phi_{d,k}: \beta \in (0, \infty) \to \lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \ln Z_\Phi(\beta) \right].
\]

The existence of the limit follows from the interpolation method \(^8\) and standard arguments show that \( \ln Z_{\Phi}(\beta) \) concentrates about \( \mathbb{E}[\ln Z_{\Phi}(\beta)] \). What makes \( \phi_{d,k} \) vicious is that the log is inside the expectation.

Since the partition function characterises the “landscape” \( E_\Phi \) in a specific cross-section of \( E_\Phi \), i.e., on the set of assignments violating a given number of clauses. Further, the maximum number of clauses that can be satisfied simultaneously equals \( m + \lim_{\beta \to \infty} \frac{2}{\beta} \ln Z_\Phi(\beta) \). Of course, we are interested in the asymptotics as \( n \to \infty \). Since \( Z_\Phi(\beta) \) scales exponentially with \( n \), we consider

\[
\phi_{d,k}: \beta \in (0, \infty) \to \lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \ln Z_\Phi(\beta) \right].
\]

1.2. Results. According to (non-rigorous) physics predictions \(^20\) for certain values of \( d \) close to the satisfiability threshold \( d_{k,SAT} \) there occurs a phase transition at a certain critical \( \beta_{cond}(d, k) > 0 \), the so-called condensation phase transition. The main result of this paper proves this conjecture. Let us postpone the precise definition of \( \beta_{cond}(d, k) \) for a moment.

**Theorem 1.1.** There exists \( k_0 \geq 3 \) such that for all \( k \geq k_0 \), \( d \leq d_{k,SAT} \) there is \( \beta_{cond}(d, k) \in (0, \infty) \) such that any \( \beta \in (0, \beta_{cond}(d, k)) \) is smooth. If \( \beta_{cond}(d, k) < \infty \), then there occurs a phase transition at \( \beta_{cond}(d, k) \).

Thus, if we fix \( d, k \) such that \( \beta_{cond}(d, k) = \infty \), then the function \( \phi_{d,k} \) is analytic on \( (0, \infty) \). But if \( d, k \) are such that \( \beta_{cond}(d, k) < \infty \), then \( \phi_{d,k} \) is non-analytic at the point \( \beta_{cond}(d, k) \). In fact, we will see that \( \beta_{cond}(d, k) < \infty \) for \( d \) exceeding a specific \( d_{cond}(k) < d_{k,SAT} \). Crucially, Theorem 1.1 identifies the precise condensation threshold \( \beta_{cond}(d, k) \); it is the first such result in any model of this kind.

Let us take a look at the precise value of \( \beta_{cond}(d, k) \). As most predictions based on the cavity method, \( \beta_{cond}(d, k) \) results from a distributional fixed point problem, i.e., a fixed point problem on the space of probability measures on the unit interval \( (0, 1) \). The fixed point problem derives mechanically from the “1RSB cavity equations” \(^{21}\). Specifically, writing \( \mathcal{P} (\Omega) \) for the set of probability measures on \( \Omega \), we define two maps

\[
\mathcal{F}_{k,d,\beta}: \mathcal{P}(0, 1) \to \mathcal{P}(0, 1), \quad \mathcal{H}_{k,d,\beta}: \mathcal{P}(0, 1) \to \mathcal{P}(0, 1)
\]

\(^2\)In the sense that \( \liminf_{n \to \infty} \mathbb{P} \{ \Phi \text{ is satisfiable} \} > 0 \) if \( d < d_{k,SAT} \) and \( \lim_{n \to \infty} \mathbb{P} \{ \Phi \text{ is satisfiable} \} = 0 \) if \( d > d_{k,SAT} \).
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 \((z(\eta)/\hat{Z}(\pi)) d \otimes_{j=1}^{k-1} \pi(\eta_j)\), where

\[
\dot{z}(\eta) = 2 - (1 - \exp(-\beta)) \prod_{j < k} \eta_j \quad \text{and} \quad \dot{Z}(\pi) = \int \dot{z}(\eta) d \otimes_{j < k} \pi(\eta_j).
\]

Then \( F_{k,d,\beta}(\pi) \) is the distribution of \((1 - (1 - \exp(-\beta)) \prod_{j=1}^{k-1} \eta_j)/\hat{Z}(\eta)\). Similarly, given \( \hat{\eta} = (\hat{\eta}_1, \ldots, \hat{\eta}_{k-1}) \) from \((z(\eta)/\hat{Z}(\pi)) d \otimes_{j=1}^{k-1} \hat{\pi}(\hat{\eta}_j)\), where

\[
z(\hat{\eta}) = \prod_{j < d/2} \hat{\eta}_j \prod_{j \geq d/2} (1 - \hat{\eta}_j) + \prod_{j < d/2} (1 - \hat{\eta}_j) \prod_{j \geq d/2} \hat{\eta}_j,
\]

Then \( \hat{F}_{k,d,\beta}(\pi) \) is the distribution of \((\prod_{j < d/2} \hat{\eta}_j \prod_{j \geq d/2} (1 - \hat{\eta}_j))/z(\hat{\eta})\). Call a distribution \( \pi \in \mathcal{P}(0, 1) \) skewed if the probability mass of the interval \((0, 1 - \exp(-4\beta))\) satisfies \( \pi(0, 1 - \exp(-4\beta)) < 2^{-0.9k} \).

**Proposition 1.2.** Let \( d_-(k) = d_{k-\text{SAT}} - k^5 \) and \( \beta_-(d, k) = k \ln 2 - 10 \ln k \). The map \( \Phi_{k,d,\beta} = F_{k,d,\beta} \circ \hat{F}_{k,d,\beta} \) has a unique skewed fixed point \( \pi_{k,d,\beta}^* \), provided that \( k \geq k_0, d \in [d_-(k), d_{k-\text{SAT}}] \) and \( \beta > \beta_-(d, k) \).

To extract \( \pi_{k,d,\beta} \text{cond} \) and \( \hat{\pi}_1 \text{cond} \) be independent random variables such that the \( v_j \) have distribution \( \pi_{k,d,\beta} \text{cond} \) and the \( \hat{v}_j \) have distribution \( \hat{F}_{k,d,\beta}(\pi_{k,d,\beta} \text{cond}) \).

Setting

\[
z_1 = \prod_{j < d/2} \hat{v}_j \prod_{j \geq d/2} (1 - \hat{v}_j) + \prod_{j < d/2} (1 - \hat{v}_j) \prod_{j \geq d/2} \hat{v}_j,
\]

and \( z_3 = v_1 + (1 - v_1)(1 - \hat{v}_1) \), we let

\[
F(k, d, \beta) = \mathbb{E}[z_1] + \frac{d}{k} \mathbb{E}[z_2] - \mathbb{E}[z_3], \quad \Phi(k, d, \beta) = \frac{\mathbb{E}[z_1 \ln z_1]}{\mathbb{E}[z_1]} + \frac{d}{k} \frac{\mathbb{E}[z_2 \ln z_2]}{\mathbb{E}[z_2]} - \frac{1}{\mathbb{E}[z_3]} \ln \mathbb{E}[z_3]. \tag{1.2}
\]

Finally, we let \( \beta_{\text{cond}}(k, d) = \infty \) if \( d < d_-(k) \) and \( \beta_{\text{cond}}(k, d) = \inf \beta > \beta_-(d, k) : F(k, d, \beta) < \Phi(k, d, \beta) \) if \( d \in [d_-(k), d_{k-\text{SAT}}] \) (with the usual convention that \( \inf \phi = \infty \)).

We proceed to highlight a few consequences of Theorem 1.1 and its proof. The following result shows that \( \beta_{\text{cond}}(d, k) < \infty \), i.e., that a condensation phase transition occurs, for degrees \( d \) strictly below the satisfiability threshold.

**Corollary 1.3.** If \( k \geq k_0 \), then \( d_{\text{cond}}(k) = \min \{ d > 0 : \beta_{\text{cond}}(d, k) < \infty \} < d_{k-\text{SAT}} - \Omega(k) \).

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 \( \beta < \beta_{\text{cond}}(d, k) \).

**Corollary 1.4.** If \( k \geq k_0, d \leq d_{k-\text{SAT}} \) and \( \beta < \beta_{\text{cond}}(d, k) \), then \( \phi_{d,k}(\beta) = \lim_{n \to \infty} \frac{1}{n} \ln \mathbb{E}[Z_\beta(\Phi)] = F(k, d, \beta) \).

Corollary 1.4 opens the door to studying the “landscape” \( E_\Phi \) for \( \beta < \beta_{\text{cond}}(d, k) \). Specifically, Corollary 1.4 extends the “planting trick” developed in [1] for the set of satisfying assignments to assignments at more general “height levels” of \( E_\Phi \). We leave a detailed discussion to future work. Finally, complementing Corollary 1.4, the following result shows that \( F(k, d, \beta) \) overshoots \( \phi_{d,k}(\beta) \) for \( \beta > \beta_{\text{cond}}(d, k) \).

**Corollary 1.5.** If \( k \geq k_0, d \leq d_{k-\text{SAT}} \) and \( \beta > \beta_{\text{cond}}(d, k) \), then there is \( \beta' < \beta \) such that \( \phi_{d,k}(\beta') < F(k, d, \beta') \).

1.3. **Outline and related work.** Admittedly, the definition of \( \beta_{\text{cond}}(d, k) \) is not exactly simple. For instance, even though the fixed point distribution from Proposition 1.2 stems from a discrete problem, it turns out to be a continuous distribution on \((0, 1)\). On the other hand, the fixed point problem just mirrors the combinatorial intricacy of the phase transition. On the other hand and perhaps despite appearances, the analytic formula [1.2] is conceptually far simpler than the definition of \( \phi_{d,k} \). For instance, we are going to see in Section 2 that the fixed point problem can be understood elegantly in terms of a Galton-Watson tree. Thus, one could say that Theorem 1.1 reduces the condensation problem on the complex random formula \( \Phi \) to a problem on a random tree.
The proof of Theorem 1.1 builds upon an abstract result from [5] that, roughly speaking, reduces the study of the partition function to analysing the Belief Propagation algorithm. The technical contribution of the present work is to actually perform that analysis. In a nutshell, Belief Propagation runs a fixed point iteration on “messages” (which are numbers in (0, 1)) that are sent back and forth between variables and clauses. The condensation phase transition hinges on the fixed point to which this algorithm converges on a random formula \( \Phi \) chosen from a reweighted distribution, the “planted model”. Crucially, we are going to reduce the study of Belief Propagation on \( \Phi \) to message passing on the Galton-Watson tree that corresponds to the operator \( G_{k,d,\beta} \) from Proposition 1.2.

The predictions of the “cavity method” typically come as distributional fixed points but there are few proofs that establish such predictions rigorously. The one most closely related to the present work is the paper of Bapst et al. [6] 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 Erdos-Rényi random graph. Conceptually, this corresponds to taking the limit \( \beta \to \infty \) in (1.1), which simplifies the problem rather substantially (see Section 2.5 below). Thus, the main result of [6] corresponds to Corollary 1.3. Other previous results on condensation, which dealt with random hypergraph 2-coloring and the Potts model on the random graph, were only approximate [7, 11, 12].

Interestingly, determining the satisfiability threshold on the random regular formula \( \Phi \) is conceptually much easier than identifying the condensation threshold [9]. This is because the local structure of the random formula \( \Phi \) 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 [14, 15], and these proofs employed Belief Propagation in this simpler setting. By contrast, we will see in Section 2 that the condensation phase transition hinges on the reweighted distribution \( \Phi \), whose local structure is random (in an “asymmetric” way).

Recent work on the \( k \)-SAT threshold in uniformly random formulas [9, 10], in particular the breakthrough paper by Ding, Sly and Sun [16], also harnessed the physicists’ Belief Propagation or Survey Propagation calculations in the uniformly random model a substantial technical difficulty is posed by the presence of variables of exceptionally high degree, an issue that is, of course, absent in the regular model. Specifically, [9, 10, 16] apply the second moment method to a random variable whose construction is guided by Belief/Survey Propagation. By contrast, here we employ Belief Propagation in the direct “algorithmic” way enabled by [5]. Let us take a closer look.

2. THE PROOF STRATEGY

2.1. Two moments do not suffice. The default approach to studying the function \( \phi_{d,k}(\beta) \) is the venerable “second moment method”. Cast on a logarithmic scale, if

\[
\limsup_{n \to \infty} \frac{1}{n} \ln \mathbb{E}[Z_\Phi(\beta)^2] \leq \lim_{n \to \infty} \frac{2}{n} \ln \mathbb{E}[Z_\Phi(\beta)],
\]

then

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

The last term is easy to study because the log is outside the expectation. In particular, the function \( \beta \in (0,\infty) \to \lim_{n \to \infty} \frac{1}{n} \ln \mathbb{E}[Z_\Phi(\beta)] \) turns out to be analytic. Consequently, the least \( \beta \in (0,\infty) \) where (2.2) fails to hold must be a phase transition.

From a bird’s eye view, both the physics intuition and the second moment are all about the geometry of the Gibbs measure of \( \Phi \) at a given \( \beta \in (0,\infty) \). Let us encode truth assignments as points \( \sigma \in \{-1,1\}^n \) with the convention that 1 stands for ‘true’ and -1 for ‘false’. Then the Gibbs measure is the probability

\[3\text{Survey Propagation can be viewed as a Belief Propagation applied to a modified constraint satisfaction problem [21].} \]
distribution on $\{\pm 1\}^n$ defined by

$$
\sigma \in \{\pm 1\}^n \mapsto \exp(-\beta E_\Phi(\sigma))/Z_\Phi(\beta).
$$

Thus, we weigh assignments according to the number of clauses that they violate, giving greater weight to 'better' assignments as $\beta$ gets larger. Let $\sigma, \sigma_1, \sigma_2, \ldots$ be independent samples from the Gibbs measure and write $\langle X(\sigma_1, \ldots, \sigma_l) \rangle_{\Phi, \beta}$ for the expectation of $X : (\{\pm 1\}^l \mapsto \mathbb{R}$. Then according to the physics picture the condensation point $\beta_{\text{cond}}(k)$ should be the supremum of all $\beta > 0$ such that $\mathbb{E}\langle \sigma_1 \cdot \sigma_2 \rangle_{\Phi, \beta} = o(n)$. In other words, if we choose a random formula $\Phi$ and then sample two assignments $\sigma_1, \sigma_2$ according to the Gibbs measure independently, then $\sigma_1, \sigma_2$ will be about orthogonal.

This decorrelation property is, roughly speaking, a necessary condition for the success of the second moment method \cite{MPW}. Therefore, the prediction that $\mathbb{E}\langle \sigma_1 \cdot \sigma_2 \rangle_{\Phi, \beta} = o(n)$ right up to $\beta_{\text{cond}}(d, k)$ may inspire confidence that the same is true of (2.1). In fact, (2.1) holds if either $d$ or $\beta$ is relatively small.

**Lemma 2.1.** If $d \leq d_{\text{crit}}(k)$ or $\beta \leq \beta_{\text{crit}}(k)$ then (2.1) is true.

However, for $\beta$ near $\beta_{\text{cond}}(d, k)$ the second moment method turns out to fail rather spectacularly. Formally, if $\beta_{\text{cond}}(d, k) < \infty$, then there exists $\epsilon > 0$ such that (2.1) is violated for all $\beta > \beta_{\text{cond}}(d, k) - \epsilon$, i.e., the second moment overshoots the square of the first moment by an exponential factor.

2.2. **Quenching the average.** To understand what goes away it is convenient to turn the second moment into a first moment under a reweighted distribution that we call the planted model. This is the distribution on formula/assignment pairs under which the probability of $(\hat{\Phi}, \hat{\sigma})$ equals $\exp(-\beta E_{\Phi}(\hat{\sigma})) \mathbb{E}[Z_{\Phi}(\beta)]$. Let $(\hat{\Phi}, \hat{\sigma})$ be a random pair drawn from this distribution. Then by symmetry the distribution of the assignment $\hat{\sigma}$ is uniform and we may assume without loss that $\hat{\sigma} = 1$ is the all-ones assignment. Further, the probability that a specific formula $\Phi$ comes up equals $\mathbb{P}[\hat{\Phi} = \Phi] = Z_\beta(\Phi)/\mathbb{E}[Z_\beta(\Phi)]$. Thus, the planted distribution weights formulas by their partition function. In effect,

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

If we trace the (elementary) calculation behind Lemma 2.1 we see that $\mathbb{E}[Z_\beta(\beta)]$ is dominated by two distinct contributions. First, assignments that are more or less orthogonal to $\hat{\sigma}$ yield a term of order $\mathbb{E}[Z_\beta(\beta)]$. Second, there is a contribution from $\sigma$ close to $\hat{\sigma} = 1$; say, $\sigma \cdot 1 \geq n(1 - 2^{-k/10})$. Geometrically, this reflects the fact that the planted assignment $1$ sits in a “valley” of the Hamiltonian $E_\Phi$ w.h.p. The valleys are officially known as clusters and we let

$$
\langle \mathcal{C}_{\Phi, \hat{\sigma}}(\beta) = Z_{\Phi}(\beta) 1[\sigma \cdot 1 > n(1 - 2^{-k/10})] \rangle_{\Phi, \beta}
$$

be the (weighted) cluster size. Performing an elementary calculation, we find that it is the expected cluster size that derails the second moment method for $\beta$ near $\beta_{\text{cond}}(d, k)$.

At a second glance, this is unsurprising. For $\mathcal{C}_{\Phi, \hat{\sigma}}(\beta)$ scales exponentially with $n$ and is therefore prone to large deviations. That is, even though typically $\hat{\sigma}$ sits in a modest crater of $E_\Phi$, there is a small chance that it lies in a giant canyon, a rare event that drives up the expectation. To suppress such “lottery effects” we ought to scale the cluster size appropriately. That is, we should investigate $\mathbb{E}[\ln \mathcal{C}_{\Phi, \hat{\sigma}}(\beta)]$ instead of $\mathbb{E}[\mathcal{C}_{\Phi, \hat{\sigma}}(\beta)]$. A similar issue (that the expected cluster size drives up the second moment) occurred in earlier work on condensation \cite{PPV, MV, MVW, Vir}. Borrowing the remedy suggested in these papers, we observe that applying the second moment method to a carefully truncated random variable yields

**Lemma 2.2.** Equation (2.2) holds iff $\limsup_{n \to \infty} n^{-1} \mathbb{E}[\ln \mathcal{C}_{\Phi, \hat{\sigma}}(\beta)] = \lim_{n \to \infty} n^{-1} \ln \mathbb{E}[Z_{\Phi}(\beta)]$.

Hence, we are left to calculate $\mathbb{E}[\ln \mathcal{C}_{\Phi, \hat{\sigma}}(\beta)]$, the “quenched average” in physics jargon. As the log and the expectation do not commute, this problem is well beyond the reach of elementary methods. Tackling it is the main achievement of this paper and our principal tool will be the Belief Propagation algorithm.
2.3. Belief Propagation. Guided by the hypothesis that \( \ln \mathcal{E}_{\Phi, \sigma}(\beta) \) is governed by the local structure of the formula rather than by "long-range effects", the physics recipe is to run the Belief Propagation message passing algorithm on the random formula \( \Phi \). To be precise, we associate a bipartite graph with the formula, the factor graph. Its vertices are the variables \( x_1, \ldots, x_n \) and the clauses \( a_1, \ldots, a_m \). There is an edge between \( x_i \) and \( a_j \) if variable \( x_i \) occurs in clause \( a_j \). We write \( \partial \nu = \partial \Phi \) \( \nu \) for the neighborhood of a vertex \( \nu \). Further, for a variable \( x \) and an adjacent clause \( a \) we let sign\( \{x, a\} = \operatorname{sign}_{\Phi}\{x, a\} = \pm 1 \) if \( x \) appears in \( a \) positively/negatively. In addition, for a (clause or variable) vertex \( \nu \) we let \( \partial_{\pm} \nu \) be the set of all neighbors \( u \in \partial \nu \) such that sign\( \{u, \nu\} = \pm 1 \). Finally, for a clause \( a \) and a \( \pm 1 \)-vector \( s = (s_y)_{y \in \partial a} \) we write \( s \models a \) if \( s_y = \operatorname{sign}(a, y) \) for some \( y \in \partial a \); that is, the truth value combination \( s \) satisfies clause \( a \).

The assumption underpinning Belief Propagation is that the cluster size can be calculated by determining an optimal "local probability distribution" for every vertex. Specifically, a marginal sequence is a family \( \mu = (\mu_{x_i}, \mu_{a_j})_{i \in [n], j \in [m]} \) where each \( \mu_{x_i}(\pm 1) \) is a distribution on \{\pm 1\} and each \( \mu_{a_j} \) is a distribution on \{\pm 1\} \( k \) such that

\[
\sum_{(s_y)_{y \in \partial a} \in \{\pm 1\}^{\partial a}} \mu_{a}(s_y)_{y \in \partial a} = \mu_x(s_x) \quad \text{for any clause } a, x \in \partial a, s_x \in \{\pm 1\}.
\]

In words, the \( x \)-marginal of \( \mu_a \) coincides with \( \mu_{x_i} \), a local consistency condition \( [21] \).

Of course, the marginals of the variables/clauses under the Gibbs measure \( \Phi \) given that \( \sigma \cdot 1 \geq n(1 - 2^{-k/10}) \) satisfy condition \( [23] \). Conversely, under the assumption that these marginals contain all the necessary information, one can derive a natural "guess" as to the cluster size. Indeed, given a marginal sequence \( \mu \), we define the Bethe free energy \( [21] \) as

\[
B_{\Phi, \beta}(\mu) = \frac{d - 1}{n} \sum_{i=1}^{n} \sum_{s = \pm 1} \mu_{x_i}(s) \ln \mu_{x_i}(s) - \frac{1}{n} \sum_{j=1}^{m} \sum_{s \in \{\pm 1\}^k} \mu_{a_j}(s) \langle \beta 1_{s \neq a_j} + \ln \mu_{a_j}(s) \rangle.
\]

Now, the hypothesis is that the cluster size can be determined by maximising the Bethe free energy.

**Hypothesis 2.3.** We have \( \frac{1}{n} \mathbb{E}[\ln \mathcal{E}_{\Phi, \sigma}(\beta)] = \max \{B_{\Phi, \beta}(\mu) : \sum_{i=1}^{n} \mu_{x_i}(1) \geq n(1 - 2^{-k/10})\} \).

Belief Propagation is an algorithm designed to solve this (non-convex) optimisation problem. The idea is to perform a fixed point iteration on "messages" whizzing along the edges of the factor graph. Each message is a number between 0 and 1. Formally, the message space \( \mathcal{M}_{\Phi} \) is the set of all families

\[
\eta = (\eta_x \models b, \eta_a \models y)_{i \in [n], j \in [m], b \in \partial x_i, y \in \partial a_j}
\]

such that \( \eta_x \models b, \eta_a \models y \) are distributions on \{\pm 1\}. Moreover, Belief Propagation is the map

\[
\text{BP}_{\Phi, \beta} : \mathcal{M}_{\Phi} \rightarrow \mathcal{M}_{\Phi}, \quad \eta \mapsto \eta' = \text{BP}_{\Phi, \beta}(\eta),
\]

where for \( s, s_x \in \{\pm 1\} \) and \( a \in \partial x \) we let

\[
\eta_{a \models x}^i(s) \propto \sum_{(s_y)_{y \in \partial a \setminus \{x\}}} \exp (-1 \langle s_y, \mu_{a \models y} \rangle) \prod_{y \in \partial a \setminus \{x\}} \eta_y \models a(s_y), \quad \eta_{x \models a}^i(s) \propto \prod_{b \in \partial x \setminus \{a\}} \eta_b \models x(s).
\]

Here '\( \propto \)' denotes the normalisation required to turn \( \eta_{x \models a}^i, \eta_{b \models x}^i \) into probability distributions. The messages give rise to "local probability distributions" associated with the variables and clauses:

\[
\eta_{x_i}(\pm 1) \propto \prod_{a \in \partial x_i} \eta_a \models x_i(\pm 1), \quad \eta_{a_j}(s) \propto \exp (-1 \langle s, \mu_{a_j} \rangle) \prod_{y \in \partial a_j} \eta_y \models a_j(s) \quad (s \in \{\pm 1\}^{\partial a_j}).
\]

Plugging the "Belief Propagation marginals" from \( [25] \) into \( [24] \), we obtain the Bethe free energy \( B_{\Phi, \beta}(\eta) \) of \( \eta \in \mathcal{M}_{\Phi} \). Further, the marginal sequences that are stationary points of the Bethe free energy correspond to the fixed points \( \eta \in \mathcal{M}_{\Phi} \) of Belief Propagation \( [29] \). Hence, we can hope to maximise the Bethe

\[\text{ Thus, if we write } p(\omega) \propto f(\omega) \text{ for } \omega \in \Omega \text{ we mean } p(\omega_0) = f(\omega_0) / \sum_{\omega \in \Omega} f(\omega) \text{ for every } \omega_0 \in \Omega.\]
free energy and thus calculate $\mathbb{E}[\ln \mathcal{E}_{\Phi, \beta}(\beta)]$ by studying these fixed points. The following proposition, whose proof is the heart of the paper, shows that this does indeed work. Call $\eta \in \mathcal{M}_\Phi$ skewed if
\[
\left| \left\{ i \in [n] : \eta_{x_i}(1) \leq 1 - \exp(-4\beta) \right\} \right| \leq n2^{-k/20}.
\] (2.6)

**Proposition 2.4.** If $d > d_-(k)$ and $\beta > \beta_-(d, k)$, then w.h.p.
\[
\frac{1}{n} \ln \mathcal{E}_{\Phi, \beta}(\beta) \sim \max \left\{ B_{\Phi, \beta}(\eta) : \eta \in \mathcal{M}_\Phi, \eta \text{ is a skewed Belief Propagation fixed point} \right\} \sim \mathcal{B}(d, k, \beta).
\]

Theorem 1.1 follows by combining Lemma 2.1, Lemma 2.2 and Proposition 2.4.

### 2.4. The fixed point problem.

Hence, we are left to prove Proposition 2.3. The proposition establishes a connection between three quantities: the cluster size, the dominant Belief Propagation fixed point on $\hat{\Phi}$ and the fixed point from Proposition 1.2. The link between the three is a random tree that mimics the local structure of the formula $\hat{\Phi}$. Indeed, although the factor graph is regular (each variable has degree $d$ and every clause has degree $k$), the signs with which the variables appear in the clauses are random in a rather non-trivial fashion.

Even though it is hardly apparent from the definition of $\hat{\Phi}$, its local structure allows for an elegant explicit description in terms of a Galton-Watson tree. To define this tree, we observe that there is a unique $q = q(k, d, \beta) \in (0, 1)$ such that
\[
1 - (1 - \exp(-\beta))q^k = 2(1 - q).
\]
The Galton-Watson process has four types: variable nodes of type $\pm 1$ and clause nodes of type $\pm 1$. The root is a variable node of type 1 with probability $1 - q$ and of type $-1$ otherwise. Further, the offspring of a variable node of type $\pm 1$ is $\epsilon - 1$ clause nodes of type $\pm 1$ and $\epsilon$ clause nodes of type $\mp 1$. In addition, a clause node spawns $k - 1$ variable nodes in total. If the clause node has type 1, the number of $1$-children has distribution Bin$(k - 1, 1 - q)$. Further, the offspring of a clause node of type $-1$ is as follows.

- with probability $\exp(-\beta)q^{k-1}/(1 - (1 - \exp(-\beta))q^{k-1})$ the offspring is $k - 1$ variables of type $-1$.
- otherwise the number of $1$-children has a conditional binomial distribution Bin$_{\leq 1}(k - 1, 1 - q)$.

Let $T = T(d, k, \beta)$ be the resulting random (infinite) tree. We can view $T$ as the factor graph of an infinite $k$-SAT formula in which the signs with which literals appear in clauses are given by the node types. More precisely, if $a$ is a clause node of type $\pm 1$, then its parent node $x$ is a positive/negative literal. Similarly, if the type of the variable node $x$ indicates whether it occurs positively or negatively in its parent clause.

By design the tree $T$ captures the local structure of $\hat{\Phi}$. More precisely, for a variable $x_i$, a clause $a \in \partial_{\Phi,x_i}$ and an integer $\omega \geq 0$ let $\partial^{\omega}_{\Phi}(x_i \rightarrow a)$ be the sub-formula of $\hat{\Phi}$ consisting of all variables and clauses that are reachable from $x_i$ by a path of length at most $\omega$ in the factor graph that does not pass through $a$. Further, for a possible outcome $T$ of the branching process obtain $\partial^{\omega}_{\Phi}T$ by deleting all vertices at distance greater than $\omega$ from the root of $T$. Let us consider the random variable
\[
\mathcal{N}(T, \omega) = \sum_{i=1}^{n} \sum_{a \in \partial_{\Phi,x_i}} 1[\partial^{\omega}_{\Phi}T \supseteq \partial^{\omega}_{\Phi}(x_i \rightarrow a)].
\] (2.7)

**Lemma 2.5.** For any $T$ and any $\omega \geq 0$ we have $\frac{\mathcal{N}(T, \omega)}{km} \sim \mathbb{P}[\partial^{\omega}_{\Phi}T \supseteq \partial^{\omega}_{\Phi}T]$ in probability as $n \to \infty$.

In addition, $T$ provides a combinatorial interpretation of the fixed point distribution from Proposition 1.2; it describes nothing but Belief Propagation on the $k$-SAT formula induced by $T$. More precisely, pick an arbitrary distribution $p$ on $(0, 1)$ such that $p(1 - \exp(-2\beta), 1) \geq 1 - 2^{-0.9k}$. Now, independently for every variable $x$ of $T$ and every $a \in \partial_{\Phi,x}$ choose $\eta^0_{T,x,a}$ from $p$ and set $\eta^0_{T,x,a}(-1) = 1 - \eta^0_{T,x,a}(1)$. Moreover, let $\eta^{\omega+1}_{T} = \text{BP}^{\omega}_{T,\beta}(\eta^0_{T})$ for any $\omega \geq 0$. Finally, with $r$ denoting the root of $T$, let
\[
\eta^0_{T}(\pm 1) \propto \prod_{a \in \partial_{r}} \eta^0_{T,a \rightarrow r}(\pm 1).
\]
be the message that the root of $T$ would send to its parent if it had one.
Proposition 2.6. If \( k \geq k_0, d > d_+(k) \) and \( \beta > \beta_-(d, k) \), then \( (\eta^*_T(1))_{\omega \geq 1} \) converges almost surely to a random variable \( \eta^*_T \) with distribution \( \pi^*_T \).

The proof of Proposition 2.6 is by carefully unravelling the fixed point equations and keeping track of contraction properties. In Section 3 we will derive Proposition 2.4 from Lemma 2.5 and Proposition 2.6.

2.5 Summary and discussion. In summary, we reduced the study of \( \phi_{d,k}(\beta) \) to the analysis of the cluster size \( \mathcal{C}_{\Phi,\sigma}(\beta) \) in the planted model. The latter is given by the Bethe free energy of the dominant skewed Belief Propagation fixed point, which converges to the Belief Propagation fixed point on the random tree that encodes the fixed point problem from Proposition 1.2.

These considerations entail the typical cluster size of an assignment chosen from the Gibbs measure of the original random formula \( \Phi \) for \( \beta < \beta_{\text{cond}}(d, k) \). In fact, such a random assignment \( \sigma \) sits in a “valley” of \( E_\Phi \) just as well as \( \hat{\sigma} \) belongs to a “valley” of \( E_\Phi \).

Corollary 2.7. If \( k \geq k_0, d \leq d_{k-\text{SAT}} \) and \( \beta < \beta_{\text{cond}}(d, k) \), then \( \lim_{n \to \infty} n^{-1} \mathbb{E} \left[ \ln \mathcal{C}_{\Phi,\sigma}(\beta) \right]_{\Phi,\beta} = \mathcal{B}(k, d, \beta) \). In particular, if \( \beta_{\text{cond}}(d, k) < \infty \), then

\[
\lim_{\beta \to \beta_{\text{cond}}(d, k)} \lim_{n \to \infty} n^{-1} \mathbb{E} \left[ \ln \mathcal{C}_{\Phi,\sigma}(\beta) \right]_{\Phi,\beta} = 0. \tag{2.8}
\]

Equation (2.8) implies that as \( \beta \) approaches \( \beta_{\text{cond}}(d, k) \) the typical size of a single cluster approaches the partition function \( Z_\Phi(\beta) \). Hence the term “condensation”. Earlier work on condensation in random graph coloring [6] was based on studying the cluster size as well, but only in the limit \( \beta \to \infty \) [6]. This is a very substantial simplification due to the presence of “frozen variables” [1, 24], i.e., vertices that take the same color in all the colorings in the cluster. Hence, the cluster size stems exclusively from vertices that are unfrozen, and these (essentially) form a subcritical graph with components of size \( O(\ln n) \). In effect, the corresponding fixed point distribution is discrete. By contrast, in our case of finite \( \beta \) there are no frozen variables. Consequently, the combinatorics of the cluster size is far more intricate, which leads to a continuos fixed point distribution in Proposition 1.2. Let us take a look how we can nevertheless get a handle on the cluster size.

3. Belief Propagation on Random Formulas

This section deals with the proof of Propositions 2.4. We assume throughout that \( k \geq k_0, d_+(k) < d \leq d_{k-\text{SAT}} \) and \( \beta > \beta_-(k) \). Although technically simpler, the proof of the upper bound contains the most important ideas. We shall therefore concentrate on the proof of

Proposition 3.1. We have \( n^{-1} \ln \mathcal{C}_{\Phi,\sigma}(\beta) \leq \mathcal{B}(k, d, \beta) + o(1) \) w.h.p.

The starting point is the following lemma, which follows directly from our general result [5, Corollary 4.2].

Lemma 3.2. W.h.p. we have

\[
\frac{1}{n} \ln \mathcal{C}_{\Phi,\sigma}(\beta) \leq \max \left\{ B_{\Phi,\beta}(\mu) : \mu \text{ is a marginal sequence s.t. } \sum_{i \in [n]} \mu_{\sigma}(1) \geq (1 - 2^{-k/10}) n \right\} + o(1).
\]

Combining Lemma 3.2 with a rough a priori bound (based on a standard first moment calculation and a simple expansion argument), we obtain the following strengthened statement. Let us call a marginal sequence \( \mu \) skewed if \( \frac{1}{n} \ln \mathcal{C}_{\Phi,\sigma}(\beta) \leq \max \left\{ B_{\Phi,\beta}(\mu) : \mu \text{ is a skewed marginal sequence} \right\} + o(1).

Corollary 3.3. W.h.p. we have \( \frac{1}{n} \ln \mathcal{C}_{\Phi,\sigma}(\beta) \leq b_{\Phi,\beta} + o(1) \).

Now, the plan is as follows. Suppose that \( \mu \) is a skewed maximiser of \( B_{\Phi,\beta} \). We are going to identify a large, explicit set of variables, the core, such that \( \mu_{\sigma}(1) \geq 1 - \exp(-4\beta \) for all \( x \) in the set. This will allow us to construct a Belief Propagation fixed point \( \eta \) whose Bethe free energy matches that of \( \mu \). Further, we will see that all messages sent out by the variables \( x \) in the core are satisfy \( \eta_{x \to o}(1) \geq 1 - \exp(-2\beta) \). Finally, by
investigating how the core is connected with the vertices outside we will be able to bring Proposition 2.6 to bear to show that \( b_{\Phi, \beta} = B_{\Phi, \beta}(\eta) \sim \mathcal{B}(d, k, \beta) \). Then Proposition 3.1 follows from Corollary 3.3.

To carry out the details, we define the \( \lambda \)-core of \( \Phi \) (in symbols: \( \text{Core}_\lambda(\Phi) \)) as the largest set of variables such that all \( x \in W \) satisfy the following conditions.

- **CR1**: there are at least \( \lambda^{-1} k^{7/8} \) clauses \( a \in \partial_1 x \) such that \( \partial_1 a = \{x\} \).
- **CR2**: there are no more than \( 3 \lambda \) clauses \( a \in \partial x \) such that \( |\partial a| = k \).
- **CR3**: for any \( 1 \leq l \leq k \) the number of \( a \in \partial_{-1} x \) such that \( |\partial a| = l \) is bounded by \( \lambda k^{l+3}/l! \).
- **CR4**: there are no more than \( \lambda^{3/4} \) clauses \( a \in \partial_1 x \) such that \( |\partial a| = 1 \) but \( \partial a \notin W \).
- **CR5**: there are no more than \( \Lambda k^{3/4} \) clauses \( a \in \partial_{-1} x \) such that \( |\partial a| < k \) and \( |\partial a \setminus W| \geq |\partial a|/4 \).

The \( \lambda \)-core is well-defined; if \( W, W' \) satisfy the above conditions, then so does \( W \cup W' \). Further, if \( \lambda < \lambda' \), then \( \text{Core}_\lambda(\Phi) \subset \text{Core}_{\lambda'}(\Phi) \). Standard arguments yield

**Lemma 3.4.** W.h.p. we have \(|\text{Core}_{1/2}(\Phi)| \geq (1 - 2^{-0.95 k}) n\).

To express that the core enjoys strong expansion properties w.h.p., we say that a set of variables \( S \) is sticky if for every \( x \in S \) one of the following conditions hold.

- **ST1**: there are at least \( k^{3/4} \) clauses \( a \in \partial_1 x \) such that \( \partial_1 a = \{x\} \) and \( \partial_1 a \cap S \neq \emptyset \).
- **ST2**: there are at least \( k^{3/4} \) clauses \( a \in \partial_{-1} x \) such that \( |\partial_{-1} a| < k \) and \( |\partial a \cap S| \geq |\partial a|/4 \).

Clearly, if \( S, S' \) are sticky, then so is \( S \cup S' \).

**Lemma 3.5.** W.h.p. any sticky set \( S \subset \text{Core}_1(\Phi) \) of size \(|S| \leq n^{2^{-k/10}} \) is empty.

After these preparations we are ready to undertake the step from marginals to messages.

**Lemma 3.6.** W.h.p. \( \Phi \) has a Belief Propagation fixed point \( \eta^* \) such that \( \eta^*_{x \rightarrow a}(1) \geq 1 - \exp(-2\beta) \) for all \( x \in \text{Core}_1(\Phi) \) and \( a \in \partial x \), and \( b_{\Phi, \beta} = B_{\Phi, \beta}(\eta^*) \).

**Proof.** Let \( \Gamma \) be the set of all \( \gamma \in \mathbb{R}^n \) such that \(|\{i : \gamma_i < 1 - \exp(-4\beta)\}| < n^{2^{-k/10}}\). A priori, the problem of maximising \( B_{\Phi, \beta}(\mu) \) is \((2n + 2^k m)\)-dimensional because we have variables \( \mu_x(\pm 1) \) for each variable \( x \) and \( (\mu_a(s))_{s \in \{\pm 1\}^k} \) for each clause. We begin by reducing the dimension to \( n \). Clearly, as \( \mu_x(1) + \mu_x(-1) = 1 \) for every variable \( x \), the \( 2n \)-dimensional vector \( (\mu_{x_i}(\pm 1))_{i \in [n]} \) is completely determined by \( (\mu_{x_i}(1))_{i \in [n]} \). Further, to reduce the problem of calculating \( b_{\Phi, \beta} \) to an optimisation problem just over the vector \( (\mu_{x_i}(1))_{i \in [n]} \) we notice that the variable marginals entail the optimal clause marginals \((\mu_{a_i})_{i \in [m]} \). Indeed, the definition (2.3) of the Bethe free energy implies that for every clause \( a \) the optimal distribution \( \mu_a \) maximises

\[
- \sum_{s \in \{\pm 1\}^k} \mu_a(s) \beta 1(s \neq a) + \ln \mu_a(s)
\]

subject to the marginalisation condition (2.3). Following [29], we tackle this constrained optimisation by introducing Lagrange multipliers \( \eta_{x \rightarrow a}(\pm 1) \) for each \( x \) with \( \partial a \) that satisfy

\[
\frac{\sum_{(s_y) \in \{\pm 1\}^k} \exp(-1(\sigma_y) \in \partial a \neq a | \beta) \prod_{y \in \partial a} \eta_{y \rightarrow a}(s_y)}{\sum_{(s_y) \in \{\pm 1\}^k} \exp(-1(\sigma_y) \in \partial a \neq a | \beta) \prod_{y \in \partial a} \eta_{y \rightarrow a}(s_y)} = \mu_x(s_x)
\]

for \( s_x = \pm 1 \).

In effect, we have reduced the Bethe free energy \( B_{\Phi, \beta}(\mu) \) to a function of the variables \((\mu_{x_i}(1))_{i \in [n]} \) only. Its partial derivative works out to be

\[
\frac{\partial}{\partial \mu_x(1)} B_{\Phi, \beta} = -\frac{1}{n} \ln \left( \frac{\mu_x(1)}{\mu_x(-1)} \right) + \frac{1}{n} \sum_{a \in \partial x} \ln \left( \frac{\mu_{x_a}(1) \eta_{x \rightarrow a}(1)}{\mu_{x_a}(-1) \eta_{x \rightarrow a}(-1)} \right).
\]

Now, suppose that \( \mu \) is a skewed marginal sequence such that \( B_{\Phi, \beta}(\mu) \) is maximum. We claim that the set \( S \) of all \( x \in \text{Core}_1(\Phi) \) such that \( \mu_x(1) \leq 1 - \exp(-4\beta) \) is sticky. For otherwise there is \( x \in S \) for which neither **ST1** nor **ST2** holds and we are going to reach a contradiction by showing that (3.3) is strictly positive. To this end, we consider four different contributions to (3.3) separately, performing elementary calculations in each case.
positive, contradicting the maximality of \( \mu \). Moreover, out of these 2\( \Delta \) clauses at most \( k^{3/4} \leq \Delta \) clause contain a variable that does not belong to \( S \) (as \( x \) violates \textbf{ST1}). The other at least \( \Delta \) clauses therefore satisfy \( \vartheta_1 a = \{ x \} \) and 
\[
\mu_y(1) \geq 1 - \exp(-4\beta)
\]
for all \( y \in \vartheta_1 a \). Each such clause therefore contributes an additive \( \Omega(\beta) \) to the derivative. Hence, the total contribution of clauses \( a \) such that \( \vartheta_1 a = \{ x \} \) comes to \( \Omega(k^{7/8}) \).

(ii) Suppose \( a \in \partial x \) satisfies \( |\vartheta_1 a| < k \) and \( l = |(\vartheta_1 a \cap \text{Core}_1(\Phi)) \setminus S| \geq |\vartheta_1 a|/2 \). Then the contribution of \( a \) to \( (3.2) \) is lower-bounded by \( -\exp(-l/3) \). Therefore, \textbf{CR3} ensures that the total contribution of all such clauses \( a \) is lower-bounded by \(-1\).

(iii) Because \textbf{ST2} is violated, \textbf{CR2} and \textbf{CR5} ensure that there remain no more than \( O(k^{3/4}) \) clauses \( a \in \partial x \) that are not covered by (i) or (ii). The contribution to \( (3.2) \) of each such clause is no smaller than 
\[
\eta_{x-a}(s)/\mu_x(s) \leq \exp(\beta).
\]
Therefore, the overall contribution of all remaining clauses is \( O(k^{7/8}/\beta) \).

(iv) The contribution of the first summand in \( (3.3) \) is \( \Omega(\beta) \) as \( \mu_x(1) \leq 1 - \exp(-4\beta) \).

The bottom line is that the derivative \( (3.2) \) is \( \Omega(k^{7/8}/\beta) \). Hence, if \( S \) fails to be sticky, then \( (3.2) \) is strictly positive, contradicting the maximality of \( B_{\Phi, \beta}(\mu) \).

Thus, \( S \) is sticky and has size \( |S| < n 2^{-k/20} \) because \( \mu \) is skewed. Consequently, by Lemma \textbf{3.5} we may assume that \( S = \emptyset \). In addition, by Lemma \textbf{3.4} we may assume that \( |\text{Core}_1(\Phi)| \geq (1 - 2^{-0.95k}) n \). Combining these two facts and observing that \( \mu_x(1) < 1 \) due to the first summand in \( (3.3) \), we see that \( \mu \) is an interior point of \( \Gamma \). Therefore, \( \mu \) is a stationary point of the Bethe free energy. Finally, to construct the desired Belief Propagation fixed point, we turn the Lagrange multipliers from \( (3.2) \) into messages by letting

\[
\eta_{x-a}^*(\pm 1) \propto \eta_{x-a}(\pm 1) \propto \sum_{(s_z)_{z \in \partial a\setminus\{x\}}} \exp(-1|\{s_z\}_{z \in \partial a \setminus \{x\}} a \neq a| \beta) \prod_{z \in \partial a \setminus \{x\}} \eta_{z-a}(s_z).
\]

As all the derivatives \( (3.2) \) vanish (\( \mu \) is stationary), \( \eta^* \) is a Belief Propagation fixed point. Further, \( \eta_{x-a}^*(\pm 1) = \mu_x(\pm 1) \) for all \( x \), whence \( B_{\Phi, \beta}(\mu) = B_{\Phi, \beta}(\eta) \) and \( \eta_{x-a}^*(1) \geq 1 - \exp(-2\beta) \) for \( x \in \text{Core}_1(\Phi) \), \( a \in \partial x \). \hfill \Box

The next lemma establishes an explicit connection between the messages from Lemma \textbf{3.6} and the interpretation of the distributional fixed point from Proposition \textbf{2.3}. Recall \( \mathcal{N}(T, \omega) \) from \textbf{2.7}.

**Lemma 3.7.** For any \( \varepsilon > 0 \) there is \( \omega > 0 \) such that the following holds. Choose \( T, \Phi \) independently and let

\[
\mathcal{N}_\varepsilon(T, \omega) = \left\{ (x_i, a) : i \in [n], a \in \partial x_i, \vartheta^\omega \Phi(x_i \rightarrow a), |\eta_{x-a}^*(1) - \eta_{x-a}^{\omega}(1)| > \varepsilon \right\}.
\]

Then \( \lim_{\omega \to \infty} \lim_{n \to \infty} \mathbb{E}[\mathcal{N}_\varepsilon(T, \omega)/\mathcal{N}(T, \omega)] = 0 \).

**Proof.** The proof is based on a switching argument. We can view the random formula \( \hat{\Phi} \) as a random bijection \( \hat{\Phi} : \{a_1, \ldots, a_m\} \times [k] \to \{x_1, \ldots, x_n\} \times [d] \), a construction known as the “configuration model”. The idea is that we generate \( \Phi \) by setting up a deck of cards that contains \( d \) numbered copies of each variable. The first \( d/2 \) copies of \( x \) are going to be its positive occurrences and the last \( d/2 \) the negative occurrences. We shuffle the \( dn \) cards and put them down one by one to fill in the clauses’ \( km \) slots.

Let us pick one variable \( x \) and one clause \( a \in \partial x \) at random. Further, let \( \omega \) be a large but fixed odd integer and let \( b_1, \ldots, b_N \) be the clause vertices at distance precisely \( \omega \) from \( x_1 \) in the formula \( \Phi \) with \( a \) removed. Moreover, let \( x_i \) be the variable node on the shortest path from \( b_i \) to \( x \), i.e., the parent of \( b_i \) in the tree \( T = \vartheta^\omega \Phi(x \rightarrow a) \). We may assume without loss that \( x_i \) occurs in the \( k \)th position of each clause \( b_i \). In addition, let \( y_1, \ldots, y_{(k-1)N} \) be the “variable cards” placed in the remaining \( (k-1)N \) positions of the clauses \( b_1, \ldots, b_N \). Now, obtain another random formula \( \Phi \) from \( \Phi \) as follows.

- Choose a random \((k-1)N\)-tuple \((z_1, \ldots, z_{(k-1)N})\) of distinct “variable cards” \( z_i \in \{x_1, \ldots, x_n\} \times [d] \).
- Swap the cards \( y_i \) and \( z_i \) for \( i = 1, \ldots, (k-1)N \).

In other words, we randomly re-connect the boundary of the tree \( T \). Clearly, the outcome \( \Phi \) is simply a uniformly random regular \( k \)-SAT instance.
Let $E_i$ be the event that $z_i$ corresponds to a variable in the 1-core of $\Phi$. Because $z_1, \ldots, z_{(k-1)N}$ are chosen randomly, Lemma 3.4 shows that $\mathbb{P}[E_i] \geq 1 - 2^{-0.95k} + o(1)$ for each $i$. In fact, the $E_i$ are asymptotically independent because $N \leq (d+k)^\omega$ remains bounded as $n \to \infty$. Further, let $\mathcal{D}$ be the event that the 1-core of $\Phi$ is contained in the 1/2-core of $\Phi$. Once more because $z_1, \ldots, z_{(k-1)N}$ are few and random, we find $\mathbb{P}[\mathcal{D}] = 1 - o(1)$. Finally, let $\mathcal{A}$ be the event that $\eta_{x-a}(1) \geq 1 - \exp(-2\beta)$ for all $x \in \text{Core}_{1/2}(\Phi)$, $a \in \partial_a$. Then $\mathbb{P}[\mathcal{A}] = 1 - o(1)$ by Lemmas 3.5 and 3.6. Hence, we may assume that $\mathcal{D}, \mathcal{A}$ occur.

If so, then Lemma 3.6 and the asymptotic independence of the events $E_i$ imply that the distribution of the "incoming messages" $\eta_{x-b_i}^\star$ for $z \neq y_i$ coincides asymptotically with the initialisation of the messages as in Proposition 2.6. Furthermore, because $x, a$ were chosen uniformly, the distribution of the tree $\partial^o \Phi(x \to a)$ asymptotically coincides with the distribution of $\partial^o T$. Hence, Proposition 2.6 implies that the distribution of $\eta_{x-a}$ converges to $\eta_{T}^\star$ if we take the double limit $n \to \infty$ and subsequently $\omega \to \infty$. \hfill \Box

Proposition 3.1 is immediate from Proposition 2.6 and Lemmas 2.5, 3.6 and 3.7 and Corollary 3.2.

Finally, a matching lower bound on $n^{-1}\ln \mathbb{E}_{\Phi, \beta}(\Phi)$ can be derived via a very similar approach. Indeed, building upon [5 Theorem 5.5] we need to study a variant of Belief Propagation for a "replicated" CSP where there are two truth values $\sigma_1(x), \sigma_2(x)$ assigned to each variable, a construction reminiscent of [14]. Each clause corresponds to the constraint of being satisfied with respect to both assignments $\sigma_1, \sigma_2$. The analysis of this version of Belief Propagation is very similar to and technically hardly more difficult than the plain version that we used for the upper bound.

REFERENCES

[1] D. Achlioptas, A. Coja-Oghlan: Algorithmic barriers from phase transitions. Proc. 49th FOCS (2008) 793–802.
[2] D. Achlioptas, C. Moore: Random $k$-SAT: two moments suffice to cross a sharp threshold. SIAM Journal on Computing 36 (2006) 740–762.
[3] D. Achlioptas, A. Naor, Y. Peres: Rigorous location of phase transitions in hard optimization problems. Nature 435 (2005) 759–764.
[4] D. Achlioptas, Y. Peres: The threshold for random $k$-SAT is $2^k \ln 2 - O(k)$. Journal of the AMS 17 (2004) 947–973.
[5] V. Bapst, A. Coja-Oghlan: Harnessing the Bethe free energy. Proc. 19th RANDOM (2015), to appear. arXiv:1504.03975 (2015).
[6] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Raßmann, D. Vilenchik: The condensation phase transition in random graph coloring. Proc. 18th RANDOM (2014) 449–464.
[7] V. Bapst, A. Coja-Oghlan, F. Raßmann: A positive temperature phase transition in random hypergraph 2-coloring. arXiv:1410.2190 (2014).
[8] M. Bayati, D. Gamarnik, P. Tetali: Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. Annals of Probability 41 (2013) 4080–4115.
[9] A. Coja-Oghlan: The asymptotic $k$-SAT threshold. Proc. 46th STOC (2014) 804–813. Full version: arXiv:1310.2728 (2014).
[10] A. Coja-Oghlan, K. Panagiotou: Going after the $k$-SAT threshold. Proc. 45th STOC (2013) 705–714.
[11] A. Coja-Oghlan, L. Zdeborová: The condensation transition in random hypergraph 2-coloring. Proc. 23rd SODA (2012) 241–250.
[12] P. Contucci, S. Dommers, C. Giardina, S. Starr: Antiferromagnetic Potts model on the Erdos-Rényi random graph. Communications in Mathematical Physics 323 (2013) 517–554.
[13] A. Dembo, A. Montanari, A. Sly, N. Sun: The replica symmetric solution for Potts models on $d$-regular graphs. Comm. Math. Phys. 327 (2014) 551–575.
[14] J. Ding, A. Sly, N. Sun: Satisfiability threshold for random regular NAE-SAT. Proc. 46th STOC (2014) 814–822.
[15] J. Ding, A. Sly, N. Sun: Maximum independent sets on random regular graphs. arXiv:1310.4787 (2013).
[16] J. Ding, A. Sly, N. Sun: Proof of the satisfiability conjecture for large $k$. arXiv:1411.0650 (2014).
[17] D. Gamarnik, T. Nowicki, 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 (2006) 76–106.
[18] D. Gamarnik, M. Sudan: Limits of local algorithms over sparse random graphs. Proc. 5th ITCS (2014) 369–376.
[19] D. Gamarnik, M. Sudan: Performance of the Survey Propagation-guided decimation algorithm for the random NAE-$k$-SAT problem. arXiv 1402.0052 (2014).
[20] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborova: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. National Academy of Sciences 104 (2007) 10318–10323.
[21] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press 2009.
[22] M. Mézard, G. Parisi, M. Virasoro: Spin Glass Theory and Beyond. World Scientific 1987.
Moreover, we let

\[ \theta^\beta_i(n) = \theta_{\beta^2}^\beta_i(n) \]

and

\[ \theta^\beta_i(n) = \theta_{\beta^2}^\beta_i(n) \]

Additionally, let

\[ M_\Phi^{(2)}(\mu) = \{ (\mu_i)_{i \in [n]} \in \mathcal{P}((-1,1)^2)^n \} \]

contain all families of probability distributions on \((-1,1)^2\) indexed by the Boolean variables of \(\Phi\). Similarly,

\[ \tilde{M}_\Phi^{(2)}(\mu) = \{ (\tilde{\mu}_i)_{i \in [m]} \in \mathcal{P}((-1,1)^{2k})^n \} \]

contains all families of probability distributions on \((-1,1)^{2k}\) indexed by the clauses of \(\Phi\). We use the shorthands \(\mu = (\mu_i)_{i \in [n]}\) and \(\tilde{\mu} = (\tilde{\mu}_i)_{i \in [m]}\). Further, we define

\[ M_0^{(2)}(\mu) = \left\{ \mu \in M_\Phi^{(2)}(\mu) : \frac{1}{n} |i \in [n], \mu_i, (1,1) \leq 1 - \exp(-4\beta)| \leq 2^{1-k/10} \right\} \]

Additionally, for a formula \(\Phi\), a clause \(a\) and a variable \(i\) we define \(\delta(a,i) = \delta_\Phi \setminus \{i\}\) (resp. \(\delta(i,a) = \delta_\Phi \setminus \{a\}\)). Moreover, we let \(s_a = (s_{\chi_x})_{x \in \partial a}\). For \(\mu \in M_\Phi^{(2)}\), we define the set \(\tilde{M}_\Phi^{(2)}(\mu)\) by

\[ \tilde{M}_\Phi^{(2)}(\mu) = \{ \tilde{\mu} \in \tilde{M}_\Phi^{(2)} : \forall a \in F, x \in \partial a : \sum_{(\chi_{(y)})_{y \in \partial a} \in (-1,1)^{2(k-1)}} \tilde{\mu}_a(s_a) = \mu_i(s) \} \]

For \(\mu \in M_0^{(2)}\) and \(\beta \in [0,\infty]\), we define the Bethe free energy associated to \(\Phi\) and \(\mu\) at inverse temperature \(\beta\) by

\[ B_{\Phi,\beta}(\mu) = \sup_{\tilde{\mu} \in \tilde{M}_\Phi^{(2)}(\mu)} \left\{ \frac{1}{n} \sum_{i \in [n]} (1 - d) H(\mu_i) - \frac{\beta}{n} \sum_{j \in [m]} \sum_{s_j \in (-1,1)^{2k}} \tilde{\mu}_a(s_a) \ln \left( \frac{\tilde{\mu}_a(s_a)}{\psi^{(2)}(\mu_i,\beta)(s_a)} \right) \right\} \]

where

\[ H(\mu) = - \sum_{s \in (-1,1)^2} \mu(s) \ln \mu(s) \]

is the entropy of \(\mu \in \mathcal{P}((-1,1)^2)\). Here and throughout we use the convention that \(\ln 0 = 0\). Further, we use the \(O\)-notation with the convention that \(O_N(...)\) refers to the limit \(N \to \infty\). Of course, a similar convention applies to \(o, \Omega, \text{etc.} \)
Throughout the appendix we assume that $k \geq k_0$ is large enough, that $d_- (k) \leq d \leq d_{k_{\text{SAT}}}$ and that $\beta \geq \beta_- (k)$.

The first step towards Theorem 1.1 is the proof of Proposition 1.2, which we carry out in Section B. Additionally, in Section B we also establish the following.

**Proposition A.1.** Assume that $d \in [d_- (k), d_{k_{\text{SAT}}}]$ and $\beta > \beta_- (k)$. Then

$$\frac{1}{n} \ln \mathbb{E} [Z_{\Phi} (\beta)] \sim \mathcal{F} (k, d, \beta).$$

The rest of the appendix deals with the proof of Proposition 2.4. We first make the following observation.

**Proposition A.2.** There exists an event $\mathcal{U}$ such that $P [\hat{\Phi} \in \mathcal{U}] = 1 - o(1)$ and

$$\frac{1}{n} \ln \mathbb{E} [\mathcal{C}_{\hat{\Phi}, \hat{\sigma}} (\beta)^2 1 (\hat{\Phi} \in \mathcal{U})] \leq \sup_{\mu \in M_{\mathcal{C}}^{(2)}} B_{\Phi, \beta}^{(2)} (\mu) + o(1).$$

**Proof.** Let $\mathcal{U}$ be the event that $|\text{Core}_1 (\hat{\Phi})| \geq (1 - 2^{-0.95}) n$ and that $\Phi$ has no sticky set of size between $2^{-0.95} n$ and $2^{-0.95} n$. Then we shall see in Section B Proposition B.3 that $P [\hat{\Phi} \in \mathcal{U}] = 1 - o(1)$.

We use the concept of a state and the corresponding notation from [5]. Let $k$ be the event that $\hat{\Phi} \in \mathcal{U}$. Then by [5] Theorem 4.4 there exists a state $s = s (\hat{\Phi})$ such that $\hat{\Phi} \models \epsilon s$ and

$$\langle 1 (\hat{\Phi}, \sigma) \models \epsilon s \rangle_{\Phi, \beta} = \Omega (1).$$

The state $s$ induces a marginal sequence $(s_i)_{i \in [n]}$ and we claim that this marginal sequence is skewed. Indeed, otherwise the set $Q = \{ \sigma \in \{ \pm 1 \}^n : (\hat{\Phi}, \sigma) \models \epsilon s \}$ of assignments $\sigma \in \{ \pm 1 \}^n$ has the following two properties:

$$\mathcal{C}_{\hat{\Phi}, \hat{\sigma}} (\beta) \leq O (1) \cdot \sum_{\sigma \in Q} \exp (-\beta E_{\Phi} (\sigma)),$$

$$\forall \sigma \in Q : \left| \frac{|\{ x \in \text{Core}_1 (\hat{\Phi}) \setminus S (\hat{\Phi}) : \sigma (x) = 1 \}|}{|\text{Core}_1 (\hat{\Phi}) \setminus S (\hat{\Phi})|} \right| \in |\exp (-\beta), 2^{-0.95} n|.$$  

For $\sigma \in Q$, let $T = \{ x \in \text{Core}_1 (\hat{\Phi}) \setminus S (\hat{\Phi}) : \sigma (x) = -1 \}$ and let $\bar{\sigma}$ be defined by $\bar{\sigma} (x) = 1$ if $x \in T$ and $\bar{\sigma} (x) = \sigma (x)$ otherwise. Let $\bar{Q} = \{ \bar{\sigma}, \sigma \in Q \}$. We claim that

$$E_{\Phi} (\bar{\sigma}) \leq E_{\Phi} (\sigma) - 2^{3/4} |T|.$$  

(A.1)

This will imply a contradiction. Indeed, we would have

$$\mathcal{C}_{\Phi, \sigma} (\beta) \leq O (1) \cdot \sum_{\sigma \in Q} \exp (-\beta E_{\Phi} (\sigma))$$

$$\leq O (1) \sum_{t = \lfloor \exp (-2 \beta n/2) \rfloor} \left( \begin{array}{c} n \\ln t \exp (-2 \beta n/2) \sum_{\sigma \in Q} \exp (-\beta E_{\Phi} (\sigma)) \end{array} \right)$$

$$\leq O (n) \sup_{t \in [\exp (-2 \beta n/2), 2^{-0.95} n]} \exp (t (\ln t - \beta k^{3/4} / 2)) \sum_{\sigma \in Q} \exp (-\beta E_{\Phi} (\sigma))$$

$$\leq \exp (\Omega (n)) \sum_{\sigma \in \{ -1 \}^n} \exp (-\beta E_{\Phi} (\sigma)) \leq \exp (\Omega (n)) \mathcal{C}_{\Phi, \sigma} (\beta),$$

which is clearly absurd.

We now proceed to prove (A.1). We consider for $\sigma \in Q$ the following process:

- Let $\sigma_0 = \sigma$, $V_0 = T$ and $U_0 = \sigma^{-1} (-1) \setminus V_0$.
- While the set of vertices $i \in V_t$ such that $E_{\Phi} (\sigma_i (t)) \leq E_{\Phi} (\sigma_t) - 2^{3/4}$ is not empty, pick one such vertex $i_t$ at uniformly at random and let $\sigma_{t+1} = \sigma_t (i_t)$ and $V_{t+1} = V_t \setminus \{i_t\}$.
Observe that
\[ E_\Phi (\sigma_t) \leq E_\Phi (\sigma) - k^{3/4} t. \]  
(\ref{eq:replicated_bethe_free_energy})

Let \( \tau \) be the stopping time of this process and assume that \( \tau < |T| \), or, in other words, that \( V_\tau \neq \emptyset \). We claim that \( V_\tau \) is a 1-sticky set. Indeed for \( i \in V_\tau \) we have
\[
-k^{3/4} \leq E_\Phi (\sigma_t^{(i)}) - E_\Phi (\sigma_t) \leq \left| \delta_{1,0}(i) \right| + \left| \{ a \in \delta_{1,0}a, \delta_{-1}a \cap (V_\tau \cup U_0) \neq \emptyset \} \right|
+ \left| \delta_{-1,0}a \right| + \left| \cup_{1 \leq a \leq k} \{ a \in \delta_{-1,0}a, \delta_{1}a \subset V_\tau \cup U_0 \} \right|
\]

Because \( i \in \text{Core}(\Phi) \) we have \( |\delta_{1,0}a| \geq k^{7/8}, |\delta_{-1,0}a| \leq 3, \left| \{ a \in \delta_{1,0}a, \delta_{-1}a \cap (V_\tau \cup U_0) \neq \emptyset \} \right| \leq k^{3/4} \) and \( \left| \{ a \in \delta_{1,0}a, \delta_{-1}a \cap U_0 \} \right| \geq |\delta_{-1}a|/4| \leq k^{3/4} \). Therefore, one of the following must hold.

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

It follows that the set \( V_\tau \) is 1-sticky. However, \( \text{Core}(\Phi) \setminus S(\Phi) \) cannot contain a 1-sticky set of size \( |V_\tau| \leq |T| \leq 2^{-k^{20}} \) as this would contradict the maximality of \( S(\Phi) \). It follows that \( \tau = |T| \), and therefore \( \sigma_\tau = \sigma^{(T)} \), from which (\ref{eq:replicated_bethe_free_energy}) follows using (\ref{eq:replicated_bethe_free_energy}).

Finally, since \( (s_x)_{x \in [n]} \) is skewed, the assertion follows from [5] Lemma 4.12 and [5] Lemma 4.13.

In Section \( \text{C} \) we are going to prove the following upper bound on the “replicated” Bethe free energy.

**Proposition A.3.** Whp. \( \sup_{\mu \in M_0^k} \mathcal{B}_\Phi^{(2)}(\mu) \leq 2 \mathcal{B}(k, d, \beta) + o_n(1). \)

Additionally, we need to verify a technical condition that links the bound from Proposition A.3 with the local structure of the factor graph. Let \( \Phi, \Phi' \) be regular \( k \)-SAT instances. We can think of \( \Phi, \Phi' \) as bijections from sets of clause clones to sets of variable clones (“configuration model”). Suppose that the variables and clauses of \( \Phi, \Phi' \) are \( x_i, x'_j, a_j, a'_j \), for \( i \in [n], j \in [m] \). We distinguish (variable or clause) clones \( r, r' \) of \( \Phi, \Phi' \), which we consider their roots. Moreover, we consider the first \( d/2 \) occurrences of each variable \( x_i \) its positive occurrences and the last \( d/2 \) its negative occurrences. An isomorphism \( \psi : \Phi \rightarrow \Phi' \) is a bijection with the following properties.

1. \( r' = \psi(r) \).
2. \( \psi \) maps variable clones to variable clones and clause clones to clause clones.
3. If \( \psi(v, h) = (w, j) \), then \( h = j \).
4. We have \( \psi \circ \Phi(v, h) = \Phi' \circ \psi(v, h) \) for all clones \( (v, h) \).

Let \( \omega \geq 0 \) and let \( T \) be a regular \( k \)-SAT formula with a distinguished (variable or clause) clone \( r \). For each variable clone \( (x, i) \) of \( \Phi \) we have a random variable \( 1[\delta^\omega T = \delta^\omega \Phi(x, i)] \) that indicates that the depth-\( \omega \) neighborhood of \( \Phi \) rooted at \( (x, i) \) is isomorphic to \( T \). Similarly, for each clause cone \( (a, j) \) of \( \Phi \) we consider the random variable \( 1[\delta^{\omega+1} T = \delta^{\omega+1} \Phi(a, j)] \). Let \( \Sigma_\omega \) be the \( \sigma \)-algebra generated by all these random variables. Thus, \( \Sigma_\omega \) captures the “local structure” of the random formula up to depth \( \omega \).

**Proposition A.4.** For any \( \epsilon > 0 \) there exists \( \omega_0 \) such that for any \( \omega \geq \omega_0 \) there is \( n_0 = n_0(\epsilon, \omega) > 0 \) such that for all \( n > n_0 \) the following is true. Let \( \mathcal{B}_\epsilon \) be the event that \( \sup_{\mu \in M_0^k} \mathcal{B}_\Phi^{(2)}(\mu) \leq 2 \mathcal{B}(k, d, \beta) + \epsilon \). Then

\[
P\left[ \ln \mathbb{E} \left[ \mathcal{C}_\Phi (\phi) 1[\mathcal{B}_\epsilon] | \Sigma_\omega \right] \geq n(\mathcal{B}(d, k, \beta) - \epsilon) \right] > 1 - \epsilon.
\]

To prove Proposition A.4, we introduce a further kind of random formula, the planted replica model, which we denote by \( \Phi = \Phi(d, k, \beta, \omega) \). Just as the planted model \( \Phi \), the random formula \( \Phi \) comes with a reference assignment \( \phi \).

To define \( \Phi = \Phi(d, k, \beta, \omega) \) we need a little preparation. Let \( T \) be a \( k \)-SAT instance whose factor graph is a tree rooted at a variable clone \( r \). Initialising all messages \( \eta^{(0)}_{T,x,a}(-1) = 1 - \eta^{(0)}_{T,x,a}(1) = 1 - \exp(-4d\beta) \), let \( \eta^{(0)}_{T,x,a}(\pm 1) \) be the messages after \( \omega \) iterations of Belief Propagation and let

\[
\mu^{(0)}_\omega(\pm 1) \propto \prod_{a \in \partial \phi} \eta^{(0)}_{T,a,r}(\pm 1).
\]
For formulas $T$ that are not trees we define $\mu_T^{(x)}(\pm 1) = 1/2$.

Further, given a regular $k$-SAT formula $\Phi$ we define the $\omega$-type of a variable clone $(x, i)$ as the isomorphism class $\theta(x, i, \Phi, \omega)$ of the sub-formula $\partial_\omega(x, i)$. In addition, the $\omega$-type $\theta(a, j, \Phi, \omega)$ of a clause clone $(a, j)$ is the isomorphism class of the sub-formula $\partial_\omega(a, j)$.

We now describe the experiment that yields $(\Phi, \hat{\sigma})$. There are four steps.

1. Choose a random formula $\Phi$.
2. For each variable $x$ let $\hat{\sigma}(x) = \pm 1$ with probability $\mu_x^{(2\omega)}(\pm 1) = \mu_x^{(2\omega)}(\pm 1)$ independently.

To proceed, let $a$ be a clause. Pick a distribution $\mu_a^{(\omega)}$ on $\{\pm 1\}^k$ that minimises $\mathbb{E}\mathbb{E}_\mu^{(\omega)}()$ subject to the condition that for each $x \in \partial_a \Phi$ the $x$-marginal of $\mu_a^{(2\omega)}$ equals $\mu_x^{(2\omega)}$.

3. For every clause $a$ choose $\hat{\sigma}(a) \in \{\pm 1\}^k$ independently from the distribution $\mu_a^{(2\omega)}$.
4. Choose a bijection $\hat{\Phi} : \{a_1, \ldots, a_m\} \times [k] \rightarrow \{x_1, \ldots, x_n\} \times [d]$ uniformly at random subject to the following conditions.
   4a) If $\Phi(a, j) = (x, i)$, then $\hat{\Phi}(a, j) = (y, i)$ for some variable $y$.
   4b) If $\Phi(a, j) = (x, i)$, then $\hat{\sigma}(a, j) = \hat{\sigma}(x)$.
   4c) If $\Phi(a, j) = (x, i)$ and $\hat{\Phi}(a, j) = (y, i)$, then $\theta(x, i, \hat{\Phi}, 4\omega) = \theta(y, i, \hat{\Phi}, 4\omega)$.

If no such bijection exists, we start over from (1).

The following observation is a consequence of [5] Section 5.3).

**Fact A.5.**

(i) The probability that there exists a bijection $\hat{\Phi}$ that satisfies (4a)–(4c) is $\exp(o(n))$.

(ii) With probability $\exp(o(n))$ we have $\theta(v, j, \Phi, 4\omega) = \theta(v, j, \hat{\Phi}, 4\omega)$ for all clones $(v, j)$.

In Section 3 we are going to prove the following fact about the planted replica model.

**Proposition A.6.** Wh.p. we have $\sup_{\mu \in M_{d_\omega}^{(\omega)}} B^{(2)}_{\Phi, \beta}(\mu) \leq 2B(k, d, \beta) + o_n(1)$.

Together with the argument from [5] Section 4.6] Proposition A.4 follows from Proposition A.6.

Finally, we will rely on the following standard concentration results.

**Lemma A.7.** Let $d \leq d_{k-SAT}$ and $\beta \in \mathbb{R}$ be fixed. For any $\alpha > 0$ there is $\delta > 0$ such that

\[
\mathbb{P}\left[\left|\frac{1}{n} \ln Z_\Phi(\beta) - \frac{1}{n} \ln Z_\Phi(\beta)\right| > \alpha\right] < \exp(-\delta n),
\]

\[
\mathbb{P}\left[\left|\frac{1}{n} \ln \mathcal{E}_\Phi(\beta) - \frac{1}{n} \ln \mathcal{E}_\Phi(\beta)\right| > \alpha\right] < \exp(-\delta n).
\]

**Proof:** The proof follows from the fact that if two formula $\Phi, \Phi'$ differ by at most one switch of edges, the associated partition functions satisfy

$|\ln Z_\Phi(\beta) - \ln Z_\Phi(\beta)| \leq 2\beta$ and, for $\sigma \in \{-1, 1\}^n$ $|\ln \mathcal{E}_\Phi(\beta) - \ln \mathcal{E}_\Phi(\beta)| \leq 2\beta$.

The stated concentration result is then a consequence of Azuma’s inequality (applied to the configuration model).

**Proof of Theorem 1.1** Propositions A.2, A.3 and A.4 together with [5] Theorem 5.5 imply that $\frac{1}{n} \ln \mathcal{E}_\Phi(\beta)$ converges to $\mathcal{B}(k, d, \beta)$ in probability. Therefore, the theorem (and its corollaries) follow from Propositions 1.3, 2.1 and A.1 and Lemmas 2.2 and A.7.

**Appendix B. The fixed point problem on trees**

In this section we prove Proposition 1.2 and Proposition A.1 We assume that $d \in [d_{-}(k), d_{k-SAT}]$ and that $\beta \geq \beta_{-}(k)$.

In the following of the paper we define $c_\beta = 1 - \exp(-\beta)$. 
B.1. The multi-type Galton-Watson branching process. We first analyse the fixed points of the operator $\Phi_{k,d,\beta}$ and explain how they relate to the Galton-Watson trees defined in Section 2.4. For $\pi, \tilde{\pi} \in \mathcal{P}(0,1)$ we define

\[
\hat{h}(\pi) = \int_{(0,1)} \eta \, d\pi(\eta), \quad \hat{\tilde{h}}(\tilde{\pi}) = \int_{(0,1)} \tilde{\eta} \, d\tilde{\pi}(\tilde{\eta}).
\]

We let $f : (0,1)^{d-1} \to (0,1)$ (resp. $\hat{f} : (0,1)^{k-1} \to (0,1)$) be defined by

\[
f(\tilde{\eta}_1, \ldots, \tilde{\eta}_{d-1}) = \frac{\prod_{j=1}^{d/2} \tilde{\eta}_j \prod_{j=d/2}^{d-1} (1 - \tilde{\eta}_j)}{\tilde{Z}(\tilde{\eta}_1, \ldots, \tilde{\eta}_{d-1})}, \quad \hat{f}(\eta_1, \ldots, \eta_{k-1}) = \frac{1 - c_\beta \prod_{j=1}^{k-1} \eta_j}{\tilde{Z}(\eta_1, \ldots, \eta_{k-1})},
\]

and $f_d, \hat{f}_{k,\beta} : [0,1] \to (0,1)$ be defined by

\[
f_d(\tilde{\eta}) = f(\tilde{\eta}), \quad \hat{f}_{k,\beta}(\eta) = \tilde{f}(\eta).
\]

We say that $(\pi, \tilde{\pi})$ is a fixed point of $(\mathcal{F}_{k,d,\beta}, \hat{\mathcal{F}}_{k,d,\beta})$ iff $\pi = \mathcal{F}_{k,d,\beta}(\tilde{\pi})$ and $\tilde{\pi} = \hat{\mathcal{F}}_{k,d,\beta}(\pi)$.

**Lemma B.1.** If $(\pi, \tilde{\pi})$ is a fixed point of $(\mathcal{F}_{k,d,\beta}, \hat{\mathcal{F}}_{k,d,\beta})$, then we have

\[
h[\pi] = f_d(\hat{h}(\tilde{\pi})), \quad \tilde{h}[\pi] = \hat{f}_{k,\beta}(h[\pi]).
\]

**Proof.** We first observe that

\[
Z[\tilde{\pi}] = \int_{(0,1)^{d-1}} \tilde{Z}(\tilde{\eta}_1, \ldots, \tilde{\eta}_{d-1}) = \int_{(0,1)^{k-1}} \tilde{Z}(\eta_1, \ldots, \eta_{k-1}) d\pi(\eta_j)
\]

Using these equations, we obtain

\[
h[\pi] = \int_{(0,1)} \eta \, d\mathcal{F}_{k,d,\beta}[\pi](\eta) = \frac{1}{Z[\tilde{\pi}]} \int_{(0,1)^{d-1}} \tilde{Z}(\tilde{\eta}_1, \ldots, \tilde{\eta}_{d-1}) \hat{f}(\tilde{\eta}_1, \ldots, \tilde{\eta}_{d-1}) \prod_{j=1}^{d-1} d\tilde{\eta}_j
\]

Similarly, we have

\[
\tilde{h}[\pi] = \int_{(0,1)} \tilde{\eta} \, d\hat{\mathcal{F}}_{k,d,\beta}[\pi](\tilde{\eta}) = \frac{1}{Z[\pi]} \int_{(0,1)^{k-1}} \tilde{Z}(\eta_1, \ldots, \eta_{k-1}) \hat{f}(\eta_1, \ldots, \eta_{k-1}) \prod_{j=1}^{k-1} d\pi(\eta_j)
\]

Recalling the definition of $q = q(d, k, \beta)$ in Section 2.4 and defining $\tilde{q} = 1 - q$, the following is a simple observation.

**Fact B.2.** The set of equations

\[
h = 1 - \hat{h}, \quad \tilde{h} = \hat{f}_{k,\beta}(h),
\]

admits for unique solution in $[0,1]^2$ the pair $(q, \tilde{q})$. 

16
We define the measures $\pi_+, \pi_-, \hat{\pi}_+$ and $\hat{\pi}_-$ over $(0,1)$ by

$$d\pi_+(\eta) = \frac{1-\eta}{1-q} d\pi(\eta), \quad d\pi_-(\eta) = \frac{\eta}{q} d\pi(\eta), \quad (B.1)$$

$$d\hat{\pi}_+(\eta) = \frac{1-\hat{\eta}}{1-q} d\hat{\pi}(\hat{\eta}), \quad d\hat{\pi}_-(\hat{\eta}) = \frac{\hat{\eta}}{q} d\hat{\pi}(\hat{\eta}). \quad (B.2)$$

**Lemma B.3.** If $(\pi, \hat{\pi})$ is a fixed point of $(\mathcal{F}_{k,d,\beta}, \hat{\mathcal{F}}_{k,d,\beta})$, we have

$$\pi_- = \int_{(0,1)^{d-1}} \delta_{\hat{\pi}(\eta_1, \ldots, \hat{\eta}_{d-1})}^{d/2-1} \bigotimes_{j=1}^{d-1} d\hat{\pi}_-(\hat{\eta}_j) \bigotimes_{j=d/2} d\pi_+(\eta_j), \quad (B.3)$$

$$\pi_+ = \int_{(0,1)^{d-1}} \delta_{\hat{\pi}(\eta_1, \ldots, \hat{\eta}_{d-1})}^{d/2-1} \bigotimes_{j=1}^{d-1} d\hat{\pi}_+(\hat{\eta}_j) \bigotimes_{j=d/2} d\pi_-(\eta_j), \quad (B.4)$$

$$\hat{\pi}_- = \sum_{r=1}^{k-1} \left( \frac{k-1}{r} \right) q^r (1-q)^{k-1-r} \int_{(0,1)^{k-1}} \delta_{\hat{\pi}(\eta_1, \ldots, \hat{\eta}_{k-1})}^{r} \bigotimes_{j=r+1}^{k-1} d\pi_-(\eta_j) \bigotimes_{j=1}^{r-1} d\pi_+(\eta_j), \quad (B.5)$$

$$\hat{\pi}_+ = \int_{(0,1)^{k-1}} \sum_{r=0}^{k-1} \left( \frac{k-1}{r} \right) q^r (1-q)^{k-1-r} \delta_{\hat{\pi}(\eta_1, \ldots, \hat{\eta}_{k-1})}^{r} \bigotimes_{j=r+1}^{k-1} d\pi_-(\eta_j) \bigotimes_{j=1}^{r-1} d\pi_+(\eta_j). \quad (B.6)$$

**Proof.** We first observe that $q Z[\hat{\pi}] = (\hat{q})^{d/2-1} (1-\hat{q})^{d/2}$. We then compute

$$\pi_- = \int_{(0,1)} \frac{\eta}{q} \delta_{\eta}^{d\pi(\eta)} = \int_{(0,1)} \frac{\eta}{q} \delta_{\eta}^{d\mathcal{F}_{k,d,\beta}[\hat{\pi}](\eta)}$$

$$= \frac{1}{Z[\hat{\pi}]} \int_{(0,1)^{d-1}} \frac{1}{q} \pi_{\hat{\pi}}^{\eta} \bigotimes_{j=1}^{d-1} f(\hat{\eta}_1, \ldots, \hat{\eta}_{d-1}) \delta_{\hat{\pi}(\eta_1, \ldots, \hat{\eta}_{d-1})}^{d-1} \bigotimes_{j=d/2} d\hat{\pi}(\hat{\eta}_j)$$

$$= \int_{(0,1)^{d-1}} \prod_{j=1}^{d/2-1} \hat{\eta}_j \prod_{j=d/2}^{d-1} (1-\hat{\eta}_j) \delta_{\hat{\pi}(\eta_1, \ldots, \hat{\eta}_{d-1})}^{d-1} \bigotimes_{j=d/2} d\hat{\pi}(\hat{\eta}_j)$$

$$= \int_{(0,1)^{d-1}} \delta_{\hat{\pi}(\eta_1, \ldots, \hat{\eta}_{d-1})}^{d/2-1} \bigotimes_{j=1}^{d/2} d\hat{\pi}_-(\hat{\eta}_j) \bigotimes_{j=d/2} d\pi_+(\eta_j)$$

The equation on $\pi_+$ is proved similarly. We also compute

$$\hat{\pi}_- = \int_{(0,1)} \frac{\hat{\eta}}{q} \delta_{\hat{\eta}}^{d\hat{\pi}(\hat{\eta})} = \int_{(0,1)} \frac{\hat{\eta}}{q} \delta_{\hat{\eta}}^{d\hat{\mathcal{F}}_{k,d,\beta}[\pi](\hat{\eta})}$$

$$= \int_{(0,1)^{k-1}} \frac{1-c_\beta \prod_{j=1}^{k} \hat{\eta}_j}{1-c_\beta \prod_{j=1}^{k} \hat{\eta}_j} \delta_{\hat{\pi}(\eta_1, \ldots, \eta_{k-1})}^{k-1} \bigotimes_{j=1}^{k-1} d\pi(\eta_j)$$

$$= \sum_{r=0}^{k-2} \left( \frac{k-1}{r} \right) q^r (1-q)^{k-1-r} \frac{1-c_\beta q^{k-1}}{17}$$
The equation on \( \tilde{\pi}_+ \) is proved in a similar manner.

The following lemma is easily verified by reversing the order of the equations above.

**Lemma B.4.** Reciprocally, let \((\pi_+, \pi_-, \tilde{\pi}_+, \tilde{\pi}_-) \in \mathcal{P}(0, 1)^4\) satisfy (B.3)-(B.6). Then \((\pi, \tilde{\pi})\) defined by

\[
\pi = (1 - q)\pi_+ + q\pi_- , \quad \tilde{\pi} = (1 - \tilde{q})\tilde{\pi}_+ + \tilde{q}\tilde{\pi}_-
\]

is a fixed point of \((\mathcal{F}_k, d, \beta, \mathcal{F}_k, d, \beta)\).

For \(\omega \geq 0\), let \(\mathcal{T}_\omega\) be the set of trees generated by restricting the Galton-Watson process described in Section 2.4 to individuals at distance at most \(2\omega\) from the root. We observe that the trees in \(\mathcal{T}_\omega\) only differ through the types \((-\) or \(+\)) of their variables and clauses. Therefore, it will prove convenient to unify the notations. For each \(\omega \geq 0\), we choose \(T^0_\omega \in \mathcal{T}_\omega\). We can see every tree \(T \in \mathcal{T}_\omega\) as a (re-)labeling of the variables and clauses of \(\mathcal{T}^0_\omega\) with labels \(+\) or \(-\). We denote by \(V_\omega\) the set of variables of \(T^0_\omega\), by \(F_\omega\) its set of clauses, and by \(E_\omega\) its set of edges. To keep the notations simple, we will use in the remaining of this section the convention that the letters \(i, j, i_1, \ldots\) denote variables and \(a, b, a_1, \ldots\) denote clauses. We denote by \(i_0\) the root of \(T^0_\omega\) and we also introduce an additional clause \(a_0\) of type \(+\). We let \(E^+_\omega\) be the set of its directed edges toward the root, i.e.

\[
E^+_\omega = \{(i_0, a_0)\} \cup \{(i, a), (i, a) \in E_\omega, \text{dist}(i_0, a) \leq \text{dist}(i_0, i)\}
\]

We denote by \(\partial V_\omega\) the subset of \(V_\omega\) formed by the variables at distance \(2\omega\) from \(i_0\) and by \(\partial E^+_\omega\) the set of \((i, a) \in E^+_\omega\) with \(i\) at distance \(2\omega\) from \(i_0\). We let \(d_\omega = (d - 1)^\omega(k - 1)^\omega = |\partial V_\omega| = |\partial E^+_\omega|\). For \(T \in \mathcal{T}_\omega\) we denote identically \(\partial T = \partial E^+_\omega\), and we use the index \(\partial T\) to denote vectors indexed by the edges \((i, a) \in \partial E^+_\omega\). For instance given for each edge \((i, a) \in \partial E^+_\omega\) a real number \(\eta_{i-a}\), \(\eta_{\partial T}\) denotes the vector \((\eta_{i-a})_{(i, a) \in \partial E^+_\omega}\).

For a fixed tree \(T \in \mathcal{T}_\omega\) and a boundary condition \(\eta_{\partial T} \in (0, 1)^{d_\omega}\), we consider the set of messages

\[
((\eta_{i-a}(T, \eta_{\partial T}))_{(i, a) \in E^+_\omega}, (\tilde{\eta}_{i-a}(T, \eta_{\partial T}))_{(a, i) \in E^+_\omega}) \in (0, 1)^{|E^+_\omega|}
\]

which coincide with \(\eta_{\partial T}\) on \(\partial E^+_\omega\) and are the unique solution to the following equations, for \((i, a) \in E^+_\omega\) (resp. \((b, i) \in E^+_\omega\))

\[
\eta_{i-a} = \frac{\prod_{b \in \partial(i, a)} \eta_{b-i}}{\prod_{b \in \partial(i, a)} \eta_{b-i} + \prod_{b \in \partial(i, a)} (1 - \eta_{b-i})} , \quad \tilde{\eta}_{b-i} = \frac{1 - \prod_{j \in \partial(b, i)} \eta_{j-b}}{2 - \prod_{j \in \partial(b, i)} \eta_{j-b}}
\]

We use the shorter notation

\[
(\eta(T, \eta_{\partial T}), \tilde{\eta}(T, \eta_{\partial T})) = ((\eta(T, \eta_{\partial T})_{i-a})_{(i, a) \in E^+_\omega}, (\tilde{\eta}(T, \eta_{\partial T})_{a-i})_{(a, i) \in E^+_\omega}),
\]

and we say that \((\eta(T, \eta_{\partial T}), \tilde{\eta}(T, \eta_{\partial T}))\) are the \textit{Belief Propagation messages on} \(T\) induced by the boundary condition \(\eta_{\partial T}\). More generically, we shall speak of messages on a tree \(T \in \mathcal{T}_\omega\) \((\eta, \tilde{\eta})\), as a solution of the \textit{Belief Propagation equations of} \(T\) if there exist a boundary condition \(\eta_{\partial T} \in (0, 1)^{d_\omega}\) on \(T\) such that \((\eta, \tilde{\eta}) = (\eta(T, \eta_{\partial T}), \tilde{\eta}(T, \eta_{\partial T}))\).

For \(T \in \mathcal{T}_\omega\), we define \(\partial_1 T\) (resp. \(\partial_{-1} T\)) as the subset of \(\partial T\) formed by variables of type \(+\) (resp. of type \(-\)) in \(T\). We also let \(p^{(\omega)}_{k, d, \beta}(T)\) be the probability that the random process \(GW(k, d, \beta)\) restricted to depth \(2\omega\) has generated the tree \(T\). The following lemma is an immediate consequence of the very construction of our random process.
Lemma B.5. Let $\pi$ be a fixed point of $\Psi_{k,d,\beta}$ and $\pi_+, \pi_-$ be defined by $\text{[B.1][B.2]}$. We have for $t \in \{0, \ldots, \omega\}$

$$\pi = \sum_{T \in \mathcal{T}_\omega} p^{(t)}_{k,d,\beta}(T) \int_{(0,1)^{d_1}} \delta_{\eta_{i_0-a_0}(T \eta_{GT})} \bigotimes_{(i,a) \in E^+_{\omega}} d\pi_+(\eta_{i-a}) \bigotimes_{(i,a) \in E^-_{\omega}} d\pi_-(\eta_{i-a}).$$

Proof. The lemma is easily proved by recurrence over $t \in \{0, \ldots, \omega\}$, using Lemma B.3 and Lemma B.4. \qed

B.2. Existence and unicity of the fixed point. A standard method to prove the existence and unicity of the fixed point of an operator acting on a Banach space is to use the contraction method, i.e. to prove that the operator is $K$-Lipschitz continuous for some $K < 1$. However, because we will need in Section C to carry out a similar fixed point analysis in the setting of an iteration operator on a formula $\Phi$, we will take a more pedantic route here that will prove useful later on.

We first introduce a different parametrization of the tree and of the messages $(\eta, \tilde{\eta})$, and a few more notations. Let $T \in \mathcal{T}_\omega$ be fixed.

- For $(i, a) \in E^+_{\omega}$, we define $b_{a,i} = -1$ if $i$ has type $+$, $b_{a,i} = 1$ otherwise.
- For $(a, i) \in E^+_{\omega}$, we define $b_{a,i} = -1$ if $a$ has type $+$, $b_{a,i} = 1$ otherwise.
- For $(i, a) \in E^-_{\omega}$ we define
  \[ \partial_1(i, a) = \{b, (b, i) \in E^+_{\omega}, b_{b,i} = -1\}, \quad \partial_{-1}(i, a) = \{i, (a, i) \in E^+_{\omega}, b_{b,i} = 1\}. \]
- For $(a, i) \in E^-_{\omega}$ we define
  \[ \partial_1(a, i) = \{j, (j, a) \in E^+_{\omega}, b_{a,j} = -1\}, \quad \partial_{-1}(a, i) = \{j, (a, j) \in E^+_{\omega}, b_{a,j} = 1\}. \]

We observe that a description of a tree $T \in \mathcal{T}_\omega$ solely in terms of variables, clauses and numbers $b_{a,i}$ is equivalent to our previous description in terms of variables aux clauses of type $+$ and $-$. Thereby in the following we identify $\mathcal{T}_\omega$ with the set of bipartite trees on individuals of type variable or clause, of depth $2\omega$, rooted at a clause, and with numbers $(b_{a,i})_{(a,i) \in E_{\omega}} \in \{-1, 1\}^{E_{\omega}}$. Accordingly, for all $(i, a) \in E^+_{\omega}$ (resp. $(b, i) \in E^+_{\omega}$), we define

$$\varepsilon_{i-a} = \frac{1}{2} - (-1)^{b_{a,i}} \eta_{i-a} - \frac{1}{2}, \quad \tilde{\varepsilon}_{b-i} = \frac{1 - \tilde{\eta}_{j-b}}{\tilde{\eta}_{j-b}}. \quad (B.7)$$

We denote by $(\varepsilon, \tilde{\varepsilon}) = ((\varepsilon_{i-a})_{(i,a) \in E^+_{\omega}}, (\tilde{\varepsilon}_{a-i})_{(a,i) \in E^-_{\omega}})$. It is easily seen that $(\eta, \tilde{\eta})$ is a solution of the Belief Propagation equation on $T$ if and only if $(\varepsilon, \tilde{\varepsilon})$ satisfy the following equations, for $(i, a) \in E^+_{\omega} \setminus \partial E^+_{\omega}$ (resp. $(b, i) \in E^+_{\omega}$)

$$\varepsilon_{i-a} = \frac{1}{1 + P_{i-a}}, \quad P_{i-a} = \prod_{b \in \partial_1(i,a)} \tilde{\varepsilon}_{b-i} \prod_{b \in \partial_{-1}(i,a)} \frac{1}{\tilde{\varepsilon}_{b-i}}, \quad (B.8)$$

$$\tilde{\varepsilon}_{b-i} = \left(1 - c_{b} \prod_{j \in \partial_1(b,i)} \varepsilon_{j-b} \prod_{j \in \partial_{-1}(b,i)} (1 - \varepsilon_{j-b})\right)^{-1}. \quad (B.9)$$

We denote by $(\varepsilon(T, \varepsilon_{\partial T}), \tilde{\varepsilon}(T, \varepsilon_{\partial T}))$ the messages defined by Eq. B.8-B.9, with the boundary condition $\varepsilon_{i-a} = (\varepsilon_{\partial T})_{i-a}$ for $(i, a) \in \partial E^+_{\omega}$. Our concern is to understand how $\varepsilon_{i-a} = (\varepsilon_{\partial T})_{i-a}$ behaves when $T$ is drawn at random from $p^{(a)}_{k,d,\beta}$ and the boundary messages are drawn i.i.d. from a fixed point $\pi$ of $\Psi_{k,d,\beta}$. To deploy our analysis, it will also be needed to consider the messages induced on a tree $T \in \mathcal{T}_\omega$ by Eq. B.8-B.9 with a different boundary condition. We denote by $\varepsilon^{(a)}(T)$ the messages $(\varepsilon, \tilde{\varepsilon})$ defined by Eq. B.8-B.9, with the boundary condition $\varepsilon_{i-a} = \exp(-4d\beta)$ for all $(i, a) \in \partial E^+_{\omega}$. Similarly, we let $\eta^{(a)}(T)$ denote the corresponding messages $(\eta, \tilde{\eta})$, and we finally define

$$\pi^{(a)} = \sum_{T \in \mathcal{T}_\omega} p^{(a)}_{k,d,\beta}(T) \delta_{\eta^{(a)}(T)}. \quad (B.10)$$

We now present a simpler construction that will allow us to better understand the convergence of $\pi^{(a)}$ toward a fixed point of $\Psi_{k,d,\beta}$. We shall consider a setting where each variable node of a tree $T \in \mathcal{T}_\omega$ is
assigned a state which can be frozen or non frozen. Let $(T, \epsilon_{\partial T}) \in \mathcal{T}_\omega \times \{\text{frozen, non frozen}\}$ be fixed. We extend $\epsilon_{\partial T}$ into an assignment of the variables of $T$ in the following way. We set $V_t^{(0)} = \{i \in \partial T, \epsilon_i = \text{frozen}\}$. Having defined $\epsilon_i$ for $i$ at distance $t - 1$ from $\partial T$, we set $V_t^{(t-1)} = \{i$ at distance $2(t - 1)$ from $\partial T$ is frozen$\}$. For $i \in T$ at distance $t$ from $\partial T$, we then set $\epsilon_i = \text{frozen if and only if the five following conditions (reminiscent of the core construction) hold true. Let $a'$ be the clause such that $(i, a') \in E_i^{(\omega)}$.}

- $|a \in \partial_1(i, a')| \geq k^{7/8}$,
- $\partial_1(i, a') \leq 3$,
- for any $1 \leq l \leq k$ the number of clauses $a \in \partial_1(i, a')$ such that $|\partial_1 a| = l$ is bounded by $k^{l+3}/l!$,
- there are no more than $k^{3/4}$ clauses $a \in \partial_1 i$ such that $|\partial_1 a| = 1$ but $\partial(a, i) \notin V_t^{(t-1)}$,
- there are no more than $k^{3/4}$ clauses $a \in \partial_1 i$ such that $|\partial_1 a| < k$ and $|\partial_1(a, i)| \leq |\partial_1 a|/4$.

We denote by $\epsilon(T, \epsilon_{\partial T})$ the element of $\{\text{frozen, non frozen}\}$ thus obtained, and by $V_t(T, \epsilon_{\partial T}) = \bigcup_{t=0}^\omega V_t^{(t)}$ the associated set of frozen variables. We say that a clause $a \in F_\omega$ is cold for the pair $(T, \epsilon_{\partial T})$ iff $|\partial_1 a \cap V_t(T, \epsilon_{\partial T})| \geq 1$. We say that a variable $j \in V_{\omega} \setminus \partial V_{\omega}$ is cold if either $\epsilon_j = \text{frozen}$ or the clause $a$ such that $(j, a) \in E_j^+ \text{ is cold.}$ For two variables $i, j \in V_{\omega}^+$, we denote by $[i \rightarrow j]$ the shortest non backtracking path from $i$ to $j$. We say that a path $[i_0 \rightarrow i_{\omega}]$ with $i_{\omega} \in \partial V_{\omega}$, is cold for the pair $(T, \epsilon_{\partial T})$ iff it contains at least $|0.4\omega|$ cold variables. Finally, we say that the pair $(T, \epsilon_{\partial T})$ is $\omega$-frozen iff all the paths $[i_0 \rightarrow i_{\omega}]$, with $i_0 \in \partial V_{\omega}$, are cold. The following proposition, that we prove in Section B.4, is crucial to our results.

**Proposition B.6.** Let $(T, \epsilon_{\partial T}) \in \mathcal{T}_\omega \times \{\text{frozen, non frozen}\}$ be distributed according to a probability distribution $\mathbb{P}$ such that the following is true.

1. The marginal distribution of $T$ is $p_{k,d,\beta}^{(\omega)}$.
2. For all $T \in \mathcal{T}_\omega$, $i \in \partial T$, and $(\epsilon_j)_{j \in \partial T \setminus \{i\}}$, we have $\epsilon_i = \text{non frozen with probability at most } 2^{-0.9k}$.

Then we have

$$\mathbb{P}\left[(T, \epsilon_{\partial T}) \text{ is } \omega\text{-frozen}\right] = 1 - o_\omega(1).$$

To each pair $(T, \epsilon_{\partial T}) \in \mathcal{T}_\omega \times (0, 1]$ we associate the pair $(T, \epsilon_{\partial T}(\epsilon_{\partial T})) \in \mathcal{T}_\omega \times \{\text{frozen, non frozen}\}$ defined by, for $(i, a) \in \partial E_i^+$, $\epsilon_i(\epsilon_{\partial T}) = \text{frozen iff } \epsilon_{i \rightarrow a} \leq \exp(-2\beta)$. We say that $(T, \epsilon_{\partial T})$ is $\omega$-good iff the associated pair $(T, \epsilon_{\partial T}(\epsilon_{\partial T}))$ is $\omega$-frozen. The following is an immediate reformulation of the above proposition.

**Corollary B.7.** Let $(T, \epsilon_{\partial T}) \in \mathcal{T}_\omega \times (0, 1]$ be distributed according to a probability distribution $\mathbb{P}$ such that the following is true.

1. The marginal distribution of $T$ is $p_{k,d,\beta}^{(\omega)}$.
2. For all $T \in \mathcal{T}_\omega$, $(i, a) \in \partial E_i^+$, and $(\epsilon_j)_{j \in \partial T \setminus \{i\}}$, we have $\epsilon_i > \exp(-2\beta)$ with probability at most $2^{-0.9k}$.

Then we have

$$\mathbb{P}\left[(T, \epsilon_{\partial T}) \text{ is } \omega\text{-good}\right] = 1 - o_\omega(1).$$

Moreover, the following result is easily proved by recurrence over $\omega$.

**Lemma B.8.** Let $(T, \epsilon_{\partial T}) \in \mathcal{T}_\omega \times (0, 1]$ be $\omega$-good, $(\epsilon, \hat{\epsilon}) = (\epsilon(T, \epsilon_{\partial T}), \hat{\epsilon}(T, \epsilon_{\partial T}))$ be the corresponding solution of the Belief Propagation equations on $T$, and $\epsilon = \epsilon(T, \epsilon_{\partial T}(\epsilon_{\partial T}))$. Let $(i, a) \in E_i^+$. Then if $i \in V_{\omega} \setminus \partial V_{\omega}$ and $\epsilon_i = \text{frozen}$, we have $\epsilon_{i \rightarrow a} < \exp(-100\beta)$ and $\epsilon_{i \rightarrow a} < \exp(-100\beta)$.

**Proof.** We first prove the statement concerning $\epsilon$. We introduce, for $t \in [0, \ldots, \omega]$, $V_t^{(t)} = V_t^{(t)}(T, \epsilon_{\partial T}(\epsilon_{\partial T}))$. We prove the result by recurrence over the distance $t \in [1, \ldots, \omega]$ from $i \in V_{\omega}$ to $\partial T$. Simultaneously, we prove by recurrence over $t \in [0, \ldots, \omega]$ that if $\epsilon_j = \text{frozen}, \epsilon_{i \rightarrow a} < \exp(-100\beta)$. For $t = 0$ this result is true by definition of $\omega$-good pairs. We assume that this result holds for $(t-1) \in [0, \omega - 1]$. Let $(i, a) \in E_i^+$ be such that $i$ is at distance $t$ from $\partial T$. Then the following estimates hold.

- For all $b \in \partial_{1,0}(i, a)$ such that $\partial(a, i) \subseteq V_{t-1}^{(t-1)}$ we have $\hat{\epsilon}_{b \rightarrow i} \geq \exp(\beta/2)$,
• For all $1 \leq l \leq k$, $b$ in $\partial_{-1,l}(i, a)$ such that $|\partial_{1}(a, i) \setminus V_{l}^{(i-1)}| \geq l/4$, we have $\hat{\epsilon}_{b-i} \geq \exp(-\exp(-l\beta/3))$. Because $i$ is frozen, there are at least $k^{7/18}/2$ clauses $b \in \partial(i, a)$ such that the first condition holds, at most $k^{1/3}/l!$ clauses $b \in \partial_{-1,l}(i, a)$ such that the second condition holds (for a given $1 \leq l \leq k$), and at most $3+2k^{3/4}$ clauses $b \in \partial(i, a)$ such that none of these conditions hold. For the latter case we use the simplest bound $\frac{1}{\epsilon_{b-i}} \geq \exp(-\beta)$. Consequently we obtain by replacing in (B.8)

$$P_{l \rightarrow a} \geq \exp\left[k^{7/18}\beta/4\right] \exp\left(\sum_{l \geq 1} (k^{1/3} \exp(-\beta l/3) / l!)\right)^{-1} \exp(-3k^{3/4}l) \geq \exp(200\beta)$$

It follows that $\epsilon_{i \rightarrow a} \leq \exp(-100\beta)$ as desired. The second part of the lemma, regarding $\epsilon^{(\omega)}$, is proved in a completely analogous manner.

**Lemma B.9.** For a sequence $(T_{\omega}, (\varepsilon_{\partial T_{\omega}})_{\omega})_{\omega \geq 0}$ such that, for all $\omega \geq 0$, $(T_{\omega}, (\varepsilon_{\partial T_{\omega}})_{\omega}) \in \mathcal{T}_{\omega} \times (0, 1)^{d_{\omega}}$ is $\omega$-good, the following is true.

$$|\varepsilon^{(\omega)}_{i \rightarrow a} - \varepsilon^{(\omega)}_{i \rightarrow a}| = o_{\omega}(1).$$

**Proof.** We fix $\omega \geq 0$ and denote $(\varepsilon^{\dagger}, \varepsilon^{\ddagger}) = (\varepsilon(T_{\omega}, (\varepsilon_{\partial T_{\omega}})_{\omega})$ and $\epsilon^{\dagger} = \epsilon(\varepsilon(\partial T_{\omega}), (\varepsilon_{\partial T_{\omega}}))$ to keep the notations simple. For $j \in T$, let $\omega_{j} = \omega - \text{dist}(i_{0}, j)$. For $(j, a) \in E_{\omega}^{+}$, we define

$$V_{j \rightarrow a} = \{(\varepsilon^{\dagger}, \forall a' \in \partial(j, a), a'_{j-} \in [\min(a'_{j-}, \varepsilon^{(\omega)}_{a'_{j-}}), \max(\varepsilon^{\dagger}_{a'_{j-}}, \varepsilon^{(\omega)}_{a'_{j-}})]\}.$$

By applying Taylor’s theorem to Eq. (B.8), we obtain for $(j, a) \in E_{\omega}^{+}$

$$|\varepsilon^{\dagger}_{j \rightarrow a} - \varepsilon^{(\omega)}_{j \rightarrow a}| \leq \sup_{(\varepsilon^{\dagger}, a'' \in \partial(j, a) \in V_{j \rightarrow a}} \frac{P_{j \rightarrow a}(\varepsilon)}{1 + P_{j \rightarrow a}(\varepsilon))^{2}} \sum_{a' \in \partial(j, a)} \left|\frac{\varepsilon^{\dagger}_{a'_{j-}} - \varepsilon^{(\omega)}_{a'_{j-}}}{\varepsilon_{a'_{j-}}}\right| \varepsilon^{\dagger}_{a'_{j-}} - \varepsilon^{(\omega)}_{a'_{j-}} (B.12)$$

We observe that

$$\sup_{(\varepsilon^{\dagger}, a'' \in \partial(j, a) \in V_{j \rightarrow a}} \frac{P_{j \rightarrow a}(\varepsilon)}{1 + P_{j \rightarrow a}(\varepsilon))^{2}} \leq \frac{1}{2}. \quad (B.13)$$

Moreover, if $\epsilon_{j}$ is frozen, we have with Lemma B.8

$$\sup_{(\varepsilon^{\dagger}, a'' \in \partial(j, a) \in V_{j \rightarrow a}} \frac{P_{j \rightarrow a}(\varepsilon)}{1 + P_{j \rightarrow a}(\varepsilon))^{2}} \leq \exp(-99\beta). \quad (B.13)$$

We further observe that, for any $(a', j) \in E_{\omega}^{+}$

$$\sup_{(\varepsilon^{\dagger}, a'' \in \partial(j, a) \in V_{j \rightarrow a}} \left|\frac{\varepsilon^{\dagger}_{a'_{j-}} - \varepsilon^{(\omega)}_{a'_{j-}}}{\varepsilon_{a'_{j-}}}\right| \leq \exp(\beta) \sum_{\varepsilon_{j \rightarrow a} \neq \text{frozen}} |\varepsilon^{\dagger}_{j \rightarrow a} - \varepsilon^{(\omega)}_{j \rightarrow a}|. \quad (B.14)$$

Moreover, if $a'$ is cold (w.r.t. $\epsilon$) we have

$$\sup_{(\varepsilon^{\dagger}, a'' \in \partial(j, a) \in V_{j \rightarrow a}} \left|\frac{\varepsilon^{\dagger}_{a'_{j-}} - \varepsilon^{(\omega)}_{a'_{j-}}}{\varepsilon_{a'_{j-}}}\right| \leq \exp(-99\beta) \sum_{\varepsilon_{j \rightarrow a} \neq \text{frozen}} |\varepsilon^{\dagger}_{j \rightarrow a} - \varepsilon^{(\omega)}_{j \rightarrow a}| + \exp(\beta) \sum_{\varepsilon_{j \rightarrow a} \neq \text{frozen}} |\varepsilon^{\dagger}_{j \rightarrow a} - \varepsilon^{(\omega)}_{j \rightarrow a}|. \quad (B.15)$$

Because $T_{\omega}$ is $\omega$-good, for all $i_{\omega} \in \partial T_{\omega}$, the path $[i_{0} \rightarrow i_{\omega}]$ contains at least $\lfloor 0.4\omega \rfloor$ variables $j$ such that $\epsilon_{j}$ is frozen or the clause $a$ such that $(j, a) \in E_{\omega}^{+}$ is cold. Using this remark while combining Eq. (B.11-B.15) and iterating these equations, we obtain

$$|\varepsilon^{\dagger}_{i_{0} \rightarrow a} - \varepsilon^{(\omega)}_{i_{0} \rightarrow a}| \leq \sum_{j \in V \text{ dist}(i, j) = 2\omega} \exp(-30\beta \omega + 6\beta \omega) \leq \exp(-20\beta \omega) \leq 2^{-k\omega}.$$
Corollary B.10. Let $\pi$ be a skewed fixed point of $\mathcal{G}_{k,d,\beta}$. Then $\pi^{(\omega)}$ weakly converges towards $\pi$. In particular $\mathcal{G}_{k,d,\beta}$ has at most one skewed fixed point.

Proof. Let $\pi_+, \pi_-$ be as associated to $\pi$ by Eq. B.1-B.2. For $\omega \geq 0$, we consider the probability distribution $\mathbb{P}$ over $\mathcal{F}_\omega \times (0,1)^{d_\omega}$ induced by the following process.

- Draw $T \in \mathcal{F}_\omega$ at random from the distribution $p^{(\omega)}_{k,d,\beta}$.
- Independently, draw for each $(i,a) \in \partial E^+_\omega$ with $i \in \partial_1(T)$ (resp. with $i \in \partial_{-1}(T)$) a random variable $\eta_{i-a}$ from the distribution $\pi_+$ (resp. $\pi_-$), and set $\varepsilon_{i-a} = \frac{1}{2} - (-1)^b_k \eta_{i-a} - \frac{\omega}{2}$ and $\varepsilon_{\partial T} = (\varepsilon_{i-a})_{(i,a) \in \partial E^+_\omega}$.

It follows from the fact that $\pi$ is skewed that $\mathbb{P}$ satisfies the hypotheses of Corollary B.7 Using Lemma B.9 we have

$$\mathbb{E} \left[ |\varepsilon_{i_0,a}(T, \varepsilon_{\partial T}) - \varepsilon_{i_0,a}^{(\omega)}(T) | \right] \leq o_{\omega}(1) + \mathbb{P} \left[ (T, \varepsilon_{\partial T}) \text{ is not $\omega$-good} \right].$$

By combining Corollary B.7 with the fact that $\pi$ is skewed, we see that the last probability is bounded by $o_{\omega}(1)$. It follows from Lemma B.5, Eq. B.10 and Eq. B.7 that $\pi^{(\omega)}$ weakly converges towards $\pi$, and that $\mathcal{G}_{k,d,\beta}$ has at most one fixed point.

Without further work, we can also prove the existence of the fixed point.

Corollary B.11. The operator $\mathcal{G}_{k,d,\beta}$ admits a skewed fixed point.

Proof. For $T \in \mathcal{F}_{\omega+1}$ fixed, we denote by $T^{(\omega)}$ the subtree of $T$ obtained by restricting $T$ to variables and clauses at distance at most $2\omega$ from the root, and by $\varepsilon_{\partial T}^{(1)} = (\varepsilon_{i-a})_{(i,a) \in E^+_\omega, i \in \partial V_\omega}$ the messages defined by $\varepsilon_{i-a}^{(1)} = \varepsilon_{i-a}^{(\omega+1)}(T)$, for $(i,a) \in E^+_\omega$ with $i \in \partial V_\omega$. We have

$$\mathbb{E} \left[ |\varepsilon_{i_0,a}^{(\omega+1)}(T) - \varepsilon_{i_0,a}^{(\omega)}(T^{(\omega)}) | \right] = \mathbb{E} \left[ |\varepsilon_{i_0,a}(T, \varepsilon_{\partial T}^{(1)}) - \varepsilon_{i_0,a}^{(\omega)}(T) | \right] \leq o_{\omega}(1) + \mathbb{P} \left[ (T, \varepsilon_{\partial T}^{(1)}) \text{ is not $\omega$-good} \right].$$

It is easily seen that $\mathbb{P}$ satisfies the hypotheses of Corollary B.7 and the last probability is therefore bounded by $o_{\omega}(1)$. It follows that the sequence $(\pi^{(\omega)})_{\omega \geq 0}$ is a Cauchy sequence for the weak convergence in the Banach space of skewed probability distributions over $(0,1)$, proving the corollary.

Proof of Proposition A.7. The proposition follows by combining Corollary B.10 and Corollary B.11.

Proof of Proposition A.7. With the notations of Section 2.4 and Lemma B.1 we compute

$$\mathbb{E} \left[ z_1(\bar{\eta}_1, \ldots, \bar{\eta}_d) \right] = \ln \left( 2q^{d/2} (1 - q)^{d/2} \right),$$

$$\mathbb{E} \left[ z_2(\eta_1, \ldots, \eta_k) \right] = \ln \left( 1 - c_\beta q^k \right),$$

$$\mathbb{E} \left[ z_3(\eta_1, \bar{\eta}_1) \right] = \ln \left( 2q(1 - q) \right).$$

Thereby we have

$$\mathcal{B}(k, d, \beta) = \ln 2 + d \ln \left( 1 - c_\beta q^k \right) - \frac{d}{2} \ln \left( \frac{1}{2q} \right) - \frac{d}{2} \ln \left( \frac{1}{2(1 - q)} \right).$$

The proposition then follows from a standard computation of $\frac{1}{n} \ln \mathbb{E} \left[ Z_{\Phi}(\beta) \right]$. □
B.3. **Finite ω approximations of \( \mathcal{B}(k, d, \beta) \).** We finally present a simple approximation of \( \mathcal{B}(k, d, \beta) \) that will be useful in the following. Let \( GW(k, d, \beta) \) denote the Galton-Watson process introduced in Section 2.4 and considered up to now. Let \( GW'(k, d, \beta) \) be the multi-type random process defined in the following way. The individuals are either variables or clauses, with types \( + \) or \( - \). The root is a variable of type \( + \) with probability \( q \) and a variable of type \( - \) with probability \( 1 - q \), and it has for offspring \( d/2 \) clauses of type \( + \) and \( d/2 \) clauses of type \( - \). Each of these clauses then generates a random process as in \( GW(k, d, \beta) \). Let \( \mathcal{T}_\omega \) be the set of trees that this process generates at depth \( 2\omega \), and for \( T \in \mathcal{T}_\omega \) let \( \tilde{p}^{(\omega)}_{k,d,\beta}(T) \) denote the probability that the process generates \( T \) at depth \( 2\omega \).

Let also \( \mathcal{F}_\omega \) denote the set of trees that the random process \( GW(k, d, \beta) \) generates at depth \( 2\omega - 1 \), starting from a clause, and for \( \tilde{T}_1, \ldots, \tilde{T}_d \in \mathcal{F}_\omega \) let \( \tilde{p}^{(\omega)}_{k,d,\beta}(\tilde{T}_1, \ldots, \tilde{T}_d) \) denote the probability that the neighborhood of the root in equal to \( (\tilde{T}_1, \ldots, \tilde{T}_d) \) under the process \( GW'(k, d, \beta) \). Denoting by \( e_1 \) the first edge exiting the root of the random process \( GW'(k, d, \beta) \), let for \( T \in \mathcal{T}_\omega \) and \( \tilde{T} \in \mathcal{F}_\omega \) let \( \tilde{p}(T, \tilde{T}) \) be the probability that the \( 2\omega \)-neighborhood of this edge is formed of the trees \( T, \tilde{T} \). Similarly, denoting by \( a_1 \) the first clause connected to the root of the random process \( GW'(k, d, \beta) \), let for \( T_1, \ldots, T_k \in \mathcal{T}_\omega \) let \( \tilde{p}^{(\omega)}_{k,d,\beta}(T_1, \ldots, T_k) \) be the probability that the neighborhood of \( a_1 \) is equal to \( (T_1, \ldots, T_k) \) under the process \( GW'(k, d, \beta) \). Finally, let \( \bar{\mathcal{T}}_\omega \) denote the set of trees seen at depth \( 2\omega - 1 \) when rerooting \( GW(k, d, \beta) \) at a clause adjacent to its root.

**Lemma B.12.** We have (with the notations of Section 1.3 and Lemma B.5)

\[
\mathcal{B}(k, d, \beta) = \sum_{(\tilde{T}_1, \ldots, \tilde{T}_d) \in \bar{\mathcal{T}}_d} \tilde{p}^{(\omega)}_{k,d,\beta}(\tilde{T}_1, \ldots, \tilde{T}_d) \int_{(0,1)^d} \ln \left[ z_1(\eta_{i_1 \rightarrow i_0}(\tilde{T}_1, \eta_{\partial \tilde{T}_1}), \ldots, \eta_{i_d \rightarrow i_0}(\tilde{T}_d, \eta_{\partial \tilde{T}_d})) \right]
\]

\[
\otimes \prod_{j=1}^d \left( \int_{(i,a) \in E^+_{\partial \tilde{T}_j}} d\pi_+((\eta_{\partial \tilde{T}_j})) \otimes \int_{(i,a) \in E^-_{\partial \tilde{T}_j}} d\pi_-((\eta_{\partial \tilde{T}_j})) \right)
\]

\[
+ \sum_{(T_1, \ldots, T_k) \in \mathcal{T}_k} \tilde{p}^{(\omega)}_{k,d,\beta}(T_1, \ldots, T_k) \int_{(0,1)^k} \ln \left[ z_2(\eta_{i_0 \rightarrow a_0}(T_1, \eta_{\partial T_1}), \ldots, \eta_{i_k \rightarrow a_0}(T_k, \eta_{\partial T_k})) \right]
\]

\[
\otimes \prod_{j=1}^k \left( \int_{(i,a) \in E^+_{\partial T_j}} d\pi_+((\eta_{\partial T_j})) \otimes \int_{(i,a) \in E^-_{\partial T_j}} d\pi_-((\eta_{\partial T_j})) \right)
\]

\[
- \sum_{T \in \mathcal{T}_\omega, \tilde{T} \in \mathcal{F}_\omega} \tilde{p}^{(\omega)}_{k,d,\beta}(T, \tilde{T}) \int_{(0,1)^2} \ln \left[ z_3(\eta_{i_0 \rightarrow a_0}(T, \eta_{\partial T}), \eta_{i_1 \rightarrow i_0}(\tilde{T}, \eta_{\partial \tilde{T}})) \right]
\]

\[
\otimes \prod_{(i,a) \in E^+_{\partial T}} d\pi_+((\eta_{\partial T})) \otimes \prod_{(i,a) \in E^-_{\partial T}} d\pi_-((\eta_{\partial T})) + o_\omega(1).
\]

**Proof.** The proof is obtained by writing the expectation values in the definition of \( \mathcal{B}(k, d, \beta) \) explicitly in terms of \( (\pi^{*}_{k,d,\beta})^+, (\pi^{*}_{k,d,\beta})^- \) (resp. \( (\tilde{\pi}^{*}_{k,d,\beta})^+, (\tilde{\pi}^{*}_{k,d,\beta})^- \)) and following steps similar to the one of the proof of Lemma B.5. \( \square \)

We define

\[
\mathcal{B}^{(\omega)}(k, d, \beta) = \sum_{(\tilde{T}_1, \ldots, \tilde{T}_d) \in \bar{\mathcal{T}}_d} \tilde{p}^{(\omega)}_{k,d,\beta}(\tilde{T}_1, \ldots, \tilde{T}_d) \ln \left[ z_1(\hat{\eta}_1^{(\omega)}(\tilde{T}_1), \ldots, \hat{\eta}_d^{(\omega)}(\tilde{T}_d)) \right]
\]

\[
+ \sum_{(T_1, \ldots, T_k) \in \mathcal{T}_k} \tilde{p}^{(\omega)}_{k,d,\beta}(T_1, \ldots, T_k) \ln \left[ z_2(\eta_1^{(\omega)}(T_1), \ldots, \eta_d^{(\omega)}(T_k)) \right]
\]

\[
- \sum_{T \in \mathcal{T}_\omega, \tilde{T} \in \mathcal{F}_\omega} \tilde{p}^{(\omega)}_{k,d,\beta}(T, \tilde{T}) \ln \left[ z_3(\eta^{(\omega)}(T), \hat{\eta}^{(\omega)}(\tilde{T})) \right].
\]
Proposition B.13. We have
\[ B(k, d, \beta) = B^{(0)}(k, d, \beta) + o_\omega(1). \]

Proof. The result easily follows from the weak convergence of \( \pi^{(0)} \) toward \( \pi^{*}_{k, d, \beta} \). \( \square \)

Finally, let us prove here a few results that will be useful in the following. In Section 2.5 it will be convenient to have a third parametrization of the messages \((\eta, \tilde{\eta})\), that we present here. Given \( T \in \mathcal{T}_\omega \) and messages \((\eta, \tilde{\eta})\) on \( T \), we define messages

\[
(v, \tilde{v}) = ((v_{-a})_{i, a \in E_\omega^+}, (\tilde{v}_{a-i})_{(a, i) \in E_\omega^+}) \in \mathcal{P}([-1, 1])^{E_\omega^+}
\]

by letting, for \((i, a) \in E_\omega^+\) (resp. \((b, i) \in E_\omega^+\) and \(x \in [-1, 1]\)

\[
v_{-a}(x) = \frac{1}{2} - xb_{a,i} \left( \frac{1}{2} - \eta_{-a} \right), \quad \tilde{v}_{a-i}(x) = \frac{1}{2} - xb_{b,i} \left( \frac{1}{2} - \tilde{\eta}_{b-i} \right)
\]

We observe that \((\eta, \tilde{\eta})\) satisfies the Belief Propagation equations on \( T \) if and only if the associated messages \((v, \tilde{v})\) satisfy, for \((i, a) \in E_\omega^+\) (resp. \((b, i) \in E_\omega^+\) and \(x \in [-1, 1]\)

\[
v_{-a}(x) = \frac{1}{2} - xb_{a,i} \left( \frac{1}{2} - \eta_{-a} \right), \quad \tilde{v}_{a-i}(x) = \frac{1}{2} - xb_{b,i} \left( \frac{1}{2} - \tilde{\eta}_{b-i} \right)
\]

We also introduce, for \( T \in \mathcal{T}_\omega \) and \( x \in [-1, 1]\)

\[
v_{-a}(x, T) = \frac{1}{2} - xb_{a,i} \left( \frac{1}{2} - \eta_{-a} \right), \quad \tilde{v}_{a-i}(x, T) = \frac{1}{2} - xb_{b,i} \left( \frac{1}{2} - \tilde{\eta}_{b-i} \right)
\]

We also define, for \( T \in \mathcal{T}_\omega \) and \( \hat{\mathcal{T}} \in \mathcal{T}_\omega^d \), \( \mu^{(0)}(\cdot, T) \) and \( \tilde{\mu}^{(0)}(\cdot, \hat{\mathcal{T}}) \) by the natural counterparts to formula 2.5. The following is then a direct consequence of Proposition B.13.

Corollary B.14. We have

\[
B(k, d, \beta) = \sum_{(T_1, \ldots, T_k) \in \mathcal{T}_\omega^k} \tilde{p}^{(0)}_{k, d, \beta}(T_1, \ldots, T_k) \ln \left[ \sum_{(x_{a \in [-1, 1]^k})} \psi_{a, \beta}(x_a) \prod_{i \in a} \psi^{(0)}(x_i, T_i) \right]
\]

\[
+ \sum_{(\hat{T}_1, \ldots, \hat{T}_d) \in \mathcal{T}_\omega^d} \tilde{p}^{(0)}_{k, d, \beta}(\hat{T}_1, \ldots, \hat{T}_d) \ln \left[ \sum_{x_{a \in [-1, 1]^d}} \prod_{i \in a} \tilde{\psi}^{(0)}(x_i, \hat{T}_a) \right]
\]

\[
- \sum_{T \in \mathcal{T}_\omega, \hat{T} \in \mathcal{T}_\omega^d} \tilde{p}^{(0)}_{k, d, \beta}(T, \hat{T}) \ln \left[ \sum_{x \in [-1, 1]} \psi^{(0)}(x, T) \tilde{\psi}^{(0)}(x, \hat{T}) \right] + o_\omega(1).
\]

Finally, we shall need the following estimate in Section D.

Remark B.15. We have

\[
\frac{1}{n} \ln H(\mu^{(0)}) \leq 2^{-0.999k} \quad \text{and} \quad \frac{1}{n} \ln B^{(0)}(k, d, \beta) \geq -2^{-0.999k}.
\]

Proof. Both estimates are obtained by a careful study of the marginals and messages \( \mu^{(0)}, \psi^{(0)} \) and \( \tilde{\psi}^{(0)} \). For the first one, we observe that the fraction of trees \( T \in \mathcal{T}_\omega \) such that \( \mu^{(0)}(-1, T) \geq 1 - \exp(-4\beta) \) is at most \( 2^{-0.9999k} \) (by simple recurrence over \( \omega \)). The second estimate is obtained in a similar manner by carefully studying the various contributions to \( B^{(0)}(k, d, \beta) \). \( \square \)
B.4. **Proof of Proposition B.6.**

Proof: Let \( \mathbb{P} \) satisfying the hypotheses of Proposition B.6 and \( (T, \epsilon_{\partial T}) \in \mathcal{T}_\omega \times \{ \text{frozen}, \text{non frozen} \}^d \) be fixed. In order to study the fraction of variables in a path \( l_0 \rightarrow l_\omega \) that are not cold, we need to understand the probability that a random variable node of the tree is not frozen. More precisely, we need the following. Given an edge \((a', i) \in E^{+}_\omega\), let \( T_{a'}\) be the subtree of \( T \backslash \{ i \} \) rooted at \( a'\). For \((a', i) \in E^{+}_\omega\), we say that \( i \) is strongly frozen with respect to \( a' \) if there exists no boundary condition \( \epsilon'_{\partial T'} \) over \( \partial T \) such that the following is true.

- \( i \) is not frozen in \((T, \epsilon'_{\partial T})\),
- \( \forall l \in \partial T \backslash \partial T_{a'}, \epsilon'_i = \epsilon_l \).

Observe that strongly frozen variables are also frozen. Let \( V^{(t)}_t \) be the set of frozen variables at distance \( t \) from \( \partial V_\omega \). If a variable node \( i \) at distance \( t > 0 \) from \( \partial V_\omega \) is not strongly frozen with respect to \( a' \) then one the following is true. Let \( a \) be the clause such that \((i, a) \in E^{+}_\omega\).

\[
\begin{align*}
(a) & \quad \partial_{1,0}(i, a) \leq k^{7/8}, \\
(b) & \quad \partial_{-1,0}(i, a) \geq 2, \\
(c) & \quad \text{there is } 1 \leq l \leq k \text{ such that } |\partial_{-1,l}(i, a)|| \geq k^{l/3}/l! - 1, \\
(d) & \quad |\{ b \in \partial_{1,0}(i, a), \partial(b, i) \not\subset V^{(t-1)}_t \}| \geq k^{3/4} - 1, \\
(e) & \quad |\{ b \in \partial_{-1,i}, |\partial_{-1,i} b| \leq k, |\partial(b, i) \setminus V^{(t-1)}_t| \geq |\partial_{1} b|/4 \}| \geq k^{3/4} - 1.
\end{align*}
\]

Let \( p_{i,a'} \) be the probability that \( i \) is not strongly frozen with respect to \( a' \) when the pair \((T, \epsilon_{\partial T})\) is drawn from the distribution \( P \). Obviously, this probability only depends on the distance from \( i \) to \( \partial V_\omega \), and for \( i \) at distance \( t \) from the boundary we denote it by \( p^{(t)} \). By definition of the random process GW\((k, d, \beta)\), (a), (b) and (c) both happen with probability at most \( 2^{-0.95k} \). We shall prove by induction on \( t \in \{1, \ldots, \omega\} \) that

\[
\forall t \in \{1, \ldots, \omega\}, p^{(t)} \leq 2^{-0.95 k}.
\]

For \( t = 0 \), we have \( p^{(t)} \leq 2^{-0.95 k} \) by definition of \( P \). We assume more generically that \( p^{(t-1)} \leq 2^{-0.95 k} \) for \( (t-1) \in \{0, \ldots, \omega - 1\} \). Let \( i \) be a variable node of \( T \) at distance \( t \) from \( \partial V_\omega \). The probability that a given clause \( a'' \in \partial_{1,0}(i, a) \) contains at least one not strongly frozen variable different from \( i \) is \( k p^{(t-1)} + \tilde{O}_k(4^{-k}) \), and therefore the probability that (d) happens is at most \( 2^{-1.5k} \). Similarly, (e) happen with probability at most \( 2^{-1.9k} \). Therefore we obtain

\[
p^{(t)} \leq 2^{-0.95k} + 3.2^{-1.5k} \leq 2^{-0.9k}.
\]

Let \((j, a) \in E^{+}_\omega \). We say that \( a \) is strongly cold with respect to \( j \) if \( |\partial_{1}(a, i) \setminus \{ j \} \cap V_{i} | \geq 1 \). Let \( j \in V_\omega \) be fixed and \( a, a' \) be such that \((a', j), (j, a) \in E^{+}_\omega \). We say that \( j \) is strongly cold with respect to \( a' \) if it is strongly frozen with respect to \( a' \) or if \( a \) is strongly cold with respect to \( j \). Accordingly, for \( j \) not to be strongly cold with respect to \( a' \), one of the following must hold.

- \( j \) is not strongly frozen with respect to \( a' \),
- \( a \) is not strongly cold with respect to \( j \).

Let us now compute the probability \( q^{(t)} \) that \( j \in V_\omega \) at distance \( t > 1 \) from \( \partial V_\omega \) is not strongly cold with respect to \( a' \) (with \((a', j) \in E^{+}_\omega \)) when the pair \((T, \epsilon)\) is drawn from the distribution \( P \). The probability of event (a) is bounded by \( 2^{-0.9k} \) by our previous estimate. The event (b) is independent of the event (a) and requires that one of the following is true.

\[
\begin{align*}
(b1) & \quad \partial_{1} a \subset \{i, j\}, \\
(b2) & \quad |\partial_{1} a \setminus \{i, j\} = \{l\} \text{ and } l \text{ is not strongly frozen} \\
(b3) & \quad |\partial_{1} a \setminus \{i, j\}| \geq 2 \text{ and } \partial_{1} a \setminus \{i, j\} \text{ contains only variables that are not strongly frozen}
\end{align*}
\]

By our previous estimate (b2) and (b3) happen with probability at most \( 2^{-1.5k} \) while (b1) happens with probability at most \( 2^{-0.9k} \). It follows that \( q^{(t)} \leq 2^{-0.89k} p^{(t)} 2^{-1.7k} \).
Let \( i_\omega \in \partial V_\omega \) be fixed. We denote sequence of variables and clauses on the path \([i_0 \rightarrow i_\omega]\) as \((i_0, a_1, i_1, \ldots, a_\omega, i_\omega)\). For the path \([i_0 \rightarrow i_\omega]\) not to be cold, there must be at least \([0.6\omega]\) variables along this path that are not strongly cold. Moreover, for \( j, j' \) along the path, the probability that \( j \) and \( j' \) are strongly cold are independent (by construction). Therefore we obtain

\[
P [[i_0 \rightarrow i_\omega] \text{ not cold}] \leq \sum_{l \geq [0.6\omega]} \sum_{0 \leq j_1 \leq \cdots \leq j_\omega} P [i_{j_1} \text{ not strongly cold and } i_{j_2} \text{ not strongly cold and} \]

\[
\ldots \text{ and } i_{j_\omega} \text{ not strongly cold}] 
\leq \sum_{l \geq [0.6\omega]} \sum_{0 \leq j_1 \leq \cdots \leq j_\omega} \prod_{k=1}^{l} q^{(j_k)} 
\leq 2^{0.6\omega} (2^{-1.7k})^{[0.6\omega]} \leq 2^{-1.02k\omega}.
\]

Consequently, we obtain with the union bound

\[
P [(T, \varepsilon_{\partial T}) \text{ is not } \omega\text{-frozen}] \leq \sum_{i_\omega \in \partial V_\omega} P [[i_0 \rightarrow i_\omega] \text{ not cold}] \leq 2^{1.01k\omega}2^{-1.02k\omega} = o_\omega(1).
\]

\[\square\]

**APPENDIX C. UPPER BOUND ON THE SECOND MOMENT**

In this section we prove Proposition[A.3] and Proposition[A.6]. We assume that \( d \in [d_{-}(k), d_{k-SAT}] \) and that \( \beta \geq \beta_{-}(k) \).

C.1. **Outline of the proof.** In order to prove Proposition[A.3] and Proposition[A.6] we will identify, for every \( \omega > 0 \), a set of formulas \( \mathcal{E}_{n,k,d,\beta,\omega} \) such that

- \( \Phi \in \mathcal{E}_{n,k,d,\beta,\omega} \) w.h.p.,
- \( \Phi \in \mathcal{E}_{n,k,d,\beta,\omega} \) w.h.p.,
- for \( \Phi \in \mathcal{E}_{n,k,d,\beta,\omega} \) we have \( \sup_{\mu \in \mathcal{M}_{\omega}^{(2)} B_{\Phi,\beta}^{(2)}(\mu) \leq 2\mathcal{B}(k,d,\beta) + \omega_\omega(1) \).

We will construct this set explicitly. First, we will need some information on the local structure of the formula \( \Phi \). For a variable node \( i \) and \( t \geq 0 \) we let \( T_{i}^{(t)} \) be the \( 2t \)-neighborhood of \( i \) (which is not necessarily a tree). For \( T \in \mathcal{T}_{\omega+1} \) (defined in Section[B.3] we define the empirical density of \( T \) by

\[
\rho_{\Phi}(T) = \frac{1}{n} \sum_{i \in [n]} 1_{P_{\Phi,\omega+1}^{i} = T}.
\]

We shall say that a random regular \( k \)-SAT formula \( \Phi \) satisfies property P1 if and only if the following is true.

**P1:**

\[
\forall T \in \mathcal{T}_{\omega+1}, \rho_{\Phi}(T) \sim P_{k,d,\beta}^{\omega+1}(T).
\]

In order to proceed further, we begin by introducing a few additional notations, similar to the ones that we used in Section[B.2] Let \( \Phi \) be fixed and \( V \) denote its set of vertices, \( F \) denote its set of edges and \( E \) denote its set of (undirected) edges. For \( i \in V \) (resp. \( a \in F \) ), we let

\[
\partial_i i = \{ a \in \partial i, b_{a,i} = -1 \}, \quad \partial_{-1} i = \{ a \in \partial i, b_{a,i} = 1 \},
\]

\[
\partial_1 a = \{ i \in \partial a, b_{a,i} = -1 \}, \quad \partial_{-1} a = \{ i \in \partial a, b_{a,i} = 1 \},
\]

\[
\partial_i i = \{ a \in \partial_i, |j \in \partial a \setminus \{i\}, b_{a,j} = 1| = l \}.
\]

We also introduce \( \partial_{1,i} = \partial_i \cap \partial_{1}, \partial_{-1,i} = \partial_i \cap \partial_{-1} \), and for \( a_0 \in \partial i \) fixed, \( \partial_1(i,a_0) = \partial_1 i \setminus \{a_0\}, \partial_{-1}(i,a_0) = \partial_{-1} i \setminus \{a_0\}, \partial_{1,i}(i,a_0) = \partial_{1,i} i \setminus \{a_0\} \).

Our aim will be to identify a large set \( V_{good} \subset V \) of vertices whose value under a typical assignment in the cluster is unlikely to be very far from the planted one \( (1,1) \). A first candidate for vertices whose
marginal may go wrong are those which do not belong to the 1-core of $\Phi$. Yet, we are not guaranteed that vertices in the core have marginals sufficiently close to $\mu^{(0)}$. For instance, if the marginals of most of the neighbors of a given vertex $i \in \text{Core}_1(\Phi)$ went astray, there would be no reason for $i$’s marginal not to go astray itself. However, we see that the vertices in the core whose marginals are not what we think they should be must clump together. We shall therefore recall the notion of sticky sets that appeared in Section B generalising it slightly. We say that a set $S \subset V$ is $\lambda$-sticky if and only if for all $i \in S$, one of the following conditions holds true.

**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:** there are at least $\lambda k^{3/4}$ clauses $a \in \partial_{-1} x$ such that $|\partial_{-1} a| < k$ and $|\partial_{-1} a \cap S| \geq |\partial_{-1} a|/4$.

Observe that if $S_1, S_2 \subset V$ are $\lambda$-sticky, then $S_1 \cup S_2$ is also a $\lambda$-sticky set. Let for a formula $\Phi$, $S_\lambda(\Phi)$ denote the union of all the $\lambda$-sticky sets $S$ of $\Phi$ of size $|S| \leq 2^{1-k/10}n$. Observe that if $\lambda < \lambda'$ then $S_{\lambda'}(\Phi) \subset S_{\lambda}(\Phi)$. We say that $\Phi$ satisfies property $P_2$ and only if

$$P_2:\quad \text{Core}_{1/2}(\Phi) \setminus S_{1/2}(\Phi) \geq \left(1 - 2^{-0.95k}\right)n.$$  

We say that a variable node $i \in V$ is safe if and only if $i \in \text{Core}_1(\Phi) \setminus S_1(\Phi)$. Let $\omega \geq 0$ be fixed. We say that a variable node $i \in V$ is $\omega$-safe if the $2\omega$-neighborhood of $i$ is tree like and the pair $(T, c_{\delta T}) \in \mathcal{T}_\omega \times \{\text{frozen}, \text{non frozen}\}$, with $(c_{\delta T})_{i \rightarrow a} = \text{frozen}$ if and only if $i$ is safe, is $\omega$-frozen. Let $V_{\text{safe,} \omega}$ denote the set of $\omega$-safe variables and let us say that $\Phi$ satisfies property $P_3$ if and only if

$$P_3:\quad |V \setminus V_{\text{safe,} \omega}| = o_w(n).$$

Let $\mathcal{S}_{n,k,d,\beta,\omega}$ be the set of formulas $\Phi$ such that $P_1$, $P_2$ and $P_3$ hold.

**Proposition C.1.** W.h.p. $\Phi \in \mathcal{S}_{n,k,d,\beta,\omega}$.

**Proposition C.2.** W.h.p. $\Phi \in \mathcal{S}_{n,k,d,\beta,\omega}$.

Because of their more combinatorial nature, we defer the proof of these propositions to Section C.2.

Let $\Phi \in \mathcal{S}_{n,k,d,\beta,\omega}$ be fixed. It will be convenient to consider not only marginals over the variables of a formula, but also messages between variables and clauses. We recall that $\mathcal{P}([-1, 1]^2)$ denotes the set of probability distributions over $[-1, 1]^2$. For a fixed formula $\Phi$, we let

$$\mathcal{N}_{\Phi}^{(2)} = \{ (\mu_i, v_{i \rightarrow a}, \bar{v}_{a \rightarrow i})_{i \in V, a \in \partial_i} \in \mathcal{P}([-1, 1]^2)^{n+2km} \}.$$  

We denote by $(\mu, v, \bar{v}) = (\mu_i, v_{i \rightarrow a}, \bar{v}_{a \rightarrow i})_{i \in V, a \in \partial i}$. We define an iteration operator $\mathcal{F}_{\Phi, \beta} : \mathcal{N}_{\Phi}^{(2)} \rightarrow \mathcal{N}_{\Phi}^{(2)}$ by letting $\mathcal{F}_{\Phi, \beta}(\mu, v, \bar{v}) = (\mu', v', \bar{v}')$ where for $i_0 \in V$, $a_0 \in \partial i$ and $(x, y) \in [-1, 1]^2$

$$\mu'_{i_0}(x, y) = \frac{\sum_{(x', y') \in [-1, 1]^2} \psi_{a, \beta}(x', y') \prod_{i \in \partial(a, i_0)} v_{i \rightarrow a}(x_i', y_i') \mathbf{1}(x_{i_0}' = x, y_{i_0}' = y)}{\sum_{(x', y') \in [-1, 1]^2} \psi_{a, \beta}(x', y') \prod_{i \in \partial(a, i_0)} v_{i \rightarrow a}(x_i', y_i')}$$  

(C.1)

$$v'_{a \rightarrow i_0}(x, y) = \frac{\prod_{a \in \partial(i_0)} \bar{v}'_{a \rightarrow i}(x, y)}{\sum_{(x', y') \in [-1, 1]^2} \prod_{a \in \partial(i_0)} \bar{v}'_{a \rightarrow i}(x', y')}$$  

(C.2)

$$\bar{v}'_{i_0 \rightarrow a_0}(x, y) = \frac{\prod_{a \in \partial(i_0, a_0)} v'_{a \rightarrow i}(x, y)}{\sum_{(x', y') \in [-1, 1]^2} \prod_{a \in \partial(i_0, a_0)} v'_{a \rightarrow i}(x', y')}.$$  

(C.3)

We say that $(\mu, v, \bar{v}) \in \mathcal{N}_{\Phi}^{(2)}$ satisfies the paired Belief Propagation equations iff it satisfies $(\mu, v, \bar{v}) = \mathcal{F}_{\Phi, \beta}(\mu, v, \bar{v})$. We will prove the following propositions in Section C.2.
Proposition C.3. For $\Phi \in \mathcal{E}_{n,k,d,\beta,\omega}$ the following is true. Any global maximizer $\mu^*$ of $B_{\Phi,\beta}^{(2)}$ over $M_0^{(2)}$ is a local maximum and there are $\nu^*$ and $\tilde{\nu}^*$ such that $(\mu^*, \nu^*, \tilde{\nu}^*)$ satisfies the paired Belief Propagation equations. Moreover, we have

$$B_{\Phi,\beta}^{(2)}(\mu^*) = \frac{1}{n} \sum_{a \in E} \ln \left( \sum_{x_a \in \{-1,1\}^E} \psi_{a,\beta}(x_a) \prod_{i \in \partial a} \nu^*_{i-a}(x_i) \right) + \frac{1}{n} \sum_{i \in V} \ln \left( \sum_{x_i \in \{-1,1\}^E} \prod_{a \in \partial i} \tilde{\nu}^*_{a-i}(x_i) \right) - \frac{1}{n} \sum_{[i,a] \in E} \ln \left( \sum_{x_{i-a} \in \{-1,1\}^E} \nu^*_{i-a}(x) \tilde{\nu}^*_{a-i}(x) \right).$$

Let $(\mu^{(0)}, \nu^{(0)}, \tilde{\nu}^{(0)})$ be defined by (for any $i \in V$, $a \in \partial i$ and $(x, y) \in \{-1,1\}^2$)

$$\mu^{(0)}_i(x, y) = \nu^{(0)}_{i-a}(x, y) = \tilde{\nu}^{(0)}_{a-i}(x, y) = (1 - \exp(-4\beta d))^{1_{x=1} + 1_{y=1}} (\exp(-4\beta d))^{1_{x=-1} + 1_{y=-1}}.$$  

For $0 \leq t \leq \omega$, let $(\mu^{(t)}, \nu^{(t)}, \tilde{\nu}^{(t)}) = \mathcal{F}_t^{(2)}(\mu^{(0)}, \nu^{(0)}, \tilde{\nu}^{(0)})$. Note that $\mu^{(t)}_i$ depends only on the $t$-neighborhood of $i$. The following proposition shows that it gives, for most vertices, an increasingly accurate estimation of the maximizer $(\mu^*, \nu^*, \tilde{\nu}^*)$.

Proposition C.4. With the notations of Proposition C.3, all but $o_\omega(n)$ of the edges $[i, a] \in E$ satisfy

$$\sup_{x \in \{-1,1\}^2} |\nu^*_{i-a}(x) - \nu^{(0)}_{i-a}(x)| = o_\omega(1),$$  

$$\sup_{x \in \{-1,1\}^2} |\tilde{\nu}^*_{a-i}(x) - \tilde{\nu}^{(0)}_{a-i}(x)| = o_\omega(1).$$

To conclude the proof of Proposition A.3 and Proposition A.6 it only remains to observe that the finite depth approximation $(\mu^{(\omega)}, \nu^{(\omega)}, \tilde{\nu}^{(\omega)})$ allow to approximate the Bethe free-energy by a summation over finite trees, whose statistics match up (by condition P1) the one of the branching process that we studied in the previous section.

Proof of Proposition A.3 and Proposition A.6. Let $\omega \geq 0$ and $\Phi \in \mathcal{E}_{n,k,d,\beta,\omega}$ be fixed. Let $\mu^*$ be a maximizer of $B_{\Phi,\beta}^{(2)}$ over $M_0^{(2)}$ and $(\nu^*, \tilde{\nu}^*)$ be as given by Proposition C.3. Let $E_\omega \subset E$ be the set of edges $[i, a]$ such that (C.4) hold and the $\omega$ neighborhood of $i$ is a tree. Let $V_\omega = \{i \in V, \forall a \in \partial i, [i, a] \in E_\omega\}$ and $F_\omega = \{a \in F, \forall i \in \partial a, [i, a] \in E_\omega\}$. By Proposition C.4 and our definition of $\mathcal{E}_{n,k,d,\beta,\omega}$ we have

$$|V \setminus V_\omega| = o_\omega(n), \quad |F \setminus F_\omega| = o_\omega(n), \quad |E \setminus E_\omega| = o_\omega(n).$$

Moreover, it is easily seen from the Belief Propagation equations (C.1), (C.3) that for every edge $[i, a] \in E$ and $x \in \{-1,1\}^2$ we have

$$\nu^*_{i-a}(x) \in [\exp(-2d\beta), 1 - \exp(-2d\beta)], \quad \tilde{\nu}^*_{a-i}(x) \in [\exp(-2\beta), 1 - \exp(-2\beta)].$$

Therefore we can write

$$B_{\Phi,\beta}^{(2)}(\mu^*) = \frac{1}{n} \sum_{a \in F_\omega} \ln \left( \sum_{x_a \in \{-1,1\}^E} \psi_{a,\beta}(x_a) \prod_{i \in \partial a} \nu^*_{i-a}(x_i) \right) + \frac{1}{n} \sum_{i \in V_\omega} \ln \left( \sum_{x_i \in \{-1,1\}^E} \prod_{a \in \partial i} \tilde{\nu}^*_{a-i}(x_i) \right) - \frac{1}{n} \sum_{[i,a] \in E_\omega} \ln \left( \sum_{x_{i-a} \in \{-1,1\}^E} \nu^*_{i-a}(x) \tilde{\nu}^*_{a-i}(x) \right) + o_\omega(1).$$

It then follows from the uniform continuity of the logarithm (resp. the sum of $a$ numbers, the product of $a$ numbers) on compact subsets of $(0,1)$ (resp. compact subsets of $\mathbb{R}^a$, for $a \geq 0$) that the following is
true.
\[
B_{\Phi, \beta}^{(2)}(\mu^*) = \frac{1}{n} \sum_{a \in E_\omega} \ln \left[ \sum_{x_a \in \{-1, 1\}^k} \psi_{a, \beta}(x_a) \prod_{i \in \partial a} \gamma_{i-a}^{(\omega)}(x_i) \right] + \frac{1}{n} \sum_{i \in V_\omega} \ln \left[ \sum_{x_i \in \{-1, 1\}^l} \prod_{a \in \partial i} \gamma_{i-a}^{(\omega)}(x_i) \right] + o_\omega(1).
\]

The messages \( \gamma_{i-a}^{(\omega)} \) (resp. \( \gamma_{a-i}^{(\omega)} \)) only depend of the \( \omega \) neighborhood of \( i \) (resp. \( a \)) in \( \Phi \), and they are closely related to the messages that we defined in Section B.3. More precisely, let \( \gamma_{a-i}^{(\omega)} \) (resp. \( \gamma_{i-a}^{(\omega)} \)) denote the 2\( \omega \)–1-neighborhood of \( a \) (resp. 2\( \omega \)-neighborhood of \( i \)) in the graph with set of edges \( E \setminus \{i, a\} \).

With the notations of Section [B.2] and introducing
\[
\rho_{\Phi}(T_1, \ldots, T_k) = \frac{1}{n} \sum_{i \in V_\omega} \prod_{a \in \partial i} \mathbf{1}_{T_{a-i}=T_i} \quad \text{for } (T_1, \ldots, T_k) \in F_{\omega}^k,
\]
\[
\tilde{\rho}_{\Phi}(\tilde{T}_1, \ldots, \tilde{T}_d) = \frac{k}{d n \sum_{a \in \partial i} \mathbf{1}_{T_{a-i}=\tilde{T}_i}} \quad \text{for } (\tilde{T}_1, \ldots, \tilde{T}_d) \in F_{\omega}^d,
\]
\[
\tilde{\tilde{\rho}}_{\Phi}(T, \tilde{T}) = \frac{1}{d n \sum_{a \in \partial i} \mathbf{1}_{T_{a-i}=\tilde{T}_i} \mathbf{1}_{\tilde{T}_{a-i}=\tilde{T}_i}} \quad \text{for } (T, \tilde{T}) \in F_{\omega} \times F_{\omega},
\]
we can rewrite the Bethe free energy as (borrowing, again, the notations of Section [B.2])
\[
\frac{1}{2} B_{\Phi, \beta}^{(2)}(\mu^*) = \sum_{(T_1, \ldots, T_k) \in \tilde{F}_{\omega}^k} \rho_{\Phi}(T_1, \ldots, T_k) \ln \left[ \sum_{x_\omega \in \{-1, 1\}^k} \psi_{a, \beta}(x_a) \prod_{i \in \partial a} \gamma_{i-a}^{(\omega)}(x_i, T_i) \right] + \frac{d}{k} \sum_{(T_1, \ldots, T_d) \in \tilde{F}_{\omega}^d} \tilde{\rho}_{\Phi}(\tilde{T}_1, \ldots, \tilde{T}_d) \ln \left[ \sum_{x_\omega \in \{-1, 1\}^l} \prod_{a \in \partial i} \gamma_{a-i}^{(\omega)}(x_i, \tilde{T}_a) \right] - \frac{d}{k} \sum_{(T, \tilde{T}) \in \tilde{F}_{\omega} \times \tilde{F}_{\omega}} \tilde{\tilde{\rho}}_{\Phi}(T, \tilde{T}) \ln \left[ \sum_{x_\omega \in \{-1, 1\}^l} \gamma_{i-a}^{(\omega)}(x, T) \gamma_{a-i}^{(\omega)}(x, \tilde{T}) \right] + o_\omega(1).
\]

Because \( \Phi \in \mathcal{F}_{n,k,d,\beta,\omega} \), we have
\[
\rho_{\Phi}(T_1, \ldots, T_k) \sim \rho_{k,d,\beta}^{(\omega)}(T_1, \ldots, T_k), \quad \tilde{\rho}_{\Phi}(\tilde{T}_1, \ldots, \tilde{T}_d) \sim \tilde{\rho}_{k,d,\beta}^{(\omega)}(\tilde{T}_1, \ldots, \tilde{T}_d), \quad \tilde{\tilde{\rho}}_{\Phi}(T, \tilde{T}) \sim \tilde{\rho}_{k,d,\beta}^{(\omega)}(T, \tilde{T}).
\]
It follows from Corollary [B.14] that we have
\[
B_{\Phi, \beta}^{(2)}(\mu^*) = 2 \mathcal{B}(k, d, \beta) + o_\omega(1).
\]

The proof of the proposition is then concluded as follows. For any \( \epsilon > 0 \), we consider \( \omega_0 > 0 \) such that for \( \omega \geq \omega_0 \) and \( \Phi \in \mathcal{F}_{n,k,d,\beta,\omega} \), we have \( B_{\Phi, \beta}^{(2)}(\mu^*) \leq 2 \mathcal{B}(k, d, \beta) + \epsilon \). Let \( \Phi = \Phi \) (resp. \( \Phi = \Phi \)). It follows from Proposition [C.1] (resp. Proposition [C.2]) that for \( n \) large enough it holds w.h.p. that \( \sup_{\mu \in M_{\omega}^{(2)}} B_{\Phi, \beta}^{(2)}(\mu) \leq 2 \mathcal{B}(k, d, \beta) + \epsilon \). \( \square \)

In the following of this section, let \( \omega \geq 0 \) and \( \Phi \in \mathcal{F}_{n,k,d,\beta,\omega} \) be fixed. Let also \( \mu^* = (\mu^*_i)_{i \in V} \) be a fixed global maximizer of \( B_{\Phi, \beta}^{(2)} \) over \( M_{\omega}^{(2)} \).

C.2. **Proof of Proposition [C.3]** In order to prove the proposition, we will first show that \( \mu^* \) cannot be too close to the boundary of \( M_{\omega}^{(2)} \). We define
\[
V_{\text{bad}}(\Phi) = \{ i \in \text{Core}_{1}(\Phi), 1 - \mu^*_i (1, 1) \geq \exp(-4 \beta) \}.
\]
We shall prove the following in Section [C.4]

**Lemma C.5.** \( V_{\text{bad}}(\Phi) \) is 1-sticky.
In particular, because $\Phi$ satisfies $\textbf{P2}$, \( |V_{\text{bad}}(\Phi)| \leq 2^{1-k/10} n \) and we have the following.

**Corollary C.6.** \( V_{\text{bad}}(\Phi) \subset S_1(\Phi) \).

**Proof of Proposition C.3.** Because $\beta$ is finite, it is easily checked that for all $i \in V$ and $(x, y) \in \{-1, 1\}$, $\mu^*_i(x, y) \in (0, 1)$. Using in addition Corollary C.6, we obtain that $\mu^* \in M^{(2)}_0$, and thus the differential of $B_{\Phi, \beta}$ vanishes at this point. The existence of $v^*$ and $\hat{v}^*$ such that $(\mu^*, v^*, \hat{v}^*)$ satisfies the paired Belief Propagation equation, and the expression of $B_{\Phi, \beta}(\mu^*)$ in terms of these messages, is then a standard result (see [21] Proposition 14.4) for a proof.

In the following, let $v^* = (v^*_{i-a})_{(i, a) \in E}$ and $\hat{v}^* = (\hat{v}^*_{a-i})_{(i, a) \in E}$ be as given by Proposition C.3 for the formula $\Phi \in \mathcal{E}_{n, k, d, \beta, \omega}$ and maximizer $\mu^* \in M^{(2)}_0$ considered here.

**C.3. Proof of Proposition C.4.** We will now show that the values of messages corresponding to variables in the core are well controlled. More precisely let $V'_{\text{bad}}(\Phi)$ be defined by

\[ V'_{\text{bad}}(\Phi) = \{ i \in \text{Core}_1(\Phi) \setminus S_1(\Phi) ; \exists a \in \partial i, 1 - v^*_{i-a}(1, 1) > \exp(-2\beta) \} \]

The following fact is a standard property of the Belief Propagation equations [21].

**Fact C.7.** We have, for any $i \in V$ and $a \in \partial i$,

\[
\frac{\mu^*_i(x, y)}{v^*_{i-a}(x, y)} = \frac{\hat{v}^*_{a-i}(x, y)}{\sum_{(x', y') \in \{-1, 1\}^2} \hat{v}^*_{a-i}(x', y')}.
\] (C.6)

In particular, the bound $\exp(-2\beta) \leq \psi_{a, \beta}(x, y) \leq 1$ (for $(x, y) \in \{-1, 1\}^{2k}$) easily implies the following.

**Fact C.8.** We have, for any $i \in V$ and $a \in \partial i$,

\[
\exp(-2\beta) \leq \left| \frac{\mu^*_i(x, y)}{v^*_{i-a}(x, y)} \right| \leq \exp(2\beta).
\]

This readily implies the following.

**Corollary C.9.** We have $V'_{\text{bad}}(\Phi) = \emptyset$.

**Proof.** By Fact C.8 we have $V'_{\text{bad}}(\Phi) \subset V_{\text{bad}}(\Phi) \setminus (V_{\text{bad}}(\Phi) \cap S_1(\Phi))$. The result then follows from Corollary C.6.

We are now in position to look at the contraction properties of the Belief Propagation iteration operator $\mathcal{T}_{\Phi, \beta}$. Our analysis will build on the one developed for the tree case. For an edge $(i, a) \in E$ such that the $2\omega$-neighborhood of $i$ is a tree, we denote by $v^*_{T^{(\omega)}_{i-a}} \in \mathcal{P}(\{-1, 1\}^d)$ the boundary condition associated to the tree $T^{(\omega)}_{i-a}$ via the maximizer $(v^*, \hat{v}^*)$. We define $\epsilon^*_{T^{(\omega)}_{i-a}}(a \in \text{frozen, non frozen}) \in \mathcal{P}(\{-1, 1\}^d)$ by (for $(j, b) \in \partial T^{(\omega)}_{i-a}(e^*_{\partial T^{(\omega)}_{i-a}} j) = \text{frozen}$ if and only if $v_{j-b}(1, 1) \leq 1 - \exp(-2\beta)$. We say that $(i, a)$ is $\omega$-good if the $2\omega$-neighborhood of $i$ is tree like and if the associated pair $(T_{i, a}, \epsilon^*_{\partial T^{(\omega)}_{i-a}}(a \in \text{frozen, non frozen}) \in \mathcal{P}(\{-1, 1\}^d)$ is $\omega$-frozen. The following easily follows Corollary C.9 and condition P3.

**Remark C.10.** All but $\alpha_\omega(n)$ of the edges $(i, a) \in E$ are $\omega$-good.

We first need a counterpoint to Lemma B.3. The following is easily proved by recurrence over $\omega$.

**Lemma C.11.** Let $(i, a) \in E$ be $\omega$-good and $\epsilon = \epsilon(T, e^*_{\partial T^{(\omega)}_{i-a}}))$. Let $(i, a) \in E^{(\omega)}$. Then if $i \in T^{(\omega)}_{i-a} \setminus \partial T^{(\omega)}_{i-a}$ and that $e_i$ is frozen, we have $1 - v^*_{i-a}(1, 1) < \exp(-100\beta)$ and $1 - v^{(\omega)}_{i-a}(1, 1) < \exp(-100\beta)$.

This allows to derive the following.
Lemma C.12. Assume that \([i, a] \in E\) is \(\omega\)-good. Then
\[
\sup_{x \in [-1, 1]^2} |v^*_{i \rightarrow a}(x) - v^{(\omega)}_{i \rightarrow a}(x)| = o_\omega(1).
\]

We defer the proof of the lemma for a second.

Proof of Proposition C.4. The first part of the proposition is a direct consequence of Lemma C.12 and Remark C.10. The second part is easily obtained by observing that for an edge \([j, a] j \in \partial(a, i)\) such that the edges \([j, a] j \in \partial(a, i)\) are \((\omega - 1)\)-good, we have \(\sup_{x \in [-1, 1]^2} |\hat{v}^{*}_{a \rightarrow i}(x) - \hat{v}^{(\omega - 1)}_{a \rightarrow i}(x)| = o_\omega(1)\). By Remark C.10, there are \(|E| - o_\omega(n)\) such edges.

Proof of Lemma C.12. The proof is very similar to the one of Lemma B.9. We will first introduce a different parametrization of the messages, and define
\[
\hat{\varepsilon}'_{a \rightarrow i}(x, y) = 1 - c_\beta \mathbf{1}_{|x| = 1} \prod_{i \in \partial(a, i)} \left( v_{i \rightarrow a}(b_{i, a}, -1) + v_{i \rightarrow a}(b_{i, a}, 1) \right)
- c_\beta \mathbf{1}_{|y| = 1} \prod_{i \in \partial(a, i)} \left( v_{i \rightarrow a}(-1, b_{i, a}) + c_\beta^2 v_{i \rightarrow a}(-1, b_{i, a}) \right) + \mathbf{1}_{(x, y) = (-1, -1)} \prod_{i \in \partial(a, i)} v_{i \rightarrow a}(b_{i, a}, b_{i, a}),
\]
\[
Q_{i \rightarrow a \rightarrow b}(\hat{\varepsilon})(x, y) = \prod_{a \in \partial_1(i, a_0)} \hat{\varepsilon}_{a \rightarrow i_0}(x, y) \prod_{a \in \partial_1(i_0, a_0)} \hat{\varepsilon}_{a \rightarrow i_0}(-x, -y)
\]

(we keep the dependence on \(\hat{\varepsilon}\) on \(\hat{\varepsilon}\) implicit in the following). With these notations the Belief Propagation equation (C.1), (C.3) can be rewritten as (C.7) and
\[
v'_{i \rightarrow a}(x, y) = \frac{Q_{i \rightarrow a \rightarrow b}(\hat{\varepsilon}')(x, y)}{\sum_{(x', y') \in [-1, 1]^2} Q_{i \rightarrow a \rightarrow b}(\hat{\varepsilon}')(x, y)}.
\]

Let \(i \in V_{\text{good}, \omega}\) be fixed. For \(j \in B_\omega(i)\), let \(\omega_j = \omega - \text{dist}(i, j)\). For \((j, a) \in E^+_{\omega}\), we define
\[
\gamma^{(2)}_{j \rightarrow a} = |\hat{\varepsilon}'|, \forall a' \in \partial(j, a), \forall x \in [-1, 1]^2,
\]
\[
\hat{\varepsilon}'_{a' \rightarrow j}(x) \in [\min(\hat{\varepsilon}'_{a' \rightarrow j}(x), \hat{\varepsilon}^{(\omega)}_{a' \rightarrow j}(x)), \max(\hat{\varepsilon}'_{a' \rightarrow j}(x), \hat{\varepsilon}^{(\omega)}_{a' \rightarrow j}(x))].
\]

By applying Taylor’s theorem to Eq. (C.7), we obtain,
\[
\sup_{x \in [-1, 1]^2} |v^*_{j \rightarrow a}(x) - v_{j \rightarrow a}^{(\omega)}(x)| \leq \sup_{(\hat{\varepsilon}'_{a' \rightarrow j}, (x)) \in \gamma^{(2)}_{j \rightarrow a}} \left( \frac{\sum_{x' \neq x} Q_{j \rightarrow a}(\hat{\varepsilon}'(x'))}{Q_{j \rightarrow a}(\hat{\varepsilon}(x))} \right) \sup_{a' \in \partial(j, a), x \in [-1, 1]^2} \left| \hat{\varepsilon}'_{a' \rightarrow j}(x) - \hat{\varepsilon}^{(\omega)}_{a' \rightarrow j}(x) \right|.
\]

We observe that for \(x \in [-1, 1]^2\) we have
\[
\sup_{(\hat{\varepsilon}'_{a' \rightarrow j}, (x)) \in \gamma^{(2)}_{j \rightarrow a}} \left( \frac{\sum_{x' \neq x} Q_{j \rightarrow a}(\hat{\varepsilon}'(x'))}{Q_{j \rightarrow a}(\hat{\varepsilon}(x))} \right) \leq \frac{1}{4}.
\]

Moreover, if \(v_j\) is frozen, we have with Lemma C.11 that for \(x \neq (1, 1)\)
\[
\sup_{(\hat{\varepsilon}'_{a' \rightarrow j}, (x)) \in \gamma^{(2)}_{j \rightarrow a}} \frac{Q_{j \rightarrow a}(\hat{\varepsilon}(x))}{Q_{j \rightarrow a}(\hat{\varepsilon}(1, 1))} \leq \exp(-100\beta)\]
and thereby, for all $x \in [-1,1]^2$

$$\sup_{(\hat{a}',j') \in \partial(a,j) \in \mathcal{Y}_{j-a}^{(2)}} \frac{|\hat{e}_{a'} \rightarrow j(x) - \hat{e}_{a}^{(a,j)}(x)|}{\hat{e}_{a'} \rightarrow j(x)} \leq \exp(-100\beta). \quad (C.11)$$

We further observe that for any $(a', j) \in E_\omega^+$ and $x \in [-1,1]^2$ we have

$$\sup_{(\hat{a}',j) \in \partial(a,j) \in \mathcal{Y}_{j-a}^{(2)}} \frac{|\hat{e}_{a'} \rightarrow j(x) - \hat{e}_{a}^{(a,j)}(x)|}{\hat{e}_{a'} \rightarrow j(x)} \leq \exp(2\beta) \sum_{j' \in \partial(a',j) \in [-1,1]^2} |v_{j' \rightarrow a}(x') - v_{j' \rightarrow a}^{(a,j)}(x')|. \quad (C.12)$$

Moreover, if $a'$ is cold (w.r.t. $\epsilon$) we have for $x \in [-1,1]$.

$$\sup_{(\hat{a}',j) \in \partial(a,j) \in \mathcal{Y}_{j-a}^{(2)}} \frac{|\hat{e}_{a'} \rightarrow j(x) - \hat{e}_{a}^{(a,j)}(x)|}{\hat{e}_{a'} \rightarrow j(x)} \leq \exp(-99\beta) \sum_{j' \in \partial(a',j) \in [-1,1]^2} \sum_{\epsilon_j \neq \text{frozen}} |v_{j' \rightarrow a}(x') - v_{j' \rightarrow a}^{(a,j)}(x')| + \exp(\beta) \sum_{j' \in \partial(a',j) \in [-1,1]^2} \sum_{\epsilon_j \neq \text{frozen}} |v_{j' \rightarrow a}(x') - v_{j' \rightarrow a}^{(a,j)}(x')|. \quad (C.13)$$

Because $T_\omega$ is good, for all $i_\omega \in \partial T_\omega$, the path $[i_0 \rightarrow i_\omega]$ contains at least $|0.4\omega|$ variables $j$ such that $\epsilon_j = \text{frozen}$ or the clause $a$ such that $(j, a) \in E_\omega^+$ is cold. Using this remark while combining Eq. (C.9) and iterating these equations, we obtain

$$\sup_{x \in [-1,1]} |v_{i_0 \rightarrow a_0}(x) - v_{i_0 \rightarrow a_0}^{(a_0)}(x)| \leq \sum_{j \in V} |j| \leq 0 + 6\beta \omega \leq \exp(-20\beta \omega) \leq 2^{-k\omega}. \quad \square$$

C.4. **Proof of Proposition C.5.** In order to prove Lemma C.5 we shall prove that if $V_{\text{bad}}(\Phi)$ was not 1-sticky, there would be $i \in V_{\text{bad}}(\Phi)$ such that, by modifying only $\mu^*_i$, we could increase the value of $B^{(2)}_{\Phi,\beta}$ while remaining in $M^{(2)}_0$. More precisely, we shall define for $0 \leq t \leq \mu^*_t(x, y)/2$, $\mu^{t,x,y} \in \mathcal{P}([-1,1]^d)$ by

$$\mu^{t,x,y}_i(x', y') = \begin{cases} 
\mu^*_i(x', y') + t & \text{if } (x', y') = (1, 1), \\
\mu^*_i(x', y') - t & \text{if } (x', y') = (x, y), \\
\mu^*_i(x', y') & \text{otherwise}.
\end{cases}$$

The following fact is easily observed.

**Fact C.13.** We have $\mu^{t,x,y} \in M^{(2)}_0$ and we have $B^{(2)}_{\Phi,\beta}(\mu^{t,x,y}) \leq B^{(2)}_{\Phi,\beta}(\mu^*)$.

We further define the left derivative of $\partial_{i,x,y} B^{(2)}_{\Phi,\beta}(\mu^*)$ by

$$\partial_{i,x,y} B^{(2)}_{\Phi,\beta}(\mu^*) = \lim_{t \searrow 0} \frac{B^{(2)}_{\Phi,\beta}(\mu^{t,x,y}) - B^{(2)}_{\Phi,\beta}(\mu^*)}{t}.$$

The following standard fact [21] shows that it is well defined and can be expressed using one half of the Belief Propagation equations.

**Fact C.14.** Let $i \in V$ and $(x, y) \in [-1,1]^2$ such that $\mu^*_i(x, y) > 0$ be fixed. There are messages $(v_{i-a})_{a \in \partial i} \in \mathcal{P}([-1,1]^d)$ such that we have

$$n \partial_{i,x,y} B^{(2)}_{\Phi,\beta}(\mu^*) = -\ln \left( \frac{\mu^*_i(1,1)}{\mu^*_i(x,y)} \right) + \sum_{a \in \partial i} \ln \left( \frac{\mu^*_i(1,1)}{v_{i-a}(x,y)} \right). \quad (C.14)$$
and such that, defining for $a \in \partial i$, $\tilde{\nu}_{a \leftarrow i} \in \mathcal{P}([-1,1]^4)$ by Eq. [C.7], $\tilde{\nu}_{a \leftarrow i}$ and $\nu_{i \rightarrow a}$ satisfies (the analogous of) Eq. [C.6].

In the remaining of this section, let for each $i \in V \ (\nu_{i \rightarrow a})_{a \in \partial i}$ be fixed as in Fact [C.14]. The following facts will allow us to evaluate the contribution of each term in the summation [C.14]. They easily follow from Fact [C.7] and a study of the cavity equation [C.1].

**Fact C.15.** Let $i \in V$ and $a \in \partial_i$ be fixed. Then

$$\ln \left( \frac{\mu_i^+(1,1) \nu_{i \rightarrow a}(x,y)}{v_i^-(1,1) \mu_i^+(x,y)} \right) \geq 0.$$

*Proof.* We have, with Eq. [C.6]

$$\frac{\mu_i^+(1,1) \nu_{i \rightarrow a}(x,y)}{v_i^-(1,1) \mu_i^+(x,y)} = \frac{\tilde{\nu}_{a \leftarrow i}(x,y)}{\nu_{a \leftarrow i}(x,y)} = \frac{\sum_{\nu(x',y') \in [-1,1]^2} \prod_{j \in \partial(a,i)} \nu_{i \rightarrow a}(x'_j, y'_j) \mathbf{1}(x'_j = 1, y'_j = 1)}{\sum_{\nu(x',y') \in [-1,1]^2} \prod_{j \in \partial(a,i)} \nu_{i \rightarrow a}(x'_j, y'_j) \mathbf{1}(x'_j = x, y'_j = y)} \geq 1.$$

We also note that Fact [C.8] still holds.

**Fact C.16.** We have, for any $i \in V$, $a \in \partial_i$ and $(x,y) \in [-1,1]^2$

$$\exp(-2\beta) \leq \left| \frac{\mu_i^+(x,y)}{v_i^-(x,y)} \right| \leq \exp(2\beta).$$

To keep the notations short, let us introduce for $a \in F$ and $i \in \partial a \ \partial_{\text{good}}(a,i) = \{ j \in \partial(a,i), j \in \text{Core}_1(\Phi) \setminus V_{\text{bad}}(\Phi) \}$.

**Fact C.17.** Let $i \in V$ and $a \in \partial_i$ be fixed. Let $p \geq 1$ be fixed. Assume that $|\partial_{\text{good}}(a,i)| \geq p$. Then we have, for any $(x,y) \in [-1,1]^2$

$$\left| \ln \left( \frac{\mu_i^+(1,1) \nu_{i \rightarrow a}(x,y)}{v_i^-(1,1) \mu_i^+(x,y)} \right) \right| \leq \exp(-2p\beta/3).$$

*Proof.* Let us first introduce the following shortcut notation, for $(x'_a, y'_a) \in [-1,1]^k$, we define

$$\mathbf{1}_{\text{bad}}(i,a)(x'_a, y'_a) = \prod_{j \in \partial_{\text{good}}(i,a)} \left( 1 - \mathbf{1}(x_j = 1, y_j = 1) \right).$$

Using that for $j \in \partial_{\text{good}}(a,i)$ we have $\nu_{j \rightarrow a}(1,1) \geq 1 - \exp(-\beta)$ (by Fact [C.16]), we have for $(x,y) \in [-1,1]^2$

$$\frac{\mu_i^+(1,1) \nu_{i \rightarrow a}(x,y)}{v_i^-(1,1) \mu_i^+(x,y)} = \frac{\tilde{\nu}_{a \leftarrow i}(1,1)}{\nu_{a \leftarrow i}(x,y)} = \frac{1 - O(1) \sum_{\nu(x',y') \in [-1,1]^2} \prod_{i \in \partial(a,i)} \nu_{i \rightarrow a}(x'_i, y'_i) \mathbf{1}(x'_i = 1, y'_i = 1) \mathbf{1}_{\text{bad}}(i,a)(x'_a, y'_a)}{1 - O(\exp(-\beta)) \sum_{\nu(x',y') \in [-1,1]^2} \prod_{i \in \partial(a,i)} \nu_{i \rightarrow a}(x'_i, y'_i) \mathbf{1}(x'_i = x, y'_i = y) \mathbf{1}_{\text{bad}}(i,a)(x'_a, y'_a)} = \frac{1 + O(k \exp(-p\beta))}{1 + O(k \exp(-p\beta))},$$

from which the fact easily follows. 

Finally, we need to show that there is actually a large discrepancy between marginal and messages in the single case where we think it should happen: when a variable is forced by a clause.
Fact C.18. Let $i \in V$ be fixed. Assume that $a \in \partial_{1,0} i$ and that $\partial_{-1} a \subseteq \text{Core}_1(\Phi) \setminus V_{\text{bad}}(\Phi)$. Then, for $(x, y) \in \{-1, 1\}^2 \setminus \{(1, 1)\}$

$$\ln \left( \frac{\mu^*_i(1, 1)}{\nu_i(1, 1)} \right) \geq \beta \frac{\nu_i(1, 1)}{\mu^*_i(1, 1)}.$$

Proof. Using that for $j \in \partial_{\text{good}}(a, i)$ we have $\nu_j(a, 1, 1) \geq 1 - \exp(-\beta)$ (by Fact C.16), we have for $(x, y) \in \{-1, 1\}^2$ and by a similar reasoning as previously

$$\frac{\mu^*_i(1, 1)}{\nu_i(1, 1)} \frac{\nu_i(1, 1)}{\mu^*_i(1, 1)} = \frac{\nu(a, i)}{\nu(i, 1)} \geq \frac{1 + O(k \exp(-\beta))}{\exp(-\beta) + O(k \exp(-\beta))},$$

from which the fact easily follows. □

Proof of Lemma C.5. Assume that $V_{\text{bad}}(\Phi) \neq \emptyset$ and that $V_{\text{bad}}(\Phi)$ is not 1-sticky. Then by definition there is $i \in V_{\text{bad}}(\Phi)$ such that the two following conditions hold.

(a) $|\{a \in \partial_{1,0} i, \partial_{-1} a \not\subseteq V_{\text{good}}(\Phi)\}| \leq k^{3/4},$
(b) $|\{a \in \partial_{-1} x, |\partial_{-1} a| < k, |\partial_{1} a \cap V_{\text{bad}}(\Phi) \cup U| \geq |\partial_1 a|/2\}| \leq k^{3/4}.$

Using conditions CR4-CR5, this further implies that the following is true. Let $U = V \setminus \text{Core}_1(\Phi)$. Then the following conditions hold.

(a') $|\{a \in \partial_{1,0} i, |\partial_{-1} a \cap (V_{\text{bad}}(\Phi) \cup U) \geq 1\}| \leq 2k^{3/4},$
(b') $|\{a \in \partial_{-1} x, |\partial_{-1} a| < k, |\partial_{1} a \cap (V_{\text{bad}}(\Phi) \cup U)| \geq |\partial_1 a|/2\}| \leq k^{3/4}.$

Let $(x, y) \in \{-1, 1\}^2 \setminus \{(1, 1)\}$ and such that $\mu_i(x, y) > 0$ be fixed. Replacing in the derivative $\partial_{x,y} B_{\Phi,\beta}(\mu^*)$, using Fact C.14 along with Fact C.16, Fact C.18, and the conditions CR1, CR2 and CR3, we obtain

$$\partial_{x,y} B_{\Phi,\beta}(\mu^*) \geq \ln \left( \frac{\mu_i(x, y)}{\mu_i(1, 1)} \right) + \frac{\beta}{2} \left( |\partial_{1,0} i| - |\{a \in \partial_{1,0} i, |\partial_{-1} a \cap (V_{\text{bad}}(\Phi) \cup U) \geq 1\}| \right)$$

$$- \frac{1}{2} \left( |\{a \in \partial_{-1} x, |\partial_{-1} a| = l, |\partial_1 a \cap (V_{\text{bad}}(\Phi) \cup U) \geq |\partial_1 a|/2\}| \exp(-l\beta/3) \right)$$

$$- 2\beta (3k^{3/4} + 3)$$

$$\geq \ln \left( \frac{\mu_i(x, y)}{\mu_i(1, 1)} \right) + \frac{\beta k^{7/8}}{2} - \sum_{l=1}^{k} l^{1.3} \exp(-l\beta/3) / l! - 4k^{3/4}$$

Because $i \in V_{\text{bad}}(\Phi)$, we have $\frac{\mu_i(x, y)}{\mu_i(1, 1)} \geq \exp(-3\beta)$ and therefore $\partial_{x,y} B_{\Phi,\beta}(\mu^*) > 0$. This contradicts the fact that $\mu^*$ is a maximum of $B_{\Phi,\beta}$ (cf. Fact C.13). Hence $V_{\text{bad}}(\Phi)$ is 1-sticky. □

Appendix D. Typical Properties of the Random Formula

In this section we prove Proposition C.1 and Proposition C.2. We assume that $d \in \{d_-(k), d_k-\text{SAT}\}$ and that $\beta \geq \beta_- (k)$. Let $\mathcal{E}_{n,k,d}$ denote the set of regular $k$-SAT formulas.

D.1. Proof of Proposition C.1 and Proposition C.2. We first deal with the easiest condition P1.

Lemma D.1. W.h.p. $\Phi$ satisfies P1.

Proof. Let $i \in V$ and $T \in \mathcal{T}_{\omega+1}$ be fixed. Let $X_i(T)$ be the number of formulas $\Phi \in \mathcal{E}_{n,k,d}$ such that $T_i^{(\omega+1)} = T$. It is straightforward to compute that there are precisely

$$\tilde{p}_{k,d,\beta}^{(\omega+1)}(T) \frac{(nd/2)^2}{(nd/2 - \epsilon_+)(nd/2 - \epsilon_-)(1 + o_n(1))}.$$
ways to construct a tree of depth $2(\omega + 1)$ around $i$, where $e_+$ (resp. $e_-$) is the number of positive (resp. negative) literals that appear in $T \setminus \partial T$. Once this has been done, it remains to connect the $(dn/2 - e_+)$ positive literals clones (resp. $(dn/2 - e_-)$ negative literals clones) together. This yield

$$
\frac{X_i(T)}{|\mathcal{E}_{n,k,d}|} = \tilde{p}^{(\omega+1)}_{k,d,\beta}(T) - \frac{(nd/2)^2 - (nd/2 - e_+)!}{(nd/2 - e_+)! (nd/2 - e_-)!} \frac{(1 + o_n(1))}{(nd/2)^2} = \tilde{p}^{(\omega+1)}_{k,d,\beta}(T).
$$

Consequently, we have

$$
\mathbb{E} \left[ \rho_{\Phi}(T) \right] = \frac{X_i(T)}{|\mathcal{E}_{n,k,d}|} = \tilde{p}^{(\omega+1)}_{k,d,\beta}(T).
$$

Moreover by standard concentration arguments $\rho_{\Phi}(T)$ is concentrated around its mean and we have w.h.p.

$$
\rho_{\Phi}(T) \sim \tilde{p}^{(\omega+1)}_{k,d,\beta}(T).
$$

This holds for any $T$ in the finite set $\mathcal{F}_{\omega+1}$, ending the proof of the lemma.

In particular, this entails the following.

**Corollary D.2.** W.h.p. $\Phi$ satisfies P1.

We will prove the following in Section 1.3

**Proposition D.3.** W.h.p. $\Phi$ and $\tilde{\Phi}$ satisfy P2.

The remaining of this section is devoted to a proof of the two following lemmas.

**Lemma D.4.** W.h.p. $\tilde{\Phi}$ satisfies P3.

**Lemma D.5.** W.h.p. $\tilde{\Phi}$ satisfies P3.

**Proof of Proposition D.4 and Proposition D.5.** Let $\Phi = \tilde{\Phi}$ (resp. $\Phi = \hat{\Phi}$). That $\Phi$ satisfies P1 w.h.p. follows from Lemma D.1 (resp. Corollary D.2). That it satisfies P2 w.h.p. follows from Proposition D.3. That it satisfies P3 w.h.p. follows from Lemma D.4 (resp. Lemma D.5). \qed

Let $\alpha \geq 0$ and $(i_1, \ldots, i_\alpha) \in V^\alpha$ be fixed, as well as a formula $\Phi \in \mathcal{E}_{n,k,d}$. Let $\Delta = \{ j \in V, \exists l \in [\alpha], j \in \partial T^{(l)}_{i_l} \}$ and for $j \in \Delta$ let $\mathcal{E}_j$ be the event that $j$ is safe. Moreover, let $\mathcal{D}$ be the event that $i_1, i_2, \ldots, i_\alpha$ are at distance strictly greater than $5\omega$ one from the other, and that their $5\omega$ neighborhoods are tree-like. For $j \in \Delta$, let also $\mathcal{F}_j$ denote the $\sigma$-algebra generated by the functions $\langle \Phi, i_1, \ldots, i_\alpha \rangle \mapsto T^{(l)}_{i_l} \cup \cdots \cup T^{(l)}_{i_\alpha}$.

**Lemma D.6.** For $j \in \Delta$, we have

$$
\mathbb{P} \left[ \neg \mathcal{E}_j | \mathcal{D}, \mathcal{F}_j \right] \leq 2^{-0.95k}.
$$

**Proof.** Let $l \in [d]$ be such that $j[l] = a_j$. Let $\hat{\Phi}'$ be obtained from $\hat{\Phi}$ by the following operations.

- Select $i \in \hat{\Phi}$ and $a_i \in \partial i$ uniformly at random.
- Replace the pair of edges $\{(j, a_j), (i, a_i)\}$ by the pair of edges $\{(j, a_i), (i, a_j)\}$.

Let $\mathcal{E}$ be the event that $\hat{\Phi}'$ satisfies $\mathcal{D}$. We observe that $\mathbb{P}[\mathcal{D}] = 1 - o_n(1)$ and $\mathbb{P}[\mathcal{E}] = 1 - o_n(1)$. Conditioned on $\mathcal{D}, \mathcal{E}$ and $\mathcal{F}_j$, $\hat{\Phi}$ and $\hat{\Phi}'$ are identically distributed. Moreover, we have

$$
\text{Core}_{1/2} \left( \hat{\Phi} \right) \setminus S_{1/2} \left( \hat{\Phi} \right) \subset \text{Core}_{1/2} \left( \hat{\Phi}' \right) \setminus S_{1/2} \left( \hat{\Phi}' \right).
$$

It follows that

$$
\mathbb{P} \left[ \neg \mathcal{E}_j | \mathcal{D}, \mathcal{F}_j \right] = \mathbb{P} \left[ \neg \mathcal{E}_j | \mathcal{D}, \mathcal{E}, \mathcal{F}_j \right] + o_n(1)
$$

$$
\leq \mathbb{P} \left[ i \not\in \text{Core}_{1/2} \left( \hat{\Phi} \right) \setminus S_{1/2} \left( \hat{\Phi} \right) | \mathcal{D}, \mathcal{E}, \mathcal{F}_j \right] + o_n(1)
$$

$$
\leq \mathbb{P} \left[ i \not\in \text{Core}_{1/2} \left( \hat{\Phi} \right) \setminus S_{1/2} \left( \hat{\Phi} \right) | \mathcal{F}_j \right] + o_n(1).
$$

(D.1)
For a fixed $\sigma$-algebra $\mathcal{F}$ generated by $(\Phi, i_1, \ldots, i_\alpha)$ \(\mapsto T_{i_1}^{(\alpha)} \cup \cdots \cup T_{i_\alpha}^{(\alpha)}\), let $\mathcal{H}$ denote the event that there is $\hat{\Phi}'$ isomorphic to $\hat{\Phi}$ such that $\mathcal{F}_j = \mathcal{F}$. Then, because $i$ is a random element of $V$, we have

\[
\mathbb{P}\left[ i \notin \text{Core}_{1/2} \left( \hat{\Phi} \right) \setminus S_{1/2} \left( \hat{\Phi} \right) \right] = \mathbb{P}\left[ i \notin \text{Core}_{1/2} \left( \hat{\Phi} \right) \setminus S_{1/2} \left( \hat{\Phi} \right) \right] = \mathbb{P}\left[ i \notin \text{Core}_{1/2} \left( \hat{\Phi} \right) \setminus S_{1/2} \left( \hat{\Phi} \right) \right] + o_n(1), \quad (D.2)
\]

where the last line used that $\mathbb{P}[\mathcal{H}] = 1 - o_n(1)$. Finally, Proposition $D.3$ implies that

\[
\mathbb{P}\left[ i \notin \text{Core}_{1/2} \left( \hat{\Phi} \right) \setminus S_{1/2} \left( \hat{\Phi} \right) \right] \leq 2^{1-0.96k} + o_n(1). \quad (D.3)
\]

Combining $(D.1)$, $(D.2)$ and $(D.3)$ concludes the proof of the lemma. \hfill \Box

**Proof of Lemma $D.4$**

Let $Y = |i \in V, T_i^{(\alpha)}$ is not $\omega$-safe in $\hat{\Phi}|$, and let $\alpha(n)$ be a slowly diverging function. We are going to show that there is a sequence $y_\omega = o_\omega(1)$ such that

\[
\mathbb{E}[Y(Y - 1) \cdots (Y - \alpha + 1)] \leq (y_\omega n)^{\alpha}. \quad (D.4)
\]

This bound implies the assertion; indeed,

\[
\mathbb{P}[Y > 3y_\omega n] \leq \mathbb{P}[Y(Y - 1) \cdots (Y - \alpha + 1) > (2y_\omega n)^{\alpha}] \leq \frac{\mathbb{E}[Y(Y - 1) \cdots (Y - \alpha + 1)]}{(2y_\omega n)^{\alpha}} \leq 2^{-\alpha}.
\]

To prove $(D.4)$, we observe that $Y(Y - 1) \cdots (Y - \alpha + 1)$ is just the number of orderer $\alpha$-tuples of variables such that $T_i^{(\alpha)}$ is not $\omega$-safe. Hence, by symmetry and linearity of expectation,

\[
\mathbb{E}[Y(Y - 1) \cdots (Y - \alpha + 1)] \leq n^\alpha \mathbb{P}[T_1, \ldots, T_\alpha$ are not $\omega$-safe],
\]

where $T_1, \ldots, T_\alpha$ are $2\omega$-neighborhoods chosen uniformly at random in $\hat{\Phi}$. Let $\mathcal{D}$ be the event that the roots of the trees $T_1, \ldots, T_\alpha$ are at distance greater than $5\omega$ from each others and have tree-like $5\omega$ neighborhoods, and let $\Delta = \partial T_1 \cup \cdots \cup \partial T_\alpha$. Then Lemma $D.6$ implies that for $j \in \Delta$

\[
\mathbb{P}[\lnot \mathcal{C}_j | \mathcal{D}, \mathcal{F}_j] \leq 2^{-0.95k}. \quad (D.5)
\]

We have

\[
\mathbb{P}[T_1, \ldots, T_\alpha$ are not $\omega$-safe] \leq \mathbb{P}[T_1$ is not $\omega$-safe$| \mathcal{D}] \mathbb{P}[T_2$ is not $\omega$-safe$| T_1$ is not $\omega$-safe$| \mathcal{D}] \ldots \mathbb{P}[T_\alpha$ is not $\omega$-safe$| T_1, \ldots, T_{\alpha-1}$ is not $\omega$-safe$| \mathcal{D}]$.
\]

Using $(D.8)$ we can apply Proposition $B.6$ yielding

\[
\mathbb{P}[T_1, \ldots, T_\alpha$ are not $\omega$-safe$| \mathcal{D}] \leq (o_\omega(1))^\alpha.
\]

Along with observation that $\mathbb{P}[\mathcal{D}] = 1 - o_n(1)$, this concludes the proof of the proposition. \hfill \Box

In order to prove Lemma $D.5$ we need to extend Lemma $D.6$ to the planted replica model. This will require a few more auxiliary results. We say that a tree $T \in \mathcal{T}_\omega$ is $\omega$-pure if and only if

\[
\eta^{(\omega)}(1, T) \geq 1 - \exp(-100\beta).
\]

Let $\mathcal{T}_\omega^+ \subset \mathcal{T}_\omega$ denote the set of pure trees. Let, as before, $\alpha \geq 0$ and $(i_1, \ldots, i_\alpha) \in V^\alpha$ be fixed, as well as a formula and an assignment $(\Phi, \sigma) \in \mathcal{E}_{n,k,d} \times [-1,1]^n$. Let $\Delta = \{j \in V, \exists l \in [\alpha], j \in \partial T_i^{(\alpha)}\}$ and for $j \in \Delta$, let $a_j$ be the unique clause in $\delta j \cap \bigcup_{l=1}^{\alpha} T_{i_l}$. For $j \in \Delta$, let

- $A_j$ be the event that $T_j^{(\alpha)}$ is $\omega$-pure,
- $B_j$ be the event that $\sigma(j) = 1$,
- $C_j$ be the event that $j$ is safe.
Moreover, let $\mathcal{D}$ be defined as previously. For $j \in \Delta$, let also $\mathcal{G}_j$ denote the sigma algebra induced by the functions

$$(\Phi, i_1, \ldots, i_a, j) \rightarrow \left( X = T_{i_1}^{(2\omega)} \cup \cdots \cup T_{i_a}^{(2\omega)} \setminus T_{i_j}^{(2\omega)} \setminus j \rightarrow a_j, \sigma_{|X}|. \right)$$

With these notations in mind, we will prove the following.

**Lemma D.7.** For $j \in \Delta$, we have

$$P \left[ \mathcal{A}_j \mid \mathcal{D}, \mathcal{G}_j \right] \leq 2^{-0.95 k}.$$

**Proof.** The proof is easily obtained by recursion over $\omega$, using in addition the fact that $P[\mathcal{D}] = 1 - o_n(1)$.

**Lemma D.8.** For $j \in \Delta$, we have

$$P \left[ \neg \mathcal{B}_j \mid \mathcal{A}_j, \mathcal{D}, \mathcal{G}_j \right] \leq 4^{-k}.$$

**Proof.** Recalling the definition of the replica planted model and introducing $l \in \lfloor k \rfloor$ such that $a_j[l] = j$, we have

$$P \left[ \neg \mathcal{B}_j \mid \mathcal{A}_j, \mathcal{D}, \mathcal{G}_j \right] = \sum_{\hat{T} \in \hat{\mathcal{T}}_\omega} P \left[ \hat{T}_{a_j}^{(\omega)} = \hat{T} \mid \mathcal{A}_j, \mathcal{D}, \mathcal{G}_j \right] \hat{\mu}^{(\omega)}(-1, \hat{T}, l),$$

where $\hat{\mu}^{(\omega)}(\cdot, \hat{T}, l)$ denotes the marginal of $\mu^{(\omega)}(\cdot, \hat{T})$ (defined in Section [B.3]) with respect to the $l$-th variable. For $\hat{T} \in \hat{\mathcal{T}}_\omega$ and $1 \leq l \leq k$, let $\hat{T}[l]$ denote the subtree of size $2\omega$ rooted at the $l$-th variable node adjacent to the root of $\hat{T}$. For $T \in \mathcal{T}_+ \omega$, let $\hat{\mathcal{T}}_\omega(T, l) \subset \hat{\mathcal{T}}_\omega$ denote the set of trees compatible with $T$ on $l$-th position:

$$\hat{\mathcal{T}}_\omega(T, l) = \left\{ \hat{T} \in \hat{\mathcal{T}}_\omega \mid \hat{T}[l] = T \right\}.$$

Then we immediately deduce from the previous equation that

$$P \left[ \neg \mathcal{B}_j \mid \mathcal{A}_j, \mathcal{F}_j \right] \leq \sup_{\hat{T} \in \hat{\mathcal{T}}_\omega} \sup_{\hat{T} \in \hat{\mathcal{T}}_\omega(T, l)} \hat{\mu}^{(\omega)}(-1, \hat{T}, l)(1 + o_n(1))$$

Using the definition of pure trees, we observe that for any $T \in \mathcal{T}_w^+$ and $\hat{T} \in \hat{\mathcal{T}}_\omega(T, l)$ we have $\hat{\mu}^{(\omega)}(-1, \hat{T}, l) \leq \exp(-50\beta) \leq 4^{-k}$, ending the proof of the lemma.

**Lemma D.9.** For $j \in \Delta$, we have

$$P \left[ \neg \mathcal{E}_j \mid \mathcal{A}_j, \mathcal{B}_j, \mathcal{D}, \mathcal{G}_j \right] \leq 2^{-0.95 k}.$$

**Proof.** Let $l \in [d]$ be such that $j[l] = a_j$. Let $\Phi'$ be obtained from $\Phi$ by the following operation.

- Select $i \in \Phi$ such that $\sigma[i] = \sigma[j]$ and $T_{i}^{(2\omega)}[l] = T_{i}^{(2\omega)}[-l]$. (D.6)

Let $\mathcal{E}$ be the event that $\Phi'$ satisfies $\Phi$. We observe that $P[\mathcal{D}] = 1 - o_n(1)$ and $P[\mathcal{E}] = 1 - o_n(1)$. Conditioned on $\mathcal{D}$, $\mathcal{E}$ and $\mathcal{G}_j$, $\Phi$ and $\Phi'$ are identically distributed. Moreover, we have

$$\text{Core}_{1/2}(\Phi) \setminus S_{1/2}(\Phi') \subset \text{Core}_{1}(\Phi) \setminus S_{1}(\Phi').$$

It follows that

$$P \left[ \neg \mathcal{E}_j \mid \mathcal{A}_j, \mathcal{B}_j, \mathcal{D}, \mathcal{G}_j \right] = P \left[ \neg \mathcal{E}_j \mid \mathcal{A}_j, \mathcal{B}_j, \mathcal{D}, \mathcal{E}, \mathcal{G}_j \right] (1 + o_n(1))$$

$$\leq \sum_{T \in \mathcal{T}_w^+} P \left[ T_{i}^{(2\omega)} = T \mid \mathcal{A}_j, \mathcal{B}_j, \mathcal{D}, \mathcal{E}, \mathcal{G}_j \right]$$

$$P \left[ i \notin \text{Core}_{1/2}(\Phi) \setminus S_{1/2}(\Phi) \mid \mathcal{B}_i, \mathcal{D}, \mathcal{E}, \mathcal{G}_j, T_{i}^{(2\omega)} \right] \left| \mathcal{B}_i, \mathcal{D}, \mathcal{E}, \mathcal{G}_j, T_{i}^{(2\omega)} \right| = T'.$$
We define, for $T' \in \mathcal{T}_{2\omega}^+$, $T(T') \in \mathcal{T}_{2\omega}$ by $T[l] = \hat{\omega}_{a_j}^{(2\omega)}$ and $T[-l] = T'$ for $l' \neq l$. It follows from the definition of pure trees that for $T' \in \mathcal{T}_{2\omega}^+$ we have $\mu^{(2\omega)}(1, T(T')) \geq 1/2$. Therefore

$$
\mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{A}_j, \mathcal{B}_j, \mathcal{D}, \mathcal{E}, \mathcal{G}_j \right] = \frac{\mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{A}_j, \mathcal{F}_j \right] \mu^{(2\omega)}(1, T(T'))}{\sum_{T'' \in \mathcal{T}_{2\omega}} \mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T'' \mid \mathcal{A}_j, \mathcal{F}_j \right] \mu^{(2\omega)}(1, T(T'))} (1 + o_n(1))
$$

$$
\leq \frac{\mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{A}_j, \mathcal{F}_j \right]}{\sum_{T'' \in \mathcal{T}_{2\omega}} \mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T'' \mid \mathcal{A}_j, \mathcal{F}_j \right]} \frac{\mu^{(2\omega)}(1, T(T'))}{1/2}
\leq 2\mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{A}_j, \mathcal{F}_j \right].
$$

Moreover

$$
\mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{A}_j, \mathcal{F}_j \right] = \mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{A}_j \right] (1 + o_n(1))
\leq \mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \right] \mathbb{P}\left[ \mathcal{A}_j \right] (1 + o_n(1))
\leq 2\mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \right].
$$

where we used Lemma [D.9] to obtain the second inequality. Using Bayes's theorem once more, we have

$$
\mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \right] = \mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{B}_j \right] \mathbb{P}\left[ \mathcal{B}_j \right] \leq \mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{B}_j \right].
$$

In order to deduce the last inequality, we used that by an argument similar to Lemma [D.8] $\mathbb{P}\left[ \mathcal{B}_j \mid \hat{T}_{j-a_j}^{(2\omega)} \right] \geq 1/2$. It follows by replacing in [D.8] that

$$
\mathbb{P}\left[ \neg \mathcal{E}_j \mathcal{A}_j, \mathcal{B}_j, \mathcal{D}, \mathcal{F}_j \right] \leq 6 \sum_{T' \in \mathcal{T}_{2\omega}} \mathbb{P}\left[ \hat{T}_{j-a_j}^{(2\omega)} = T' \mid \mathcal{B}_j \right] \mathbb{P}\left[ i \notin \text{Core}_{1/2}(\Phi) \setminus S_{1/2}(\Phi) \mid \mathcal{B}_i, \mathcal{D}, \mathcal{E}, T_{j-i}[l] = T' \right]
$$

$$
\leq 6\mathbb{P}\left[ i \notin \text{Core}_{1/2}(\Phi) \setminus S_{1/2}(\Phi) \mid \mathcal{B}_i, \mathcal{D}, \mathcal{E} \right]
\leq 6\mathbb{P}\left[ i \notin \text{Core}_{1/2}(\Phi) \setminus S_{1/2}(\Phi) \mid \mathcal{B}_i, \mathcal{D}, \mathcal{E} \right] \mathbb{P}[\mathcal{B}_i]^{-1} \mathbb{P}[\mathcal{D}]^{-1} \mathbb{P}[\mathcal{E}]^{-1}
\leq 8\mathbb{P}\left[ i \notin \text{Core}_{1/2}(\Phi) \setminus S_{1/2}(\Phi) \right].
$$

We used Lemma [D.8] to deduce the last inequality. Along with Proposition [D.3] this ends the proof of the lemma.

**Proof of Proposition [C.2]** We take a path similar to the proof of Proposition [C.1]. Let

$$
Y = \{|i \in V, T_i^{(w)} \text{ is not } \omega \text{-safe in } \Phi|\},
$$

and let $\alpha(n)$ be a slowly diverging function. We are going to show that there is a sequence $y_\omega = o_\omega(1)$ such that

$$
\mathbb{E}[Y(Y - 1) \ldots (Y - \alpha + 1)] \leq y_\omega n^\alpha.
$$

This bound implies the assertion as previously. As before, we observe that

$$
\mathbb{E}[Y(Y - 1) \ldots (Y - \alpha + 1)] \leq n^\alpha \mathbb{P}[T_1, \ldots, T_\alpha \text{ are not } \omega \text{-safe}],
$$

where $T_1, \ldots, T_\alpha$ are $2\omega$-neighborhoods of depth $\omega$ chosen uniformly at random in $\Phi$. Let $\mathcal{D}$ be the event that $T_1, \ldots, T_\alpha$ are tree-like and disjoints, and let $\Delta = \partial T_1 \cup \cdots \cup \partial T_\alpha$. By combining Lemma [D.7] Lemma [D.8] Lemma [D.9] we obtain that for $j \in \Delta$

$$
\mathbb{P}[\neg \mathcal{E}_j \mid \mathcal{D}, \mathcal{F}_j] \leq 2^{-0.94k}.
$$
We have
\[ \mathbb{P}[T_1, \ldots, T_\alpha \text{ are not } \omega\text{-safe}|\mathcal{D}] \leq \mathbb{P}[T_1 \text{ is not } \omega\text{-safe}|\mathcal{D}]. \]
\[ \mathbb{P}[T_2 \text{ is not } \omega\text{-safe}|T_1 \text{ is not } \omega\text{-safe}, \mathcal{D}] \ldots \]
\[ \mathbb{P}[T_\alpha \text{ is not } \omega\text{-safe}|T_1, \ldots, T_{\alpha-1} \text{ is not } \omega\text{-safe}]. \]
Using (D.8) we can apply Proposition B.6, yielding
\[ \mathbb{P}[T_1, \ldots, T_\alpha \text{ are not } \omega\text{-safe} | \mathcal{D}] \leq (a_\alpha(1))^\alpha. \]
Along with observation that \( \mathbb{P}[\mathcal{D}] = 1 - a_\alpha(1), \) this concludes the proof of the proposition.

D.2. Proof of Proposition D.3. In order to prove Proposition D.3, we will identify a set of simpler events that will imply the proposition. We will first need to control the number of vertices with unusual 1-neighborhood. To this end, we let for \( \Phi \in \mathcal{E}_{n,k,d}, U_0 \) be the set of variables such that \( |a| \leq 2^{k/20} n \) and such that \( |a| \geq 2 \). Let \( \Phi \) be the resulting subset of \( U_0 \). Our first condition will ensure that \( U_0 \) is not too large:
\[ |U_0| \leq 2^{-0.98k} n. \tag{C0} \]
We now turn to expansion properties of \( \Phi \). We define, for a set \( T \subset V \) the sets
\[ F_0(T) = \{ a \in \mathcal{E}_{n,k,d} : |a| = 1, |a| \leq |T| \} \]
\[ F_1(T) = \{ a \in \mathcal{E}_{n,k,d} : a \cap T \neq \emptyset, |a| \leq |T| \} \]
for each \( 1 \leq l \leq k \).

The following conditions encompass bounds on the sizes of the sets \( F_1(T) \) when \( T \) has moderate size.

There is no set \( T \subset V \) of size \( |T| \leq 2^{-0.97k} n, 2^{-k/20} n \) and such that \( |F_0(T)| \leq 2^{3/4} |T| / 100 \).

For each \( 1 \leq l \leq k \), there is no set \( T \subset V \) of size \( |T| \leq 2^{-0.97k} n, 2^{-k/20} n \)
and such that \( |F_1(T)| \leq 2^{3/4} |T| / 100 n^2 \).

Lemma D.10. Assume that \( \Phi \in \mathcal{E}_{n,k,d} \) satisfies \([C0],[C2]\). Then it satisfies \([P2]\).

Proof: Let \( \Phi \in \mathcal{E}_{n,k,d} \) be such that it satisfies \([C0],[C2]\). We first prove that \( \Phi \) does not admit a 1-sticky set \( S \) with \( |S| \leq 2^{-0.97k} n, 2^{-k/20} n \). Indeed, let \( S \subset V \) be a 1-sticky set for \( \Phi \) and let
\[ S_0 = \{ i \in S : |a| = 1, |a| \leq 2^{k/20} n \} \]
\[ S_i = \{ i \in S : |a| = 1, |a| \leq 2^{k/20} n \}. \]
We first observe that
\[ |F_0(S)| \geq 2^{3/4} |S_0| / 2 \]
and for that \( 1 \leq l \leq k - 1 \) \( |F_1(S)| \geq 2^{3/4} |S_l| / (2l) \).

Because \( S \) is 1-sticky, we have \( S \subset S_0 \cup \cup_{i=1}^{k-1} S_i \) and therefore either \( |S_0| \geq 2^{k/20} n \) or there is \( 1 \leq l \leq k - 1 \) such that \( |S_l| \geq 2^{k/20} n \). In either case, it follows from \([D3],[C1],[C2]\) that \( |S_l| \leq 2^{-0.98k} n, 2^{-k/20} n \).

Using that \( |F_0(S)| \leq 2^{-0.97k} n, 2^{-k/20} n \) shows that \( S \) has size outside the range \( |F_1(S)| \leq 2^{-0.97k} n, 2^{-k/20} n \).

We now turn to the study of the strong core of \( \Phi \in \mathcal{E}_{n,k,d} \). Given \( \Phi \), we consider the following whitening process. Let \( U = U_0 \) initially. While there is a variable \( i \notin U \) such that one of the following conditions occurs, add \( i \) to \( U \).

(a) \( |a| \neq 1, |a| \neq 2^k \).
(b) \( |a| \geq 2, |a| \leq 2^k \).

It is easily seen that the process converges. Let \( U_\infty \) be the resulting subset of \( V \), then we have
\[ \text{Core}(\Phi) = V \setminus U_\infty. \tag{D.10} \]
We are going to show that \( U_\infty \) cannot be too large. By condition \([C0]\), we can assume that \( |U_0| \leq 2^{0.98k} n \).

Assume for contradiction that \( |U_\infty| \geq 2^{-0.97k} n \) and let \( U \) be the set obtained when precisely \( 2^{-0.97k} n |U_0| \) variables have been added to \( U_0 \). By construction each variable \( i \in U \) has one of the following properties.

(00) \( i \) belongs to \( U_0 \),
Replacing with Remark B.15 yields \( W .v .h.p. \) as desired.

Proof of Proposition D.3. The propositions follow from combining Lemma D.10 combined with Lemma D.11. The case is impossible by a similar reasoning as previously and we obtained that \( |U_\infty| \leq 2^{-0.97k} n \) w.h.p. \( \square \)

Studying \( \Phi \) will be enough to obtain the information needed about \( \Phi \). Indeed, we shall obtain sufficiently strong estimates of the probability of events under the random formula \( \Phi \) to transfer them into high probability statements for the biased distribution generating \( \Phi \). More precisely, say that \( \Phi \) satisfies a property \( (\mathcal{P}) \) with very high probability (w.v.h.p.) iff \( (\mathcal{P}) \) has probability larger than \( 1 - \exp(-2^{-0.99k} n) \) under \( \Phi \). Then we can infer that \( (\mathcal{P}) \) has a large probability under the random formula \( \Phi \).

Lemma D.11. Let \( \mathcal{A} \) be an event. Assume that \( \Phi \) satisfies \( \mathcal{A} \) w.v.h.p. Then \( \Phi \) satisfies \( \mathcal{A} \) w.h.p.

Proof. Reformulating the definition of the planted replica model, we see that

\[
\Pr[\neg \mathcal{A}] = \frac{\sum_{\Phi \in \mathcal{E}_{n,k,d}} 1[\Phi \notin \mathcal{E}][\Phi = \Phi] \exp(nB_{\Phi,\beta}(\mu^{(o)}))}{\sum_{\Phi \in \mathcal{E}_{n,k,d}} \Pr[\Phi = \Phi] \exp(nB_{\Phi,\beta}(\mu^{(o)}))} \leq \frac{\sup_{\Phi \in \mathcal{E}_{n,k,d}} \exp(nB_{\Phi,\beta}(\mu^{(o)}))}{\sum_{\Phi \in \mathcal{E}_{n,k,d}} \Pr[\Phi = \Phi] \exp(nB_{\Phi,\beta}(\mu^{(o)}))} \Pr[\neg \mathcal{A}].
\]

(D.11)

We observe that, with the help of Remark B.15,

\[
\sup_{\Phi \in \mathcal{E}_{n,k,d}} \exp(nB_{\Phi,\beta}(\mu^{(o)})) \leq \exp(nH(\mu^{(o)})) \leq \exp(2^{-0.99k} n),
\]

and that, returning to the definition of \( \mathcal{E}^{(o)}(k, d, \beta) \) in Section B.3

\[
\sum_{\Phi \in \mathcal{E}_{n,k,d}} \Pr[\Phi = \Phi] \exp(nB_{\Phi,\beta}(\mu^{(o)})) \geq 1/2 \exp(n\mathcal{E}^{(o)}(k, d, \beta)) \geq \exp(-2^{-0.99k} n).
\]

Replacing with Remark B.15 yields

\[
\Pr[\neg \mathcal{A}] \leq 2 \exp(2^{-0.999k} n) \Pr[\neg \mathcal{A}] \leq 2 \exp(2^{-0.999k} n - 2^{-0.99k} n) = o(n),
\]

as desired. \( \square \)

In order to obtain our result, we are thus left with proving the following proposition.

Proposition D.12. W.v.h.p. \( \Phi \) satisfies \( \mathcal{C} \), \( \mathcal{C} \), and \( \mathcal{C} \).

Proof of Proposition D.3. The propositions follow from combining Lemma D.10, combined with Lemma D.11 and Proposition D.12. \( \square \)

D.3. Proof of Proposition D.12. In this section we shall study typical properties of the random formula \( \Phi \). We first need a new view on the way \( \Phi \) is generated. For \( \Phi \in \mathcal{E}_{n,k,d} \) and \( 0 \leq l \leq k \), we let \( m_l(\Phi) \) count the number of clauses \( a \) of \( \Phi \) such that \( |\delta_1 a| = l \).

We define a probability distribution \( r_{k,d,\beta} \) over \( [m_l]^{k+1} \) by

\[
r_{k,d,\beta}(m_0, \ldots, m_k) = \frac{\exp(-\beta m_0(\Phi)) |\Phi \in \mathcal{E}_{n,k,d}, \forall i \in [0, \ldots, k], m_i(\Phi) = m_i|}{\sum_{(m_0, \ldots, m_k) \in [m_l]^{k+1}} \exp(-\beta m_0(\Phi)) |\Phi \in \mathcal{E}_{n,k,d}, \forall i \in [0, \ldots, k], m_i(\Phi) = m_i|}.
\]

We can generate \( \Phi \) in the following way.

- We first draw a vector \( (m_0, \ldots, m_k) \) from the distribution \( r_{k,d,\beta} \).
- Then we draw \( \Phi \) from \( \mathcal{E}_{n,k,d} \) uniformly at random, conditioned on \( m_i(\Phi) = m_i \) for \( 0 \leq i \leq k \).
In view of the above discussion, given \((m_0, \ldots, m_k)\) we can also see \(\Phi(n, k, d, \beta)\) as being generated from the following configuration model. We let \(\mathcal{L} = \bigcup_{j \in [d]} \{i\} \times \{d\}\) and \(\mathcal{L}' = \bigcup_{j=0}^{k} \bigcup_{a_j \in [m_j]} \{j\} \times \{a_j\} \times \{k\}\) (resp. \(\mathcal{I} = \bigcup_{a \in [m]} \{j\} \times \{a\} \times \{k\}\)). Let \(\tilde{\Phi}'\) be a uniformly random bijection from \(\mathcal{I}' \to \mathcal{L}'\), subject to \(\Phi(j, a_j, l) = (i, r)\) with \(r \leq d/2\) if and only if \(l \leq j\). Given \(\tilde{\Phi}'\), we see it as a bijection \(\tilde{\Phi}\) from \(\mathcal{I} \to \mathcal{L}'\) and let \(\tilde{\Phi}\) denote the formula defined by letting

\[
\Phi_{a,l} = \begin{cases} x_i & \text{if } \Phi(a, l) = (i, r) \text{ with } r \leq d/2, \\
-x_i & \text{if } \Phi(a, l) = (i, r) \text{ with } j > d/2,
\end{cases}
\]

\[
\tilde{\Phi} = \bigvee_{a \in [m]} \bigwedge_{l \in [k]} \Phi_{a,l}.
\]

To simplify the discussion, we shall call each element \((i, j) \in \mathcal{L}\) a positive (resp. negative) clone of the variable \(i\) if \(j \leq d/2\) (resp. \(j > d/2\)), and each element \((j, a, l) \in \mathcal{I}'\) as a positive (resp. negative) clone of the clause \(a\) if \(l \leq j\) (resp. \(l > j\)). Given the numbers \(m_0, \ldots, m_k\) and the choice of the associated clauses, the formula \(\tilde{\Phi}\) can therefore be seen as a random matching of the positive (resp. negative) clauses clones with the positive (resp. negative) variable clones.

**Lemma D.13.** W.v.h.p. \(\tilde{\Phi}\) satisfies \([E] \Omega\).

**Proof.** Let \(Y_i\) denote the number of variables \(i \in V\) such that \(|a_1 \in \partial_i| \leq 4k^{7/8}\). Let \(p\) denote the probability that a binomial of parameters \((k - 1, 1 - q)\) takes values 0. Using Lemma D.1 and recalling the definition of \(p_k^{(0)}\) in Section \([L]\) gives

\[
\mathbb{E}[Y_i] \leq \sum_{r=0}^{4k^{7/8} - 1} \binom{d/2}{r} p'(1 - p)^{d/2 - r}.
\]

A simple computation reveals that \(p = 2^{1-k} + \tilde{O}_k(4^{-k})\). This implies that the summand is maximal for \(r = 4k^{7/8}\) and allows to bound \(\mathbb{E}[Y_i]\) as

\[
\mathbb{E}[Y_i] \leq \tilde{O}_k(4^{-k}) = \tilde{O}_k(2^{-k}).
\]

A standard concentration argument then yields that \(Y_i \leq 2^{-0.99k}\) w.v.h.p.

Similarly, let \(Y_2\) denote the number of variables \(i \in V\) such that \(\partial_1 i \geq 2\). Let \(Q\) denote the probability that a binomial of parameter \((d/2, \exp(-\beta) q^{-1}/(1 - c_v q^{-1})\) takes a value larger than 2. By another simple computation, we find that \(Q = \tilde{O}_k(2^{-k})\). It follows from Lemma D.1 that \(\mathbb{E}[Y_2] = Q = \tilde{O}_k(2^{-k})\). Again, by concentration this implies \(Y_2 \leq 2^{-0.99k}\) w.v.h.p.

Finally, for \(1 \leq l \leq k\) let \(Y_3(l)\) be the number of variables \(i \in V\) with \(|\partial_1 i| \geq k^{l+3} / l!\). By similar computations, we obtain

\[
\mathbb{E}[Y_3(l)] \leq \sum_{r=k^{l+3} / l!}^{d/2} \binom{d/2}{r} \binom{k - 1}{l} \frac{1}{2^k} \big(1 + \tilde{O}_k(2^{-k})\big) = \tilde{O}_k(2^{-k}).
\]

It follows that \(\mathbb{E}[Y_3] = \tilde{O}_k(2^{-k})\), and by the same concentration argument as previously \(Y_3 \leq 2^{-0.99k}\) w.v.h.p.

The proof of the lemma is completed by noting that \(|U_0| \leq Y_1 + Y_2 + Y_3\). \(\square\)

We define \(m'_l = \frac{4}{2e} k^{l+3} / l!\). The previous estimates can easily be (slightly extended and) recast as follows.

**Remark D.14.** W.v.h.p. we have for all \(0 \leq l \leq k\), \(m_l(\tilde{\Phi}) \leq m'_l\).

We are now ready to complete

**Lemma D.15.** W.v.h.p. \(\tilde{\Phi}\) satisfies \([E] \Omega\).
Proof. Given \( \hat{\Phi} \), let \( X_0(t, r, y) \) count the number of sets \( T \in V \) of size \( |T| = tn \), such that

- \( |F_0(T)| = rt n \),
- \( \sum_{a \in F_0(T)} |\partial^{-1} a \cap T| = yrtkn \).

By definition of \( F_1(T) \), \( X_0(t, r, y) = 0 \) if \( y < k^{-1} \). The expected value of \( X_0(t, r, y) \) can be computed in the following manner. First choose the sets \( T \) and \( F_0(T) \). The latter has to be chosen among the \( m_0(\hat{\Phi}) \) satisfied clauses. Among the \( tdn \) literal clones from \( T \), choose the \( rt n \) positive literal clones that will be connected to the positive literal clones of clauses in \( F_0(T) \), and the \( ytdnk \) literal clones that will be connected to negative literal clones of clauses in \( F_0(T) \). Make the same choices among the negative and positive literal clones of the clauses in \( F_0(T) \). Then match these \( rt n \) positive literal clones (resp. \( yrtkn \) negative literal clones) at random, and then match the remaining \( dtn/2 - rt n \) remaining positive literal clones (resp. \( dtn/2 - rt n \) remaining negative literal clones) at random. The normalizing factor is the total number of graphs that can be obtained from the configuration model, \( (dtn/2)!^2 \). Without words, and using in addition Remark\[D.14\] to observe that we can assume \( m_0(\hat{\Phi}) \leq m'_0 = \frac{d}{2k}k^3 \), this gives

\[
\mathbb{E}[X_0(t, r)] \leq \left( \frac{n}{tn} \right) \left( \frac{m'_0}{rt n} \right) \left( \frac{tdn}{yrtk n} \right) \left( \frac{rt n}{yrtk n} \right) \left( \frac{yrtk n}{yrtk n} \right) \frac{(rt n)!((dtn/2 - rt n)!((dtn/2)!)^2}{(dtn/2)!^2}
\]

We shall bound this quantity by using the bounds, for \( 1 \leq a \leq b \) and \( n > 0 \)

\[
b \ln \left( \frac{a}{b} \right) \leq \frac{1}{n} \ln \left( \frac{an}{bn} \right) \leq b \ln \left( \frac{ae}{b} \right).
\]

This yields

\[
\frac{1}{n} \ln \mathbb{E}[X_0(t, r, y)] \leq t \ln \left( \frac{e}{t} \right) + rt \ln \left( \frac{d^3k}{2ktr} \right) + rt \ln \left( 2ke^2t \right) + yrtk \ln \left( \frac{2e^2t}{y} \right).
\]

In particular for \( r \geq k^{3/4} \), \( t \in [2^{-0.98k}n, 2^{-k/20}n] \), and \( y \geq 1/k \), we get

\[
\frac{1}{n} \ln \mathbb{E}[X_0(t, r, y)] \leq -t \ln t + t + k^{3/4}t \ln \left( k^{10} \right) \leq -2^{-0.98k}n.
\]

In particular

\[
\sum_{r \in [2^{-0.98k}n, 2^{-k/20}n]} \sum_{t \in [k^{3/4}/d]} \sum_{y \in [0, 1]} \mathbb{E}[X_0(t, r, y)] \leq \exp \left[ -2^{-0.985k}n \right].
\]

This implies by Markov’s inequality that w.v.h.p. there are no sets \( T \) of size \( |T| \in [2^{-0.98k}n, 2^{-k/20}n] \) such that \( |F_0(T)| \geq k^{3/4}|T| \). \( \square \)

Lemma D.16. W.v.h.p. \( \hat{\Phi} \) satisfies (E2).

Proof. Given \( \hat{\Phi} \) and \( 1 \leq l \leq k \), let \( X_l(t, r, x, y) \) count the number of sets \( T \in V \) of size \( |T| = tn \) and such that the following condition are true.

- \( |F_l(T)| = rt n \),
- \( \sum_{a \in F_l(T)} |\partial^{-1} a \cap T| = xrtkn \),
- \( \sum_{a \in F_l(T)} |\partial^{-1} a \cap T| = yrtkn \).

By definition of \( F_l(T) \), \( X_l(t, r, x, y) = 0 \) if \( x < lk^{-1}/4 \) or \( y < k^{-1} \). With Remark\[D.14\] we can assume

\[
m_l(\hat{\Phi}) \leq m'_l, \quad \text{with } m'_l = \frac{d}{2^k}k^{l+3}.
\]
Reasoning as before, we obtain
\[ \mathbb{E}[X_I(t, r, x, y)] \leq \left( \frac{n}{tn} \right)^{m'_t} \left( \frac{tdn}{xtn} \right) \left( \frac{yrtn}{xron} \right)^{rtn} \left( \frac{dtn}{rtn} \right)^{-1} \left( \frac{dtn}{yn} \right)^{-1} \cdot \]

Taking logarithm and using (D.12), we obtain
\[ \frac{1}{n} \ln \mathbb{E}[X_I(t, r, x, y)] \leq -t \ln \left( \frac{e}{t} \right) + r t \ln \left( \frac{d k^{l+3}}{2^k l r t} \right) + x r t k \ln \left( \frac{2 e^2 t}{x} \right) + y r t k \ln \left( \frac{2 e^2 t}{y} \right). \]

In particular, for \( r \geq k^{3/4}/(100 l^2), t \in [2^{-0.98k}, 2^{-k/20}] \) and \( x \geq l k^{-1}/4, y \geq k^{-1} \), we have
\[ \frac{1}{n} \ln \mathbb{E}[X_I(t, r, x, y)] \leq -t \ln \left( \frac{e}{t} \right) + r t \ln \left( \frac{d k^{l+3} t^{l/4+1}}{2^k l^2 l!} \right) \leq -t \ln t + t + \frac{k^{3/4}}{100 l^2} t \ln \left( k^{l+8} l^{l/4+1} \right) \]

For any \( 1 \leq l \leq k \) we have \( k^{l+6} l^{l/4+1} \leq k^{l^2} \) and we thereby obtain that
\[ \frac{1}{n} \ln \mathbb{E}[X_I(t, r, x, y)] \leq -2^{-0.985k} n. \]

This entails that, for any \( 1 \leq l \leq k \),
\[ \sum_{t \in [2^{-0.98k}, 2^{-k/20}]} \sum_{r \in [k^{3/4}/(100 l^2), d]} \sum_{x \in [0, 1]} \sum_{y \in [0, 1]} \sum_{x \in [0, 1]} \sum_{y \in [0, 1]} \mathbb{E}[X_I(t, r, x, y)] \leq \exp \left[ -2^{-0.985k} n \right]. \]

This implies by Markov’s inequality that w.v.h.p. there are no \( 1 \leq l \leq k \) and no sets \( T \) of size \( |T| \in [2^{-0.98k} n, 2^{-k/20} n] \) such that \( |F_l(T)| \geq k^{3/4} |T|/(100 l^2) \). \( \square \)

Amin Coja-Oghlan, acoghlan@math.uni-frankfurt.de, Goethe University, Mathematics Institute, 10 Robert Mayer St, Frankfurt 60325, Germany.

Victor Bapst, bapst@math.uni-frankfurt.de, Goethe University, Mathematics Institute, 10 Robert Mayer St, Frankfurt 60325, Germany.