SIAM J. DISCRETE MATH.
Vol. 26, No. 4, pp. 1471–1509

THE DECIMATION PROCESS IN RANDOM k-SAT*

AMIN COJA-OGHLAN† AND ANGELICA Y. PACHON-PINZON‡

Abstract. Let $\mathbf{\Phi}$ be a uniformly distributed random k-SAT formula with n variables and m clauses. Nonrigorous statistical mechanics ideas have inspired a message passing algorithm called *belief propagation guided decimation* for finding satisfying assignments of $\mathbf{\Phi}$. This algorithm can be viewed as an attempt at implementing a certain thought experiment that we call the *decimation process*. In this paper we identify a variety of phase transitions in the decimation process and link these phase transitions to the performance of the algorithm.

Key words. random structures, phase transitions, k-SAT, belief propagation

AMS subject classifications. Primary, 68Q87; Secondary, 60C05, 05C85, 05C65

DOI. 10.1137/110842867

1. Introduction and results.

Let $k \geq 3$ and $n > 1$ be integers, let $r > 0$ be a real, and set $m = \lceil rn \rceil$. Let $\mathbf{\Phi} = \mathbf{\Phi}_k(n, m)$ be a propositional formula obtained by choosing a set of m clauses of length k over the variables $V = \{x_1, \ldots, x_n\}$ uniformly at random. To be precise, we consider $\mathbf{\Phi}$ as a sequence of m clauses, each of which is a set of k distinct, possibly negated variables, i.e., $\mathbf{\Phi}$ is chosen uniformly at random from among all $(2^k\binom{n}{k})^m$ such sequences. For k, r fixed we say that $\mathbf{\Phi}$ has some property $\mathcal{P}$ *with high probability* (w.h.p.) if $\lim_{n\to\infty} \mathrm{P}[\mathbf{\Phi} \in \mathcal{P}] = 1$.

1.1. Background and motivation.

The interest in random k-SAT originates from the experimental observation that for certain densities r the random formula $\mathbf{\Phi}$ is satisfiable w.h.p. while a large class of algorithms, including and particularly the workhorses of practical SAT solving such as sophisticated solvers based on the Davis–Putnam–Logemann–Loveland (DPLL) algorithm, fail to find satisfying assignments efficiently [20]. Over the past decade, a fundamentally new class of algorithms has been proposed on the basis of sophisticated but non-rigorous ideas from statistical physics [7, 19]. Experiments performed for $k = 3, 4, 5$ indicate that these new "message passing algorithms," namely, *belief propagation guided decimation* and *survey propagation guided decimation* (BP and SP decimation), excel on random k-SAT instances [16]. Indeed, the experiments suggest that BP and SP decimation find satisfying assignments for r close to the threshold where $\mathbf{\Phi}$ becomes unsatisfiable w.h.p. Generally, SP decimation is deemed conceptually superior to BP decimation.

For example, in the case $k = 4$ the threshold for the existence of satisfying assignments is conjectured to be $m/n \sim r_4 \approx 9.93$ [18]. According to experiments from [16], SP decimation finds satisfying assignments for densities up to $r = 9.73$. Experiments from [23] suggest that the "vanilla" version of BP decimation succeeds up to $r = 9.05$. Another version of BP decimation (with a different decimation strategy from [7]) succeeds up to $r = 9.24$, again according to experimental data from [16]. By comparison,

*Received by the editors July 29, 2011; accepted for publication (in revised form) July 30, 2012; published electronically October 2, 2012. This work was supported by EPSRC grant EP/G039070/2. An extended abstract version of this work appeared in *Proceedings of the* 38*th ICALP*, 2011, pp. 305–316.

http://www.siam.org/journals/sidma/26-4/84286.html

†Goethe University, Mathematics Institute, Postfach 111932, Fach 187, 60054 Frankfurt, Germany (acoghlan@math.uni-frankfurt.de).

‡TU Graz, Institut für Optimierung und Diskrete Mathematik (Math B), 8010 Graz, Steyrergasse 30/II, Austria (apachon@math.tugraz.at).

the currently best rigorously analyzed algorithm is efficient up to $r = 5.54$ [12], while `zChaff`, a prominent practical SAT solver, becomes ineffective beyond $r = 5.35$ [16].

Since random k-SAT instances have widely been deemed extremely challenging benchmarks, the stellar experimental performance of the physicists' message passing algorithms has stirred considerable excitement. However, the statistical mechanics ideas that BP and SP decimation is based on are highly nonrigorous. Thus, a rigorous analysis of these message passing algorithms is an important but challenging open problem. A first step was made in [9], where it was shown that BP decimation does not outperform far simpler combinatorial algorithms for sufficiently large clause lengths k. More precisely, the main result of [9] is that there is a constant $\rho_0 > 0$ (independent of k) such that the "vanilla" version of BP decimation fails to find satisfying assignments w.h.p. if $r > \rho_0 2^k/k$. By comparison, nonconstructive arguments show that w.h.p. $\mathbf{\Phi}$ is satisfiable if $r < r_k = 2^k \ln 2 - k$ and unsatisfiable if $r > 2^k \ln 2$ [4, 5]. This means that for $k \gg \rho_0$ sufficiently large, BP decimation fails to find satisfying assignments w.h.p. already for densities a factor of (almost) k below the threshold for satisfiability.

The analysis performed in [9] is based on an intricate method for directly tracking the execution of BP decimation. Unfortunately this argument does little to illuminate the conceptual reasons for the algorithms' demise. In particular, [9] does not provide a link to the statistical mechanics ideas that inspired the algorithm.

The present paper aims to remedy these defects. Here we study the *decimation process*, an idealized thought experiment that inspired BP decimation and that the algorithm strives to implement. We show that this experiment undergoes a variety of phase transitions that explain the failure of BP decimation for densities $r > \rho_0 \cdot 2^k/k$. Our results identify phase transitions jointly in terms of the clause/variable density r and with respect to the time parameter of the decimation process. The latter dimension was ignored in the original statistical mechanics work on BP [7, 19] but turns out to have a crucial impact on the performance of the algorithm. On a nonrigorous basis, this has been pointed out recently by Ricci-Tersenghi and Semerjian [23], and our results can be viewed as providing a rigorous albeit quantitatively less precise version of their main results. The results of this paper can also be seen as a generalization of the ones obtained in [2] for random k-SAT, and indeed our proofs build upon the techniques developed in that paper.

1.2. The decimation process. For a propositional formula Φ we let $\mathcal{S}(\Phi)$ signify the set of all satisfying assignments. Furthermore, if x is a variable of Φ and $\mathcal{S}(\Phi) \neq \emptyset$, then we let

$$M_x(\Phi) = \frac{|\{\sigma \in \mathcal{S}(\Phi) : \sigma(x) = 1\}|}{|\mathcal{S}(\Phi)|}$$

denote the marginal probability that x takes the value "true" in a random satisfying assignment of Φ.

BP decimation is a polynomial-time algorithm that aims to (heuristically) implement the following "thought experiment," which we call the *decimation process* [21, 23].[1]

[1] Several different versions of BP decimation have been suggested. In this paper we refer to the simplest but arguably most natural one, also considered in [9, 21, 23]. Other versions decimate the variables in a different order, allowing for slightly better experimental results [7, 16].

EXPERIMENT 1 (decimation process). *Input:* A satisfiable k-CNF Φ.
Result: A satisfying assignment $\sigma : V \to \{0,1\}$ (with 0/1 representing "false"/"true").

0. Let $\Phi_0 = \Phi$.
1. For $t = 1, \ldots, n$ do
2. Compute the fraction $M_{x_t}(\Phi_{t-1})$ of all satisfying assignments of Φ_{t-1} in which the variable x_t takes the value 1.
3. Assign $\sigma(x_t) = 1$ with probability $M_{x_t}(\Phi_{t-1})$, and let $\sigma(x_t) = 0$ otherwise.
4. Obtain the formula Φ_t from Φ_{t-1} by substituting the value $\sigma(x_t)$ for x_t and simplifying (i.e., delete all clauses that were satisfied by assigning x_t, and omit x_t from all other clauses).
5. Return the assignment σ.

A moment's reflection reveals that, given a satisfiable input formula Φ, the decimation process outputs a uniform sample from the set of all satisfying assignments of Φ. The obvious obstacle to actually implementing this experiment is the computation of the marginal probability $M_{x_t}(\Phi_{t-1})$ that x_t takes the value "true" in a random satisfying assignment of Φ_{t-1}, a $\#P$-hard problem in the worst case. Yet the key hypothesis underlying BP decimation is that these marginals can be approximated efficiently on *random* formulas by means of a message passing algorithm. We will discuss the BP decimation algorithm and its connection to Experiment 1 below.

We are going to study the decimation process when applied to a random formula $\boldsymbol{\Phi}$ for densities $r < 2^k \ln 2 - k$, i.e., in the regime where $\boldsymbol{\Phi}$ is satisfiable w.h.p. More precisely, conditioning on $\boldsymbol{\Phi}$ being satisfiable, we let $\boldsymbol{\Phi}_t$ be the (random) formula obtained after running the first t iterations of Experiment 1. The variable set of this formula is $V_t = \{x_{t+1}, \ldots, x_n\}$, and each clause of $\boldsymbol{\Phi}_t$ consists of *at most* k literals. We say that *almost all* $\sigma \in \mathcal{S}(\boldsymbol{\Phi}_t)$ have a certain property $\mathcal{A}$ if $|\mathcal{A} \cap \mathcal{S}(\boldsymbol{\Phi}_t)| = (1 - o(1))|\mathcal{S}(\boldsymbol{\Phi}_t)|$.

We will identify various phase transition that the formulas $\boldsymbol{\Phi}_t$ undergo as t grows from 1 to n. As it turns out, these can be characterized in terms of two simple parameters. The first is the clauses density $r \sim m/n$. Actually, it will be most convenient to work with

$$\rho = kr/2^k,$$

so that $m/n \sim \rho \cdot 2^k/k$. We will be interested in the regime $\rho_0 \le \rho \le (1 - \varepsilon_k)k \ln 2$, where ρ_0 is a constant (independent of k) and $\varepsilon_k \to 0$ as k gets large. The upper bound $(1 - \varepsilon_k)k \ln 2$ marks the point where satisfying assignments cease to exist [5]. The second parameter is the fraction

$$\theta = 1 - t/n$$

of "free" variables (i.e., variables not yet set by time t).

1.3. Results. Let Φ be a k-CNF on $V = \{x_1, \ldots, x_n\}$, let $1 \le t < n$, and let Φ_t be the formula obtained after t steps of the decimation process.

The symmetric phase. Suppose that $\sigma \in \mathcal{S}(\Phi_t)$. A variable $x \in V_t$ is *loose* in σ if there is $\tau \in \mathcal{S}(\Phi_t)$ such that $\sigma(x) \neq \tau(x)$ and $\mathrm{dist}(\sigma, \tau) \le \ln n$, where $\mathrm{dist}(\cdot, \cdot)$ denotes the Hamming distance.

THEOREM 2. *There are constants $k_0, \rho_0 > 0$ such that for $k \ge k_0$, $\rho_0 \le \rho \le k \ln 2 - 2 \ln k$, and*

$$k \cdot \theta > \exp\left[\rho\left(1 + \frac{\ln\ln\rho}{\rho} + \frac{10}{\rho}\right)\right]$$

the random formula $\mathbf{\Phi}_t$ has the following properties w.h.p.:

1. *In almost all satisfying assignments $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ at least $0.99\theta n$ variables are loose.*
2. *At least $\theta n/3$ variables $x \in V_t$ satisfy $M_x(\mathbf{\Phi}_t) \in [0.01, 0.99]$.*
3. *For almost all $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ we have*

$$|\{\tau \in \mathcal{S}(\mathbf{\Phi}_t) : dist(\sigma, \tau) \leq 0.49\theta n\}| \leq \exp(-\Omega(n))\,|\mathcal{S}(\mathbf{\Phi}_t)|\,.$$

Intuitively, Theorem 2 can be summarized as follows. In the early stages of the decimation process (while θ is "big"), most variables in a typical $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ are loose. Hence, the correlations among the variables are mostly local: if we "flip" one variable in σ, then we can "repair" the unsatisfied clauses that this may cause by simply flipping another $\ln n$ variables. Furthermore, for at least a good fraction of the variables, the marginals $M_x(\mathbf{\Phi}_t)$ are bounded away from 0/1. Finally, the third statement says that for a "typical" $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ only an exponentially small fraction of $\mathcal{S}(\mathbf{\Phi}_t)$ lies within distance $0.49\theta n$ off σ. This means that the set $\mathcal{S}(\mathbf{\Phi}_t)$ is "well spread" over the Hamming cube $\{0,1\}^{V_t}$.

Shattering and rigidity. Let Φ be a k-CNF and let $\sigma \in \mathcal{S}(\Phi_t)$. For an integer $\omega \geq 1$ we call a variable $x \in V_t$ *ω-rigid* if any $\tau \in \mathcal{S}(\Phi_t)$ with $\sigma(x) \neq \tau(x)$ satisfies $\mathrm{dist}(\sigma, \tau) \geq \omega$.

Furthermore, we say that a set $S \subset \{0,1\}^{V_t}$ is *(α, β)-shattered* if it admits a decomposition $S = \bigcup_{i=1}^{N} R_i$ into pairwise disjoint subsets such that the following two conditions are satisfied:

SH1. We have $|R_i| \leq \exp(-\alpha\theta n)|S|$ for all $1 \leq i \leq N$.

SH2. If $1 \leq i < j \leq N$ and $\sigma \in R_i$, $\tau \in R_j$, then $\mathrm{dist}(\sigma, \tau) \geq \beta\theta n$.

THEOREM 3. *There are constants $k_0, \rho_0 > 0$ such that for $k \geq k_0$, $\rho_0 \leq \rho \leq k \ln 2 - 2 \ln k$, and*

$$\frac{\rho}{\ln 2}(1 + 2\rho^{-2}) \leq k\theta \leq \exp\left[\rho\left(1 - \frac{\ln \rho}{\rho} - \frac{2}{\rho}\right)\right] \tag{1}$$

there exist $\alpha = \alpha(k, \rho) > 0, \beta = \beta(k, \rho) > 0$ such that the random formula $\mathbf{\Phi}_t$ has the following properties w.h.p.:

1. *In almost all $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ at least $0.99\theta n$ variables are $\Omega(n)$-rigid.*
2. *$\mathcal{S}(\mathbf{\Phi}_t)$ is (α, β)-shattered.*
3. *At least $\theta n/3$ variables $x \in V_t$ satisfy $M_x(\mathbf{\Phi}_t) \in [0.01, 0.99]$.*
4. *For almost all $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ we have*

$$|\{\tau \in \mathcal{S}(\mathbf{\Phi}_t) : dist(\sigma, \tau) \leq 0.49\theta n\}| \leq \exp(-\Omega(n))\,|\mathcal{S}(\mathbf{\Phi}_t)|\,.$$

Thus, if the fraction θ of free variables lies in the regime (1), then in most satisfying $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ the values assigned to 99% of the variables are linked via long-range correlations: to "repair" the damage done by flipping a single rigid variable it is inevitable to reassign a *constant fraction* of all variables. This is mirrored in the geometry of the set $\mathcal{S}(\mathbf{\Phi}_t)$: it decomposes into exponentially many exponentially tiny subsets, which are mutually separated by a linear Hamming distance $\Omega(n)$. Yet as in the symmetric phase, the marginals of a good fraction of the θn remaining variables are bounded away from 0/1, and the set $\mathcal{S}(\mathbf{\Phi}_t)$ remains well spread over the Hamming cube $\{0,1\}^{V_t}$.

The ferromagnetic phase. Let $\alpha > 0$. We say that a set $S \subset \{0,1\}^{\theta n}$ is *α-ferromagnetic* if for any $\sigma, \tau \in S$ we have $\mathrm{dist}(\sigma, \tau) \leq \alpha\theta n$.

THEOREM 4. *There are constants $k_0, \rho_0 > 0$ such that for $k \geq k_0$, $\rho_0 \leq \rho \leq k \ln 2 - 2 \ln k$, and*

$$\ln \rho < k \cdot \theta < (1 - \rho^{-2}) \cdot \frac{\rho}{\ln 2} \tag{2}$$

the random formula $\mathbf{\Phi}_t$ has the following properties w.h.p.:

1. *In almost all $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ at least $0.99\theta n$ variables are $\Omega(n)$-rigid.*
2. *The set $\mathcal{S}(\mathbf{\Phi}_t)$ is $\rho \exp(4 - \rho)$-ferromagnetic.*
3. *At least $0.99\theta n$ variables $x \in V_t$ satisfy $M_x(\mathbf{\Phi}_t) \in \left[0, 2^{-k/2}\right] \cup \left[1 - 2^{-k/2}, 1\right]$.*
4. *There is a set $R \subset V_t$ of size $|R| \geq 0.99\theta n$ such that for any $\sigma, \tau \in \mathcal{S}(\mathbf{\Phi}_t)$ we have*

$$|\{x \in R : \sigma(x) \neq \tau(x)\}| \leq k 2^{2-k} n.$$

In other words, as the decimation process progresses to a point that the fraction θ of free variables satisfies (2), the set of satisfying assignments shrinks into a subset of $\{0,1\}^{V_t}$ of tiny diameter, in contrast to a well-spread shattered set as in Theorem 3. Furthermore, most marginals $M_x(\mathbf{\Phi}_t)$ are either extremely close to 0 or extremely close to 1. In fact, there is a large set R of variables on which all satisfying assignments virtually agree. (More precisely, any two can't disagree on more than $k2^{2-k}n$ variables in R.)

The forced phase. We call a variable x *forced* in the formula Φ_t if Φ_t has a clause that only contains the variable x (a "unit clause"). Clearly, in any satisfying assignment x must be assigned so as to satisfy this clause.

THEOREM 5. *There are constants $k_0, \rho_0 > 0$ such that for $k \geq k_0$, $\rho_0 \leq \rho \leq k \ln 2 - 2 \ln k$, and*

$$\ln n/n \leq k \cdot \theta < \ln(\rho)(1 - 10/\ln \rho) \tag{3}$$

the random formula $\mathbf{\Phi}_t$ is such that at least $0.99\theta n$ variables are forced w.h.p.

Theorem 5 shows that in the final phase of the decimation process, most variables are not just rigid but indeed their values can be read off due to the existence of many unit clauses. In other words, for most variables there is a *local* reason why the variable must take a particular value.

1.4. Belief propagation. As mentioned earlier, the BP decimation algorithm is an attempt at implementing the decimation process by means of an efficient algorithm. The key issue with this is the computation of the marginals $M_{x_t}(\Phi_{t-1})$ in step 2 of the decimation process. Indeed, the problem of computing these marginals is $\#P$-hard in the worst case. Thus, instead of working with the "true" marginals, BP decimation uses certain numbers $\mu_{x_t}(\Phi_{t-1}, \omega)$ that can be computed efficiently, where $\omega \geq 1$ is an integer parameter. The precise definition of the $\mu_{x_t}(\Phi_{t-1}, \omega)$ can be found below, but basically, the $\mu_{x_t}(\Phi_{t-1}, \omega)$ are the result of a "local" dynamic programming algorithm ("belief propagation") that depends upon the assumption of a certain correlation decay property. For given k, ρ, the key hypothesis underpinning the BP decimation algorithm can be stated as follows.

HYPOTHESIS 6. *For any $\varepsilon > 0$ there is $\omega = \omega(\varepsilon, k, \rho, n) \geq 1$ such that w.h.p. for all $1 \leq t \leq n$ we have $|\mu_{x_t}(\mathbf{\Phi}_{t-1}, \omega) - M_{x_t}(\mathbf{\Phi}_{t-1})| < \varepsilon$.*

In other words, Hypothesis 6 states that throughout the entire decimation process, the "BP marginals" $\mu_{x_t}(\Phi_{t-1}, \omega)$ are a good approximation to the true marginals $M_{x_t}(\Phi_{t-1})$.

THEOREM 7. *There exist constants $c_0, k_0, \rho_0 > 0$ such that for all $k \geq k_0$ and $\rho_0 \leq \rho \leq k \ln 2 - 2 \ln k$ the following is true for any integer $\omega = \omega(k, \rho, n) \geq 1$. Suppose that*

$$c_0 \ln(\rho) < k \cdot \theta < \rho / \ln 2. \tag{4}$$

Then for at least $0.99\theta n$ variables $x \in V_t$ we have $\mu_x(\mathbf{\Phi}_t, \omega) \in [0.49, 0.51]$.

Comparing Theorem 4 with Theorem 7, we see that w.h.p. for θ satisfying (4) most of the true marginals $M_x(\mathbf{\Phi}_t)$ are very close to either 0 or 1, whereas the BP marginals lie in $[0.49, 0.51]$. Thus, in the regime described by (4) the BP marginals do *not* provide a good approximation to the actual marginals. Intuitively, this means that in the ferromagnetic phase the hypothesis that the marginals $M_{x_t}(\Phi_t)$ are determined by local effects is wrong. Indeed, our proof of Theorem 4 is based on genuinely "global" expansion-type arguments. Thus, as a direct consequence of Theorems 4 and 7 we obtain the following.

COROLLARY 8. *There exist constants $c_0, k_0, \rho_0 > 0$ such that for all $k \geq k_0$, $\rho_0 \leq \rho \leq k \ln 2 - 3 \ln k$ Hypothesis 6 is untrue.*

Let us state the precise definition of the BP marginals; this follows [7, 9, 21, 22, 23], which also contain more background on BP. (Strictly speaking, we include the definition of the $\mu_{x_t}(\Phi_{t-1}, \omega)$ only for the sake of completeness. We won't actually need the definition in our proof of Theorem 7, as we will simply invoke a result from [9] that reduces the analysis of the BP marginals to the proof of a certain quasi-randomness property of the input formula.)

Let Φ_{t-1} be a Boolean formula on $V_{t-1} = \{x_t, \ldots, x_n\}$. The *factor graph* of Φ_{t-1} is the bipartite graph $G = G(\Phi_{t-1})$ whose vertices are the variables V_{t-1} and the clauses of Φ_{t-1}. Each clause is adjacent to the variables that occur in it. Let $N(v)$ denote the neighborhood of a vertex v in G. For a variable $x \in V_t$ and a clause $a \in N(x)$ we will denote the ordered pair (x, a) by $x \to a$. Similarly, $a \to x$ stands for the pair (a, x). Furthermore, we let $\mathrm{sign}(x, a) = 1$ if x occurs in a positively and $\mathrm{sign}(x, a) = -1$ otherwise.

The *message space* $\mathcal{M}(\Phi_{t-1})$ is the set of all tuples $(\mu_{x \to a}(\zeta))_{x \in V_t,\, a \in N(x),\, \zeta \in \{0,1\}}$ such that $\mu_{x \to a}(\zeta) \in [0, 1]$ and $\mu_{x \to a}(0) + \mu_{x \to a}(1) = 1$ for all x, a, ζ. For $\mu \in \mathcal{M}(\Phi_{t-1})$ we define $\mu_{a \to x}(\zeta) = 1$ if $\zeta = (1 + \mathrm{sign}(x, a))/2$, and

$$\mu_{a \to x}(\zeta) = 1 - \prod_{y \in N(a) \setminus \{x\}} \mu_{y \to a}\left(\frac{1 - \mathrm{sign}(y, a)}{2}\right) \tag{5}$$

otherwise. Furthermore, we define the *belief propagation operator* BP as follows: for any $\mu \in \mathcal{M}(\Phi_{t-1})$ we define $\mathrm{BP}(\mu) \in \mathcal{M}(\Phi_{t-1})$ by letting

$$(\mathrm{BP}(\mu))_{x \to a}(\zeta) = \frac{\displaystyle\prod_{b \in N(x) \setminus \{a\}} \mu_{b \to x}(\zeta)}{\displaystyle\prod_{b \in N(x) \setminus \{a\}} \mu_{b \to x}(0) + \prod_{b \in N(x) \setminus \{a\}} \mu_{b \to x}(1)} \tag{6}$$

unless the denominator equals zero, in which case $(\mathrm{BP}(\mu))_{x \to a}(\zeta) = \frac{1}{2}$.

Finally, the values $\mu_x(\Phi_t, \omega)$ are defined as follows. Let $\mu[0] = \frac{1}{2} \cdot \mathbf{1} \in \mathcal{M}(\Phi_{t-1})$ be the vector with all entries equal to $\frac{1}{2}$. Moreover, define inductively

$$\mu[\ell] = \mathrm{BP}(\mu[\ell - 1]) \qquad \text{for } 1 \leq \ell \leq \omega.$$

Then

$$\mu_x(\Phi_{t-1},\omega)=\frac{\prod_{b\in N(x)}\mu_{b\to x}(1)\,[\omega]}{\prod_{b\in N(x)}\mu_{b\to x}(0)\,[\omega]+\prod_{b\in N(x)}\mu_{b\to x}(1)\,[\omega]} \tag{7}$$

for any $x\in V_t$, unless the denominator is zero, in which case we set $\mu_x(\Phi_{t-1},\omega)=\frac{1}{2}$.

Let us briefly discuss the motivation behind this formal definition. For an integer $\omega\geq 1$ let $N^\omega(x_t)$ signify the set of all vertices of G that have distance at most 2ω from x_t. Then the induced subgraph $G\,[N^\omega(x_t)]$ of the factor graph corresponds to the subformula of Φ_{t-1} obtained by removing all clauses and variables at distance more than 2ω from x_t. Note that all vertices at distance precisely 2ω are variables, so that any satisfying assignment of Φ induces a satisfying assignment of the subformula. Let us denote by $M_{x_t}(\Phi_{t-1},\omega)$ the marginal probability that x_t takes the value 1 in a random satisfying assignment of this subformula.

It is not difficult to verify that if the induced subgraph $G\,[N^\omega(x_t)]$ is a tree, then $\mu_{x_t}(\Phi_{t-1},\omega)=M_{x_t}(\Phi_{t-1},\omega)$ [7]. Indeed, the definition of the $\mu_{x_t}(\Phi_{t-1},\omega)$ can be viewed as a dynamic programming algorithm for computing $M_{x_t}(\Phi_{t-1},\omega)$ if $G\,[N^\omega(x_t)]$ is a tree. Moreover, standard arguments show that in a random formula $\boldsymbol{\Phi}$ actually $G\,[N^\omega(x_t)]$ is a tree w.h.p. (so long as $\omega=o(\ln n)$). Thus, the basic idea behind BP decimation is to approximate the true marginal $M_{x_t}(\Phi_{t-1})$ by the local marginal $M_{x_t}(\Phi_{t-1},\omega)$ in the tree $G\,[N^\omega(x_t)]$.

1.5. Summary and discussion. Fix $k\geq k_0$ and $\rho\geq\rho_0$. Theorems 2–5 show how the space of satisfying assignments of $\boldsymbol{\Phi}_t$ evolves as the decimation process progresses; this is depicted schematically in Figure 1. In the *symmetric phase* $k\theta\geq\exp((1+o_\rho(1))\rho)$, where there still is a large number of free variables, the correlations among the free variables are purely local ("loose variables"). As the number of free variables enters the regime $(1+o_\rho(1))\rho/\ln 2\leq k\theta\leq\exp((1-o_\rho(1))\rho)$, the set $\mathcal{S}(\boldsymbol{\Phi}_t)$ of satisfying assignments shatters into exponentially many tiny "clusters," each of which comprises only an exponentially small fraction of all satisfying assignments. Most satisfying assignments exhibit long-range correlations among the possible values that can be assigned to the individual variables ("rigid variables"). This phenomenon goes by the name *dynamic replica symmetry breaking* in statistical mechanics [17].

While in the previous phases the set of satisfying assignments is scattered all over the Hamming cube, in the *ferromagnetic phase* $(1-o_\rho(1))\ln\rho\leq k\theta\leq(1-o_\rho(1))\rho/\ln 2$ the set of satisfying assignments has a tiny diameter. This is mirrored by the fact that the marginals of most variables are extremely close to either 0 or 1. Furthermore, in (most of) this phase the estimates of the marginals resulting from belief propagation are off (Theorem 7). More precisely, even if BP decimation would emulate the decimation process perfectly up until the ferromagnetic phase commences, the mistaken BP marginals would then lead BP to assign rigid variables "wrongly" (i.e., differently than they are assigned in any satisfying assignment). In effect, BP decimation would fail to find a satisfying assignment, regardless of its subsequent decisions. Finally, in the forced phase $k\theta\leq(1-o_\rho(1))\ln\rho$ there is an abundance of unit clauses that make it easy to read off the values of most variables. However, getting stuck in the ferromagnetic phase, BP decimation won't even reach this regime.

These results suggest that the reason for the failure of BP decimation is the existence of the ferromagnetic phase. Intuitively, in the ferromagnetic phase the marginals

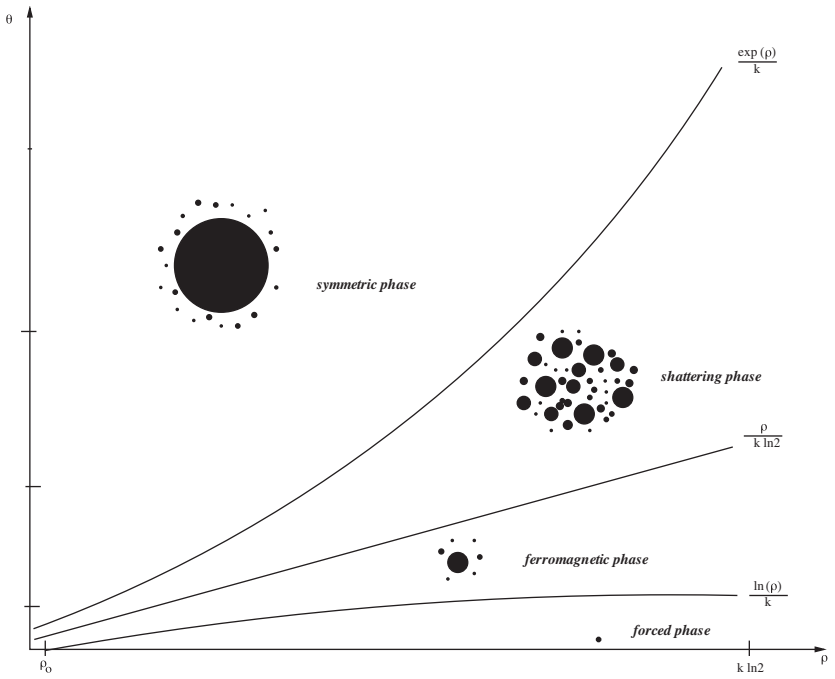

FIG. 1. *The phase diagram of the decimation process with schematic represenations of the "shape" of the solution space $\mathcal{S}(\mathbf{\Phi}_t)$ in the various phases.*

are governed by genuinely global phenomena (essentially expansion properties) that elude the inherently local BP computation. By contrast, it is conceivable that BP does indeed yield the correct marginals in the previous phases. Verifying or falsifying this remains an important open problem.

1.6. Related work.

The statistical mechanics perspective. BP and SP decimation are inspired by a generic but highly nonrigorous analysis technique from statistical mechanics called the *cavity method* [7]. This technique is primarily destined for the *analysis* of phase transitions. It is based on the (unproven) *replica symmetry breaking hypothesis*, which aims to characterize the possible types of correlations among the variables [17].

In [7, 17] the cavity method was used to study the structure of the set $\mathcal{S}(\mathbf{\Phi})$ of satisfying assignments (or, more accurately, properties of the Gibbs measure) of the *undecimated* random formula $\mathbf{\Phi}$. Thus, the results obtained in that (nonrigorous) work identify phase transitions solely in terms of the formula density ρ. On the basis of these results, it was hypothesized that (certain versions of) BP decimation should find satisfying assignments up to $\rho \sim \ln k$ or even up to $\rho \sim k \ln 2$. The argument given for the latter scenario in [17] is that the key obstacle for BP to approximate the true marginals is *condensation*, a phenomenon that is conjectured to be very similar to ferromagnetism with respect to its effect on belief propagation. In terms of the parameter ρ, the condensation threshold was (nonrigorously) estimated to occur at $\rho = k \ln 2 - 3k2^{-k-1} \ln 2$. However, [9] shows that (the basic version of) BP decimation fails to find satisfying assignments already for $\rho \geq \rho_0$ with ρ_0 a constant independent of k.

The explanation for this discrepancy is that [7, 17] neglect the time parameter $\theta = 1 - t/n$ of the decimation process. As Theorem 4 shows, even for *fixed* $\rho \geq \rho_0$ (independent of k) ferromagnetism occurs as the decimation process proceeds to θ in the regime (2). This means that decimating variables has a similar effect on the geometry of the set of satisfying assignments as increasing the clause/variable density. On a nonrigorous basis an analysis both in terms of the formula density ρ and the time parameter θ was carried out in [23]. Thus, our results can be viewed as a rigorous version of [23] (with proofs based on completely different techniques). In addition, Theorem 7 confirms rigorously that for ρ, θ in the ferromagnetic phase, BP does not yield the correct marginals.

The present results have no immediate bearing on the conceptually more sophisticated SP decimation algorithm. However, we conjecture that SP undergoes a similar sequence of phase transitions and that the algorithm will not find satisfying assignments for densities $\rho \geq \rho_0$ with ρ_0 a certain constant independent of k.

Rigorous work. Theorem 3 can be viewed as a generalization of the results on random k-SAT obtained in [2] (which additionally deals with further problems such as random graph/hypergraph coloring). In [2] we rigorously proved a substantial part of the results hypothesized in [17] on shattering and rigidity in terms of the clause/variable density ρ; this improved prior work [3, 6, 10]. The new aspect of the present work is that we identify not only a transition for shattering/rigidity, but also for ferromagnetism and forcing in terms of *both* the density ρ and the time parameter θ of the decimation process. As explained in the previous paragraph, the time parameter is crucial to link these phase transitions to the performance of algorithms such as BP decimation.

In particular, from Theorem 3 we can recover the main result of [2] on random k-SAT. Namely, if $\rho \geq \ln k + 2 \ln \ln k + 2$, then (1) is satisfied even for $\theta = 1$, i.e., the *undecimated* random formula $\mathbf{\Phi}$ has the properties 1–4 stated in Theorem 3 w.h.p. Technically, the present paper builds upon the methods developed in [2]. But in addition, new arguments are needed to accommodate the time parameter θ to prove the statements on the marginals of the variables in Theorems 2–4 and to establish ferromagnetism (Theorem 4).

The best current rigorous algorithmic results for random k-SAT are [8, 12, 13, 15]. For general k, the best current algorithm succeeds up to $\rho \sim \ln k$ [8].

Many algorithmic results for random k-SAT, including [13, 15], are derived via the method of differential equations [24]; an excellent account of this is provided in [1]. Such algorithms typically assign one variable at a time based on some "local" rule, simplify the formula accordingly, and proceed to the next variable until either a satisfying assignment is found or an empty clause (i.e., a contradiction) is generated. In particular, there is usually no (or essentially no) backtracking or reassigning variables at a later time. In a sense, BP guided decimation can be viewed as a far-reaching generalization of these algorithms (cf. [6]). However, in contrast to the decimation process that we are studying here, in differential equations analyses of algorithms the distribution of the formula that occurs after t steps typically has a relatively simple characterization. More precisely, this distribution is normally characterized by a handful of parameters that capture the past decisions of the algorithm, such as the number of clauses of a given length remaining.

By contrast, in the decimation process the distribution that we need to study is inherently complicated. In particular, the analysis of this distribution hinges upon the study of uniformly random satisfying assignments of the random formula $\mathbf{\Phi}$. Indeed, characterizing this distribution precisely is perhaps one of the most difficult open

problems in the theory of random constraint satisfaction problems. However, as we will see we can harness a key insight from [2] in order to get an approximation of this distribution that is sufficiently good for our purposes. We will return to this problem in section 3.

1.7. Organization of the paper. Before we come to the proofs of our main results, we need to go over some notation and preliminaries in section 2. Moreover, in section 3 we will introduce a further probability distribution (the "planted model") that will facilitate the analysis of the decimation process.

Theorems 2–7 are organized according to the various phases of the decimation process. However, to prove them it is more natural to proceed according to the techniques needed to establish the various types of statements. Therefore, we will prove the statements on loose/rigid/forced variables in Theorems 2–4 in section 4. Then, in section 5 we will prove the claims about shattering and ferromagnetism. Furthermore, section 6 deals with the statements on the marginals $M_x(\Phi_{t-1})$. Thus, Theorems 2, 3, 4, and 5 follow by combining the corresponding statements of Theorems 17, 26, and 37 below. Finally, in section 7 we will show how Theorem 7 follows from some auxiliary results from [9].

2. Preliminaries and notation. *Throughout the paper, particularly in the various computations, we will always assume that $k \geq k_0$ and $\rho \geq \rho_0$ for sufficiently large constants $k_0, \rho_0 > 0$, unless specified otherwise. We will also assume implicitly that n is large.*

Recall that $V_t = \{x_{t+1}, \ldots, x_n\}$ is the set of variables of the formula obtained after the first t steps of the decimation process. In addition, we let

$$L_t = \{x_{t+1}, \bar{x}_{t+1}, \ldots, x_n, \bar{x}_n\}$$

be the corresponding set of literals. For a literal l let $|l|$ denote the underlying variable. For a formula Φ on $V = \{x_1, \ldots, x_n\}$, an assignment $\sigma \in \{0,1\}^V$, and $1 \leq t \leq n$ we let $\Phi_{t,\sigma}$ denote the formula obtained by substituting $\sigma(x_s)$ for x_s for all $1 \leq s \leq t$ and simplifying (i.e., omitting all clauses that are satisfied and omitting all false literals from the remaining clauses).

We need the following *Chernoff bound* on the tails of a binomially distributed random variable X with mean λ (see [14, pp. 26–28]): for any $t > 0$,

$$\mathrm{P}(X \geq \lambda + t) \leq \exp\left(-t \cdot \varphi(t/\lambda)\right) \text{ and } \mathrm{P}(X \leq \lambda - t) \leq \exp\left(-t \cdot \varphi(-t/\lambda)\right), \tag{8}$$

where

$$\varphi(x) = (1+x)\ln(1+x) - x. \tag{9}$$

The following theorem provides a lower bound on the typical number of satisfying assignments.

THEOREM 9 (see [3]). *Assume $k \geq 4$ and $\rho \leq k \ln 2 - k^2/2^k$. Then*

$$\frac{1}{n} \ln |\mathcal{S}(\mathbf{\Phi})| \geq \ln 2 + r \ln(1 - 2^{-k}) - 0.99\rho/2^k \qquad w.h.p.$$

There is a natural way to associate a bipartite graph with a propositional formula Φ, known as the *factor graph* (cf. section 1.4). Its vertices are the variables and the clauses of Φ, and each clause is adjacent to all the variables it contains. For a variable x we let $N_3(x)$ be the subgraph of the factor grpah that is spanned by all vertices at distance at most 3 from x. A variable x is *tame* if $N_3(x)$ is acyclic and contains no more than $\ln(n)$ variables. The following is a well-known fact about random k-CNFs.

PROPOSITION 10 (see [2]). *Suppose $k \geq 3$ and $0 < r \leq 2^k \ln 2$. W.h.p. all but $o(n)$ variables are tame in $\mathbf{\Phi}$.*

Finally, the following lemma expresses an elementary expansion property of the random formula $\mathbf{\Phi}$.

LEMMA 11. *For any $k \geq 3$ there is $\chi = \chi(k) > 0$ such that for all $0 < r \leq 2^k$ the random formula $\mathbf{\Phi}$ has the following property w.h.p.:*

(10) *There is no set Q of $1 \leq |Q| \leq \chi n$ variables such that the number of clauses containing at least two variables from Q is at least $2|Q|$.*

Proof. We use a first moment argument. Let $1 \leq q \leq \chi n$ and let $Q_0 = \{x_1, \ldots, x_q\}$ be a fixed set of size q. For any set Q we let $Y(Q)$ be the number of clauses containing at least two variables from Q. Moreover, let X_q be the number of sets Q of size q such that $Y(Q) \geq 2q$. Since the distribution of $\mathbf{\Phi}$ is symmetric with respect to permutations of the variables, we have

$$\mathrm{E}[X_q] \leq \binom{n}{q} \cdot \mathrm{P}[Y(Q_0) \geq 2q] \leq \exp[q(1 + \ln(n/q))] \cdot \mathrm{P}[Y(Q_0) \geq 2q]. \tag{11}$$

Furthermore, the probability that a random k-clause contains two variables from Q_0 is at most $\binom{k}{2}(q/n)^2$ (because for each of the $\binom{k}{2}$ pairs of "slots" in the clauses the probability that both of them are occupied by variables from Q_0 is at most $(q/n)^2$). As $F_k(n,m)$ consists of m independent k-clauses, $Y(Q_0)$ is stochastically dominated by a binomial random variable $\mathrm{Bin}(m, \binom{k}{2}(q/n)^2)$. Consequently, assuming that $q/n \leq \chi$ is sufficiently small, we get

$$\begin{aligned}
&\mathrm{P}[Y(Q_0) \geq 2q] \\
&\quad \leq \mathrm{P}\left[\mathrm{Bin}\left(m, \binom{k}{2}\left(\frac{q}{n}\right)^2\right) \geq 2q\right] \\
&\quad \leq \exp\left[-1.9q \cdot \left[\ln\left(\frac{2q}{\binom{k}{2}(q/n)^2 m}\right) - 1\right]\right] \qquad \text{(by the Chernoff bound (8))} \\
&\quad \leq \exp\left[-1.9q \cdot \ln\left(\frac{4}{\mathrm{e}k^2 r} \cdot \frac{n}{q}\right)\right] \leq \exp\left[-1.9q \cdot \ln\left(\frac{4}{\mathrm{e}k^2 2^k} \cdot \frac{n}{q}\right)\right]. \qquad (12)
\end{aligned}$$

Choosing $\chi = \chi(k)$ sufficiently small, we can ensure that $(q/n)^{1/4} \leq \chi^{1/4} \leq 4/(\mathrm{e}k^2 2^k)$. Plugging this bound into (12), we get

$$\mathrm{P}[Y(Q_0) \geq 2q] \leq \exp[-1.1q \cdot \ln(n/q)]. \tag{13}$$

Combining (11) and (13), we get $\mathrm{E}[X_q] \leq \exp[-0.1q \ln(n/q)]$. In effect,

$$\mathrm{E} \sum_{1 \leq q \leq \chi n} X_q = O(n^{-0.1}).$$

Hence, Markov's inequality implies that w.h.p. $\sum_{1 \leq q \leq \chi n} X_q = 0$, in which case (10) holds. ∎

3. Analyzing the decimation process. Here we perform some groundwork to facilitate our analysis of the decimation process. The goal is to get a handle on the following experiment:

D1. Generate a random formula $\mathbf{\Phi}$, conditioned on $\mathbf{\Phi}$ being satisfiable.
D2. Run the decimation process for t steps to obtain $\mathbf{\Phi}_t$.
D3. Choose a satisfying assignment $\boldsymbol{\sigma}_t \in \mathcal{S}(\mathbf{\Phi}_t)$ uniformly at random.
D4. The result is the pair $(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$.

Because throughout the paper we will only work with densities m/n where $\mathbf{\Phi}$ is satisfiable w.h.p., the conditioning in step D1 is essentially void. Recalling that the outcome of the decimation process is a uniformly random satisfying assignment of the input formula $\mathbf{\Phi}$, we see that the following experiment is equivalent to D1–D4:

U1. Generate a random formula $\mathbf{\Phi}$, conditioned on $\mathbf{\Phi}$ being satisfiable.
U2. Choose $\boldsymbol{\sigma} \in \mathcal{S}(\mathbf{\Phi})$ uniformly at random.
U3. Substitute $\boldsymbol{\sigma}(x_i)$ for x_i for $1 \le i \le t$ and simplify to obtain $\mathbf{\Phi}_t = \mathbf{\Phi}_{t,\boldsymbol{\sigma}}$.
U4. The result is the pair $(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$, where $\boldsymbol{\sigma}_t : V_t \to \{0,1\},\ x \mapsto \boldsymbol{\sigma}(x)$.

FACT 12. *The two probability distributions induced on formula/assignment pairs by the two experiments* D1–D4 *and* U1–U4 *are identical.*

Still, an analysis of U1–U4 seems difficult because of U2: it is unclear how to analyze (or implement) this step directly. Following [2], we will surmount this problem by considering yet another experiment.

P1. Choose an assignment $\boldsymbol{\sigma}' \in \{0,1\}^V$ uniformly at random.
P2. Choose a formula $\mathbf{\Phi}'$ with m clauses that is satisfied by $\boldsymbol{\sigma}'$ uniformly at random.
P3. Substitute $\boldsymbol{\sigma}'(x_i)$ for x_i for $1 \le i \le t$ and simplify to obtain $\mathbf{\Phi}'_t = \mathbf{\Phi}'_{t,\boldsymbol{\sigma}'}$.
P4. The result is the pair $(\mathbf{\Phi}'_t, \boldsymbol{\sigma}'_t)$, where $\boldsymbol{\sigma}'_t : V_t \to \{0,1\},\ x \mapsto \boldsymbol{\sigma}'(x)$.

The distribution P1–P4 is easy to sample from and, in effect, it is also amenable to a rigorous analysis. For given the assignment $\boldsymbol{\sigma}'$, there are $(2^k-1)\binom{n}{k}$ clauses in total that evaluate to true under $\boldsymbol{\sigma}'$, and to generate $\mathbf{\Phi}'$ we merely choose m out of these uniformly and independently. Unfortunately, it is *not* true that the experiment P1–P4 is equivalent to U1–U4. However, we will employ a result from [2] that establishes a connection between these two experiments that is strong enough to extend many results from P1–P4 to U1–U4.

To state this result, observe that P1–P4 and U1–U4 essentially only differ in their first two steps. Thus, let $\Lambda_k(n,m)$ denote the set of all pairs (Φ,σ), where Φ is a k-CNF on $V = \{x_1,\ldots,x_n\}$ with m clauses, and $\sigma \in \mathcal{S}(\Phi)$. Let $\mathcal{U}_k(n,m)$ denote the probability distribution induced on $\Lambda_k(n,m)$ by U1–U2, and let $\mathcal{P}_k(n,m)$ signify the distribution induced by P1–P2; this distribution is sometimes called the *planted model*.

THEOREM 13 (see [2]). *Suppose $k \ge 4$ and $0 < \rho < k\ln 2 - k^2/2^k$. Let $\mathcal{E} \subset \Lambda_k(n,m)$. If $\mathrm{P}_{\mathcal{P}_k(n,m)}[\mathcal{E}] \ge 1 - \exp(-\rho n/2^k)$, then $\mathrm{P}_{\mathcal{U}_k(n,m)}[\mathcal{E}] = 1 - o(1)$.*

We will need the following consequence of Theorem 9.

COROLLARY 14. *Let $1 \le t \le n$. Let $(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$ be a pair chosen from the experiment* U1–U4. *Then w.h.p.*

$$\frac{1}{n}\ln|\mathcal{S}(\mathbf{\Phi}_t)| \ge \theta\ln 2 + r\ln(1-2^{-k}) - \frac{\rho}{2^k}. \tag{14}$$

Proof. Theorem 9 shows that $\frac{1}{n}\ln|\mathcal{S}(\mathbf{\Phi})| \ge \ln 2 + r\ln(1-2^{-k}) - 0.99\rho/2^k$ w.h.p. Thus, it suffices to show that for *any fixed* formula Φ such that

$$\frac{1}{n}\ln|\mathcal{S}(\Phi)| \ge \ln 2 + r\ln(1-2^{-k}) - 0.99\rho/2^k \tag{15}$$

for almost all $\sigma \in \mathcal{S}(\Phi)$ we have

$$\frac{1}{n}\ln|\mathcal{S}(\Phi_{t,\sigma})| \ge \theta\ln 2 + r\ln(1-2^{-k}) - \frac{\rho}{2^k}.$$

To show this, let $\mathcal{I} = \{0,1\}^t$. For each $\sigma \in \{0,1\}^n$ let $\sigma|_t$ denote the vector $(\sigma(x_1),\ldots,\sigma(x_t)) \in \mathcal{I}$. For each $\sigma_* \in \mathcal{I}$ let $Z(\sigma_*)$ be the number of assignments

$\sigma \in \mathcal{S}(\Phi)$ such that $\sigma|_t = \sigma_*$. Then our goal is to show that $\frac{1}{n}\ln Z(\sigma|_t) \geq \theta \ln 2 + r\ln(1-2^{-k}) - \frac{\rho}{2^k}$ for almost all $\sigma \in \mathcal{S}(\Phi)$. Thus, let

$$q = \mathrm{P}_{\sigma\in\mathcal{S}(\Phi)}\left[\frac{1}{n}\ln Z(\sigma|_t) < \theta\ln 2 + r\ln(1-2^{-k}) - \frac{\rho}{2^k}\right].$$

If $\sigma \in \mathcal{S}(\Phi)$ is chosen uniformly at random, then for any $\sigma_* \in \mathcal{I}$ we have

$$\mathrm{P}\left[\sigma|_t = \sigma_*\right] = Z(\sigma_*)/Z, \text{ where } Z = \sum_{\tau\in\mathcal{I}} Z(\tau) = |\mathcal{S}(\Phi)|.$$

Therefore, by (15)

$$q = \sum_{\sigma_*\in\mathcal{I}: Z(\sigma_*)\leq Z/(2^t\exp(0.01\rho n/2^k))} \frac{Z(\sigma_*)}{Z} \leq \frac{2^t}{Z}\cdot\frac{Z}{2^t\exp(0.01\rho n/2^k)} \leq \exp(-\Omega(n)),$$

as claimed. □

In addition to the experiment P1–P4, which led to the planted model $\mathcal{P}_k(n,m)$, we need the following variant:

P1′. Choose an assignment $\boldsymbol{\sigma}' \in \{0,1\}^V$ uniformly at random.

P2′. Choose a formula $\boldsymbol{\Phi}'$ by including each of the $(2^k-1)\binom{n}{k}$ possible clauses that are satisfied under $\boldsymbol{\sigma}'$ with probability $p = m/((2^k-1)\binom{n}{k})$ independently.

P3′. Substitute $\boldsymbol{\sigma}'(x_i)$ for x_i for $1 \leq i \leq t$ and simplify to obtain a formula $\boldsymbol{\Phi}'_t = \boldsymbol{\Phi}'_{t,\boldsymbol{\sigma}'}$.

P4′. The result is the pair $(\boldsymbol{\Phi}'_t, \boldsymbol{\sigma}'_t)$, where $\boldsymbol{\sigma}'_t : V_t \to \{0,1\},\ x \mapsto \boldsymbol{\sigma}'(x)$.

Steps P1′–P2′ of this experiment induce a probability distribution $\mathcal{P}'_k(n,m)$ on formula/assignment pairs. The following corollary establishes a connection between this distribution and $\mathcal{U}_k(n,m)$.

COROLLARY 15 (see [2]). *Suppose that $k \geq 4$ and $0 < r < 2^k\ln 2 - k$. Let $\mathcal{E}$ be any property of formula/assignment pairs. If $\mathrm{P}_{\mathcal{P}'_k(n,m)}[\mathcal{E}] \geq 1 - \exp(-\rho n/2^k)$, then $\mathrm{P}_{\mathcal{U}_k(n,m)}[\mathcal{E}] = 1 - o(1)$.*

Let Φ be a propositional formula, let $\sigma \in \mathcal{S}(\Phi)$, and let l be a literal. We say that l *supports* a clause C of Φ under σ if l is the only literal of C that is true under σ. We will need the following elementary observation about the distribution $\mathcal{P}'_k(n,m)$.

LEMMA 16. *Let $(\boldsymbol{\Phi}', \boldsymbol{\sigma}')$ be a pair chosen from the distribution $\mathcal{P}'_k(n,m)$.*

1. *For each literal l that is true under $\boldsymbol{\sigma}'$ the number of clauses supported by l is binomially distributed* $\mathrm{Bin}(\frac{k}{n}\cdot\binom{n}{k}, m/((2^k-1)\binom{n}{k}))$.
2. *For any integer D the number of literals l that support fewer than D clauses is binomially distributed with mean*

$$n\cdot\mathrm{P}\left[\mathrm{Bin}\left(\frac{k}{n}\cdot\binom{n}{k}, \frac{m}{(2^k-1)\binom{n}{k}}\right) < D\right]. \tag{16}$$

Proof. Without loss of generality we may condition on $\boldsymbol{\sigma}'$ assigning the value true to all variables. For any variable x let $\mathcal{S}_x$ be the set of all possible clauses in which x is the only positive literal. Then $|\mathcal{S}_x| = \binom{n-1}{k-1} = \frac{k}{n}\cdot\binom{n}{k}$. (Choose the $k-1$ other variables occurring in the clause; the signs are prescribed by x being the unique positive literal.) Moreover, let S_x be the number of clauses from $\mathcal{S}_x$ that actually appear in the random formula F. As each of the clauses in $\mathcal{S}_x$ is included in F with probability $p = m/((2^k-1)\binom{n}{k})$ independently, S_x has a binomial distribution $\mathrm{Bin}(\frac{k}{n}\cdot\binom{n}{k}, p)$. This establishes 1.

Since for any two variables x, y we have $\mathcal{S}_x \cap \mathcal{S}_y = \emptyset$, the random variables S_x are mutually independent for all variables x. Therefore, the number $S = \sum_x 1_{\{S_x < D\}}$ of variables supporting fewer than D clauses in F is binomially distributed as well. ∎

4. Loose, rigid, and forced variables.

4.1. Overview. In this section we prove the statements made in Theorems 2–5 on loose/rigid/forced variables. The following theorem summarizes them.

THEOREM 17. *There exist constants $k_0, \rho_0 > 0$ such that for all $k \geq k_0$ and $\rho_0 \leq \rho \leq k \ln 2 - 2 \ln k$ the following three statements hold for a random pair $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ chosen from the experiment* U1–U4 *w.h.p.:*

1. *If $k\theta > \exp[\rho(1 + \frac{\ln\ln\rho}{\rho} + \frac{10}{\rho})]$, then at least $0.99\theta n$ variables $x \in V_t$ are loose w.h.p.*
2. *If $1 < k\theta < \exp[\rho(1 - \frac{3\ln\rho}{\rho})]$, then at least $(1 - \frac{\rho^3}{\exp(\rho)})\theta n$ variables $x \in V_t$ are $\Omega(n)$-rigid w.h.p.*
3. *If $\ln(n)/n < \theta < (\ln(\rho) - 10)/k$, then at least $0.99\theta n$ variables are forced w.h.p.*

Let $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ be the (random) outcome of U1–U4; let us assume without loss of generality that $\boldsymbol{\sigma}_t$ is the all-true assignment. The notions of loose/rigid/forced describe how easy or difficult it is to "repair" the damage done by flipping the value assigned to a single variable x from $\boldsymbol{\sigma}_t(x)$ to $1 - \boldsymbol{\sigma}_t(x)$ (i.e., to satisfy the resulting unsatisfied clauses). To get a combinatorial intuition, observe that we can simply assign x the opposite value $1 - \boldsymbol{\sigma}_t(x)$, unless $\boldsymbol{\Phi}_t$ has a clause $\mathcal{C}$ in which x underlies the *only* literal that is true under $\boldsymbol{\sigma}_t$. If there is such a clause $\mathcal{C}$, we say that x *supports* $\mathcal{C}$.

But even if x supports a clause $\mathcal{C}$, x might be easy to flip. For instance, if $\mathcal{C}$ features some variable $y \neq x$ that does not support any clause, then we could just flip both x, y simultaneously. Thus, to understand for what θ, ρ variables are rigid we need to analyze the distribution of the number of clauses that a variable supports, the probability that these clauses only consists of variables that support further clauses, the probability that the same is true of those clauses, etc.

It does not seem easy to analyze this kind of property directly on the random outcome $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ of U1–U4. By contrast, as we will see the analysis can be carried out quite neatly on the random pair $(\boldsymbol{\Phi}'_t, \boldsymbol{\sigma}'_t)$ generated by the experiment P1–P4. Hence, we will work with $(\boldsymbol{\Phi}'_t, \boldsymbol{\sigma}'_t)$ and then extend the result to $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ via Theorem 13.

For a variable $x \in V_t$ let S_x be the number of clauses supported by x. Given the assignment $\boldsymbol{\sigma}'$ chosen in step P1, there are a total of $\binom{n-1}{k-1}$ possible clauses that x supports. Since in step P2 we include m out of the $(2^k - 1)\binom{n}{k}$ possible clauses satisfied under $\boldsymbol{\sigma}'$ uniformly and independently, we get

$$\mathrm{E}[S_x] = m\binom{n-1}{k-1}\left((2^k - 1)\binom{n}{k}\right)^{-1} = \rho/(1 - 2^{-k}) > \rho.$$

In fact, S_x is binomially distributed. Hence, $\mathrm{P}[S_x = 0] \leq \exp(-\rho)$. Thus, the *expected* number of variables $x \in V_t$ with $S_x = 0$ is $\leq \theta n \exp(-\rho)$. Furthermore, if we condition on $S_x = j \geq 1$, then the actual *clauses* $\mathcal{C}_1, \ldots, \mathcal{C}_j$ supported by x are just uniformly distributed over the set of all $\binom{n-1}{k-1}$ possible clauses that x supports. Therefore, the *expected* number of variables $y \in V_t$ with $S_y = 0$ occurring in any one of these clauses $\mathcal{C}_i$ is

$$(k-1) \cdot \theta \exp(-\rho) + o(1). \tag{17}$$

Hence, (17) suggests that for $\theta \gg \exp(\rho)/(k-1)$, there is a good chance that $\mathcal{C}_i$ might contain a variable y that does not support another clause, and hence that x is loose. Indeed, this observation provides the starting point for the proof of the first part of Theorem 17 in section 4.2. By contrast, if $(k-1)\cdot\theta\exp(-\rho) \ll 1$, then (17) indicates that $\mathcal{C}_i$ is unlikely to contain another nonsupporting variable. In section 4.3 below we will turn this into a proof of the second part of Theorem 17. Finally, section 4.4 contains the proof of the existence of forced variables.

4.2. Loose variables. Let σ be a satisfying assignment of a k-CNF Φ. Remember that a literal l *supports* a clause C of Φ if l is the only literal in C that is true under σ. Moreover, we say that a literal l is 1-*loose* if it is true under σ and supports no clause. In addition, l is 2-*loose* if l is true under σ and each clause that l supports contains a 1-loose literal from L_t. Thus, any 1-loose literal is 2-loose as well. The key step of the proof is to establish the following.

PROPOSITION 18. *Suppose that $\theta \geq 3\exp(\rho)(\ln\rho+10)/k$ and $r \leq 2^k\ln 2 - k$. Let $(\boldsymbol{\Phi},\boldsymbol{\sigma})$ be a random pair chosen from the distribution $\mathcal{U}_k(n,m)$. Then there are at least $0.999\theta n$ 2-loose literals in L_t w.h.p.*

To obtain Theorem 2 from Proposition 18, we will show that any 2-loose literal is loose w.h.p. (Of course, in principle one could generalize this argument by defining/studying t-loose variables for $t > 2$. However, to prove Theorem 2 it is sufficient to consider 2-loose literals.) To prove Proposition 18, we start by estimating the number of 1-loose literals.

LEMMA 19. *Suppose that $\theta \geq \exp(\rho)/k$ and $\rho \leq k\ln 2$. Let $(\boldsymbol{\Phi}',\boldsymbol{\sigma}')$ be a random pair chosen from the distribution $\mathcal{P}'_k(n,m)$. With probability at least $1-\exp(-k2^{2-k}n)$ the number of 1-loose literals in L_t is at least $\theta n\cdot\exp(-\rho)/2$.*

Proof. By Lemma 16 the number X of 1-loose literals in L_t has a binomial distribution with mean

$$\begin{aligned}\mathrm{E}[X] &= \theta n\cdot\mathrm{P}\left[\mathrm{Bin}\left(\frac{k}{n}\cdot\binom{n}{k},\frac{m}{(2^k-1)\binom{n}{k}}\right)=0\right]\\ &= \theta n\cdot\left(1-\frac{m}{(2^k-1)\binom{n}{k}}\right)^{\frac{k}{n}\cdot\binom{n}{k}} \sim \theta n\cdot\exp\left(-\frac{kr}{2^k-1}\right)\\ &= \theta n\exp(-\rho-\rho/(2^k-1)) \geq 0.99\theta n\exp(-\rho) \quad \text{as } \rho\leq k\ln 2.\end{aligned}$$

Hence, the Chernoff bound (8) shows that

$$\begin{aligned}\mathrm{P}[X<\theta n\exp(-\rho)/2] &\leq \mathrm{P}[X<\mathrm{E}[X]-0.49\theta n\exp(-\rho)]\\ &\leq \exp[-\mathrm{E}[X]\cdot\phi(-0.49\theta n\exp(-\rho)/\mathrm{E}[X])]\\ &\leq \exp\left[-\frac{(0.49\theta n\exp(-\rho))^2}{2\mathrm{E}[X]}\right]\\ &\leq \exp\left[-\frac{(0.49)^2\theta\exp(-\rho)n}{2\cdot 0.99}\right]\\ &\leq \exp[-\theta\exp(-\rho)n/9]\\ &\leq \exp[-n/(9k)] \qquad [\text{since } \theta\geq\exp(\rho)/k]\\ &\leq \exp(-k2^{2-k}n),\end{aligned}$$

as desired. ∎

With Lemma 19 in place, we can now estimate the number of 2-loose literals in a pair chosen from the planted distribution $\mathcal{P}'_k(n,m)$.

LEMMA 20. *Suppose that* $\theta \geq 3\exp(\rho)(\ln\rho+10)/k$ *and that* $\rho \leq k\ln 2$. *Let* $(\boldsymbol{\Phi}', \boldsymbol{\sigma}')$ *be a random pair chosen from the distribution* $\mathcal{P}'_k(n,m)$. *Then with probability at least* $1-\exp(-k2^{1-k}n)$ *the number of* 2*-loose literals in* L_t *is at least* $0.999\theta n$.

Proof. To simplify the notation, we are going to condition on $\boldsymbol{\sigma}'$ being the all-true assignment; this is without loss of generality (by symmetry). For each variable $x \in V_t$ we let S_x be the number of clauses supported by x. Moreover, let $S = \sum_{x\in V_t} S_x$ and let X be the number of variables $x \in V_t$ such that $S_x = 0$. Thus, X equals the number of 1-loose variables.

Let $\mathcal{E}$ be the event that $X \geq \theta n\exp(-\rho)/2$ and $S \leq 2\rho\theta n$. Since the number of possible clauses with precisely one positive literal in L_t is $\theta n\binom{n-1}{k-1}$, S has a binomial distribution, i.e.,

$$S \sim \mathrm{Bin}\left[\theta n\binom{n-1}{k-1}, \frac{m}{(2^k-1)\binom{n}{k}}\right].$$

Therefore, Lemma 19 implies that

$$\begin{aligned}\mathrm{P}\left[\neg\mathcal{E}\right] &\leq \mathrm{P}\left[X < \theta n\exp(-\rho)/2\right] + \mathrm{P}\left[S > 2\rho\theta n\right]\\ &\leq \exp\left[-k2^{2-k}n\right] + \mathrm{P}\left[\mathrm{Bin}\left(\theta n\binom{n-1}{k-1}, \frac{m}{(2^k-1)\binom{n}{k}}\right) > 2\rho\theta n\right]. \qquad (18)\end{aligned}$$

We have

$$\theta n\binom{n-1}{k-1}\cdot\frac{m}{(2^k-1)\binom{n}{k}} = \frac{2^k}{2^k-1}\cdot\rho\theta n.$$

Hence, combining (18) with the Chernoff bound (8), we obtain for sufficiently large k

$$\mathrm{P}\left[\neg\mathcal{E}\right] \leq \exp\left[-k2^{2-k}n\right] + \exp\left[-0.99\rho\theta n\right] \leq 2\exp\left[-k2^{2-k}n\right], \qquad (19)$$

where in the last step we used the assumption that $\rho \geq \rho_0$ for a fixed constant $\rho_0 > 0$.

Let us now condition on the event that $S = s$ for some number $s \leq 2\rho\theta n$, and on the event $\mathcal{E}$. In this conditional distribution for each of the s clauses supported by some variable in V_t the $k-1$ negative literals are just drawn uniformly without replacement, independently for the s clauses. Therefore, for each such clause the probability to not contain a negative literal $\bar{y}$ whose underlying variable y is 1-loose is $(1+o(1))(1-X/n)^{k-1}$. Consequently, the number T of clauses supported by some variable in V_t in which no 1-loose variable occurs negatively has a binomial distribution with mean $(1+o(1))s(1-X/n)^{k-1}$. Hence,

$$\mathrm{E}\left[T|\mathcal{E}\right] \leq 2\rho\theta n\cdot(1-\theta\exp(-\rho)/2)^{k-1} \leq 2\rho\theta n\exp(-\theta\exp(-\rho)k/3) \leq 2\exp(-10)\theta n.$$

Thus, the Chernoff bound (8) implies that for $k \geq k_0$ large enough

$$\mathrm{P}\left[T > 0.001\theta n|\mathcal{E}\right] \leq \exp(-0.001\theta n) \leq \exp\left[-k2^{2-k}n\right]. \qquad (20)$$

Finally, the assertion follows from (19) and (20). ∎

Proof of Proposition 18. Let $\mathcal{E}$ be the event that a pair $(\Phi, \sigma) \in \Lambda_{n,m}$ has at least $0.999\theta n$ 2-loose literals. Lemma 20 shows that

$$\mathrm{P}_{\mathcal{P}'_k(n,m)}[\mathcal{E}] \geq 1 - \exp(-k2^{1-k}n) \geq 1 - \exp(-\rho n/2^k). \tag{21}$$

Hence, Corollary 15 and (21) imply that $\mathrm{P}_{\mathcal{U}_k(n,m)}[\mathcal{E}] = 1 - o(1)$ as desired. ∎

Proof of Theorem 17, *part* 1. By Fact 12 it suffices to prove the desired statement for the experiment U1–U4. Thus, let $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ be a pair chosen from the distribution $\mathcal{U}_k(n,m)$. Without loss of generality we may condition on $\boldsymbol{\sigma}$ being the all-true assignment. Let $\mathcal{L}$ be the set of all tame variables in V_t that are 2-loose. Then by Propositions 10 and 18 we have $|\mathcal{L}| \geq (0.999 - o(1))\theta n \geq 0.99\theta n$ w.h.p. Assuming that this is indeed the case, we are going to show that if $x \in \mathcal{L}$, then there is a satisfying assignment τ such that $\tau(x) \neq \boldsymbol{\sigma}(x)$ and $\mathrm{dist}(\tau, \boldsymbol{\sigma}) \leq \ln(n)$.

Thus, fix a variable $x \in \mathcal{L}$. If x is 1-loose, then we can just set $\tau(x) = 1 - \boldsymbol{\sigma}(x) = 0$ and $\tau(y) = \boldsymbol{\sigma}(y) = 1$ for all $y \neq x$ to obtain a satisfying assignment with $\mathrm{dist}(\tau, \boldsymbol{\sigma}) = 1$, because x does not support any clauses. Hence, assume that x is 2-loose but not 1-loose. Let $\mathcal{C}$ be the set of all clauses supported by x in $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$. Any clause $C \in \mathcal{C}$ contains a negative occurrence of a 1-loose variable $x_C \in V_t$ in C (by the very definition of 2-loose). Define $\tau(x) = 0$, $\tau(x_C) = 0$ for all $C \in \mathcal{C}$, and $\tau(y) = \boldsymbol{\sigma}(y) = 1$ for all other variables y.

We claim that τ is a satisfying assignment. To see this, assume for contradiction that there is a clause U that is unsatisfied under τ. Then U contains a variable from $\{x\} \cup \{x_C : C \in \mathcal{C}\}$ positively, while none of these variables occurs negatively in U. Hence, $U \notin \mathcal{C}$. Moreover, since the variables x_C, $C \in \mathcal{C}$, do not support any clauses, U indeed contains two variables from the set $\{x\} \cup \{x_C : C \in \mathcal{C}\}$ positively. There are two possible cases.

Case 1: *x occurs in U*. Let $C \in \mathcal{C}$ such that x_C occurs in U as well. Then the factor graph contains the cycle x, C, x_C, U, x, in contradiction to our assumption that x is tame.

Case 2: *x does not occur in U*. There exist $C_1, C_2 \in \mathcal{C}$ such that x_{C_1}, x_{C_2} occur in U. Hence, the factor graph contains the cycle $x, C_1, x_{C_1}, U, x_{C_2}, C_2, x$, once more in contradiction to the assumption that x is tame.

Hence, there is no clause U that is unsatisfied under τ. Finally, since all the variable x_C with $C \in \mathcal{C}$ have distance two from x in the factor graph, and as x is tame, we have $\mathrm{dist}(\sigma, \tau) \leq \ln n$. ∎

4.3. Rigid variables. Let $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ be chosen from the distribution $\mathcal{U}_k(n,m)$ and assume that $\boldsymbol{\sigma}$ is the all-true assignment. As we saw in section 4.1, in order to show that for $\theta \ll \exp(\rho)/k$ most variables are $\Omega(n)$-rigid, we need to show that most variables support at least a few clauses that consist exclusively of variables that support other clauses. To get started, we prove the following lower bound on the number of literals that support at least three clauses. Recall from (9) that $\varphi(x) = (1+x)\ln(1+x) - x$.

Proposition 21. *Suppose that $k \geq 6$ and $0 < r \leq 2^k \ln 2 - k$. Let*

$$\mu = \rho \cdot 2^k/(2^k - 1) \text{ and } \zeta = (1 + \mu + \mu^2/2)/\exp(\mu),$$

and assume that $2^k\theta\zeta\varphi(1) > \rho$. Then w.h.p. in a random pair $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ chosen from the distribution $\mathcal{U}_k(n,m)$ no more than $2\zeta\theta n$ literals in L_t support fewer than three clauses.

Proof. Let S be the number of literals $l \in L_t$ that support fewer than three clauses. We are going to show that

$$\mathrm{P}_{\mathcal{P}'_k(n,m)}[S > 2\zeta\theta n] \leq \exp(-\rho n/2^k). \tag{22}$$

Then Corollary 15 implies the assertion.

By the second part of Lemma 16, in the distribution $\mathcal{P}'_k(n,m)$ the random variable S is binomially distributed with mean

$$\begin{aligned}\mathrm{E}[S] &= |L_t| \cdot \mathrm{P}\left[\mathrm{Bin}\left(\frac{k}{n}\binom{n}{k}, \frac{m}{(2^k-1)\binom{n}{k}}\right) < 3\right] \\ &\sim \theta n \cdot \mathrm{P}\left[\mathrm{Po}\left(\frac{k}{2^k-1} \cdot \frac{m}{n}\right) < 3\right] \\ &= \theta n \mathrm{P}[\mathrm{Po}(\mu) < 3] = \theta\zeta n.\end{aligned}$$

Hence, the Chernoff bound (8) shows that

$$\mathrm{P}_{\mathcal{P}'_k(n,m)}[S > 2\zeta n] \leq \exp\left(-(1+o(1))\theta\zeta\varphi(1)n\right). \tag{23}$$

By the assumptions on μ and θ we have $\theta\zeta\varphi(1) > \rho/2^k$. Hence, (22) follows from (23). □

Let $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ be chosen from the distribution $\mathcal{U}_k(n,m)$ and assume that $\boldsymbol{\sigma}$ is the all-true assignment. Even if many variables in V_t support several clauses, this does not necessarily mean that these variables are $\Omega(n)$-rigid, because these clauses might contain other variables from V_t that do not support any clauses (cf. section 4.1). However, as we shall see next, if there is a set $\mathcal{S}$ of variables that each support at least two clauses that each consist of variables from $\mathcal{S}$ only, then indeed the variables in $\mathcal{S}$ are $\Omega(n)$-rigid w.h.p.

To be precise, let $(\Phi, \sigma) \in \Lambda_k(n,m)$ be a formula/satisfying assignment pair. We say that a set $\mathcal{S} \subset L_t$ of literals is *t-self-contained* if each literal $l \in \mathcal{S}$ is true under σ and supports at least two clauses that contain literals from $\{x_1, \bar{x}_1, \ldots, x_t, \bar{x}_t\} \cup \mathcal{S} \cup \bar{\mathcal{S}}$ only, where $\bar{\mathcal{S}}$ is the set of all negations of literals in $\mathcal{S}$.

Proposition 22. *For any $k \geq 3$ there is a number $\chi = \chi(k) > 0$ such that for any $0 < r \leq 2^k \ln 2 - k$ the following is true. Let $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ be a random pair chosen from the distribution $\mathcal{U}_k(n,m)$. Then w.h.p. for any $0 \leq t \leq n$, in any t-self-contained set $\mathcal{S}$ all variables $x \in \mathcal{S} \cup \bar{\mathcal{S}}$ are χn-rigid.*

Proof. Let $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ be a random pair chosen from the distribution $\mathcal{U}_k(n,m)$. Without loss of generality we may condition on σ being the all-true assignment. By Lemma 11 there is a number $\chi = \chi(k) > 0$ such that (10) is satisfied w.h.p., and we are going to assume that this is the case.

Let $\mathcal{S}$ be a t-self-contained set. Suppose that τ is a satisfying assignment such that the set Q of all variables $x \in \mathcal{S} \cup \bar{\mathcal{S}}$ such that $\tau(x) \neq \sigma(x)$ is nonempty. For each variable $x \in Q$ there are two clauses $C_1(x), C_2(x)$ that are supported by x in $\boldsymbol{\sigma}$ and that consist of literals from $\{x_1, \bar{x}_1, \ldots, x_t, \bar{x}_t\} \cup \mathcal{S} \cup \bar{\mathcal{S}}$ only (because $\mathcal{S}$ is self-contained). Since τ is satisfying and $\tau(x) \neq \sigma(x)$, both $C_1(x)$ and $C_2(x)$ contain another variable from Q. Hence, there are at least $2|Q|$ clauses that contain at least two variables from Q. Thus, (10) implies that $|Q| > \chi n$, and consequently $\mathrm{dist}(\sigma, \tau) \geq |Q| > \chi n$. □

Now we will show that under certain assumptions on the parameters ρ, θ, a random pair $(\boldsymbol{\Phi}, \sigma)$ chosen from $\mathcal{U}_k(n,m)$ has a large t-self-contained set w.h.p.

PROPOSITION 23. *Suppose that $k \geq 4$ and $0 < r \leq 2^k \ln 2 - k$ and that $0 \leq \theta \leq 1$. Set*

$$\mu = \frac{\rho 2^k}{2^k - 1},\ \zeta = \frac{1 + \mu + \mu^2/2}{\exp(\mu)},\ \lambda = 1 - (1 - 3\theta\zeta)^{k-1},\ \gamma = \frac{\mu \cdot (\exp(\lambda\mu) - 1 - \lambda\mu)}{(1-\zeta)\exp(\mu)}$$

and let $h(x) = -x \ln x - (1-x)\ln(1-x)$. If $\zeta < 1/3$,

$$2^k \theta\zeta\varphi(1) > \rho,\ \text{and } \theta(\zeta \ln(\gamma) + h(\zeta)) + \rho/2^k < 0, \tag{24}$$

then a random pair $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ chosen from $\mathcal{U}_k(n, m)$ has a t-self-contained set of size $\geq (1 - 3\zeta)\theta n$ w.h.p.

Proof. Consider a random pair $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$; without loss of generality, suppose that $\boldsymbol{\sigma}$ is the all-true assignment. In order to construct a t-self-contained set, we could run the following process, somewhat reminiscent of the "whitening" from [6]. First, let $\mathcal{Z}$ be the set of all variables that support fewer than three clauses. Then, we add to $\mathcal{Z}$ all variables $x \in V_t \setminus \mathcal{Z}$ that do not support at least two clauses. Finally, we remove all clauses that contain a variable from $\mathcal{Z}$. This process is repeated until no further variables are added to $\mathcal{Z}$. Clearly, the set of all variables $x \in V_t \setminus \mathcal{Z}$ that support at least two clauses after this process halts is t-self-contained.

The basic idea behind the proof is to analyze the above process, i.e., to estimate the number of variables remaining that support at least two clauses. To carry this analysis out, we will work with the planted model. Let $p = m/((2^k - 1)\binom{n}{k})$. Let $(\boldsymbol{\Phi}', \boldsymbol{\sigma}')$ be chosen from the distribution $\mathcal{P}'_k(n, m)$. Without loss of generality we may condition on $\boldsymbol{\sigma}'$ being the all-true assignment. Furthermore, fix a set $Z \subset V_t$ of at most $2\zeta\theta n$ variables and condition on the event $\mathcal{E}_Z$ that Z is the set of all variables in V_t that support at most two clauses (cf. Proposition 21).

The analysis of the process depends on two probabilistic phenomena. The first of these is the distribution of the number of clauses supported by variables $x \notin Z$. Clearly, the more clauses a variable $x \notin Z$ supports, the more likely it is that this variable will "survive" the above process (i.e., that it still supports two or more clauses in the end). The second phenomenon is the probability that a random clause supported by a variable $x \notin Z$ gets removed, i.e., that it contains a variable from $\mathcal{Z}$.

To study the first phenomenon, for any variable $x \notin Z$ we let S_x be the number of clauses supported by x. Then the first part of Lemma 16 implies that S_x has a binomial distribution $\mathrm{Bin}(\binom{n-1}{k-1}, p)$, conditioned on the outcome being at least three. As a consequence, for any $j \geq 3$ we have

$$\begin{aligned}
\mathrm{P}_{\mathcal{P}'_k(n,m)}[S_x = j | \mathcal{E}_Z] &= \frac{\mathrm{P}\left[\mathrm{Bin}(\binom{n-1}{k-1}, p) = j\right]}{\mathrm{P}\left[\mathrm{Bin}(\binom{n-1}{k-1}, p) \geq 3\right]} \leq \frac{1 + o(1)}{1 - \zeta} \cdot \binom{\binom{n-1}{k-1}}{j} p^j (1-p)^{\binom{n-1}{k-1} - j} \\
&\leq \frac{1 + o(1)}{1 - \zeta} \cdot \frac{1}{j!} \left(\binom{n-1}{k-1} p\right)^j \exp\left[-p\left(\binom{n-1}{k-1} - j\right)\right] \\
&\leq \frac{(1 + o(1))\mu^j}{j! \exp(\mu)(1-\zeta)} \cdot \exp(jp). \qquad (25)
\end{aligned}$$

To study the second phenomenon, let $X \subset V_t \setminus Z$ be a set of $\zeta\theta n$ variables. For each $x \in X$ we let $T_x(X)$ be the number of clauses supported by x in which a variable from $X \cup Z$ occurs negatively. (Hence, the set $X \cup Z$ corresponds to the $\mathcal{Z}$ in the above description of the process.) In a random clause supported by x the variables

underlying the $k-1$ negative literals in that clause are distributed uniformly over $V = \{x_1, \ldots, x_n\}$. Therefore, given $\mathcal{E}_Z$ the probability that such a clause contains at least one variable from $X \cup Z$ is

$$1 - (1 - |X \cup Z|/n)^{k-1} + o(1) \leq 1 - (1 - 3\theta\zeta)^{k-1} + o(1) \sim \lambda.$$

Hence, if we condition on both $\mathcal{E}_Z$ and $S_x = j$, then the probability that $T_x(X) \geq S_x - 1$ is at most $j \cdot (\lambda + o(1))^{j-1}$. Thus, letting $\gamma = \mu \cdot (\exp(\lambda\mu) - 1 - \lambda\mu)/(1-\zeta)\exp(\mu)$, we obtain from (25)

$$\begin{aligned}\mathrm{P}_{\mathcal{P}'_k(n,m)}[T_x(X) \geq S_x - 1 | \mathcal{E}_Z] &= \sum_{j \geq 3} \mathrm{P}[T_x(X) \geq S_x - 1 | \mathcal{E}_Z \text{ and } S_x = j] \cdot \mathrm{P}[S_x = j | \mathcal{E}_Z] \\ &\leq (1 + o(1)) \sum_{j \geq 3} \frac{j\lambda^{j-1}\mu^j \exp(jp)}{j! \exp(\mu)(1-\zeta)} \sim \gamma.\end{aligned}$$

Given that $\mathcal{E}_Z$ occurs the events $T_x(X) \geq S_x - 1$ are mutually independent for all $x \in X$ (because each clause satisfied under σ is present in $\mathbf{\Phi}'$ with probability p independently). Therefore,

$$\mathrm{P}_{\mathcal{P}'_k(n,m)}[\forall x \in X : T_x(X) \geq S_x - 1 | \mathcal{E}] \leq (\gamma + o(1))^{\theta\zeta n}. \tag{26}$$

Let $\mathcal{S}_Z$ denote the event that there is no t-self-contained set $\mathcal{S} \subset V_t \setminus Z$ such that $|\mathcal{S}| \geq |V_t \setminus Z| - \zeta\theta n \geq (1 - 3\zeta)\theta n$.

If $\mathcal{E}_Z$ occurs but $\mathcal{S}_Z$ does not, then there is a set $X \subset V_t \setminus Z$ of size ζn such that $T_x(X) \geq S_x - 1$ for all $x \in X$. To see this, consider the following formal description of our process. Initially, let $X = \emptyset$. Then, while there is a variable $x \notin X \cup Z$ such that $T_x(X) \geq S_x - 1$, add x to X. If a t-self-contained set $\mathcal{S} \subset V_t \setminus Z$ of size $|\mathcal{S}| \geq |V_t \setminus Z| - \zeta\theta n$ exists, then this process will stop with a set $X \subset V_t \setminus (\mathcal{S} \cup Z)$ of size $|X| \leq |V_t \setminus (\mathcal{S} \cup Z)| < \zeta\theta n$. Hence, by the union bound and (26)

$$\begin{aligned}&\mathrm{P}_{\mathcal{P}'_k(n,m)}[\neg\mathcal{S}_Z | \mathcal{E}_Z] \\ &\quad\leq \mathrm{P}_{\mathcal{P}'_k(n,m)}[\exists X \subset V_t \setminus Z, |X| = \zeta\theta n : \forall x \in X : T_x(X) \geq S_x - 1 | \mathcal{E}_Z] \\ &\quad\leq \sum_{X \subset V_t \setminus Z, |X| = \theta\zeta n} \mathrm{P}_{\mathcal{P}'_k(n,m)}[\forall x \in X : T_x(X) \geq S_x - 1 | \mathcal{E}_Z] \\ &\quad\leq \binom{(1-2\zeta)\theta n}{\theta\zeta n} \cdot (\gamma + o(1))^{\theta\zeta n} \\ &\quad\leq \exp[(1-2\zeta)h(\zeta/(1-2\zeta)) \cdot \theta n] \cdot (\gamma + o(1))^{\zeta\theta n} \\ &\quad\leq \exp[\theta n \cdot ((1-2\zeta)h(\zeta/(1-2\zeta)) + \zeta \cdot \ln\gamma + o(1))] \\ &\quad\overset{(24)}{<} \exp(-\rho n/2^k).\end{aligned} \tag{27}$$

Now, let $\mathcal{E}$ be the event that there is a set Z of size $|Z| \leq 2\zeta\theta n$ such that $\mathcal{E}_Z$ occurs, and let $\mathcal{S}$ be the event that there is such a Z such that $\mathcal{S}_Z$ occurs. Then (27) and the law of total probability imply $\mathrm{P}[\mathcal{S}|\mathcal{E}] \geq 1 - \exp(-\rho n/2^k)$, and thus

$$\mathrm{P}_{\mathcal{P}'_k(n,m)}[\mathcal{S} \vee \neg\mathcal{E}] \geq 1 - \exp(-\rho n/2^k) \tag{28}$$

Combining (28) with Corollary 15, we thus see that $\mathrm{P}_{\mathcal{U}_k(n,m)}[\mathcal{S} \vee \neg\mathcal{E}] = 1 - o(1)$. As Proposition 21 and (24) ensure that $\mathrm{P}_{\mathcal{U}_k(n,m)}[\neg\mathcal{E}] = o(1)$, we thus obtain $\mathrm{P}_{\mathcal{U}_k(n,m)}[\mathcal{S}] = 1 - o(1)$, as desired. ∎

Proof of Theorem 17*, part* 2. Propositions 22 and 23 reduce the problem of proving the existence of a large set of rigid variables to a problem in calculus. Namely, assuming that $1 < k\theta \le \exp(\rho)/(\rho^3)$ and $\rho \le k\ln 2 - 2\ln k$ we need to verify (24). Our assumption that $\rho \le k\ln 2 - 2\ln k$ and that $k\theta > 1$ imply that $\zeta \ge \frac{1}{3}\rho^2 k^2/2^k$. Therefore,

$$2^k\theta\zeta\varphi(1) > \frac{\varphi(1)}{3}\rho^2k^2\theta > \rho,$$

and thus the first inequality in (24) is valid. Moreover, since $h(\zeta) \le \zeta(1-\ln\zeta)$, proving the second inequality in (24) reduces to showing $\theta\zeta\,[\ln\gamma + 1 - \ln\zeta] < -\rho/2^k$, i.e.,

$$\theta\zeta\ln(\mathrm{e}\gamma/\zeta) < -\rho/2^k. \tag{29}$$

Plugging in the definitions of γ and ζ, we see that for μ not too small

$$\ln\left(\frac{\mathrm{e}\gamma}{\zeta}\right) = \ln\left[\frac{\mathrm{e}\mu(\exp(\lambda\mu)-\lambda\mu-1)}{(1-\zeta)(1+\mu+\mu^2/2)}\right] \le \ln\left(\frac{5\mathrm{e}\,[\exp(\lambda\mu)-\lambda\mu-1]}{\mu}\right).$$

Since $3\theta\zeta \le 4/(k\rho)$ for $\rho \ge \rho_0$ sufficiently large, we have $\lambda = 1-(1-3\theta\zeta)^{k-1} \le 4k\theta\zeta$. Hence,

$$\lambda\mu \le 4k\mu\theta\zeta \le \frac{4\mu^2}{\exp(\mu)}\cdot\frac{\exp(\mu)}{\mu^3} \le 4/\mu.$$

Therefore, we obtain for $\mu \ge \rho \ge \rho_0$ large

$$\ln\left(\frac{\mathrm{e}\gamma}{\zeta}\right) \le \ln\left(\frac{4\mathrm{e}(\lambda\mu)^2}{\mu}\right) \le \ln\left(64\mathrm{e}/\mu^3\right) \le -1.$$

As $\theta\zeta \ge \zeta/k \ge \frac{1}{2}\mu^2\exp(-\mu)/k \ge \frac{1}{3}\rho^2/2^k$ for $k \ge k_0$ and $\rho \ge \rho_0$ not too small, we thus obtain (29). □

Let us make the following statement from the proof of Theorem 17 explicit to facilitate the proof of Theorem 26.

Corollary 24. *Suppose that* $\ln(\rho) - 10 \le k\theta < (1-\rho^{-2})\cdot\rho/(\ln 2)$. *Let* $(\boldsymbol{\Phi},\boldsymbol{\sigma})$ *be a pair chosen from the distribution* U1–U4. *Then there is a* t*-self-contained set* $R \subset V_t$ *of size* $|R| \ge (1-3\zeta)\theta n \ge 0.99\theta n$ *w.h.p.*

4.4. Forced variables. Let (Φ,σ) be a formula/assignment pair. A clause C *forces* a variable $x \in V_t$ if C contains $k-1$ literals from $\{x_1,\bar{x}_1,\ldots,x_t,\bar{x}_t\}$, none of which satisfies C under σ, and x underlies the remaining literal of C, which satisfies the clause under σ.

Lemma 25. *Suppose that* $\rho \ge \rho_0$, $k \ge k_0$, *and* $k\theta \sim \ln(\rho) - 10$. *Then w.h.p. in a pair* $(\boldsymbol{\Phi},\boldsymbol{\sigma})$ *chosen from the distribution* $\mathcal{U}_k(n,m)$ *at least* $0.991\theta n$ *variables in* V_t *are forced.*

Proof. Let $\mathcal{F}$ be the event that at least $0.991\theta n$ variables in V_t are forced. We are going to show that

$$\mathrm{P}_{\mathcal{P}'_k(n,m)}[\mathcal{F}] \ge 1-\exp(-\rho n/2^k), \tag{30}$$

so that the assertion follows from Corollary 15.

Thus, let $(\boldsymbol{\Phi}',\boldsymbol{\sigma}')$ be a pair chosen from the distribution $\mathcal{P}'_k(n,m)$. We may assume without loss of generality that $\boldsymbol{\sigma}'$ is the all-true assignment. By Lemma 16 for

each variable $x \in V_t$ the number of clauses that x supports has a binomial distribution with mean $\mu = \rho \cdot 2^k/(2^k - 1)$. Furthermore, if C is a random clause supported by x, then C contains $k-1$ random negative literals; the probability that all of these are in $V \setminus V_t$ equals $(1 - \theta + o(1))^{k-1}$. Hence, the number F_x of forcing clauses for x is binomially distributed with mean

$$\begin{aligned}
\mathrm{E}[F_x] &= \mu(1-\theta+o(1))^{k-1} \\
&\geq \rho(1-\theta)^{k-1} \ [\text{as } \mu = \rho \cdot 2^k/(2^k-1) > \rho] \\
&\geq \rho \exp\left[-(\theta+\theta^2)(k-1)\right] \ [\text{as } \theta \leq \ln(k)/k \text{ and thus } 1-\theta \geq \exp(-\theta-\theta^2)] \\
&\geq \rho \exp\left[-\theta k - \theta^2 k\right] = \rho \exp\left[(1-\theta)\cdot(10-\ln(\rho)) + o(1)\right] \ [\text{as } k\theta \sim \ln(\rho) - 10] \\
&= \exp\left[10 + \theta \cdot (10 - \ln(\rho)) + o(1)\right] = \exp\left[10 - \theta^2 k + o(1)\right] \geq \exp(5),
\end{aligned}$$

because $\theta \leq \ln(k)/k$ by our assumptions that $\rho \leq k \ln 2$ and $k\theta \sim \ln(\rho) - 10$. Therefore, for any $x \in V_t$ we have $\mathrm{P}[F_x = 0] \leq \exp(-\exp(5))$, and the events $(\{F_x = 0\})_{x \in V_t}$ are mutually independent. Hence, the number Z of variables $x \in V_t$ with $F_x = 0$ is binomially distributed with mean $\exp(-\exp(-5))\theta n$. Thus, the Chernoff bound (8) and our assumption that $k\theta \sim \ln(\rho) - 10$ yield

$$\mathrm{P}[Z \geq 0.009\theta n] \leq \exp(-0.009\theta n) \leq \exp(-\rho n/2^k),$$

whence (30) follows. □

Proof of Theorem 17, *part* 3. Lemma 25 directly implies the third part of Theorem 17 in the case $k\theta \sim \ln(\rho) - 10$. The remaining task is to prove the claim for smaller values of θ, namely, for any $\ln n/n \leq \theta \leq \theta_0 = (\ln(\rho) - 10)/k$. Let $t = (1-\theta)n$ and $t_0 = (1-\theta_0)n$. To obtain a pair $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ from the distribution U1–U4, we could proceed as follows. First, choose a pair $(\boldsymbol{\Phi}_{t_0}, \boldsymbol{\sigma}_{t_0})$ from the distribution U1–U4 with t_0 variables decimated. Then, assign the variables in $x \in V_{t_0} \setminus V_t$ the truth values $\boldsymbol{\sigma}_{t_0}(x)$, simplify the formula, and let $\boldsymbol{\sigma}_t(y) = \boldsymbol{\sigma}_{t_0}(y)$ for all $y \in V_t$. We are going to use this "two-round" experiment to analyze the number of forced variables in $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$.

The above experiment shows that any variable $x \in V_t$ that is forced in $(\boldsymbol{\Phi}_{t_0}, \boldsymbol{\sigma}_{t_0})$ remains forced in $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$. Let $\mathcal{F}$ be the set of forced variables in $(\boldsymbol{\Phi}_{t_0}, \boldsymbol{\sigma}_{t_0})$. Given that $|\mathcal{F}| = j$, the set $\mathcal{F}$ is a uniformly random subset of V_{t_0}, just by the symmetry of the distribution $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ with respect to permutations of the variables. Hence, if we condition on the event that $|\mathcal{F}| \geq 0.991\theta_0 n$, then $|\mathcal{F} \cap V_t|$ has a hypergeometric distribution with mean at least $0.991\theta n$. (This is because we are interested in the number of "special" elements that we catch upon selecting without replacement θn elements out of a total $\theta_0 n$ elements, where $|\mathcal{F}| \geq 0.991\theta_0 n$ among all $\theta_0 n$ elements are special.) Since the variance of the hypergepmetric distribution is bounded by its expectation, we can use Chebyshev's inequality to get $|\mathcal{F} \cap V_t| \geq (0.991\theta - o(1))n \geq 0.99\theta n$ w.h.p. (here we use that $\theta n \geq \ln n \gg 1$). Thus, part 3 of Theorem 17 follows from Lemma 25. □

5. Shattering, pairwise distances, and ferromagnetism.

5.1. Overview. In this section we establish the various statements on the global "shape" of the set $\mathcal{S}(\boldsymbol{\Phi}_t)$ from Theorems 2–4. The following theorem summarizes them.

THEOREM 26. *There exist constants $k_0, \rho_0 > 0$ such that for all $k \geq k_0$, and $\rho_0 \leq \rho \leq k \ln 2 - 2 \ln k$ the following three statements hold:*

1. *If*
$$\frac{\rho}{\ln 2}(1+\rho^{-2}+2^{2-k}) \leq k\theta \leq \exp\left[\rho\left(1-\frac{\ln\rho}{\rho}-\frac{2}{\rho}\right)\right],$$
then $\mathcal{S}(\mathbf{\Phi}_t)$ *is* $(\exp(2-\rho)-\varepsilon, \exp(2-\rho)+\varepsilon)$*-shattered w.h.p. for some* $0 < \varepsilon = \varepsilon(k,\rho) < \exp(2-\rho)$.
2. *If* $\ln\rho < k\theta < (\rho - 1/\rho)/\ln 2$*, then* $\mathcal{S}(\mathbf{\Phi}_t)$ *is* $\rho\exp(4-\rho)$*-ferromagnetic w.h.p. Furthermore, there is a set* $R \subset V_t$ *of size* $|R| \geq 0.99\theta n$ *such that for any* $\sigma, \tau \in \mathcal{S}(\mathbf{\Phi}_t)$ *we have*
$$|\{x \in R : \sigma(x) \neq \tau(x)\}| \leq k2^{2-k}n. \tag{31}$$
3. *If* $\theta > \rho(1+2/\rho^2)/(k\ln 2)$*, then for almost all* $\sigma \in \mathcal{S}(\mathbf{\Phi}_t)$ *we have*
$$|\{\tau \in \mathcal{S}(\mathbf{\Phi}_t) : \mathit{dist}(\sigma,\tau) \leq 0.49\theta n\}| \leq \exp(-\Omega(n))\,|\mathcal{S}(\mathbf{\Phi}_t)|.$$

To prove Theorem 26, we adapt arguments from [2, 3, 10, 11] to the setting with the *two* parameters θ, ρ (rather than just the density ρ). Let $(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$ be the (random) outcome of the experiment U1–U4. For $0 \leq \alpha \leq 1$ let $X_\alpha(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$ denote the number of satisfying assignments $\tau \in \mathcal{S}(\mathbf{\Phi}_t)$ with Hamming distance $\mathrm{dist}(\boldsymbol{\sigma}_t, \tau) = \alpha\theta n$. To establish the shattering part of Theorem 3, we basically need to study the random variables $X_\alpha(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$.

As this seems difficult to accomplish directly, we are going to work with the planted model again. That is, let $(\mathbf{\Phi}'_t, \boldsymbol{\sigma}'_t)$ be the outcome of the experiment P1–P4. It is fairly easy to get a handle on the *expected* number $\mathrm{E}\,[X_\alpha(\mathbf{\Phi}'_t, \boldsymbol{\sigma}'_t)]$ of satisfying assignments $\tau \in \mathcal{S}(\mathbf{\Phi}'_t)$ at distance $\mathrm{dist}(\boldsymbol{\sigma}'_t, \tau) = \alpha\theta n$. Let
$$\psi(\alpha) = \psi_{k,\rho}(\alpha) = -\alpha\theta\ln\alpha - (1-\alpha)\theta\ln(1-\alpha) + \frac{2^k\rho}{k}\ln\left(1-\frac{1-(1-\alpha\theta)^k}{2^k-1}\right). \tag{32}$$

LEMMA 27. *For any* $0 < \alpha < 1$ *we have* $\frac{1}{n}\ln \mathrm{E}\,[X_\alpha(\mathbf{\Phi}'_t, \boldsymbol{\sigma}'_t)] \leq \psi(\alpha)$.

Proof. The total number of assignments τ at Hamming distance $\alpha\theta n$ from $\boldsymbol{\sigma}'_t$ such that $\boldsymbol{\sigma}'_t(x_i) = \tau(x_i)$ for all $1 \leq i \leq t$ equals $\binom{\theta n}{\alpha\theta n}$. (Just choose the $\alpha\theta n$ variables from V_t where the two assignments disagree.) For any such assignment, the probability that a random clause that has $1 \leq l \leq k$ satisfied literals under $\boldsymbol{\sigma}'_t$ is unsatisfied under τ equals $(\alpha\theta)^l(1-\alpha\theta)^{k-l}$ (for τ has to disagree with $\boldsymbol{\sigma}'_t$ on all l satisfied literals but on none of the $k-l$ unsatisfied ones). Hence, the probability that a random clause that is satisfied under $\boldsymbol{\sigma}'_t$ happens to be unsatisfied under τ is
$$\frac{1}{2^k-1}\sum_{l=1}^{k}\binom{k}{l}(\alpha\theta)^l(1-\alpha\theta)^{k-l} = \frac{1-(1-\alpha\theta)^k}{2^k-1}.$$
Since the m clauses of $\mathbf{\Phi}'_t$ are chosen independently from among the clauses satisfied under $\boldsymbol{\sigma}'_t$, we obtain that
$$\mathrm{P}\,[\tau \text{ satisfies } \mathbf{\Phi}'_t] = \left(1-\frac{1-(1-\alpha\theta)^k}{2^k-1}\right)^m$$
for any τ with $\boldsymbol{\sigma}'_t(x_i) = \tau(x_i)$ for all $1 \leq i \leq t$ and $\mathrm{dist}(\boldsymbol{\sigma}'_t, \tau) = \alpha\theta n$. Finally, using the linearity of the expectation, we get
$$\mathrm{E}\,[X_\alpha(\mathbf{\Phi}'_t, \boldsymbol{\sigma}'_t)] = \sum_{\tau \text{ as above}} \mathrm{P}\,[\tau \text{ satisfies } \mathbf{\Phi}'_t] = \binom{\theta n}{\alpha\theta n}\left(1-\frac{1-(1-\alpha\theta)^k}{2^k-1}\right)^m.$$

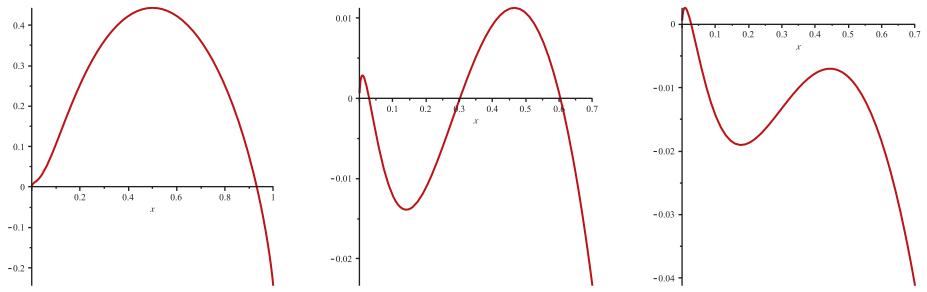

FIG. 2. *The three possible shapes of the function ψ (illustrated for $k = 16$, $\rho = 5$, and $\theta = 1, 0.37, 0.34$).*

Taking logarithms and bounding the binomial coefficient via Stirling's formula yields the assertion. □

Combining Theorem 13 and Lemma 27, we can essentially reduce the analysis of the random variables $X_\alpha(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$ to the study of the function ψ, i.e., a problem in calculus. As we will see, depending on the choice of ρ, θ the function ψ takes one of three possible shapes; see Figure 2. The left plot illustrates the shape of ψ for "large" θ in the symmetric regime $k\theta > \exp(\rho)$, where ψ is strictly positive in the entire interval $(0, 1/2)$. The second plot corresponds to the shattered regime $\rho/\ln 2 < k\theta < \exp(\rho)$. In this case, there is an interval (α, β) between 0 and $\frac{1}{2}$ where ψ is negative. Hence, Lemma 27 shows that for the outcome $(\mathbf{\Phi}'_t, \boldsymbol{\sigma}'_t)$ of P1–P4 the *expected* number of satisfying assignments τ such that $\alpha\theta n \le \mathrm{dist}(\tau, \boldsymbol{\sigma}'_t) \le \beta\theta n$ is exponentially small, and thus by Markov's inequality there are no such τ w.h.p. This observation in combination with the lower bound on the number of satisfying assignments of $\mathbf{\Phi}'_t$ from Corollary 14 will allow us to show that w.h.p. $\boldsymbol{\sigma}'_t$ lies in a tiny cluster of satisfying assignments that is separated from the remaining set $\mathcal{S}(\mathbf{\Phi}'_t)$ by a linear Hamming distance. Finally, the third plot corresponds to the ferromagnetic phase $k\theta < \rho/\ln 2$, where the expected number of satisfying assignments beyond a certain small distance $\alpha\theta n$ from $\boldsymbol{\sigma}'_t$ is exponentially small. Hence, in this phase only the small cluster of satisfying assignments around $\boldsymbol{\sigma}'_t$ remains.

In section 5.2 we will show that shattering occurs when the shape of ψ changes from the first to the second shape. Moreover, in section 5.3 we will show that ferromagnetism corresponds to the third shape. Finally, ψ also contains the necessary information to prove the statement about pairwise distances (see section 5.4).

5.2. Shattering. In this section we prove the first part of Theorem 26. The main step of the proof is summarized in the following proposition.

PROPOSITION 28. *Let $k \ge 6$ and $r > 0$ be fixed. Let $0 < \theta \le 1$. Suppose that there is a number $a \in (0, 1)$ such that*

$$\psi(a) + \rho/2^k < 0 \quad \text{and} \quad \sup_{0<\alpha<a} \psi(\alpha) < \theta \ln 2 + 2^k \rho \ln(1 - 2^{-k})/k - \rho/2^{k-1}. \tag{33}$$

Then there is $\varepsilon = \varepsilon(k, \rho)$ such that for a pair $(\mathbf{\Phi}_t, \boldsymbol{\sigma}_t)$ generated by the experiment U1–U4, *the set $\mathcal{S}(\mathbf{\Phi}_t)$ is $(a - \varepsilon, a + \varepsilon)$-shattered w.h.p.*

The proof of Proposition 28 appears at the end of this section. Before we come to that, let us show how Proposition 28 implies the first part of Theorem 26. To this end, we need to verify (33).

LEMMA 29. *Assume that $0 \le \theta \le \exp(\rho - 2)/(\rho k)$. Let $a = \exp(2 - \rho)$. Then $\psi(a) < -a\theta/2$.*

Proof. We have

$$\psi(a) \le a\theta(1-\ln a) - \frac{\rho}{k}\left(1-(1-a\theta)^k\right) \le a\theta(1-\ln a) - \frac{\rho}{k}\left(1-\exp(-ak\theta)\right)$$
$$\le a\theta(1-\ln a) - \frac{\rho}{k}\left(ak\theta - (ak\theta)^2/2\right) = a\theta\left[1-\ln a - \rho(1-ak\theta/2)\right],$$

where we used the elementary inequality $\exp(-z) \le 1 - z + z^2/2$ for $z \ge 0$. Since $k\theta\rho \le \exp(\rho-2)$ by assumption, our choice of a implies that

$$\psi(a) \le a\theta\left[1-\ln a - \rho + a\exp(\rho-2)/2\right] = -a\theta/2,$$

as claimed. ∎

LEMMA 30. *Assume that* $0 \le \theta \le \exp(\rho-2)/(\rho k)$. *Let* $a = \exp(2-\rho)$. *Then* $\sup_{\alpha<a}\psi(\alpha) \le \frac{3}{2\mathrm{e}^2 k\rho}$.

Proof. Let $0 \le \alpha < a$. We have

$$\psi(\alpha) \le \theta(\alpha - \alpha\ln\alpha - \alpha\rho(1-\alpha k\theta/2)).$$

Let $\psi_1(\alpha)$ be the expression on the right-hand side (r.h.s.). Then

$$\frac{d}{d\alpha}\psi_1(\alpha) = \theta\left[-\ln\alpha - \rho + \alpha k\rho\theta\right], \frac{d^2}{d\alpha^2}\psi_1(\alpha) = \theta\left[k\rho\theta - 1/\alpha\right].$$

Thus, our assumption on θ implies that $\frac{d^2}{d\alpha^2}\psi_1(\alpha) < 0$ for all $0 < \alpha < a$, and therefore ψ_1 has a unique local maximum in the interval $(0,\alpha)$. To pinpoint this maximum, note that for $\alpha_0 = \exp(-\rho)$ the first derivative $\frac{d}{d\alpha}\psi_1(\alpha_0)$ is positive. Moreover, at $\alpha_1 = \exp(1-\rho)$ we have $\frac{d}{d\alpha}\psi_1(\alpha_1) < 0$. Hence, the unique local maximum of ψ_1 lies in the interval (α_0,α_1). To study the maximum value, consider the function $\psi_2 : \alpha \mapsto \alpha - \alpha\ln\alpha - \alpha\rho$. Its derivative is $d/d\alpha\,\psi_2(\alpha) = -\rho - \ln\alpha$, so that the maximum of this function occurs at α_0. Furthermore, the quadratic term $\alpha \mapsto \alpha^2\theta\rho k/2$ is monotonically increasing in α. Therefore,

$$\sup_{0<\alpha<a}\psi(\alpha) \le \sup_{0<\alpha<a}\psi_1(\alpha) = \sup_{\alpha_0<\alpha<\alpha_1}\psi_1(\alpha) \le \theta(\psi_2(\alpha_0) + \alpha_1^2\theta\rho k/2) \le 3\theta\exp(-\rho)/2.$$

Finally, the assertion follows from the assumed bound on θ. ∎

Proof of Theorem 26, *part* 1. Assume that $\rho \le k\ln 2 - \ln k$ and

$$\frac{\rho}{k\ln 2}(1+\rho^{-2}+2^{2-k}) \le \theta \le \exp(\rho-2)/(\rho k).$$

Let $a = \exp(2-\rho)$. Lemma 29 shows that

$$\psi(a) + \rho/2^k \le \rho/2^k - \exp(2-\rho)\theta/2 \le \rho/2^k - \frac{\exp(2-\rho)\rho}{2k\ln 2}$$
$$= \frac{\rho}{2^k}\left(1 - \frac{2^k\exp(2-\rho)}{2k\ln 2}\right).$$

Since $\rho \le k\ln 2 - \ln k$, the r.h.s. is negative. By Lemma 30 we have

$$\theta\ln 2 + \frac{2^k\rho}{k}\ln(1-2^{-k}) - \rho/2^{k-1} \ge \theta\ln 2 - \frac{\rho}{k} - \rho/2^{k-2}$$
$$\ge \frac{1}{k\rho} > \sup_{\alpha<a}\psi(\alpha). \tag{34}$$

Thus, the assertion follows from Proposition 28. ∎

Proof of Proposition 28. In the rest of this section we keep the notation and the assumptions from Proposition 28. Let

$$b = \theta \ln 2 + 2^k \rho \ln(1 - 2^{-k})/k.$$

We will mostly be working with a formula/assignment pair $(\mathbf{\Phi}', \boldsymbol{\sigma}')$ chosen from the planted distribution $\mathcal{P}_k(n, m)$. Recall that $\mathbf{\Phi}'_{t,\boldsymbol{\sigma}'}$ denotes the formula obtained by substituting the value $\boldsymbol{\sigma}'(x_i)$ for the variables x_i for $1 \le i \le t$, and simplifying.

LEMMA 31. *There exist numbers $\xi > 0$, $0 < a_1 < a_2 < 1$ such that a pair $(\mathbf{\Phi}', \boldsymbol{\sigma}')$ chosen from the distribution $\mathcal{P}_k(n, m)$ has the following two properties with probability at least $1 - \exp(-(\xi + \rho/2^k)n)$.*

1. *$\mathbf{\Phi}'_{t,\boldsymbol{\sigma}'}$ does not have a satisfying assignment τ with $a_1 n < dist(\boldsymbol{\sigma}', \tau) < a_2 n$.*
2. *$|\{\tau \in \mathcal{S}(\mathbf{\Phi}'_{t,\boldsymbol{\sigma}'}) : dist(\boldsymbol{\sigma}', \tau) < a_2 n\}| \le \exp((b - \xi - \rho/2^k)n)$.*

Proof. For $\alpha > 0$ we let

$$X_\alpha = \left|\left\{\tau \in \mathcal{S}(\mathbf{\Phi}'_{t,\boldsymbol{\sigma}'}) : \mathrm{dist}(\boldsymbol{\sigma}', \tau) = \alpha\theta n\right\}\right|.$$

By Lemma 27,

$$\frac{\ln \mathrm{E}[X_\alpha]}{n} \le \psi(\alpha). \tag{35}$$

Let $a \in (0, 1)$ be such that $\psi(a) + \rho/2^k < 0$ (cf. (33)). As ψ is continuous there exist $0 < a_1 < a < a_2 < 1$ and $\xi_1 > 0$ such that

$$\sup_{a_1 \le \alpha \le a_2} \psi(\alpha) < -\rho/2^k - 2\xi_1. \tag{36}$$

Combining (35) and (36), we conclude that $\mathrm{E}[X_\alpha] \le \exp\left[-n(\rho/2^k + 2\xi_1)\right]$ for all $a_1 \le \alpha \le a_2$. Summing over integers $a_1 n \le j \le a_2 n$, we see that for large n

$$\sum_{a_1 n \le j \le a_2 n} \mathrm{E}\left[X_{j/n}\right] \le n \exp\left[-n(\rho/2^k + 2\xi_1)\right] \le \exp\left[-n(\rho/2^k + \xi_1)\right].$$

Hence, by Markov's inequality the probability that there is a satisfying assignment τ that coincides with $\boldsymbol{\sigma}'$ on the first t variables such that $a_1 n \le \mathrm{dist}(\boldsymbol{\sigma}', \tau) \le a_2 n$ is bounded by $\exp(-n(\rho/2^k + \xi_1))$. This proves the first assertion.

Since we are assuming that $\sup_{0<\alpha<a} \psi(\alpha) < b - \rho/2^{k-1}$, and as (36) shows that $\psi(\alpha) < -\rho/2^k - 2\xi_1 < b - \rho/2^{k-1} - 2\xi_1$ for all $a \le \alpha < a_2$, there is a number $\xi_2 > 0$ such that

$$\sup_{0 < \alpha \le a_2} \psi(\alpha) < b - \rho/2^{k-1} - 3\xi_2.$$

Hence, (35) implies that

$$\mathrm{E}[X_\alpha] \le \exp(n\psi(\alpha)) \le \exp(n(b - \rho/2^{k-1} - 3\xi_2)) \qquad \text{for all } 0 < \alpha \le a_2.$$

Taking the sum over integers $0 \le j \le a_2 n$, we get for large enough n

$$\sum_{0 \le j \le a_2 n} \mathrm{E}\left[X_{j/n}\right] \le n \exp(n(b - \rho/2^k - 3\xi_2)) \le \exp(n(b - \rho/2^{k-1} - 2\xi_2)).$$

That is, the expected number of assignments $\tau \in \mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}'})$ such that $\mathrm{dist}(\boldsymbol{\sigma}', \tau) \le a_2 n$ is bounded by $\exp(n(b - \rho/2^{k-1} - 2\xi_2))$. Hence, Markov's inequality entails that with

probability at least $1-\exp(-n(\rho/2^k+\xi_2))$ there are at most $\exp(n(b-\rho/2^k-\xi_2))$ such satisfying assignments τ. This proves the second assertion. □

COROLLARY 32. *There exist numbers $\xi>0$, $0<a_1<a_2<1$ such that a pair $(\mathbf{\Phi},\boldsymbol{\sigma})$ chosen from the distribution $\mathcal{U}_k(n,m)$ enjoys the two properties stated in Lemma 31 with probability at least $1-\exp(-\xi n)$.*

Proof. This follows directly from Lemma 31 and Theorem 13. □

Proof of Proposition 28. Let ξ,a_1,a_2 be the numbers provided by Corollary 32 and let $(\mathbf{\Phi},\boldsymbol{\sigma})$ be a pair chosen from the distribution $\mathcal{U}_k(n,m)$. With each assignment $\tau\in\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})$ we associate a set

$$\mathcal{C}(\tau)=\{\chi\in\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}}):\mathrm{dist}(\chi,\tau)\le a_1n\}.$$

Moreover, we call $\tau\in\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})$ *good* if $|\mathcal{C}(\tau)|\le\exp((b-\rho/2^k-\xi)n)$ and there is no $\chi\in\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})$ such that $a_1n\le\mathrm{dist}(\chi,\tau)\le a_2n$. Let $\mathcal{S}_{good}$ be the set of all good $\tau\in\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})$ and $\mathcal{S}_{bad}=\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})\setminus\mathcal{S}_{good}$. Corollary 32 and our choice of b ensure that $\mathbf{\Phi}$ has the following two properties w.h.p.:

$$|\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})|\ge\exp(n(b-\rho/2^k)), \tag{37}$$

$$|\mathcal{S}_{good}|\ge(1-\exp(-\xi n))\cdot|\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})|. \tag{38}$$

Assuming that (37) and (38) hold and that n is sufficiently large, we are going to construct a decomposition of $\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})$ into subsets as required by SH1–SH2. To this end, choose some $\sigma_1\in\mathcal{S}_{good}$. Having defined $\sigma_1,\dots,\sigma_l$, we choose an arbitrary $\sigma_{l+1}\in\mathcal{S}_{good}\setminus\bigcup_{j=1}^{l}\mathcal{C}(\sigma_j)$, unless this set is empty, in which case we stop. Let $\sigma_1,\dots,\sigma_N$ be the resulting sequence and define

$$R_l=\mathcal{C}(\sigma_l)\setminus\bigcup_{j=1}^{l-1}\mathcal{C}(\sigma_j)\qquad\text{for }1\le l\le N\quad\text{and }R_0=\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})\setminus\bigcup_{l=1}^{N}R_l.$$

Then $\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})=R_0\cup\cdots\cup R_N$. (Observe that possibly $R_0=\emptyset$, while $R_l\neq\emptyset$ for all $1\le l\le N$ as $\sigma_l\in R_l$.) Furthermore, for each $1\le l\le N$ we have $R_l\subset\mathcal{C}(\sigma_l)$ and thus

$$\begin{aligned}|R_l|&\le|\mathcal{C}(\sigma_l)|\le\exp((b-\rho/2^k-\xi)n)&&\text{(because }\sigma_l\text{ is good)}\\&\le|\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})|\cdot\exp(-\xi n)&&\text{(by (37)).}\end{aligned}\tag{39}$$

Furthermore, as $R_0\subset\mathcal{S}_{bad}$, (38) implies

$$|R_0|\le|\mathcal{S}_{bad}|\le\exp(-\xi n))\cdot|\mathcal{S}(\mathbf{\Phi}_{t,\boldsymbol{\sigma}})|. \tag{40}$$

Combining (39) and (40) we see that the decomposition $R_0,\dots,R_N$ satisfies SH1. Furthermore, SH2 is satisfied by construction. □

5.3. Ferromagnetism. Here we prove the second part of Theorem 26. The following proposition reduces that task to a problem in calculus.

PROPOSITION 33. *Let $k\ge3$ and $r>0$ be fixed. Let $0<\theta\le1$. If there is a number $a\in(0,1)$ such that*

$$\sup_{a<\alpha\le1}\psi(\alpha)+\rho/2^k<0, \tag{41}$$

then in the result $(\mathbf{\Phi}_t,\boldsymbol{\sigma}_t)$ of the experiment U1–U4 *the formula $\mathbf{\Phi}_t$ is $2a$-ferromagnetic w.h.p.*

Proof. Let $(\mathbf{\Phi}', \boldsymbol{\sigma}')$ be a pair chosen from the planted distribution $\mathcal{P}_k(n, m)$. For $\alpha > 0$ we let

$$X_\alpha = \left|\left\{\tau \in \mathcal{S}(\mathbf{\Phi}'_{t,\boldsymbol{\sigma}'}) : \mathrm{dist}(\boldsymbol{\sigma}', \tau) = \alpha\theta n\right\}\right|.$$

Then $\frac{1}{n}\ln \mathrm{E}\left[X_\alpha\right] \leq \psi(\alpha)$ by Lemma 27. Hence, $\frac{1}{n}\ln \mathrm{E}\left[X_\alpha\right] < -\rho/2^k$ for $\alpha > a$ by (41). Thus, Markov's inequality yields

$$\mathrm{P}\left[\exists \tau \in \mathcal{S}_t(\mathbf{\Phi}'_{t,\boldsymbol{\sigma}'}) : \mathrm{dist}(\boldsymbol{\sigma}', \tau) \geq a\theta n\right] \leq \theta n \cdot \exp(-(\Omega(1) + \rho/2^k)n) < \exp(-\rho n/2^k).$$

Therefore, the assertion follows from Theorem 13. ∎

LEMMA 34. *Suppose that $\rho \leq k\ln 2 - 2\ln k$ and $\theta = (1 - 1/\rho^2)\frac{\rho}{k\ln 2}$. Moreover, assume that $\rho \geq \rho_0$ and $k \geq k_0$ for certain constants ρ_0, k_0. Let $a = \exp(2 - \rho)$. Then* (41) *is satisfied.*

Proof. Let $h(x) = -x\ln x - (1-x)\ln(1-x)$ be the entropy function. We have

$$\begin{aligned}
\psi(\alpha) &= \theta h(\alpha) + \frac{2^k\rho}{k}\ln\left(1 - \frac{1-(1-\alpha\theta)^k}{2^k - 1}\right)\\
&\leq \theta h(\alpha) - \frac{2^k\rho}{k(2^k-1)}\left(1-(1-\alpha\theta)^k\right) \qquad (\text{using } \ln(1-z) \leq -z)\\
&\leq \theta h(\alpha) - \frac{\rho}{k}\left(1-(1-\alpha\theta)^k\right) \leq \theta h(\alpha) - \frac{\rho}{k}(1-\exp(-\alpha k\theta)).
\end{aligned}$$

To bound the r.h.s., we are going to consider several cases, assuming each time that $\rho \geq \rho_0, k \geq k_0$ are sufficiently large.

Case 1: $\alpha \leq 1/(k\rho\theta)$. As $\alpha \geq a = \exp(2-\rho)$, we have

$$h(\alpha) \leq \alpha(1 - \ln\alpha) \leq \alpha(1 - \ln a) \leq \alpha(\rho - 1).$$

Furthermore, using $\exp(-z) \leq 1 - z + z^2/2$, we get

$$-\frac{\rho}{k}(1-\exp(-\alpha k\theta)) \leq -\alpha\theta\rho\left(1 - \alpha k\theta/2\right). \tag{42}$$

Hence, as $\alpha k\rho\theta \leq 1$ by assumption,

$$\psi(\alpha) \leq \alpha\theta\left[\frac{\alpha k\rho\theta}{2} - 1\right] \leq -\alpha\theta/2.$$

The assumption $\rho \leq k\ln 2 - 2\ln k$ ensures that the last term is smaller than $-\rho/2^k$.

Case 2: $1/(k\rho\theta) < \alpha < 1/(k\theta)$. Using the lower bound on α, we get

$$h(\alpha) \leq \alpha(1 - \ln\alpha) \leq \alpha(1 + \ln(\rho k\theta)).$$

Furthermore, using (42), we obtain

$$\begin{aligned}
\psi(\alpha) &\leq \alpha\theta\left[1 + \ln(k\rho\theta) - \rho + \frac{\alpha k\rho\theta}{2}\right]\\
&\leq \alpha\theta\left[1 + \ln(k\rho\theta) - \rho/2\right] \qquad (\text{as } \alpha < 1/(k\theta))\\
&\leq \alpha\theta\left[1 - \ln\ln 2 + 2\ln\rho - \rho/2\right] \qquad (\text{as } \theta \leq \tfrac{\rho}{k\ln 2})\\
&\leq -\alpha\theta\rho/4.
\end{aligned}$$

The assumption $\rho \leq k\ln 2 - 2\ln k$ ensures that the last term is smaller than $-\rho/2^k$.

Case 3: $1/(k\theta) < \alpha \le \alpha_0 = 0.15$. Since $h(\alpha)$ is monotonically increasing for $\alpha < 1/2$, we have

$$\begin{aligned}\psi(\alpha) &\le \theta h(\alpha) - \frac{\rho}{k}(1-\exp(-\alpha k\theta)) \le \theta h(\alpha_0) - \frac{\rho}{k}(1-1/\mathrm{e}) \\ &\le \frac{\rho}{k}\left[\frac{h(\alpha_0)}{\ln 2} - 1 + 1/\mathrm{e}\right].\end{aligned}$$

The choice of α_0 ensures that the last term is smaller than $-\rho/2^k$.

Case 4: $\alpha_0 < \alpha$. As $k\theta = (1-1/\rho^2)\rho/\ln 2$, we get

$$\begin{aligned}\psi(\alpha) &\le \theta h(\alpha) - \frac{\rho}{k}(1-\exp(-\alpha k\theta)) \\ &\le \theta \ln 2 - \frac{\rho}{k}(1-\exp(-\alpha_0(1-1/\rho^2)\rho/\ln 2)) \\ &\le \frac{\rho}{k}\left[\exp(-\alpha_0\rho) - 1/\rho^2\right].\end{aligned}$$

The last term is smaller than $-\rho/2^k$ because $\rho \le k\ln 2$.

Proof of Theorem 26, *part* 2a. Let $\theta_0 = (1-1/\rho^2)\rho/(k\ln 2)$ and $t_0 = (1-\theta_0)n$. Lemma 34 and Proposition 33 show that for the *specific* value θ_0 the formula $\mathbf{\Phi}_t$ is $\exp(2-\rho)$-ferromagnetic w.h.p. Now, suppose that $t > t_0$ and thus $\theta < \theta_0$, while $k\theta > \ln\rho$. Then $\mathbf{\Phi}_t$ is obtained from $\mathbf{\Phi}_{t_0}$ by assigning some further variables. Therefore,

$$\max\{\mathrm{dist}(\sigma,\tau) : \sigma,\tau \in \mathcal{S}(\mathbf{\Phi}_t)\} \le \max\{\mathrm{dist}(\sigma,\tau) : \sigma,\tau \in \mathcal{S}(\mathbf{\Phi}_{t_0})\}.$$

Hence, Proposition 33 and Lemma 34 imply that $\mathbf{\Phi}_t$ is $\rho\exp(4-\rho)$-ferromagnetic w.h.p. □

We are left to show part 2b of Theorem 26, i.e., the existence of a large set R of variables on which all satisfying assignments virtually agree. For this we need the following lemma.

LEMMA 35. *Suppose that* $1 \le k\theta \le \rho/(\ln 2)$. *Let* $(\mathbf{\Phi},\boldsymbol{\sigma})$ *be a pair chosen from the distribution* $\mathcal{U}_k(n,m)$. *Then w.h.p. the following statement is true:*

(43) *There is no set of variables* $Z \subset V_t$ *of size* $2kn/2^k \le |Z| \le (\mathrm{e}\rho)^{-4}\theta n$ *such that each variable in* Z *supports two clauses under* $\boldsymbol{\sigma}$, *each of which contains an occurrence of a variable in* Z *that evaluates to false under* $\boldsymbol{\sigma}$.

Proof. Let $(\mathbf{\Phi}',\boldsymbol{\sigma}')$ be chosen from the planted model $\mathcal{P}'_k(n,m)$. For a set $Z \subset V_t$ let $X_Z = 1$ if each variable in Z supports two clauses under $\boldsymbol{\sigma}'$, each of which contains an occurrence of a variable in Z that evaluates to false under $\boldsymbol{\sigma}'$. As each clause that is satisfied under $\boldsymbol{\sigma}'$ is present in $\mathbf{\Phi}'$ with probability p independently, and as for each variable $x \in Z$ there are no more than $|Z|\binom{n-2}{k-2}$ possible clauses supported by x that contain another variable from Z, we obtain

$$\mathrm{P}_{\mathcal{P}'_k(n,m)}[X_Z = 1] \le \left(|Z|\binom{n-2}{k-2}p\right)^{2|Z|} \le (k\rho|Z|/n)^{2|Z|}.$$

Thus, for any $0 < \alpha \le (\mathrm{e}\rho)^{-4}$ we can bound the probability that there is a set Z of

size $|Z| = \alpha\theta n$ with $X_Z = 1$ as follows:

$$\begin{aligned}
&\mathrm{P}_{\mathcal{P}'_k(n,m)}\left[\exists Z, |Z| = \alpha\theta n : X_Z = 1\right] \\
&\leq \sum_{Z:|Z|=\alpha\theta n} \mathrm{P}_{\mathcal{P}'_k(n,m)}\left[X_Z = 1\right] \leq \binom{\theta n}{\alpha\theta n}(\alpha k\theta\rho)^{2\alpha\theta n} \\
&\leq \left(\frac{\mathrm{e}}{\alpha}\right)^{\alpha\theta n}(\alpha k\theta\rho)^{2\alpha\theta n} = \left[\mathrm{e}\alpha(k\theta\rho)^2\right]^{\alpha\theta n} \\
&\leq (\mathrm{e}\alpha\rho^4/\ln^2 2)^{\alpha\theta n} \qquad (\text{as } \theta \leq \rho/\ln 2) \\
&\leq \exp(-\alpha\theta n) \qquad (\text{as } \alpha \leq (\mathrm{e}\rho)^{-4}) \\
&\leq \exp(-2kn/2^k), \qquad (44)
\end{aligned}$$

because $|Z| = \alpha\theta n \geq 2kn/2^k$. Using the union bound, we obtain from (44) that

$$\begin{aligned}
\mathrm{P}_{\mathcal{P}'_k(n,m)}\left[\exists Z, 2kn/2^k \leq |Z| \leq (\mathrm{e}\rho)^{-4}\theta n : X_Z = 1\right] &\leq n \cdot \exp(-2kn/2^k) \\
&< \exp(-\rho n/2^k), \qquad (45)
\end{aligned}$$

because we assume that $\rho \leq k\ln 2$. Thus, (45) and Corollary 15 imply the assertion. ∎

Proof of Theorem 26, *part* 2b. Assume that $k\theta < (1-\rho^{-2})\cdot\rho/(\ln 2)$. Let $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ be a pair chosen from the distribution U1–U4. We are going to show that w.h.p. there is a set $R \subset V_t$ of size $|R| \geq 0.99\theta n$ such that (31) holds for any $\sigma, \tau \in \mathcal{S}(\boldsymbol{\Phi}_t)$. If $k\theta < \ln(\rho) - 10$, then this will follow from the third part of Theorem 17 by letting R be the set of forced variables.

Thus, suppose that $\ln(\rho) - 10 \leq k\theta < (1-\rho^{-2})\cdot\rho/(\ln 2)$. Let $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ be a pair chosen from the distribution U1–U4. From Corollary 24, Lemma 35, and part 2a of Theorem 26 we know that w.h.p. $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ has the following three properties:

1. There is a t-self-contained set $R \subset V_t$ of size $|R| \geq (1-3\zeta)\theta n \geq 0.99\theta n$.
2. Condition (43) holds.
3. $\mathcal{S}(\boldsymbol{\Phi}_t)$ is $\rho\exp(4-\rho)$-ferromagnetic.

Let $\tau \in \mathcal{S}(\boldsymbol{\Phi})$ be a satisfying assignment and let $Z(\tau) = \{x \in R : \boldsymbol{\sigma}(x) \neq \tau(x)\}$. Since R is t-self-contained, each $x \in Z$ supports two clauses under $\boldsymbol{\sigma}$. Furthermore, each of these clauses contains a further variable from Z, because τ is a satisfying assignment. Therefore, (43) shows that either $|Z(\tau)| < 2kn/2^k$ or $|Z(\tau)| > (\mathrm{e}\rho)^{-4}\theta n$. But if $|Z(\tau)| > (\mathrm{e}\rho)^{-4}\theta n$, then

$$\mathrm{dist}(\boldsymbol{\sigma}, \tau) \geq |Z(\tau)| \geq (\mathrm{e}\rho)^{-4}\theta n > \rho\exp(4-\rho)\theta n,$$

and thus $\mathcal{S}(\boldsymbol{\Phi}_t)$ is not $\rho\exp(4-\rho)$-ferromagnetic. This shows that w.h.p. we have $|Z(\tau)| \leq 2kn/2^k$, and thus R satisfies (31). ∎

5.4. Pairwise distances. In this section we prove the third part of Theorem 26. Once more by analyzing the function ψ, we can show that w.h.p. in a pair $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ chosen from the distribution U1–U4 only a small fraction of all satisfying assignments are close to $\boldsymbol{\sigma}_t$.

LEMMA 36. *Suppose that* $\theta \geq \frac{\rho}{k\ln 2}(1 + 1/\rho^2 + k/2^{k-2})$. *Let* $(\boldsymbol{\Phi}, \boldsymbol{\sigma})$ *be a pair chosen from the distribution* $\mathcal{U}_k(n,m)$. *W.h.p. we have*

$$\left|\{\tau \in \mathcal{S}(\boldsymbol{\Phi}_{t,\boldsymbol{\sigma}}) : \mathit{dist}(\tau, \boldsymbol{\sigma}_t) \leq 0.49\theta n\}\right| \leq \exp(-\Omega(n))\left|\mathcal{S}(\boldsymbol{\Phi}_{t,\boldsymbol{\sigma}})\right|.$$

Proof. By Corollary 14, w.h.p.

$$\frac{1}{n}\ln|\mathcal{S}(\boldsymbol{\Phi}_{t,\boldsymbol{\sigma}})| \geq \theta\ln 2 + r\ln\left(1-2^{-k}\right) - \rho/2^k > 0. \qquad (46)$$

We are going to show that

$$\sup_{\alpha \le 0.49} \psi(\alpha) - \theta \ln 2 - r \ln\left(1 - 2^{-k}\right) < -\rho/2^{k-1}. \tag{47}$$

Then the assertion follows from Corollary 15, Lemma 27, and (46).

Thus, we are left to show (47). Since $-z - z^2 \le \ln(1-z) \le -z$ for $|z| \le 1/2$, we have

$$\begin{aligned}
&\psi(\alpha) - \theta \ln 2 - \frac{2^k \rho}{k} \ln\left(1 - 2^{-k}\right) \\
&\qquad = \theta(h(\alpha) - \ln 2) + \frac{2^k \rho}{k} \left[\ln\left(1 - \frac{1 - (1-\alpha\theta)^k}{2^k - 1}\right) - \ln\left(1 - 2^{-k}\right)\right] \\
&\qquad \le \theta(h(\alpha) - \ln 2) - \frac{2^k \rho}{k} \cdot \left[\frac{1 - (1-\alpha\theta)^k}{2^k - 1} - 2^{-k} - 4^{-k}\right] \\
&\qquad \le \theta(h(\alpha) - \ln 2) + \frac{\rho}{k}(1 - \alpha\theta)^k + 2^{-k} \qquad [\text{as } \rho \le k \ln 2] \\
&\qquad \le \theta(h(\alpha) - \ln 2) + \frac{\rho}{k} \exp(-\alpha k \theta) + 2^{-k}.
\end{aligned}$$

The differential of the last expression with respect to θ is negative, and thus the function is monotonically decreasing in θ. Therefore, it suffices to consider the minimum value $\theta = \rho/(k \ln 2)$. Thus, we obtain

$$\psi(\alpha) - \theta \ln 2 - \frac{2^k \rho}{k} \ln\left(1 - 2^{-k}\right) \le \frac{\rho}{k}\left(\frac{h(\alpha)}{\ln 2} - 1 + \exp\left(-\frac{\alpha\rho}{\ln 2}\right)\right) + 2^{-k}. \tag{48}$$

We consider a few different cases, assuming each time that $\rho \ge \rho_0, k \ge k_0$ are sufficiently large.

Case 0: $\alpha < \exp(2 - \rho)$. Lemma 30 shows that $\psi(\alpha) \le 1/(k\rho)$ and (34) shows that

$$\theta \ln 2 + 2^k \frac{\rho}{k} \ln(1 - 2^{-k}) \ge \theta \ln 2 - \rho/k - \rho/2^k.$$

Hence,

$$\psi(\alpha) - \theta \ln 2 - \frac{2^k \rho}{k} \ln\left(1 - 2^{-k}\right) \le \frac{1}{k\rho} - \theta \ln 2 + \frac{\rho}{k} + \rho/2^k.$$

Since we are assuming that $\theta \ge \frac{\rho}{k \ln 2}(1 + 1/\rho^2 + k/2^{k-2})$, the r.h.s. is smaller than $\rho/2^{k-1}$.

Case 1: $\exp(2 - \rho) \le \alpha \le \exp(-\rho/2)$. Since $\exp(-z) \le 1 - z + z^2/2$ for $z > 0$, we get from (48)

$$\begin{aligned}
\psi(\alpha) - \theta \ln 2 - r \ln\left(1 - 2^{-k}\right) &\le \frac{\alpha\rho}{k \ln 2}\left[1 - \ln\alpha - \rho + \frac{\alpha\rho^2}{2 \ln 2}\right] + 2^{-k} \\
&\le \frac{\alpha\rho}{k \ln 2}\left[-1 + \frac{\alpha\rho^2}{2 \ln 2}\right] + 2^{-k} < -\rho/2^{k-1},
\end{aligned}$$

provided that $\rho_0 \le \rho \le k \ln 2 - 2 \ln k$.

Case 2: $\exp(-\rho/2) \leq \alpha \leq 1/(2\rho)$. Bounding the exponential by a quadratic function, we get

$$\begin{aligned}\psi(\alpha) - \theta\ln 2 - r\ln\left(1-2^{-k}\right) &\leq \frac{\alpha\rho}{k\ln 2}\left[1-\ln\alpha-\rho+\frac{\alpha\rho^2}{2\ln 2}\right]+2^{-k}\\ &\leq \frac{\alpha\rho}{k\ln 2}\left[-\frac{\rho}{2}+\frac{\rho}{4\ln 2}\right]+2^{-k} < -\rho/2^{k-1},\end{aligned}$$

provided that $\rho_0 \leq \rho \leq k\ln 2 - 2\ln k$.

Case 3: $1/(2\rho) < \alpha \leq 10\ln(2)/\rho$. Suppose that $\alpha = x\ln(2)/\rho$ for some $1/2 \leq x \leq 10\ln 2$. Then

$$\begin{aligned}&\psi(\alpha) - \theta\ln 2 - r\ln\left(1-2^{-k}\right)\\ &\quad\leq \frac{\rho}{k}\left[\frac{x\ln 2}{\rho}(1-\ln x-\ln\ln 2+\ln\rho)-1+\exp(-x)\right]+2^{-k}.\end{aligned}$$

As x remains bounded away from 0, the term $\exp(-x)-1$ is strictly negative. Thus, the entire expression is smaller than $-\rho/2^{k-1}$ for $\rho \geq \rho_0$ sufficiently large.

Case 4: $10\ln(2)/\rho < \alpha \leq 0.49$. We have

$$\psi(\alpha) - \theta\ln 2 - r\ln\left(1-2^{-k}\right) \leq \frac{\rho}{k}\left(\frac{h(0.49)}{\ln 2}-1+\exp(-10)\right)+2^{-k}.$$

The r.h.s. is clearly smaller than $-\rho/2^{k-1}$. □

Lemma 36 directly implies the third part of Theorem 26.

6. Estimating the actual marginals. In this section we will prove the various statements from Theorems 2–4 about the marginals of the truth values of individual variables.

THEOREM 37. *There exist constants $k_0, \rho_0 > 0$ such that for all $k \geq k_0$ and $\rho_0 \leq \rho \leq k\ln 2 - 2\ln k$, the following two statements hold:*

1. *If $\theta \geq \frac{\rho}{k\ln 2}(1+1/\rho^2+k/2^{k-2})$, then w.h.p. for at least $\theta n/3$ variables $x \in V_t$ we have*
$$M_x(\boldsymbol{\Phi}_t) \in [0.01, 0.99].$$
2. *If $\ln(n)/n < \theta < \rho(1-1/\rho^2)/(k\ln 2)$, then w.h.p. for all but $\exp(-\rho)\theta n$ variables $x \in V_t$ we have*
$$M_x(\boldsymbol{\Phi}_t) \in \left[0, 2^{-k/2}\right] \cup \left[1-2^{-k/2}, 1\right].$$

To get an intuition why Theorem 37 might be true, observe that Theorem 26 implies that for θ as under 1, the pairwise distance between satisfying assignments is typically big. In other words, the satisfying assignment are spread out over the Hamming cube, which seems incompatible with many variables having marginals extremely close to either 0 or 1. By contrast, for θ as under 2, i.e., in the "ferromagnetic" phase, the set of satisfying assignments is confined to a tiny part of the Hamming cube (cf. Theorem 26). Thus, it is not surprising that the marginals of most variables are close to either 0 or 1 in this case.

Proof of Theorem 37. Let us first assume that $\theta \geq \frac{\rho}{k\ln 2}(1+1/\rho^2+k/2^{k-2})$. Let $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ be a pair chosen from the distribution U1–U4. We set up an auxiliary graph $\mathcal{G}$ whose vertices are all pairs (x,τ) of variables $x \in V_t$ and assignments $\tau \in \mathcal{S}(\boldsymbol{\Phi}_t)$.

A pair (x,τ) is connected by an edge with another pair (x,τ') if $\tau(x)=\tau'(x)$. We claim that w.h.p.

$$2|E(\mathcal{G})| \le 0.511|\mathcal{S}(\mathbf{\Phi}_t)|^2\theta n. \tag{49}$$

To see this, let us call an assignment $\tau\in\mathcal{S}(\mathbf{\Phi}_t)$ *sparse* if the total number of $\tau'\in\mathcal{S}(\mathbf{\Phi}_t)$ such that $\mathrm{dist}(\tau,\tau')<0.49\theta n$ is $o(|\mathcal{S}(\mathbf{\Phi}_t)|)$. Then Lemma 36 implies that w.h.p. all but $o(|\mathcal{S}(\mathbf{\Phi}_t)|)$ assignments $\tau\in\mathcal{S}(\mathbf{\Phi}_t)$ are sparse. Therefore, we obtain

$$\begin{aligned}2|E(\mathcal{G})| &= \sum_{(x,\tau)\in V_t\times\mathcal{S}(\mathbf{\Phi}_t)}\deg(x,\tau)\\ &\le \sum_{\tau\in\mathcal{S}(\mathbf{\Phi}_t)\text{ sparse}}\ \sum_{\tau'\in\mathcal{S}(\mathbf{\Phi}_t)}(\theta n-\mathrm{dist}(\tau,\tau'))+\theta n\,|\mathcal{S}(\mathbf{\Phi}_t)|\\ &\quad\cdot|\{\tau\in\mathcal{S}(\mathbf{\Phi}_t):\tau\text{ is not sparse}\}|\\ &\le 0.51\theta n|\mathcal{S}(\mathbf{\Phi}_t)|^2+\theta n\cdot o(|\mathcal{S}(\mathbf{\Phi}_t)|^2)\le 0.511\theta n|\mathcal{S}(\mathbf{\Phi}_t)|^2,\end{aligned}$$

as claimed.

Now, assume that $\mathbf{\Phi}_t$ does indeed satisfy (49), i.e., $2|E(\mathcal{G})|\le 0.511|\mathcal{S}_t(\Phi)|^2\theta n$. Suppose that x is a variable whose marginal does not lie in $(0.01,0.99)$. Then the set $S_x=\{(\tau,x):\tau\in\mathcal{S}(\Phi_t)\}$ spans at least $\binom{0.99\mathcal{S}(\Phi_t)}{2}$ edges. Hence, if we let ν be the number of variables x whose marginal does not lie in $(0.01,0.99)$, then

$$\nu\binom{0.99\mathcal{S}(\Phi_t)}{2}\le|E(\mathcal{G})|\le 0.511|\mathcal{S}_t(\Phi)|^2\theta n/2.$$

Therefore, $\nu\le\frac{0.511+o(1)}{0.99}\theta n\le\frac{2}{3}\theta n$, whence the first part of Theorem 37 follows.

We come to the proof of the second assertion. If $k\theta<\ln(\rho)/2$, then part 3 of Theorem 17 (i.e., the existence of forced variables) immediately implies part 2 of Theorem 37. Thus, let us assume that $\ln(\rho)/2\le k\theta\le\rho/\ln 2$. Let $(\mathbf{\Phi},\boldsymbol{\sigma})$ be a pair chosen from the distribution $\mathcal{U}_k(n,m)$. By the second part of Theorem 17, w.h.p. there is a set R of size $|R|\ge 0.99\theta n$ for which (31) holds, i.e., for any $\tau\in\mathcal{S}(\mathbf{\Phi}_t)$ we have

$$|\{x\in R:\sigma(x)\ne\tau(x)\}|\le k2^{-k}n. \tag{50}$$

Assuming that there is such a set R, we define an auxiliary bipartite graph $\mathcal{B}$ whose vertices are the variables in R and the satisfying assignments $\mathcal{S}(\mathbf{\Phi}_t)$. Each variable $x\in R$ is connected with all $\tau\in\mathcal{S}(\mathbf{\Phi}_t)$ such that $\tau(x)\ne\sigma(x)$. The number of edges of $\mathcal{B}$ satisfies

$$|E(\mathcal{B})|\le k2^{-k}n\,|\mathcal{S}(\mathbf{\Phi}_t)|\,, \tag{51}$$

for (50) implies that each $\tau\in\mathcal{S}(\mathbf{\Phi}_t)$ has degree at most $k2^{-k}n$ in $\mathcal{B}$.

For each variable $x\in R$ the marginal probability that x takes the truth value $1-\boldsymbol{\sigma}(x)$ equals $d(x)/|\mathcal{S}(\mathbf{\Phi}_t)|$. Let ν denote the number of $x\in R$ such that

$$d(x)/|\mathcal{S}(\mathbf{\Phi}_t)|>2^{-k/2}.$$

Then (51) entails that $\nu\le k2^{-k/2}n\le 0.01\theta n$, where the last inequality follows from our assumption that $\theta\ge\ln(\rho)/(2k)$ and that $k\ge k_0,\rho\ge\rho_0$ are sufficiently large. □

7. Belief propagation. The proof of Theorem 7 builds strongly on results from [9]. The main technical contribution of [9] is an analysis of the computation of the belief propagation marginals defined in (7). We will be able to use that analysis largely as a "black box." This is because [9] reduces the analysis of the BP computation to the proof that the outcome $\mathbf{\Phi}_t$ of the experiment U1–U4 enjoys certain quasi-randomness properties. We begin by stating the necessary properties. To this end, we define

$$\delta_t = \exp(-c(1-t/n)k) \text{ for any } 0 \le t \le n, \text{ and } T = \left(1 - \frac{\ln \rho}{c^2 k}\right) n, \tag{52}$$

where $0 < c < 0.1$ is a small absolute constant (independent of k, ρ, t, n).

Fix a k-CNF Φ and an assignment $\sigma \in \{0,1\}^V$. Let $G = G(\Phi, \sigma, t)$ denote the factor graph of $\Phi_{t,\sigma}$. For a vertex v of G we let $N(v)$ denote the neighborhood of v in G. Furthermore, for a variable $x \in V_t$ and a set $Q \subset V_t$ we let

$$N_{\le 1}(x, Q) = \{b \in N(x) : |N(b) \cap Q \setminus \{b\}| \le 1 \wedge 0.1\theta k \le |N(b)| \le 10\theta k\}. \tag{53}$$

Thus, $N_{\le 1}(x, Q)$ is the set of all clauses that contain x (which may or may not be in Q) and at most one other variable from Q. In addition, there is a condition on the *length* $|N(b)|$ of the clause b in the decimated formula $\Phi_{t,\sigma}$. (Observe that having assigned the first t variables, we should expect the average clause length to be θk.)

We call a clause a of Φ *redundant* if there is another clause b such that $|N(a) \cap N(b)| \ge 2$. For a clause b and a variable $x \in N(b)$ we let $\mathrm{sign}(x, b) = 1$ if x occurs in b positively and $\mathrm{sign}(x, b) = -1$ otherwise. Moreover, for a linear map $\Lambda : \mathbf{R}^{V_t} \to \mathbf{R}^{V_t}$ let $\|\Lambda\|_\square$ signify the norm

$$\|\Lambda\|_\square = \max_{\zeta \in \mathbf{R}^{V_t} \setminus \{0\}} \frac{\|\Lambda\zeta\|_1}{\|\zeta\|_\infty}.$$

Definition 38. *Let $\delta > 0$. We say that (Φ, σ) is* (δ, t)-quasi-random *if Φ satisfies* Q0 *and $\Phi_{t,\sigma}$ satisfies* Q1–Q4 *below.*

- Q0. *There are no more than $\ln\ln n$ redundant clauses. Moreover, no variable occurs in more than $\ln n$ clauses of Φ.*
- Q1. *No more than $10^{-5}\delta\theta n$ variables occur in clauses of length less than $\theta k/10$ or greater than $10\theta k$. Moreover, there are at most $10^{-4}\delta\theta n$ variables $x \in V_t$ such that*

$$(\theta k)^3 \delta \cdot \textstyle\sum_{b \in N(x)} 2^{-|N(b)|} > 1.$$

- Q2. *If $Q \subset V_t$ has size $|Q| \le \delta\theta n$, then there are no more than $10^{-4}\delta\theta n$ variables x such that either*

$$\sum_{b \in N(x) : |N(b) \cap Q \setminus \{x\}| = 1} 2^{-|N(b)|} > \rho(\theta k)^5 \delta, \text{ or} \tag{54}$$

$$\sum_{b \in N(x) : |N(b) \cap Q \setminus \{x\}| > 1} 2^{|N(b) \cap Q \setminus \{x\}| - |N(b)|} > \frac{\delta}{\theta k}, \text{ or} \tag{55}$$

$$\left| \sum_{b \in N_{\le 1}(x, Q)} \frac{\mathrm{sign}(x, b)}{2^{|N(b)|}} \right| > \frac{\delta}{1000}. \tag{56}$$

Q3. *For any* $0.01 \leq z \leq 1$ *and any set* $Q \subset V_t$ *of size* $0.01\delta(n-t) \leq |Q| \leq 100\delta(n-t)$ *we have*

$$\sum_{b:|N(b)\cap Q|\geq z|N(b)|} z|N(b)| \leq 1.01|Q|.$$

Q4. *For any set* $Q \subset V_t$ *of size* $|Q| \leq 10\delta(n-t)$ *the linear operator*

$$\Lambda_Q \ : \ \mathbf{R}^{V_t} \to \mathbf{R}^{V_t}, \tag{57}$$
$$\Gamma \mapsto \left(\sum_{b\in N_{\leq 1}(x,Q)} \sum_{y\in N(b)\setminus\{x\}} 2^{-|N(b)|} \cdot \mathrm{sign}(x,b)\mathrm{sign}(y,b)\Gamma_y \right)_{x\in V_t}$$

has norm $\|\Lambda_Q\|_\square \leq \delta^4\theta n$.

Although we will just employ the analysis of the BP mechanics from [9] in a black-box manner, it may be helpful to get an intuitive understanding of the above properties. Condition Q0 is just a well-known general property of random formulas. A necessary condition for the BP formalism to apply is that the factor graph does not have (many) short cycles. Since redundant clause induce cycles of length four, it is unsurprising that we need to impose that there are few of them. Moreover, the condition on the maximum degree of a variable is needed for technical reasons.

We already observed that the average clause length is $k\theta$, and condition Q1 requires that most clauses have about this length. The second condition in Q1 ensures that only few variables occur in many short clauses, where there is a trade-off between the actual length of the clause and the number of clauses of that length that we tolerate.

Conditions Q2 and Q3 are essentially expansion properties of the factor graph of $\Phi_{t,\sigma}$. More precisely, (54) and (55) require that there is no small set $Q \subset V_t$ of variables that is overly strongly connected with the rest of the factor graph. Moreover, (56) imposes that the clauses that x touches and that do not contain another variable from the small forbidden set Q do not have a strong preference as to the value that x should take. That is, the difference between the number of positive and negative occurrences of x in these clauses (weighted by their lengths) is not too big.

Condition Q3 requires that there does not exist a small set Q from which very many clauses contain many variables. Roughly speaking, if we picture the factor graph of $\Phi_{t,\sigma}$ as a sparse random (bipartite) graph, then we would expect that the average degree of a small set is not much more than two. The purpose of Q3 is to impose a suitable formulation of such a condition.

Finally, Q4 is perhaps the most important property. We can think of Λ_Q as a signed, weighted "adjacency operator" on the variables V_t, in which clauses that are exposed to some forbidden set Q are excluded. More precisely, each clause b connecting two variables x, y contributes $2^{-|N(b)|}$ if both occur with the same sign in b, and $-2^{-|N(b)|}$ otherwise. Hence, once more clauses are weighted by their lengths. The operator Λ_Q corresponds to (a projection of the) total derivative of the BP operator at the point where all messages are $1/2$. Hence, Q4 expresses that the derivative of the BP operator is essentially "flat" at that point.

Combined, properties Q0–Q4 basically imply that on $\Phi_{t,\sigma}$ iterating the BP operator ω times for any $\omega \geq 1$ will not produce a large set of variables whose BP marginals deviate significantly from $1/2$. More precisely, these properties imply that any substantial deviations from $1/2$ cannot spread through the formula because such biases will average out due to the expansion properties Q1–Q3 and because the derivative

of the BP operator at the point where all messages are 1/2 (essentially) vanishes due to Q4.

To formalize the notion that most BP marginals are close to 1/2, we need a few more definitions. Let Φ be a k-CNF and let $\delta > 0$. For a number $\delta > 0$ and an index $l > t$ we say that x_l is (δ, t)-*biased* if the result $\mu_{x_l}(\Phi_{t,\sigma}, \omega)$ of the BP computation on $\Phi_{t,\sigma}$ differs from $\frac{1}{2}$ by more than δ, i.e.,

$$|\mu_{x_l}(\Phi_{t,\sigma}, \omega) - 1/2| > \delta.$$

Moreover, (Φ, σ) is (δ, t)-*balanced* if no more than $\delta\theta n$ variables are (δ, t)-biased.

THEOREM 39 (see [9]). *There is $\rho_0 > 0$ such that for any k, r satisfying $\rho_0 \cdot 2^k/k \leq r \leq 2^k \ln 2$ and n sufficiently large the following is true. Suppose (Φ, σ) is (δ_t, t)-quasi-random for some $1 \leq t \leq T$, with δ_t, T as in (52). Then (Φ, σ) is (δ_t, t)-balanced.*

To put Theorem 39 to work, we need to verify that for a pair $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ chosen from the distribution U1–U4 quasi-randomness holds w.h.p.

PROPOSITION 40. *There exists constants $\rho_0 > 0$, $k_0 > 0$ such that for any $k \geq k_0$ and r satisfying $\rho_0 \cdot 2^k/k \leq r \leq 2^k \ln 2$ a pair $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ chosen from the distribution* U1–U4 *is (δ_t, t)-quasi-random w.h.p. for any $t \leq T$ such that $\theta = 1 - t/n \leq \rho/(k \ln 2)$.*

Theorem 7 follows directly by combining Theorem 39 and Proposition 40. Thus, the remaining task is to prove Proposition 40, i.e., to establish the quasi-randomness properties detailed in Definition 38. The following lemma deals with condition Q0.

LEMMA 41 (see [9]). *The random formula $\boldsymbol{\Phi}$ satisfies condition* Q0 *w.h.p. for any density $0 < r = m/n \leq 2^k \ln 2$.*

Furthermore, to prove that a random pair $(\boldsymbol{\Phi}_t, \boldsymbol{\sigma}_t)$ chosen from the distribution U1–U4 satisfies Q1–Q4 w.h.p., we once more take a detour via the planted model. More precisely, below we will show the following.

PROPOSITION 42. *There exists a constant $\rho_0 > 0$ such that for any k, r satisfying $\rho_0 \cdot 2^k/k \leq r \leq 2^k \ln 2$ there is $\xi = \xi(k, r) > 0$ so that for n large and δ_t, T in (52) the following is true. Let $(\boldsymbol{\Phi}', \boldsymbol{\sigma}')$ be a pair chosen from the planted model $\mathcal{P}'_k(n, m)$. Then*

$$\mathrm{P}\left[(\boldsymbol{\Phi}', \boldsymbol{\sigma}') \text{ is } (\delta_t, t)\text{-quasi-random} | \mathrm{Q0}\right] \geq 1 - \exp\left[-\rho 2^{1-k} n\right] \tag{58}$$

for any $t \leq T$ such that $\theta = 1 - t/n \leq \rho/(k \ln 2)$.

Combining Lemma 41, Proposition 42, and Corollary 15 yields Proposition 40.

Proof of Proposition 42. Let $(\boldsymbol{\Phi}', \boldsymbol{\sigma}')$ be a pair chosen from the planted model $\mathcal{P}'_k(n, m)$. Our goal is to establish (58). To simplify the notation, we may and will assume without loss that $\boldsymbol{\sigma}' = \mathbf{1}$ is the all-true assignment. In order to prove Proposition 42, it will be convenient to use a further auxiliary result from [9]. This result deals with a random formula $\boldsymbol{\Phi}''$ on $V = \{x_1, \ldots, x_n\}$ obtained by including each possible clause with probability $p = m/((2^k - 1)\binom{n}{k})$ independently. Thus, $\boldsymbol{\Phi}''$ is a random formula in which the total number of clauses is binomially distributed with mean $\frac{2^k}{2^k - 1} \cdot rn$.

PROPOSITION 43 (see [9, Appendix E]). *There exists a constant $\rho_0 > 0$ such that for any k, r satisfying $\rho_0 \cdot 2^k/k \leq r \leq 2^k \ln 2$ for n large and δ_t, T as in (52) the following properties hold for the random formula $\boldsymbol{\Phi}''$ with probability at least $1 - \exp[-10 \sum_{s \leq t} \delta_s]$ for any $1 \leq t \leq T$ given that $\boldsymbol{\Phi}''$ satisfies* Q0.

1. Q1 *and* Q3 *are satisfied.*
2. *For any set* Q *of size* $|Q| \leq \delta_t \theta n$ *there are at most* $10^{-5}\delta\theta n$ *variables* x *that satisfy either* (54), (55), *or*

$$\left| \sum_{b \in N_{\leq 1}(x,Q)} \frac{\operatorname{sign}(x,b)}{2^{|N(b)|}} \right| > \frac{\delta_t}{2000}. \tag{59}$$

3. *For any* $Q \subset V$ *the operator* Λ_Q *from* (57) *satisfies* $\|\Lambda_Q\|_\square \leq \delta_t^4(n-t)/2$

The following lemma will allow us to establish a connection between the two random formulas $\mathbf{\Phi}'$ and $\mathbf{\Phi}''$.

LEMMA 44. *There exists a constant* $\rho_0 > 0$ *such that for any* k, r *satisfying* $\rho_0 \cdot 2^k/k \leq r \leq 2^k \ln 2$ *the following is true for the random formula* $\mathbf{\Phi}''$ *with probability at least* $1 - \exp(-\rho 2^{2-k} n)$.

1. *The total number of all-negative clauses is bounded by* $2^{1-k}m$.
2. *For each variable* $x \in V_t$ *let* N_x *be the number of all-negative clauses in which* x *appears. Then the number of variables* $x \in V_t$ *with* $N_x > 2^{0.01\theta k}$ *is bounded by* $\delta_t^2 \theta n$.

Proof. The first assertion simply follows from Chernoff bounds. With respect to the second assertion, assume that the first claim occurs, i.e., the total number of all-negative clauses is bounded by $2^{1-k}m = 2\rho n/k$. Then for each variable the average number of occurrences in such clauses is bounded by 2ρ. Therefore, the total number of variables that occur more than $2^{0.01\theta k}$ times is bounded by $2\rho \cdot 2^{-0.01\theta k} n$. By symmetry, the number of such variables that are among the last θn variables is (asymptotically) binomially distribution with mean $2\rho \cdot 2^{-0.01\theta k}\theta n$. Therefore, the second assertion follows from Chernoff bounds as well. ∎

Proof of Proposition 42. Assume that $t \leq T$ is such that $\theta \leq \rho/(k \ln 2)$. As the δ_s from (52) form a geometric sequence, we have $\sum_{s \leq t} \delta_s \sim \frac{n}{ck \exp(c\theta k)}$. Since we are assuming that $\rho \leq k \ln 2$, that $0 < c < 0.1$, $\theta \leq \rho/(k \ln 2)$, we have

$$\sum_{s \leq t} \delta_s \sim \frac{n\left(\exp(ck(1-\theta+1/n)) - 1\right)}{ck \exp(ck)} \geq \frac{\rho n}{2^{k-1}}. \tag{60}$$

The formula $\mathbf{\Phi}'$ is obtained by including each clause that is satisfied under the all-true assignment with probability $p = m/((2^k - 1)\binom{n}{k})$ independently. Hence, we can think of $\mathbf{\Phi}''$ as being obtained by just adding to $\mathbf{\Phi}'$ each of the $\binom{n}{k}$ clauses that are unsatisfied under the all-true assignment (i.e., all-negative clauses) independently with probability p. Thus, with probability at least

$$1 - \exp\left[-10 \sum_{s \leq t} \delta_s\right] \overset{(60)}{\geq} 1 - \exp(-\rho 2^{1-k} n)$$

the formula $\mathbf{\Phi}''$ has the properties 1–3 from Proposition 43. Let us condition on this event.

Since $\mathbf{\Phi}''$ contains $\mathbf{\Phi}'$ as a subformula, the fact that $\mathbf{\Phi}''$ enjoys properties Q1 and Q3 implies directly that the same is true of $\mathbf{\Phi}'$. Furthermore, any variable x for which either (54) or (55) is true in $\mathbf{\Phi}'$ has the same property in $\mathbf{\Phi}''$ (because the expressions on the left-hand side are monotone with respect to the addition of clauses). With respect to the expression in (56), let $Q \subset V_t$ be a set of size $|Q| \leq \delta_t \theta n$. For any

$x \in V_t$ let

$$S_x(\mathbf{\Phi}') = \sum_{b \in N_{\leq 1,\mathbf{\Phi}'}(x,Q)} \frac{\mathrm{sign}(x,b)}{2^{|N_{\mathbf{\Phi}'}(b)|}}, S_x(\mathbf{\Phi}') = \sum_{b \in N_{\leq 1,\mathbf{\Phi}''}(x,Q)} \frac{\mathrm{sign}(x,b)}{2^{|N_{\mathbf{\Phi}''}(b)|}}$$

and let

$$R_x = S_x(\mathbf{\Phi}'') - S_x(\mathbf{\Phi}').$$

Due to Q1, we may assume that all clauses $b \in N_{\leq 1,\mathbf{\Phi}''}(x,Q) \setminus N_{\leq 1,\mathbf{\Phi}'}(x,Q)$ are of length at least $0.1\theta k$. Thus, letting N_x denote the number of clauses in $\mathbf{\Phi}'' \setminus \mathbf{\Phi}'$ containing x, we get $|R_x| \leq 2^{-0.1\theta k} N_x$. The second part of Lemma 44 implies that for all but $\delta_t^2 \theta n$ variables we have $N_x \leq 2^{0.01\theta k}$. Hence,

$$|R_x| \leq 2^{-0.99\theta k} \qquad \text{for all but } \delta^2 \theta n \text{ variables.}$$

This shows in combination with the second part of Proposition 43 that $\mathbf{\Phi}'$ satisfies Q2.

With respect to Q4, let D be the difference of the two linear operators for $\mathbf{\Phi}'$ and $\mathbf{\Phi}''$. Only clauses of length at least $0.1\theta k$ and at most $10\theta k$ contribute to D. Hence, letting N denote the number of all-negative clauses, we have

$$\|D\|_\square \leq 2^{-0.1\theta k} (10\theta k)^2 N.$$

Since $N \leq 2^{1-k} m = 2\rho n/k$ by Lemma 44, we thus get

$$\|D\|_\square \leq 200\theta n (\theta k) 2^{-0.1\theta k}.$$

This implies together with the third part of Proposition 43 that $\mathbf{\Phi}'$ satisfies Q4. ∎

REFERENCES

[1] D. Achlioptas, *Lower bounds for random 3-SAT via differential equations*, Theoret. Comput. Sci., 265 (2001), pp. 159–185.

[2] D. Achlioptas and A. Coja-Oghlan, *Algorithmic barriers from phase transitions*, in Proceedings of the 49th FOCS, 2008, pp. 793–802.

[3] D. Achlioptas, A. Coja-Oghlan, and F. Ricci-Tersenghi, *On the solution space geometry of random formulas*, Random Structures Algorithms, 38 (2011), pp. 251–268.

[4] D. Achlioptas and C. Moore, *Random k-SAT: Two moments suffice to cross a sharp threshold*, SIAM J. Comput., 36 (2006), pp. 740–762.

[5] D. Achlioptas and Y. Peres, *The threshold for random k-SAT is* $2^k \ln 2 - O(k)$, J. AMS, 17 (2004), pp. 947–973.

[6] D. Achlioptas and F. Ricci-Tersenghi, *Random formulas have frozen variables*, SIAM J. Comput., 39 (2009), pp. 260–280.

[7] A. Braunstein, M. Mézard, and R. Zecchina, *Survey propagation: An algorithm for satisfiability*, Random Structures Algorithms, 27 (2005), pp. 201–226.

[8] A. Coja-Oghlan, *A better algorithm for random k-SAT*, SIAM J. Computing, 39 (2010), pp. 2823–2864.

[9] A. Coja-Oghlan, *On belief propagation guided decimation for random k-SAT*, in Proceedings of the 22nd SODA, 2011, pp. 957–966.

[10] H. Daudé, M. Mézard, T. Mora, and R. Zecchina, *Pairs of SAT-assignments in random Boolean formulae*, Theoret. Comput. Sci., 393 (2008), pp. 260–279.

[11] U. Feige, A. Flaxman, and D. Vilenchik, *On the diameter of the set of satisfying assignments in random satisfiable k-CNF formulas*, SIAM J. Discrete Math., 25 (2011), pp. 736–749.

[12] A. Frieze and S. Suen, *Analysis of two simple heuristics on a random instance of k-SAT*, J. Algorithms, 20 (1996), pp. 312–355.

[13] M. Hajiaghayi and G. Sorkin, *The Satisfiability Threshold of Random 3-SAT is at Least 3.52*, Research report RC22942, IBM, Armonk, NY, 2003.
[14] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.
[15] A. Kaporis, L. Kirousis, and E. Lalas, *The probabilistic analysis of a greedy satisfiability algorithm*, Random Structures Algorithms, 28 (2006), pp. 444–480.
[16] L. Kroc, A. Sabharwal, and B. Selman, *Message-passing and local heuristics as decimation strategies for satisfiability*, in Proceedings of the 24th SAC, 2009, pp. 1408–1414.
[17] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborova, *Gibbs states and the set of solutions of random constraint satisfaction problems*, Proc. Natl. Acad. Sci. USA, 104 (2007), pp. 10318–10323.
[18] S. Mertens, M. Mézard, and R. Zecchina, *Threshold values of random K-SAT from the cavity method*, Random Structures Algorithms, 28 (2006), pp. 340–373.
[19] M. Mézard, G. Parisi, and R. Zecchina, *Analytic and algorithmic solution of random satisfiability problems*, Science, 297 (2002), pp. 812–815.
[20] D. Mitchell, B. Selman, and H. Levesque, *Hard and easy distribution of SAT problems*, in Proceedings of the 10th AAAI, 1992, pp. 459–465.
[21] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian, *Solving constraint satisfaction problems through Belief Propagation-guided decimation*, in Proceedings of the 45th Annual Allerton Conference on Communication, Control, and Computing, 2007.
[22] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Francisco, 1988.
[23] F. Ricci-Tersenghi and G. Semerjian, *On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms*, J. Stat. Mech., (2009), P09001.
[24] N. Wormald, *The differential equation method for random graph processes and greedy algorithms*, in Lectures on Approximation and Randomized Algorithms, M. Karoński and H.J. Prömel, eds., Polish Scientific Publishers, 1999, pp. 73–155.