

The Decimation Process in Random k -SAT

Amin Coja-Oghlan and Angelica Y. Pachon-Pinzon*

University of Warwick, Mathematics and Computer Science,
Coventry CV4 7AL, UK
{a.coja-oghan,a.y.pachon-pinzon}@warwick.ac.uk

Abstract. Non-rigorous statistical mechanics ideas have inspired a message passing algorithm called *Belief propagation guided decimation* for finding satisfying assignments of random k -SAT instances. 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.

1 Introduction

Let $k \geq 3$ and $n > 1$ be integers, let $r > 0$ be a real, and set $m = \lceil rn \rceil$. Let $\Phi = \Phi_k(n, m)$ be a propositional formula obtained by choosing a set of m clauses of length k over the variables $V = \{x_1, \dots, x_n\}$ uniformly at random. For k, r fixed we say that Φ has some property \mathcal{P} with high probability ('w.h.p.') if $\lim_{n \rightarrow \infty} \Pr[\Phi \in \mathcal{P}] = 1$.

The interest in random k -SAT originates from the experimental observation that for certain densities r the random formula Φ is satisfiable w.h.p. while a large class of algorithms, including and particularly the workhorses of practical SAT solving such as sophisticated DPLL-based solvers, fail to find a satisfying assignment efficiently [14]. Over the past decade, a fundamentally new class of algorithms have been proposed on the basis of ideas from statistical physics [6,13]. 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/SP decimation’), excel on random k -SAT instances [10]. Indeed, the experiments indicate that BP/SP decimation find satisfying assignments for r close to the threshold where Φ becomes unsatisfiable w.h.p. Generally, SP is deemed conceptually superior to BP.

For example, in the case $k = 4$ the threshold for the existence of satisfying assignments is conjectured to be $m/n \sim r_4 \approx 9.93$ [12]. According to experiments from [10], SP decimation finds satisfying assignments for densities up to $r = 9.73$. Experiments from [16] 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 [6]) succeeds up to $r = 9.24$, again according to experimental data from [10]. By comparison, the currently best rigorously

* Supported by EPSRC grant EP/G039070/2.

analyzed algorithm is efficient up to $r = 5.54$ [9], while **zChaff**, a prominent practical SAT solver, becomes ineffective beyond $r = 5.35$ [10].

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/SP decimation are based on are highly non-rigorous, and thus a rigorous analysis of these message passing algorithms is an important but challenging open problem. A first step was made in [7], 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 [7] 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, non-constructive arguments show that w.h.p. Φ is satisfiable if $r < r_k = 2^k \ln 2 - k$, and unsatisfiable if $r > 2^k \ln 2$ [3,4]. 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 [7] 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, [7] 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 the BP decimation algorithm aims 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 [6,13] but turns out to have a crucial impact on the performance of the algorithm. On a non-rigorous basis, this has been pointed out recently by Ricci-Tersenghi and Semerjian [16], and our results can be viewed as providing a rigorous version of (parts of) their main results. The results of this paper can also be seen as a generalization of the ones obtained in [1] for random k -SAT, and indeed our proofs build heavily upon the techniques developed in that paper.

2 Results

BP decimation is a polynomial-time algorithm that aims to (heuristically) implement the ‘thought experiment’ shown in Fig. 1 [15,16], which we call the *decimation process*.¹ 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

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 [7,15,16]. Other versions decimate the variables in a different order, allowing for slightly better experimental results [6,10].

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 computed efficiently on *random* formulas by means of a message passing algorithm. We will return to the discussion of BP decimation and its connection to Experiment 1 below.

Experiment 1 (‘decimation process’). *Input:* A satisfiable k -CNF Φ .

Result: A satisfying assignment $\sigma : V \rightarrow \{0, 1\}$ (with 0/1 representing ‘false’/‘true’).

0. Let $\Phi_0 = \Phi$.
1. For $t = 1, \dots, n$ 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 got satisfied by assigning x_t , and omit x_t from all other clauses).
5. Return the assignment σ .

Fig. 1. The decimation process

We are going to study the decimation process when applied to a random formula Φ for densities $r < 2^k \ln 2 - k$, i.e., in the regime where Φ is satisfiable w.h.p. More precisely, conditioning on Φ being satisfiable, we let Φ_t be the (random) formula obtained after running the first t iterations of Experiment 1. The variable set of this formula is $V_t = \{x_{t+1}, \dots, x_n\}$, and each clause of Φ_t consists of *at most* k literals. Let $\mathcal{S}(\Phi_t) \subset \{0, 1\}^{V_t}$ be the set of all satisfying assignments of Φ_t . We say that *almost all* $\sigma \in \mathcal{S}(\Phi_t)$ have a certain property \mathcal{A} if $|\mathcal{A} \cap \mathcal{S}(\Phi_t)| = (1 - o(1))|\mathcal{S}(\Phi_t)|$.

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

$$\rho = kr/2^k \quad \text{and} \quad \theta = 1 - t/n,$$

so that $m/n \sim \rho \cdot 2^k/k$. We will be interested in the regime $\rho_0 \leq \rho \leq k \ln 2$, where ρ_0 is a constant (independent of k). The upper bound $k \ln 2$ marks the point where satisfying assignments cease to exist [4]. The second parameter θ is the fraction of ‘free’ variables (i.e., variables not yet assigned by time t).

The symmetric phase. Let Φ be a k -CNF on V , let $1 \leq t < n$, let Φ_t be the formula obtained after t steps of the decimation process, and suppose that $\sigma \in \mathcal{S}(\Phi_t)$. A variable $x \in V_t$ is *loose* if there is $\tau \in \mathcal{S}(\Phi_t)$ such that $\sigma(x) \neq \tau(x)$ and $d(\sigma, \tau) \leq \ln n$, where $d(\cdot, \cdot)$ denotes the Hamming distance. For any $x \in V_t$ we let $M_x(\Phi_t) = |\{\sigma \in \mathcal{S}(\Phi_t) : \sigma(x) = 1\}| / |\mathcal{S}(\Phi_t)|$ be the marginal probability that x takes the value ‘true’ in a random satisfying assignment of Φ_t .

Theorem 2. *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*

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

the random formula Φ_t has the following properties w.h.p.

1. *In almost all $\sigma \in \mathcal{S}(\Phi_t)$ at least $0.99\theta n$ variables are loose.*
2. *At least $\theta n/3$ variables $x \in V_t$ satisfy $M_x(\Phi_t) \in [0.01, 0.99]$.*
3. *The average distance of two random satisfying assignments is $\geq 0.49\theta n$.*

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 $\sigma \in \mathcal{S}(\Phi_t)$ are loose. Hence, the correlations amongst 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 $M_x(\Phi_t)$ are bounded away from 0/1. Finally, as the distance between satisfying assignments is large on average, the set $\mathcal{S}(\Phi_t)$ 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 $d(\sigma, \tau) \geq \omega$. 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 $\text{dist}(\sigma, \tau) \geq \beta\theta n$.

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*

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

the random formula Φ_t has the following properties w.h.p.

1. *In almost all $\sigma \in \mathcal{S}(\Phi_t)$ at least $0.99\theta n$ variables are $\Omega(n)$ -rigid.*
2. *There exist $\alpha = \alpha(k, \rho) > 0$, $\beta = \beta(k, \rho) > 0$ such that $\mathcal{S}(\Phi_t)$ is (α, β) -shattered.*
3. *At least $\theta n/3$ variables $x \in V_t$ satisfy $M_x(\Phi_t) \in [0.01, 0.99]$.*
4. *The average distance of two random satisfying assignments is at least $0.49\theta n$.*

Thus, if the fraction θ of free variables lies in the regime (1), then in most satisfying $\sigma \in \mathcal{S}(\Phi_t)$ 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 $\mathcal{S}(\Phi_t)$: 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 free variables remain bounded away from 0/1, and the set $\mathcal{S}(\Phi_t)$ 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 $\text{dist}(\sigma, \tau) \leq \alpha n$.

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*

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

the random formula Φ_t has the following properties w.h.p.

1. *In almost all $\sigma \in \mathcal{S}(\Phi_t)$ at least $0.99\theta n$ variables are $\Omega(n)$ -rigid.*
2. *The set $\mathcal{S}(\Phi_t)$ is $\exp(2 - \rho)/k$ -ferromagnetic.*
3. *At least $0.99\theta n$ variables $x \in V_t$ satisfy $M_x(\Phi_t) \in [0, 2^{-k/2}] \cup [1 - 2^{-k/2}, 1]$.*
4. *There is a set $R \subset V_t$ of size $|R| \geq 0.99\theta n$ such that for any $\sigma, \tau \in \mathcal{S}(\Phi_t)$ we have $|\{x \in R : \sigma(x) \neq \tau(x)\}| \leq k2^{-k}n$.*

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 ferromagnetic subset of $\{0, 1\}^{V_t}$ of tiny diameter, in contrast to a well-spread shattered set as in Theorem 3. Furthermore, most marginals $M_x(\Phi_t)$ 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 $k2^{-k}n$ 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*

$$1/n \ll k \cdot \theta < \ln(\rho)(1 - 10/\ln \rho) \quad (3)$$

the random formula Φ_t has the following properties w.h.p.

1. *At least $0.99\theta n$ variables are forced.*
2. *The set $\mathcal{S}(\Phi_t)$ is $\exp(2 - \rho)/k$ -ferromagnetic.*

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 in [6]. Basically, they 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 is

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 $|\mu_{x_t}(\Phi_{t-1}, \omega) - M_{x_t}(\Phi_{t-1})| < \varepsilon$.*

In other words, Hypothesis 6 states that throughout the 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*

$$c_0 \ln(\rho) < k \cdot \theta < \rho / \ln 2. \quad (4)$$

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

The proof is based on the techniques developed in [7]; the details are omitted from this extended abstract. Comparing Theorem 4 with Theorem 7, we see that w.h.p. for θ satisfying (4) most of the ‘true’ marginals $M_x(\Phi_t)$ 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.

Corollary 1. *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.*

Summary and discussion. Fix $k \geq k_0$ and $\rho \geq \rho_0$. Theorems 2–5 show how the space of satisfying assignments of Φ_t evolves as the decimation process progresses. 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 amongst 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 $\mathcal{S}(\Phi_t)$ 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 amongst the possible values that can be assigned to the individual variables (‘rigid variables’). This phenomenon goes by the name of *dynamic replica symmetry breaking* in statistical mechanics [11].

While in the previous phases the set of satisfying assignments is scattered all over the Hamming cube (as witnessed by the average Hamming distance of two satisfying assignments), 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). As part 4 of Theorem 4 shows, the mistaken estimates of the Belief Propagation computation would make it impossible for BP decimation to penetrate the ferromagnetic phase. More precisely, even if BP decimation would emulate the decimation process perfectly up until the ferromagnetic phase commences, with probability $1 - \exp(-\Omega(n))$ BP decimation would then assign at least $k2^{-k}n$ variables in the set R from part 4 of Theorem 4 ‘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 reach this regime.

3 Related Work

BP/SP decimation are inspired by a generic but highly non-rigorous analysis technique from statistical mechanics called the *cavity method* [6]. This technique is primarily destined for the *analysis* of phase transitions. In [6,11] the cavity method was used to study the structure of the set $\mathcal{S}(\Phi)$ of satisfying assignments (or, more accurately, properties of the Gibbs measure) of the *undecimated* random formula Φ . Thus, the results obtained in that (non-rigorous) 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 $\rho \sim k \ln 2$ [11]. The argument given for the latter scenario in [11] is that the key obstacle for BP to approximate the true marginals is *condensation*, a phenomenon that from the viewpoint of BP is very similar to ferromagnetism. In terms of the parameter ρ , the condensation threshold was (non-rigorously) estimated to occur at $\rho = k \ln 2 - 3k2^{-k-1} \ln 2$. However, [7] 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 [6,11] 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 non-rigorous basis an analysis both in terms of the formula density ρ and the time parameter θ was carried out in [16]. Thus, our results can be viewed as a rigorous version of parts of [16] (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 constant independent of k .

Theorem 3 can be viewed as a generalization of the results on random k -SAT obtained in [1] (which additionally deals with further problems such as random graph/hypergraph coloring). In [1] we rigorously proved a substantial part of the results hypothesized in [11] on shattering and rigidity in terms of the clause/variable density ρ ; this improved prior work [2,5,8]. 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 [1] 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 Φ has the properties 1.–4. stated in Theorem 3. Technically, the present paper builds upon the methods developed in [1].

4 Analyzing the Decimation Process

In the rest of the paper, we are going to sketch the proofs of the main results. In this section we perform some groundwork to facilitate a rigorous analysis of the decimation process. The key problem is to get a handle on the following experiment:

- D1.** Generate a random formula Φ , conditioned on Φ being satisfiable.
- D2.** Run the decimation process for t steps to obtain Φ_t .
- D3.** Choose a satisfying assignment $\sigma_t \in \mathcal{S}(\Phi_t)$ uniformly at random.
- D4.** The result is the pair (Φ_t, σ_t) .

As throughout the paper we only work with densities m/n where Φ 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 Φ , we see that the following experiment is equivalent to **D1–D4**:

- U1.** Generate a random formula Φ , conditioned on Φ being satisfiable.
- U2.** Choose $\sigma \in \mathcal{S}(\Phi)$ uniformly at random.
- U3.** Substitute $\sigma(x_i)$ for x_i for $1 \leq i \leq t$ and simplify to obtain a formula Φ_t .
- U4.** The result is the pair (Φ_t, σ_t) , where $\sigma_t : V_t \rightarrow \{0, 1\}$, $x \mapsto \sigma(x)$.

Fact 8. *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 [1], we will surmount this problem by considering yet another experiment.

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

The experiment **P1–P4** is easy to implement and, in effect, also amenable to a rigorous analysis. For given the assignment σ' , there are $(2^k - 1)\binom{n}{k}$ clauses in total that evaluate to ‘true’ under σ' , and to generate Φ' 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 [1] 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 $V = \{x_1, \dots, x_n\}$ 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 9 ([1]). *Suppose $k \geq 4$ and $0 < \rho < k \ln 2 - k^2/2^k$. Let $\mathcal{E} \subset \Lambda_k(n, m)$. If $\Pr_{\mathcal{P}_k(n, m)}[\mathcal{E}] \geq 1 - \exp(-\rho n/2^k)$ then $\Pr_{\mathcal{U}_k(n, m)}[\mathcal{E}] = 1 - o(1)$.*

5 Shattering, Pairwise Distances, and Ferromagnetism

To prove shattering and ferromagnetism, we adapt arguments from [1,2,8] to the situation where we have the *two* parameters θ, ρ (rather than just ρ). Let (Φ_t, σ_t) be the (random) outcome of the experiment **U1–U4**. For $0 \leq \alpha \leq 1$ let $X_\alpha(\Phi_t, \sigma_t)$ denote the number of satisfying assignments $\tau \in \mathcal{S}(\Phi_t)$ with Hamming distance $d(\sigma_t, \tau) = \alpha \theta n$. To establish the ‘shattering’ part of Theorem 3, we are going to prove the following

Claim 10. *Under the assumptions of Theorem 3 there exist $a_1 < a_2 < 0.49$, $a_3 > 0$ depending only on k, ρ such that w.h.p. we have*

$$X_\alpha(\Phi_t, \sigma_t) = 0 \quad \text{for all } a_1 < \alpha < a_2, \text{ and} \tag{5}$$

$$\max_{\alpha \leq 0.49} X_\alpha(\Phi_t, \sigma_t) < \exp(-a_3 n) \cdot |\mathcal{S}(\Phi_t)|. \tag{6}$$

Claim 10 implies that for the outcome Φ_t of the first t steps of the decimation process the set $\mathcal{S}(\Phi_t)$ shatters w.h.p. For by Fact 8 Claim 10 implies that w.h.p. almost all $\sigma_t \in \mathcal{S}(\Phi_t)$ are such that (5) and (6) hold. Choose any such $\sigma_{t,1} \in \mathcal{S}(\Phi_t)$ and let $R_1 = \{\tau \in \mathcal{S}(\Phi_t) : d(\tau, \sigma_{t,1}) \leq a_1 n\}$. Then, choose $\sigma_{t,2} \in \mathcal{S}(\Phi_t) \setminus R_1$ satisfying (5) and (6), let $R_2 = \{\tau \in \mathcal{S}(\Phi_t) \setminus R_1 : d(\tau, \sigma_{t,2}) \leq a_1 n\}$, and proceed inductively until all remaining satisfying assignments violate either (5) or (6). Let R_1, \dots, R_N be the classes constructed in this way and let $R_0 = \mathcal{S}(\Phi_t) \setminus \bigcup_{i=1}^N R_i$. An additional (simple) argument is needed to show that $|R_0| \leq \exp(-\Omega(n))|\mathcal{S}(\Phi_t)|$ w.h.p. The decomposition R_0, \dots, R_N witnesses that $\mathcal{S}(\Phi_t)$ shatters.

With respect to pairwise distances of satisfying assignments, (6) implies that w.h.p. only an exponentially small fraction of all satisfying assignments of Φ_t lies within distance $\leq 0.49\theta n$ of σ_t . It is not difficult to derive the statement made in Theorem 3 on the average pairwise distance from this. In addition, the fact that the average pairwise distance of satisfying assignments is $\geq 0.49\theta n$ w.h.p. implies in combination with a double counting argument the claim about the marginals $M_x(\Phi_t)$ in Theorems 2 and 3.

To establish Claim 10 we will work with the experiment **P1–P4** and use Theorem 9 to transfer the result to the experiment **U1–U4**. Thus, let (Φ'_t, σ'_t) be the (random) outcome of experiment **P1–P4**, and assume that k, ρ, θ are as in Theorem 3. To prove (5) we need to bound $X_\alpha(\Phi'_t, \sigma'_t)$ from above, for which we use the ‘first moment method’. Indeed, by standard arguments (similar to those used in [2]) the expectation of $X_\alpha(\Phi'_t, \sigma'_t)$ satisfies $\frac{1}{n} \ln \mathbb{E} X_\alpha(\Phi'_t, \sigma'_t) \leq \psi(\alpha)$, with

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

Thus, in order to prove that $\max_{a_1 < \alpha < a_2} X_\alpha(\Phi'_t, \sigma'_t) = 0$ w.h.p. we would just have to prove that $\max_{a_1 < \alpha < a_2} \psi(\alpha) < 0$ (so that Markov’s inequality implies that $X_\alpha = 0$ w.h.p.). But as our goal is to prove a result about the $X_\alpha(\Phi_t, \sigma_t)$ (i.e., the experiment **U1–U4**), we need to prove a slightly stronger bound, namely $\max_{a_1 < \alpha < a_2} \psi(\alpha) < -\rho/2^k$. Then Markov’s inequality and Theorem 9 imply the first part of Claim 10. Via elementary calculus, one can show that the aforementioned bound holds with $a_1 = \exp(2 - \rho) - \varepsilon$ and $a_2 = \exp(2 - \rho) + \varepsilon$ for a sufficiently small $\varepsilon > 0$.

To prove (6) we bound $\mathbb{E} X_\alpha$ from above by a similar first moment argument. But in addition, we need a lower bound on $|\mathcal{S}(\Phi_t)|$. To derive this, we need

Theorem 11 ([2]). *Assume $k \geq 4$ and $\rho \leq k \ln 2 - k^2/2^k$. Then w.h.p. $\frac{1}{n} \ln |\mathcal{S}(\Phi)| \geq \ln 2 + 2^k \frac{\rho}{k} \ln(1 - 2^{-k}) - 0.99\rho/2^k$.*

Together with a double counting argument, Theorem 11 implies the part 2 of Claim 10. The ‘ferromagnetism’ bit of Theorem 4 follows from similar arguments.

6 Rigid Variables

Assume that k, ρ, θ satisfy the assumptions of Theorem 3. Let (Φ_t, σ_t) be the (random) outcome of **U1–U4**. Our goal is to show that w.h.p. most variables $x \in V_t$ are rigid.

What is the basic obstacle that makes it difficult to ‘flip’ the value of x ? Observe that we can simply assign x the opposite value $1 - \sigma_t(x)$, unless Φ_t has a clause \mathcal{C} in which either x or \bar{x} is the *only* literal that is true under σ_t . If there is such a clause, we say that x *supports* \mathcal{C} . But even if x supports a clause \mathcal{C} it might be easy to flip. For instance, if \mathcal{C} features some variable $y \neq x$ that does not support a clause, then we could just flip both x, y simultaneously. Thus, to establish the existence of $\Omega(n)$ -rigid variables 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.

This analysis can be performed fairly neatly for the outcome (Φ'_t, σ'_t) of the experiment **P1–P4**. Let us sketch how this works, and why rigidity occurs at $k\theta = \exp((1+o(1))\rho)$ (cf. (1)). For a variable $x \in V_t$ we let S_x be the number of

clauses supported by x . Given the assignment σ' 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 σ' uniformly and independently, we get $E[S_x] = m\binom{n-1}{k-1} / ((2^k - 1)\binom{n}{k}) = \rho/(1 - 2^{-k}) \geq \rho$. In fact, S_x is binomially distributed. Hence, $P[S_x = 0] \leq \exp(-\rho)$. 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 $\mathcal{C}_1, \dots, \mathcal{C}_j$ supported by x are just independently 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 one of these clauses \mathcal{C}_i is $(1 + o(1))(k-1) \cdot \theta \exp(-\rho) \leq k\theta \exp(-\rho)$. Hence, if θ is as in (1), then this number is $\leq \exp(-2)/\rho$, i.e., ‘small’ for $\rho \geq \rho_0$ sufficiently big. Thus, we would expect that *most* clauses supported by x indeed consist exclusively of variables that support other clauses. Hence, for θ as in (1) we can expect most variables to be rigid.

Let us now indicate how this argument can be carried out in detail. Analyzing the distribution of the variables S_x in the experiment **P1–P4** and extending the result to the experiment **U1–U4** via Theorem 9, and setting $\zeta = \rho^2 / \exp(\rho)$, we obtain the following.

Proposition 1. *Suppose that k, ρ, θ satisfy the assumptions of Theorem 3. Then w.h.p. in a random pair (Φ_t, σ_t) generated by the experiment **U1–U4** no more than $2\zeta\theta n$ variables in V_t support fewer than three clauses,*

To establish rigidity, we need to show that most variables support clauses in which only variables occur that support other clauses. To express this, we say that $S \subset V_t$ is *t-self-contained* if each $x \in S$ supports at least two clauses of Φ_t that contain variables from S only. From Proposition 1 we can derive

Proposition 2. *Suppose that k, ρ, θ satisfy the assumptions of Theorem 3. The outcome (Φ_t, σ_t) of **U1–U4** has a t-self-contained set of size $(1 - 3\zeta)\theta n$ w.h.p.*

Suppose that (Φ_t, σ_t) has a self-contained set S of size $(1 - 3\zeta)\theta n$. To flip the value of a variable $x \in S$ we need to also flip one other variable from each of the (at least two) clauses that x supports and that consist of variables from S only. As each of these two variables, in turn, supports at least two clauses comprised of variables from S only, we need to also flip further variables in those. But these variables are again contained in S . This suggests that attempting to flip x will entail an avalanche of further flips. Indeed, the expansion properties of the random formula Φ_t imply the following.

Proposition 3. *With k, ρ, θ as in Theorem 3 there is $\chi = \chi(k, \rho) > 0$ such that the outcome (Φ_t, σ_t) of **U1–U4** has the following property w.h.p.: all variables that are contained in a t-self-contained set are χn -rigid.*

Propositions 2 and 3 directly imply part 1 of Theorem 3. Self-contained sets also play a key role in the proof of Theorem 4 (details omitted).

References

1. Achlioptas, D., Coja-Oghlan, A.: Algorithmic barriers from phase transitions. In: Proc. 49th FOCS, pp. 793–802 (2008)
2. Achlioptas, D., Coja-Oghlan, A., Ricci-Tersenghi, F.: On the solution space geometry of random formulas. *Random Structures and Algorithms* 38, 251–268 (2011)
3. Achlioptas, D., Moore, C.: Random k -SAT: two moments suffice to cross a sharp threshold. *SIAM Journal on Computing* 36, 740–762 (2006)
4. Achlioptas, D., Peres, Y.: The threshold for random k -SAT is $2^k \ln 2 - O(k)$. *Journal of the AMS* 17, 947–973 (2004)
5. Achlioptas, D., Ricci-Tersenghi, F.: Random formulas have frozen variables. *SIAM J. Comput.* 39, 260–280 (2009)
6. Braunstein, A., Mézard, M., Zecchina, R.: Survey propagation: an algorithm for satisfiability. *Random Structures and Algorithms* 27, 201–226 (2005)
7. Coja-Oghlan, A.: On belief propagation guided decimation for random k -SAT. In: Proc. 22nd SODA, pp. 957–966 (2011)
8. Daudé, H., Mézard, M., Mora, T., Zecchina, R.: Pairs of SAT-assignments in random Boolean formulae. *Theoretical Computer Science* 393, 260–279 (2008)
9. Frieze, A., Suen, S.: Analysis of two simple heuristics on a random instance of k -SAT. *Journal of Algorithms* 20, 312–355 (1996)
10. Kroc, L., Sabharwal, A., Selman, B.: Message-passing and local heuristics as decimation strategies for satisfiability. In: Proc. 24th SAC, pp. 1408–1414 (2009)
11. Krzakala, F., Montanari, A., Ricci-Tersenghi, F., Semerjian, G., Zdeborova, L.: Gibbs states and the set of solutions of random constraint satisfaction problems. *Proc. National Academy of Sciences* 104, 10318–10323 (2007)
12. Mertens, S., Mézard, M., Zecchina, R.: Threshold values of random K -SAT from the cavity method. *Random Struct. Alg.* 28, 340–373 (2006)
13. Mézard, M., Parisi, G., Zecchina, R.: Analytic and algorithmic solution of random satisfiability problems. *Science* 297, 812–815 (2002)
14. Mitchell, D., Selman, B., Levesque, H.: Hard and easy distribution of SAT problems. In: Proc. 10th AAAI, pp. 459–465 (1992)
15. Montanari, A., Ricci-Tersenghi, F., Semerjian, G.: Solving constraint satisfaction problems through Belief Propagation-guided decimation. In: Proc. 45th Allerton (2007)
16. Ricci-Tersenghi, F., Semerjian, G.: On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms. *J. Stat. Mech.*, 09001 (2009)