On the Solution-Space Geometry of Random Constraint Satisfaction Problems

Dimitris Achlioptas[*]
Department of Computer Science
University of California Santa Cruz
optas@cs.ucsc.edu

Federico Ricci-Tersenghi
Department of Physics
University of Rome "La Sapienza"
federico.ricci@roma1.infn.it

ABSTRACT

For a number of random constraint satisfaction problems, such as random k-SAT and random graph/hypergraph coloring, there are very good estimates of the largest constraint density for which solutions exist. Yet, all known polynomial-time algorithms for these problems fail to find solutions even at much lower densities. To understand the origin of this gap we study how the structure of the space of solutions evolves in such problems as constraints are added. In particular, we prove that much before solutions disappear, they organize into an exponential number of clusters, each of which is relatively small and far apart from all other clusters. Moreover, inside each cluster most variables are frozen, i.e., take only one value. The existence of such frozen variables gives a satisfying intuitive explanation for the failure of the polynomial-time algorithms analyzed so far. At the same time, our results establish rigorously one of the two main hypotheses underlying Survey Propagation, a heuristic introduced by physicists in recent years that appears to perform extraordinarily well on random constraint satisfaction problems.

Categories and Subject Descriptors

F.2.2 [**Nonnumerical Algorithms and Problems**]: Computations on discrete structures; G.3 [**Probability and Statistics**]: Stochastic Processes

General Terms

Algorithms, Theory

Keywords

Random formulas, Satisfiability, Survey Propagation

1. INTRODUCTION

For a number of random Constraint Satisfaction Problems (CSP), by now we have very good estimates of the largest constraint density (ratio of constraints to variables) for which typical problems have solutions. For example [3], a random graph of average degree d is with high probability[1] k-colorable if $d < (2k-2)\ln(k-1)$, but w.h.p. non-k-colorable if $d > (2k-1)\ln k$. This implies that for every $d > 0$, the chromatic number of a random graph with average degree d is either k_d or $k_d + 1$, where k_d is the smallest integer k such that $d < 2k \ln k$.

Algorithmically, it is very easy to get a factor-2 approximation for the graph coloring problem on random graphs. The algorithm "repeatedly pick a random vertex and assign it a random available color" will w.h.p. succeed in coloring a random graph of average degree d if originally each vertex has $2k_d$ available colors. Alternatively, k colors suffice when $d < k \ln k$. In spite of significant efforts over the last 30 years, no improvement has been made over this trivial algorithm. Specifically, no polynomial-time algorithm is known that k-colors random graphs of average degree $d = (1+\epsilon)k \ln k$, for some fixed $\epsilon > 0$ and arbitrarily large k.

In the random k-SAT problem one asks if a random k-CNF formula, $F_k(n, m)$, with n variables and m clauses is satisfiable. It is widely believed that the probability that such a formula is satisfiable exhibits a sharp threshold. Specifically, the *Satisfiability Threshold Conjecture* asserts that $r_k = r_k^*$ for all $k \geq 3$, where

$$\begin{aligned} r_k &\equiv \sup\{r : F_k(n, rn) \text{ is satisfiable w.h.p.}\} , \\ r_k^* &\equiv \inf\{r : F_k(n, rn) \text{ is unsatisfiable w.h.p.}\} . \end{aligned}$$

It is easy to see that $r_k^* \leq 2^k \ln 2$, since the probability that at least one assignment satisfies $F_k(n, rn)$ is bounded by $2^n(1 - 2^{-k})^{rn}$, a quantity that tends to 0 for $r \geq 2^k \ln 2$. Recently, it was shown that random k-CNF formulas have satisfying assignments for densities very close to this upper bound [5]. Specifically, it was proven that for all $k \geq 3$,

$$r_k > 2^k \ln 2 - \frac{(k+1)\ln 2 + 3}{2} \quad . \tag{1}$$

As for the k-coloring problem, the lower bound of [5] on the largest density for which solutions provably exist is non-constructive, based on the second moment method. Here, the gap relative to algorithms is ever greater: no polynomial algorithm is known that finds satisfying assignments in a random k-CNF formula when $r = \omega(k)\, 2^k/k$, for any function $\omega(k) \to \infty$ (arbitrarily slowly). In Table 1, we illustrate this gap for some small values of k. For $k = 3$, the upper bound on r_k^* comes from [12], while for $k > 3$ from [11, 18]. The best algorithmic lower bound for $k = 3$ is from [17], while for $k > 3$ it is from [14].

Similar results (and gaps) exist for a number of other constraint satisfaction problems, such as random NAE k-SAT and hypergraph

[*] Work supported by National Science Foundation Faculty Early CAREER Development Award CCF–0546900.

STOC'06, May 21–23, 2006, Seattle, Washington, USA.
Copyright 2006 ACM 1-59593-134-1/06/0005 ...$5.00.

[1] We will say that a sequence of events $\mathcal{E}_n$ occurs with high probability (w.h.p.) if $\lim_{n\to\infty} \Pr[\mathcal{E}_n] = 1$.

k	3	4	7	10	20	21
Best known upper bound for r_k^*	4.506	10.23	87.88	708.94	$726,817$	$1,453,635$
Best known lower bound for r_k	3.52	7.91	84.82	704.94	$726,809$	$1,453,626$
Best known algorithmic lower bound	3.52	5.54	33.23	172.65	$95,263$	$181,453$

2-coloring, regular random graph coloring, random Max k-SAT, and others (for example, see [4]). Indeed, this phenomenon seems to occur in nearly all random CSP in which the underlying constraint graph is sparse and random, making it natural to ask if there is a common underlying cause. (The bipartite graphs where constraints are adjacent to the variables they bind are also known as factor graphs.)

As it turns out, sparse random CSP have been systematically studied by physicists in the past few decades under the name "mean-field diluted spin-glasses". Spins here are the variables (reflecting the notion that variables have small, discrete domains), the term glass refers to the fact that the system has not been allowed to relax to a configuration in which spins interact in a mutually agreeable way (reflecting that different constraints prefer different values for the variables), diluted refers to the fact that the factor graph is sparse (reflecting that each spin interacts with only a few other spins), while "mean field" refers to the fact that the factor graph is random, i.e., there is no underlying geometry mandating the interactions. The interest in such "unphysical" systems is partly motivated by the fact that in many statistical mechanics problems where the variables do lie on a lattice such as $\mathbb{Z}^d$, for d sufficiently large (but finite), the effect of the underlying geometry vanishes.

Perhaps more surprising is the fact that in the last few years, motivated by ideas developed for the study of materials, physicists have put forward a hypothesis for the origin of the aforementioned algorithmic gap in random CSP and, most remarkably, a method for overcoming it. Specifically, Mézard, Parisi, and Zecchina [22] developed an extremely efficient algorithm, called Survey Propagation (SP), for finding satisfying assignments of random formulas in the satisfiable regime. For example, their algorithm typically finds a satisfying truth assignment of a random 3-CNF formula with $n = 10^6$ variables and $4.25n$ clauses in minutes (and appears to scale as $O(n \log n)$). No other algorithm practically solves formulas of such density with $n = 10^4$.

Our original motivation for this work was to see if some (any!) of the physically-motivated ideas underlying SP can be made mathematically rigorous. More generally, we believe that understanding the geometry of the space of satisfying truth assignments in random formulas is essential for understanding the behavior of algorithms on them. This is particularly true for the case of random-walk type algorithms, which we view as the first natural class to target armed with such an understanding and for which very little is known rigorously, with the notable exception of [7].

We make significant progress in this goal by proving that already much below the satisfiability threshold, the set of satisfying assignments fragments into exponentially many connected components. These components are relatively small in size, far apart from one another, and inside each one the majority of variables are "frozen", i.e., take only one value. As the formula density is then increased towards the threshold, the fraction of frozen variables in each component tends to 1, causing the connected components to in volume and grow further apart from one another.

Our results are in perfect agreement with the picture hypothesized by the physicists. Moreover, as we discuss below, the existence of frozen variables provides a good explanation for the origin of the barrier faced by all analyzed algorithms on random CSP, i.e., "local", DPLL-like algorithms. Finally, we show that one of the two main assumptions underlying SP regarding the structure of the set of solutions is essentially correct. This brings us closer to a rigorous analysis of SP and answers affirmatively the main open question raised by Maneva, Mossel and Wainwright in [19].

Specifically, we prove that for all $k \geq 9$, the connected components of the set of satisfying assignments of random formulas have non-trivial cores, as assumed by SP (see Definition 9). We point out that it is not clear whether this is true for small k. Indeed, [19] gave experimental evidence that for $k = 3$, random formulas do *not* have non-trivial cores. As we will see, our methods also give evidence in that direction, suggesting that the requirement $k > 3$, is not an artifact of our analysis. This gives additional motivation for the "core-like" objects introduced in [19] whose existence would relate to the success of SP for small k (we discuss this point further in Section 4.1).

In the next section we give an informal discussion relating the performance of DPLL-like algorithms on random formulas to notions such as Gibbs sampling and long-range correlations. This is meant to provide intuition for the empirical success of SP and motivate our results. We emphasize that while both the discussion and the results are about random k-SAT, this is not strictly necessary: our ideas and proofs are quite generic, and should generalize readily to many other random CSP, e.g., graph coloring.

1.1 DPLL algorithms, Belief Propagation, and Frozen Variables

Given a satisfiable formula F on variables $v_1, v_2, \ldots, v_n$ it is easy to see that the following simple procedure samples uniformly from the set of all satisfying assignments of F:

```
Start with the given formula $F$

For $i = 1$ to $n$ do:

    1. Compute the fraction, $p_i$, of satisfying assignments
       of the current formula in which $v_i$ takes the value 1.

    2. Set $v_i$ to 1 with probability $p_i$ and to 0 otherwise.

    3. Simplify the formula.
```

Clearly, the first step in the loop above is meant only as a thought experiment. Nevertheless, it is worth making the following two observations. The first is that if we are only interested in finding *some* satisfying assignment, as opposed to sampling a uniformly random one, then we do not need to compute exact marginals. For example, if we always set v_i to 1 iff $p_i \geq 1/2$, then we simply need to ensure that if a variable takes the same value x in *all* satisfying assignments, then x should be the majority value in its computed marginal. The second observation is that the order in which we set the variables does not need to be determined a priori. That is, we can imagine that in each step we compute marginals for all remaining variables and that for each marginal we have an associated confidence. To improve our chances of avoiding a fatal error, we can then set only the variable for which we have highest confidence.

The above two elementary observations actually capture all algorithms that have been analyzed so far on random formulas (and, in fact, most DPLL-type algorithms used in practice). Observe, for example, that both the *unit-clause* and the *pure literal* heuristics follow immediately from the above considerations. In the case of unit-clause, the participation of a variable v in a unit clause c allows us to infer its marginal with perfect confidence and thus setting v is an "obvious" choice. In the case of a pure literal ℓ, again we can infer with certainty the majority marginal of the underlying variable v (it is the value that satisfies ℓ). In the absence of such obvious choices, all DPLL-type algorithms attempt to identify a variable whose marginal can be determined with some confidence. For example, below are the choices made in the absence of unit clauses and pure literals by some of the algorithms that have been analyzed on random 3-CNF formulas. In order of increasing performance:

> UNIT-CLAUSE [10]: select a random variable and assign it a random value.
>
> 3-CLAUSE MAJORITY [9]: select a random variable and assign it its majority value among the 3-clauses.
>
> SHORT-CLAUSE[14]: select a random shortest clause c, a random variable v in c, and set v so as to satisfy c.
>
> HAPPIEST LITERAL [16]: select a variable whose majority literal appears in most clauses.

Each of the above heuristics attempts to compute marginals based on a different set of evidence, the content of which ranges from completely empty [10], to considering all the clauses containing each variable [16]. Correspondingly, the largest density for which these algorithms succeed on random 3-CNF formulas ranges from $8/3$ for [10] to 3.42 for [16]. UNIT-CLAUSE, in fact, succeeds for every k as long as $r < 2^k/k$ and, as we mentioned earlier, no algorithm is known to beat this bound asymptotically. Given that improving upon the empty set of evidence is rather easy, it is tempting to think that by considering a larger set of evidence for each variable one can do significantly better. For example, consider an algorithm $\mathcal{A}_d$ which computes a marginal for each variable v based on the clauses that appear in the depth-d neighborhood of v in the factor graph. One could hope that as d grows, such an algorithm would do very well, perhaps even reach the satisfiability threshold.

Physicists say it is not so. The hope that local algorithms could do well on random formulas rests on the presumption that the influence exerted on a variable v by other variables, diminishes rapidly with their distance from v in the factor graph. This corresponds to there being "no long range correlations" in random formulas, e.g., the joint probability distribution of a random finite subset of the variables should be, essentially, the product of their marginals.

Unfortunately, the existence of numerous, far-apart, connected components of satisfying assignments can induce long-range correlations among the variables, eliminating such hopes. The reason is that if one considers the variables in the boundary of a particular tree, then at sufficiently large densities their joint behavior is dominated by only a small number of connected components of satisfying assignments: the ones in which most of these variables are frozen, freeing up other variables to take multiple values and amplify the contribution of that particular joint collection of values to the variable marginals. In other words, in such a setting, assuming that the variables in the boundary of a variable's neighborhood behave independently with respect to the rest of the formula, can be very far off from the truth.

To overcome the above issue, physicists hypothesized that the above clustering is the only significant source of long-range correlations. (Very) roughly speaking, this amounts to modeling each connected component of satisfying assignments as a subcube that results by selecting a large fraction of the variables and freezing them independently at random, while leaving the rest (largely) free. Our results imply that this simplified view of clusters is not very far off the truth.

1.2 Organization

In the next section we give mathematical statements of our main results, both for the existence of exponentially many well-separated clusters and for the existence of frozen variables in each cluster. In Section 3 we outline the proof of the results on the existence of clusters and explain their relationship to the work of Mora, Mézard, and Zecchina [20, 21]. In Section 4 we provide some background on Survey Propagation and explain how our results on frozen variables relate to the implicit hypothesis made by Mézard, Parisi, and Zecchina in their derivation of Survey Propagation [22], and how our results answer the main open question of Maneva, Mossel and Wainwright [19]. In Section 5 we introduce the probabilistic setup for our analysis and in Section 6 we discuss how it relates to the case $k < 8$ and to the "planted assignment" model. Proving our main result on the existence of frozen variables boils down to a question in large deviations and an associated multi-dimensional optimization problem. While we cannot include the solution to this optimization problem here due to lack of space, in Section 7 we show how we arrive at it and give an overview of how we solve it.

2. STATEMENT OF RESULTS

We first need to introduce some definitions. Throughout, we assume that we are dealing with a CNF formula F, defined over variables $X = x_1, \ldots, x_n$, and we let $\mathcal{S}(F) \subseteq \{0,1\}^n$ denote the satisfying assignments of F.

DEFINITION 1. *Given two arbitrary sets $X, Y \subseteq \{0,1\}^n$, their distance is the minimum Hamming distance between any $x \in X$ and $y \in Y$. We will say that $x, y \in \{0,1\}^n$ are adjacent if they have Hamming distance 1.*

- *The **clusters** of a formula F are the connected components of $\mathcal{S}(F)$.*
- *The **diameter** of a cluster C is the maximum Hamming distance between any pair $\sigma_1, \sigma_2 \in C$.*

Our first set of results is captured by the following:

THEOREM 2. *For every $k \geq 8$, there exist $a_k < b_k < 1/2$, and $r < r_k$ such that w.h.p. in $F_k(n, rn)$ all of the following hold:*

1. *Every cluster has diameter at most $a_k n$,*
2. *Every pair of satisfying assignments has distance either less than $a_k n$ or more than $b_k n$, and*
3. *There are $(1+\zeta)^n$ clusters that have distance at least $b_k n$ from one another, for some $\zeta > 0$ independent of k.*

Theorem 2 implies that for all $k \geq 8$, at some point below the satisfiability threshold, the set of satisfying assignments consists of exponentially many "cluster-regions" (unions of clusters), such that each region has small diameter (at most $a_k n$), and is well-separated from all other cluster-regions (by a distance of at least $b_k n$). To see this start with any cluster C and consider the set $R(C) \subseteq \{0,1\}^n$ of truth assignments that have distance at most $a_k n$ from C and the set $B(C) \subseteq \{0,1\}^n$ of truth assignments that have distance at most $b_k n$ from $R(C)$. Observe now that the set $B(C) \setminus R(C)$ cannot contain any satisfying truth assignments, as any such assignment would be at distance $a_k n < d < b_k n$ from some assignment in C. Thus, the set of satisfying assignments in $R(C)$ is a union of clusters (cluster-region), all of which have distance at least $b_k n$ from any cluster not in the region.

The picture suggested by Theorem 2 comes in sharper focus for large k. Specifically, Theorem 3 below asserts that for sufficiently large k, sufficiently close to the threshold, the cluster regions become arbitrarily small and maximally far apart.

THEOREM 3. *For every $\delta > 0$, if $r = (1 - \epsilon(\delta))r_k$, then for all $k \geq k_0(\delta)$, Theorem 2 holds with $a_k = 1/k$ and $b_k = 1/2 - \delta$, where $\zeta > 0$ is independent of δ and k.*

Theorems 2 and 3 follow from considerations similar to those first developed in [20, 21]. Indeed, part 2 of our Theorem 2 was already proven in [20], while part 1 follows easily from part 2. The new contribution is part 3, i.e., establishing the existence of exponentially many clusters (as opposed to "at least two"). We only state all three parts here for ease of reference and because our derivation makes them fit well together.

Our main results in this paper come from taking an approach that allows us to "look inside" clusters. This allows to prove, for example, that significantly below the satisfiability threshold the majority of variables in every cluster are frozen. We note that [20, 21] do not give any results on frozen variables. In proving this result we also prove that random formulas have non-trivial cores for all $k \geq 9$, thus also answering the main question of [19] (we postpone the definition of cores until Section 4).

DEFINITION 4. *The* ***projection*** *of a variable x_i in a set of satisfying assignments C, denoted as $\pi_i(C)$, is the union of the values taken by x_i over the assignments in C. If $\pi_i(C) \neq \{0, 1\}$ we say that x_i is* ***frozen*** *in C.*

A strength of our approach is that it allows to prove not just the existence, but the pervasiveness of frozen variables. Specifically, Theorem 5 below asserts that for sufficiently large k, as we approach the satisfiability threshold, the fraction of frozen variables in *every* cluster gets arbitrarily close to 1.

THEOREM 5 (**Main Result**). *For every $\alpha > 0$ and all $k \geq k_0(\alpha)$, there exists $c_k^\alpha < r_k$, such that for all $r \geq c_k^\alpha$, w.h.p.* **every** *cluster of $F_k(n, rn)$ has at least $(1 - \alpha)n$ frozen variables.*

As k grows,

$$\frac{c_k^\alpha}{2^k \ln 2} \rightarrow \frac{1}{1 + \alpha(1 - \alpha)} \quad .$$

By taking $\alpha = 1/2$ in Theorem 5 we see that for sufficiently large k, every single cluster already has a majority of frozen variables at $r = (4/5 + \delta_k)2^k \ln 2$, where $\delta_k \rightarrow 0$, i.e., for a constant fraction of the satisfiable regime. More generally, Theorem 5 asserts that as k grows and the density approaches the threshold, clusters shrink in volume and grow further apart by having smaller and smaller internal entropy.

The analysis that establishes Theorem 5 also allows us to show

COROLLARY 6. *For every $k \geq 9$, there exists $r < r_k$ such that w.h.p.* **every** *cluster of $F_k(n, rn)$ has frozen variables.*

It remains open whether frozen variables exist for $k \leq 8$. As we mentioned above, [19] reported experimental evidence suggesting that frozen variables do *not* exist for $k = 3$. We will see that our proof also gives evidence in this direction for small values of k.

3. CLUSTERING: PROOF SKETCH AND RELATED WORK

There are two main ingredients for proving Theorems 2 and 3. The first excludes the possibility of pairs of truth assignments at certain Hamming distances. Specifically, it is easy to show that the expected number of pairs of satisfying assignments in $F_k(n, rn)$ with Hamming distance z is at most $\Lambda(z/n, k, r)^n$, where

$$\Lambda(\alpha, k, r) = \frac{2(1 - 2^{1-k} + 2^{-k}(1 - \alpha)^k)^r}{\alpha^\alpha (1 - \alpha)^{1-\alpha}} \quad .$$

Therefore, if for some k, r and $z = \alpha n$ we have $\Lambda(\alpha, k, r) < 1$, it immediately follows by the union bound that w.h.p. in $F_k(n, rn)$ there are no pairs of satisfying assignments that have distance z. In Figure 1 we draw the function Λ (upper curve), and a related function Λ_b (lower curve, to be discussed shortly), for $\alpha \in [0, 3/4]$ with $k = 8$ and $r = 169$.

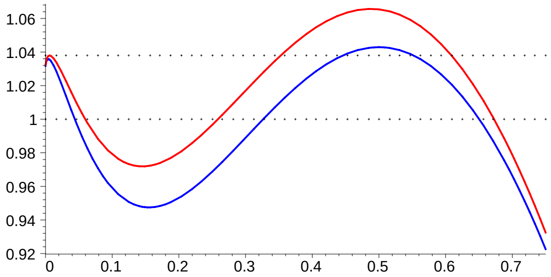

Figure 1: $\Lambda(\alpha, 8, 169)$ **and** $\Lambda_b(\alpha, 8, 169)$ **for** $\alpha \in [0, 3/4]$**.**

We see that $\Lambda(\alpha, 8, 169) < 1$ for $\alpha \in [0.06, 0.26] \cup [0.68, 1]$, implying that w.h.p. there is no pair of satisfying assignments with Hamming distance αn in $F_8(n, 169n)$ for such α. Note that by the results in [5], we know that $F_8(n, 169n)$ is w.h.p. satisfiable and, thus, excluding the possibility of satisfying pairs at certain distances is a non-vacuous statement.

Part 2 of Theorems 2 and 3 is established by applying exactly the above line of reasoning for an appropriate choice of $r < r_k$ and estimating the corresponding location of the two smallest roots a_k, b_k of the equation $\Lambda(\alpha, k, r) = 1$. To prove Part 1 of both theorems, we additionally note that if a cluster C has diameter d, then it must contain pairs of solutions at every distance $1 \leq t \leq d$. To see this, take any pair $\sigma_1, \sigma_2 \in C$ that have distance d, any path from σ_1 to σ_2, and observe that the sequence of distances from σ_1 along the vertices of the path must contain every integer in $\{1, \ldots, d\}$. Therefore, if a_k is the smallest root of $\Lambda(\alpha, k, r) = 1$, w.h.p. every cluster in $F_k(n, rn)$ has diameter at most $a_k n$.

Proving Part 3 of Theorems 2 and 3 requires greater sophistication and leverages in a strong way the results of [5]. This is because having $\Lambda(\alpha, k, r) > 1$ for some α, k, r does *not* imply that pairs of satisfying assignments exist for such α, k, r: in principle, the behavior of Λ could be determined by a tiny minority of solution-rich formulas. Hence the need for the second moment method [2, 5]. Specifically, say that a satisfying assignment is *balanced* if its number of satisfied literal occurrences is in the range $km/2 \pm \sqrt{n}$, and let X be the number of balanced assignments in $F_k(n, rn)$. In [5], it was shown that $\mathbb{E}[X]^2 = \Lambda_b(1/2, k, r)^n$ and

$$\mathbb{E}[X^2] < D \times \max_{\alpha \in [0,1]} \Lambda_b(\alpha, k, r)^n \quad ,$$

for some explicit function Λ_b and constant $D = D(k) > 0$. It was also shown that for all $r < 2^k \ln 2 - k$, the maximum of Λ_b occurs at $\alpha = 1/2$, implying that for such k, r we have $\mathbb{E}[X^2] < D \times \mathbb{E}[X]^2$. By the Payley-Zigmund inequality, this last fact implies that for any $t \leq \mathbb{E}[X]$,

$$\Pr[X > t] \geq \frac{(\mathbb{E}[X] - t)^2}{\mathbb{E}[X^2]} \ . \tag{2}$$

In [5], inequality (2) was applied with $t = 0$, per the "second moment method", establishing that $F_k(n, rn)$ has at least one (balanced) satisfying assignment for all $r < 2^k \ln 2 - k$. If instead we take $t = \mathbb{E}[X]/\mathrm{poly}(n)$, we get that, in fact, with constant probability, X is within a polynomial factor of its expectation $\Lambda_b(1/2, k, r)^{n/2}$. Since the property "has more than q satisfying assignments" has a sharp threshold [13], our last assertion implies that, in fact, for every $r < 2^k \ln 2 - k$, $F_k(n, rn)$ has at least $\Lambda_b(1/2, k, r)^{n/2}/\mathrm{poly}(n)$ satisfying assignments w.h.p.

To prove that there are exponentially many clusters, we divide the above lower bound for the total number of satisfying assignments with the following upper bound for the number of truth assignments in each cluster. Recall that a_k is the smallest root of $\Lambda(\alpha, k, r) = 1$ and let

$$h(k, r) = \max_{\alpha \in [0, a_k]} \Lambda(\alpha, k, r) \ .$$

If B is the expected number of pairs of truth assignments with distance at most $a_k n$ in $F_k(n, rn)$, it follows that $B < \mathrm{poly}(n) \times h(k, r)^n$, since the expected number of pairs at each distance is at most $\Lambda(\alpha, k, r)^n$ and there are no more than $n + 1$ possible distances. By Markov's inequality, this implies that w.h.p. the number of pairs of truth assignments in $F_k(n, rn)$ that have distance at most $a_k n$ is $\mathrm{poly}(n) \times h(k, r)^n$. Recall now that w.h.p. every cluster in $F_k(n, rn)$ has diameter at most $a_k n$. Therefore, w.h.p. the total number of pairs of truth assignments in each cluster is at most $\mathrm{poly}(n) \times h(k, r)^n$. Thus, if $h(k, r) < \Lambda_b(1/2, k, r)$, we can conclude that $F_k(n, rn)$ has at least

$$1/\mathrm{poly}(n) \times \left(\frac{\Lambda_b(1/2, k, r)}{h(k, r)} \right)^{n/2}$$

clusters. Indeed, the higher of the two horizontal lines in Figure 1 highlights that $h(8, 169) < \Lambda_b(1/2, 8, 169)$.

More formally, to establish Part 3 of Theorem 2 for each $k \geq 8$, we give $r < r_k$ such that $h(k, r) < \Lambda_b(1/2, k, r)$. In particular, taking $r = 2^{k-1}(1+\delta_k)$ for some suitably chosen $\delta_k \to 0$ suffices. For Part 3 of Theorem 3 we need to show that for every $\epsilon > 0$, if $r = (1 - \epsilon)2^k \ln 2$, then for all $k \geq k_0(\epsilon)$, we have $h(k, r) < \Lambda_b(1/2, k, r)$. Again, this reduces to a (rather tedious) asymptotic analysis which we omit in this extended abstract.

3.1 Related Work

The observation that if $\Lambda(\alpha, k, r) < 1$ then w.h.p. $F_k(n, rn)$ has no pairs of satisfying assignments at distance αn was first made and used in [20]. Moreover, in [21] the authors gave an expression $\Lambda_\ell(\alpha, k, r)$ for the expected number of *locally maximal* pairs of satisfying assignments at each distance, where a pair σ, τ is locally maximal if there is no variable which has value 0 in σ and flipping its value in both σ and τ yields a new pair of satisfying assignments. (If a formula has a pair of satisfying assignments at distance d, then it always has a locally maximal pair at distance d). Clearly, $\Lambda_\ell(\alpha, k, r) < \Lambda(\alpha, k, r)$ always, but for large k and $r = \Theta(2^k)$ the difference is minuscule for all α.

Additionally, in [20, 21] the authors derive an expression for the second moment of the number of *pairs of* balanced assignments at distance αn, for each $\alpha \in [0, 1]$. Whenever, for some α, k, r, the dominant contribution to this second moment comes from uncorrelated pairs of pairs of balanced assignments, this implies that with *constant* probability $F_k(n, rn)$ contains at least one (balanced) pair of assignments at distance αn. We note that determining the dominant contribution to the above second moment rigorously, given α, k, r, is a highly non-trivial problem which the authors tackle numerically for small k, and heuristically for general k, i.e., they simply make a guess for the locus of the maximizer. In particular, this "fourth moment" optimization problem is *much* harder than the already complicated second moment analysis of [5].

Finally, the authors prove that the property "has a pair of satisfying assignments at distance q" has a sharp threshold, thus boosting their constant probability result for having a pair of satisfying assignments at a given distance to a high probability one. To the best of our understanding, these three are the only results discussed in [21]. Combined, they imply that for every $k \geq 8$, there is $r < r_k$ and constants $a_k < b_k < c_k < 1/2 < d_k$, such that in $F_k(n, rn)$:

- W.h.p. every pair of satisfying assignments has distance either less than $a_k n$ or more than $b_k n$.
- For every $d \in [c_k, d_k] \cdot n$, w.h.p. there is a pair of truth assignments that have distance d.

We note that even if the maximizer in the second moment computation was determined rigorously and coincided with the heuristic guess of [21], the strongest statement that can be inferred from the above two assertions in terms of "establishing clustering" is: for every $k \geq 8$, there is $r < r_k$, such that w.h.p. $\mathcal{S}(F_k(n, rn))$ has at least two clusters.

In contrast, our Theorem 2 establishes that w.h.p. $\mathcal{S}(F_k(n, rn))$ consists of *exponentially* many, well-separated cluster regions, each region containing at least one cluster. Additionally, Theorem 3 establishes that as k grows and r approaches the threshold, these regions grow maximally far apart and their diameter vanishes.

4. FROZEN VARIABLES: SURVEY PROPAGATION AND RELATED WORK

For a cluster C, the string $\pi(C) = \pi_1(C), \pi_2(C), \ldots, \pi_n(C)$ is the **projection** of C and we will use the convention $\{0, 1\} \equiv *$, so that $\pi(C) \in \{0, 1, *\}^n$. Imagine for a moment that given a formula F we could compute the marginals of variables over the cluster projections, i.e., that for each variable we could compute the fraction of clusters in which the projection is $0, 1$, and $*$. Then, by repeatedly setting variables sensibly, i.e., never setting to $1 - x$ a variable whose cluster projection marginal puts all the mass on x, we are guaranteed to find a satisfying assignment: after each step there is at least one cluster consistent with our choices so far.

Clearly, being able to perform such marginalization seems quite far fetched given that even if we are handed a truth assignment σ in a cluster C, it is not at all clear how to compute $\pi(C)$ in time less than $|C|$. Survey Propagation (SP) is an attempt to compute marginals over cluster projections by making a number of approximations. One fundamental assumption underlying SP is that, unlike the marginals over truth assignments, the marginals over cluster projections essentially factorize, i.e., if two variables are far apart in the formula, then their joint distribution over cluster projections is essentially the product of their cluster projection marginals. Determining the validity of this assumption remains an outstanding open problem.

The other fundamental assumption underlying SP is that *approximate* cluster projections can be encoded as the solutions of a CSP whose factor graph can be readily derived from the input formula. Our results are closely related to this second assumption and establish that, indeed, the approximate cluster projections used in SP retain a significant amount of information from the cluster projections. To make this last notion concrete and enhance intuition, we give below a self-contained, but brisk discussion of Survey Propagation. For the sake of presentation this discussion is historically inaccurate. We attempt to restore history in Section 4.1.

As we said above, even if we are given a satisfying assignment σ, it is not obvious how to determine the projection of its cluster $C(\sigma)$. To get around this problem SP sacrifices information in the following manner.

DEFINITION 7. *Given a string $x \in \{0,1,*\}^n$, we will say that variable x_i is **free** in x if in every clause c containing x_i or $\overline{x}_i$, at least one of the other literals in c is assigned true or $*$.*
We will refer to the following as a

***coarsening-step:** if a variable is free, assign it $*$.*

Given $x, y \in \{0,1,*\}^n$ say that x is dominated by y, written $x \preceq y$, if for every i, either $x_i = y_i$ or $y_i = *$.

Consider now the following process:

start at σ and apply coarsening until a fixed point is reached.

LEMMA 8. *For every formula F and truth assignment $\sigma \in \mathcal{S}(F)$, every sequence of coarsening steps that reaches a fixed point reaches the same fixed point $w(\sigma)$. If σ_1, σ_2 belong to the same cluster C, then $w(\sigma_1) = w(\sigma_2) \succeq \pi(C)$.*

PROOF. Trivially, applying a coarsening step to a string x produces a string y such that $x \preceq y$. Moreover, if x_i was free in x, then y_i will be free in y. As a result, if both $y, z \in \{0,1,*\}^n$ are reachable from $x \in \{0,1,*\}^n$ by coarsening steps, so is the string that results by starting at x, concatenating the two sequences of operations and removing all but the first occurrence of each coarsening step. This implies that there is a unique fixed point $w(x)$ for each $x \in \{0,1,*\}^n$ under coarsening. Observe now that if $\sigma, \sigma' \in \mathcal{S}(F)$ differ only in the i-th coordinate, then the i-th variable is free in both σ, σ' and coarsening it in both yields the same string τ. By our earlier argument, $w(\sigma) = w(\tau) = w(\sigma') = w_C$, where $C \subseteq \mathcal{S}(F)$ is the cluster containing σ, σ'. Considering all adjacent pairs in C, we see that $w_C \succeq \pi(C)$. □

DEFINITION 9. *The **core** of a cluster C is the unique coarsening fixed point of the truth assignments in C.*

By Lemma 8, if a variable takes either the value 0 or the value 1 in the core of a cluster C, then it is frozen to that value in C. To prove Theorem 5 we prove that the core of every cluster has many non-$*$ variables.

THEOREM 10. *For any $\alpha > 0$, let $k_0(\alpha)$ and c_k^α be as in Theorem 5. If $k \geq k_0$ and $r \geq c_k^\alpha$, then w.h.p. in the coarsening fixed point of* **every** *$\sigma \in \mathcal{S}(F_k(n, rn))$ fewer than $\alpha \cdot n$ variables take the value $*$.*

To prove Theorem 10 (which implies Theorem 5) we derive sharp bounds for the large deviations rate function of the coarsening process applied to a fixed satisfying assignment. As a result, we also prove that in the planted-assignment model the cluster containing the planted assignment already contains frozen variables at $r \sim (2^k/k)\ln k$. Also, we will see that our proof gives a strong hint that for small values of k, such as $k = 3$, and for all densities in the corresponding satisfiable regime, most satisfying assignments **do** converge to $(*, \cdots, *)$ upon coarsening.

We can think of coarsening as an attempt to estimate the projection of $C(\sigma)$ by starting at σ and being somewhat reckless. To see this, consider a parallel version of coarsening in which given $x \in \{0,1,*\}^n$ we coarsen all free variables in it simultaneously. While the first round of such a process will only assign $*$ to variables whose projection in $C(\sigma)$ is indeed $*$, subsequent rounds might not: a variable v is deemed free, if in every clause containing it there is some other variable satisfying the clause, *or* a variable assigned $*$. This second possibility is equivalent to assuming that the $*$-variables in the clauses containing v, call them Γ_v, can take joint values that allow v to not contribute in the satisfaction of any clause. In general formulas this is, of course, not a valid assumption. On the other hand, the belief that in random formulas there are no long-range correlations *among the non-frozen* variables of each cluster makes this is a reasonable statistical assumption: since the formula is random, the variables in Γ_v are probably far apart from one another in the factor graph that results after removing the clauses containing v. Thus, indeed, any subset of variables of Γ_v that do not co-occur in a clause should be able to take *any* set of joint values. Our results can be seen as evidence of the utility of this line of reasoning, since we prove that for sufficiently large densities, the coarsening fixed point of a satisfying assignment is *never* $(*, \ldots, *)$. Indeed, as we approach the satisfiability threshold, the fraction of frozen variables in it tends to 1.

Of course, while the core of a cluster C can be easily derived given some $\sigma \in C$, such a σ is still hard to come by. The last leap of approximation underlying SP is to define a set $Z(F) \subseteq \{0,1,*\}^n$ that includes all cluster cores, yet is such that membership in $Z(F)$ is "locally checkable", akin to membership in $\mathcal{S}(F)$. Specifically,

DEFINITION 11. *A string $x \in \{0,1,*\}^n$ is a **cover** of a CNF formula F if: (i) under x, every clause in F contains a satisfied literal or at least two $*$, and (ii) every free variable in x is assigned $*$, i.e., x is $*$–maximal.*

Cores trivially satisfy (ii) as fixed points of coarsening; it is also easy to see, by induction, that any string that results by applying coarsening steps to a satisfying assignment satisfies (i). Thus, a core is always a cover. On the other hand, checking whether $x \in \{0,1,*\}^n$ satisfies (i) can be done trivially by examining each clause in isolation. For (ii) it is enough to check that for each variable v assigned 0 or 1 in x, there is at least one clause satisfied by v and dissatisfied by all other variables in it. Again, this amounts to n simple checks, each check done in isolation by considering the clauses containing the corresponding variable. The price we pay for dealing with locally-checkable objects is that the set of all covers $Z(F)$ can be potentially much bigger than the set of all cores, e.g., $(*, \cdots, *)$ is always a cover, even if F is unsatisfiable.

The Survey Propagation algorithm can now be stated as follows.

Repeat until all variables are set:

1. Compute the marginals of variables over covers.
2. Select a variable with least mass on $*$ and assign it the 0/1 value on which it puts most mass.
3. Simplify the formula.

The computation of marginals over covers in the original derivation [22] of SP was, in fact, done via a message passing procedure that runs on the factor graph of the original formula (more on this in Section 4.1). Also, in [22], if a configuration is reached in which all variables put (nearly) all their mass on $*$, the loop is stopped and a local search algorithm is invoked. The idea is that when such a configuration is reached, the algorithm has "arrived" at a cluster and finding a solution inside that cluster is easy since only non-frozen variables remain unset.

4.1 Related Work

The original presentation of Survey Propagation motivated the algorithm in terms of a number of physical notions (cavities, magnetic fields, etc.). Specifically, the algorithm was derived by applying the "cavity method" within a "1-step Replica Symmetry Breaking" scheme, with no reference whatsoever to notions such as cluster projections, cores, or covers (in fact, even clusters were only specified as the connected components that result when satisfying assignments at "finite Hamming distance" are considered adjacent). On the other hand, a very definitive message-passing procedure was specified on the factor graph of the original formula and the computer code accompanying the paper and implementing that procedure worked spectacularly well. Moreover, a notion foreshadowing cores was included in the authors' discussion of "Warning Propagation".

Casting SP as an attempt to compute marginals over cores was done independently by Braunstein and Zecchina in [8] and Maneva, Mossel, and Wainwright in [19]. In particular, in both papers it is shown that the messages exchanged by SP over the factor graph of the input formula are the messages implied by the Belief Propagation formalism [6] applied to a factor graph encoding the set of all covers. The first author and Thorpe [1] have additionally shown that for every formula F, there is a factor graph G_F encoding the set of F's covers which inherits the cycle structure of F's factor graph, so that if the latter is locally tree-like so is G_F.

In [19], the authors give a number of formal correspondences between SP, Markov random fields and Gibbs sampling and note that a cover $\sigma \in \{0,1,*\}^n$ can also be thought of as partial truth assignment in which every unsatisfied clause has length at least 2, and in which every variable v assigned 0 or 1 has some clause c for which it is essential in σ, i.e., v satisfies c but all other variables in c are set opposite to their sign in c. This last view motivates a generalization of SP in which marginals are computed not only over covers, but over all partial assignments in which every unsatisfied clause has length at least 2, weighted exponentially in the number of non-essential 0/1 variables and the number of $*$-variables. One particular motivation for this generalization is that while SP appears to work very well on random 3-CNF, [19] gives experimental evidence that such formulas do not have non-trivial cores, i.e., upon coarsening truth assignments end up as $(*,\ldots,*)$. This apparent contradiction is reconciled by attributing the success of SP to the existence of "near-core" strings allowed under the proposed generalization.

While [19] provided a framework for studying SP by connecting it to concrete mathematical objects such as cores and Markov random fields, it did not provide results on the actual structure of the solution space of random k-CNF formulas. Indeed, motivated by the experimental absence of cores for $k = 3$, the authors asked whether random formulas have non-trivial cores for any k. Our results, establish a positive answer to this question for all $k \geq 9$.

5. THE PROBABILISTIC FRAMEWORK

Theorem 5 follows from Theorem 10 and Lemma 8. To prove Theorem 10 we say that a satisfying assignment σ is α-coreless if its coarsening fixed point $w(\sigma)$ has at least αn $*$-variables. Let X be the random variable equal to the number of α-coreless satisfying assignments in a random k-CNF formula $F_k(n, rn)$. By symmetry,

$$\mathbb{E}[X] = 2^n \cdot \left(1 - \frac{1}{2^k}\right)^{rn} \cdot \Pr[\mathbf{0} \text{ is } \alpha\text{-coreless} \mid \mathbf{0} \text{ is satisfying}] \ . \tag{3}$$

Observe that conditioning on "$\mathbf{0}$ is satisfying" is exactly the same as "planting" the $\mathbf{0}$ solution, and amounts to selecting the $m = rn$ random clauses in our formula, uniformly and independently from amongst all clauses having at least one negative literal. We will see that for every $k \geq 3$, there exists t_k^α such that

$$\Pr[\mathbf{0} \text{ is } \alpha\text{-coreless} \mid \mathbf{0} \text{ is satisfying}] = \begin{cases} 1 - o(1) & \text{if } r < t_k^\alpha \ , \\ \\ o(1) & \text{if } r > t_k^\alpha \ . \end{cases} \tag{4}$$

We will see that $t_k^1 \sim (2^k/k)\ln k$, i.e., for densities up to that point, w.h.p. a fixed satisfying assignment does not have any frozen variables. We find it interesting (and speculate that it's not an accident) that all analyzed algorithms work up to densities just below t_k^1, i.e., up to some $r < c\,2^k/k$, where only the constant c increases with the sophistication of the algorithm.

To prove $\mathbb{E}[X] = o(1)$ we will derive a strong upper bound for the probability in (4) when $r > t_k^\alpha$. Specifically, we will prove that $\Pr[\mathbf{0} \text{ is } \alpha\text{-coreless} \mid \mathbf{0} \text{ is satisfying}] < e^{-f(r)n}$ for a function f such that for all $r \geq c_k^\alpha$,

$$2 \cdot \left(1 - \frac{1}{2^k}\right)^r \cdot e^{-f(r)} < 1 \ . \tag{5}$$

Thus, by (3), for all such r we have $\mathbb{E}[X] = o(1)$ and Theorem 10 follows.

5.1 Coarsening as Hypergraph Stripping

Given any CNF formula F and any $\sigma \in \mathcal{S}(F)$ it is easy to see that $w(\sigma)$ is completely determined by the set of clauses $U(\sigma)$ that have precisely one satisfied literal under σ. This is because after any sequence of coarsening steps applied to σ, a clause that had two or more satisfied literals under σ, will have at least one satisfied literal or at least two $*$ and thus never prevent a variable from being free. Therefore, to coarsen a truth assignment σ it is enough to consider the clauses in $U(\sigma)$. Let us say that a variable v is unfrozen if there is no clause in which it is the unique satisfying variable and let us say that a clause is unfrozen if it contains an unfrozen variable. It is now easy to see that coarsening σ is equivalent to starting with U and removing unfrozen clauses, one by one, in an arbitrary order until a fixed point is reached, i.e., no unfrozen clauses remain. Variables occurring in any remaining (frozen) clauses are, thus, frozen in $w(\sigma)$ (to their value in σ), while all other variables are assigned $*$. This view of coarsening as repeated removal of clauses from $U(\sigma)$ will be very useful in our probabilistic analysis below.

To estimate $\Pr[\mathbf{0} \text{ is } \alpha\text{-coreless} \mid \mathbf{0} \text{ is satisfying}]$ we consider a random k-CNF formula with rn clauses chosen uniformly among those satisfying $\mathbf{0}$. To determine $w(\mathbf{0})$, by our discussion above, it suffices to consider the clauses in our formula that have precisely one satisfied (negative) literal. The number of such clauses is distributed as

$$m = \operatorname{Bin}\left(rn, \frac{k}{2^k - 1}\right) \ .$$

It will be convenient to work in a model where each of these m clauses is formed by choosing 1 negative literal and $k-1$ positive literals, uniformly, independently *and with replacement*. (Since $m = O(n)$, by standard arguments, our results then apply when replacement is not allowed and the original number of clauses is $rn - o(n)$.) We think of the k literals in each clause as k balls; we paint the single satisfied literal of each clause red, and the $k-1$ unsatisfied literals blue. We also have one bin for each of the n variables and we place each literal in the bin of its underlying variable. We will use the term "blue bin" to refer to a bin that has at least one blue ball and no red balls. With this picture in mind, we see that the $*$-variables in $w(\mathbf{0})$ correspond precisely to the set of empty bins when the following process terminates:

1. Let v be any blue bin; if none exists exit.
   %Identify an unfrozen variable v if one exists.
2. Remove any ball from v.
   %Remove the occurrence of v in some (unfrozen) clause c.
3. Remove $k-2$ random blue balls.
   %Remove the other $k-2$ unsatisfied literals of c.
4. Remove a random red ball.
   %Remove the satisfied literal in c.

Note that the above process removes exactly one clause (1 red ball and $k-1$ blue balls) in each step and, therefore, if we pass the condition in Step 1, there are always suitable balls to remove. To give a lower bound on the probability that the process exits before m steps (thus, reaching a non-trivial fixed point), we will give a lower bound on the probability that it exits within the first $i = \alpha m$ steps, for some carefully chosen $\alpha = \alpha(k, r) \in (0, 1)$. In particular, observe that if the process does not exit within the first i steps this means that:

$$\text{At the beginning of each of the first } i \text{ steps}$$
$$\text{there is at least one blue bin.} \tag{6}$$

In order to bound the probability of the event in (6) it will be convenient to bound the probability of the same event occurring in the following simplified process. The point is that this modified process is significantly easier to analyze, while the event in (6) is only slightly more likely for the values of k, r of interest to us.

(a) Let v be any blue bin; if none exists go to Step (c).

(b) Remove any ball from v.

(c) Remove a random red ball.

LEMMA 12. *The event in* (6) *is no less likely in the modified process than in the original process.*

We prove Lemma 12 below. To bound the probability of the event in (6) in the modified process we argue as follows. Let q be the number of bins which do not contain any red ball after i steps and let b be the original number of blue balls in these q bins. If $b < i$, then after b steps of the modified process every non-empty bin will contain at least one red ball, since up to that point we remove precisely one blue ball per step. Therefore, the probability of the event in (6) is bounded above by the probability that $b \geq i$. To bound this last probability we observe that the red balls in the modified process evolve completely independent of the blue balls. Moreover, since we remove exactly one red ball in each step, the state of the red balls after i steps is distributed exactly as if we had simply thrown $m - i$ red balls into the n bins.

So, all in all, given a random k-CNF formula F with rn clauses, conditional on **0** satisfying F and for a fixed $\alpha \in (0, 1)$, the probability that the coarsening process started at **0** fails to reach a fixed point within i steps is bounded by the probability that $b \geq i$, where

$$b = \mathrm{Bin}\left((k-1)m, \frac{q}{n}\right) \quad , \quad \text{where} \tag{7}$$
$$m = \mathrm{Bin}\left(rn, \frac{k}{2^k - 1}\right) \quad , \quad \text{and} \tag{8}$$
$$q = \mathrm{Occ}\,(m - i, n) \quad , \tag{9}$$

where $\mathrm{Occ}(x, y)$ is the distribution of the number of empty bins when we throw x balls into y bins.

As a result, given k, r, our goal is to determine a value for i that minimizes $\Pr[b \geq i]$. Before we delve into the probabilistic calculations, in the next section we comment on how our analysis relates to the planted assignment problem and to the existence of non-trivial cores for small values of k.

PROOF OF LEMMA 12. Consider a process which is identical to the original process except with Step 3 removed. We will call this the intermediate process. We begin by proving that the original and the intermediate processes can be coupled so that whenever the event in (6) occurs in the original process it also occurs in the intermediate process.

First, observe that the evolution of the red balls in both processes is purely random and therefore can be assumed to be identical, i.e., we can think of the original process as making a genuine random choice in Step 4 and the intermediate process as mimicking that choice. (We think of all balls as carrying a distinct identifier.) Similarly, we can assume that originally, the placement of the blue balls in bins is identical for the two processes.

Let us say that a pair of blue ball placements is good if in every bin the set of blue balls in the original process is a subset of the set of blue balls in the intermediate process. Clearly, whenever we are in a good configuration, since the placement of the red balls is identical in the two processes, any choice of bin and ball of the original process in Steps 1,2, respectively is an available choice for the intermediate process. Moreover, if the intermediate process mimics these choices, this results is a new good pair of blue ball placements. Therefore, since the original pair of blue ball placements is good, if the event in (6) occurs in the original process it also occurs in the intermediate process.

Next, we compare the intermediate process to the modified process observing that they are identical except that in the event that we run out of bins containing only blue balls the intermediate process stops, while the modified process carries on. Therefore, we couple the two as follows: the modified process mimics the intermediate process for as long as the event in (6) does not occur, and makes its own random choices afterwards. Therefore, if the event in (6) occurs in the intermediate process it also occurs in the modified process. □

6. THE PLANTED ASSIGNMENT MODEL AND SMALL VALUES OF K

Conditional on **0** being satisfying, analyzing $w(\mathbf{0})$ is exactly the same as working in the "planted assignment model" and analyzing the core of the cluster containing the planted assignment. This is rather easy to do if we are content with results holding with probability $1 - o(1)$. Specifically, by (7),(8),(9) and standard concentration results it follows immediately that if $i = \alpha m$ then w.h.p.

$$m = \lambda \cdot n + o(n), \qquad \text{where } \lambda = \frac{rk}{2^k - 1} \tag{10}$$
$$q = \gamma \cdot n + o(n), \qquad \text{where } \gamma = \exp\left(-\lambda(1-\alpha)\right) \tag{11}$$
$$b = \beta \cdot n + o(n), \qquad \text{where } \beta = (k-1)\gamma\lambda \ . \tag{12}$$

With these conditionals in place, we can in fact go ahead and determine the mean path of the coarsening process using the method of differential equations [23], i.e., the number of red and blue balls after each step, up to $o(n)$. In particular, this allows us to prove that

THEOREM 13. *For every $k \geq 3$, there exists a critical value t_k^1 such that if $r < t_k^1$ then w.h.p. $w(\mathbf{0}) = (*, \ldots, *)$, while if $r > t_k^1$ then w.h.p. a bounded fraction of the variables in $w(\mathbf{0})$, and therefore in $C(\mathbf{0})$, are frozen.*

In the table below we give the value of t_k^1 for some small values of k (rounding to two decimals). By lower/upper below we mean the best known lower/upper bound for satisfiability threshold.

k	3	4	5	6	7
Lower	3.52	7.91	18.79	40.62	84.82
Upper	4.51	10.23	21.33	43.51	87.88
t_k^1	5.72	11.58	21.75	40.13	73.88
u_k	6.25	12.34	22.90	41.95	76.84

We see that for $k = 3, 4, 5$, the probability that **0** has a non-trivial coarsening fixed point conditional on being satisfying, tends to 0 for all densities in the satisfiable regime. Clearly, conditioning on "**0** is satisfying", is not the same as picking a "typical" satisfying assignment. Nevertheless, the gap between t_k^1 and the best threshold upper bound for $k = 3$ is sufficiently large to strongly suggest that below the satisfiability threshold, most satisfying assignments do arrive at $(*, \ldots, *)$ upon coarsening. This is consistent with the experimental results of [19], who first raised this possibility.

We now comment on the couple of simplifications of the original process that we introduced in the previous section in order to get a process that is easier to analyze. As we showed, these simplifications only increase the probability of the event in (6) and it is natural to wonder if this increase is significant, allowing for the possibility that our analysis can be made much sharper. Below we give evidence that this is not the case. In particular, if each of m, q, b is within $o(n)$ of its expected value, then the inequality $b \geq i$ in the modified process is equivalent to

$$r < \frac{2^k - 1}{k} \cdot \frac{\ln\left(\frac{k-1}{\alpha}\right)}{1-\alpha} \equiv u_k(\alpha) \ .$$

In the table above we give the value of $u_k = \min_\alpha u_k(\alpha)$ for some small values of k. As we can see, these values are quite close to t_k^1 and get relatively closer as k is increased. In other words, considering the modified process does not cause too big a loss in the analysis. Indeed, taking e.g. $\alpha = 1/\ln k$, already gives $u_k \to (2^k/k) \ln k$, which is consistent with the physics prediction that $t_k^1 \to (2^k/k) \ln k$.

Of course, if one is interested in establishing that certain properties of $w(\mathbf{0})$ hold with exponentially small failure probability, as we do, then conditioning as above is not an option. One has to do a large deviation analysis of all the underlying variables and their interactions in the coarsening process and determine the dominant source of fluctuations. This is precisely what we do with respect to the event $b \geq i$ in the modified process.

7. LARGE DEVIATIONS

It is well-known that if $np > 0$ then for every $\delta \geq -1$,

$$\Pr[\mathrm{Bin}(n, p) = (1 + \delta)np] \leq F(np, \delta) \ ,$$

where

$$F(x, y) = \exp(-x[(1 + y)\ln(1 + y) - y]) \ .$$

A similar large deviations bound was shown in [15] for the number of empty bins in a balls-in-bins experiment (Theorem 3). That is, for every $\delta \geq -1$,

$$\Pr[\mathrm{Occ}(m, n) = (1 + \delta)e^{-m/n}] \leq F(ne^{-m/n}, \delta) \ .$$

7.1 Application

Write $r = \lambda(2^k - 1)/k$ and fix $\delta, \epsilon, \zeta \geq -1$. Write $\rho = \lambda(1 + \delta)(1 - \alpha)$ in order to compress the expressions below. The probability that

$$\begin{aligned} m &= (1+\delta)\mathbb{E}[m] \\ &= (1+\delta)\frac{rk}{2^k-1} \cdot n \\ &= (1+\delta)\lambda \cdot n \end{aligned} \qquad (13)$$

$$\begin{aligned} q &= (1+\zeta)\mathbb{E}[q|m] \\ &= (1+\zeta)\exp\left(-\frac{m-i}{n}\right) \cdot n \\ &= (1+\zeta)e^{-\rho} \cdot n \end{aligned} \qquad (14)$$

$$\begin{aligned} b &= (1+\epsilon)\mathbb{E}[b|q, m] \\ &= (1+\epsilon)(k-1)m \cdot \frac{q}{n} \\ &= (1+\delta)(1+\epsilon)(1+\zeta)\lambda(k-1)e^{-\rho} \cdot n \ , \end{aligned} \qquad (15)$$

is bounded by

$$F(\lambda n, \delta) \cdot F(e^{-\rho}n, \zeta) \cdot F((1+\delta)(1+\zeta)\lambda(k-1)e^{-\rho}n, \epsilon) \ . \qquad (16)$$

We will write this as $e^{-n\Omega}$, where

$$\Omega \equiv \lambda\omega(\delta) + e^{-\rho}\omega(\zeta) + \lambda(k-1)(1+\delta)(1+\zeta)e^{-\rho}\omega(\epsilon) \ ,$$

with $\omega(x) = (1 + x)\ln(1 + x) - x$.

Conditional on the events in (13)–(15) we see from (15) that the condition $b \geq i$ becomes $B \geq 0$, where

$$B \equiv (1 + \epsilon)(1 + \zeta)(k - 1)e^{-\rho} - \alpha \ .$$

For any fixed k, r and α define $\Phi \equiv \{(\delta, \zeta, \epsilon) : B \geq 0\}$. Thus,

$$\Pr[\mathbf{0} \text{ is } \alpha\text{-coreless} \mid \mathbf{0} \text{ is satisfying}] < \exp(-n \cdot \min_{\Phi} \Omega) \times \mathrm{poly}(n)$$

and to prove that the expected number of α-coreless assignments in $o(1)$, it suffices to prove

$$\min_{\Phi} \Omega > \ln 2 + r\ln(1 - 2^{-k}) \equiv s \ . \qquad (17)$$

7.2 Optimization Overview

To establish (17) it is enough to prove that the maximum of B in the variables δ, ζ and ϵ under the condition $\Omega \leq s$ is negative. Considering that the function B is monotone in the three variables δ, ζ and ϵ, the maximizer of B in the region $\Omega \leq s$ has to be on the boundary, that is for $\Omega = s$. The maximum of B under the condition $\Omega = s$ corresponds to the extremum of the function $G = B - \mu(\Omega - s)$, where μ is a Lagrange multiplier.

From the above we can derive

LEMMA 14. *For any fixed k, r and $\alpha \in (0, 1)$, at the extremum of the function G, the following assertions hold: ϵ is non-negative; ζ is non-negative; δ is non-positive.*

We next prove some bounds on δ and ϵ, at the maximizer.

LEMMA 15. *Fix any $r, k, \alpha \in (0, 1)$. At the maximizer of B,*

$$\begin{aligned} \delta_0 &\equiv -\sqrt{\frac{2s}{\lambda}} \leq \delta \leq 0 \\ \epsilon &< \frac{1-\alpha}{k-1} + \frac{\ln 3}{\lambda(1+\delta_0)(k-1)} \equiv \epsilon_0 \ . \end{aligned}$$

Thus, the stationary point of G must occur in the region $\Lambda = \{(\delta, \epsilon) : \delta_0 \leq \delta \leq 0, 0 \leq \epsilon \leq \epsilon_0\}$. Using these bounds, we derive analytical results for our optimization problem for all $k \geq 14$. For $9 \leq k \leq 13$, the stationary point of G is found numerically.

7.3 Proving the existence of frozen variables for $k \geq 14$ analytically

For any fixed values of δ and ϵ, the requirement $B \geq 0$ implies

$$\zeta \geq \frac{\alpha\, e^{\rho}}{(k-1)(1+\epsilon)} - 1 \ . \tag{18}$$

Plugging this lower bound in the second term of Ω, we see that the requirement $\Omega = s$ implies

$$\frac{\alpha}{(k-1)(1+\epsilon)} \left[\ln\left(\frac{\alpha\, e^{\rho}}{(k-1)(1+\epsilon)}\right) - 1\right] + e^{-\rho} \leq s \ . \tag{19}$$

Therefore, it suffices to find λ and α such that (19) cannot be satisfied by any $\delta_0 \leq \delta \leq 0$ and $0 \leq \epsilon \leq \epsilon_0$. This is certainly true if a lower bound to the l.h.s. of (19) makes such an equation unsatisfied, that is if

$$\frac{\alpha}{(k-1)(1+\epsilon_0)} \left[\ln\left(\frac{\alpha\, e^{\lambda(1+\delta_0)(1-\alpha)}}{(k-1)(1+\epsilon_0)}\right) - 1\right] + e^{-\lambda(1-\alpha)} > s$$

and the term within the squared brackets above is positive. Thus, to summarize, it suffices to find λ and α such that

$$\frac{\alpha\left[\ln\left(\frac{\alpha\, e^{\lambda(1+\delta_0)(1-\alpha)}}{(k-1)(1+\epsilon_0)}\right) - 1\right]}{(k-1)(1+\epsilon_0)} > s - e^{-\lambda(1-\alpha)} > 0 \ .$$

For this we perform the change of variable $\lambda = c\, k \,\ln 2$ and use Lemma 15 to introduce simpler bounds for δ and ϵ, which hold for all $c \geq 4/5$ and $k \geq 2$, namely $|\delta_0| \leq k^{-1/2}$ and $\epsilon_0 \leq 2/(k-1)$. Using these simpler bounds and solving with respect to c, we get

$$c \geq \frac{1 + \frac{\alpha}{k+1}\left[1 - \ln\left(\frac{\alpha}{k+1}\right)\right] / \ln 2}{1 + \alpha(1-\alpha)\frac{1-1/\sqrt{k}}{1+1/k}} \equiv g_c(k, \alpha) \ .$$

For any fixed $\alpha \in (0, 1)$, $g_c(k, \alpha)$ is a decreasing function of k. Moreover, it is easy to see that as $k \to \infty$

$$g_c(k, \alpha) \to \frac{1}{1 + \alpha(1-\alpha)} \ ,$$

implying the desired result for all $k \geq k_0(\alpha)$.

Indeed, this analysis allows us to prove that for all $k \geq 14$, there exists a choice of α such that $\min_{\Phi} \Omega > s$ for some $r < r_k$. Specifically, we first rescale the lower bound for r_k from (1) as

$$\tau_k \equiv \frac{2^k \ln 2 - \frac{(k+1)\ln 2 + 3}{2}}{(2^k - 1)\ln 2} = \frac{1}{1 - 2^{-k}} - \frac{(k+1)\ln 2 + 3}{2(2^k - 1)\ln 2} \ ,$$

and observe that τ_k is increasing in k. It is not hard to check that the function $g_c(14, \alpha)$ dips below τ_{14} for a certain range of α, implying that the left endpoint of the range is an upper bound on the fraction of unfrozen clauses. For larger values of k things only get better since $g_c(k, \alpha)$ is decreasing with k, whereas τ_k is increasing.

Acknowledgments

We are grateful to the anonymous referees for a number of suggestions on the presentation of our results.

8. REFERENCES

[1] D. Achlioptas, J. Thorpe, manuscript.

[2] D. Achlioptas, C. Moore, *The asymptotic order of the random k-SAT threshold*, In 43rd FOCS, 126–127, 2002.

[3] D. Achlioptas, A. Naor, *The two possible values of the chromatic number of a random graph*, Annals of Mathematics, **162** (3) (2005), 1333–1349.

[4] D. Achlioptas, A. Naor, and Y. Peres, *Rigorous location of phase transitions in hard optimization problems*, Nature, **435**, June 9th 2005, 759–764.

[5] D. Achlioptas, Y. Peres, *The threshold for random k-SAT is $2^k \ln 2 - O(k)$*, Journal of American Mathematical Society, **17** (2004), 947–973.

[6] S. Aji, R.J. McEliece, *The Generalized Distributive Law*, IEEE Trans. Inform. Theory, **46** (2000), 325–343.

[7] M. Alekhnovich, E. Ben-Sasson, *Linear upper bounds for random walk on small density random 3CNFs*, In 44th FOCS, 352–361, 2003.

[8] A. Braunstein, R. Zecchina, *Survey propagation as local equilibrium equations*, J. Stat. Mech., (2004), P06007.

[9] M.-T. Chao, J. Franco. *Probabilistic analysis of two heuristics for the 3-satisfiability problem*, *SIAM J. Comput.*, **15** (4) (1986), 1106–1118.

[10] V. Chvátal, B. Reed. *Mick gets some (the odds are on his side)*, In Proc. 33th STOC, 620–627, 1992.

[11] O. Dubois and Y. Boufkhad. *A general upper bound for the satisfiability threshold of random r-SAT formulae*, Journal of Algorithms, **24** (1997), 395–420.

[12] O. Dubois, Y. Boufkhad, and J. Mandler. *Typical random 3-SAT formulae and the satisfiability threshold*, Electronic Colloquium on Computational Complexity, **10** (2003).

[13] E. Friedgut, personal communication.

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

[15] A. Kamath, R. Motwani, K. Palem, and P. Spirakis, *Tail bounds for occupancy and the satisfiability threshold conjecture*, Rand. Str. & Alg., **7** (1995), 59–80.

[16] A. Kaporis, L. M. Kirousis, and E. G. Lalas. *The probabilistic analysis of a greedy satisfiability algorithm*, In Proc. 10th ESA, 574–585, 2002.

[17] A. Kaporis, L. M. Kirousis, and E. G. Lalas. *Selecting complementary pairs of literals*, In Proc. LICS '03 workshop on Typical Case Complexity and Phase Transitions, 2003.

[18] L. M. Kirousis, E. Kranakis, D. Krizanc, and Y. Stamatiou. *Approximating the unsatisfiability threshold of random formulas*, Rand. Str. & Alg., **12** (1998), 253–269.

[19] E. Maneva, E. Mossel, and M. J. Wainwright, *A New look at Survey Propogation and its Generalizations*, In 16th SODA, 1089–1098, 2005.

[20] M. Mézard, T. Mora, and R. Zecchina, *Clustering of Solutions in the Random Satisfiability Problem*, Phys. Rev. Lett. **94** (2005), 197205. Also, cond-mat/0504070.

[21] M. Mézard, T. Mora, and R. Zecchina, *Pairs of SAT Assignments and Clustering in Random Boolean Formulae*, cond-mat/0506053, June 2nd 2005.

[22] M. Mézard, G. Parisi, and R. Zecchina, *Analytic and Algorithmic Solution of Random Satisfiability Problems*, Science **297** (2002), 812–815.

[23] N. C. Wormald, *Differential equations for random processes and random graphs*, Ann. App. Prob. , **5** (1995), 1217–1235.