SIAM J. DISCRETE MATH.
Vol. 25, No. 2, pp. 736–749

ON THE DIAMETER OF THE SET OF SATISFYING ASSIGNMENTS IN RANDOM SATISFIABLE k-CNF FORMULAS[*]

URIEL FEIGE[†], ABRAHAM D. FLAXMAN[‡], AND DAN VILENCHIK[§]

Abstract. It is known that random k-CNF formulas have a so-called satisfiability threshold at a density (namely, clause-variable ratio) of roughly $2^k \ln 2$: at densities slightly below this threshold almost all k-CNF formulas are satisfiable, whereas slightly above this threshold almost no k-CNF formula is satisfiable. In the current work we consider satisfiable random formulas and inspect another parameter—the diameter of the solution space (that is, the maximal Hamming distance between a pair of satisfying assignments). It was previously shown that for all densities up to a density slightly below the satisfiability threshold the diameter is almost surely at least roughly $n/2$ (and n at much lower densities). At densities very much higher than the satisfiability threshold, the diameter is almost surely zero (a very dense satisfiable formula is expected to have only one satisfying assignment). In this paper we show that for all densities above a density that is slightly above the satisfiability threshold (more precisely, at ratio $(1+\varepsilon)2^k \ln 2$, $\varepsilon = \varepsilon(k)$ tending to 0 as k grows) the diameter is almost surely $\mathcal{O}(k2^{-k}n)$. This shows that a relatively small change in the density around the satisfiability threshold (a multiplicative $(1+\varepsilon)$ factor) makes a dramatic change in the diameter. This drop in the diameter cannot be attributed to the fact that a larger fraction of the formulas are not satisfiable (and hence have diameter 0), because the nonsatisfiable formulas are excluded from consideration by our conditioning that the formula be satisfiable.

Key words. random k-SAT, phase transition

AMS subject classifications. 05–06

DOI. 10.1137/090749323

1. Introduction.

The computational complexity of Boolean formula satisfiability has been the focus of intensive research for decades. Recently, a promising approach to understanding the algorithmic difficulty of k-SAT has emerged in the form of a rigorous analysis of the structural properties of formulas drawn at random from certain distributions. For example, a natural distribution which has been studied extensively is the uniform distribution over k-CNF formulas with exactly m clauses over n variables. We denote this distribution by $\mathcal{F}_{n,m,k}$. Despite its simple description, many fundamental properties of this model have yet to be understood. For example, the computational complexity of deciding if a random formula is satisfiable and of finding a satisfying assignment are both major open problems [14], [21].

The clause-to-variable ratio m/n of a formula is referred to as the *density* of the formula. The random model $\mathcal{F}_{n,m,k}$ exhibits a "phase transition" in satisfiability, where sparse formulas are likely to be satisfiable, whereas dense formulas are unlikely to be satisfiable. Moreover, this phase transition happens at a very short density interval.

---

[*] Received by the editors February 11, 2009; accepted for publication (in revised form) April 8, 2010; published electronically July 1, 2011.
http://www.siam.org/journals/sidma/25-2/74932.html

[†] Department of Computer Science and Applied Mathematics, The Weizmann Institute, 76100 Rehovot, Israel (uriel.feige@weizmann.ac.il).

[‡] Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA 98121 (abie@u.washington.edu).

[§] Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, CA 90095 (vilenchik@math.ucla.edu).

There exists a satisfiability threshold $d_k = d_k(n)$ such that k-CNF formulas with density $m/n > d_k$ are not satisfiable *whp*,[1] while formulas with $m/n < d_k$ are satisfiable *whp* [16]. A first-moment-method calculation provides an upper bound of $d_k \leq 2^k \ln 2$, and the threshold is conjectured to be within a constant distance of this upper bound (for all values of k). A lower bound of $2^k \ln 2 - \mathcal{O}(k)$ was established rigorously using a weighted second-moment method in [3]. For small values of k these rigorous results are rather irrelevant. For example, for $k = 3$, the best rigorously known lower and upper bounds on the satisfiability threshold are 3.52 [18] and 4.506 [19]. In such cases, one option to "determine" the threshold is by running a complete SAT-solver on random instances with varying densities and checking when such a solver fails to find a satisfying assignment most of the time. Another option is using tools from statistical physics of disordered systems, and in particular the cavity method. Using this method the threshold for 3SAT was calculated to be 4.2667 [23], [22], a value which also agrees with experimental results [12].

For a satisfiable k-CNF formula F, let $r_{\max}(F)$ be the maximal Hamming distance between a pair of satisfying assignments of F. In this paper we study the behavior of $r_{\max}(F)$ as a function of the density. Specifically, we will consider random satisfiable formulas, and ask what the typical value of $r_{\max}$ is likely to be at various densities. Observe that as one adds more clauses to a formula, the set of satisfying assignments can only decrease, and hence also $r_{\max}$ can only decrease. This indicates that the typical value of $r_{\max}$ should decrease as the density increases. However, when the formula becomes unsatisfiable, the formula is discarded from consideration. Since the formulas of lowest diameter (diameter 0) are those discarded from consideration, and their proportion increases as the density increases, this may conceivably lead to a situation in which as the density increases, the expected diameter increases rather than decreases. In particular, there does not seem to be an a priori reason why the threshold for satisfiability should correspond to a change in typical value of the diameter of *satisfiable* formulas.

Let us review what is known about $r_{\max}(F)$ at densities below the satisfiability threshold. For $m/n \leq 2^{k-1} \ln 2$ we know that all but an $o(1)$-fraction of the formulas satisfy $r_{\max}(F) = n$ (this is because they are satisfied as NAE-k-SAT instances [2]). The results in [4] imply that (for k sufficiently large) close to the satisfiability threshold, most formulas have $r_{\max}(F) \sim n/2$ (more precisely, if $m/n = (1-\delta)2^k \ln 2,\ \delta \in (0, 1/3)$, then for most formulas $r_{\max}(F) \geq (\frac{1}{2} - \frac{5\delta^{1/2}}{6})n$). This large diameter is attributed to the existence of many small clusters of satisfying assignments, which are "spread" in the space of all 2^n possible assignments. Currently, it is widely believed that at the threshold there is more than one cluster left. Therefore there might be no reason to expect a significant change in the typical value of $r_{\max}(F)$. For densities much higher than the satisfiability threshold (by a factor of roughly $\log n$), the typical value of $r_{\max}(F)$ is 0, because such formulas, if satisfiable, are likely to have only one satisfying assignment (see, for example, [7] for the case of 3-CNF). This shows that the diameter of random satisfiable formulas undergoes a transition as the density increases (starting at n, and eventually reaching 0), but it is not clear whether there is any density that serves as a threshold around which there is a sharp drop in diameter.

In this paper we show the following.

THEOREM 1. *For all $k \geq 23$ and $m/n \geq (1 + 0.99^k)2^k \ln 2$, all but an $o(1)$-fraction of satisfiable k-CNF formulas F with m clauses over n variables satisfy*

---

[1] We say a sequence of events holds *with high probability* (*whp*) to mean with probability tending to 1 as n tends to infinity.

$$r_{\max}(F) \leq 2050k2^{-k}n.$$

Our result proves that there occurs a transition from a typical structure of satisfying assignments which are widespread in the n-dimensional binary cube to a structure where all satisfying assignments are typically contained in a ball of small diameter. The window in which this transition occurs is contained in $[(1-\varepsilon_1)2^k \ln 2, (1+\varepsilon_2)2^k \ln 2]$, where both ε_1 and ε_2 tend to 0 as k grows.

Here are a few interesting observations regarding this transition.

1. The transition in $r_{\max}$ occurs at a window of densities that lies around $2^k \ln 2$, and whose width is a low-order term w.r.t. $2^k \ln 2$. Since we are considering only satisfiable k-CNF formulas (below or above the threshold), there is no a priori reason for this transition to be found in the vicinity of the satisfiability threshold (as the latter is irrelevant for such formulas). It was brought to our attention by one of the referees that this kind of behavior may be expected using the cavity method (but since we are not deeply familiar with this method, we could not verify that point). At any rate, there is no simple convincing argument for the existence of such a transition and its location around the satisfiability threshold.
2. Since we are looking at satisfiable formulas, this is not a product distribution. Therefore some methods for establishing threshold behaviors (such as [16]) are not applicable.
3. Consider the property of having a diameter of at least r. This is not necessarily a monotone property of the density (at least we are not aware of an easy proof that it is). For example, after the satisfiability threshold, the probability is redistributed amongst satisfiable and unsatisfiable formulas, the latter receiving the majority of the support. It might be the case that the rare satisfiable formulas are the ones with large diameter. Again, this shows that some approaches to proving the existence of sharp thresholds (such as [16], which requires monotonicity) may not be applicable.
4. Typically $r_{\max} = n$ for $m/n < 2^k \ln 2/2$. This is because at those ratios most formulas are satisfiable as NAE-k-SAT formulas [2] (in which case for every satisfying assignment in the NAE (not all equal) manner, its complement at distance n is also satisfying). Numerical calculations using tools from statistical physics predict that at $2^k \ln k/k$ there is a phase transition from a typical structure of a big connected ball of satisfying assignments into many small balls of satisfying assignments (which are called clusters). Observe that $2^k \ln k/k < 2^k \ln 2/2$ for all $k \geq 3$; therefore while there is a major change in the structure of the solution space, $r_{\max}$ is not affected.

Let us briefly discuss what happens for $k < 23$. Our approach assumes that $(2 \cdot 0.99)^k$ is a low-order term compared with 2^k. This is, however, not true (or not relevant) when k is small. Also, the fact that we have a constant like 2050 in the bound on $r_{\max}$ makes the result trivial for small values of k. On the other hand, for fixed k (say $k = 3$) one can numerically estimate the value of $r_{\max}$ (via the same methods used in the proof of Theorem 1, just figuring out the exact numerics instead of a rigorous, less tight estimation that we perform). For example, for $k = 3$ the numerics show that typically $r_{\max} < 0.2n$ for density $m/n = 7.625$ (which is $\sim 1.375 \cdot 2^k \ln 2$ for $k = 3$).

Questions regarding the structure of the solution space guided the development of algorithms in similar contexts in the past (two such examples are algorithms that were developed for 3CNF formulas with a planted solution, and the intuition that served the development of the Survey Propagation algorithm [9]). In this paper we limit our study to some structural properties of the solution space and do not address algorithmic aspects, though hopefully our new insights can serve the algorithmic perspective at some point as well.

1.1. Techniques. One reasonable approach to proving Theorem 1 is to consider the uniform distribution over satisfiable k-CNF formulas with m clauses over n variables, and study $r_{\max}(F)$ of a random instance in that distribution. Throughout $\mathcal{U}_{n,m,k}$ denotes the uniform distribution. More specifically, we consider a random formula F from $\mathcal{U}_{n,m,k}$ and estimate the expected number of pairs of satisfying assignments at distance xn from each other. A similar approach was used, for example, in [3], [13], [4] for random formulas in the below-threshold regime.

The major additional challenge that we face in this present work is the fact that the uniform distribution $\mathcal{U}_{n,m,k}$ is not a product space; clause appearances are dependent, and it is unclear how to quantify this dependence. On the other hand, in the below-threshold regime, since *whp* a random k-CNF formula is satisfiable, one can study random k-CNF formulas instead of satisfiable ones. This distribution, which we denoted above by $\mathcal{F}_{n,m,k}$, is very "close" to a product space (compare with the distribution where every clause is chosen independently at random with probability $p = m/(2^k\binom{n}{k})$, which is already a product space).

One demonstration of this technical challenge is the difficulty of answering the following question: given a fixed assignment ψ, what is the probability that it satisfies a random F? If F is drawn from $\mathcal{F}_{n,m,k}$, then the answer is simple, $\Pr[\psi \models F] = (1-2^{-k})^m$. If F is drawn from $\mathcal{U}_{n,m,k}$, then giving an explicit expression (as a function of m, n, k) for $\Pr[\psi \models F]$ is still an open question.

We will show that for $x \geq 2050k2^{-k}$ the expected number of pairs of satisfying assignments at distance xn from each other is much smaller than $1/n$. Since there are at most n possible ways to choose x, we can use the union bound to prove that *whp* F has the desired properties (since $\mathcal{U}_{n,m,k}$ is the uniform distribution, showing that the property holds *whp* translates immediately to a deterministic statement about all but a vanishing fraction of satisfiable formulas).

To derive our estimate on the expected number of pairs of satisfying assignments at distance xn we first analyze a different distribution which is commonly called the *planted distribution*, and we shall denote it by $\mathcal{P}_{n,m,k}$. To generate a formula according to $\mathcal{P}_{n,m,k}$, fix an assignment uniformly at random, then include m clauses uniformly at random out of $(2^k-1)\binom{n}{k}$ clauses that are consistent with the "planted" assignment.

When working with $\mathcal{P}_{n,m,k}$, the clauses are nearly independent and calculation is much easier. We then relate the planted model and the uniform model to obtain the desired result. The idea of translating bounds from the planted to the uniform model was used in [1], [4], [13] for the below-threshold regime, and also in [10], [11] but in a different context.

The reader may wonder at this point what happens when $m/n < (1+0.99^k)2^k \ln 2$? Do typically all satisfying assignments lie in a low-diameter Hamming ball all the way down to the satisfiability threshold (or even below it)? We believe (based on some numerical calculations that we performed) that Theorem 1 can be extended (maybe with some changes in the upper bound on $r_{\max}$) down to $m/n = 2^k \ln 2 + \mathcal{O}(k)$ (which is an $\mathcal{O}(k)$-additive term from the satisfiability threshold). This extension is done using the same technique of going through the planted distribution. However, when $m/n = 2^k \ln 2 + \mathcal{O}(k)$ this technique breaks. In section 5 we discuss this issue and suggest another technique that may prove useful when our first technique fails.

2. Relating the uniform and the planted distributions.

Let u_x be a random variable counting the number of pairs of satisfying assignments at distance xn from each other that a random formula in $\mathcal{U}_{n,m,k}$ has. Let T denote the expected number of satisfying assignments that a random formula in $\mathcal{U}_{n,m,k}$ has, and f_x a random variable which denotes the number of satisfying assignments at distance xn from the planted assignment, had F been distributed according to $\mathcal{P}_{n,m,k}$. The following proposition allows us to upper bound $E[u_x]$ via the more accessible quantity $E[f_x]$.

Proposition 2. *Let F be a random formula sampled according to $\mathcal{U}_{n,m,k}$, $u_x = u_x(F)$, $f_x = f_x(F)$; then*

$$E[u_x] = T \cdot E[f_x]/2.$$

(A similar approach for relating the uniform and the planted distribution can be found in [13], though in that case the uniform distribution was the nonconditioned one.)

Proof. For two satisfying assignments φ_i, φ_j we use $\delta(\varphi_i, \varphi_j)$ to denote their Hamming distance. Consider some ordering on the 2^n possible assignments, and let A_i be an indicator variable which is 1 if φ_i satisfies F. Using this terminology,

$$u_x = \frac{1}{2} \sum_{i,j\,:\,\delta(\varphi_i,\varphi_j)=xn} A_i \cdot A_j.$$

Linearity of expectation gives

$$E[u_x] = \frac{1}{2} \sum_{i,j\,:\,\delta(\varphi_i,\varphi_j)=xn} \Pr[A_i \wedge A_j] = \frac{1}{2} \sum_{\delta(\varphi_i,\varphi_j)=xn} \Pr[A_i|A_j]\Pr[A_j].$$

By symmetry, the latter equals

$$2^n \cdot \frac{\Pr[A_j]}{2} \cdot \sum_{i\,:\,\delta(\varphi_i,\varphi_j)=xn} \Pr[A_i|A_j].$$

It remains to estimate $\Pr[A_i|A_j]$. Conditioning on the event A_j means conditioning on the fixed assignment φ_j to be satisfying. In turn, $\mathcal{U}_{n,m,k}$ conditioned on φ_j being a satisfying assignment means that only clauses which are satisfied by φ_j can be included, and by symmetry, every set of t clauses satisfied by φ_j has the same probability of being included. Observe that for $t = m$ this is exactly the definition of the planted distribution $\mathcal{P}_{n,m,k}$. Therefore $\sum_i \Pr[A_i|A_j] = E[f_x]$ when summing over all assignments φ_i at distance xn from φ_j. Furthermore, $T = \sum_j \Pr[A_j]$ (now we are summing over all 2^n assignments), and hence $\Pr[A_j] = T/2^n$. Putting everything together, we derive

$$E[u_x] = T \cdot E[f_x]/2. \qquad \square$$

In [13] this sort of proposition was already enough to estimate $E[u_x]$ since T can be easily calculated when m/n is below the satisfiability threshold. However, in $\mathcal{U}_{n,m,k}$, m/n above the satisfiability threshold, it is not clear how to calculate T. The following lemma is then useful (the proof can also be found in [11] and is given here for completeness).

Lemma 3. *Let W be the expected number of satisfying assignments of a random $\mathcal{P}_{n,m,k}$ instance. Then always $T \leq W$.*

Proof. Let t_i be the number of formulas on n variables and m clauses which have exactly i satisfying assignments. Let p_i be the probability that a formula with exactly i satisfying assignments is sampled from $\mathcal{U}_{n,m,k}$, and let q_i be defined similarly for $\mathcal{P}_{n,m,k}$. Observe that due to symmetry, sampling a formula from $\mathcal{P}_{n,m,k}$ is exactly equivalent to sampling a pair (φ, F) uniformly at random from all pairs such that φ is an assignment and F is a formula satisfied by φ. Hence

$$p_i = \frac{t_i}{\sum_{j=1}^{2^n} t_j}, \qquad q_i = \frac{i \cdot t_i}{\sum_{i=1}^{2^n} i \cdot t_i}$$

and

$$T = \sum_{i=1}^{2^n} i \cdot p_i = \frac{\sum_{i=1}^{2^n} i \cdot t_i}{\sum_{i=1}^{2^n} t_i},$$

$$W = \sum_{i=1}^{2^n} i \cdot q_i = \frac{\sum_{i=1}^{2^n} i^2 \cdot t_i}{\sum_{i=1}^{2^n} i \cdot t_i}.$$

Therefore to prove $T \leq W$, it suffices to show

$$\left(\sum_{i=1}^{2^n} i \cdot t_i\right)^2 \leq \left(\sum_{i=1}^{2^n} t_i\right) \cdot \left(\sum_{i=1}^{2^n} i^2 \cdot t_i\right).$$

This is just Cauchy–Schwarz, $(\sum a_i \cdot b_i)^2 \leq (\sum a_i^2) \cdot (\sum b_i^2)$, with $a_i = \sqrt{t_i}$ and $b_i = i \cdot \sqrt{t_i}$. $\square$

3. The planted setting. In this section we analyze W and $E[f_x]$. Recall that we use W to denote the expected number of satisfying assignments that a random formula in $\mathcal{P}_{n,m,k}$ has, and f_x counts the number of satisfying assignments at distance xn from the planted assignment, had F belonged to $\mathcal{P}_{n,m,k}$.

Our analysis of $E[f_x]$ is composed of two regimes. The first is the case $x \in [0, 1/k]$. In this regime we know that $E[f_x]$ changes from $\omega(1)$ to $o(1)$. This phenomenon is depicted in Figure 1. The y-axis in the plot is $f^\star(x)$ such that $E[f_x] = e^{f^\star(x)n}$, and the x-axis is the

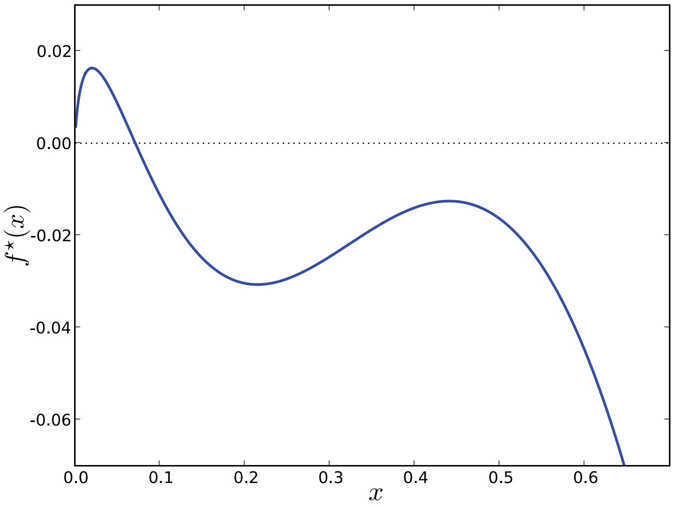

FIG. 1. *Plot of $f^\star(x)$ for $k = 6$ and $\varepsilon = 2^{-k}$.*

Hamming distance from the planted assignment. Therefore the transition from $E[f_x] = \omega(1)$ to $E[f_x] = o(1)$ corresponds to $f^\star(x)$ changing sign from positive to negative.

To translate our results to the uniform setting, it turns out that we need to have a more precise control on the rate in which $E[f_x]$ decreases once changing to $o(1)$. Therefore the analysis of that regime is more careful (Proposition 6). Then we analyze the case $x \in [1/k, 1]$. In this regime, for a suitable choice of ε (recall $m/n = (1+\varepsilon)2^k \ln 2$), $E[f_x]$ is constantly $o(1)$ (in fact, exponentially small in n). Therefore a more crude analysis will suffice (Proposition 5). This corresponds in Figure 1 to the fact that the curve is bounded away below the x-axis in that range.

In this section we consider a slight modification of $\mathcal{P}_{n,m,k}$. Instead of choosing m clauses uniformly at random, we choose m clauses with repetitions. However, for $m/n = \mathcal{O}(1)$, the expected number of pairs of identical clauses in F (in the modified model) is $\mathcal{O}(m^2/n^k)$. Thus, for $k \geq 3$ this quantity is $o(1)$. Therefore, as standard calculations show, every property that holds with probability q in the modified model holds with probability $q(1+\mathcal{O}(1))$ in $\mathcal{P}_{n,m,k}$. Somewhat abusing notation, we will denote the modification also by $\mathcal{P}_{n,m,k}$.

Let us start by formulating $E[f_x]$ in a way which is convenient to work with.

LEMMA 4.

$$E[f_x] \leq \binom{n}{xn} \cdot \left(1 - \frac{1-(1-x)^k}{2^k-1}\right)^m.$$

Proof. Fix an assignment ψ at distance xn from the planted assignment φ. The probability that ψ also satisfies F can be calculated in the following manner. Let A be the set of variables on which both ψ and φ agree. $|A| = (1-x)n$. Consider a random clause C satisfied by φ; if all k variables in that clause fall in A, then C is surely satisfied by ψ. The probability for that is $q = \binom{(1-x)n}{k}/\binom{n}{k}$. If at least one variable falls out of A, which happens with probability $1-q$, then the clause is satisfied only with probability $\frac{2^k-2}{2^k-1}$. This is because there is one way to complement the variables which is not consistent with ψ but is consistent with φ. There are $\binom{n}{xn}$ ways to fix ψ, and therefore

$$\begin{aligned} E[f_x] &= \binom{n}{xn}\left(q \cdot 1 + (1-q)\cdot\frac{2^k-2}{2^k-1}\right)^m \\ &= \binom{n}{xn}\left(\frac{2^k-2+q}{2^k-1}\right)^m = \binom{n}{xn}\left(1 - \frac{1-q}{2^k-1}\right)^m. \end{aligned}$$

Finally, observing that $q \leq (1-x)^k$ proves the lemma. ☐

It will be more convenient to work with the following quantity:

$$f^\star(x) \equiv \frac{\ln E[f_x]}{n}. \tag{3.1}$$

One can verify that (using Lemma 4 and Stirling's formula)

$$f^\star(x) \leq H(x)\ln 2 + c\,\ln\left(1 - \frac{1-(1-x)^k}{2^k-1}\right), \tag{3.2}$$

where $H(x)$ denotes the binary entropy measure,

$$H(x) = -(1-x)\log_2(1-x) - x\log_2 x,$$

and $c = m/n = (1+\varepsilon)2^k \ln 2$.

To make use of Proposition 2 we need to obtain tight bounds on W and $E[f_x]$. In terms of $f^\star(x)$, $E[f_x] = e^{f^\star(x)n}$; therefore to prove $E[f_x] = o(1)$ it suffices to prove $f^\star(x) < 0$. This is exactly what the following two propositions formally establish.

Proposition 5. *For any $k \geq 23$, $\varepsilon \geq 0.99^k$, and $x \in [1/k, 1]$,*

$$f^\star(x) \leq -2050k2^{-k}.$$

Proof. We break the interval $[1/k, 1]$ into two subintervals. Let us first consider $x \in [0.3, 1]$. Always $H(x)\ln 2 \leq \ln 2$, and on the other hand, using $\ln(1-x) \leq -x$,

$$\begin{aligned} c\ln\left(1 - \frac{1-(1-x)^k}{2^k-1}\right) &\leq -\frac{(1+\varepsilon)2^k \ln 2}{2^k-1}(1-(1-x)^k) \\ &\leq -(1+\varepsilon)\cdot(1-(1-x)^k)\cdot \ln 2. \end{aligned}$$

Therefore it suffices to prove that $(1+\varepsilon)(1-(1-x)^k) \geq 1 + (2050k2^{-k}/\ln 2)$ for every $x \in [0.3, 1]$. Indeed,

$$(1-(1-x)^k) \geq (1-0.7^k), \qquad (1+\varepsilon) \geq (1+0.99^k).$$

One can verify that for $k \geq 23$, multiplying these two quantities is always greater than $1 + (2050k2^{-k}/\ln 2)$.

Let us now move to the case $x \in [1/k, 0.3]$. $H(x)$ is monotonically increasing until $x = 0.5$; therefore, it takes its maximal value in this interval at $x = 0.3$. $1-(1-x)^k$ takes its minimal value at $1/k$. Observe that $(1-1/k)^k \leq e^{-1}$, and therefore

$$1-(1-x)^k \geq 1 - 1/e.$$

In this case we have $f^\star(x) \leq H(0.3)\ln 2 - (1-1/e) \leq -0.02 < -2050k2^{-k}$ for every $k \geq 23$. ∎

Proposition 6. *For any $k \geq 23$, $\varepsilon \geq 0$, and $\lambda \in [2900, 2^k/k]$, if $x = \lambda 2^{-k}$, then $f^\star(x) \leq -\lambda 2^{-k}$.*

Proof. For any $x \in [0, 1/k]$ we have

$$\ln(1-x) \geq -2x,$$

and, for $0 \leq x \leq 1$,

$$1-(1-x)^k \geq kx - \frac{k^2x^2}{2}.$$

Thus,

$$\begin{aligned}
&H(x)\ln 2 + c\ln\left(1-\frac{1-(1-x)^k}{2^k-1}\right)\\
&\quad= -x\ln x-(1-x)\ln(1-x)+(1+\varepsilon)2^k(\ln 2)\ln\left(1-\frac{1-(1-x)^k}{2^k-1}\right)\\
&\quad\le -x\ln x+2x(1-x)-(1+\varepsilon)2^k(\ln 2)\left(\frac{1-(1-x)^k}{2^k-1}\right)\\
&\quad\le -x\ln x+2x-(1+\varepsilon)(\ln 2)\left(kx-\frac{k^2x^2}{2}\right).
\end{aligned}$$

Substituting $\lambda 2^{-k}$ for x, this upper bound becomes

$$\begin{aligned}
&-x\ln x+2x-(1+\varepsilon)(\ln 2)\left(kx-\frac{k^2x^2}{2}\right)\\
&\quad= \lambda 2^{-k}(k(\ln 2)-\ln\lambda)+2\lambda 2^{-k}-(1+\varepsilon)(\ln 2)(k\lambda 2^{-k}-k^2\lambda^2 2^{-2k-1})\\
&\quad= -(\lambda\ln\lambda)2^{-k}+2\lambda 2^{-k}-\varepsilon(\ln 2)(k\lambda 2^{-k}-k^2\lambda^2 2^{-2k-1})+(\ln 2)k^2\lambda^2 2^{-2k-1}\\
&\quad= -\lambda 2^{-k}((\ln\lambda)-2+\varepsilon(\ln 2)(k-k^2\lambda 2^{-k-1})-(\ln 2)k^2\lambda 2^{-k-1})\\
&\quad= -\lambda 2^{-k}((\ln\lambda)\left(1-(\ln 2)k^2\frac{\lambda}{\ln\lambda}2^{-k-1}\right)-2+\left(\varepsilon(\ln 2)(k-k^2\lambda 2^{-k-1}))\right).
\end{aligned}$$

Observe that $\lambda \le 2^k/k$ and thus

$$k-k^2\lambda 2^{-k-1}\ge 0,$$

and since $\varepsilon \ge 0$ it suffices to prove that

$$(\ln\lambda)\left(1-(\ln 2)k^2\frac{\lambda}{\ln\lambda}2^{-k-1}\right)-2\ge 1.$$

Since $\lambda \le 2^k/k$ and $k \ge 23$, we have

$$\begin{aligned}
(\ln 2)k^2\frac{\lambda}{\ln\lambda}2^{-k-1} &\le (\ln 2)k^2\frac{2^k/k}{k(\ln 2)-\ln k}2^{-k-1}\\
&= (\ln 2)\frac{1}{2((\ln 2)-(\ln k)/k)}\le 0.623,
\end{aligned}$$

and so it suffices to verify that

$$\ln\lambda \ge 3/(1-0.623),$$

which is always true for $\lambda \in [2900, 2^k/k]$. ∎

4. Proof of Theorem 1. Recall Proposition 2 and Lemma 3, which together establish

$$E[u_x]\le W\cdot E[f_x]/2.$$

W is the expected number of satisfying assignments in the planted model, $W = \sum_x E[f_x]$.

The idea of the proof is to use Propositions 5 and 6 to upper bound W by looking at the largest x s.t. $E[f_x]$ contributes to W (that is, $E[f_x]$ is not vanishing with n). We shall use x_0 to denote this number (regardless, observe that x_0 is an upper bound on the diameter of the cluster region in the *planted* setting). Then, to beat W, we take $x_1 > x_0$, so that for every $x \geq x_1$, $E[f_x] \cdot W \ll 1$. Respectively, x_1 upper bounds the diameter of the cluster region in the uniform setting. It turns out that $x_1/x_0 = \mathcal{O}(k)$, and since x_0 scales down with 2^{-k}, this additional factor is manageable.

Formally, Propositions 5 and 6 assert that only $x \leq 2900 \cdot 2^{-k}$ may contribute to the value of W. Indeed, take $x_0 = 2900 \cdot 2^{-k}$; then $E[f_x] = o(n^{-1})$ for every $x > x_0$. For $x \leq x_0$, the total number of possible assignments (which obviously bounds the expected number of satisfying assignments) at distance xn from the planted is

$$\binom{n}{xn} \leq \left(\frac{en}{xn}\right)^{xn} \leq e^{(1-\ln x)xn}.$$

This quantity is maximized for $x \leq x_0$ at x_0; therefore, for sufficiently large n,

$$W \leq o(1) + \sum_{x \leq x_0} \binom{n}{xn} \leq n e^{(k \ln 2 + 1 - \ln 2900) 2900 \cdot 2^{-k} n} \leq n e^{2020 k 2^{-k} n}.$$

Now take $x_1 = 2050k2^{-k}$; applying Propositions 5 and 6 once more gives that for $x \geq x_1$,

$$E[f_x] \leq e^{-2050k2^{-k}n}.$$

In turn, for $x \geq x_1$

$$E[u_x] \leq W \cdot E[f_x]/2 \leq e^{2020k2^{-k}n} \cdot e^{-2050k2^{-k}n} = e^{-30k2^{-k}n}.$$

Using Markov's inequality, for $x \geq x_1$,

$$\Pr[u_x > 0] \leq e^{-30k2^{-k}n} = e^{-\Omega(n)}.$$

Applying the union bound,

$$\Pr[\exists x \geq 2050k2^{-k}, u_x > 0] \leq n \cdot e^{-\Omega(n)}.$$

5. Moving even closer to the threshold.

In the previous sections we showed that when $m/n \geq (1 + 0.99^k)2^k \ln 2$ for $k \geq 23$, *whp* there are no pairs of satisfying assignments at distance greater than $\mathcal{O}(k2^{-k})$ from each other (Theorem 1). Our approach was to consider the planted distribution and estimate $E[f_x]$, the expected number of satisfying assignments at distance xn from the planted assignment. Then we used Proposition 2 to relate this quantity to $E[u_x]$, the expected number of pairs of satisfying assignments at distance xn from each other (in the uniform setting). The relation we established was given (in Proposition 2) by

$$E[u_x] \leq W \cdot E[f_x].$$

W is the expected number of satisfying assignments in $\mathcal{P}_{n,m,k}$.

Observe that W is always at least 1, and therefore using this relation to show that $E[u_x] = o(1)$ makes sense only when $E[f_x] = o(1)$. However, using (rather tedious) calculations one can show that when $m/n = 2^k \ln 2 + \mathcal{O}(k)$ there exists $x \in$

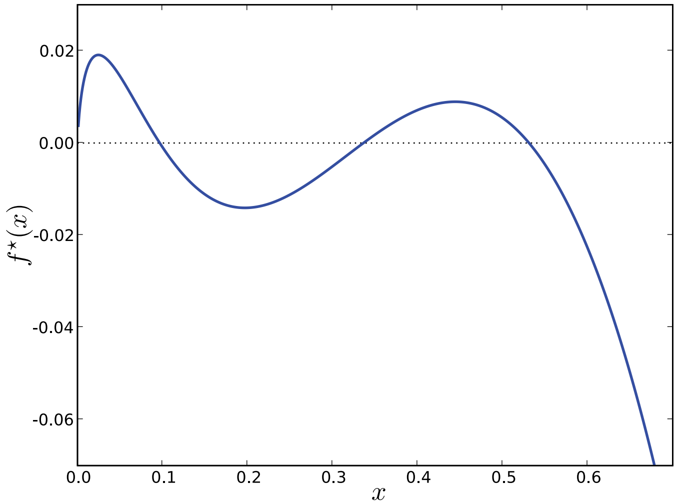

FIG. 2. *Plot of* $f^\star(x)$ *for* $k = 6$ *and* $\varepsilon = -2^{-k}$.

$[0.5 - \mathcal{O}(2^{-k}), 0.5]$ such that $E[f_x]$ is exponentially large in n (details omitted). This phenomenon is depicted in Figure 2. Therefore from this density downwards our method fails (observe that $E[f_x]$ is monotonically decreasing and continuous in m/n).

Compare the plots in Figures 1 and 2. Both depict on the y-axis $f^\star(x) \equiv \frac{\ln E[f_x]}{n}$ and the distance from the planted assignment on the x-axis. To generate the plots we used the estimate on $E[f_x]$ given in Lemma 4. Although Lemma 4 establishes an upper bound on $E[f_x]$, in fact for x bounded away from 0 equality holds (up to an $o(1)$ additive factor inside the parenthesis). Since $E[f_x]$ is monotonically decreasing in m/n and continuous, as m/n gets smaller, the "hunchback" around $x = 1/2$ gets closer to the x-axis, and at some ratio crosses it to become positive. This ratio occurs at $m/n = 2^k \ln 2 + \mathcal{O}(k)$. As k grows, the hunchback (regardless of whether it is above or below the x-axis) becomes narrower and in general is concentrated in an interval of width $\mathcal{O}(2^{-k})$ around $1/2$, with the maximum occurring at $1/2 - \mathcal{O}(2^{-k})$. We have validated these claims using a combination of numerical and rigorous calculations (details omitted here).

In this section we suggest a new technique which in some sense refines the one we used. Using our refined technique we can prove, for example, that at some settings, even though $E[f_x]$ is exponential in n (which means that our original technique fails), in fact *whp* $f_x = 0$. Hopefully this refinement can benefit the uniform distribution as well. We do not discuss this point in the present paper. The key to the refinement is to replace f_x with another quantity which counts maximal satisfying assignments at distance xn from the planted assignment—$f_x^{\max}$. This notion is similar to the notion of minimal satisfying assignments used in [20]. To demonstrate the power of this new technique we describe a setting where $E[f_x] \geq 1$ (which means that our original technique fails) for some $x \in [0.3, 0.6]$, but $E[f_x^{\max}] = o(1)$ for all $x \in [0.3, 0.6]$, and in that setting this will imply that *whp* $f_x = 0$. To do this we shall assume that the following is true.

Assumption. Let F be distributed according to $\mathcal{P}_{n,m,k}$, with k a sufficiently large constant. There exists an $\varepsilon > 0$ so that if $m/n \geq (1+\varepsilon)2^k \ln 2$, then *whp* F has no satisfying assignments at distance xn from the planted assignment for $0.2 \leq x \leq 0.3$ or $0.6 \leq x \leq 1.0$.

This assumption can be proven rigorously. Relying on this assumption, we can prove the following proposition.

PROPOSITION 7. *There exists a nonempty interval* $(\varepsilon_2, \varepsilon_1)$ *in which for every* $\varepsilon \in (\varepsilon_2, \varepsilon_1)$ *and* F *distributed according to* $\mathcal{P}_{n,m,k}$, $m = (1+\varepsilon)2^k \ln 2$, *there exists* $x \in [0.3, 0.6]$ *so that* $E[f_x] \geq 1$, *while whp* $f_x = 0$ *for every* x *in that interval.*

We defer the proof of Proposition 7 to the end of this section. Let us now formally define the notion of maximal satisfying assignments. For two assignments φ, φ' let $A(\varphi, \varphi')$ be the set of variables on which both assignments agree.

DEFINITION 8. *Given a planted instance* F *with a planted assignment* φ, *we say that a satisfying assignment* φ' *of* F *is maximal if flipping any* $x \in A(\varphi, \varphi')$ *results in an assignment that does not satisfy* F.

In that sense φ' is in a maximal Hamming distance from φ. For example, if the complement of the planted also satisfies F, then it is maximal (in a vacuous way).

Let ε_1 be the maximal value such that for $m/n = (1+\varepsilon_1)2^k \ln 2$ and some $x \in [0.3, 0.6]$,

$$E[f_x] \geq 1.$$

Let ε_2 be the minimal value such that for $m/n = (1+\varepsilon_2)2^k \ln 2$ and every $x \in [0.3, 0.6]$

$$E[f_x^{\max}] \leq n^{-2}.$$

The proofs of Propositions 5 and 6 show that ε_2 always exists, and we have verified the existence of ε_1 numerically. The condition $E[f_x^{\max}] \leq n^{-2}$ for $x \in [0.3, 0.6]$ easily translates to the following claim: *whp* there are no maximal satisfying assignments at distance xn for $x \in [0.3, 0.6]$. This follows from Markov's inequality, which gives an upper bound of n^{-2} on the probability that $f_x^{\max} > 0$ (for a fixed x). Now take the union bound over at most n possible values of x.

Before proving Proposition 7, we still need to show that the interval $(\varepsilon_2, \varepsilon_1)$ is not empty.

PROPOSITION 9. $\varepsilon_2 < \varepsilon_1$.

Proof. Fix $x \in [0.3, 0.6]$ and consider a random formula F from $\mathcal{P}_{n,m,k}$. Let M_i be the event that φ_i at distance xn from the planted assignment φ is maximal, and A_i the event that φ_i satisfies F. Using this terminology,

$$E[f_x^{\max}] = \sum_{i:\, \delta(\varphi_i)=xn} \Pr[A_i \wedge M_i] = \sum_i \Pr[M_i|A_i] \Pr[A_i] = \Pr[M_i|A_i] E[f_x]. \tag{5.1}$$

In the last step we used the fact that $\Pr[M_i|A_i]$ is the same for every φ_i by symmetry, and therefore we can pull it out in front of the summation. It remains to estimate $\Pr[M_i|A_i]$. Conditioning on the event A_i in the planted model means conditioning on the fixed assignment φ_i to be satisfying in addition to the planted assignment. In other words this means that only clauses which are satisfied by both φ_i and φ can be included. By symmetry, every set of t clauses satisfied by both has the same probability of being included. Observe that for $t = m$ this is exactly the definition of the doubly planted distribution (the distribution where one chooses randomly among clauses satisfied by *both* assignments).

A standard approach is to consider the following variation of the doubly planted model: pick every clause satisfied by both φ_i and φ with probability p, where p satisfies $p = m/|S|$, S being the set of clauses which are satisfied by both φ_i, φ. For the properties that interest us, it is straightforward to translate results between these two models. It is also easy to see that $|S| \geq (2^k - 2)\binom{n}{k}$.

Now consider a variable s in φ_i whose assignment agrees with φ, and without loss of generality, assume it is TRUE. We call a clause C s-qualifying for φ_i if it takes the form $(s \vee \ell_{y_1} \vee \ell_{y_2} \vee \cdots \vee \ell_{y_{k-1}})$, where ℓ_{y_j} is a FALSE literal (over the variable y_j) under φ_i. If φ_i is maximal, then at least one of the $\binom{n}{k-1}$ s-qualifying clauses had to be included. The probability that at least one such clause is included is at most

$$1 - (1-p)^{\binom{n}{k-1}} \leq 1 - e^{-km/(n(2^k-2))}.$$

Next we observe that φ_i has at least $(1-x)n$ variables which are assigned according to φ. Also observe that the set of s-qualifying clauses is disjoint from the set of q-qualifying clauses. Finally, for φ_i to be maximal there must be at least one s-qualifying clause in F for every variable s. The probability for that is at most

$$\Pr[M_i|A_i] \leq (1 - e^{-km/(n(2^k-2))})^{(1-x)n} \leq (1 - (1-x)e^{-km/(n(2^k-2))})^n \equiv a^n \tag{5.2}$$

for some $a = a(k) < 1$ (here we assumed that $x \in [0.3, 0.6]$ and therefore $(1-x) \in [0.4, 0.7]$). Combining (5.1) and (5.2) we derive

$$E[f_x^{\max}] \leq E[f_x] \cdot a^n. \tag{5.3}$$

We claim that this implies $\varepsilon_1 - \varepsilon_2 \geq h$ for some $h = h(k) > 0$ (h actually depends on a, but a depends only on k). Fix some $b = b(k) > 1$ s.t. $b \cdot a < 1$ (since $a = a(k) < 1$, such b exists). Since $E[f_x]$ is continuous and decreasing in m/n, and by the maximality of ε_1, we can find $h = h(k) > 0$ s.t. $E[f_x] \leq b^n$ for all $x \in [0.3, 0.6]$ when $m/n \leq (1 + \varepsilon_1 - h)2^k \ln 2$. On the other hand, as (5.3) implies, $E[f_x^{\max}] \leq b^n \cdot a^n = (ab)^n \leq n^{-2}$ (for sufficiently large n) for all $x \in [0.3, 0.6]$. By the minimality of ε_2 this in particular implies that $\varepsilon_2 \leq \varepsilon_1 - h$.

Proof of Proposition 7. Fix some $\varepsilon \in (\varepsilon_2, \varepsilon_1)$ and consider a random formula F in $\mathcal{P}_{n,m,k}$ so that $m/n = (1+\varepsilon)2^k \ln 2$. By the choice of $\varepsilon > \varepsilon_2$, it holds that *whp* F has no maximal satisfying assignments at distance xn from the planted assignment for $x \in [0.3, 0.6]$. Assume that indeed this is the case, and also assume that the assumption we made holds.

By the choice of $\varepsilon < \varepsilon_1$ and the maximality of ε_1, for some $x_1 \in [0.3, 0.6]$ indeed $E[f_{x_1}] \geq 1$. We shall now show that $f_x = 0$ for all $x \in [0.3, 0.6]$. Assume by contradiction that $f_x > 0$ for some $x \in [0.3, 0.6]$. Namely, there exists a satisfying assignment ψ at distance xn from the planted assignment, φ. Construct the assignment ψ' in the following manner: while possible, flip the assignment of a variable that agrees with φ that leaves the assignment satisfying. By construction it is clear that ψ' is maximal. The crucial observation now is that at each iteration of the process we increase the distance between the current assignment and the planted assignment by exactly one. Specifically, we start the procedure with an assignment at distance xn for $x \in [0.3, 0.6]$, and we keep increasing the distance. If the final distance yn is s.t. $y \not\in [0.3, 0.6]$, then at some point we've reached a satisfying assignment at distance $\geq 0.6n + 1$. This contradicts our assumption above. Therefore we have that ψ', a maximal satisfying assignment already, is at distance yn for $y \in [0.3, 0.6]$. This, however, contradicts our assumption that no maximal satisfying assignments exist at that range.

Acknowledgment. Large parts of this work were performed at Microsoft Research, Redmond, Washington.

REFERENCES

[1] D. Achlioptas and A. Coja-Oghlan, *Algorithmic barriers from phase transitions*, in Proceedings of the 49th IEEE Symposium on Foundations of Computer Science, 2008, pp. 793–802.

[2] D. Achlioptas and C. Moore, *Almost all graphs with average degree* 4 *are* 3*-colorable*, in Proceedings of the 34th ACM Symposium on Theory of Computing, 2002, pp. 199–208.

[3] D. Achlioptas and Y. Peres, *The threshold for random k-SAT is* $2^k \log 2 - \mathcal{O}(k)$, J. Amer. Math. Soc., 17 (2004), pp. 947–973.

[4] D. Achlioptas and F. Ricci-Tersenghi, *On the solution-space geometry of random constraint satisfaction problems*, in Proceedings of the 38th ACM Symposium on Theory of Computing, 2006, pp. 130–139.

[5] N. Alon and N. Kahale, *A spectral technique for coloring random* 3*-colorable graphs*, SIAM J. Comput., 26 (1997), pp. 1733–1748.

[6] N. Alon, M. Krivelevich, and B. Sudakov, *Finding a large hidden clique in a random graph*, Random Structures Algorithms, 13 (1998), pp. 457–466.

[7] E. Ben-Sasson, Y. Bilu, and D. Gutfreund, *Finding a Randomly Planted Assignment in a Random* 3*CNF*, manuscript, 2002.

[8] A. Blum and J. Spencer, *Coloring random and semi-random k-colorable graphs*, J. Algorithms, 19 (1995), pp. 204–234.

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

[10] H. Chen, *An algorithm for SAT above the threshold*, in Proceedings of the 6th International Conference on Theory and Applications of Satisfiability Testing, 2003, pp. 14–24.

[11] A. Coja-Oghlan, M. Krivelevich, and D. Vilenchik, *Why almost all satisfiable k-CNF formulas are easy*, in Proceedings of the 13th Conference on Analysis of Algorithms (DMTCS), 2007, pp. 89–102.

[12] J. M. Crawford and L. D. Auton, *Experimental results on the crossover point in random* 3*-SAT*, Artificial Intelligence, 81 (1996), pp. 31–57.

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

[14] U. Feige, *Relations between average case complexity and approximation complexity*, in Proceedings of the 34th ACM Symposium on Theory of Computing, 2002, pp. 534–543.

[15] U. Feige and R. Krauthgamer, *Finding and certifying a large hidden clique in a semirandom graph*, Random Structures Algorithms, 16 (2000), pp. 195–208.

[16] E. Friedgut, *Sharp thresholds of graph properties, and the k-SAT problem*, J. Amer. Math. Soc., 12 (1999), pp. 1017–1054.

[17] C. Hui and A. M. Frieze, *Coloring bipartite hypergraphs*, in Proceedings of the 5th International Conference on Integer Programming and Combinatorial Optimization, 1996, pp. 345–358.

[18] A. C. Kaporis, L. M. Kirousis, and E. G. Lalas, *The probabilistic analysis of a greedy satisfiability algorithm*, in Proceedings of the 10th Annual European Symposium on Algorithms, Lecture Notes in Comput. Sci. 2461, Springer-Verlag, Berlin, 2002, pp. 574–585.

[19] A. C. Kaporis, L. M. Kirousis, Y. C. Stamatiou, M. Vamvakari, and M. Zito, *Coupon collectors, q-binomial coefficients and the unsatisfiability threshold*, in Proceedings of the 7th Italian Conference on Theoretical Computer Science, Lecture Notes in Comput. Sci. 2202, Springer-Verlag, Berlin, 2001, pp. 328–338.

[20] L. Kirousis, E. Kranakis, D. Krizanc, and Y. Stamatiou, *Approximating the unsatisfiability threshold of random formulas*, Random Structures Algorithms, 12 (1998), pp. 253–269.

[21] L. A. Levin, *Average case complete problems*, SIAM J. Comput., 15 (1986), pp. 285–286.

[22] S. Mertens, M. Mézard, and R. Zecchina, *Threshold values of random k-SAT from the cavity method*, Random Structures Algorithms, 28 (2006), pp. 340–373.

[23] M. Mézard and R. Zecchina, *Random k-satisfiability problem: From analytic solution to an efficient algorithm*, Phys. Rev. E, 66 (2002), 056126.