Theoretical Computer Science 407 (2008) 359–369

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

On the satisfiability threshold and clustering of solutions of random 3-SAT formulas

Elitza Maneva [a,*], Alistair Sinclair [b]

[a] IBM Almaden Research Center, San Jose, CA, United States

[b] Computer Science Division, University of California at Berkeley, United States

ARTICLE INFO

Article history:
Received 28 January 2008
Received in revised form 20 June 2008
Accepted 27 June 2008
Communicated by J. Diaz

Keywords:
Random 3-SAT
Phase transitions
Survey propagation
Threshold of satisfiability
First-moment method

ABSTRACT

We study the structure of satisfying assignments of a random 3-SAT formula. In particular, we show that a random formula of density $\alpha \geq 4.453$ almost surely has no non-trivial "core" assignments. Core assignments are certain partial assignments that can be extended to satisfying assignments, and have been studied recently in connection with the Survey Propagation heuristic for random SAT. Their existence implies the presence of clusters of solutions, and they have been shown to exist with high probability below the satisfiability threshold for k-SAT with $k \geq 9$ [D. Achlioptas, F. Ricci-Tersenghi, On the solution-space geometry of random constraint satisfaction problems, in: Proc. 38th ACM Symp. Theory of Computing, STOC, 2006, pp. 130–139]. Our result implies that either this does not hold for 3-SAT, or the threshold density for satisfiability in 3-SAT lies below 4.453. The main technical tool that we use is a novel simple application of the first moment method.

1. Introduction

The study of random instances of 3-SAT has been a major research focus in recent years, both because of its inherent interest, and because it is a natural test case for the wider understanding of the complexity of computational tasks on random inputs. In random 3-SAT the input is a formula drawn uniformly at random from all formulas of fixed density α, i.e., formulas with αn clauses on n variables. Friedgut [7] proved that there exists a function $\alpha_c(n)$, known as the *satisfiability threshold*, such that for any positive ϵ, random formulas of density $\alpha_c(n) - \epsilon$ have satisfying assignments with high probability, and random formulas of density $\alpha_c(n) + \epsilon$ have no satisfying assignment with high probability. It is conjectured that $\alpha_c(n)$ is a constant (and that its value is about 4.27), but currently all that is known is that, for large n, $3.520 \leq \alpha_c(n) \leq 4.506$ [11,8,6].

In the range of densities for which the formula is satisfiable with high probability, the interesting algorithmic question is whether we can find even one of the many satisfying assignments in polynomial time. The lower bound on $\alpha_c(n)$ is a result of the analysis of such a polynomial time algorithm [11,8]. This algorithm belongs to a family of algorithms known as "myopic" because they assign variables greedily, one by one, in an order that is based only on the number of positive and negative occurrences of each variable.

An apparently much more powerful algorithm is Survey Propagation [17,18]. In experiments on very large instances (say, with $n = 10^6$ variables), it finds solutions for formulas of densities only just below the conjectured threshold value $\alpha = 4.27$; however, a rigorous analysis of its performance is still far from our reach. Like the myopic algorithms, it also assigns variables, one by one, in a greedy manner, but its choices are based on more global information about the role of each variable in the formula. That information is provided by the fixed point of a sophisticated message passing dynamics between variables and clauses.

* Corresponding author.
E-mail addresses: elitza.maneva@gmail.com (E. Maneva), sinclair@cs.berkeley.edu (A. Sinclair).

0304-3975/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2008.06.053

The message passing procedure is based on an intriguing (informal) picture of the properties of the solution space of a random formula, which is derived from the 1-step Replica Symmetry Breaking ansatz of statistical physics [16]. A key postulate of this physical picture is that, for formulas of density higher than a certain value (estimated to be 3.92), the space of solutions is split into "clusters". Within the same cluster it is possible to reach any satisfying assignment from any other by flipping one variable at a time, while always keeping the formula satisfied. On the other hand, in order to get from a solution in one cluster to a solution in another, the values of a linear number of variables have to be flipped at the same time. Loosely speaking, the message passing procedure of Survey Propagation was proposed as a way to collect information about the *clusters* of assignments, rather than individual assignments.

The remarkable performance of Survey Propagation is a compelling reason to explore properties of the solution space of typical formulas, as a way to further our understanding of random 3-SAT, its hardness, and the satisfiability threshold, and also (looking much further ahead) in order to systematically design algorithms for more general problems with distributional inputs.

A detailed study of the Survey Propagation algorithm undertaken in [14,3] led to an interpretation of the message passing procedure as the more familiar Belief Propagation algorithm [19] applied to a particular probability distribution on *partial* assignments, i.e., assignments of values from the set $\{0, 1, *\}$. Here $*$ is to be interpreted as "unassigned", and a variable is allowed to be unassigned only if it is not forced to be assigned 0 or 1 in order to satisfy a clause.

Among these partial assignments, the set of "core" assignments plays a central role. These are partial assignments that are obtained from a satisfying assignment by successively replacing each unconstrained variable by $*$. (A variable is *unconstrained* if changing its value does not make any clause unsatisfied.) Any satisfying assignment has a unique corresponding core. Moreover, since all assignments in a cluster have the same core, cores can be viewed (informally) as "summaries" of clusters. Of course, this view is useful only if different clusters tend to have distinct non-trivial cores. (The trivial core assignment is the one without any assigned variables.) Recently, Achlioptas and Ricci-Tersenghi [1] showed that, in random k-SAT for $k \geq 9$, for some range of densities up to the satisfiability threshold, with high probability *every satisfying assignment has a non-trivial core assignment associated with it*. This implies that clusters have a large number of frozen variables; indeed, for large k the fraction of frozen variables comes arbitrarily close to 1. (The clustering picture has also been confirmed for 8-SAT by [15,1] using a different method that does not say anything about cores.)

The above results hold only for random k-SAT with $k \geq 8$ or 9. In this paper we investigate similar questions for the apparently harder case of random 3-SAT. Our main result is the following:

Theorem 1. *For random instances of* 3-SAT *with density greater than* 4.453*, with high probability there exist* no *non-trivial core assignments.*

This theorem requires some interpretation. Note first that the density 4.453 lies above the conjectured threshold value of 4.27, but below the current best known upper bound of 4.506. Thus we may deduce:

Corollary 2. *One of the following statements holds for random* 3-SAT*:*

- $\alpha_c(n) \leq 4.453$*; or*
- *there is a range of densities immediately below the satisfiability threshold for which, with high probability, there are no non-trivial core assignments.*

One interpretation of Theorem 1 is as evidence for an improved upper bound on the threshold of the form $\alpha_c(n) \leq 4.453$. On the other hand, if in fact $\alpha_c(n) > 4.453$ then the theorem establishes a range of densities immediately below the threshold for which with high probability the formula is satisfiable but has no non-trivial core assignments. This would represent a surprising difference between the properties of random 9-SAT and random 3-SAT. Interestingly, experiments with 3-SAT and solutions of large formulas found by Survey Propagation do not find cores (see [14]). It is quite conceivable that both of the statements in the above corollary hold.

We now briefly discuss our proof technique, which involves a novel application of the first moment method and is, we believe, perhaps more noteworthy than the result itself. A straightforward application of the first moment method to the set of core assignments only allows us to show that, with high probability, there are no cores of small size or of large size. To handle core assignments of intermediate size, it is further necessary to bound the probability that a partial assignment can be extended to a satisfying assignment. This probability is equivalent to the probability that a random formula with a given density of 2-clauses and 3-clauses is satisfiable. For this sub-problem, it is natural to use one of the methods that have been previously introduced for bounding the probability of satisfiability of random 3-SAT formulas [10,5,13,9,12,6]. However, the most powerful method, from [6], is very heavy numerically, and it does not seem possible to carry it out for the whole range of required densities; on the other hand, simpler methods such as those in [13] are apparently not powerful enough. Instead, we introduce a new method, which we now briefly outline (for a detailed description, see Section 3.2).

Traditionally, bounds on the probability of satisfiability, such as those mentioned above, are obtained by applying the first moment method to a random variable which counts the number of satisfying assignments of a particular kind. Indeed, much of the work on bounding the satisfiability threshold has been directed towards identifying a set of satisfying assignments that is a strict subset of the set of all satisfying assignments (so that its expected size is much smaller), but is always non-empty if the formula is satisfiable. The novelty in our approach lies in identifying a new random variable which depends not only on satisfying assignments but also on partial assignments. This random variable is at least 1 for every satisfiable formula, it exploits the clustering structure of the solution space, and most importantly, it is significantly easier to compute than many alternative approaches.

Finally we note that the proof of our theorem depends on a claim that a particular analytic function takes only negative values in a given range. We do not provide a complete proof of this numerical claim, but we outline the steps needed for completing the proof, and give very strong numerical evidence that all the corresponding statements hold. The difficulty with making the proof completely rigorous is that this would require too much computational effort, which we feel is not justified at this stage given that the bound we obtain is still some distance from the satisfiability threshold. The results presented here should rather be considered as a proof of concept.

The remainder of this paper is structured as follows. In Section 2 we give necessary background, and precise definitions of the various concepts used in the paper. Section 3 is devoted to the proof of our main result, Theorem 1. We conclude in Section 4 with some final remarks and suggestions for future work.

2. Technical definitions

Let $x_1, \dots, x_n$ be a set of n Boolean variables. A literal is either a variable or its negation. A 3-SAT formula is a formula in CNF, where each clause is a disjunction of 3 distinct literals (on different variables). For every clause c we will denote the set of variables that appear in the clause by $V(c)$. The distribution that we consider is the following: for a given density α we choose uniformly at random and with replacement $m = \lfloor \alpha n \rfloor$ clauses out of all possible clauses on 3 distinct variables. For a clause c and a variable $x_i \in V(c)$ we denote by $s_{c,i}$ and $u_{c,i}$ the value for variable x_i that is respectively satisfying and unsatisfying for clause c.

Suppose that the variables $\mathbf{x} = (x_1, \dots, x_n)$ are allowed to take values in $\{0, 1, *\}$, which we refer to as a *partial assignment*. A variable taking value $*$ (star) should be thought of as unassigned.

Definition 1. A partial assignment to $\mathbf{x}$ is *invalid* for a clause c if either (a) all variables are unsatisfying; or (b) all variables are unsatisfying except for one index $j \in V(c)$, for which $x_j = *$. Otherwise, the partial assignment is *valid* for clause c. We say that a partial assignment is valid for a formula (or just "valid") if it is valid for all of its clauses.

For a valid partial assignment, the subset of variables that are assigned either 0 or 1 values can be divided into *constrained* and *unconstrained* variables in the following way:

Definition 2. We say that a variable x_i is the *unique satisfying variable* for a clause c if it is assigned $s_{c,i}$, whereas all other variables in the clause (i.e., the variables $\{x_j : j \in V(c)\backslash\{i\}\}$) are assigned $u_{c,j}$. A variable x_i is *constrained* by clause c if it is the unique satisfying variable for c.

A variable is *unconstrained* if it has 0 or 1 value, and is not constrained by any clause. Thus for any partial assignment $\sigma \in \{0, 1, *\}^n$, the variables are divided into stars, constrained and unconstrained variables. Let $S^*(\sigma)$ be the set of unassigned variables, and $n_*(\sigma)$ and $n_o(\sigma)$ denote respectively the number of stars and the number of unconstrained variables. We define the *weight* of a valid partial assignment to be

$$W(\sigma) := \rho^{n_*(\sigma)}(1-\rho)^{n_o(\sigma)}, \tag{1}$$

where ρ is a parameter in the interval $[0, 1]$. The weight of an invalid partial assignment is 0.

In [14] the Survey Propagation algorithm is interpreted as a special case of a larger family of Belief Propagation algorithms applied to a family of distributions on valid partial assignments. This family of distributions, parameterized by ρ, is defined as $\Pr[\sigma] \propto W(\sigma)$. At one extreme, $\rho = 0$, this becomes just the uniform distribution over (full) satisfying assignments. The other extreme, $\rho = 1$, corresponds to Survey Propagation. The pure version of Survey Propagation corresponds to setting $\rho = 1$. Intermediate values of ρ interpolate between these extremes.

Next we define a natural partial order (represented by an acyclic directed graph) on valid partial assignments. The vertex set of the directed graph G consists of all valid partial assignments. The edge set is defined in the following way: for a given pair of valid partial assignments σ and τ, the graph includes a directed edge from σ to τ if there exists an index $i \in \{1, \dots, n\}$ such that (i) $\sigma_j = \tau_j$ for all $j \neq i$; and (ii) $\tau_i = *$ and $\sigma_i \neq *$.

Valid partial assignments can be separated into levels based on their number of star variables, i.e., the assignment σ is in level $n_*(\sigma)$. Thus every edge goes from an assignment in level $l-1$ to one in level l, where $1 \leq l \leq n$. G is acyclic and we write $\tau < \sigma$ if there is a directed path in G from σ to τ. In this case we will also say that the assignment τ is *consistent* with the assignment σ. The outgoing edges of any valid partial assignment σ correspond to its unconstrained variables, and therefore its outdegree is equal to $n_o(\sigma)$. The minimal assignments in this ordering are the assignments without unconstrained variables, i.e., the positive weight assignments for $\rho = 1$.

Definition 3. A *core* of a satisfying assignment σ is a minimal assignment τ such that $\tau < \sigma$.

The following proposition about cores is proved in [14].

Proposition 3 (*[14]*). *Any satisfying assignment σ has a unique core. Furthermore, if satisfying assignments $\sigma^1, \sigma^2 \in \{0, 1\}^n$ belong to the same cluster of solutions, then they have the same core.*

In the above proposition, a cluster is simply a connected component of the graph on solutions, in which two solutions are connected by an edge if and only if they are at Hamming distance 1.

Definition 4. A *cover* is a valid partial assignment that contains no unconstrained variables.

In particular, the core of any satisfying assignment is a cover. On the other hand not all cover assignments are cores (because they may not be extendable to satisfying assignments). We say that a cover assignment τ is *non-trivial* if $n_*(\tau) < n$, so that it has at least one assigned variable.

The proof of Theorem 1 uses a surprising property of the weights (1), which was observed in [14] but was not utilized there. (A more general statement and a connection to a combinatorial object known as "convex geometry" was developed in [2].) Specifically, the total weight of partial assignments consistent with a given satisfying assignment is exactly 1. This fact implies that the probability of satisfiability is at most the expected total weight of partial assignments.

Theorem 4 ([14]). *For every $\rho \in [0, 1]$, $\sum_{\tau \le \sigma} W(\tau) = \rho^{n_*(\sigma)}$ for any valid partial assignment $\sigma \in \{0, 1, *\}^n$. In particular, if σ is a satisfying assignment then $\sum_{\tau \le \sigma} W(\tau) = 1$.*

In our proof, we will apply Theorem 4 with various values for $\rho \in (0.8, 1)$. In each application, we will chose the value of ρ to get the best bound possible for the probability that a formula chosen from some distribution has a satisfying assignment. In particular, if a satisfying assignment exists, the total weight of valid partial assignments is at least 1. Therefore, the probability of satisfiability is at most as large as the expected value of the total weight of partial assignments. Applying this idea directly to the original distribution on 3-Sat formulas leads to an upper bound on the threshold α_c that is weaker than the currently best known bound of 4.506. However, the derivation is simpler than other approaches, which makes it possible to apply the same method to bound the probability that a fixed valid partial assignment can be extended to a satisfying assignment. (This amounts to applying the method to random formulas coming from a variety of distributions on formulas with both 2-clauses and 3-clauses.) This allows us to estimate the probability of the existence of a non-trivial core, and thus to prove the theorem.

3. Proof of the main theorem

This section contains a proof of Theorem 1. We begin with an overview of the entire proof, in the course of which we will state various technical lemmas; these lemmas will be proved in the three subsections that follow.

Note first that for $\alpha \ge 4.506$ the statement of the main theorem follows from the fact that random 3-Sat formulas of density at least 4.506 are known to be unsatisfiable with high probability [6]. Hence from now on we focus on the case that $\alpha \in [4.453, 4.506]$. Our goal is to prove that, for densities above 4.453, with high probability there are no non-trivial covers that can be extended to satisfying assignments, i.e., there are no non-trivial cores.

To this end, define the *size* of a cover (or of a core) to be the number of variables assigned value 0 or 1. The following lemma, proved in [14], establishes that with high probability all non-trivial covers (and consequently cores) are of linear size.

Lemma 5 ([14]). *For a random* 3-Sat *formula of density α, with high probability, there are no non-trivial covers of size strictly less than $\frac{1}{\alpha e^2} n$.*

This lemma implies that it is sufficient to consider core assignments of size an for $a \in [1/(\alpha e^2), 1]$. Let X_a denote the number of cover assignments of size $\lfloor an \rfloor$, and Y_a denote the number of core assignments of size $\lfloor an \rfloor$.

In addition to the density α, the size measure a, and the weight parameter $\rho \in [0, 1]$ from Theorem 4, we will use two further parameters, d and b, whose precise definitions will be given in the proofs in the following subsections. Roughly speaking, d is the fraction of clauses that are constraining with respect to a given partial assignment, and b is the fraction of constrained variables with respect to a given partial assignment. The ranges of these parameters are $d \in (a/\alpha, 1]$ and $b \in [0, 1 - a]$. We will also make use of a derived parameter r which is defined by d, a and α. It appears in connection to the event that the constraining clauses succeed in constraining all constrained variables. Specifically, r is the value satisfying the equation $d = \frac{ar}{\alpha} \ln \frac{r}{r-1}$. (Note that such a value r always exists and is unique as the right-hand side is monotonic in r.) In fact, because of the form of this expression, it will be more convenient to think of d as being determined by r, a, and α. Whenever we take the supremum over one of these parameters, we always mean the supremum over its allowed range.

We now define two functions that play a central role in our analysis:

$$\begin{aligned} f(\alpha, a, r) \quad &:= a \ln(2) + H(a) + \alpha H(d) + \alpha d \ln\left(\frac{3a^3}{8}\right) + \alpha(1-d)\ln\left(1 - \frac{a^2(3-a)}{4}\right) \\ &\quad + \alpha d \ln(r/e) - a\ln(r-1); \\ h(\alpha, a, r, \rho, b) &:= b\ln(2) + (1-a-b)\ln(\rho) + (1-a)H(b/(1-a)) \\ &\quad - \alpha(1-d)b\,\frac{b(6-5b-15a) + 12(1-a)a}{2(4-a^2(3-a))} + b\ln\left(1 - \rho e^{-3\alpha(1-d)\frac{b(b+2a)}{2(4-a^2(3-a))}}\right). \end{aligned}$$

Here H denotes the entropy function $H(x) = -x\ln(x) - (1-x)\ln(1-x)$.

The first ingredient in the proof of Theorem 1 is the following lemma, whose proof is presented in Section 3.1.

Lemma 6. *For a random* 3-Sat *formula of density $\alpha \ge 1$, and for every $a \in [0, 1]$,*

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(E[X_a]\right) \le \sup_r f(\alpha, a, r).$$

A simple application of Markov's inequality immediately yields:

Corollary 7. *If $\alpha \geq 1$ and $a \in [0, 1]$ are such that, for every $r > 1$ with $d = \frac{ar}{\alpha} \ln \frac{r}{r-1} \leq 1$, it holds that $f(\alpha, a, r) < 0$, then with high probability random 3-*Sat *formulas of density α do not have covers of size an.*

The second ingredient in the proof of Theorem 1 is the following lemma, which is proved in Section 3.2.

Lemma 8. *For a random* 3-Sat *formula of density $\alpha \geq 1$, and for every $a \in [0, 1]$ and $\rho \in [0, 1)$,*

$$\lim_{n\to\infty} \frac{1}{n} \ln (E[Y_a]) \leq \sup_r (f(\alpha, a, r) + \min\{0, \sup_b h(\alpha, a, r, \rho, b)\}).$$

Remark. This lemma can be strengthened by allowing ρ to depend on r; however, we will not need to use this stronger version.

Since every core is a cover, we know that $Y_a \leq X_a$. Hence (in light of Lemma 6) Lemma 8 is interesting only when $\sup_b h(\alpha, a, r, \rho, b)$ is negative. In fact, we shall prove Lemma 8 by showing that this supremum bounds the logarithm of $1/n$ times the probability that a particular cover of size a can be extended to a satisfying assignment. (For a precise statement, see Lemma 11 in Section 3.2.) An immediate corollary of Lemma 8 is the following.

Corollary 9. *If $\alpha \geq 1$, $a \in [0, 1]$ and there exists $\rho \in [0, 1)$ such that for every $r > 1$ with $d = \frac{ar}{\alpha} \ln \frac{r}{r-1} \leq 1$, and for every $b \in [0, 1-a]$, it holds that $f(\alpha, a, r) + h(\alpha, a, r, \rho, b) < 0$, then with high probability random* 3-Sat *formulas of density α do not have cores of size an.*

The final ingredient in the proof of Theorem 1 is the following numerical claim, whose proof we discuss in Section 3.3.

Claim 10. *For every $\alpha \in [4.453, 4.506]$, $a \in [1/(4.506e^2), 1]$, and $r > 1$ with $d = \frac{ar}{\alpha} \ln \frac{r}{r-1} \leq 1$, it holds that either $f(\alpha, a, r) < 0$ or for every $b \in [0, 1-a]$, $f(\alpha, a, r) + h(\alpha, a, r, 0.4a + 0.7, b) < 0$.*

Remark. In Claim 10 we have used the weight $\rho \equiv \rho(a) = 0.4a + 0.7$. We arrived at this particular choice of ρ by first optimizing numerically for ρ at 100 values for $a \in [1/4.506, 1]$, and then fitting a simple function to the values for ρ that were found. Since Lemma 8 holds for every value of ρ, we may choose any convenient function. Using a simple analytic function guarantees that h is also analytic, which makes the numerical analysis of h easier.

Finally, combining Claim 10 with Corollary 9 completes the proof of Theorem 1.

3.1. The expected number of covers: Proof of Lemma 6

Let $s = \lfloor an \rfloor$. Then we have

$$\begin{aligned} E[X_a] &= E\left[\sum_{\sigma\in\{0,1,*\}^n} \mathrm{Ind}[\sigma \text{ is valid } \cap\ (n_*(\sigma) = n-s) \cap\ (n_o(\sigma) = 0)]\right] \\ &= \sum_{\sigma\in\{0,1,*\}^n : n_*(\sigma)=n-s} \Pr[\sigma \text{ is valid } \cap\ (n_o(\sigma) = 0)] \\ &= \binom{n}{s} 2^s \Pr[\sigma = (0^s *^{n-s}) \text{ is valid and } x_1, \ldots, x_s \text{ are constrained}]. \end{aligned}$$

We denote by P the probability that $\sigma = (0^s *^{n-s})$ is valid and all of its assigned variables are constrained. P is equivalent to the probability of the following event in an experiment with $m = \alpha n$ balls thrown uniformly and independently at random into $2^3\binom{n}{3}$ bins. There are 3 kinds of bins:

1. Bins of type 1 should be empty. These correspond to clauses that are not allowed:
   - $(x_{i_1} \vee x_{i_2} \vee x_{i_3})$, with $i_1, i_2, i_3 \in \{1, 2, \ldots, s\}$;
   - $(x_{i_1} \vee x_{i_2} \vee \bar{x}_j)$, with $i_1, i_2 \in \{1, 2, \ldots, s\}$ and $j > s$;
   - $(x_{i_1} \vee x_{i_2} \vee x_j)$, with $i_1, i_2 \in \{1, 2, \ldots, s\}$ and $j > s$.

   The total number of these is
   $$\binom{s}{3} + 2(n-s)\binom{s}{2} = n^3 \left(\frac{a^3}{6} + a^2(1-a)\right) + o(n^3).$$
2. Bins of type 2 correspond to constraining clauses: $(x_{i_1} \vee x_{i_2} \vee \bar{x}_t)$, with $i_1, i_2, t \in \{1, 2, \ldots, s\}$. For each variable x_t there are $\binom{s-1}{2} = n^2 \frac{a^2}{2} + o(n^2)$ clauses that could constrain it and at least one has to be included. Equivalently, there has to be at least one ball in one of those bins for every x_t with $t \in \{1, 2, \ldots, s\}$. The total number of these clauses is $s\binom{s-1}{2} = n^3 \frac{a^3}{2} + o(n^3)$.

3. There are no constraints for the remaining bins, of type 3. Their total number is

$$2^3\binom{n}{3} - n^3\left(\frac{a^3}{6} + a^2(1-a)\right) - n^3\,\frac{a^3}{2} + o(n^3) = \frac{n^3}{3}\,(4 - a^2(3-a)) + o(n^3).$$

Suppose $m' = dm$ of the clauses we choose are of type 2, and the remaining $m - m'$ are of type 3. The probability of this event is

$$p_{m'} = \binom{m}{m'}\left(\frac{3\,a^3}{8} + o(1)\right)^{m'}\left(1 - \frac{a^2\,(3-a)}{4} + o(1)\right)^{m-m'}.$$

The probability that the m' clauses of type 2 are such that there is at least one of each kind is the same as the coupon collectors probability of success, with $s = \lfloor an \rfloor$ different coupons, and $m' = (d\alpha)n$ trials. We will use the following general fact from [4], which was previously used in a very similar context in [12]: Let $q(cN, N)$ denote the probability of collecting N coupons within cN trials. If $c < 1$, $q(cN, N) = 0$. Otherwise, as N goes to infinity $q(cN, N)$ grows like $g(c)^N$, where $g(c) = \left(\frac{r_0}{e}\right)^c \frac{1}{r_0 - 1}$, and r_0 is the solution of $r \ln\left(\frac{r}{r-1}\right) = c$. More precisely, $\lim_{N\to\infty} \frac{1}{N} \ln(q(cN, N)) = \ln(g(c))$. We have

$$\begin{aligned}
P &= \sum_{m'=0}^{m}\binom{m}{m'}\left(\frac{3\,a^3}{8} + o(1)\right)^{m'}\left(1 - \frac{a^2\,(6-2a)}{8} + o(1)\right)^{m-m'} q(m', s)\\
&\le m\,\max_{m'}\left\{\binom{m}{m'}\left(\frac{3\,a^3}{8} + o(1)\right)^{m'}\left(1 - \frac{a^2\,(3-a)}{4} + o(1)\right)^{m-m'} q(m', s)\right\}.
\end{aligned}$$

Finally,

$$E[X_a] \le \binom{n}{s} 2^s m \max_{m'}\left\{\binom{m}{m'}\left(\frac{3\,a^3}{8} + o(1)\right)^{m'}\left(1 - \frac{a^2\,(3-a)}{4} + o(1)\right)^{m-m'} q(m', s)\right\},$$

and hence

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\ln(E[X_a]) &\le \ln\left(\frac{2^a}{a^a(1-a)^{1-a}}\right) + \sup_d\left\{\alpha \ln\left(\frac{\left(\frac{3\,a^3}{8}\right)^d\left(1 - \frac{a^2\,(3-a)}{4}\right)^{1-d}}{d^d\,(1-d)^{1-d}}\right) + a\ln\left(g\left(\frac{d\alpha}{a}\right)\right)\right\}\\
&= \sup_d f(\alpha, a, r) = \sup_r f(\alpha, a, r).
\end{aligned}$$

This completes the proof of Lemma 6.

3.2. *The probability that a cover assignment is a core: Proof of Lemma 8*

For a partial assignment σ, it will be convenient to denote by $\varphi_\sigma(x_{S^*(\sigma)})$ the formula that is obtained by substituting the variables that have 0/1 assignments in σ, i.e., removing from φ clauses that are satisfied by at least one of the assigned variables and removing all remaining appearances of assigned variables. This is a formula on $n^*(\sigma)$ variables. Notice that if σ is a valid assignment for φ, then the formula φ_σ contains no empty clauses and no unit clauses. Furthermore all clauses of type 2 (from the previous section) are removed, because they are satisfied by their constrained variables. Among the clauses of type 3, there are clauses that are removed, there are clauses that become two-variable clauses, and there are clauses that remain untouched. Since there is no simple way to describe the resulting distribution on formulas, we will keep referring to the set of all clauses of type 3, even the ones that are removed in φ_σ. When we condition on the fact that σ is a cover and m' of the clauses are of type 2, as in the previous section, we know that the set of clauses we are interested in are distributed exactly as a uniform set of $(m - m')$ clauses of type 3. Thus we can express the expected number of cores as

$$\begin{aligned}
E[Y_a] &= \sum_{\sigma\in\{0,1,*\}^n : n_*(\sigma) = n-s} \Pr[\sigma \text{ is a cover}] \times \Pr[\sigma \text{ is a core} \mid \sigma \text{ is a cover}]\\
&= \binom{n}{s} 2^s\ \Pr[\sigma = (0^s *^{n-s}) \text{ is a cover}] \times \Pr[\sigma = (0^s *^{n-s}) \text{ is a core} \mid \sigma \text{ is a cover}]\\
&= \binom{n}{s} 2^s \times \sum_{m'=s}^{m} p_{m'}\, q(m', s)\\
&\quad \times \Pr[\varphi_\sigma(x_{s+1}, \ldots, x_n) \text{ is satisfiable} \mid \sigma = (0^s *^{n-s}),\ m - m' \text{ clauses are of type 3}].
\end{aligned}$$

We will bound this probability using the Poisson approximation, which is a standard technique in this area. There are two related random models. In the first model, which we call the *exact model*, $(m - m')$ clauses are chosen uniformly at random with replacement from all $M = n^3(4-a^2(3-a))/3 + o(n^3)$ clauses of type 3. In the second model, which we call the *Poisson model*, each of the M clauses is included in the formula with probability $p = (m-m')/M - 1/(n^2\sqrt{\log n}) = \frac{3\alpha(1-d)}{n^2(4-a^2(3-a))} + o\left(\frac{1}{n^2}\right)$. The expected number of clauses in both models is the same up to a term $\delta = M/(n^2\sqrt{\log n}) = \Theta(n/\sqrt{\log n})$, which is sub-linear in n.

The Poisson model has been studied before in the context of random 3-Sat. It can be shown, as the example below demonstrates, that whenever a property holds with high probability in the exact model, it also holds with high probability in the corresponding Poisson model. Applying first moment techniques to the Poisson model is usually easier, because the clauses are independently chosen; however the bounds obtained are usually weaker.

Next, we relate the probability that φ_σ is satisfiable under the two models. Let $\Pr_p$ denote probability in the Poisson model, and $\Pr_e$ denote probability in the exact model. Let the random variable J denote the number of *distinct* clauses included in the formula. Then

$$\begin{aligned}\Pr_p[\varphi_\sigma \text{ is satisfiable}] &= \sum_{i=0}^{M} \Pr_p[J=i] \times \Pr[\varphi_\sigma \text{ is satisfiable} \mid J=i] \\ &\geq \sum_{i=0}^{m-m'-\delta} \Pr_p[J=i] \times \Pr[\varphi_\sigma \text{ is satisfiable} \mid J=i].\end{aligned}$$

Since this conditional probability decreases as i increases, and $\Pr_p[J \leq E_p[J]] \geq 1/2$, where $E_p[J] = m - m' - \delta$, we have

$$\begin{aligned}\Pr_p[\varphi_\sigma \text{ is satisfiable}] &\geq \Pr[\varphi_\sigma \text{ is satisfiable} \mid J = m-m'-\delta] \times \sum_{i=0}^{m-m'-\delta} \Pr[J=i] \\ &\geq \Pr[\varphi_\sigma \text{ is satisfiable} \mid J = m-m'-\delta] \times \frac{1}{2}.\end{aligned}$$

On the other hand, for the exact model

$$\begin{aligned}\Pr_e[\varphi_\sigma \text{ is satisfiable}] &= \sum_{i=0}^{m-m'} \Pr_e[J=i] \times \Pr[\varphi_\sigma \text{ is satisfiable} \mid J=i] \\ &\leq \left(\sum_{i=0}^{m-m'-\delta} \Pr_e[J=i]\right) + \Pr[\varphi_\sigma \text{ is satisfiable} \mid J = m-m'-\delta] \\ &\leq \Pr_e[\text{at least } \delta \text{ clauses repeated}] + 2\Pr_p[\varphi_\sigma \text{ is satisfiable}] \\ &\leq \binom{m-m'}{\delta}\left(\frac{m-m'}{M}\right)^{\delta} + 2\Pr_p[\varphi_\sigma \text{ is satisfiable}] \\ &\leq \left(\frac{(m-m')^2}{M}\right)^{\delta} + 2\Pr_p[\varphi_\sigma \text{ is satisfiable}] \\ &= \theta(2^{-n\sqrt{\log n}}) + 2\Pr_p[\varphi_\sigma \text{ is satisfiable}].\end{aligned}$$

Thus if the probability of satisfiability in the Poisson model is bounded above by c^{-n}, for some constant c, then $\lim_{n\to\infty} \frac{1}{n} \ln(\Pr_e[\varphi_\sigma \text{ is satisfiable}]) \leq c$. Therefore, it suffices to get a bound for the Poisson model. Any of the first moment techniques for bounding the satisfiability threshold of 3-Sat can be adapted to bound the probability that φ_σ is satisfiable. Of the ones that are technically applicable (i.e., result in an explicit analytic expression for every setting of the parameters s and m'), we obtain the strongest result using the novel approach of applying the first moment method to the distribution on partial assignments defined by the weights in Eq. (1).

Lemma 11. *In the Poisson model with parameters n, $m = \alpha n$, $m' = dm$, $s = \lfloor an \rfloor$ and for $\sigma = (0^s *^{n-s})$, and r such that $d = \frac{ar}{\alpha} \ln \frac{r}{r-1}$,*

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(\Pr_p[\,\varphi_\sigma(x_{s+1},\ldots,x_n) \text{ is satisfiable}]\right) \leq \inf_{\rho\in[0,1)} \sup_{b\in[0,1-a]} h(\alpha, a, r, \rho, b).$$

Proof. We will apply Theorem 4 to the formula $\varphi_\sigma(x_{s+1},\ldots,x_n)$. The theorem implies that if φ_σ has a satisfying assignment then $\sum_{\tau\in V} W(\tau) \geq 1$ where $W(\tau) = \rho^{n_*(\tau)}(1-\rho)^{n_o(\tau)}$ and the sum is over the set V of all partial assignments $\tau \in \{0,1,*\}^{n-s}$ that are valid for φ_σ. Thus the probability of satisfiability is bounded from above by the expected value of $\sum_{\tau\in V} W(\tau)$. This holds for any value of $\rho \in [0,1]$. We bound this expectation.

For any $t \in \{0, 1, \ldots, n-s\}$ let Z_t denote the sum of the weights of valid assignments τ for $\varphi_\sigma(x_{s+1}, \ldots, x_n)$ such that $n_*(\tau) = n - s - t$. Then

$$E[Z] = \sum_{t=0}^{n-s} E[Z_t] \le n \max_{t \in \{0,1,\ldots,n-s\}} E[Z_t].$$

$$\begin{aligned}
E[Z_t] &= \sum_{u=0}^{t} \rho^{n-s-t}(1-\rho)^u \sum_{\tau \in \{0,1,*\}^{n-s}} \Pr_p[\tau \text{ is valid } \cap (n_*(\tau) = n-s-t) \cap (n_o(\tau) = u)] \\
&= \rho^{n-s-t} \sum_{u=0}^{t}(1-\rho)^u \sum_{\tau \in \{0,1,*\}^{n-s} \,:\, n_*(\tau)=n-s-t} \Pr_p[\tau \text{ is valid} \cap (n_o(\tau) = u)] \\
&= \rho^{n-s-t} \sum_{u=0}^{t}(1-\rho)^u \binom{n-s}{t}\binom{t}{u} 2^t \Pr_p[\tau = (0^t *^{n-s-t}) \text{ is valid and } x_{s+1}, \ldots, x_{s+u} \text{ are unconstrained}]
\end{aligned}$$

Next we derive the probability that the assignment $(x_{s+1}, x_{s+2}, \ldots, x_n) = (0^t *^{n-s-t})$ is valid and the first u variables are unconstrained. Recall that φ_σ is obtained from φ by setting its first s variables according to σ. Only clauses of type 3 influence φ_σ and according to the Poisson model, each of them is included independently with probability p. The probability that the assignment $(x_{s+1}, x_{s+2}, \ldots, x_n) = (0^t *^{n-s-t})$ is valid and the first u variables are unconstrained is equivalent to the probability that among the included clauses of type 3, there are no clauses of the following kinds:

- $(x_{i_1} \vee x_{i_2} \vee x_{i_3})$, with $i_1 \in \{1, 2, \ldots, s+t\}$, $i_2, i_3 \in \{s+1, s+2, \ldots, s+t\}$,
- $(x_{i_1} \vee x_{i_2} \vee \bar{x}_j)$, with $i_1 \in \{1, 2, \ldots, s+t\}$, $i_2 \in \{s+1, s+2, \ldots, s+t\}$ and $j > s+t$,
- $(x_{i_1} \vee x_{i_2} \vee x_j)$, with $i_1 \in \{1, 2, \ldots, s+t\}$, $i_2 \in \{s+1, s+2, \ldots, s+t\}$ and $j > s+t$,
- $(x_{i_1} \vee x_{i_2} \vee \bar{x}_j)$, with $i_1 \in \{1, 2, \ldots, s+t\}$, $i_2 \in \{s+1, s+2, \ldots, s+t\}$ and $j \in \{s+1, \ldots, s+u\}$,

and for every $j \in \{s+u+1, \ldots, s+t\}$, the formula contains a clause $(x_{i_1} \vee x_{i_2} \vee \bar{x}_j)$, with $i_1 \in \{1, 2, \ldots, s+t\}$, $i_2 \in \{s+1, \ldots, s+t\}$.

In the Poisson model all clauses are independent, so it is easy to put these events together to obtain:

$$\begin{aligned}
Q &= \Pr_p[\tau = (0^t *^{n-s-t}) \text{ is valid and } x_{s+1}, \ldots, x_{s+u} \text{ are unconstrained}] \\
&= (1-p)^{\left(\binom{t}{3}+s\binom{t}{2}\right) + 2(n-s-t)\left(\binom{t}{2}+st\right) + u\left(\binom{t}{2}+st\right)} \times \left(1-(1-p)^{\binom{t}{2}+st}\right)^{t-u}.
\end{aligned}$$

The expression for the expectation can be simplified:

$$\begin{aligned}
E[Z_t] &= \rho^{n-s-t}\binom{n-s}{t} 2^t \sum_{u=0}^{t}(1-\rho)^u \binom{t}{u} Q \\
&= \rho^{n-s-t}\binom{n-s}{t} 2^t (1-p)^{\binom{t}{3} + s\binom{t}{2} + 2(n-s-t)\left(\binom{t}{2}+st\right)} \\
&\quad \times \sum_{u=0}^{t}\binom{t}{u}(1-\rho)^u (1-p)^{u\left(\binom{t}{2}+st\right)} \times \left(1-(1-p)^{\binom{t}{2}+st}\right)^{t-u} \\
&= \rho^{n-s-t}\binom{n-s}{t} 2^t (1-p)^{\binom{t}{3} + s\binom{t}{2} + 2(n-s-t)\left(\binom{t}{2}+st\right)} \\
&\quad \times \left((1-\rho)(1-p)^{\binom{t}{2}+st} + \left(1-(1-p)^{\binom{t}{2}+st}\right)\right)^t \\
&= \rho^{n-s-t}\binom{n-s}{t} 2^t (1-p)^{\binom{t}{2}(6n-5t-3s-2)/3+2(n-s-t)st}\left(1-\rho(1-p)^{\binom{t}{2}+st}\right)^t.
\end{aligned}$$

Let $t = bn$, and recall that $s = \lfloor an \rfloor$, $p = \frac{3\alpha(1-d)}{n^2(4-a^2(3-a))} + o\left(\frac{1}{n^2}\right)$. Then

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\ln(E[Z_t]) &= \ln\left(\frac{(1-a)^{1-a}\, 2^b\, \rho^{1-a-b}}{b^b\,(1-a-b)^{1-a-b}}\right) - \frac{\alpha\,(1-d)\,b\,(b(6-5b-3a)+12(1-a-b)a)}{2(4-a^2(3-a))} \\
&\quad + b\ln\left(1-\rho e^{-\frac{3\alpha(1-d)b(b+2a)}{2(4-a^2(3-a))}}\right) \\
&= h(\alpha, a, r, \rho, b).
\end{aligned}$$

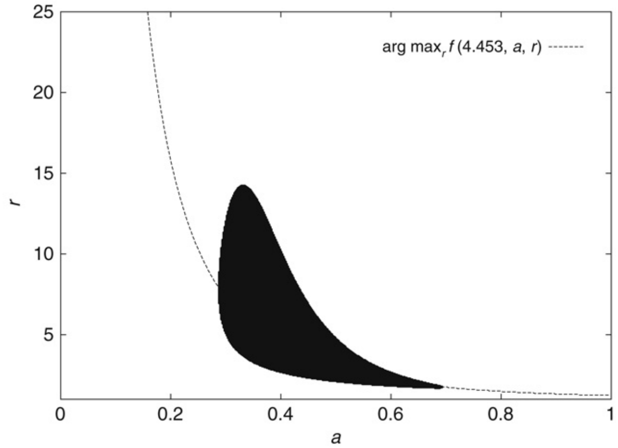

Fig. 1. The pairs (a, r) such that $f(4.453, a, r) > -0.0001$, and the value of r that maximizes $f(4.453, a, r)$ for every a.

Finally,

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(\Pr_p[\,\varphi_\sigma(x_{s+1}, \ldots, x_n) \text{ is satisfiable}]\right) \le \lim_{n\to\infty} \frac{1}{n} \ln(E[Z]) \le \sup_{b\in[0,1-a]} h(\alpha, a, r, \rho, b),$$

which is the statement of Lemma 11. ∎

Substituting the bound of Lemma 11 into the expression for the expectation yields

$$\lim_{n\to\infty} \frac{1}{n} \ln(E[Y_a]) \le \sup_r (f(\alpha, a, r) + \min\{0, \sup_b h(\alpha, a, r, \rho(a, r), b)\}),$$

completing the proof of Lemma 8.

3.3. Numerical analysis: Steps to the proof of Claim 10

What remains is to verify the numerical claim that for every $\alpha \in [4.453, 4.506]$, $a \in [1/(4.506e^2), 1]$, and $r > 1$ with $d = \frac{ar}{\alpha} \ln \frac{r}{r-1} \le 1$, it holds that either $f(\alpha, a, r) < 0$, or for every $b \in [0, 1-a]$, $f(\alpha, a, r) + h(\alpha, a, r, 0.4a + 0.7, b) < 0$. We outline the steps towards a proof.

There are two stages. First, we identify the (a, r) pairs such that for every $\alpha \in [4.453, 4.506]$ it holds that $f(\alpha, a, r) < 0$. Second, for the remaining range of values of (a, r) we show that for every $b \in [0, 1-a]$, $f(\alpha, a, r) + h(\alpha, a, r, 0.4a+0.7, b) < 0$.

The derivative of f with respect to r is

$$\frac{\partial f}{\partial r} = a\left(\ln\left(1 + \frac{1}{r-1}\right) - \frac{1}{r-1}\right) \times \ln\left(\frac{3\left(\alpha - ar\ln\left(\frac{r}{r-1}\right)\right)}{2\left(\frac{4}{a^2} - 3 + a\right)\ln\left(\frac{r}{r-1}\right)}\right).$$

Since the first factor is always negative, the derivative is 0 when

$$3\alpha = \ln\left(\frac{r}{r-1}\right)\left(\frac{8}{a^2} - 6 + 2a + 3ar\right). \tag{2}$$

This equation has a unique root for every $a \in [1/(4.506e^2), 1]$, $\alpha \in [4.453, 4.506]$ (the derivative of the right-hand side is negative). Therefore we can conclude that for every a and α in the above range, f is at first increasing with r then decreasing. Its maximum is achieved at the root of Eq. (2). Fig. 1 shows the values for r where f is maximized, with respect to a, when $\alpha = 4.453$.

Since the root of (2) is monotone decreasing with respect to a and α, we can find values r_1 and r_2 such that for every a and α in the range, if $r < r_1$ then $\partial f/\partial r > 0$, and if $r > r_2$ then $\partial f/\partial r < 0$. Values satisfying this condition are $r_1 = 1.2$ and $r_2 = 670$. Thus if we show that $f(\alpha, a, r)$ is negative for $r = r_1$ and $r = r_2$ then it is negative for any r outside the range $(1.2, 670)$.

We will use the shorthand notation $q := r \ln \frac{r}{r-1}$. For $r \in (1.2, 670)$ we have $q \in (1.0007, 2.16)$. The whole range satisfies the condition that $d = \frac{ar}{\alpha} r \ln \frac{r}{r-1} \le 1$.

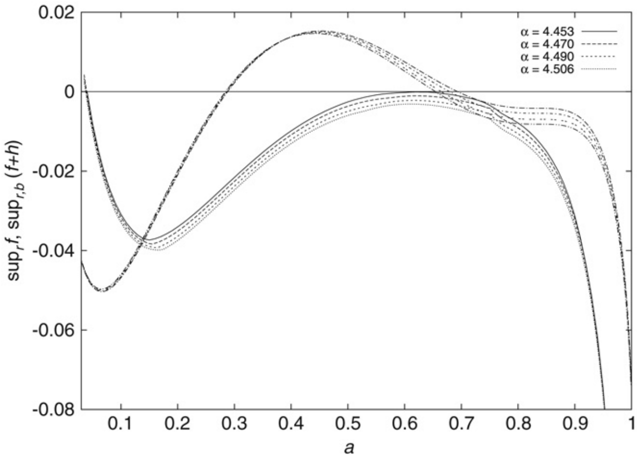

Fig. 2. The values of $\sup_r f(\alpha, a, r)$ and $\sup_{r,b}(f(\alpha, a, r) + h(\alpha, a, r, 0.4a + 0.7, b))$ for $\alpha = 4.453, 4.470, 4.490$, and 4.506.

Next we take care of the boundary region with respect to a. Notice that the derivative of f is negative for a large enough (close to 1), because of the entropy term in f. The derivative of f with respect to a is

$$\frac{\partial f}{\partial a} = \ln(2) - \ln\frac{a}{1-a} - q\ln\frac{3a^2(\alpha - aq)}{2q(4-a^2(3-a))} + 3q - \frac{3a(\alpha - aq)(2-a)}{4-a^2(3-a)} + q\ln(r) - \ln(r-1).$$

The following observations are helpful for bounding this derivative:

- $3q + q\ln(r) - \ln(r-1)$ is maximized in the interval $r \in [1.2, 670]$ at $r = 1.2$.
- $a^2(\alpha - aq)/(4 - a^2(3-a))$ is an increasing function of a.
- $a(\alpha - aq)(2-a)/(4 - a^2(3-a))$ is an increasing function of a.

For $a > 0.999$, using the above facts, the derivative is negative. Thus if we show that f is negative for $a = 0.999$, $r \in (1.2, 670)$, $\alpha \in [4.453, 4.506]$ then f is negative also for every $a > 0.999$.

We are left with the region $a \in [1/(4.506e^2), 0.999]$, $r \in (1.2, 670)$. In this region, all the derivatives of f can be bounded, and a sufficiently fine grid can be chosen over which to evaluate f in order to identify the grid sections where the function can take positive values. In particular, the derivative with respect to a is at most 28.2, and with respect to α it is at most 1. Furthermore, since we know that for every a and α, f is maximized as a function of r at the root of Eq. (2), one can find this maximum and the range of r where $f(\alpha, a, r)$ is positive using binary search. The points with $f(4.453, a, r) > -0.0001$ are depicted in Fig. 1.

In the second stage we need to analyze $f(\alpha, a, r) + h(\alpha, a, r, 0.4a + 0.7, b)$ for the remaining region of values for (a, r).

First, notice that we can take care of the boundary regions with respect to b by taking advantage of the entropy term, which has very large slope for b close to 0, and very steep negative slope for b close to $1 - a$. Specifically, the derivative with respect to b is

$$\begin{aligned}\frac{\partial h}{\partial b} &= \ln(2) - \ln(\rho) - \ln\left(\frac{b}{1-a-b}\right) - \alpha(1-d)\frac{12b - 15b^2 - 30ab + 12a - 12a^2}{2(4-a^2(3-a))} \\ &\quad + \ln(1-\rho e^{-A}) + 3\alpha(1-d)b\,\frac{\rho e^{-A}(a+b)}{(1-\rho e^{-A})(4-a^2(3-a))}\end{aligned}$$

where $A = \frac{3\alpha(1-d)b(b+2a)}{2(4-a^2(3-a))}$.

Using the bounds on all parameters: $\alpha \in [4.453, 4.506]$, $a \in (0.28, 0.7)$, $r \in (1.5, 14.3)$, and the ones that follow from those: $d \in (0.06, 0.26)$, $A \in [0, 1.59]$, $\rho \in (0.02, 0.21)$, we can conclude that

$$-3.04 - \ln\left(\frac{b}{1-a-b}\right) < \frac{\partial h}{\partial b} < 5.38 - \ln\left(\frac{b}{1-a-b}\right).$$

The lower bound can be made positive by setting $b/(1-a-b) \le 0.04$ and the upper bound can be made negative by setting $b/(1-a-b) \ge 220$. Therefore it suffices to show that h is negative for b in the range $[0.01, 0.996(1-a)]$.

Again, using the bounds for the parameters, all the derivatives of $f + h$ can be bounded, and a sufficiently fine grid can be chosen over which to evaluate $f + h$. We did not perform the evaluation on a grid that is as fine as is required for the rigorous proof because, based on the current bounds on the derivatives, we would need to evaluate the function at more than 10^{10} points. (As we discussed in the introduction, we feel that the computational effort is not yet justified.) For illustration, Fig. 2 shows the estimated values of $\sup_r f(\alpha, a, r)$ and of $\sup_{r,b}(f(\alpha, a, r) + h(\alpha, a, r, \rho(a), b))$ for $\alpha = 4.453, 4.470, 4.490$, and 4.506. The maximum is taken over evaluations at a grid of step size 0.001 for the parameters r and b. We estimated with better precision the location of the maximum of $f + h$ by evaluating the function at a fine grid in the region where the evaluations on the coarse grid give the largest values. The maximum found in this way is at $\alpha = 4.453$, $a = 0.62566$, $b = 0.03568$ and $r = 2.00134$ ($d = 0.19473$). The value of $f + h$ at this point is -0.000058.

4. Concluding remarks

As mentioned in the introduction, all of the known rigorous upper bounds for the satisfiability threshold of 3-SAT are based on the first moment method [10,5,13,9,12,6]; the corresponding upper bounds in this sequence of results are: 4.758, 4.64, 4.601, 4.596, 4.571, 4.506. The general method is to consider a random variable Z that is equal to the number of satisfying assignments of a particular kind. These satisfying assignments are such that at least one exists if the formula is satisfiable. Showing that above a certain density $E[Z] \to 0$ thus implies (by Markov's inequality) that $\Pr[Z = 0] \to 1$, and consequently the probability that the formula is unsatisfiable also goes to 1. For example, [5] takes Z to be the number of *negatively prime solutions*, i.e., solutions for which every variable assigned 1 is constrained.

Here, we have used the same idea but with Z taken to be the total weight of partial assignments under the weight function (1) inspired by Survey Propagation. Given the dramatic success of Survey Propagation for random 3-SAT, it seems plausible that this approach can potentially yield rather tight upper bounds on the threshold. We were able to achieve only partial progress in this direction, but we are quite hopeful that extensions of our approach could lead to further progress.

One natural extension would be to use a different weight ρ for each variable, depending for example on the number of positive and negative occurrences of the variable. The corresponding generalization of Theorem 4 is proved in [2]. It is quite possible that the value of $E[Z]$ in this case is significantly smaller.

Acknowledgments

We would like to thank the anonymous referees for their detailed and very useful suggestions.

Some of first author's research was done while the first author was a Ph.D. student at UC Berkeley, supported in part by NSF grant DMS-0528488. The second author was supported in part by NSF grants DMS-0528488 and CCR-0635153.

References

[1] D. Achlioptas, F. Ricci-Tersenghi, On the solution-space geometry of random constraint satisfaction problems, in: Proc. 38th ACM Symp. Theory of Computing, STOC, 2006, pp. 130–139.
[2] F. Ardila, E. Maneva, Pruning processes and a new characterization of convex geometries, Technical Report, 2007 (submitted for publication). Preprint at: arXiv:0706.3750.
[3] A. Braunstein, R. Zecchina, Survey propagation as local equilibrium equations, J. Stat. Mech. Theory Exp. (JSTAT) (2004) 06007.
[4] V. Chvatal, Almost all graphs with 1.44 edges are 3-colourable, Random Structures Algorithms 2 (1991) 11–28.
[5] O. Dubois, Y. Boufkhad, A general upper bound for the satisfiability threshold of random r-SAT formulae, J. Algorithms 24 (1997) 395–420.
[6] O. Dubois, Y. Boufkhad, J. Mandler, Typical random 3-SAT formulae and the satisfiability threshold, in: Proc. 11th ACM-SIAM Symp. on Discrete Algorithms, SODA, 2000, pp. 126–127. Extended version at: arXiv:cs/0211036.
[7] E. Friedgut, Neccesary and sufficient conditions for sharp threhsolds of graph properties and the k-problem, J. Amer. Math. Soc. 12 (1999) 1017–1054.
[8] M. Hajiaghayi, G. Sorkin, The satisfiability threshold of random 3-SAT is at least 3.52, Technical Report, 2003. Preprint at: arXiv:math/0310193.
[9] S. Janson, Y.C. Stamatiou, M. Vamvakari, Bounding the unsatisfiability threshold of random 3-SAT, Random Structures Algorithms 17 (2) (2000) 103–116.
[10] Anil Kamath, Rajeev Motwani, Krishna V. Palem, Paul G. Spirakis, Tail bounds for occupancy and the satisfiability threshold conjecture, Random Structures Algorithms 7 (1) (1995) 59–80.
[11] Alexis C. Kaporis, Lefteris M. Kirousis, Efthimios G. Lalas, The probabilistic analysis of a greedy satisfiability algorithm, Random Structures Algorithms 28 (4) (2006) 444–480.
[12] Alexis C. Kaporis, Lefteris M. Kirousis, Yannis C. Stamatiou, Malvina Vamvakari, Michele Zito, The unsatisfiability threshold revisited, Discrete Appl. Math. 155 (12) (2007) 1525–1538.
[13] L.M. Kirousis, E. Kranakis, D. Krizanc, Y.C. Stamatiou, Approximating the unsatisfiability threshold of random formulas, Random Structures Algorithms 12 (3) (1998) 253–269.
[14] E. Maneva, E. Mossel, M.J. Wainwright, A new perspective on survey propagation, J. ACM 54 (4) (2007) 2–41.
[15] M. Mézard, T. Mora, R. Zecchina, Clustering of solutions in the random satisfiability problem, Phys. Rev. Lett. 94 (2005) 197205.
[16] M. Mézard, G. Parisi, M.A. Virasoro, Spin Glass Theory and Beyond, World Scientific, Singapore, 1987.
[17] M. Mézard, G. Parisi, R. Zecchina, Analytic and algorithmic solution of random satisfiability problems, Science 297 (2002) 812. Scienceexpress published online June 27, 2002; 10.1126/science.1073287.
[18] M. Mézard, R. Zecchina, Random k-satisfiability: From an analytic solution to an efficient algorithm, Phys. Rev. E 66 (2002).
[19] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, Palo Alto, CA, 1988.