The unsatisfiability threshold revisited

Alexis C. Kaporis [a,1,2] Lefteris M. Kirousis [a,1,2]
Yannis C. Stamatiou [a,b,1,3] Malvina Vamvakari [a,b,3] Michele Zito [c,4]

[a] *University of Patras, Department of Computer Engineering and Informatics, University Campus, GR-265 04 Patras, Greece.*
e-mail: {kaporis,kirousis,stamatiu,mvamv}@ceid.upatras.gr.

[b] *Computer Technology Institute, 61 Riga Feraiou Str., GR-262 21 Patras, Greece.*

[c] *Department of Computer Science, University of Liverpool, U.K.*
e-mail: michele@csc.liv.ac.uk.

Abstract

The problem of determining the unsatisfiability threshold for random 3-SAT formulas consists in determining the clause to variable ratio that marks the experimentally observed abrupt change from almost surely satisfiable formulas to almost surely unsatisfiable. Up to now, there have been rigorously established increasingly better lower and upper bounds to the actual threshold value. In this paper, we consider the problem of bounding the threshold value from above using methods that, we believe, are of interest on their own right. More specifically, we show how the method of *local maximum satisfying truth assignments* can be combined with results for the *occupancy problem* in random allocation schemes of balls into bins in order to achieve an upper bound for the unsatisfiability threshold less than 4.571. Thus we improve over the best, with an available complete proof, previous upper bound, which was 4.596. In order to obtain this value, we also establish a bound on the q-binomial coefficients (a generalization of the binomial coefficients) which, we believe, is of independent interest.

Key words: phase transition, complexity, satisfiability, probabilistic analysis.

[1] Research partially supported by the University of Patras - Research Committee, project C. Carathéodory.
[2] Research partially supported by the Computer Technology Institute.
[3] Research supported by the Computer Technology Institute.
[4] Research supported by EPSRC grant GR/L/77089.

1 Introduction

Let ϕ be a random 3-SAT formula constructed by selecting uniformly and with replacement m clauses from the set of all possible clauses with three literals over n variables. It has been observed experimentally that as the numbers n, m of variables and clauses respectively tend to infinity, while the ratio m/n remains equal to a constant r, the random formulas exhibit a threshold behaviour: if $r > 4.17$ (approximately) then almost all random formulas are unsatisfiable while the opposite is true if $r < 4.17$. The constant r is called the *density* of the formula. On the theoretical side, Friedgut [8] has proved that there exists a sequence γ_n such that for any $\epsilon > 0$, if finally for all n, $r \leq \gamma_n - \epsilon$, then the probability of a random formula being satisfiable approaches 0, while if finally for all n, $r \geq \gamma_n + \epsilon$, then this probability approaches 1. It has not been rigorously proved that the sequence γ_n converges. Thus, finding the *exact* value of the threshold point or even proving that a threshold value exists is still a major problem in probability and complexity theory. Up to now, only upper and lower bounds have been rigorously established for the threshold value (rigorously, for the terms of the sequence γ_n). The best lower bound has been recently proved by D. Achlioptas and G. Sorkin in [1] and it is 3.26. Concerning the upper bound, Dubois et al. in [5] announced that they have obtained the value 4.506, but no complete proof has been made available yet.

In this paper, we address the upper bound question for the unsatisfiability threshold from a new perspective that combines the idea of *local maximum* satisfying truth assignments proposed by Kirousis et al. in [12] with the sharp probability estimates for the occupancy problem in random allocation schemes of balls into bins given by Kamath et al. in [11] (for an excellent introduction to the occupancy problem see [6,15]). With this approach we obtain an upper bound of 4.571, thus improving over the best, with an available complete proof, previous upper bound (4.596 given in [10]). The last author, following a similar approach, gives in his Ph.D. thesis [21] a bound of 4.5793 but without resorting to q-binomial coefficients (a generalization of the binomial coefficients). To obtain the value of 4.571, we also had to establish an upper bound to the q-binomial coefficients. Despite the extensive literature on q-binomial coefficients (see, e.g., [7,9,14]), no such bound was, to the best of our knowledge, known.

2 The method of local maxima

In this section, we will state briefly the methodology followed in [12] and state an inequality that bounds from above the probability that a random formula is satisfiable. This inequality will be the starting point of our considerations.

Let $\mathbf{S}$ be the class of all truth assignments to n variables and $\mathbf{A}_n$ the (random) class of truth assignments that satisfy a random formula ϕ. For a given $A \in \mathbf{S}$, a *single flip* sf is the change in A of exactly one specified FALSE value to TRUE and by A^{sf} we denote the truth assignment that results from this change. We define as $\mathbf{A}_n^1 \subseteq \mathbf{A}_n$ the random class of truth assignments with the following two properties:

- $A \models \phi$,
- for every single flip sf, it holds $A^{sf} \not\models \phi$.

A partial order can be defined on $\mathbf{S}$: a truth assignment A is smaller than a truth assignment A' iff there exists an i such that both A and A' assign the same value to all variables x_j, for all $j < i$ while A assigns FALSE to x_i and A' assigns TRUE to it. The random class $\mathbf{A}_n^1$ coincides with the set of satisfying truth assigments that are *local maxima* with respect to the partial order defined above among satisfying truth assignments that differ in one bit.

A more restricted random class of truth assignments results from $\mathbf{A}_n^1$ if we extend the scope of locality in obtaining a local maximum. A *double flip* is the change of exactly two specified variables x_i and x_j (with $i < j$) where x_i is changed from FALSE to TRUE and x_j from TRUE to FALSE. In analogy with single flips, by A^{df} we denote the truth assignment that results from A if we apply the double flip df. Let $\mathbf{A}_n^2$ be defined as the set truth of assignments A that have the following properties:

- $A \models \phi$,
- for all single flips sf, it holds $A^{sf} \not\models \phi$,
- for all double flips df, it holds $A^{df} \not\models \phi$.

Our starting point is the following inequality:

LEMMA 1 [12]

$$
\begin{aligned}
&\Pr[\phi \text{ is satisfiable}] \leq \mathrm{E}[|\mathbf{A}_n^2|] = \\
&\quad \Pr[A \models \phi] \textstyle\sum_{A \in \mathbf{S}} \Pr[A \in \mathbf{A}_n^1 \mid A \models \phi] \cdot \Pr[A \in \mathbf{A}_n^2 \mid A \in \mathbf{A}_n^1] = \\
&\quad (7/8)^{rn} \textstyle\sum_{A \in \mathbf{S}} \Pr[A \in \mathbf{A}_n^1 \mid A \models \phi] \cdot \Pr[A \in \mathbf{A}_n^2 \mid A \in \mathbf{A}_n^1]. \qquad (1)
\end{aligned}
$$

In order to find an upper bound for the unsatisfiability threshold, it suffices to find the smallest possible value for r for which the right-hand side of (1) tends to 0. We call $\Pr[A \in \mathbf{A}_n^1 \mid A \models \phi]$ the probability of single flips and $\Pr[A \in \mathbf{A}_n^2 \mid A \in \mathbf{A}_n^1]$ the probability of double flips. In the sections to follow, we will describe the sequence of steps that will lead us to the determination of upper bounds for these probabilities.

3 Calculation of the probability of single flips

In this section, we will find an exact asymptotic expression for $\Pr[A \in \mathbf{A}_n^1 \mid A \models \phi]$ using a sharp estimate for the occupancy problem provided in [11].

Given a truth assignment A, the set of *uniquely satisfied clauses* of A is the set of clauses that are satisfied by A and contain exactly one TRUE literal under A. The set of *blocking clauses* of A, $\mathbf{B}_A$, is the set of uniquely satisfied clauses whose unique TRUE literal is negated. For notational convenience, we will also consider a formula ϕ as a set of clauses. Thus, the expression $\phi \cap \mathbf{A}$, with $\mathbf{A}$ a set of clauses, has the meaning of set intersection with the additional requirement that a clause that appears in the intersection appears as many times as it appears in ϕ. Then the following is obviously true:

$$\Pr[A \in \mathbf{A}_n^1 \,|A \models \phi] = \\ \sum_{l=0}^{rn} \Pr[A \in \mathbf{A}_n^1 \,|A \models \phi, |\phi \cap \mathbf{B}_A| = l] \cdot \Pr[|\phi \cap \mathbf{B}_A| = l \,|A \models \phi]. \tag{2}$$

We will now compute the probabilities that appear in the right-hand side of (2). To compute $\Pr[A \in \mathbf{A}_n^1 \,|A \models \phi, |\phi \cap \mathbf{B}_A| = l]$, we fix a satisfying truth assignment $A \in \mathbf{A}_n$ with k FALSE values for some $k = 0, \ldots, n$ and condition on the event that the random formula ϕ contains l blocking clauses of A. We model the k FALSE variables of A as, initially empty, *bins* and the l blocking clauses that appear in ϕ as *distinct* balls. Each time a blocking clause appears in the process of forming the conjuncts of ϕ, its unique TRUE literal determines the bin into which the ball/clause is placed. For each variable that is FALSE under A, there correspond $\binom{n-1}{2}$ possible blocking clauses. Therefore, each time a blocking clause appears, it chooses its bin/variable uniformly. Thus, the above experiment may be viewed as a *random allocation* of l balls (blocking clauses) into k bins (variables that are FALSE under A). The event $\{A \in \mathbf{A}_n^1 \mid A \models \phi, |\phi \cap \mathbf{B}_A| = l\}$ occurs if, after the selection of the rn conjuncts of ϕ and under the condition that ϕ contains l blocking clauses of A, the following is true: for each variable x assigned FALSE by A, there exists in ϕ at least one blocking clause containing $\neg x$. But in the ball allocation scheme described above, this is equivalent to the event that *no bin remains empty*. Therefore, the computation of the first probability is reduced to the computation of the probability that after the random placement of l balls into k bins, no bin remains empty. This is an instance of the *occupancy problem*.

Given two functions F and G of n, we use the notation $F \asymp G$ to denote that fact that $\ln(F) \sim \ln(G)$. The following theorem by Kamath et al. ([11]) gives a sharp estimate for the probability that w bins remain empty:

Theorem 1 [11] *Let W denote the number of empty bins after the placement, uniformly and independently, of l balls into k bins, where both l and k are constant* *multiples of n. Let $c = \frac{l}{k} \geq 1$. If we denote by $H(l,k,w)$ the probability that $W = w$ and if, in addition, $|w - \mathrm{E}[W]| = \Omega(k)$ then*

$$H(l,k,w) \asymp e^{-k(\int_0^{1-\frac{w}{k}} \ln(\frac{s_w - x}{1-x})dx - c\ln(s_w))},$$

where s_w is the solution of the equation

$$w = k(1 - s_w(1 - e^{-c/s_w}). \tag{3}$$

Since we require to have at least one blocking clause for each FALSE value of A, or equivalently no bin to remain empty, for our purposes we set $w = 0$. Then we have:

$$|w - \mathrm{E}[W]| = |0 - k(1 - \frac{1}{k})^l| = \Omega(k).$$

Let now k, the number of FALSE values of A, be αn and l, the number of blocking clauses, be $\beta r n$, for some $\alpha, \beta \in (0,1)$ such that $\beta \geq \frac{\alpha}{r}$. From (3) we get:

$$0 = k[1 - s_0(1 - \exp^{-\frac{c}{s_0}})] \Leftrightarrow \ln(s_0 - 1) = \ln(s_0) - \frac{c}{s_0}.$$

It can be verified that because of the above equality

$$s_0 = \frac{c}{c + \mathbf{W}(-c\exp(-c))},$$

where $\mathbf{W}$ is a special function known as Lambert $\mathbf{W}$ function (for details about this function see [4]). In addition it can be easily verified that:

$$\int_0^1 \ln(\frac{s_0 - x}{1 - x})dx = s_0\ln(s_0) - s_0\ln(s_0 - 1) + \ln(s_0 - 1) = c + \ln(s_0 - 1).$$

Thus,

$$H(l,k,0) \asymp \exp\left[-k(c + \ln(s_0 - 1) - c\ln(s_0))\right].$$

Since we have fixed $A \in \mathbf{A}_n$ with k FALSE values and conditioned on the event that the random formula ϕ contains l blocking clauses for A, we obtain:

$$\begin{aligned}\Pr[A \in \mathbf{A}_n^1 \mid A \models \phi, |\phi \cap \mathbf{B}_A| = l] &= H(k,l,0) \\ &\asymp \exp\left[-k(c + \ln(s_0 - 1) - c\ln(s_0))\right].\end{aligned} \tag{4}$$

Now to compute $\Pr[|\phi \cap \mathbf{B}_A| = l \; |A \models \phi]$, i.e. the second probability appearing in the right-hand side of (2), we consider the sequence of clause selections for ϕ, drawn from the set of all clauses satisfied by A, as a sequence of $m = rn$ Bernoulli trials. Success occurs whenever a clause belongs to $\mathbf{B}_A$, i.e. it is a blocking clause.

The probability of this event is equal to $\frac{k\binom{n-1}{2}}{7\binom{n}{3}} = \frac{3k}{7n} = \frac{3\alpha}{7}$. Now, we have the following asymptotic expansion of a Binomial distribution with constant probability of success: :

$$\Pr[|\phi \cap \mathbf{B}_A| = l \mid A \models \phi] = \binom{rn}{l}\left(\frac{3k}{7n}\right)^l\left(1-\frac{3k}{7n}\right)^{rn-l}$$
$$\asymp \left[\frac{(3\alpha)^\beta(7-3\alpha)^{(1-\beta)}}{7\beta^\beta(1-\beta)^{(1-\beta)}}\right]^{rn}, \tag{5}$$

where we used $\binom{rn}{l} \asymp \left[(\frac{rn}{l})^{\frac{l}{rn}}(\frac{rn}{rn-l})^{(1-\frac{l}{rn})}\right]^{rn}$. Let now $E(\alpha,\beta,r,n)$ be given by

$$\left[\exp\left[-\alpha(c+\ln(s_0-1)-c\ln s_0)\right]\left(\frac{(3\alpha)^\beta(7-3\alpha)^{1-\beta}}{7\beta^\beta(1-\beta)^{1-\beta}}\right)^r\right]^n \tag{6}$$

Combining (2), (4), (5) and (6), we obtain the following (recall that $\alpha = \frac{k}{n}$ and $\beta = \frac{l}{rn}$):

Theorem 2

$$\Pr[A \in \mathbf{A}_n^1 | A \models \phi] \asymp \sum_{l=k}^{rn} E(\alpha,\beta,r,n). \tag{7}$$

4 Probability models for random formulas

A random 3-SAT formula ϕ with $m = rn$ clauses is usually formed according to one of the following probability models (Ω is the set of $8\binom{n}{3}$ possible 3-SAT clauses, and $p(n) = \frac{rn}{8\binom{n}{3}} \sim \frac{3r}{4n^2}$):

- Select the m clauses of ϕ, drawing each clause uniformly and independently from Ω, with replacement (model G_{mm}).
- Select the m clauses of ϕ, drawing each clause uniformly and independently from Ω, without replacement (model G_m).
- Independently, each clause in Ω is chosen with probability $p(n)$ for inclusion in ϕ (model G_p).

The probability of an event Q concerning a random formula formula ϕ generated according to model G_m, G_{mm}or $G_p$is denoted by $\Pr_m[\phi \in Q], \Pr_{mm}[\phi \in Q]$ and $\Pr_p[\phi \in Q]$, respectively. Notice that the probabilities in (1) are all in G_{mm}, since the model we considered until now allows clause repetitions when forming a formula. Because the second probability in (1) has been computed in [12] in the model G_p, we introduce below some machinery for going from one probability model to another.

Consider again an arbitrary but fixed truth assignment A. As all probabilities which will undergo a change in the probabilistic model are conditional on $A \models \phi$, in the considerations below we assume that the universe of all clauses is restricted to those that are satisfied by A, and consequently that $p(n) = \frac{rn}{7\binom{n}{3}} \sim \frac{6r}{7n^2.}$

First let P be the event that ϕ has no two clauses identical and let $\bar{P}$ its complement. Then, because the order of the number of all possible clauses is $\Theta(n^3)$ and the order of the number of the clauses contained in ϕ is $\Theta(n)$, $\lim_{n\to\infty} \Pr_{mm}[\bar{P}] = 0$.

Now let Q_1 and Q_2 be two arbitrary events such that the following two conditions, which we call *regularity* conditions hold:

- For some $\epsilon > 0$ and finally for all n, $\ln (\Pr_{mm}[Q_2|Q_1]) < -\epsilon$, i.e. $\Pr_{mm}[Q_2|Q_1]$ is bounded away from 1.
- $\lim_{n\to\infty} \Pr_{mm}[\bar{P}|Q_1, Q_2] = \lim_{n\to\infty} \Pr_{mm}[\bar{P}|Q_1] = 0$.

Notice that the events we consider in this paper have probabilities (conditional or not) that are exponentially small, so the first of the two regularity conditions is satisfied. Also, the second regularity condition is true when Q_1 and Q_2 are a the events $A \in \mathbf{A}_n^1$ and $A \in \mathbf{A}_n^2$, respectively. Indeed both these events and their conjunction are negatively correlated with $\bar{P}$, so $\Pr_{mm}[\bar{P}|Q_1] \leq \Pr_{mm}[\bar{P}] \to 0$ and similarly for $\Pr_{mm}[\bar{P}|Q_1, Q_2]$. To prove the negative correlation claim for, say, Q_1 and $\bar{P}$, observe that the correlation claim is equivalent to $\Pr_{mm}[Q_1|P] \geq \Pr_{mm}[Q_1]$, which in turn is equivalent to $\Pr_m[Q_1] \geq \Pr_{mm}[P]$. This last inequality is intuitively obvious under the assumption that $A \models \phi$, because the probability to get blocking clauses for all FALSE values of the satisfying truth assignment A increases when the clauses of the formula are assumed to be different. For a formal proof of this for general increasing and reducible properties (like Q_1 and Q_2), we refer to [13]. Therefore, the second regularity condition is also true for the probabilities we will consider below.

Under the above regularity conditions, we have that:

$$\Pr_m[Q_2|Q_1] \asymp \Pr_{mm}[Q_2|Q_1]. \tag{8}$$

Indeed,

$$\Pr_m[Q_2 \, |Q_1] = \Pr_{mm}[Q_2 \, |Q_1, P] = \frac{\Pr_{mm}[Q_2 \, |Q_1] - \Pr_{mm}[Q_2 \wedge \bar{P} \, |Q_1]}{1 - \Pr_{mm}[\bar{P} \, |Q_1]}$$

$$= \Pr_{mm}[Q_2 \mid Q_1] \frac{1 - \Pr_{mm}[\bar{P} \mid Q_2, Q_1]}{1 - \Pr_{mm}[\bar{P} \mid Q_1]}.$$

Now first taking logarithms, then dividing both sides with $\ln\left(\Pr_{mm}[Q_2|Q_1]\right)$ and finally letting $n \to \infty$, we get the required (the regularity conditions are needed in the computation of the limits).

On the other hand, it follows easily from Theorem II.2 (iii) in [2] that:

$$\Pr_m[Q_2 \mid Q_1] \leq 3m^{1/2}\Pr_p[Q_2 \mid Q_1], \tag{9}$$

even without assuming that the regularity conditions for Q_1 and Q_2 hold. Notice that the proof in [2] is purely algebraic, therefore we can legitimately apply it for conditional probabilities.

Now, (8) (9) and (1) establish the following inequality (ignoring polynomial factors):

$$\begin{aligned} &\Pr_{mm}[\phi \text{ is satisfiable}] \leq \\ &\quad \left(\frac{7}{8}\right)^{rn} \sum_{A \in \mathbf{S}} \Pr_{mm}[A \in \mathbf{A}_n^1 \mid A \models \phi] \cdot \Pr_p[A \in \mathbf{A}_n^2 \mid A \in \mathbf{A}_n^1]. \end{aligned} \tag{10}$$

We are now in a position to use the probability calculations performed earlier in this paper using the ball allocation analogy for the method of local maxima with the probability calculations in [12] in order to derive a value for r that makes the right-hand side of (10) converge to 0. It is perhaps interesting to note that $\Pr_p[A \models \phi] \not\asymp \Pr_{mm}[A \models \phi]$ (the former is larger). Therefore, as both $Pr_p[A \models \phi]$ and $Pr_{mm}[A \models \phi]$ are easily computable, it was advantageous to retain in the first factor of the right hand side of (10) the value of $\Pr[A \models \phi]$ computed in the model G_{mm} ($Pr_{mm}[A \models \phi] = (7/8)^{rn}$) rather than replace it with its value in G_p ($\Pr_p[A \models \phi] \asymp (e^{-(1/8)rn}$). Also, $\Pr_p[A \in \mathbf{A}_n^1 \mid A \models \phi] \not\asymp \Pr_{mm}[A \in \mathbf{A}_n^1 \mid A \models \phi]$. In order to exploit the advantage of the model G_{mm} for the probability of single flips as well, we had to develop the balls in bins scheme (the calculations would have been much easier if performed in G_p). We were unable to do the same for the computation of $\Pr_{mm}[A \in \mathbf{A}_n^2 \mid A \in \mathbf{A}_n^1]$, so we switch to the "easier", but worse, model G_p (in contrast, by Equation (8), the models G_{mm} and G_m are not asymptotically distinguishable, ignoring polynomial factors).

5 Calculation of the probability for double flips

In this section, we will find an asymptotic upper bound for $\Pr_p[A \in \mathbf{A}_n^2 \mid A \in \mathbf{A}_n^1]$. To this end, we establish an asymptotic upper bound for the q-binomial coefficients and combine it with the results in [12].

Let $sf(A) = k = \alpha n$ denote the number of FALSE values assigned by the truth assignment A and let $df(A)$ be the number of double flips of A. Let also:

$$X(sf(A)) = \Pr_{mm}[A \in \mathbf{A}_n^1 \mid A \models \phi] \asymp \sum_{l=k}^{rn} E(\alpha, \beta, r, n). \tag{11}$$

We also define the following functions of r:

$$u = e^{-r/7},$$
$$z = -\frac{6u^6 \ln(1/u)}{1-u^3} + \frac{18u^9 \ln^2(1/u)}{(1-u^3)^2} - \frac{\mathbf{W}\left(-\frac{6u^6\ln(1/u)}{1-u^3}\right)}{\frac{6u^6\ln(1/u)}{1-u^3}}, \tag{12}$$
$$Y = 1 + z\frac{1}{n} + o\left(\frac{1}{n}\right). \tag{13}$$

In [12], it was proved that for $3 \leq r \leq 5$, $z < 0$ and that:

$$\Pr_p[A \in \mathbf{A}_n^2 \mid A \in \mathbf{A}_n^1] \leq Y^{df(A)}. \tag{14}$$

Therefore, using (11) and (14), Inequality (10) may be written as follows:

$$\Pr_{mm}[\phi \text{ is satisfiable}] \leq (7/8)^{rn} \textstyle\sum_{A\in\mathbf{S}} X(sf(A))Y^{df(A)}. \tag{15}$$

Furthermore the following equality can be derived (see [12]) by induction on n:

$$\textstyle\sum_{A\in\mathbf{S}} X(sf(A))Y^{df(A)} = \displaystyle\sum_{k=0}^{n} \binom{n}{k}_Y X(k), \tag{16}$$

where $\binom{n}{k}_q$ denotes the q*-binomial* or *Gaussian* coefficients (see [9]). From Inequality (15), the definition of $X(sf(A))$ in (11), Equality (16) and Theorem 2, we obtain the following:

$$\Pr_{mm}[\phi \text{ is satisfiable}] \leq \left(\frac{7}{8}\right)^{rn} \sum_{k=0}^{n} \sum_{l=k}^{rn} \binom{n}{k}_Y E(\alpha, \beta, r, n). \tag{17}$$

We will now consider an arbitrary term of the double sum that appears in (17) and examine for which values of r it converges to 0. If we find a condition on r that forces all such terms to converge to 0, then the whole sum will converge to 0 since it contains polynomially many terms, all of which vanish exponentially fast. This technique avoids the problem of finding a closed-form upper bound for the sum itself. However, in order to handle an arbitrary term, we need an upper bound for the q-binomial coefficients. To establish such a bound we need the following standard result:

LEMMA 2 [17] *Let* $f(z) = \sum_{i=0}^{\infty} f_i z^i$ *be the generating function for the sequence* f_i, $i \geq 0$. *Then if* $f(z)$ *is analytic in* $|z| < R$ *and if* $f_i \geq 0$ *for all* $i \geq 0$, *then for any* t, $0 < t < R$, *and any* $n \geq 0$, *it holds that* $f_n \leq t^{-n} f(t)$.

Using this lemma, we can prove the following:

Theorem 3 *Let* $\binom{n}{\alpha n}_q$ *denote the q-binomial coefficients for* $\alpha, q \in (0,1)$. *Then the following inequality holds:*

$$\binom{n}{\alpha n}_q \leq 2q^{-\binom{\alpha n}{2}} x_0^{-\alpha n} e^{\frac{1}{\ln q}[\mathrm{dilog}(1+x_0)-\mathrm{dilog}(1+x_0 q^{n-1})]}, \tag{18}$$

where $x_0 = \frac{1-q^{\alpha n}}{q^{\alpha n}-q^{n-1}}$ *and* $\mathrm{dilog}(x) = \int_1^x \frac{\ln t}{1-t} dt$.

Proof. For the ordinary generating function of $q^{\binom{i}{2}}\binom{n}{i}_q$ the following holds [3, p.118]:

$$\sum_{i=0}^{n} q^{\binom{i}{2}} \binom{n}{i}_q x^i = \prod_{i=1}^{n}(1+xq^{i-1}) = e^{\sum_{i=1}^{n}\ln(1+xq^{i-1})} = (1+x)e^{\sum_{i=2}^{n}\ln(1+xq^{i-1})}.$$

Since $\ln(1+xq^{i-1})$ is decreasing in i,

$$\sum_{i=0}^{n} q^{\binom{i}{2}} \binom{n}{i}_q x^i \leq (1+x)e^{\int_1^n \ln(1+xq^{i-1})di} = (1+x)e^{\frac{1}{\ln q}[\mathrm{dilog}(1+x)-\mathrm{dilog}(1+xq^{n-1})]}.$$

Applying Lemma 2, we have that for all $x \in (0,1)$

$$q^{\binom{i}{2}} \binom{n}{i}_q \leq x^{-i}(1+x)e^{\frac{1}{\ln q}[\mathrm{dilog}(1+x)-\mathrm{dilog}(1+xq^{n-1})]}. \tag{19}$$

The above inequality holds for any value of $x \in (0,1)$. Therefore, we may optimize it by choosing the value $x_0 = \frac{1-q^i}{q^i-q^{n-1}}$ that minimizes the expression on the right-hand side of (19). The required inequality is then obtained by setting $i = \alpha n$. ■

Setting $q = Y = 1 + z/n$ in (18) and using the approximation $\ln(1+z/n) \sim z/n$, as $n \to \infty$, the following can be derived:

$$\binom{n}{\alpha n}_q \leq 2\left[\left(\frac{1}{x_0}\right)^{\alpha} \cdot e^{-\frac{\alpha^2 z}{2}+\frac{1}{z}[\mathrm{dilog}(1+x_0)-\mathrm{dilog}(1+x_0 e^z)]}\right]^n, \tag{20}$$

where $x_0 = \frac{1-e^{\alpha z}}{e^{\alpha z}-e^z}$, which is expedient in the proof of the following:

Theorem 4 *An arbitrary term of the double sum in (17) is asymptotically (ignoring polynomial multiplicative factors) bounded from above by the following expression F raised to n:*

$$F = \exp\left[-\alpha(c+\ln(s_0-1)-c\ln s_0)\right]\left(\frac{(3\alpha)^{\beta}(7-3\alpha)^{1-\beta}}{8\beta^{\beta}(1-\beta)^{1-\beta}}\right)^r \times \frac{e^{-\frac{\alpha^2 z}{2}+\frac{1}{z}[\mathrm{dilog}(1+x_0)-\mathrm{dilog}(1+x_0 e^z)]}}{x_0{}^{\alpha}},$$

where, $\alpha = \frac{k}{n}$, $\beta = \frac{l}{rn}$, $x_0 = \frac{1-e^{\alpha z}}{e^{\alpha z}-e^z}$, z as given in (12) and $s_0 = \frac{c}{c+\mathbf{W}(-ce^{-c})}$, with $c = r\frac{\beta}{\alpha}$.

An immediate consequence of this result is that *any* value of r for which F is smaller than 1 for all α, β over the domain $\mathbf{D} = \{\alpha, \beta \in (0,1)$ and $\frac{\beta r}{\alpha} \geq 1\}$ is an upper bound for the unsatisfiability threshold. In other words, any value of r for which the maximum of the function $\ln(F)$ over $\mathbf{D}$ is negative is an upper bound for the threshold.

We finally claim that for any value of r, the expression $\ln(F)$ is an upwards convex function of α, β over the domain $\mathbf{D}$. For a proof of this claim see Appendix.

Since for any fixed r, $\ln(F)$ is upwards convex and continuously differentiable, there is a unique point in $\mathbf{D}$ where $\ln(F)$ attains its maximum, and this point can be computed by setting the partial derivatives of $\ln(F)$ equal to 0. Due to the complicated form of the expression $\ln(F)$, we maximized it numerically over $\mathbf{D}$ for $r = 4.571$ using a Maple [16] implementation of *Downhill Simplex*. This implementation is based on the method and the code described in [18] and it is freely distributed by F.J. Wright in his Web page [20]. Guided by the plot of $\ln(F)$ given by Maple, we chose as a starting set of values for downhill simplex $\alpha = 0.42$ and $\beta = 0.21$. We set the accuracy and the scale parameters equal to 10^{-50}. In addition, we set the *Digits* parameter of Maple (accuracy of floating point numbers) equal to 100. We ran downhill simplex and it returned as the maximum value of $\ln(F)$ over $\mathbf{D}$ the number -0.0000884. We then computed all the partial derivatives of $\ln(F)$ at the point of $\mathbf{D}$ where $\ln(F)$ takes the value -0.0000884. They were found to be numerically equal to 0. As a final check, we generated 30000 random points close to the point of $\mathbf{D}$ where $\ln(F)$ takes the value -0.0000884 and we confirmed that at all these points, the value of $\ln(F)$ is not greater than -0.0000884. All these considerations show that the maximum of $\ln(F)$ over $\mathbf{D}$ is negative for $r = 4.571$. (For larger values of r, the downhill simplex returns a positive maximum). Thus, the value $r = 4.571$ is established as an upper bound to the unsatisfiability threshold.

References

[1] D. Achlioptas, G. B. Sorkin, “Optimal Myopic Algorithms for Random 3-SAT” in: *Proc. 41st Annual Symposium on Foundations of Computer Science (FOCS)*, pp 590–600, 2000.

[2] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[3] L. Comtet, *Advanced Combinatorics; The Art of Finite and Infinite Expansions*, D. Reidel, 1974.

[4] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, D.E. Knuth, “On the Lambert **W** function,” *Advances in Computational Mathematics* 5, pp. 329–359, 1996.

[5] O. Dubois, Y. Boufkhad, J. Mandler, “Typical random 3-SAT formulae and the satisfiability threshold,” in: *Proc. 11th Symposium on Discrete Algorithms (SODA)*, pp 126–127, 2000.

[6] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., 1968.

[7] N.J. Fine, *Basic Hypergeometric Series and Applications*, Mathematical Surveys and Monographs, Number 27, 1980.

[8] E. Friedgut, appendix by J. Bourgain, “Sharp thresholds of graph properties, and the k-sat problem,” *J. Amer. Math. Soc.* 12, pp 1017–1054, 1999.

[9] G. Gasper, M. Rahman, *Basic Hypergeometric Series*, Encyclopedia of Mathematics and its Applications, vol 35, Cambridge University Press, 1990.

[10] S. Janson, Y.C. Stamatiou, M. Vamvakari, “Bounding the unsatisfiability threshold of random 3-SAT,” *Random Structures and Algorithms* 17, pp 108–116, 2000.

[11] A. Kamath, R. Motwani, K. Palem, P. Spirakis, “Tail bounds for the occupancy and the satisfiability threshold conjecture,” *Random Structures and Algorithms* 7, pp 59–80, 1995.

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

[13] L.M. Kirousis and Y.C. Stamatiou, *An inequality for reducible, increasing properties of randomly generated words*, Technical Report TR-96.10.34, Computer Technology Institute, University of Patras, Patras Greece, 1996.

[14] D.E. Knuth, *Fundamental Algorithms*, The Art of Computer Programming, 2nd ed., Addison-Wesley, 1973.

[15] V.F. Kolchin, *Random Mappings*, Translation series in Mathematics and Engineering, Optimization Software Inc., 1986.

[16] M.B. Monagan, K.O. Geddes, K.M. Heal, G. Labahn, S.M. Vorkoetter, *Programming Guide*, Maple V Release 5, Springer-Verlag, 1998.

[17] A.M. Odlyzko, "Asymptotic Enumeration Methods," in: R.L. Graham, M. Grötschel, and L. Lovász, eds. *Handbook of Combinatorics*, Chapter 22, pp. 1063–1229, Elsevier, 1995.

[18] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, 1993.

[19] A.E. Taylor, W.R Mann, *Advanced Calculus*, 3rd ed., John Wiley & Sons, 1983.

[20] F.J. Wright: `http://centaur.maths.qmw.ac.uk/Computer_Algebra/`. A Maple (V-5) implementation of *Downhill Simplex* based on code given in [18].

[21] M. Zito, Randomised Techniques in Combinatorial Algorithmics, PhD Thesis, Department of Computer Science, University of Warwick, November 1999.

Appendix

We establish the upwards convexity of $\ln(F)$ for a any fixed r over the domain $\mathbf{D} = \{\alpha, \beta \in [0,1]$ and $\frac{\beta r}{\alpha} \geq 1\}$.

$$\begin{aligned}\ln(F) = & -\alpha(c + \ln(s_0 - 1) - c\ln s_0) \\ & -\beta r \ln\beta - (1-\beta) r \ln(1-\beta) + \beta r \ln(3\alpha) - r\ln 8 + r(1-\beta)\ln(7-3\alpha) \\ & +\alpha \ln\tfrac{1}{x_0} - \alpha^2 \tfrac{z}{2} + \tfrac{1}{z}[\mathrm{dilog}(1+x_0) - \mathrm{dilog}(1+x_0 e^z)].\end{aligned}$$

A. The expression $-\beta r \ln\beta - (1-\beta) r \ln(1-\beta) + \beta r \ln(3\alpha) - r\ln 8 + r(1-\beta)\ln(7-3\alpha)$ is an upwards convex function of α, β. Indeed, the quadratic form of its Hessian (see e.g. [19]) computed at an arbitrary vector $(\alpha', \beta') \in \mathbf{R}^2$ is:

$$\begin{aligned} & -(\alpha')^2 r\left[\frac{\beta}{\alpha^2} + \frac{9(1-\beta)}{(7-3\alpha)^2}\right] + 2\alpha'\beta' r\left[\frac{1}{\alpha} + \frac{3}{7-3\alpha}\right] - (\beta')^2 r\left[\frac{1}{\beta} + \frac{1}{1-\beta}\right] \\ & = -(\alpha')^2 r\left[\frac{\beta(7-3\alpha)^2 + 9(1-\beta)r}{\alpha^2(7-3\alpha)^2}\right] + 2\alpha'\beta' r\left[\frac{7}{\alpha(7-3\alpha)}\right] - (\beta')^2 r\left[\frac{1}{\beta(1-\beta)}\right] \\ & \leq -(\alpha')^2 r\left[\frac{\beta(7-3\alpha)^2 + 9(1-\beta)r}{\alpha^2(7-3\alpha)^2}\right] + 2\alpha'\beta' r\left[\frac{\sqrt{\beta(7-3\alpha)^2 + 9(1-\beta)\alpha^2}}{\sqrt{\beta(1-\beta)}\alpha(7-3\alpha)}\right] \\ & -(\beta')^2 r\left[\frac{1}{\beta(1-\beta)}\right] \\ & = -\left[\alpha'\sqrt{r\frac{\beta(7-3\alpha)^2 + 9(1-\beta)r}{\alpha^2(7-3\alpha)^2}} - \beta'\sqrt{r\frac{1}{\beta(1-\beta)}}\right]^2. \end{aligned}$$

Therefore, the Hessian is negative semi-definite and so the function $-\beta r\ln\beta - (1-\beta) r \ln(1-\beta) + \beta r \ln(3\alpha) - r\ln 8 + r(1-\beta)\ln(7-3\alpha)$ is upwards convex.

B. The expression $\alpha \ln\frac{1}{x_0} - \alpha^2\frac{z}{2} + \frac{1}{z}[\mathrm{dilog}(1+x_0) - \mathrm{dilog}(1+x_0 e^z)]$, where $x_0 = \frac{1-e^{\alpha z}}{e^{\alpha z} - e^z}$, $z < 0$, is an upwards convex function of α, β. Indeed, first observe that it is actually a function of α alone since z (recall its definition in the beginning of Section 5) and x_0 depend only on α (for fixed r). Therefore, the quadratic form of its Hessian computed at an arbitrary $(\alpha', \beta') \in \mathbf{R}^2$ is: $z\left[-\frac{(e^z-1)e^{\alpha z}}{(1-e^{\alpha z})(e^{\alpha z}-e^z)} - 1\right](\alpha')^2$. Since $z < 0$, in order to show that the last expression is non-positive it suffices to show that $\frac{(e^z-1)e^{\alpha z}}{(1-e^{\alpha z})(e^{\alpha z}-e^z)} < -1$. Since $(1-e^{\alpha z})(e^{\alpha z} - e^z) > 0$, it is sufficient to show that $(e^z - 1)e^{\alpha z} < -(1-e^{\alpha z})(e^{\alpha z} - e^z) \Leftrightarrow e^z - 2e^{\alpha z}e^z + (e^{\alpha z})^2 > 0$. But the last inequality holds, since $e^z - 2e^{\alpha z}e^z + (e^{\alpha z})^2 > (e^z)^2 - 2e^{\alpha z}e^z + (e^{\alpha z})^2 = (e^z - e^{\alpha z})^2 > 0$.

C. Finally, let us consider the expression $M = -\alpha(c + \ln(s_0 - 1) - c\ln s_0)$ as a function of α, β. The quadratic form of its Hessian computed at an arbitrary vector $(\alpha', \beta') \in \mathbf{R}^2$ is:

$$\frac{\beta r e^{-\beta r/(\alpha s_0)}}{e^{-\beta r/(\alpha s_0)} - \alpha/(\beta r)} \left[\frac{\alpha'}{\alpha} - \frac{\beta'}{\beta}\right]^2. \tag{21}$$

We prove (21) by the following steps: Using that $\ln(s_0 - 1) = \ln s_0 - \frac{c}{s_0}$, $c = \frac{\beta r}{\alpha}$ and $s_0(1 - e^{-c/s_0}) = 1$, we first obtain:

$$M = -r\beta - \alpha \ln s_0 + \frac{r\beta}{s_0} + r\beta \ln s_0. \tag{22}$$

Setting in Expression (22), $s_0 = \frac{c}{x^*}$, where x^* is an arbitrary positive variable, we then obtain:

$$M = -r\beta - \alpha \ln \frac{r\beta}{\alpha} + \alpha \ln x^* + \alpha x^* + r\beta \ln \frac{r\beta}{\alpha} - r\beta \ln x^*. \tag{23}$$

Notice now that:

$$s_0(1 - e^{-c/s_0}) = 1 \Leftrightarrow 1 - e^{x^*} - \frac{\alpha}{r\beta}x^* = 0 \Leftrightarrow \frac{\alpha e^{x^*}}{e^{x^*} - 1} = \frac{r\beta}{x^*}. \tag{24}$$

Therefore the partial derivatives of Expression (23) are given by:

$$\frac{\partial M}{\partial \alpha} = x^* + \ln x^* + \ln \alpha - \frac{r\beta}{\alpha} - \ln r\beta, \qquad \frac{\partial M}{\partial \beta} = r \ln r\beta - r\ln\alpha - r\ln x^* .$$

$$\frac{\partial^2 M}{\partial \alpha^2} = \frac{x^*+1}{x^*}\frac{\partial x^*}{\partial \alpha} + \frac{1}{\alpha} + \frac{r\beta}{\alpha^2}, \quad \frac{\partial^2 M}{\partial \beta^2} = -\frac{r}{x^*}\frac{\partial x^*}{\partial \beta} + \frac{r}{\beta}, \quad \frac{\partial^2 M}{\partial \alpha \partial \beta} = -\frac{r}{x^*}\frac{\partial x^*}{\partial \alpha} - \frac{r}{\alpha}.$$

We then use once more Expression (24) in order to obtain:

$$\frac{\partial x^*}{\partial \alpha} = \frac{x^*}{r\beta}\frac{1}{e^{-x^*} - \alpha/(r\beta)}, \qquad \frac{\partial x^*}{\partial \beta} = -\frac{\alpha x^*}{r\beta^2}\frac{1}{e^{-x^*} - \alpha/(r\beta)}.$$

Using the partial derivatives above, we conclude that (21) holds.

Now using (24) again, we observe that

$$e^{-\beta r/(\alpha s_0)} - \frac{\alpha}{\beta r} = e^{-x^*} - \frac{\alpha}{\beta r} = \frac{1 + x^* - e^{x^*}}{x^* e^{x^*}} < 0,$$

since $1 + x^* < e^{x^*}$. Therefore, expression (21) is non-positive and the proof of the upwards convexity of $\ln(F)$ over $\mathbf{D}$ is complete.