*Approximating the Unsatisfiability Threshold of Random Formulas**, †

Lefteris M. Kirousis,[1] Evangelos Kranakis,[2] Danny Krizanc,[2] Yannis C. Stamatiou[1]

[1] *Department of Computer Engineering and Informatics, University of Patras, 265 00 Patras, Greece; e-mail: {kirousis, stamatiu}@ceid.upatras.gr*

[2] *School of Computer Science, Carleton University, Ottawa, Ont. K1S 5B6, Canada; e-mail: {kranakis, krizanc}@scs.carleton.ca*

Received 14 June 1996; accepted 27 September 1997

ABSTRACT: Let ϕ be a random Boolean formula that is an instance of 3-SAT. We consider the problem of computing the least real number κ such that if the ratio of the number of clauses over the number of variables of ϕ strictly exceeds κ, then ϕ is almost certainly unsatisfiable. By a well-known and more or less straightforward argument, it can be shown that $\kappa \leq 5.191$. This upper bound was improved by Kamath et al. to 4.758 by first providing new improved bounds for the occupancy problem. There is strong experimental evidence that the value of κ is around 4.2. In this work, we define, in terms of the random formula ϕ, a decreasing sequence of random variables such that, if the expected value of any one of them converges to zero, then ϕ is almost certainly unsatisfiable. By letting the expected value of the first term of the sequence converge to zero, we obtain, by simple and elementary computations, an upper bound for κ equal to 4.667. From the expected value of the second term of the sequence, we get the value 4.601+. In general, by letting the

---

* This work was performed while the first author was visiting the School of Computer Science, Carleton University, and was partially supported by NSERC (Natural Sciences and Engineering Research Council of Canada), and by a grant from the University of Patras for sabbatical leaves. The second and third authors were supported in part by grants from NSERC (Natural Sciences and Engineering Research Council of Canada). During the last stages of this research, the first and last authors were also partially supported by EU ESPRIT Long-Term Research Project ALCOM-IT (Project No. 20244).

†An extended abstract of this paper was published in the *Proceedings of the Fourth Annual European Symposium on Algorithms, ESA'96*, September 25–27, 1996, Barcelona, Spain (Springer-Verlag, LNCS, pp. 27–38). That extended abstract was coauthored by the first three authors of the present paper.

Correspondence to: L. M. Kirousis

expected value of further terms of this sequence converge to zero, one can, if the calculations are performed, obtain even better approximations to κ. This technique generalizes in a straightforward manner to k-SAT for $k > 3$. Random Struct. Alg., 12, 253–269, 1998

Key Words: approximations; probabilistic method; random formulas; satisfiability; threshold point

1. INTRODUCTION

Let ϕ be a random 3-SAT formula on n Boolean variables $x_1, \ldots, x_n$. Let m be the number of clauses of ϕ. The clauses-to-variables ratio of ϕ is defined to be the number m/n. We denote this ratio by r. The problem we consider in this paper is to compute the least real number κ such that if r strictly exceeds κ, then the probability of ϕ being satisfiable converges to 0 as n approaches infinity. We say in this case that ϕ is asymptotically almost certainly unsatisfiable. Experimental evidence suggests that the value of κ is around 4.2. Moreover, experiments suggest that if r is strictly smaller than κ, then ϕ is asymptotically almost certainly satisfiable. Thus, experimentally, κ is not only the lower bound for unsatisfiability, but it is a threshold value where "suddenly," probabilistically certain unsatisfiability yields to probabilistically certain satisfiability (for a review of the experimental results, see [13]).

In the literature for this problem, the most common model for random 3-SAT formulas is the following: from the space of clauses with *exactly three* literals of three *distinct* variables from $x_1, \ldots, x_n$, uniformly, independently, and with replacement select m clauses that form the set of conjuncts of ϕ (thus, a clause may be selected more than once). We adopt this model in this paper; however, the results can be generalized to any of the usual models for random formulas. The total number N of all possible clauses is $8\binom{n}{3}$, and given a truth assignment A, the probability that a random clause is satisfied by A is $7/8$. Also, given three distinct variables x_i, x_j, x_k, there is a unique clause on the variables x_i, x_j, x_k which is *not* satisfied by A. There are $\binom{n}{3}$ such clauses, and they constitute exactly the set of clauses not satisfied by A.

A proposition stating that if r exceeds a certain constant, then ϕ is asymptotically almost certainly unsatisfiable has as an immediate corollary that this constant is an upper bound for κ. We use this observation in our technique to improve the upper bound for κ.

A well-known "first moment" argument shows that

$$\kappa \leq \log_{8/7} 2 = 5.191.$$

To prove it, observe that the expected value of the number of truth assignments that satisfy ϕ is $2^n(7/8)^{rn}$; then let this expected value converge to zero and use Markov's inequality (this argument is expanded below). According to Chvátal and Reed [5], this observation is due to Franco and Paull [8], Simon et al. [19], Chvátal and Szemerédi [6], and possibly others.

Let $\mathscr{A}_n$ be the set of all truth assignments on the n variables $x_1, \ldots, x_n$, and let $\mathscr{S}_n$ be the set of truth assignments that satisfy the random formula ϕ. Thus, the cardinality $|\mathscr{S}_n|$ is a random variable. Also, for an instantiation ϕ of the random formula, let $|\mathscr{S}_n(\phi)|$ denote the number of truth assignments that satisfy ϕ. (A word of caution: in order to avoid overloading the notation, we use the same symbol ϕ to denote the random formula and an instantiation of it.) We give below a rough outline of the simplest case of our technique.

By definition, the expected value of the number of satisfying truth assignments of a random formula, i.e., $\mathbf{E}[|\mathscr{S}_n|]$, satisfies the following relation:

$$\mathbf{E}[|\mathscr{S}_n|] = \sum_{\phi} \left(\Pr[\phi] \cdot |\mathscr{S}_n(\phi)|\right). \tag{1}$$

On the other hand, the probability of a random formula being satisfiable is given by the equation

$$\Pr[\text{the random formula is satisfiable}] = \sum_{\phi} \left(\Pr[\phi] \cdot I_{\phi}\right) \tag{2}$$

where

$$I_{\phi} = \begin{cases} 1 & \text{if } \phi \text{ is satisfiable} \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

From Eqs. (1) and (2), the following Markov's inequality follows immediately:

$$\Pr[\text{the random formula is satisfiable}] \leq \mathbf{E}[|\mathscr{S}_n|]. \tag{4}$$

It is easy to find a condition on κ under which $\mathbf{E}[|\mathscr{S}_n|]$ converges to zero. Such a condition, by Markov's inequality (4), implies that ϕ is asymptotically almost certainly unsatisfiable (this elementary technique is known as the "first moment method"). However, on the right-hand side of Eq. (1), we may have small probabilities multiplied by large cardinalities; therefore, such a condition may be unnecessarily strong for guaranteeing only that ϕ is almost certainly unsatisfiable. In this work, instead of considering the random class $\mathscr{S}_n$ that may have a large cardinality for certain instantiations of the random formula with small probability, we consider a subset of it obtained by taking truth assignments that satisfy a local maximality condition. Thus, the condition obtained by letting the expected value of this new class converge to zero is weakened, and consequently, the upper bound for κ is lowered.

As we show in the next section, the bound for κ obtained by this sharpened first moment technique is equal to 4.667. This improves the previous best bound due to Kamath et al. [12] of 4.758, which was obtained by nonelementary means. Moreover, our method is not computational, i.e., it does not use any mechanical computations that do not have provable accuracy and correctness (the fact that, in our method, we use a computer program to find a solution of an equation with *one* unknown does not render our proof computational because the algorithms that find solutions to such equations have provable accuracy). The bound that Kamath et al. [12] attain with a noncomputational proof is equal to 4.87. In Section 3, we show how to further improve the bound to 4.601+ (a value between 4.601 and 4.60108) by defining an even smaller subset of $\mathscr{S}_n$. This is achieved by increasing the range of locality when selecting the local maxima that represent $\mathscr{S}_n$. We define a decreasing sequence of subsets of $\mathscr{S}_n$ by selecting from $\mathscr{S}_n$ truth assignments that satisfy a condition of local maximality with increasing range of locality. From this sequence, if we perform the calculations, we can obtain a sequence of improving approximations to κ. In the last section, we discuss the case of letting this range of locality become unboundedly large. Dubois and Boufkhad [7] have independently announced the upper bound of 4.64.

Our bounds can be possibly improved even further if one uses not the Markov type inequality mentioned above, but an analog of the "harmonic mean formula" given by Aldous [2], and then applies the technique that is used in Kamath et al. [12]. This is discussed in the last section. Our method readily generalizes to k-SAT for $k > 3$.

2. SINGLE FLIPS

Recall that $\mathscr{A}_n$ is the class of all truth assignments, and $\mathscr{S}_n$ is the random class of truth assignments that satisfy a random formula ϕ. We now define a class even smaller than $\mathscr{S}_n$.

Definition 1. *For a random formula ϕ, $\mathscr{S}_n^{\sharp}$ is defined to be the random class of truth assignments A such that: (i) $A \vDash \phi$, and (ii) any assignment obtained from A by changing exactly one* FALSE *value of A to* TRUE *does not satisfy ϕ.*

Notice that the truth assignment with all of its values equal to TRUE vacuously satisfies condition (ii) of the previous definition. Consider the lexicographic ordering among truth assignments, where the value FALSE is considered smaller than TRUE, and the values of variables with higher index are of lower priority in establishing the way two assignments compare. It is not hard to see that $\mathscr{S}_n^{\sharp}$ is the set of elements of $\mathscr{S}_n$ that are local maxima in the lexicographic ordering of assignments, where the neighborhood of determination of local maximality is the set of assignments that differ from A in at most one position.

We now prove the following.

Lemma 1. *The following Markov type inequality holds for $\mathscr{S}_n^{\sharp}$:*

$$\Pr[\textit{the random formula is satisfiable}] \leq \mathbf{E}\left[\left|\mathscr{S}_n^{\sharp}\right|\right]. \tag{5}$$

Proof. From the previous definition, we easily infer that if an instantiation ϕ of the random formula is satisfiable, then $\mathscr{S}_n^{\sharp}(\phi) \neq \emptyset$. (Recall that $\mathscr{S}_n^{\sharp}(\phi)$ is the instantiation of the random class $\mathscr{S}_n^{\sharp}$ at the instantiation ϕ.) We also have that

$$\Pr[\text{the random formula is satisfiable}] = \sum_{\phi} \left(\Pr[\phi] \cdot I_{\phi}\right)$$

where

$$I_\phi = \begin{cases} 1 & \text{if } \phi \text{ is satisfiable} \\ 0 & \text{otherwise}. \end{cases} \tag{6}$$

On the other hand,

$$\mathbf{E}\left[\left|\mathscr{S}_n^\sharp\right|\right] = \sum_\phi \left(\Pr[\phi]\cdot\left|\mathscr{S}_n^\sharp(\phi)\right|\right).$$

The lemma now immediately follows from the above. ■

We also have the following.

Lemma 2. *The expected value of the random variable* $|\mathscr{S}_n^\sharp|$ *is given by the formula*

$$\mathbf{E}\left[\left|\mathscr{S}_n^\sharp\right|\right] = (7/8)^{rn} \sum_{A\in\mathscr{A}_n} \Pr\left[A \in \mathscr{S}_n^\sharp \middle| A \in \mathscr{S}_n\right]. \tag{7}$$

Proof. First observe that the random variable $|\mathscr{S}_n^\sharp|$ is the sum of indicator variables, and then condition on $A \vDash \phi$ (recall that r is the number of clauses-to-number-of-variables ratio of ϕ, so $m = nr$). ■

We call a change of *exactly one* FALSE value of a truth assignment A to TRUE a *single flip*. The number of possible single flips, which is, of course, equal to the number of FALSE values of A, is denoted by $sf(A)$. The assignment obtained by applying a single flip sf on A is denoted by A^{sf}.

We now prove the following.

Theorem 1. *The expected value* $\mathbf{E}[|\mathscr{S}_n^\sharp|]$ *is at most* $(7/8)^{rn}(2 - e^{-3r/7} + o(1))^n$. *It follows that the unique positive solution of the equation*

$$(7/8)^r(2 - e^{-3r/7}) = 1$$

is an upper bound for κ (*this solution is less than* 4.667).

Proof. Fix a single flip sf_0 on A, and assume that $A \vDash \phi$. Observe that the assumption that $A \vDash \phi$ excludes $\binom{n}{3}$ clauses from the conjuncts of ϕ, i.e., there remain $7\binom{n}{3}$ clauses from which to choose the conjuncts of ϕ. Now, consider the clauses that are not satisfied by A^{sf_0} and contain the flipped variable. There are $\binom{n-1}{2}$ of them. Under the assumption that $A \vDash \phi$, in order to have that $A^{sf_0} \nvDash \phi$, it is necessary and sufficient that at least one of these $\binom{n-1}{2}$ clauses be a conjunct of ϕ. Therefore, for each of the m clause selections for ϕ, the probability of being one that guarantees that $A^{sf_0} \nvDash \phi$ is $\binom{n-1}{2}/7\binom{n}{3} = 3/(7n)$. Therefore, the probability that $A^{sf_0} \nvDash \phi$ (given that $A \vDash \phi$) is equal to $1 - (1 - 3/(7n))^m$. Now, there are $sf(A)$ possible flips for A. The events that ϕ is not satisfied by the assignment A^{sf} for *each* single flip sf (under the assumption that $A \vDash \phi$) refer to disjoint sets of clauses. Therefore, the dependencies among them are such that

$$\Pr\left[A \in \mathscr{S}_n^{\sharp} \middle| A \models \phi\right] \leq \left(1 - \left(1 - \frac{3}{7n}\right)^m\right)^{sf(A)} = \left(1 - e^{-3r/7} + o(1)\right)^{sf(A)}. \quad (8)$$

Petr Savický has supplied us with a formal proof of the above inequality. In addition, a result that implies it is presented in [17]. Indeed, in the notation of the main theorem in [17], it is enough, in order to obtain the above inequality, to let: (i) $V = \{1, \dots, m\}$, (ii) $I = \{1, \dots, sf(A)\}$, (iii) $X_v = i$ iff the vth clause of ϕ is satisfied by A but not satisfied by A^{sf_i}, where A^{sf_i} is the truth assignment obtained from A by flipping the ith FALSE value of A, and (iv) for all $i, \mathscr{F}_i$ be the "increasing" collection of nonempty subsets of V.

Now, recall that $sf(A)$ is equal to the number of FALSE values of A. Therefore, by Eq. (7) and by Newton's binomial formula, $\mathbf{E}[|\mathscr{S}_n^{\sharp}|]$ is bounded above by $(7/8)^{rn}(2 - (1 - 3/(7n))^{rn})^n$, which proves the first statement of the theorem.

It also follows that $\mathbf{E}[|\mathscr{S}_n^{\sharp}|]$ converges to zero for values of r that strictly exceed the unique positive solution of the equation $(7/8)^r(2 - e^{-3r/7}) = 1$. By Lemma 1, this solution is an upper bound for κ. As can be seen by any program that computes roots of equations with accuracy of at least four decimal digits (we used Maple [18]), this solution is less than 4.667. ∎

The generalization of the previous result to the case of k-SAT, for an arbitrary $k \geq 3$ is immediate.

Theorem 2. *For the case of k-SAT* $(k \geq 3)$, *the expected value* $\mathbf{E}[|\mathscr{S}_n^{\sharp}|]$ *is at most* $((2^k - 1)/2^k)^{rn}(2 - e^{-kr/(2^k-1)} + o(1))^n$. *It follows that the unique positive solution of the equation*

$$\left(\frac{2^k - 1}{2^k}\right)^r \left(2 - e^{-kr/(2^k - 1)}\right) = 1$$

is an upper bound for κ (*as defined for k-SAT*).

3. THE GENERAL METHOD AND DOUBLE FLIPS

In this section, we generalize the previous method to an arbitrary range of locality when selecting the subset of $\mathscr{S}_n$. We start with a definition.

Definition 2. *Given a random formula* ϕ *and a nonnegative integer* l, $\mathscr{A}_n^l (l \leq n)$ *is defined to be the random class of truth assignments A such that*: (*i*) $A \models \phi$, *and* (*ii*) *any assignment that differs from A in at most l variables and is lexicographically strictly larger than A does not satisfy* ϕ.

Observe that $\mathscr{S}_n$ of the previous section, i.e., the class of truth assignments satisfying the random formula, is now redefined as $\mathscr{A}_n^0$ and $\mathscr{S}_n^{\sharp}$ is redefined as $\mathscr{A}_n^1$. $\mathscr{A}_n^l$ is the subset of $\mathscr{S}_n$ that consists of the lexicographic local maxima of it where the neighborhood of locality for an assignment A is the set of assignments that differ from A in at most l places. Moreover, $\mathscr{A}_n^l$ is a sequence of classes which is nonincreasing (with respect to set inclusion).

Now, exactly as in Lemma 1, the following can be proved.

Lemma 3. *The following Markov-type inequalities hold for the classes* $\mathscr{A}_n^l$:

$$\Pr[\phi \textit{ is satisfiable}] = \mathbf{E}[|\mathscr{A}_n^n|] \leq \mathbf{E}\big[|\mathscr{A}_n^{n-1}|\big] \leq \cdots \leq \mathbf{E}\big[|\mathscr{A}_n^1|\big] \leq \mathbf{E}\big[|\mathscr{A}_n^0|\big]. \quad (9)$$

It follows from the above that for a fixed l, by letting $\lim_n \mathbf{E}[|\mathscr{A}_n^l|] = 0$, we obtain upper bounds for κ which decrease as l increases. We concentrate below on the case $l = 2$.

A change of exactly two values of a truth assignment A that gives a truth assignment which is lexicographically greater than A must be of one of the following kinds: (1) a change of the value FALSE of a variable to TRUE and a change of the value TRUE of a higher indexed variable to FALSE, or (2) a change of two variables both of value FALSE to TRUE. From these two possible kinds of changes, we consider only the first since the calculations become easier, while the final result remains the same. We call such changes *double flips*. Define A^{df} and $df(A)$ in a way analogous to the single-flip case (notice that if A is considered as a sequence of the Boolean values 0 and 1, then $df(A)$ is equal to the number of order inversions as we move along A from high-indexed variables to low-indexed ones, i.e., from right to left). Let $\mathscr{A}_n^{2\sharp}$ be the set of assignments A such that $A \vDash \phi$, and for all single flips sf, $A^{sf} \nvDash \phi$ and for all double flips df, $A^{df} \nvDash \phi$. It can be easily seen that $\mathscr{A}_n^2$ is a subset of $\mathscr{A}_n^{2\sharp}$ [in general, a proper one because in the definition of $\mathscr{A}_n^{2\sharp}$, we did not take into account the changes of kind (2)]. Therefore, a value of r that makes the expected value $\mathbf{E}[|\mathscr{A}_n^{2\sharp}|]$ converge to zero is, by Lemma 3, an upper bound for κ. Actually, it can be proved that both $\mathbf{E}[|\mathscr{A}_n^{2\sharp}|]$ and $\mathbf{E}[|\mathscr{A}_n^2|]$ converge to zero for the same values of r, but we will not use this fact, so we omit its proof.

Now, in analogy to Lemma 2, we have the following.

Lemma 4.

$$\mathbf{E}\Big[\big|\mathscr{A}_n^{2\sharp}\big|\Big] = (7/8)^{rn} \sum_{A \in \mathscr{A}_n} \Pr\Big[A \in \mathscr{A}_n^{2\sharp} \Big| A \vDash \phi\Big]$$

$$= (7/8)^{rn} \sum_{A \in \mathscr{A}_n} \Pr\big[A \in \mathscr{A}_n^1 \big| A \vDash \phi\big] \cdot \Pr\Big[A \in \mathscr{A}_n^{2\sharp} \Big| A \in \mathscr{A}_n^1\Big]. \quad (10)$$

Therefore, by the remarks at the beginning of the current section, an upper bound for κ can be found by computing a value (the smaller the better) for r for which the right-hand side of the equation above converges to zero. We will do this in two steps. First, we will compute an upper bound for each term of the second sum in the equation above; then, we will find an upper bound for $\mathbf{E}[|\mathscr{A}_n^{2\sharp}|]$ which will be a closed expression of r and n. Letting this closed expression converge to zero with n, we will get an equation in terms of r that gives the required bound for κ.

To compute an upper bound for the terms of the sum, we will make use of an inequality that appears as [11, Theorem 7], which gives an estimate for the probability of the intersection of dependent events. We give the details in the first subsection of the present section. The computations that will then give a closed expression that is an upper bound for $\mathbf{E}[|\mathscr{A}_n^{2\sharp}|]$ are carried out in the second subsection.

3.1. Probability Calculations

Given a fixed A, we will now find an upper bound for $\Pr[A \in \mathscr{A}_n^1 | A \vDash \phi] \cdot \Pr[A \in \mathscr{A}_n^{2\sharp} | A \in \mathscr{A}_n^1]$. We assume for the rest of this subsection that the condition $A \vDash \phi$ holds. This is equivalent to assuming that the space of all clauses from which we uniformly, independently, and with replacement choose the ones that form ϕ is equal to the set of all clauses satisfied by A. This set of clauses has cardinality $7\binom{n}{3}$. Also notice that under the condition $A \vDash \phi$, the event $A \in \mathscr{A}_n^1$ is equivalent to the statement that for any single flip sf, $A^{sf} \nvDash \phi$. In the sequel, all computations of probabilities, analyses of events, etc., will be carried out assuming that $A \vDash \phi$, usually without explicit mention of it.

To compute $\Pr[A \in \mathscr{A}_n^{2\sharp}]$, it is more convenient to work in another model for random formulas. In the next paragraphs, we give the necessary definitions and notations.

Consistent with the standard notation of the theory of random graphs [4], let $\mathscr{G}_p$ be the model for random formulas where each clause has an independent probability p to appear in the formula, let $\mathscr{G}_m$ be the model where the random formula is obtained by uniformly and independently selecting m clauses *without* replacement, and, finally, let $\mathscr{G}_{mm}$ be the model that we use in this paper, where the formula is obtained by uniformly and independently selecting m clauses *with* replacement (recall that, according to our assumption, we only refer to clauses that are satisfied by A).

The probabilities of an event E in $\mathscr{G}_p$ ($\mathscr{G}_m$) will be denoted by $\Pr_p[E]$ ($\Pr_m[E]$, respectively). In order not to change our notation, we continue to denote the probability of E in the model $\mathscr{G}_{mm}$ by $\Pr[E]$. Set $p = m/\left(7\binom{n}{3}\right) \sim 6r/(7n^2)$. From Bollobás [4, Chap. 3, p. 35, Theorem 2(iii)], we have that for any property Q of formulas, $\Pr_m[Q] \le 3m^{1/2}\Pr_p[Q]$. Additionally, if Q is monotonically increasing (i.e., if it holds for a formula, it also holds for any formula containing more clauses) and *reducible* (i.e., it holds for a formula iff it holds for the formula where multiple occurrences of clauses have been omitted), then $\Pr[Q] \le \Pr_m[Q]$. Intuitively, this is so because, by the assumptions of increasing monotonicity and reducibility for Q, when selecting the clauses to be included in ϕ, we increase the probability to satisfy Q by selecting a "new" clause rather than by selecting one that has already been selected. A formal proof of this property can be found in [15]. Therefore, as nonsatisfiability is both monotonically increasing and reducible, we conclude that

$$\begin{aligned}
\Pr\left[A \in \mathscr{A}_n^{2\sharp}\right] \\
&\le 3m^{1/2}\Pr_p\left[A \in \mathscr{A}_n^{2\sharp}\right] \\
&= 3m^{1/2}\Pr_p\left[A \in \mathscr{A}_n^{2\sharp} \wedge A \in \mathscr{A}_n^1\right] \quad \left(\text{because } A \in \mathscr{A}_n^1 \text{ is implied by } A \in \mathscr{A}_n^{2\sharp}\right) \\
&= 3m^{1/2}\Pr_p\left[A \in \mathscr{A}_n^1\right] \cdot \Pr_p\left[A \in \mathscr{A}_n^{2\sharp} \middle| A \in \mathscr{A}_n^1\right]. \qquad (11)
\end{aligned}$$

It is easy to see, carrying the corresponding argument in the proof of Theorem 1 within the model $\mathscr{G}_p$, that

$$\Pr_p\left[A \in \mathscr{A}_n^1\right] = \left(1-(1-p)^{\binom{n-1}{2}}\right)^{sf(A)} = \left(1-e^{-3r/7}+o(1)\right)^{sf(A)}. \qquad (12)$$

So, by Eqs. (10)–(12), to find an upper bound for $\Pr[A \in \mathscr{A}_n^1]\cdot \Pr[A \in \mathscr{A}_n^{2\sharp} | A \in \mathscr{A}_n^1]$, it is enough to find an upper bound for $\Pr_p[A \in \mathscr{A}_n^{2\sharp} | A \in \mathscr{A}_n^1]$. Computing this last probability is equivalent to computing the probability that for all double flips df, $A^{df} \not\models \phi$, under the condition that for all single flips sf, $A^{sf} \not\models \phi$. In the next lemma, given a fixed double flip df_0, we will compute the probability that $A^{df_0} \not\models \phi$ under the same condition. We will then compute the joint probability for all double flips.

At this point, it is convenient to introduce the following notation to be used in the sequel: for a variable x_i, x_i^A is the literal x_i if the value of x_i in A is TRUE, and it is the literal $\neg x_i$, otherwise. Also, let $q = 1-p$.

First, fix a double flip df_0. Then we have the following.

Lemma 5. *The following holds*:

$$\Pr_p\left[A^{df_0} \not\models \phi | A \in \mathscr{A}_n^1\right] = 1 - \frac{q^{(n-2)^2}\left(1-q^{n-2}\right)}{1-q^{\binom{n-1}{2}}} = 1 - \frac{6e^{-6r/7}r}{7(1-e^{-3r/7})}\frac{1}{n} + o\left(\frac{1}{n}\right). \qquad (13)$$

Proof. Assume without loss of generality that df_0 changes the values of x_1 and x_2, and that these values are originally FALSE and TRUE, respectively. Also, let sf_0 be the *unique* single flip that changes a value which is also changed by df_0. In this case, sf_0 is the flip that changes the value of x_1 from FALSE to TRUE.

Notice that, because all single flips that are distinct from sf_0 change values which are not changed by df_0,

$$\Pr_p\left[A^{df_0} \not\models \phi | A \in \mathscr{A}_n^1\right] = \Pr_p\left[A^{df_0} \not\models \phi | A^{sf_0} \not\models \phi\right].$$

To compute the "negated" probability on the right-hand side of the above inequality, we proceed as follows.

It is easy to see, carrying the corresponding argument in the proof of Theorem 1 within the model $\mathscr{G}_p$, that $\Pr_p[A^{sf_0} \not\models \phi] = 1 - q^{\binom{n-1}{2}}$. We now first compute the "positive" (with respect to A^{df_0}) probability:

$$\Pr_p\left[A^{df_0} \models \phi \wedge A^{sf_0} \not\models \phi\right].$$

Observe that, in order to have that $A^{df_0} \models \phi$, any clause that contains *at least one* of the literals $\neg x_1, x_2$ and its remaining literals belong to $\neg x_i^A$, $i > 2$, *must not be* among the conjuncts of ϕ. The number of these clauses is equal to $2\binom{n-2}{2} + n - 2 = (n-2)^2$. However, the additional requirement that $A^{sf_0} \not\models \phi$, in conjunction with the requirement that $A^{df_0} \models \phi$, makes it necessary that at least one clause that contains *both* $\neg x_1, \neg x_2$ and one of $\neg x_i^A$, $i > 2$, *is* among the conjuncts of ϕ (the number of such clauses is $n-2$). The probability for these events to occur simultaneously is equal to $q^{(n-2)^2}(1-q^{n-2})$. This last expression gives the probability $\Pr_p[A^{df_0} \models \phi \wedge A^{sf_0} \not\models \phi]$.

From the above, it follows that

$$\Pr_p\left[A^{df_0} \not\models \phi | A^{sf_0} \not\models \phi\right] = 1 - \frac{q^{(n-2)^2}\left(1-q^{n-2}\right)}{1-q^{\binom{n-1}{2}}}.$$

This concluses the proof. ■

Unfortunately, we cannot just multiply the probabilities in the previous lemma to compute $\Pr_p[A \in \mathscr{A}_n^{2\sharp} | A \in \mathscr{A}_n^1]$ because these probabilities are not independent. This is so because two double flips may have variables in common. Fortunately, we can apply a variant of Suen's inequality that was proved by Janson (this inequality appears as [11, Theorem 7]; for the original Suen's inequality, see [20] or [3]) and gives an estimate for the probability of the intersection of dependent events. In what follows, we will first present the inequality as well as the assumptions under which it is applicable, and then apply it in the context of our problem.

Let $\{I_i\}_{i \in \mathscr{I}}$ be a finite family of indicator random variables defined on a probability space. Also let Γ be a dependency graph for $\{I_i\}_{i \in \mathscr{I}}$, i.e., a graph with vertex set $\mathscr{I}$ such that if A and B are two disjoint subsets of $\mathscr{I}$, and Γ contains no edge between A and B, then the families $\{I_i\}_{i \in A}$ and $\{I_i\}_{i \in B}$ are *independent*. If $p_i = \Pr[I_i = 1]$, $\Delta = \frac{1}{2}\sum_{(i,j):i \sim j}\mathbf{E}[I_i I_j]$ (summing over ordered pairs (i,j)), $\delta = \max_{i \in \mathscr{I}} \sum_{j \sim i} p_j$ and $\epsilon = \max_{i \in \mathscr{I}} p_i$ and, in addition, $\delta + \epsilon \leq e^{-1}$, then

$$\Pr\left[\sum_{i \in \mathscr{I}} I_i = 0\right] \leq e^{\Delta \phi_2(\delta+\epsilon)} \prod_{k \in \mathscr{I}} (1-p_k) \tag{14}$$

where $\phi_2(x)$ is the smallest root of the equation $\phi_2(x) = e^{x\phi_2(x)}$, given x such that $0 \leq x \leq e^{-1}$ (ϕ_2 is increasing in this range).

Now, in our context, given a truth assignment, let DF (the index set $\mathscr{I}$ above) be the class of all double flips. For an element df of $\mathscr{I}$, let $I_{df} = 1$ iff $A^{df} \models \phi$, given that $A \in \mathscr{A}_n^1$. Then, $p_{df} = \Pr_p[I_{df} = 1] = \Pr_p[A^{df} \models \phi | A \in \mathscr{A}_n^1]$, and from Lemma 5, it is equal to

$$\frac{6e^{-6r/7}r}{7(1-e^{-3r/7})}\frac{1}{n} + o\left(\frac{1}{n}\right).$$

Also, it holds that

$$\delta \leq \frac{6e^{-6r/7}r}{7(1-e^{-3r/7})} + o(1)$$

since a double flip may share a flipped variable with at most n other double flips. Also, from Lemma 5, $\epsilon = o(1)$. Therefore, for the range of r that is of concern to us ($r > 3.003$, which is the best known lower bound for the threshold; see [9]), $\delta + \epsilon \leq 6e^{-6r/7}r/(7(1-e^{-3r/7})) + o(1) \leq e^{-1}$ (for sufficiently large n).

For two elements df and df' of DF, let $df \sim df'$ denote that df and df' are distinct double flips sharing a flipped variable. Then

$$\Delta = \frac{1}{2} \sum_{(df, df'):\, df \sim df'} \mathbf{E}\left[I_{df} I_{df'} \right] = \frac{1}{2} \sum_{(df, df'):\, df \sim df'} \Pr_p\left[A^{df} \vDash \phi,\, A^{df'} \vDash \phi \middle| A \in \mathscr{A}_n^1 \right]. \tag{15}$$

Before calculating the probability that is involved in Eq. (15), we show that the events that we are considering, i.e., the events that $A^{df} \vDash \phi$, df a double flip, conditional on $A \in \mathscr{A}_n^1$, form a dependency graph. In other words, we must check whether the following property holds: for any two sets J_1 and J_2 of double flips such that no flip in J_1 shares a variable with a flip in J_2, any Boolean combination of conditional events corresponding to flips in J_1 is independent of any Boolean combination of conditional events corresponding to flips in J_2. Suppose that the conditional was not $A \in \mathscr{A}_n^1$, but $A \vDash \phi$. Then the resulting space is a $\mathscr{G}_p$ space, i.e., each clause satisfied by A has an independent probability to appear in the random formula. Then the mutual independence required to obtain the above inequality would be obviously satisfied, as the two Boolean combinations that must be shown independent correspond to distinct clauses. In our case, however, where the conditional is $A \in \mathscr{A}_n^1$, the probability space is not a $\mathscr{G}_p$ space. Nevertheless, the required independence still holds. To prove this, let B_1 and B_2 be two Boolean combinations of *unconditional* events corresponding to two sets of double flips that do not share a variable. The *conditional* independence that is required to obtain the above inequality is equivalent to

$$\Pr_p\left[B_1, B_2, \bigwedge_{sf \in SF} A^{sf} \nvDash \phi \middle| A \vDash \phi \right] \cdot \Pr_p\left[\bigwedge_{sf \in SF} A^{sf} \nvDash \phi \middle| A \vDash \phi \right]$$
$$= \Pr_p\left[B_1, \bigwedge_{sf \in SF} A^{sf} \nvDash \phi \middle| A \vDash \phi \right] \cdot \Pr_p\left[B_2, \bigwedge_{sf \in SF} A^{sf} \nvDash \phi \middle| A \vDash \phi \right].$$

Notice that, because the conditional in the probabilities in the above equality is $A \vDash \phi$, the resulting space is from the model $\mathscr{G}_p$. Now, the above equation is trivial to prove using the fact that, in such a space, combinations of events corresponding to either single or double flips with no common variables are independent.

We now compute the exponential correlation factor that appears in inequality (14). The computation is a bit tedious. In the following lemmata, we give the results of the various steps, hoping that the interested (and patient) reader can carry them out by herself. The method to be used is very similar to that of the proof Lemma 5. In order to save a little more on notation, we set

$$u = e^{-r/7}.$$

Notice that, then, by Lemma 5,

$$\Pr_p\left[A^{df_0} \vDash \phi \middle| A \in \mathscr{A}_n^1 \right] = \frac{6u^6 \ln(1/u)}{1 - u^3} \frac{1}{n} + o\left(\frac{1}{n} \right). \tag{16}$$

Lemma 6. *Let df_0 and df_1 be two double flips that share the variable that they change from* FALSE *to* TRUE. *Then*

$$\Pr_p\left[A^{df_0} \models \phi,\, A^{df_1} \models \phi \middle| A \in \mathscr{A}_n^1\right] = \frac{q^{2(n-2)} q^{3\binom{n-2}{2}} q^{n-3} p}{1 - q^{\binom{n-1}{2}}}$$

$$= \frac{6u^9 \ln(1/u)}{1-u^3} \frac{1}{n^2} + o\left(\frac{1}{n^2}\right). \tag{17}$$

Lemma 7. *Let df_0 and df_1 be two double flips that share the variable that they change from* TRUE *to* FALSE. *Then*

$$\Pr_p\left[A^{df_0} \models \phi,\, A^{df_1} \models \phi \middle| A \in \mathscr{A}_n^1\right] = \frac{q^{2(n-2)} q^{3\binom{n-2}{2}} q^{n-3} \left(1 - q^{n-2}\right)^2}{\left(1 - q^{\binom{n-1}{2}}\right)^2}$$

$$= \frac{36u^9 \ln^2(1/u)}{(1-u^3)^2} \frac{1}{n^2} + o\left(\frac{1}{n^2}\right). \tag{18}$$

Now, observe that the number of intersecting ordered pairs of double flips is at most $df(A) \cdot n$. Finally, it is easy to see that the probability in Lemma 6 is smaller than the probability in Lemma 7. From these observations, and by substituting in Eq. (15) the right-hand side of Eq. (18), we get that

$$\Delta = \frac{1}{2} \sum_{(df,\, df'):\, df \sim df'} \mathbf{E}\left[I_{df} I_{df'} \right]$$
$$\le df(A) \cdot \left(\frac{18u^9 \ln^2(1/u)}{(1-u^3)^2} \frac{1}{n} + o\left(\frac{1}{n}\right) \right) \phi_2\left(\frac{6u^6 \ln(1/u)}{1-u^3} + o(1) \right).$$

From this and Eq. (16), it follows, by inequality (14), that

$$\Pr_p\left[A \in \mathscr{A}_n^{2\sharp} \middle| A \in \mathscr{A}_n^1\right] \le \left[1 - \frac{6u^6 \ln(1/u)}{1-u^3} \frac{1}{n} + \frac{18u^9 \ln^2(1/u)}{(1-u^3)^2} \frac{1}{n} \times \phi_2\left(\frac{6u^6 \ln(1/u)}{1-u^3} + o(1) \right) + o\left(\frac{1}{n}\right) \right]^{df(A)}. \tag{19}$$

It is easy to see (e.g., by using Maple, or by a bit tedious analytical computations) that the expression at the base of the right-hand side of the above inequality is at most 1 for $3 \le r \le 5$. Now, by Eqs. (10)–(12), and (19), we get that

$$\mathbf{E}\left[\left| \mathscr{A}_n^{2\sharp} \right| \right] \le 3(rn)^{1/2} (7/8)^{rn} \sum_A X^{sf(A)} Y^{df(A)} \tag{20}$$

where

$$X = 1 - u^3 + o(1) \tag{21}$$

and

$$Y = 1 - \frac{6u^6 \ln(1/u)}{1-u^3}\left(1 - \frac{3u^3 \ln(1/u)}{1-u^3} \cdot \phi_2\left(\frac{6u^6 \ln(1/u)}{1-u^3} + o(1)\right)\right)\frac{1}{n} + o\left(\frac{1}{n}\right). \tag{22}$$

In the next subsection, we give an estimate for the sum in inequality (20).

3.2. Estimates

Lemma 8. *If* $0 \le X^2 \le Y \le 1$, *then*

$$\sum_A X^{sf(A)} Y^{df(A)} \le \prod_{i=0}^{n-1} (1 + XY^{i/2}). \tag{23}$$

Notice that, in our case, the condition $X^2 \le Y$ holds, as by Eqs. (21) and (22), we have that $Y = 1 - o(1)$ and $X = 1 - u^3 + o(1)$. The easiest way to prove the inequality in the lemma is first to show that

$$\sum_A X^{sf(A)} Y^{df(A)} = \sum_{k=0}^{n} \binom{n}{k}_Y X^k$$

where

$$\binom{n}{k}_q = \frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-k+1})}{(1-q^k)(1-q^{k-1})\cdots(1-q^1)}$$

are the so-called q-binomial or Gauss coefficients (see Knuth's book [16, p. 64]), and then proceed inductively on n. Complete information on such techniques can be found in a book on basic hypergeometric (or q-hypergeometric) series by Gasper and Rahman [10]. A direct proof is also possible, but it is rather involved. We do not give the details, as they do not offer anything new to our problem (for a proof, see [14]).

Now, recall that

$$u = e^{-r/7} \tag{24}$$

$$X = 1 - u^3 + o(1) \tag{25}$$

and

$$Y = 1 - \frac{6u^6 \ln(1/u)}{1-u^3}\left(1 - \frac{3u^3 \ln(1/u)}{1-u^3} \cdot \phi_2\left(\frac{6u^6 \ln(1/u)}{1-u^3} + o(1)\right)\right)\frac{1}{n} + o\left(\frac{1}{n}\right). \tag{26}$$

Also set $Z = n \ln Y$, and observe that from Eq. (26), it follows that

$$Z = \frac{6u^6 \ln(u)}{1-u^3}\left(1 + \frac{3u^3 \ln(u)}{1-u^3} \cdot \phi_2\left(\frac{6u^6 \ln(1/u)}{1-u^3} + o(1)\right)\right) + o(1). \tag{27}$$

Our estimate for $\mathbf{E}[|\mathscr{A}_n^{2\sharp}|]$ will be given in terms of the dilogarithm function (see the book by Abramowitz and Stegun [1]) which is defined as

$$\text{dilog}(x) = -\int_1^x \frac{\ln(t)}{t-1}\, dt.$$

Finally, let $df\_eq(r)$ be the expression that we get if we substitute in

$$\ln(7/8) r(Z/2) + \text{dilog}(1+X) - \text{dilog}(1 + Xe^{Z/2})$$

the values of X and Z without their asymptotic terms, and then set $u = e^{-r/7}$ (it will shortly become clear why we introduce the above expression of X, Z, and r).

We now state the concluding result.

Theorem 3. *If $df\_eq(r) > 0$, then $\lim_n \mathbf{E}[|\mathscr{A}_n^{2\sharp}|] = 0$, and therefore*

$$\lim_n \Pr[\phi \textit{ is satisfiable}] = 0.$$

It follows that $\kappa < 4.601+$.

Proof. From inequalities (20) and (23), we conclude that, in order to have

$$\lim_{n\to\infty} \mathbf{E}\left[\left|\mathscr{A}_n^{2\sharp}\right|\right] = 0,$$

it is sufficient to show that the expression

$$3(rn)^{1/2}(7/8)^{rn}\left(\prod_{i=0}^{n-1}(1 + XY^{i/2})\right)$$

converges to zero. Raising this last expression to the power $1/n$, then taking the logarithm, and finally making the standard approximation of a sum by an integral (for the case of a decreasing function), we conclude that a sufficient condition for $\lim_n \mathbf{E}[|\mathscr{A}_n^{2\sharp}|] = 0$ is that

$$r \ln(7/8) + \lim_n \left((1/n)\int_{-1}^{n-1} \ln(1 + XY^{\tau/2})\, d\tau\right) < 0.$$

However,

$$\int \ln(1 + XY^{\tau/2})\, d\tau = -\frac{\mathrm{dilog}(1 + XY^{\tau/2})}{\ln(Y^{1/2})}.$$

The first assertion of the theorem now follows by elementary calculations taking into account that $Y^{n/2} = e^{Z/2}$ and $Y = 1 + o(1)$. The second assertion follows by Lemma 3. The estimate for κ is obtained by computing the unique positive solution of the equation $df\_eq(r) = 0$. We obtained the value 4.601+ by using Maple [18]. ■

4. DISCUSSION

Our technique can be extended to triple, or even higher order, flips. To do that, first observe that

$$\mathbf{E}\left[\left|\mathscr{A}_n^l\right|\right] = (7/8)^{rn} \sum_{A \in \mathscr{A}_n} \Pr\left[A \in \mathscr{A}_n^1 \middle| A \vDash \phi\right] \cdot \Pr\left[A \in \mathscr{A}_n^2 \middle| A \in \mathscr{A}_n^1\right] \cdots \Pr\left[A \in \mathscr{A}_n^l \middle| A \in \mathscr{A}_n^{l-1}\right],$$

and then obtain upper bounds for the factors in the terms of the above sum. Thus, we can get increasingly better estimates of κ. Furthermore, if r_l is the infimum of the values of r that make $\lim_n \mathbf{E}[\mathscr{A}_n^l] = 0$, we conjecture that $\lim_l r_l = \kappa$. The equation $\Pr[\phi$ is satisfiable$] = \mathbf{E}[|\mathscr{A}_n^n|]$ of Lemma 3 is an indication that this is indeed so.

Finally, observe that the estimate obtained by fixed-order flips can be possibly improved further if, instead of the Markov-type inequalities in Lemma 3, we use a "harmonic mean formula." To be specific, first notice that the following result can be easily proved in exactly the same way as the original harmonic mean formula given by Aldous [2].

Proposition 1. *For every $l \geq 0$,*

$$\Pr[\textit{the random formula is satisfiable}] = \sum_A \left(\Pr\left[A \in \mathscr{A}_n^l\right] \cdot \mathbf{E}\left[\frac{1}{\left|\mathscr{A}_n^l\right|} \middle| A \in \mathscr{A}_n^l\right]\right).$$

Proof. Let I_ϕ be the indicator variable defined in Eq. (6) of the proof of Lemma 1. Let also I_ϕ^A be the following indicator variable (with the random ϕ—not the nonrandom A—as its argument):

$$I_\phi^A = \begin{cases} 1 & \text{if } A \in \mathscr{A}_n^l \\ 0 & \text{otherwise.} \end{cases}$$

Now, observe that

$$
\begin{aligned}
&\Pr[\text{the random formula is satisfiable}]\\
&\quad = \sum_{\phi} \left(\Pr[\phi]\cdot I_{\phi}\right)\\
&\quad = \sum_{\phi \text{ is SAT}} \left(\Pr[\phi]\cdot \sum_{A} \frac{I_{\phi}^{A}}{|\mathscr{A}_n^l|}\right)\\
&\quad = \sum_{A} \left(\Pr[A \in \mathscr{A}_n^l]\cdot \sum_{\phi \text{ is SAT}} \frac{\Pr[\phi | A \in \mathscr{A}_n^l]}{|\mathscr{A}_n^l|}\right)\\
&\quad = \sum_{A} \left(\Pr[A \in \mathscr{A}_n^l]\cdot \mathbf{E}\left[\frac{1}{|\mathscr{A}_n^l|}\middle| A \in \mathscr{A}_n^l\right]\right). \qquad \blacksquare
\end{aligned}
$$

It is now conceivable that the techniques introduced by Kamath et al. in [12] can be applied to estimate $\mathbf{E}[1/|\mathscr{A}_n^l|\,|A \in \mathscr{A}_n^l]$ for an arbitrary fixed $A \in \mathscr{A}_n^l$. Kamath et al. give such an estimate for the case $l = 0$. The generalization at least to the case $l = 1$ should be possible. Given that in Section 2 we have computed the probability $\Pr[A \in \mathscr{A}_n^1]$, such a generalization in conjunction with the above proposition would improve the single-flips estimate.

ACKNOWLEDGMENTS

Several people read earlier drafts of this paper and made valuable suggestions for improvements in the presentation. Some found mistakes (not all of which were trivial), and suggested ways to avoid them. In alphabetical order, they are: D. Achlioptas, two anonymous referees, A. Kaporis, M. Molloy, P. Savický, and P. Spirakis. We are grateful to them. We thank Mizan Rahman for the very informative discussions concerning the sum $\sum_A X^{sf(A)} Y^{df(A)}$ and its relation to basic hypergeometric series. We also thank Svante Janson for his suggestion to use his improved version of Suen's inequality in the subsection on double flips.

REFERENCES

[1] M. Abramowitz and I. E. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th printing, U.S. Department of Commerce, National Bureau of Standards, Washington, DC, 1972.

[2] D. J. Aldous, The harmonic mean formula for probabilities of unions: Applications to sparse random graphs, *Discrete Math.*, **76**, 167–176 (1989).

[3] N. Alon, J. H. Spencer, and P. Erdös, *The Probabilistic Method*, John Wiley & Sons, New York, 1992.

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

[5] V. Chvátal and B. Reed, Mick gets some (the odds are on his side), *Proceedings of the Thirty-Third IEEE Symposium on Foundations of Computer Science*, 1992, pp. 620–627.

[6] V. Chvátal and E. Szemerédi, Many hard examples for resolution, *J. Assoc. Comput. Mach.*, **35**, 759–768 (1988).

[7] O. Dubois and Y. Boufkhad, A general upper bound for the satisfiability threshold of random r-SAT formulae, preprint, LAFORIA, CNRS-Université Paris 6, 1996.

[8] J. Franco and M. Paull, Probabilistic analysis of the Davis-Putnam procedure for solving the satisfiability problem, *Discrete Appl. Math.*, **5**, 77–87, 1983.

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

[10] G. Gasper and M. Rahman, *Basic Hypergeometric Series, Encyclopedia of Mathematics and Its Applications*, Cambridge University Press, Cambridge, 1990, Vol. 35.

[11] S. Janson, New versions of Suen's correlation inequality, unpublished manuscript.

[12] A. Kamath, R. Motwani, K. Palem, and P. Spirakis, Tail bounds for occupancy and the satisfiability threshold conjecture, *Random Struct. Alg.*, **7**, 59–80 (1995); also, *Proceedings of the Thirty-Fifth FOCS*, IEEE, 1994, pp. 592–603.

[13] S. Kirkpatrick and B. Selman, Critical behavior in the satisfiability of random Boolean expressions, *Science*, **264**, 1297–1301 (1994).

[14] L. Kirousis, E. Kranakis, and D. Krizanc, An upper bound for a basic hypergeometric series, Technical Report TR-96-07, School of Computer Science, Carleton University, Canada, 1996.

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

[16] D. Knuth, *Fundamental Algorithms, The Art of Computer Programming*, 2nd ed., Addison-Wesley, Reading, MA, 1973, Vol. 1.

[17] C. McDiarmid, On a correlation inequality of Farr, *Combinatorics, Prob., Comput.*, **1**, 157–160 (1992).

[18] D. Redfern, *The Maple Handbook: Maple V Release 3*, Springer-Verlag, New York, 1994.

[19] J.-C. Simon, J. Carlier, O. Dubois, and O. Moulines, Étude statistique de l'existence de solutions de problèmes SAT, application aux systèmes-experts, *C. R. Acad. Sci. Paris. Sér. I Math.*, **302**, 283–286 (1986).

[20] W. C. Suen, A correlation inequality and a Poisson limit theorem for nonoverlapping balanced subgraphs of a random graph, *Random Struct. Alg.*, **1**, 231–242 (1990).