Approximating the Unsatisfiability Threshold of Random Formulas[*]

Lefteris M. Kirousis[†][§]
(kirousis@fryni.ceid.upatras.gr)

Evangelos Kranakis[‡][§]
(kranakis@scs.carleton.ca)

Danny Krizanc[‡][§]
(krizanc@scs.carleton.ca)

Yannis C. Stamatiou[†][§]
(stamatiu@fryni.ceid.upatras.gr)

[†] University of Patras, Department of Computer Engineering and Informatics, Rio, 265 00 Patras, Greece.

[‡] Carleton University, School of Computer Science, Ottawa, ON K1S 5B6, Canada.

[§] Research performed while the first author was visiting the School of Computer Science of 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 author 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 author were also partially supported by EU ESPRIT Long-term Research Project ALCOM-IT (Project nr. 20244).

[*] An extended abstract of the paper has appeared 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 co-authored by the first three authors of the present paper.

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, Motwani, Palem, and Spirakis 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 4.598. In general, by letting the 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 $k > 3$.

AMS Mathematics Subject Classification (1991): *Primary* 68Q25, *Secondary* 03B05, 60B99.
ACM Computing Classification System (1991): F.2.2, F.4.1.
Key Words and Phrases: Approximations, Probabilistic method, Random formulas, Satisfiability, Threshold point.
Carleton University, School of Computer Science: TR-96-27

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 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 $\mathcal{A}_n$ be the set of all truth assignments on the n variables $x_1, \ldots x_n$, and let $\mathcal{S}_n$ be the set of truth assignments that satisfy the random formula ϕ. The cardinality $|\mathcal{S}_n|$ is thus a random variable. Also, for an instantiation ϕ of the random formula, let $|\mathcal{S}_n(\phi)|$ 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., $\mathbf{E}[|\mathcal{S}_n|]$, satisfies the following relation

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

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

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

where

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

From equations (1) and (2) the following Markov's inequality follows immediately:

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

It is easy to find a condition on κ under which $\mathbf{E}[|\mathcal{S}_n|]$ 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, as in the right-hand side of equation (1) we may have small probabilities multiplied with large cardinalities, 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 $\mathcal{S}_n$ 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, Motwani, Palem, and Spirakis [12] of 4.758, which was obtained by non-elementary 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 non-computational proof is equal to 4.87. Dubois and Boufkhad [7] have independently obtained the upper bound of 4.64 by a method similar to the one presented in Section 2.

In Section 3 we show how to further improve the bound to 4.598 by defining an even smaller subset of $\mathcal{S}_n$. This is achieved by increasing the range of locality when selecting the local maxima that represent $\mathcal{S}_n$. Actually, we define a decreasing sequence of subsets of $\mathcal{S}_n$ by selecting from $\mathcal{S}_n$ 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.

Moreover, 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 apply the technique that is used in Kamath et al. [12]. This is discussed in the last section. Finally, our method readily generalizes to k-SAT, for $k > 3$.

2 Single Flips

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

DEFINITION **1** *For a random formula ϕ, $\mathcal{S}_n^\sharp$ is defined to be the random class of truth assignments A such that (i) $A \models \phi$, 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 its values equal to TRUE vacuously satisfies condition (ii) of the previous definition. Consider the lexicographic ordering among truth assignments, where, as usual, 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 $\mathcal{S}_n^\sharp$ is the set of elements of $\mathcal{S}_n$ 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:

LEMMA **1** *The following Markov type inequality holds for $\mathcal{S}_n^\sharp$:*

$$\mathbf{Pr}[\text{the random formula is satisfiable}] \leq \mathbf{E}[|\mathcal{S}_n^\sharp|]. \tag{5}$$

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

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

where

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

On the other hand,

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

The lemma now immediately follows from the above. □

We also have the following:

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

$$\mathbf{E}[|\mathcal{S}_n^\sharp|] = (7/8)^{rn} \textstyle\sum_{A \in \mathcal{A}_n} \mathbf{Pr}[A \in \mathcal{S}_n^\sharp \mid A \in \mathcal{S}_n]. \tag{7}$$

Proof First observe that the random variable $|\mathcal{S}_n^\sharp|$ is the sum of indicator variables and then condition on $A \models \phi$ (recall, 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 that

Theorem 1 *The expected value $\mathbf{E}[|\mathcal{S}_n^\sharp|]$ 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 $A \models \phi$. Observe that the assumption that $A \models \phi$ excludes $\binom{n}{3}$ clauses from the conjuncts of ϕ, i.e., there remain $7\binom{n}{3}$ clauses to choose the conjuncts of ϕ from. Consider now 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 $A \models \phi$, in order to have that $A^{sf_0} \not\models \phi$, 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 $A^{sf_0} \not\models \phi$ is $\binom{n-1}{2}/7\binom{n}{3} = 3/(7n)$. Therefore, the probability that $A^{sf_0} \not\models \phi$ (given that $A \models \phi$) 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 $A \models \phi$) refer to disjoint sets of clauses. Therefore, as it is intuitively clear, the dependencies among them are such that:

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

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) $V = \{1, \ldots, m\}$, (ii) $I = \{1, \ldots, sf(A)\}$, (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, $\mathcal{F}_i$ 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 equation (7) and by Newton's binomial formula, $\mathbf{E}[|\mathcal{S}_n^\sharp|]$ 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 $\mathbf{E}[|\mathcal{S}_n^\sharp|]$ 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 it 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 $\mathbf{E}[|\mathcal{S}_n^\sharp|]$ 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*

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

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 $\mathcal{S}_n$. We start with a definition:

DEFINITION **2** *Given a random formula ϕ and a nonnegative integer l, $\mathcal{A}_n^l (l \leq n)$ 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 $\mathcal{S}_n$ of the previous section, i.e., the class of truth assignments satisfying the random formula is, in the notation of the previous definition, equal to $\mathcal{A}_n^0$, whereas $\mathcal{S}_n^\sharp$ is equal to $\mathcal{A}_n^1$. In general, $\mathcal{A}_n^l$ is the subset of $\mathcal{S}_n$ 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, obviously, $\mathcal{A}_n^l$ is a sequence of classes which is decreasing relative to l (with respect to set inclusion).

Now, exactly as in Lemma 1, it can be proved that:

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

$$\mathbf{Pr}[\phi \text{ is satisfiable}] = \mathbf{E}[|\mathcal{A}_n^n|] \leq \mathbf{E}[|\mathcal{A}_n^{n-1}|] \leq \cdots \leq \mathbf{E}[|\mathcal{A}_n^1|] \leq \mathbf{E}[|\mathcal{A}_n^0|]. \tag{9}$$

It follows from the above that for a fixed l, by letting $\lim_n \mathbf{E}[|\mathcal{A}_n^l|] = 0$, we obtain upper bounds for κ which decrease as l increases. In other words, if r_l denotes the infimum of the values of r that make the expression $\mathbf{E}[|\mathcal{A}_n^l|]$ converge to zero (as $n \to \infty$), then r_l is an upper bound for κ, and the larger l is, the better the bound. 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 strictly larger 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 $\mathcal{A}_n^{2\sharp}$ be the set of assignments A such that $A \models \phi$ and for all single flips sf, $A^{sf} \not\models \phi$ and for all double flips df, $A^{df} \not\models \phi$. It can be easily seen that $\mathcal{A}_n^2$ is a subset of $\mathcal{A}_n^{2\sharp}$ (in general a proper one, because in the definition of $\mathcal{A}_n^{2\sharp}$ we did not take into account the changes of kind (2)). Therefore a value of r that makes the expected value $\mathbf{E}[|\mathcal{A}_n^{2\sharp}|]$ converge to zero is, by Lemma 3, an upper bound for κ. Actually, it can be proved that both $\mathbf{E}[|\mathcal{A}_n^{2\sharp}|]$ and $\mathbf{E}[|\mathcal{A}_n^2|]$ 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

Lemma **4**

$$\mathbf{E}[|\mathcal{A}_n^{2\sharp}|] = (7/8)^{rn} \sum_{A \in \mathcal{A}_n} \mathbf{Pr}[A \in \mathcal{A}_n^1 \mid A \models \phi] \cdot \mathbf{Pr}[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1]. \tag{10}$$

Therefore, by the remarks in 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 equality above converges to zero. We will do this in two steps. First we will compute an upper bound for each term of the sum in the equality above; then we will find an upper bound for $\mathbf{E}[|\mathcal{A}_n^{2\sharp}|]$ 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 Janson's inequality [11], 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 $\mathbf{E}[|\mathcal{A}_n^{2\sharp}|]$ are carried out in the second subsection.

3.1 Probability Calculations

In this subsection, given a fixed A, we find an upper bound for $\mathbf{Pr}[A \in \mathcal{A}_n^1 \mid A \models \phi] \cdot \mathbf{Pr}[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1]$. Because, by definition, $\mathcal{A}_n^1 \supset \mathcal{A}_n^{2\sharp}$, we have

$$\mathbf{Pr}[A \in \mathcal{A}_n^1 \mid A \models \phi] \cdot \mathbf{Pr}[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1] = \mathbf{Pr}[A \in \mathcal{A}_n^{2\sharp} \mid A \models \phi]. \tag{11}$$

We assume for the rest of this subsection that the condition $A \models \phi$ 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 $A \models \phi$, the event $A \in \mathcal{A}_n^1$ is equivalent to the statement that for any single flip sf, $A^{sf} \not\models \phi$. In the sequel, all computations of probabilities, analyses of events, etc. will be carried out assuming that $A \models \phi$, usually without explicit mention of it.

To compute $\mathbf{Pr}[A \in \mathcal{A}_n^{2\sharp}]$, 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 $\mathcal{G}_p$ be the model for random formulas where each clause has an independent probability p to appear in the formula, let $\mathcal{G}_m$ be the model where the random formula is obtained by uniformly and independently selecting m clauses *without* replacement, and, finally, let $\mathcal{G}_{mm}$ 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 $\mathcal{G}_p$ ($\mathcal{G}_m$) will be denoted by $\mathbf{Pr}_p[E]$ ($\mathbf{Pr}_m[E]$, respectively). In order not to change our notation, we continue to denote the probability of E in the model $\mathcal{G}_{mm}$ by $\mathbf{Pr}[E]$. Set $p = m/(7\binom{n}{3}) \sim 6r/(7n^2)$. By Theorem 2 (iii) of Chapter 3 (page 35) of Bollobás' book [4], we have that for any property Q of formulas, $\mathbf{Pr}_m[Q] \leq 3m^{1/2}\mathbf{Pr}_p[Q]$. 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 $\mathbf{Pr}[Q] \leq \mathbf{Pr}_m[Q]$. 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 non-satisfiability is both monotonically increasing and reducible, we conclude that

$$\begin{aligned} \mathbf{Pr}[A \in \mathcal{A}_n^{2\sharp}] &\leq 3m^{1/2}\mathbf{Pr}_p[A \in \mathcal{A}_n^{2\sharp}] \\ &= 3m^{1/2}\mathbf{Pr}_p[A \in \mathcal{A}_n^{2\sharp} \wedge A \in \mathcal{A}_n^1] \text{ (because } A \in \mathcal{A}_n^1 \text{ is implied by } A \in \mathcal{A}_n^{2\sharp}) \\ &= 3m^{1/2}\mathbf{Pr}_p[A \in \mathcal{A}_n^1] \cdot \mathbf{Pr}_p[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1]. \end{aligned} \tag{12}$$

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

$$\mathbf{Pr}_p[A \in \mathcal{A}_n^1] = \left(1 - (1-p)^{\binom{n-1}{2}}\right)^{sf(A)} \sim (1 - e^{-3r/7})^{sf(A)}. \qquad (13)$$

So, by equations (11), (12), and (13) to find an upper bound for $\mathbf{Pr}[A \in \mathcal{A}_n^1] \cdot \mathbf{Pr}[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1]$, it is enough to find an upper bound for $\mathbf{Pr}_p[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1]$. 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:

LEMMA **5** *The following holds:*

$$\mathbf{Pr}_p[A^{df_0} \not\models \phi \mid A \in \mathcal{A}_n^1] = 1 - \frac{q^{(n-2)^2}(1-q^{n-2})}{1 - q^{\binom{n-1}{2}}} = 1 - \frac{6e^{-6r/7}r}{7(1-e^{-3r/7})}\frac{1}{n} + o\left(\frac{1}{n}\right). \qquad (14)$$

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,

$$\mathbf{Pr}_p[A^{df_0} \not\models \phi \mid A \in \mathcal{A}_n^1] = \mathbf{Pr}_p[A^{df_0} \not\models \phi \mid A^{sf_0} \not\models \phi].$$

To compute the "negated" probability in 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 $\mathcal{G}_p$, that $\mathbf{Pr}_p[A^{sf_0} \not\models \phi] = 1 - q^{\binom{n-1}{2}}$. We now first compute the "positive" (with respect to A^{df_0}) probability:

$$\mathbf{Pr}_p[A^{df_0} \models \phi \wedge A^{sf_0} \not\models \phi].$$

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 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 $\mathbf{Pr}_p[A^{df_0} \models \phi \wedge A^{sf_0} \not\models \phi]$.

From the above, it follows that

$$\mathbf{Pr}_p[A^{df_0} \not\models \phi \mid A^{sf_0} \not\models \phi] = 1 - \frac{q^{(n-2)^2}(1-q^{n-2})}{1-q^{\binom{n-1}{2}}}.$$

This concludes the proof. □

Unfortunately, we cannot just multiply the probabilities in the previous lemma to compute $\mathbf{Pr}_p[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1]$, because these probabilities are not independent. This is so because two double flips may have variables in common. Fortunately, we can apply Janson's inequality [11] that gives an estimate for the probability of the intersection of dependent events. For a detailed presentation of this theorem we refer to the 2nd edition of Spencer's book [20].

Let DF be the class of all double flips. 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. Let $\Delta = \sum_{df \sim df'} \mathbf{Pr}_p[A^{df} \models \phi \wedge A^{df'} \models \phi \,|\, A \in \mathcal{A}_n^1]$, and finally let $\epsilon \geq \mathbf{Pr}_p[A^{df} \models \phi \,|\, A \in \mathcal{A}_n^1], \forall df \in DF$. We claim that if $r = m/n$ exceeds 3 (more accurately, 2.93—these values of r are within the range that is of interest to us, because it is known [9] that for $r < 3.003$, ϕ is asymptotically almost certainly satisfiable), then:

$$\mathbf{Pr}_p[\bigwedge_{df \in DF} A^{df} \not\models \phi \,|\, A \in \mathcal{A}_n^1] \leq$$

$$\left(\prod_{df \in DF} \mathbf{Pr}_p[A^{df} \not\models \phi \,|\, A \in \mathcal{A}_n^1]\right) \cdot e^{\Delta/[2(1-\epsilon)]}. \quad \text{(Janson's inequality)}$$

Consider two arbitrary fixed double flips df_0 and df_1. Also, for notational convenience, for any (conditional or unconditional) event E, let $\mathbf{Pr}_p^{A \in \mathcal{A}_n^1}[E]$ denote $\mathbf{Pr}_p[E \mid A \in \mathcal{A}_n^1]$. According to [20], in order to prove Janson's inequality, it suffices to prove that double flips that do not share a flipped variable are mutually independent under the condition $A \in \mathcal{A}_n^1$ and that the following two correlation inequalities hold:
(a) For all $I \subset DF$ with $df_0 \notin I$,

$$\mathbf{Pr}_p^{A \in \mathcal{A}_n^1}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \leq \mathbf{Pr}_p^{A \in \mathcal{A}_n^1}[A^{df_0} \models \phi]. \tag{15}$$

(b) For all $I \subset DF$ with $df_0, df_1 \notin I$,

$$\mathbf{Pr}_p^{A \in \mathcal{A}_n^1}[A^{df_0} \models \phi \wedge A^{df_1} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \leq \mathbf{Pr}_p^{A \in \mathcal{A}_n^1}[A^{df_0} \models \phi \wedge A^{df_1} \models \phi]. \tag{16}$$

Observe that in contrast to the condition $A \models \phi$, which simply excludes some clauses from the probability space, the condition $A \in \mathcal{A}_n^1$, when assumed, does not give rise to a probability space that belongs to the model $\mathcal{G}_p$. Nevertheless, double flips that do not share a flipped variable remain mutually independent even under the condition $A \in \mathcal{A}_n^1$. This follows immediately by writing the conditional probabilities as ratios of unconditional ones (see also the Appendix). However, the aforementioned correlation inequalities cannot be trivially reduced to the corresponding ones without the condition $A \in \mathcal{A}_n^1$, neither can they be proved by the usual method of the FKG inequality (for a presentation of this inequality and its uses, see [3]). Nevertheless, they still hold when $r = m/n \geq 3$. Indeed, if we assume that I is a singleton and that the double flip in it shares with df_0 the variable they flip from FALSE to TRUE, then by direct computations of the probabilities, carried out as in Lemma 5 (see also the subsequent Lemma 6), we conclude that inequality (15) holds when the value of $r = m/n$ exceeds 3. If the unique double flip in the singleton I shares with df_0 the variable they flip from TRUE to false FALSE, or if they share no variable, then again by similar probability computations, we conclude that inequality (15) holds for any value of r. The formal proofs for the general case of both (15) and (16) are given in the Appendix.

We now conclude, making use of Janson's inequality, that:

$$\mathbf{Pr}_p[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1] \leq \left(\mathbf{Pr}_p[A^{\mathit{df}_0} \not\models \phi \mid A \in \mathcal{A}_n^1]\right)^{\mathit{df}(A)} \cdot e^{\Delta/[2(1-\epsilon)]}, \tag{17}$$

where Δ and ϵ are defined above. By equation (14), $\epsilon = o(1)$, so we ignore ϵ. The computation of Δ is a bit tedious. In the following lemmata, we give the results of the various steps in this computation, 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

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

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

$$\mathbf{Pr}_p[A^{\mathit{df}_0} \models \phi, A^{\mathit{df}_1} \models \phi \mid A \in \mathcal{A}_n^1] = \frac{q^{2(n-2)} q^{3\binom{n-2}{2}} q^{n-3} p}{1 - q^{\binom{n-1}{2}}} = \frac{6u^9 \ln(1/u)}{1-u^3} \frac{1}{n^2} + o\left(\frac{1}{n^2}\right). \tag{18}$$

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

$$\begin{aligned} \mathbf{Pr}_p[A^{\mathit{df}_0} \models \phi, A^{\mathit{df}_1} \models \phi \mid A \in \mathcal{A}_n^1] &= \frac{q^{2(n-2)} q^{3\binom{n-2}{2}} q^{n-3} (1-q^{n-2})^2}{1 - q^{\binom{n-1}{2}}} \\ &= \frac{36u^9 \ln^2(1/u)}{(1-u^3)^2} \frac{1}{n^2} + o\left(\frac{1}{n^2}\right). \end{aligned} \tag{19}$$

Now observe that the number of pairs of flips described in Lemma 6 is at most $df(A) \cdot (n - sf(A))$, while the number of pairs described in Lemma 7 is at most $df(A) \cdot sf(A)$. Also, it is easy to see that the probability in Lemma 6 is smaller than the probability in Lemma 7. Therefore, we obtain the estimate:

$$\Delta \leq df(A) \cdot \left(\frac{36u^9 \ln^2(1/u)}{(1-u^3)^2} \frac{1}{n} + o\left(\frac{1}{n}\right) \right).$$

From this, by inequality (17), it follows that:

$$\begin{aligned} &\mathbf{Pr}_p[A \in \mathcal{A}_n^{2\sharp} \mid A \in \mathcal{A}_n^1] \leq \\ &\quad \left(1 - \frac{6u^6 \ln(1/u)}{1-u^3} \frac{1}{n} + \frac{18u^9 \ln^2(1/u)}{(1-u^3)^2} \frac{1}{n} + o\left(\frac{1}{n}\right)\right)^{df(A)}. \end{aligned} \tag{20}$$

It is easy to see that the expression at the base of the right-hand side of the above inequality is at most 1, for $u \in (0,1)$. Now, by equations (10), (11), (12), (13), and (20), we get that:

$$\mathbf{E}[|\mathcal{A}_n^{2\sharp}|] \leq 3(rn)^{1/2}(7/8)^{rn} \sum_A X^{sf(A)} Y^{df(A)}, \tag{21}$$

where

$$X = 1 - u^3 + o(1) \tag{22}$$

and

$$Y = 1 - \frac{6u^6 \ln(1/u)}{1-u^3} \left(1 - \frac{3u^3 \ln(1/u)}{1-u^3}\right) \frac{1}{n} + o\left(\frac{1}{n}\right). \tag{23}$$

In the next subsection, we give an estimate for the sum in inequality (21).

3.2 Estimates

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

$$\sum_A X^{sf(A)} Y^{df(A)} \leq \prod_{i=0}^{n-1} (1 + XY^{i/2}). \tag{24}$$

Notice that in our case the condition $X^2 \leq Y$ holds, as by equations (22) and (23), 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-nomial or Gauss coefficients (see Knuth's book [16], page 64), and then proceed inductively on n. Complete information on such techniques can be found in a book on basic 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:

$$u = e^{-r/7}, \tag{25}$$

$$X = 1 - u^3 + o(1), \tag{26}$$

and

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

Set also $Z = n \ln Y$ and observe that from equation (27) it follows that:

$$Z = -\frac{6u^6 \ln(1/u)}{1-u^3}\left(1 - \frac{3u^3\ln(1/u)}{1-u^3}\right) + o(1). \tag{28}$$

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

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

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

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

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 $\lim_n \mathbf{E}[|\mathcal{A}_n^{2\sharp}|] = 0$, and therefore*

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

It follows that $\kappa < 4.598$.

Proof From inequalities (21) and (24), we conclude that in order to have

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

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 $\lim_n \mathbf{E}[|\mathcal{A}_n^{2\sharp}|] = 0$, is that:

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

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 that $Y = 1 + o(1)$. 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 4.598 by using Maple [18]. □

4 Discussion

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

$$\begin{aligned}&\mathbf{E}[|\mathcal{A}_n^l|] = \\ &\quad (7/8)^{rn}{\textstyle\sum}_{A\in\mathcal{A}_n}\mathbf{Pr}[A\in\mathcal{A}_n^1 \mid A\models\phi]\cdot\mathbf{Pr}[A\in\mathcal{A}_n^2 \mid A\in\mathcal{A}_n^1]\cdots\mathbf{Pr}[A\in\mathcal{A}_n^l \mid A\in\mathcal{A}_n^{l-1}],\end{aligned}$$

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 $\lim_n \mathbf{E}[\mathcal{A}_n^l] = 0$, we conjecture that $\lim_l r_l = \kappa$. The equality $\mathbf{Pr}[\phi$ is satisfiable$] = \mathbf{E}[|\mathcal{A}_n^n|]$ 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 $l \geq 0$,*

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

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

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

Now observe that:

$$\mathbf{Pr}[\text{the random formula is satisfiable}] = \sum_\phi (\mathbf{Pr}[\phi] \cdot I_\phi) = \sum_\phi \left(\mathbf{Pr}[\phi] \cdot \sum_A \frac{I_\phi^A}{|\mathcal{A}_n^l|} \right) =$$

$$\sum_A \left(\mathbf{Pr}[A \in \mathcal{A}_n^l] \cdot \sum_\phi \frac{\mathbf{Pr}[\phi \mid A \in \mathcal{A}_n^l]}{|\mathcal{A}_n^l|} \right) = \sum_A \left(\mathbf{Pr}[A \in \mathcal{A}_n^l] \cdot \mathbf{E}\left[\frac{1}{|\mathcal{A}_n^l|} \middle| A \in \mathcal{A}_n^l \right] \right). \quad \square$$

It is now conceivable that the techniques introduced by Kamath et al. in [12] can be applied to estimate $\mathbf{E}[1/|\mathcal{A}_n^l| \mid A \in \mathcal{A}_n^l]$, for an arbitrary fixed $A \in \mathcal{A}_n^l$. 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 $\mathbf{Pr}[A \in \mathcal{A}_n^1]$, such a generalization in conjunction with the above Proposition would improve the single flips estimate.

Acknowledgments

We thank Dimitris Achlioptas and Mike Molloy for reading an earlier draft of this paper and noticing an error it contained. The discussions that followed were very helpful to us. We also thank Paul Spirakis. His remarks were crucial in improving our notation. We are grateful to 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 thank Petr Savický. He pinpointed to us an error in the way we used Janson's inequality in a previous draft of this paper. To avoid it, we had to make use of the $\mathcal{G}_p$ model.

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, 1972.

[2] D.J. Aldous, "The harmonic mean formula for probabilities of unions: applications to sparse random graphs," *Discrete Mathematics* 76, pp 167–176, 1989.

[3] N. Alon, J.H. Spencer, and P. Erdös, *The Probabilistic Method,* J. Wiley, 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)," *Proc. 33rd IEEE Symposium on Foundations of Computer Science*, pp 620–627, 1992.

[6] V. Chvátal and E. Szemerédi, "Many hard examples for resolution," *Journal of the Association for Computing Machinery* 35, pp 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 Applied Mathematics* 5, pp 77–87, 1983.

[9] A. Frieze and S. Suen, *Analysis of Two Simple Heuristics on a Random Instance of k-SAT*, Preprint, Department of Mathematics, Carnegie Mellon University, 1995.

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

[11] S. Janson, "Poisson approximation for large deviations," *Random Structures and Algorithms* 1, pp 221–230, 1990.

[12] A. Kamath, R. Motwani, K. Palem, and P. Spirakis, "Tail bounds for occupancy and the satisfiability threshold conjecture," *Random Structures and Algorithms* 7, pp 59–80, 1995. Also in: *Proc. 35th FOCS*, IEEE, pp 592–603, 1994.

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

[14] L. Kirousis, E. Kranakis, and D. Krizanc, *An Upper Bound for a Basic Hypergeometric Series*, Carleton University, School of Computer Science, Technical Report TR-96-07, 1996.

[15] 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.

[16] D. Knuth, *Fundamental Algorithms*, The Art of Computer Programming, vol. 1, 2nd edition, Addison-Wesley, Reading, Massachusetts, 1973.

[17] C. McDiarmid, "On a correlation inequality of Farr," *Combinatorics, Probability and Computing* 1, pp 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 systémes-experts," *C.R. Acad. Sci. Paris. Sér. I Math.* 302, pp 283–286, 1986.

[20] J. H. Spencer, *Ten Lectures on the Probabilistic Method*, 2nd edition, SIAM, Philadelphia, 1994

Appendix

In this section, we formally prove that the assumptions of Janson's inequality, as presented in [20], hold in the context described in Subsection 3.1. Before providing the details, we introduce certain useful definitions, notational abbreviations and assumptions.

Let A be a fixed truth assignment. All clauses considered in this section are satisfied by A. The condition $A \models \phi$ is assumed without explicit mention and the probability space we work is the $\mathcal{G}_p$ space of formulas formed by clauses satisfied by A.

First notice that all unconditional events we consider in this paper belong to one of the following two classes:

1. Events asserting that the random formula ϕ contains a clause from a certain set of clauses (we refer to this set as the set of clauses *associated* with the event). For example, if sf is a single flip that changes the value of the variable x_i from FALSE to TRUE, then $A^{sf} \not\models \phi$ is such an event and its associated set of clauses consists of the clauses that contain the literal $\neg x_i$ and two literals $\neg y^A$ and $\neg z^A$, where y, z are variables distinct from x_i (recall, x^A denotes the literal on the variable x that is satisfied by the truth assignment A).

2. Boolean combinations of events belonging to the aforementioned class. The set of clauses associated with such a Boolean combination is, by definition, the union of the sets of clauses associated with each event in the Boolean combination.

It is immediate, because the probability space belongs to the model $\mathcal{G}_p$, that two events are independent iff the sets of clauses associated with them are disjoint. Also if an event is independent from each event from a given collection of events, then it is also independent from any Boolean combination of events from this collection. In other words, the notion of independence coincides with the notion of mutual independence as defined in [20, Lecture 8]. We will often make use of this fact.

Also notice that by our assumption that $A \models \phi$, the condition $A \in \mathcal{A}_n^1$ is equivalent to the condition that for all single flips sf, $A^{sf} \not\models \phi$.

For notational convenience, for a conditional or unconditional event E, $\mathbf{Pr}[E]$ will denote, in this Appendix only, the conditional probability $\mathbf{Pr}_p[E \mid A \in \mathcal{A}_n^1]$. The reader is warned though that the condition $A \in \mathcal{A}_n^1$ does not create a probability space of the type $\mathcal{G}_p$. Nevertheless, the events we consider are such that any independency between any two of them that holds in the unconditional probability space carries over to the conditional space. This can be verified at each particular instance, by writing the conditional probabilities as ratios of unconditional ones. A general argument showing this is the following: Consider a particular single flip sf and let F be the set of clauses associated with the the event $A^{sf} \not\models \phi$. Any pair of events E_1 and E_2 that we consider in this paper has the property that if the sets of clauses accociated with E_1 and E_2 are disjoint, then F does *not* intersect both of them. In addition, the sets of clauses associated with $A^{sf} \not\models \phi$ for each sf are pairwise disjoint. From these "separability" properties, it can be immediately concluded that if E_1 and E_2 are independent in the underlying $\mathcal{G}_p$ space, they are also independent under the condition $\forall sf, A^{sf} \not\models \phi$. We will often implicitly use this fact. Notice that this remark proves the mutual independence of events that are related to non-intersecting double flips, a prerequisite to Janson's inequality due to our definition of Δ in Subsection 3.1.

Let DF and SF be the sets of double and, respectively, single flips on A, and let df_0 and df_1 be two arbitrary elements of DF, which we consider fixed. We prove below that if $r = m/n$ exceeds 3, then the following two statements are true:

(a) For all $I \subset DF$ with $df_0 \notin I$,

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \leq \mathbf{Pr}[A^{df_0} \models \phi].$$

(b) For all $I \subset DF$ with $df_0, df_1 \notin I$,

$$\mathbf{Pr}[A^{df_0} \models \phi \wedge A^{df_1} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \leq \mathbf{Pr}[A^{df_0} \models \phi \wedge A^{df_1} \models \phi].$$

The above two inequalities would follow from the FKG inequality [3] if the implicit condition

$$\bigwedge_{sf \in SF} A^{sf} \not\models \phi$$

were not assumed.

Let $df \in DF$, and assume that df changes the values of the variable x_i from FALSE to TRUE and the value of the variable x_j $(j > i)$ from TRUE to FALSE. Consider the set of clauses C_{df} that either (i) contain the literal x_j and two literals $\neg x_k^A$ and $\neg x_l^A$, with $k, l \notin \{i, j\}$, or (ii) contain the literals $\neg x_i$ and x_j and a literal $\neg x_k^A$ with $k \neq i, j$. These clauses are satisfied by

A, and not satisfied by A^{df} (notice that there is also a third group of clauses satisfied by A but not satisfied by A^{df}, namely, the clauses that contain the literal $\neg x_i$ and two literals $\neg x_k^A$ and $\neg x_l^A$ with $k, l \notin \{i, j\}$—we do not include this third group in C_{df}). Let $C = \cup_{df \in DF} C_{df}$. Also let $\mathcal{C}$ denote the event that none of the clauses in C belong to ϕ. It is immediate that under the condition $\mathcal{C}$, for two arbitrary double flips df' and df'', the events $A^{df'} \not\models \phi$ and $A^{df''} \not\models \phi$ are dependent iff the flips df' and df'' share the variable that they flip from FALSE to TRUE. Notice also that the clauses in C are *not* involved with the implicit condition

$$\bigwedge_{sf \in SF} A^{sf} \not\models \phi.$$

We first give an informal outline of the main idea in the proof of the statements (a) and (b): First, performing the corresponding probability computations, verify that for any value of $r = m/n$ exceeding 3 the inequalities in the statements (a) and (b) hold when I is a singleton. Also show that the statements (a) and (b) follow from the corresponding statements but with the probabilities involved having the additional condition $\mathcal{C}$ (an admittedly superficial "explanation" of this fact is the following: (a) and (b) hold when the condition $\bigwedge_{sf \in SF} A^{sf} \not\models \phi$ is not present; but as this condition, which presumably may destroy the validity of (a) or (b), refers to clauses not in C, it is plausible to expect that if (a) or (b) were false, they would also be false under the condition $\mathcal{C}$, as this last condition excludes none of the clauses involved in $\bigwedge_{sf \in SF} A^{sf} \not\models \phi$). Now, under the condition $\mathcal{C}$, to reduce statements (a) and (b) for an arbitrary I to the case where I is a singleton is not difficult, as, under $\mathcal{C}$, we have nice independencies among the events $A^f \not\models \phi$, where f ranges over double or single flips.

We give below the formal proof.

Claim I. To prove statement (a), it is sufficient to prove that it holds with the probabilities involved having the additional condition $\mathcal{C}$.

PROOF Let c be a clause in $C = \cup_{df \in DF} C_{df}$. First notice that the event $c \in \phi$ is independent from each of the events $A^{sf} \not\models \phi, sf \in SF$. We obviously have:

$$\begin{aligned}
&\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] = \\
&\quad \mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, c \in \phi] \cdot \mathbf{Pr}[c \in \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \\
&\quad + \mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, c \notin \phi] \cdot \mathbf{Pr}[c \notin \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi]. \qquad (29)
\end{aligned}$$

Assume first that $c \notin C_{df_0}$. Then the event $c \notin \phi$ is independent from the event $A^{df_0} \models \phi$. Therefore, as the event $c \notin \phi$ is also independent from each of the events $A^{sf} \not\models \phi$, we have that:

$$\mathbf{Pr}[A^{df_0} \models \phi] = \mathbf{Pr}[A^{df_0} \models \phi \mid c \notin \phi]. \qquad (30)$$

Also notice that the event $c \in \phi$ implies the event $A^{df} \not\models \phi$ for some $df \in I$ (i.e., for those $df \in I$ for which $c \in C_{df}$), while it is independent from the events $A^{df} \not\models \phi$ for the remaining $df \in I$. Let I' be the subset of I consisting of the double flips $df \in I$ for which the events $c \in \phi$ and $A^{df} \not\models \phi$ are independent (i.e., those $df \in I$ for which $c \notin C_{df}$). Then

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, c \in \phi] = \mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I'} A^{df} \not\models \phi, c \notin \phi]. \qquad (31)$$

From equalities (29), (30), and (31), it follows that proving the inequality

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \leq \mathbf{Pr}[A^{df_0} \models \phi]$$

for a particular I (when $c \notin C_{df_0}$) is equivalent to proving the following two inequalities:

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I'} A^{df} \not\models \phi, c \notin \phi] \leq \mathbf{Pr}[A^{df_0} \models \phi \mid c \notin \phi], \qquad (32)$$

and

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, c \notin \phi] \leq \mathbf{Pr}[A^{df_0} \models \phi \mid c \notin \phi]. \qquad (33)$$

Assume now $c \in C_{df_0}$. Because, in this case, the event $c \in \phi$ is incompatible with $A^{df_0} \models \phi$, we have that

$$\mathbf{Pr}[A^{df_0} \models \phi] = \mathbf{Pr}[A^{df_0} \models \phi \mid c \notin \phi] \cdot \mathbf{Pr}[c \notin \phi], \qquad (34)$$

and

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] = \mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, c \notin \phi] \cdot \mathbf{Pr}[c \notin \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi]. \qquad (35)$$

Because now the probability of $c \notin \phi$ *without* the implicit condition $\bigwedge_{sf \in SF} A^{sf} \not\models \phi$ is equal to the (implicitly conditional) probability $\mathbf{Pr}[c \notin \phi]$, we deduce by FKG, that

$$\mathbf{Pr}[c \notin \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \leq \mathbf{Pr}[c \notin \phi].$$

Therefore, we conclude from equations (34), and (35) that proving

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi] \leq \mathbf{Pr}[A^{df_0} \models \phi]$$

for a particular I (when $c \in C_{df_0}$) is equivalent to proving:

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, c \notin \phi] \leq \mathbf{Pr}[A^{df_0} \models \phi \mid c \notin \phi]. \qquad (36)$$

Repeating recursively this procedure for all clauses $c \in C$, we conclude that to prove the statement (a), it is sufficient to prove it when the probabilities involved have the additional condition $\mathcal{C}$. Notice that in this recursive proof, I is not fixed but, at each recursive step, it ranges over all subsets of DF. □

Claim II. To prove statement (b), it is sufficient to prove that it holds with the probabilities involved having the additional condition $\mathcal{C}$.

Proof The proof is identical to the proof of the previous claim, except that the two cases that we distinguish with respect to clause c are (i) $c \notin C_{df_0} \cup C_{df_1}$ and (ii) $c \in C_{df_0} \cup C_{df_1}$. □

Now, let I_0 be the subset of I consisting of the elements of I that share with df_0 the value that they flip from FALSE to TRUE. Also, let sf_0 be the single flip that shares a flipped variable with df_0. Notice that the condition $\mathcal{C}$ excludes from the probability space certain clauses, so this condition alone reduces the probability space to one which again belongs to the model $\mathcal{G}_p$. Also, by the independencies that hold under $\mathcal{C}$, namely from the fact that under $\mathcal{C}$ the events associated with two (double or single) flips are dependent iff they share the variable they flip from FALSE to TRUE, we conclude that the probability

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, \mathcal{C}]$$

is equal to the probability of $A^{df_0} \models \phi$ in the $\mathcal{G}_p$ space of clauses that are satisfied by A and for which $\mathcal{C}$ holds, conditioned on $\bigwedge_{df \in I_0} A^{df} \not\models \phi$ and $A^{sf_0} \not\models \phi$. Finally, from the fact that the event $\bigwedge_{sf \neq sf_0} A^{sf} \not\models \phi$ is independent (under $\mathcal{C}$, also) from $A^{df_0} \models \phi$ and from $\bigwedge_{df \in I_0} A^{df} \not\models \phi$, we conclude that

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I} A^{df} \not\models \phi, \mathcal{C}] = \mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I_0} A^{df} \not\models \phi, \mathcal{C}].$$

Therefore, by Claim I, to prove (a) it is sufficient to prove that for any $I_0 \subset I$ whose elements are double flips that share with df_0 the variable that is flipped from FALSE to TRUE and $df_0 \notin I_0$:

$$\mathbf{Pr}[A^{df_0} \models \phi \mid \bigwedge_{df \in I_0} A^{df} \not\models \phi, \mathcal{C}] \leq \mathbf{Pr}[A^{df_0} \models \phi \mid \mathcal{C}],$$

or equivalently,

$$\mathbf{Pr}[A^{df_0} \models \phi, \bigwedge_{df \in I_0} A^{df} \not\models \phi \mid \mathcal{C}] \leq \mathbf{Pr}[A^{df_0} \models \phi \mid \mathcal{C}] \cdot \mathbf{Pr}[\bigwedge_{df \in I_0} A^{df} \not\models \phi \mid \mathcal{C}]. \tag{37}$$

Notice now that it is not compatible to have $\bigwedge_{sf \in SF} A^{sf} \not\models \phi$ and at the same time

$$A^{df'}, A^{df''}, A^{df'''} \models \phi$$

for three (or more) double flips $A^{df'}, A^{df''}, A^{df'''}$ all sharing the variable they flip from FALSE to TRUE. So, if we expand by inclusion-exclusion both sides of inequality (37), with respect to the conjunction $\bigwedge_{df \in I_0} A^{df} \not\models \phi$, we conclude that to prove inequality (37), it suffices to prove that if an arbitrary double flip df shares with df_0 the variable it flips from FALSE to TRUE, then

$$\mathbf{Pr}[A^{df_0} \models \phi, A^{df} \models \phi \mid \mathcal{C}] \geq \mathbf{Pr}[A^{df_0} \models \phi \mid \mathcal{C}] \cdot \mathbf{Pr}[A^{df} \models \phi \mid \mathcal{C}]. \tag{38}$$

Now, by carrying out the calculations in Lemmata 5 and 6 under the additional condition $\mathcal{C}$, we get that the left-hand side of inequality (38) is equal to

$$\frac{q^{\binom{n-2}{2}} q^{n-3} p}{1 - q^{\binom{n-1}{2}}},$$

while its right-hand side is equal to

$$\left(\frac{q^{\binom{n-2}{2}}(1 - q^{n-2})}{1 - q^{\binom{n-1}{2}}} \right)^2 .$$

From the above, by elementary algebraic manipulations, it follows that inequality (38) is equivalent to

$$\frac{q^{\binom{n-2}{2}}(1 - q^{n-2})^2}{(1 - q^{\binom{n-1}{2}}) q^{n-3} p} \leq 1,$$

which, asymptotically, is equivalent to

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

where $p \sim 6r/(7n^2)$ and $q = 1 - p$. By using Maple (or by elementary calculus), the last inequality is true when r exceeds 3 (with more accuracy, 2.93). This concludes the proof that statement (a) is true. The continuation of the proof of statement (b) is completely analogous.