Technical Reports on Mathematical and Computing Sciences: TR-C141
title: On the Maximum Satisfiability of Random 3-CNF Formulae
author: Mitsuo Motoki
affiliation:
Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology
Meguro-ku Ookayama, Tokyo 152-8552.
email: mmotoki@is.titech.ac.jp

Abstract.

For Maximum 3-Satisfiability Problem (MAX-3SAT), we discuss about the expected number of unsatisfiable clauses on optimal assignment.

We proved two lower bounds theoretically. One is $\Omega(m/\ln m)$ for $m \geq 5.19n$ and the other is $m/8 - O(\sqrt{mn})$ for $m \geq 9.7n$ where n is the number of variables and m is the number of clauses.

We also compare these theoretical lower bounds with experimental results.

1 Introduction

The Maximum Satisfiability Problem (MAX-SAT), the problem of estimating maximum number of satisfiable clauses simultaneously in a given conjunctive normal form (CNF) Boolean formula, is optimization version of Satisfiability Problem (SAT), one of the well-known NP-complete problem. Maximum 3-Satisfiability Problem (MAX-3SAT) is restricted version of MAX-SAT to take only 3-CNF formula as input.

Recently, approximation algorithms for optimization problem have been improved, especially on the ratio of approximation solution to optimal solution on the worst case. On the other hand, experimental analysis has some problems about instance generation. It is hard to generate test instances at random with their optimal solution for optimization problem. Furthermore it is not clear that what kind of instances is required and how many unsatisfiable clauses on optimal is required? For such a problem, we aim to analyze behavior of the number of unsatisfiable clauses.

In this paper, we consider the expected number of unsatisfiable clauses. On the expected number of satisfiable clauses, Broder et al.[1] proved the probability that the difference from the expected maximum number of satisfiable clause is larger than $\sqrt{2m \log m}$ is less than $1/m$ by using Martingale technique. But they did not determine the actual expected number. We prove two lower bounds of the expected number of the unsatisfiable clauses.

We recall some notions on the 3SAT problem and define some notations. In this paper, we consider a *3-conjunctive normal form* (3-CNF) formula f over n *variables*, $x_1, x_2, \ldots, x_n$ generated uniform at random. That is, f is a sequence $c_1, c_2, \ldots, c_m$ of *clauses*, where each clause is a set of exactly 3 *literals*, i.e., variable x_i or its complement $\overline{x_i}$. Here we assume that any variable appears at most once in each clause. Throughout this paper, we use n and m to denote the number of variables and clauses respectively.

2 Theoretical analysis

Let f be a random 3CNF formula with n variables and m clauses. For any truth assignment t over n variables, we denote by $u(f,t)$ the number of unsatisfiable clauses on t. Furthermore we use $u(f)$ to denote $\min_t(u(f,t))$ that is the number of unsatisfiable clauses on optimal assignment of f.

First, we prove the lower bound of $\mathrm{E}[u(f)]$ in a naive way.

Theorem 1. *For any $m \geq n \ln 2/\ln(8/7) \approx 5.19n$, the expected number of unsatisfiable clauses on optimal assignment is larger than*

$$\frac{m \ln 8/7 - n \ln 2 + \ln(m/7-1)}{\ln m/7}(1-7/m).$$

Proof. For the any assignment t, the probability that the number of unsatisfiable clauses in f on t is exactly v is estimated as follows. (Through out this section, the probability is taken under the random choice of f.)

$$\Pr[u(f,t)=v] = \binom{m}{v}\left\{\frac{1}{8}\right\}^v \left\{\frac{7}{8}\right\}^{m-v} = \left\{\frac{7}{8}\right\}^m \binom{m}{v} 7^{-v}.$$

Thus we can bound the probability that the number of unsatisfiable clauses in f on its optimal assignment is larger than v as follows.

$$\begin{aligned}
\Pr[u(f) \geq v] &= 1 - \Pr[u(f) < v] \\
&= 1 - \sum_{i=0}^{v-1} \Pr[u(f) = i] \\
&\geq 1 - \sum_{i=0}^{v-1} \Pr\left[\bigvee_t u(f,t) = i\right] \\
&\geq 1 - \sum_{i=0}^{v-1} 2^n \Pr\left[u(f,t) = i\right] \\
&\geq 1 - 2^n (7/8)^m \sum_{i=0}^{v-1} (m/7)^v \\
&\geq 1 - 2^n (7/8)^m \frac{(m/7)^v}{m/7-1}
\end{aligned}$$

Because $m \geq \frac{n \ln 2}{\ln 8/7}$, $2^n(7/8)^m \leq 1$, that is, the bound of $\Pr[u(f) \geq v]$ is larger than 0. By using Markov's inequality $\Pr[X \geq a] \leq \mathrm{E}[X]/a$, we obtain

$$\mathrm{E}[u(f)] \geq v\left\{1 - 2^n(7/8)^m \frac{(m/7)^v}{m/7-1}\right\}.$$

By setting $v = \frac{m \ln 8/7 - n \ln 2 + \ln(m/7-1)}{\ln m/7} - 1$, we have

$$\mathrm{E}[u(f)] \geq \frac{m \ln 8/7 - n \ln 2 + \ln(m/7-1)}{\ln m/7}(1-7/m).$$

□

In the Theorem 1, rounding $\binom{m}{v}$ by m^v causes the denominator $\ln m/7$. Thus we need to estimate the probability that $u(f)$ is larger than v more tightly. This observation leads us to the next result.

Theorem 2. *For any* $m \geq 7n\ln 2 + \sqrt{7n\ln 2}\sqrt{7n\ln 2 - 16} - 8 \approx 9.7n - 8$, *the expected number of unsatisfiable clauses on optimal assignment is larger than*

$$\frac{1}{8}\left(m - \sqrt{14mn\ln 2} + 8\right)\left(1 - \frac{1}{\sqrt{2\pi n\ln 2}}\right).$$

Proof. We already have $\Pr[u(f) \geq v] \geq 1 - 2^n \sum_{i=0}^{v-1}\binom{m}{i}(1/8)^i(7/8)^{m-i}$ in the proof of Theorem 1. This probability includes partial sum of binomial distribution. For sufficiently large m, we can approximate binomial distribution to normal distribution by using central limit theorem. Thus

$$\begin{aligned}
\Pr[u(f) \geq v] &\geq 1 - 2^n \sum_{i=0}^{v-1}\binom{m}{i}\left\{\frac{1}{8}\right\}^i\left\{\frac{7}{8}\right\}^{m-i} \\
&\approx 1 - 2^n \int_{-\infty}^{\frac{8v-8-m}{\sqrt{7m}}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \\
&= 1 - 2^n \Phi\left(-\frac{m - 8v + 8}{\sqrt{7m}}\right).
\end{aligned}$$

We set v to $m/8 - \sqrt{14mn\ln 2}/8 + 1$. Because $m \geq 7n\ln 2 + \sqrt{7n\ln 2}\sqrt{7n\ln 2 - 16} - 8$, v is larger than 0, and furthermore $(m - 8v + 8)/\sqrt{7m}$ tends to infinity. Thus applying $\Phi(-x) < \exp(-x^2/2)/\sqrt{2\pi}x$ if $x > 0$[2], we can bound

$$\Pr[u(f) \geq v] \geq 1 - 2^n \frac{\sqrt{7m}}{\sqrt{2\pi}(m - 8v + 8)} \exp\left(-\frac{(m - 8u + 8)^2}{14m}\right).$$

Because $v \leq m/8 - \sqrt{14mn\ln 2}/8 + 1$, we have $2^n \exp\left(-\frac{(m-8u+8)^2}{14m}\right) < 1$. Also it follows from $v \leq m/8 - \sqrt{14mn\ln 2}/8 + 1 \leq m/8 - \sqrt{7m/2\pi}/8 + 1$, that $\sqrt{7m/2\pi}/(m - 8v + 8)$ is less than 1. Thus, $\Pr[u(f) \geq v] \geq 0$.

Combine all, by using Markov's inequality, we get

$$\mathrm{E}[u(f)] \geq \frac{1}{8}\left(m - \sqrt{14mn\ln 2} + 8\right)\left(1 - \frac{1}{\sqrt{2\pi n\ln 2}}\right).$$

□

3 Experimental results

In theoretical analyses, we show $\Omega(m/\ln m)$ and $\Omega(m)$ lower bounds for the average of $u(f)$, where $u(f)$ is the minimum number of unsatisfiable clauses in f simultaneously. We compare these theoretical lower bounds with the average number of unsatisfiable clauses on optimal assignment that we obtained through some experiments.

First, we describe the algorithm used in the experiment. The algorithm generates formula f at random and finds its optimal assignment by a backtrack algorithm explained below.

Let f be a generated formula to our backtrack algorithm. Let u_* be the current minimum number of unsatisfiable clauses. We use u_* to prune our backtrack branches. Initially u_* is defined by the assignment computed by walksat.

Now our backtrack algorithm is executed as follows. Let x be the variable in f that occurs most frequently. Startly this x, the algorithm determine its assignment based on the number of satisfied clauses. That is, x is assigned 0 if the number of clauses containing $\overline{x}$ is larger than that of x, and x is assigned 1 otherwise. The formula is reduced based on this assignment. The algorithm then chooses the variable that occurs most frequently in the reduced formula. During this process, if the number of unsatisfiable clauses is not less than u_*, the algorithm takes backtrack, and if all clauses are satisfied or unsatisfiable and the number of unsatisfiable clauses is less than u_*, it renews u_*.

To reduce the size of formula, we apply following rules similar to a resolution that do not change the number of satisfiable (resp., unsatisfiable) clauses on any assignment. In following rules, we denote $f_1 \equiv f_2$ if the number of satisfiable clauses of f_1 is equal to that of f_2 on any assignment. We denote always satisfiable clause by T and empty (unsatisfiable) clause by E. We use x_i to denote literal, and if $i \neq j$, variable of x_i is distinct to that of x_j.

First group is normal resolutions.

- $(x)(\overline{x}) \equiv TE$
- $(x_1 \vee x_2)(x_1 \vee \overline{x_2}) \equiv T(x_1)$
- $(x_1 \vee x_2 \vee x_3)(x_1 \vee x_2 \vee \overline{x_3}) \equiv T(x_1 \vee x_2)$

Second group is a chain of implies and two unit clauses. This group can contain any length of chain, but we use length at most eight chain.

- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2}) \equiv T(x_1 \vee \overline{x_2})E$
- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2} \vee x_3)(\overline{x_3}) \equiv T(x_1 \vee \overline{x_2})(x_2 \vee \overline{x_3})E$
- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2} \vee x_3)(\overline{x_3} \vee x_4)(\overline{x_4}) \equiv T(x_1 \vee \overline{x_2})(x_2 \vee \overline{x_3})(x_3 \vee \overline{x_4})E$
- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2} \vee x_3)(\overline{x_3} \vee x_4)(\overline{x_4} \vee x_5)(\overline{x_5}) \equiv T(x_1 \vee \overline{x_2})(x_2 \vee \overline{x_3})(x_3 \vee \overline{x_4})(x_4 \vee \overline{x_5})E$
- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2} \vee x_3)(\overline{x_3} \vee x_4)(\overline{x_4} \vee x_5)(\overline{x_5} \vee x_6)(\overline{x_6}) \equiv T(x_1 \vee \overline{x_2})(x_2 \vee \overline{x_3})(x_3 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_5 \vee \overline{x_6})E$

  $\vdots$

Third group is like a three pointed star, that is, a clause whose size is three and three unit clauses with complement literal, each literal is distinct each other and appears in the size three clause or is connected by at most two implies for each literal.

- $(x_1)(x_2)(x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3}) \equiv E(x_1 \vee x_2)(x_1 \vee x_3)(\overline{x_1} \vee x_2 \vee x_3)$
- $(x_1)(\overline{x_1} \vee x_4)(x_2)(x_3)(\overline{x_2} \vee \overline{x_3} \vee \overline{x_4}) \equiv E(x_1 \vee \overline{x_4})(x_2 \vee x_4)(x_3 \vee x_4)(x_2 \vee x_3 \vee \overline{x_4})$
- $(x_1)(\overline{x_1} \vee x_4)(x_2)(x_3)(\overline{x_2} \vee \overline{x_3} \vee \overline{x_4}) \equiv E(x_1 \vee \overline{x_4})(x_2 \vee x_4)(x_2 \vee x_3)(\overline{x_2} \vee x_3 \vee x_4)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(x_3)(\overline{x_2} \vee \overline{x_3} \vee \overline{x_5}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee x_5)(x_3 \vee x_5)(x_2 \vee x_3 \vee \overline{x_5})$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(x_3)(\overline{x_2} \vee \overline{x_3} \vee \overline{x_5}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee x_5)(x_2 \vee x_3)(\overline{x_2} \vee x_3 \vee x_5)$
- $(x_1)(\overline{x_1} \vee x_4)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee \overline{x_4} \vee \overline{x_6}) \equiv E(x_1 \vee \overline{x_4})(x_2 \vee \overline{x_6})(x_4 \vee x_6)(x_3 \vee x_4)(x_3 \vee \overline{x_4} \vee x_6)$
- $(x_1)(\overline{x_1} \vee x_4)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee \overline{x_4} \vee \overline{x_6}) \equiv E(x_1 \vee \overline{x_4})(x_2 \vee \overline{x_6})(x_4 \vee x_6)(x_3 \vee x_6)(x_3 \vee x_4 \vee \overline{x_6})$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee \overline{x_5} \vee \overline{x_6}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_3 \vee x_5)(x_3 \vee x_6)(\overline{x_3} \vee x_5 \vee x_6)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee \overline{x_5} \vee \overline{x_6}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_3 \vee x_5)(x_5 \vee x_6)(x_3 \vee \overline{x_5} \vee x_6)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee \overline{x_5} \vee \overline{x_6}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_5 \vee x_6)(x_3 \vee x_6)(x_3 \vee x_5 \vee \overline{x_6})$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(\overline{x_6} \vee x_7)(x_3)(\overline{x_3} \vee \overline{x_5} \vee \overline{x_7}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_6 \vee \overline{x_7})(x_3 \vee x_5)(x_3 \vee x_7)(\overline{x_3} \vee x_5 \vee x_7)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(\overline{x_6} \vee x_7)(x_3)(\overline{x_3} \vee \overline{x_5} \vee \overline{x_7}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_6 \vee \overline{x_7})(x_3 \vee x_5)(x_5 \vee x_7)(x_3 \vee \overline{x_5} \vee x_7)$
- $(x_1)(\overline{x_1} \vee x_4)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee x_8)(\overline{x_4} \vee \overline{x_6} \vee \overline{x_8}) \equiv E(x_1 \vee \overline{x_4})(x_2 \vee \overline{x_6})(x_3 \vee \overline{x_8})(x_4 \vee x_6)(x_4 \vee x_8)(\overline{x_4} \vee x_6 \vee x_8)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee x_8)(\overline{x_5} \vee \overline{x_6} \vee \overline{x_8}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_3 \vee \overline{x_8})(x_5 \vee x_6)(x_5 \vee x_8)(\overline{x_5} \vee x_6 \vee x_8)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(x_3)(\overline{x_3} \vee x_8)(\overline{x_5} \vee \overline{x_6} \vee \overline{x_8}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_3 \vee \overline{x_8})(x_5 \vee x_6)(x_6 \vee x_8)(x_5 \vee \overline{x_6} \vee x_8)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(\overline{x_6} \vee x_7)(x_3)(\overline{x_3} \vee x_8)(\overline{x_5} \vee \overline{x_7} \vee \overline{x_8}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_6 \vee \overline{x_7})(x_3 \vee \overline{x_8})(x_5 \vee x_7)(x_5 \vee x_8)(\overline{x_5} \vee x_7 \vee x_8)$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(\overline{x_6} \vee x_7)(x_3)(\overline{x_3} \vee x_8)(\overline{x_5} \vee \overline{x_7} \vee \overline{x_8}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_6 \vee \overline{x_7})(x_3 \vee \overline{x_8})(x_5 \vee x_8)(x_7 \vee x_8)(x_5 \vee x_7 \vee \overline{x_8})$
- $(x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee x_5)(x_2)(\overline{x_2} \vee x_6)(\overline{x_6} \vee x_7)(x_3)(\overline{x_3} \vee x_8)(\overline{x_8} \vee x_9)(\overline{x_5} \vee \overline{x_7} \vee \overline{x_9}) \equiv E(x_1 \vee \overline{x_4})(x_4 \vee \overline{x_5})(x_2 \vee \overline{x_6})(x_6 \vee \overline{x_7})(x_3 \vee \overline{x_8})(x_8 \vee \overline{x_9})(x_5 \vee x_7)(x_5 \vee x_9)(\overline{x_5} \vee x_7 \vee x_9)$

Forth group is an unit clause and a circle with three or four implies.

- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2} \vee x_3)(\overline{x_3} \vee \overline{x_1}) \equiv ET(x_1 \vee \overline{x_2} \vee x_3)(\overline{x_1} \vee x_2 \vee \overline{x_3})$
- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2} \vee x_3)(\overline{x_3} \vee x_4)(\overline{x_4} \vee \overline{x_1}) \equiv E(x_1 \vee \overline{x_2} \vee x_3)(\overline{x_1} \vee x_2 \vee \overline{x_3})(x_1 \vee \overline{x_3} \vee x_4)(\overline{x_1} \vee x_3 \vee \overline{x_4})$
- $(x_1)(\overline{x_1} \vee x_2)(\overline{x_2} \vee x_3)(\overline{x_3} \vee x_4)(\overline{x_4} \vee \overline{x_1}) \equiv E(x_1 \vee \overline{x_2} \vee x_4)(\overline{x_1} \vee x_2 \vee \overline{x_4})(x_2 \vee \overline{x_3} \vee x_4)(\overline{x_2} \vee x_3 \vee \overline{x_4})$

Fifth group is two circles with three or four implies

- $(x_1 \vee x_2)(\overline{x_2} \vee \overline{x_3})(x_3 \vee x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee \overline{x_5})(x_5 \vee \overline{x_1}) \equiv ET(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_1} \vee x_4 \vee x_5)(x_1 \vee \overline{x_4} \vee \overline{x_5})$
- $(x_1 \vee x_2)(\overline{x_2} \vee \overline{x_3})(x_3 \vee x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee \overline{x_5})(x_5 \vee \overline{x_6})(x_6 \vee \overline{x_1}) \equiv E(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_1} \vee x_4 \vee x_5)(x_1 \vee \overline{x_4} \vee \overline{x_5})(x_1 \vee \overline{x_6} \vee x_5)(\overline{x_1} \vee x_6 \vee \overline{x_5})$
- $(x_1 \vee x_2)(\overline{x_2} \vee \overline{x_3})(x_3 \vee x_1)(\overline{x_1} \vee x_4)(\overline{x_4} \vee \overline{x_5})(x_5 \vee \overline{x_6})(x_6 \vee \overline{x_1}) \equiv E(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_1} \vee x_4 \vee x_6)(x_1 \vee \overline{x_4} \vee \overline{x_6})(x_4 \vee x_5 \vee \overline{x_6})(\overline{x_4} \vee \overline{x_5} \vee x_6)$

Last group is to reduce a pair of size three clauses that consist from same pair of variables and are not resolvable. We use these rules because some rules in the above make such a pair of size three clauses.

- $(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_2} \vee x_1)(\overline{x_1} \vee x_3) \equiv TT(\overline{x_2})(x_2 \vee x_3)$
- $(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_1} \vee x_2)(\overline{x_1} \vee x_3) \equiv TT(\overline{x_1})(x_2 \vee x_3)$
- $(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_2} \vee \overline{x_4})(x_4 \vee x_1)(\overline{x_1} \vee x_3) \equiv T(\overline{x_2})(x_2 \vee x_3)(x_1 \vee x_2 \vee x_4)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_4})$
- $(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_1} \vee \overline{x_4})(x_4 \vee x_2)(\overline{x_1} \vee x_3) \equiv T(\overline{x_1})(x_2 \vee x_3)(x_1 \vee x_2 \vee x_4)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_4})$
- $(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_2} \vee \overline{x_4})(x_4 \vee x_1)(\overline{x_1} \vee \overline{x_5})(x_5 \vee x_3) \equiv (\overline{x_2})(x_2 \vee x_3)(x_1 \vee x_2 \vee x_4)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_4})(x_1 \vee x_3 \vee x_5)(\overline{x_1} \vee \overline{x_3} \vee \overline{x_5})$
- $(x_1 \vee x_2 \vee x_3)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_3})(\overline{x_1} \vee \overline{x_4})(x_4 \vee x_2)(\overline{x_1} \vee \overline{x_5})(x_5 \vee x_3) \equiv (\overline{x_1})(x_2 \vee x_3)(x_1 \vee x_2 \vee x_4)(\overline{x_1} \vee \overline{x_2} \vee \overline{x_4})(x_1 \vee x_3 \vee x_5)(\overline{x_1} \vee \overline{x_3} \vee \overline{x_5})$

The last two rules in the last group do not reduce the size of formulae, but we use them to make an unit clause.

Furthermore, we use the following variable elimination rules[3].

Pure literal rule : If variable x appears only positive (resp., negative) form in f, we assign 1 (resp., 0) to x.

Dominating unit clause rule : If the number of unit clause (x)s (resp., $(\overline{x})$s) is greater than or equal to the total number of literal $\overline{x}$s (resp., xs) in f, we assign 1 (resp., 0) to x.

Next we describe the setup used in the experiments. We fixed the number of variables to 50, and we selected the numbers of clauses being from 250 to 600 at intervals of 50. For each

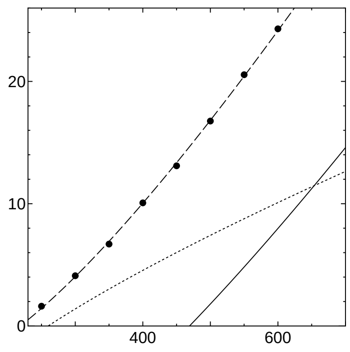

Figure 1: the average number of unsatisfiable clauses on optimal assignment

point, we checked the average number of unsatisfiable clauses on optimal to 100 formulae chosen uniformly at random.

We show the results in Figure 1, along the horizontal axis we give the number of clauses m, and along the vertical axis we give the average number of unsatisfiable clauses on optimal. Big black dots are the results of experiments, and doted line and solid line is the lower bounds of Theorem 1 and Theorem 2. With this graph, it seems that behavior of experimental results is similar to that of lower bound of Theorem 2. Thus we assume that the relation of the average number of unsatisfiable clauses with the number of clauses forms $m/8 - \sqrt{c_1 m} + c_2$, and get $m/8 - \sqrt{5.87m} + 8.51$(dashed line).

References

[1] Andrei Z. Broder, Alan M. Frieze and Eli Upfal, On the Satisfiability and Maximum Satisfiability of Random 3-CNF Formulas, in *Proc. SODA'93*, pp. 322-330.

[2] William Feller, An Introduction to Probability Theory and Its Applications, Volume 1.

[3] Rolf Niedermeier and Peter Rossmanith, New Upper Bounds for MaxSat, in *Proc. ICALP'99*, Lecture Note in Computer Science 1644, pp. 576–584.