Probabilistic Analysis of Random Mixed Horn Formulas

Andreas Wotzlaw
Institut für Informatik
Universität zu Köln
Weyertal 121, 50931 Köln
Germany

Ewald Speckenmeyer
Institut für Informatik
Universität zu Köln
Weyertal 121, 50931 Köln
Germany

Stefan Porschen
Fachgruppe Mathematik, FB4
HTW-Berlin
Treskowallee 8, 10318 Berlin
Germany

Abstract

We present a probabilistic analysis of random mixed Horn formulas (MHF), i.e., formulas in conjunctive normal form consisting of a positive monotone part of quadratic clauses and a part of Horn clauses, with m clauses, n variables, and up to n literals per Horn clause. For MHFs parameterized by n and m with uniform distribution of instances and for large n, we derive upper bounds for the expected number of models. For the class of random negative MHFs, where only monotone negative Horn clauses are allowed to occur, we give a lower bound for the probability that formulas from this class are satisfiable. We expect that the model studied theoretically here may be of interest for the determination of hard instances, which are conjectured to be found in the transition area from satisfiability to unsatisfiability of the instances from the parameterized classes of formulas.

Introduction

In this paper we study probabilistically random mixed Horn formulas. A conjunctive normal form (CNF) formula is a *mixed Horn formula* (MHF) if each of its clauses is either a monotone positive clause of length 2, or it is a Horn clause. This class of Boolean formulas has received some attention recently. Reducing many NP-complete problems to the satisfiability problem (SAT) of CNF formulas results in a natural way into MHFs (Porschen, Schmidt, and Speckenmeyer 2009; Wotzlaw, Speckenmeyer, and Porschen 2012). Hence, dedicated algorithms could be developed with good worst-case upper bounds for solving SAT of MHFs (Porschen and Speckenmeyer 2007; Kottler, Kaufmann, and Sinz 2008).

We consider not only general MHFs but also negative MHFs, where only monotone negative Horn clauses are allowed to occur. For both classes the size of Horn clauses is at most n. It is known that the satisfiability problem of both classes is NP-complete (Porschen and Speckenmeyer 2007). For the class of general MHFs as well as of negative MHFs, both with uniform distribution of instances and for large n, we derive, by means of a probabilistic analysis, the upper bounds for the expected number of solutions. Furthermore, in the case of negative MHFs we give a lower bound for the probability that formulas parameterized by n and m are satisfiable. To this end, we apply non-algorithmic techniques involving the computation of the second moment. A similar approach has already been used for random k-SAT by Achlioptas et al. (2004; 2006) and recently by Schuh (2012).

Our study is closely related to the research on random k-SAT, the satisfiability problem of random k-CNF formulas. A random k-CNF formula, $F_k(n,m)$, is formed by selecting uniformly, independently, and without replacement m k-clauses from the set of all $2^k\binom{n}{k}$ possible clauses on n variables and taking their conjunction. Random formulas have been studied extensively in probabilistic combinatorics in the last three decades. The mathematical investigation of random k-SAT began with the works of Goldberg et al. (1982), Franco and Paull (1983), Purdom (1983), and Chao and Franco (1990).

In the early 1990s, random instances of the k-SAT problem have been understood to undergo a phase transition as a ratio of the number of k-clauses to the number of variables passes through some critical threshold. That is, for a given number of variables, the probability that a random instance is satisfiable decreases rapidly within a small interval of clauses. This sharp threshold phenomenon was first observed experimentally for $F_3(n,m)$ by Cheeseman, Kanefsky, and Taylor (1991) and Mitchell, Selman, and Levesque (1992). By setting $m = rn$, they have found that while for $r < 4.1$ all formulas are satisfiable, for $r > 4.3$ almost all are unsatisfiable. Moreover, as n increases, this transition narrows around $r \approx 4.2$. This has led to a popular *satisfiability threshold conjecture*, saying that for each $k \geq 3$, there exists a constant r_k such that

$$\lim_{n\to\infty} \Pr(F_k(n,rn) \text{ is satisfiable}) = \begin{cases} 1 & \text{if } r < r_k, \\ 0 & \text{if } r > r_k. \end{cases}$$

More recently, Namasivayam and Truszczyński (2010) proved that the SAT problem for a simple class of MHFs, of which the Horn part is allowed to contain only negative clauses of length $k+1$, remains NP-complete. Moreover, they identified experimentally for that MHF class regions of low and high satisfiability depending on the density $\frac{m_p}{n}$ and the clause size k, where m_p denotes the number of positive 2-clauses. They observed also that the hardness of the instances in the phase transition region shows the well-known easy-hard-easy pattern as a function of k. They stated as an open problem to determine analytically tight bounds on the location of the phase transition for their model. Motivated by their experimental results and inspired by the analysis of random unrestricted k-SAT by Schuh (2012), the goal of our study is to bring more understanding in the phase transition phenomenon in the context of random MHFs and random negative MHFs first.

We expect that the model studied theoretically here, may be of interest for the determination of hard instances, which are conjectured to be found in the transition area from satisfiability to unsatisfiability of the instances from the parameterized classes of formulas.

Preliminaries

Let $V = \{x_1, ..., x_n\}$ be a set of n Boolean variables. Each variable induces a positive literal (variable x) or a negative literal (negated variable, $\overline{x}$). A clause c is considered as a disjunction of different literals over V and is represented as a set $c = \{l_1, ..., l_{|c|}\}$. A clause is termed a k-clause if it contains at most $k > 0$ literals. A clause containing at most one positive literal is termed a Horn clause. It is a *definite* Horn clause if it contains exactly one positive literal. A propositional formula C over V is considered as a clause set $C = \{c_1, ..., c_{|C|}\}$. The number of clauses in C is denoted by $|C|$. A formula is called a mixed Horn formula if it is composed only of Horn or monotone positive 2-clauses. In this paper we consider also mixed Horn formulas with monotone negative Horn clauses. We call them *negative* mixed Horn formulas.

Let $t : V \to \{0, 1\}$ be a truth assignment. We say that t satisfies a literal l iff $t(l) = 1$. A clause is said to be satisfied by a truth assignment t iff t satisfies any of its literals. t is said to satisfy formula C iff all clauses of C are satisfied by t. In this case t is called a model of C. We write $t(C) = 1$ iff t is a model of C, $t(C) = 0$ otherwise. A formula C is said to be satisfiable iff it has at least one model. Otherwise it is unsatisfiable.

Definition 1. *Let $H(V)$ ($H_-(V)$) denote the set of (monotone negative) Horn clauses of length at most $n > 0$ on V, and let $P(V)$ denote the set of all monotone positive 2-clauses on V. Furthermore, both $H(V) \cup P(V)$ and $H_-(V) \cup P(V)$ contain no empty clauses.*

Definition 2. *Let $MH(n, m)$ ($MH_-(n, m)$) denote the set of all (negative) mixed Horn formulas of m clauses on n variables. Each formula from $MH(n, m)$ ($MH_-(n, m)$) is formed by selecting uniformly, independently, and* without *replacement m clauses from $H(V) \cup P(V)$ ($H_-(V) \cup P(V)$) and taking their conjunction.*

Note that each formula from $MH(n, m)$ and $MH_-(n, m)$ is free of duplicate clauses. Hence there are $\binom{|P(V) \cup H(V)|}{m}$ and $\binom{|P(V) \cup H_-(V)|}{m}$ formulas in $MH(n, m)$ and $MH_-(n, m)$, respectively. We do not consider empty formulas, i.e., $m > 0$.

Proposition 1. *For a set V of n Boolean variables:*

1. $|P(V)| = \binom{n}{2}$
2. $|H_-(V)| = 2^n - 1$
3. $|H(V)| = |H_-(V)| + n2^{n-1}$

Proof. Observe that in $|H(V)|$ there are $\binom{n}{i}$ monotone negative i-clauses and $n\binom{n-1}{i-1}$ definite Horn i-clauses, for $i = 1, ..., n$. □

Proposition 2. *Let t be some truth assignment on V and w.l.o.g. assume that $\lambda := |\{x \in V : t(x) = 0\}|$. Let $P^\lambda(V) \subseteq P(V)$, $H_-^\lambda(V) \subseteq H_-(V)$, and $H^\lambda(V) \subseteq H(V)$ denote the sets of clauses satisfied by t. It holds that*

1. $|P^\lambda(V)| = |P(V)| - \binom{\lambda}{2}$
2. $|H_-^\lambda(V)| = |H_-(V)| - 2^{n-\lambda} + 1$
3. $|H^\lambda(V)| = |H(V)| - (1 + \lambda)2^{n-\lambda} + 1$

Proof. Observe that in order to obtain $|H^\lambda(V)|$ we must remove from $H(V)$ all monotone negative clauses containing only variables from $\{x \in V : t(x) = 1\}$ and all definite Horn j-clauses containing one positive literal l_p such that $t(l_p) = 0$ and $j - 1$ negative literals assigned by t to 1, for $j = 1, ..., n$. □

Expectation value for the number of models

The number of formulas from $MH(n, m)$ satisfied by a truth assignment $t \in \{0, 1\}^n$ is given by

$$N(t) := \sum_{C \in MH(n,m)} t(C).$$

Let N denote the number of all models of formulas from $MH(n, m)$. Since $N = \sum_{t \in \{0,1\}^n} N(t)$ and due to the linearity of expectation, the expected value of N is given by

$$\mathrm{E}[N] = \mathrm{E}\left[\sum_t N(t)\right] = \sum_t \sum_C \mathrm{E}[t(C)] = \sum_t \sum_C p(C)t(C).$$

In the last term the summations run over all truth assignments $t \in \{0, 1\}^n$ and all formulas $C \in MH(n, m)$, respectively, whereas $p(C)$ denotes the occurrence probability of formula C.

Similar to the models for random k-SAT, we assume that the formulas from $MH(n, m)$ are distributed uniformly, i.e., for all $C \in MH(n, m)$

$$p(C) = \frac{1}{|MH(n, m)|} = \binom{|P(V)| + |H(V)|}{m}^{-1} =: p$$

for $m \leq |P(V)| + |H(V)|$. Under this assumption we have

$$\mathrm{E}[N] = p \sum_t \sum_C t(C).$$

According to Proposition 2, for a truth assignment t assigning exactly λ variables from V to 0, there are at most

$$N(\lambda) := \binom{|P^\lambda(V)| + |H^\lambda(V)|}{m}$$

formulas of length m in $MH(n,m)$ satisfied by t. Thus, the expected value of N for $MH(n,m)$ is given by

$$\begin{aligned}
\mathrm{E}[N] &= p\sum_{\lambda=0}^{n}\binom{n}{\lambda}N(\lambda)\\
&= \binom{|P(V)|+|H(V)|}{m}^{-1}\\
&\quad \sum_{\lambda=0}^{n}\binom{n}{\lambda}\binom{|P^{\lambda}(V)|+|H^{\lambda}(V)|}{m}\\
&= \sum_{\lambda=0}^{n}\binom{n}{\lambda}\prod_{j=0}^{m-1}\frac{|P^{\lambda}(V)|+|H^{\lambda}(V)|-j}{|P(V)|+|H(V)|-j}\\
&=: \sum_{\lambda=0}^{n}\binom{n}{\lambda}\prod_{j=0}^{m-1}\beta(\lambda,j) \qquad (1)
\end{aligned}$$

for $m \le |P^{\lambda}(V)|+|H^{\lambda}(V)|$, since $|P^{\lambda}(V)|+|H^{\lambda}(V)| \le |P(V)|+|H(V)|$.

We analyze now $\mathrm{E}[N]$ for large values of n. To this end, we first estimate the value of $\beta(\lambda,j)$. For $n\to\infty$ and by Proposition 1 and 2, we have

$$\begin{aligned}
\beta(\lambda,j) &= \frac{|P(V)|+|H(V)|-\binom{\lambda}{2}-(1+\lambda)2^{n-\lambda}+1-j}{|P(V)|+|H(V)|-j}\\
&= 1-\frac{\binom{\lambda}{2}+(1+\lambda)2^{n-\lambda}-1}{\binom{n}{2}+2^{n}-1+n2^{n-1}-j}\\
&\to 1-\frac{2(1+\lambda)}{(n+2)2^{\lambda}} < \mathrm{e}^{\frac{-2(1+\lambda)}{(n+2)2^{\lambda}}}.
\end{aligned}$$

By applying this result to (1), we obtain finally

$$\mathrm{E}[N] \le \sum_{\lambda=0}^{n}\binom{n}{\lambda}\mathrm{e}^{\frac{-2m(1+\lambda)}{(n+2)2^{\lambda}}} \le 2^{n}\mathrm{e}^{\frac{-2m(1+n)}{(n+2)2^{n}}}.$$

Analogously, we obtain the expected value of the number of all models N_- of formulas from $MH_-(n,m)$.

$$\begin{aligned}
\mathrm{E}[N_-] &= \sum_{\lambda=0}^{n}\binom{n}{\lambda}\prod_{j=0}^{m-1}\frac{|P^{\lambda}(V)|+|H_-^{\lambda}(V)|-j}{|P(V)|+|H_-(V)|-j}\\
&=: \sum_{\lambda=0}^{n}\binom{n}{\lambda}\prod_{j=0}^{m-1}\beta_-(\lambda,j)
\end{aligned}$$

for $m \le |P^{\lambda}(V)|+|H_-^{\lambda}(V)|$, since $|P^{\lambda}(V)|+|H_-^{\lambda}(V)| \le |P(V)|+|H_-(V)|$.

For $\mathrm{E}[N_-]$ we can show that $\beta_-(\lambda,j)\to 1-\frac{1}{2^{\lambda}}$ with n approaching infinity what can be bounded from above by $\mathrm{e}^{-\frac{1}{2^{\lambda}}}$. Thus we obtain

$$\mathrm{E}[N_-] \le \sum_{\lambda=0}^{n}\binom{n}{\lambda}\mathrm{e}^{-\frac{m}{2^{\lambda}}} \le 2^{n}\mathrm{e}^{-\frac{m}{2^{n}}}.$$

Theorem 1. *The expected number of models of mixed Horn formulas in $MH(n,m)$ and $MH_-(n,m)$ is bounded from above by respectively*

$$\mathrm{E}[N] \le 2^{n}\mathrm{e}^{\frac{-2m(1+n)}{(n+2)2^{n}}} \ \textit{and}\ \ \mathrm{E}[N_-] \le 2^{n}\mathrm{e}^{-\frac{m}{2^{n}}}.$$

Observe that since N is a non-negative integer-valued random variable, we can apply Markov's inequality in order to obtain an upper bound for the probability that $N>0$, i.e.,

$$\Pr(N>0)=\Pr(N\ge 1)\le \mathrm{E}[N].$$

Unfortunately, the bounds from Theorem 1 are too rough to obtain useful upper bounds for $\Pr(N>0)$. Hence, is it an interesting question whether both $\mathrm{E}[N]$ and $\mathrm{E}[N_-]$ can be bounded more tightly from above.

Lower bound for $\Pr(N_->0)$

Large values for $\mathrm{E}[N]$ do not imply that $\Pr(N=0)$ is small. Since the first moment $\mathrm{E}[N]$ gives only a rough upper bound for satisfiability of $MH(n,m)$, one can investigate the lower bound for $\Pr(N>0)$ by considering the second moment method. More specifically, by a direct application of the Cauchy-Schwartz inequality for non-negative random variables N and X ($X=1$ if $N>0$, otherwise 0), we obtain that

$$\Pr(N>0)\ge \frac{\mathrm{E}[N]^2}{\mathrm{E}[N^2]}$$

where the second moment

$$\begin{aligned}
\mathrm{E}[N^2] &= p\sum_{C}\left(\sum_{t}t(C)\right)^2\\
&= p\sum_{C}\left(\sum_{t}t(C)+2\sum_{t_1\ne t_2}t_1(C)t_2(C)\right)\\
&= \mathrm{E}[N]+2p\sum_{C}\sum_{t_1\ne t_2}t_1(C)t_2(C)\\
&=: \mathrm{E}[N]+2\alpha. \qquad (2)
\end{aligned}$$

Using the same arguments, we can also derive the lower bound for the probability $\Pr(N_->0)$ that most of the formulas in $MH_-(n,m)$ are satisfiable. That is

$$\Pr(N_->0)\ge \frac{\mathrm{E}[N_-]^2}{\mathrm{E}[N_-]+2\alpha_-} \qquad (3)$$

where α_- is defined in a similar way as in (2).

We deliver the lower bound only for $\Pr(N_->0)$. To this end, we proceed first with the computation of α_-. Again, we estimate it for large values of n. Thus we write

$$\begin{aligned}
\alpha_- &= \frac{1}{|MH_-(n,m)|}\sum_{t_1\ne t_2}\sum_{C}t_1(C)t_2(C)\\
&= \frac{1}{|MH_-(n,m)|}\sum_{t_1\ne t_2}N(t_1,t_2),
\end{aligned}$$

where $N(t_1,t_2)$ denotes the number of formulas from $MH_-(n,m)$ satisfied simultaneously by a pair of truth assignments t_1 and t_2.

In order to compute $N(t_1,t_2)$ for a fixed pair of unequal truth assignments t_1 and t_2, imagine each of them as a sequence of n bits, where each bit corresponds to some variable x_i from V. Without loss of generality, the order of bits is

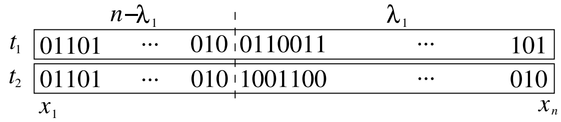

Figure 1: Example of two truth assignments with $n - \lambda_1$ matching bits.

defined by the vector $(x_1, ..., x_n)$, and is the same for all sequences. For any pair of bit sequences we denote by $\lambda_1 \le n$ the number of non-matching (unequal) bits and by $n - \lambda_1$ the number of matching bits (see Figure 1). Note that the matching (non-matching) bits do not have to form a consecutive block of size $n - \lambda_1$ (λ_1) within the bit sequence they belong to. Furthermore, denote by λ_2 the number of bits equal to 0 among the $n - \lambda_1$ equal bits of t_1. Observe that λ_2 has the same value for t_2. Similarly, denote by λ_3 the number of bits equal to 1 among the λ_1 unequal bits of t_1. Note, that the number of bits equal to 1 among the λ_1 unequal bits in t_2 is $\lambda_1 - \lambda_3$.

Obviously, the triples $(\lambda_1, \lambda_2, \lambda_3)$ provide a partition into classes of the set of all pairs of distinct truth assignments. The number of formulas from $MH_-(n, m)$ satisfied by any pair of a fixed class is constant. Thus we can write

$$\alpha_- = \frac{1}{|MH_-(n,m)|} \sum_{\lambda_1=1}^{n} \binom{n}{\lambda_1} \sum_{\lambda_2=0}^{n-\lambda_1} \binom{n-\lambda_1}{\lambda_2} \sum_{\lambda_3=0}^{\lambda_1-1} \binom{\lambda_1-1}{\lambda_3} N(\lambda_1, \lambda_2, \lambda_3),$$

where $N(\lambda_1, \lambda_2, \lambda_3)$ denotes the number of formulas from $MH_-(n, m)$ satisfied by any pair of distinct truth assignments specified by λ_1, λ_2, and λ_3. This number can be traced back to the number of clauses from $P(V) \cup H_-(V)$ satisfied by those truth assignments. More specifically,

$$N(\lambda_1, \lambda_2, \lambda_3) = \binom{|P(V)| + |H_-(V)| - \Delta(\lambda_1, \lambda_2, \lambda_3)}{m},$$

where $\Delta(\lambda_1, \lambda_2, \lambda_3)$ denotes the number of clauses from $P(V) \cup H_-(V)$ not satisfied by those truth assignments. Here, $m \le |P(V)| + |H_-(V)| - \Delta(\lambda_1, \lambda_2, \lambda_3)$. Furthermore, note that

$$|MH_-(n,m)| = \binom{|P(V)| + |H_-(V)|}{m}. \qquad (4)$$

It can now be determined by elementary considerations that

$$\begin{aligned}\Delta(\lambda_1, \lambda_2, \lambda_3) = & \binom{\lambda_2}{2} + \sum_{i=1}^{n-\lambda_1-\lambda_2} \binom{n-\lambda_1-\lambda_2}{i} + \\ & \binom{\lambda_3}{2} + \binom{\lambda_1-\lambda_3}{2} + \sum_{i=1}^{\lambda_3} \binom{\lambda_3}{i} + \\ & \sum_{i=1}^{\lambda_1-\lambda_3} \binom{\lambda_1-\lambda_3}{i} + \lambda_1\lambda_2 + \\ & \sum_{i=1}^{n-\lambda_1-\lambda_2} \binom{n-\lambda_1-\lambda_2}{i} \left(\sum_{j=1}^{\lambda_3} \binom{\lambda_3}{j} \right. \\ & \left. + \sum_{j=1}^{\lambda_1-\lambda_3} \binom{\lambda_1-\lambda_3}{j} \right),\end{aligned}$$

The minimal value of this function is

$$\Delta_{\min} := \Delta(1, n-1, 0) = \Delta(2, n-2, 1) = \frac{n(n-1)}{2} + 1.$$

as determined by a computer algebra system. Inserting these results and (4) into α_-, we obtain

$$\begin{aligned}\alpha_- = & \sum_{\lambda_1=1}^{n} \binom{n}{\lambda_1} \sum_{\lambda_2=0}^{n-\lambda_1} \binom{n-\lambda_1}{\lambda_2} \sum_{\lambda_3=0}^{\lambda_1-1} \binom{\lambda_1-1}{\lambda_3} \\ & \prod_{j=0}^{m-1} \left(1 - \frac{\Delta(\lambda_1, \lambda_2, \lambda_3)}{|P(V)| + |H_-(V)| - j}\right) \\ \le & \sum_{\lambda_1=1}^{n} \binom{n}{\lambda_1} \sum_{\lambda_2=0}^{n-\lambda_1} \binom{n-\lambda_1}{\lambda_2} \sum_{\lambda_3=0}^{\lambda_1-1} \binom{\lambda_1-1}{\lambda_3} \\ & \mathrm{e}^{-\frac{m\Delta_{\min}}{|P(V)|+|H_-(V)|}} \\ = & 2^{n-1}(2^n - 1)\mathrm{e}^{-\frac{m\Delta_{\min}}{|P(V)|+|H_-(V)|}}.\end{aligned}$$

In order to proceed with the estimation of (3), we need the lower bound for $\mathrm{E}[N_-]$. According to the discussion on $\mathrm{E}[N_-]$ from the previous section, we have for large values of n:

$$\mathrm{E}[N_-] = \sum_{\lambda=0}^{n} \binom{n}{\lambda} \left(1 - \frac{1}{2^\lambda}\right)^m \ge \frac{2^n - 1}{2^m}.$$

By inserting this result and the estimation for α_- into (3), we obtain for large n:

$$\begin{aligned}\frac{\mathrm{E}[N_-]^2}{\mathrm{E}[N_-] + 2\alpha_-} & \ge \frac{\mathrm{E}[N_-]^2}{\mathrm{E}[N_-] + 2^n(2^n-1)\mathrm{e}^{-\frac{m\Delta_{\min}}{|P(V)|+|H(V)_-|}}} \\ & \ge \frac{2^n - 1}{2^m\left(1 + 2^n 2^m \mathrm{e}^{\frac{-m\Delta_{\min}}{|P(V)|+|H(V)_-|}}\right)} \\ & \ge \frac{2^n - 1}{2^m(1 + 2^n 2^m)}.\end{aligned}$$

Finally, we are ready to give the lower bound for $\Pr(N_- > 0)$.

Theorem 2. *The probability that most of the negative Horn formulas from $MH_{-}(n, m)$ are satisfiable is bounded from below by*

$$\frac{2^n - 1}{2^m(1 + 2^n 2^m)} \leq \Pr(N_{-} > 0).$$

Conclusion and open problems

We have investigated by means of a probabilistic analysis the upper bounds for the expected number of models for general random MHFs as well as for random negative MHFs. We have also derived a lower bound for the probability that formulas from the latter class with uniform distribution of instances and parameterized by n and m are satisfiable. With our theoretical study, we hope to shed more light onto random MHFs. However, for the localization of the phase transition in MHFs by the methodology of this paper, one need better bounds on the expected number of models than the ones proved here. Whether such bound estimations exist remains an interesting open question.

Furthermore, similar to the more general random k-SAT, we can formulate the satisfiability threshold conjecture for $MH_k^{-}(n, m)$, where the length of the negative Horn clauses is restricted by $k \leq n$, as follows: For each $k > 0$ there exists a constant h_k such that

$$\lim_{n \to \infty} \Pr(MH_k^{-}(n, hn) \text{ is satisfiable}) = \begin{cases} 1 & \text{if } h < h_k, \\ 0 & \text{if } h > h_k. \end{cases}$$

It would be interesting to establish the existence of h_k both for general k as well as for some specific k. Moreover, proving an upper and a lower bound for h_k is also a highly desirable goal from the theoretical point of view. We believe that our results should shed more light on the sharp threshold phenomenon observed already experimentally for other subclasses of MHFs by Namasivayam and Truszczyński (2010).

References

Achlioptas, D., and Moore, C. 2006. Random k-SAT: Two moments suffice to cross a sharp threshold. *SIAM Journal on Computing* 36(3):740–762.

Achlioptas, D., and Peres, Y. 2004. The threshold for random k-SAT is $2^k \log 2 - o(k)$. *Journal of ACM* 17(4):947–973.

Chao, M.-T., and Franco, J. 1990. Probabilistic analysis of a generalization of the unit-clause literal selection heuristics for the k satisfiability problem. *Information Processing Letters* 51(3):289–314.

Cheeseman, P.; Kanefsky, B.; and Taylor, W. M. 1991. Where the really hard problems are. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, 331–337.

Franco, J., and Paull, M. 1983. Probabilistic analysis of the Davis Putnam procedure for solving the satisfiability problem. *Discrete Applied Mathematics* 5:77–87.

Goldberg, A.; Purdom, P. W.; and Brown, C. A. 1982. Average time analyses of simplified Davis-Putnam procedures. *Information Processing Letters* 15:72–75.

Kottler, S.; Kaufmann, M.; and Sinz, C. 2008. A new bound for an NP-hard subclass of 3-SAT using backdoors. In *Proceedings of 11th International Conference on Theory and Applications of Satisfiability Testing (SAT'08)*, volume 4996 of *Lecture Notes in Computer Science*, 161–167.

Mitchell, D.; Selman, B.; and Levesque, H. 1992. Hard and easy distributions of SAT problems. In *Proceedings of the 10th National Conference on Artificial Intelligence*, 459–465.

Namasivayam, G., and Truszczyński, M. 2010. Simple but hard mixed Horn formulas. In *Proceedings of the 13th International Conference on Theory and Applications of Satisfiability Testing (SAT'10)*, volume 6175 of *Lecture Notes in Computer Science*, 382–387.

Porschen, S., and Speckenmeyer, E. 2007. Satisfiability of mixed Horn formulas. *Discrete Applied Mathematics* 155(11):1408–1419.

Porschen, S.; Schmidt, T.; and Speckenmeyer, E. 2009. On some aspects of mixed Horn formulas. In *Proceedings of the 12th International Conference on Theory and Applications of Satisfiability Testing (SAT'09)*, volume 5584 of *Lecture Notes in Computer Science*, 86–100.

Purdom, P. W. 1983. Search rearrangement backtracking and polynomial average time. *Artificial Intelligence* 21:117–133.

Schuh, B. R. 2012. Phase transition in unrestricted random SAT. *CoRR* abs/1204.1656.

Wotzlaw, A.; Speckenmeyer, E.; and Porschen, S. 2012. Generalized k-ary tanglegrams on level graphs: A satisfiability-based approach and its evaluation. *Discrete Applied Mathematics* 160(16-17):2349–2363.