Approximating Highly Satisfiable Random Max-2-SAT

Andrei A. Bulatov, Cong Wang
Simon Fraser University

SAT 2014, Vienna

Random Max-k-SAT

Let $\Phi_k(n, dn)$, d constant denote the random distribution on the set of k-CNFs with exactly dn clauses

Max-k-SAT:

Instance: A k-CNF sampled according to $\Phi_k(n, dn)$

Objective: Satisfy as many clauses as possible

Phase Transition and Friends

Phase Transition Conjecture: For any k there is d(k) s.t. for any d

- if $d < d(k)$ then $\varphi \in \Phi_k(n, dn)$ is satisfiable whp
- if $d > d(k)$ then $\varphi \in \Phi_k(n, dn)$ is unsat whp

Problem 1:

If $d > d(k)$, how many clauses are satisfiable whp?

A k-CNF $\varphi \in \Phi_k(n, dn)$ is said to be **p-satisfiable**, if there is an assignment satisfying $\left((1 - 2^{-k})d + p\right)n$ clauses

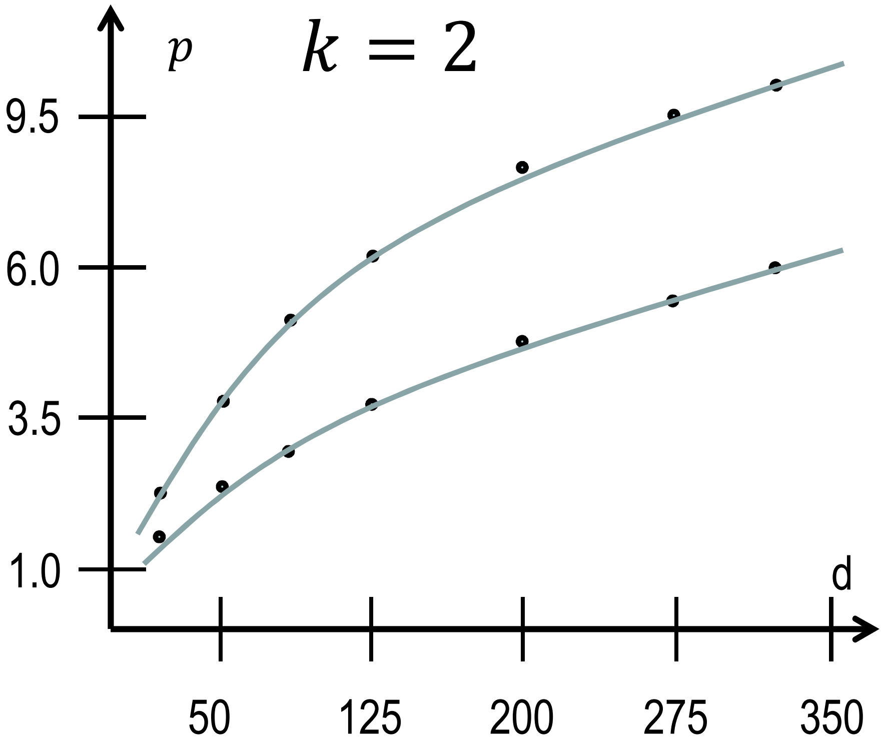
Let $p(k, d)$ be the number that

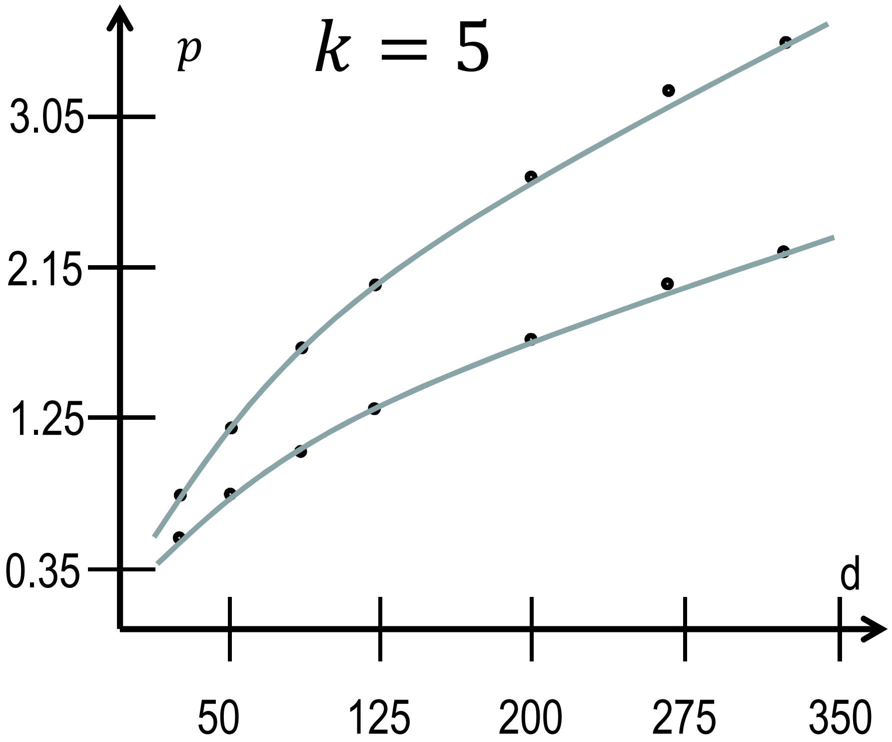
- for $p < p(k, d)$ a CNF $\varphi \in \Phi_k(n, dn)$ is p-sat whp
- for $p > p(k, d)$ a CNF $\varphi \in \Phi_k(n, dn)$ is not p-sat whp

Phase Transition and Friends II

Achlioptas et al.

$$(1-\delta_k)\Psi\left(\frac{2^k \log 2}{d}\right) d2^{-k} \leq p(k,d) \leq \Psi\left(\frac{2^k \log 2}{d}\right) d2^{-k}$$

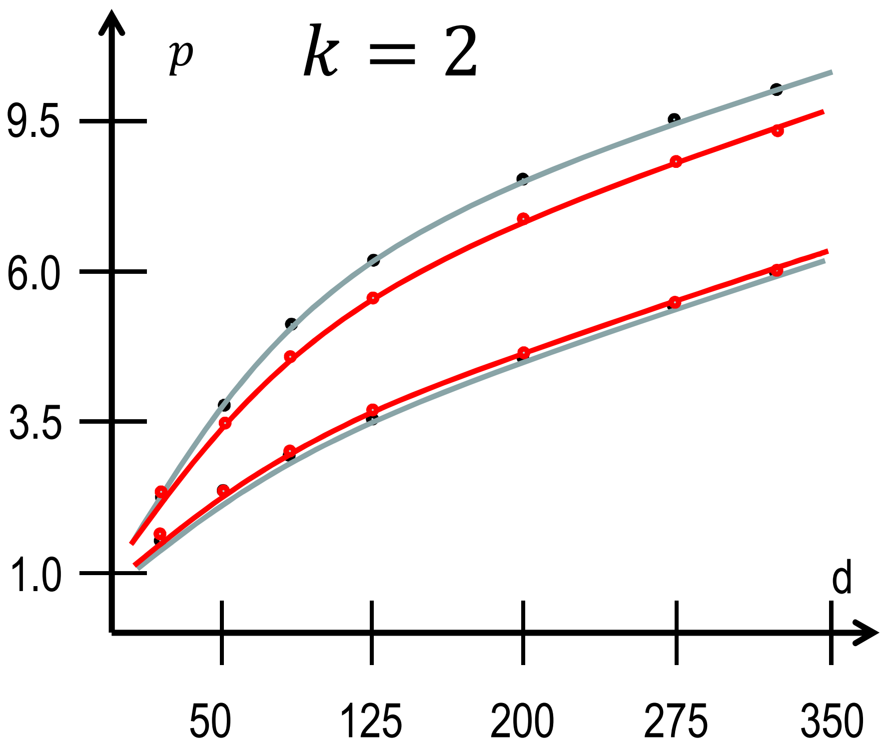




Phase Transition and Friends III

Gamarnik et al.

$$\frac{1}{3}\sqrt{d} \leq p(2,d) \leq \frac{1}{2}\sqrt{d}$$



Solving Random (Max)-SAT

$\Phi_k^{sat}(n, dn)$ is the uniform distribution of satisfiable CNFs
$\Phi_k^{pl}(n, dn)$ is the planted distribution

	k-SAT	Max-k-SAT
Worst case	NP-hard	NP-hard to approximate $\frac{22}{21} - \varepsilon$ for Max-2-SAT $\frac{8}{7} - \varepsilon$ for Max-3-SAT (Hastad, 2001)
Random uniform	Trivial whp, except the area around the threshold	Max-3-SAT is approximable within 1.0957 (de la Vega, Karpinski, 2007)
Satisfiable, planted	The two distribution are similar Can be solved in polytime whp (Coja-Oghlan et al., 2007)	? Problem 2

New Distributions for Max-2-SAT

$\Phi_k^{sat}(n, dn, p)$, $p \in [0, 2^{-k}d]$, the uniform distribution of p-satisfiable CNFs

$\Phi_k^{pl}(n, dn, p)$, $p \in [0, 2^{-k}d]$, the uniform distribution of planted CNFs

Remarks:

- $\Phi_k^{sat}(n, dn, 0) = \Phi_k(n, dn)$,
- $\Phi_k^{sat}\left(n, dn, 2^{-k}d\right) = \Phi_k^{sat}(n, dn)$,
- $\Phi^{sat}(n, dn, p) = \Phi_2^{sat}(n, dn, p)$,

same for planted

- A planted assignment in $\Phi_k^{pl}(n, dn, p)$ is assumed all-ones

Sat vs. Planted

$\mathrm{P}^{sat}_{d,p}(A)$ is the probability of event A under $\Phi^{sat}(n, dn, p)$

$\mathrm{P}^{pl}_{d,p}(A)$ is the probability of event A under $\Phi^{pl}(n, dn, p)$

Theorem

For any $p \in [0, \frac{d}{4}]$ and any event A there is $p' \in [p, \frac{d}{4}]$

$$\mathrm{P}^{sat}_{d,p}(A) \leq \left(\frac{d}{4} - p\right)n \cdot \exp\left[n \cdot \log\left(1 + \exp\left[-\frac{\tau p'^2}{d}\right]\right)\right] \mathrm{P}^{pl}_{d,p'}(A)$$

Approximation Algorithm: Outline

Given a 2-CNF φ:

Step 1. Take some reasonably good assignment of φ
We use Majority Vote to do that

Step 2. Unassign those variables that look suspicious

Suspicious are those that don't have strong support, the number of clauses satisfied only due to those variables

Step 3. Brute force the set of unassigned variables

It is possible, since this set is small and the subformula they induce whp has connected components of size $\log n$

Support and Cores

We consider $p > \gamma\sqrt{d \log d}$, $\gamma = O(1)$, not necessarily a constant

Variable x **supports** clause C w.r.t. an assignment f, if C is satisfied, but only by x

A set of variables W is a **core** w.r.t. assignment f, if every variable in W supports at least $\frac{d}{2}$ clauses with var's from W

W is a **large core** if $|W| \geq (1 - d^{-\nu})n$, $\nu > 2$.

Majority Vote

Majority Vote:

Given CNF φ, for each variable x count the number of positive and negative occurrences of x, assign correspondingly

Break ties at random

Proposition

Let $\varphi \in \Phi^{sat}(n, dn, p)$. Then

(a) whp Majority Vote produces a `highly satisfiable' assignment

(b) There is a large core w.r.t. to the assignment found

Unassigning

Let f be the assignment found by Majority Vote

Unassign all var's that support substantially fewer than $\frac{d}{2}$ clauses

Let h be the resulting partial assignment defined on W

Proposition

Let $\varphi \in \Phi^{pl}(n, dn, p)$ and f_0 the planted assignment.

(a) W contains a large core w.r.t. f_0

(b) W whp contains no more than $d^{-1/2}n$ var's assigned wrongly

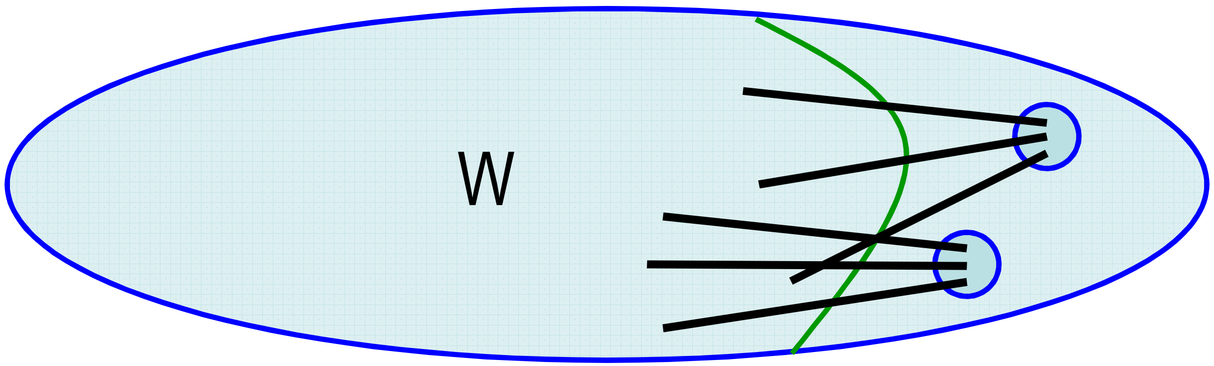
Corollary

The same is true for $\varphi \in \Phi^{sat}(n, dn, p)$ and for some p-satisfying assignment instead of the planted one.

Complement of a Core

Proposition

Let $\varphi \in \Phi^{sat}(n, dn, p)$, $p = \gamma\sqrt{d \log d}$, $\gamma > 2$, where d is sufficiently large. Let W be a core w.r.t. a p-satisfying assignment and such that $|W| \geq (1 - d^{-\nu})n$, $\nu > 2$. Then whp the largest connected component in subformula of induced by W is of size at most $\log n$.



Results

Theorem 1

There is a poytime algorithm that for any $\varepsilon > 0$ and any $p > 2\sqrt{2}\sqrt{d \log d}$, given $\varphi \in \Phi^{pl}(n, dn, p)$ whp finds a $(p - \varepsilon)$-satisfying assignment.

Theorem 2

There is an algorithm that for any $p = \gamma\sqrt{d \log d}$, $\gamma > 2\sqrt{2}$ given $\varphi \in \Phi^{sat}(n, dn, p)$ whp finds a p'-satisfying assignment, $p' \geq p - \varepsilon$, where ε is a polynomial in d of degree $1 - O(-\gamma^2)$.

Majority Vote and Uniform Distribution

Let $p_0 = p(2, d)$. Then $\Phi(n, dn)$ is almost $\Phi^{sat}(n, dn, p_0)$.

The number of clauses satisfied by MV on $\varphi \in \Phi(n, dn)$ is highly concentrated around some value $\beta(d)$.

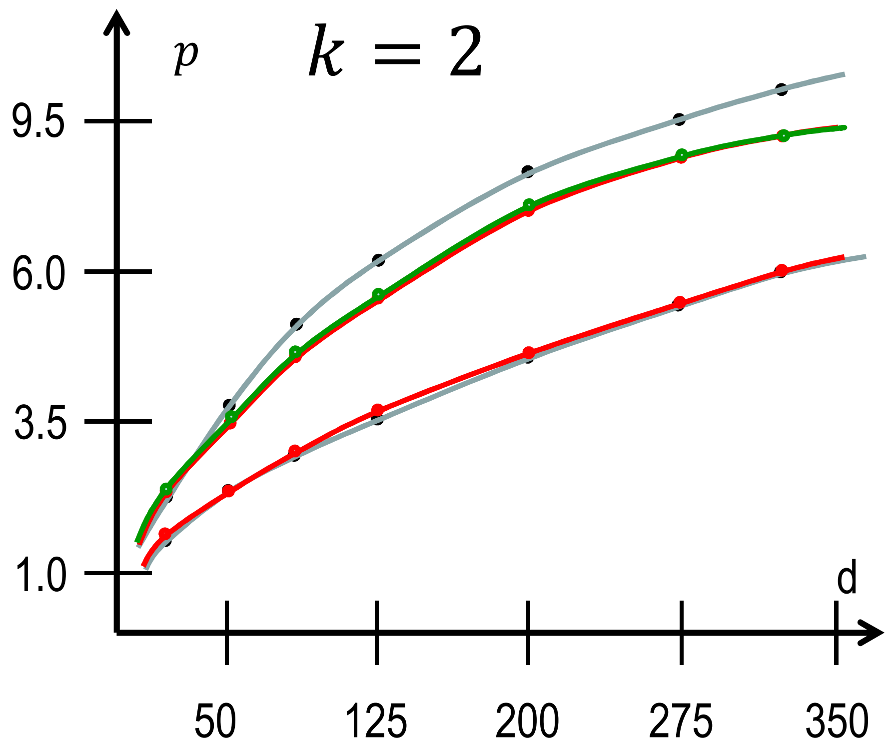
The number of clauses satisfied by MV on $\varphi \in \Phi^{\mathrm{pl}}(n, dn, p)$ is also highly concentrated around some value $\beta(d, p)$.
And so is the number of clauses satisfied by MV on $\varphi \in \Phi^{\mathrm{sat}}(n, dn, p)$

We should have

$$\beta(d) = \beta(d, p_0)$$

Number of Satisfiable Clauses

This is what we have:



Thank you!