

Approximating Maximum Constraint Satisfaction Problems

Johan Håstad
some slides by Per Austrin



**KTH Computer Science
and Communication**

Shafi and Silvio celebration, December 10, 2013

Today we are celebrating a Turing Award.

December 10th is the day of the Nobel prizes.

Today we are celebrating a Turing Award.

December 10th is the day of the Nobel prizes.

Maybe the Israeli dress code is a bit too relaxed for a merger.

The personal angle

You do not appreciate your parents until your own kids have grown up.

You do not appreciate your advisor until you have graduated a number of your own students

What was the best?

- Freedom to do what I wanted.
- Comments on talks and papers.

Back to business.

This is a survey talk but I focus on the story.

Will fail to acknowledge some critical contributions.

Essentially no technical details.

The MAX k -SAT problem:

Given: k -CNF formula

Goal: satisfy as many clauses as possible

(**Note:** *exactly* k literals in every clause)

E.g. MAX 3-SAT:

$$\begin{aligned}(x_1 \vee x_2 \vee x_3) &\wedge (\bar{x}_1 \vee x_2 \vee \bar{x}_4) &&\wedge \\(x_1 \vee x_3 \vee \bar{x}_4) &\wedge (x_2 \vee \bar{x}_3 \vee \bar{x}_4) &&\wedge \\(\bar{x}_2 \vee x_3 \vee \bar{x}_4) &\wedge (\bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4) &&\wedge \\(x_2 \vee x_3 \vee x_5) &\wedge (\bar{x}_2 \vee \bar{x}_3 \vee x_5) &&\wedge \\(\bar{x}_2 \vee x_4 \vee x_5) &\wedge (x_3 \vee x_4 \vee \bar{x}_5) &&\wedge \\(\bar{x}_3 \vee x_4 \vee x_5) &\wedge (\bar{x}_3 \vee x_4 \vee \bar{x}_5)\end{aligned}$$

MAX k -SAT

The MAX k -SAT problem:

Given: k -CNF formula

Goal: satisfy as many clauses as possible

(*Note:* exactly k literals in every clause)

E.g. MAX 3-SAT:

$$\begin{array}{llll} (x_1 \vee x_2 \vee x_3) \wedge (\bar{x}_1 \vee x_2 \vee \bar{x}_4) \wedge & & x_1 & = & \text{FALSE} \\ (x_1 \vee x_3 \vee \bar{x}_4) \wedge (x_2 \vee \bar{x}_3 \vee \bar{x}_4) \wedge & & x_2 & = & \text{FALSE} \\ (\bar{x}_2 \vee x_3 \vee \bar{x}_4) \wedge (\bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4) \wedge & & x_3 & = & \text{TRUE} \\ (x_2 \vee x_3 \vee x_5) \wedge (\bar{x}_2 \vee \bar{x}_3 \vee x_5) \wedge & & x_4 & = & \text{FALSE} \\ (\bar{x}_2 \vee x_4 \vee x_5) \wedge (x_3 \vee x_4 \vee \bar{x}_5) \wedge & & x_5 & = & \text{FALSE} \\ (\bar{x}_3 \vee x_4 \vee x_5) \wedge (\bar{x}_3 \vee x_4 \vee \bar{x}_5) & & & & \end{array}$$

NP-hard; **Approximation ratio**,

$$\alpha = \frac{\text{Value}(\text{Found solution})}{\text{Value}(\text{Best solution})}$$

worst case over all instances.

$\alpha = 1$ the same as finding optimal solution, otherwise $\alpha < 1$.

Approximating MAX k -SAT

Trivial algorithm: Pick random assignment.

Approximation ratio $3/4$ for **MAX 2-SAT**.

In fact satisfies fraction $3/4$ of clauses.

Approximating MAX k -SAT

Trivial algorithm: Pick random assignment.

Approximation ratio $3/4$ for **MAX 2-SAT**.

In fact satisfies fraction $3/4$ of clauses.

Surely something smarter can be done?

Approximating MAX k -SAT

Trivial algorithm: Pick random assignment.

Approximation ratio $3/4$ for **MAX 2-SAT**.

In fact satisfies fraction $3/4$ of clauses.

Surely something smarter can be done?

Yes: Goemans-Williamson [GW95] used semi-definite programming to obtain **0.878** approximation for MAX 2-SAT.

Approximating MAX k -SAT

Trivial algorithm: Pick random assignment.

Approximation ratio $7/8$ for MAX 3-SAT.

Approximating MAX k -SAT

Trivial algorithm: Pick random assignment.

Approximation ratio $7/8$ for MAX 3-SAT.

Surely something smarter can be done?

Approximating MAX k -SAT

Trivial algorithm: Pick random assignment.

Approximation ratio $7/8$ for MAX 3-SAT.

Surely something smarter can be done?

No: NP-hard to achieve ratio $7/8 + \epsilon$ [H01]

- Random assignment gives optimal approximation ratio!

CSPs defined by a predicate P

For predicate $P : \{0, 1\}^k \rightarrow \{0, 1\}$, the MAX CSP(P) problem:

Given: set of constraints, each of the form

$$P(l_1, l_2, \dots, l_k) = 1, \text{ literals } l_1, \dots, l_k$$

Goal: satisfy as many constraints as possible

E.g. MAX CSP($x_1 \vee x_2 \vee x_3$) = MAX 3-SAT

$$\begin{aligned}
& P(x_1, \bar{x}_2, x_6) \wedge P(x_1, x_2, x_3) \wedge P(\bar{x}_1, x_2, \bar{x}_4) \wedge \\
& P(\bar{x}_1, x_4, x_7) \wedge P(x_1, x_3, \bar{x}_4) \wedge P(x_2, \bar{x}_3, \bar{x}_4) \wedge \\
& P(x_1, \bar{x}_3, x_6) \wedge P(\bar{x}_2, x_3, \bar{x}_4) \wedge P(\bar{x}_2, \bar{x}_3, \bar{x}_4) \wedge \\
& P(\bar{x}_4, x_6, \bar{x}_7) \wedge P(x_2, x_3, x_5) \wedge P(\bar{x}_2, \bar{x}_3, x_5) \wedge \\
& P(x_3, \bar{x}_6, x_7) \wedge P(\bar{x}_2, x_4, x_5) \wedge P(x_3, x_4, \bar{x}_5) \wedge \\
& P(x_4, x_5, x_7) \wedge P(\bar{x}_3, x_4, x_5) \wedge P(\bar{x}_3, x_4, \bar{x}_5)
\end{aligned}$$

Satisfy as many as possible.

Approximation Resistance

Trivial algorithm: Pick random assignment.

Approximation ratio $|P^{-1}(1)|/2^k$.

P is *approximation resistant* if hard to do better.

Alternative formulation

P is approximation resistant iff, for any $\epsilon > 0$, it is hard to distinguish

$(1 - \epsilon)$ -satisfiable instances

and

$(|P^{-1}(1)|/2^k + \epsilon)$ -satisfiable instances.

Understanding Approximation Resistance

Overall goal: understand structure of resistant predicates.

– when is non-trivial approximation possible?

To make life simple:

- Boolean variables.
- Same predicate in each constraint.

Alternative view, non-resistant

Given a set of k -tuples of literals.

Promise: There is an assignment to the variables such that $(1 - \epsilon)$ fraction of the resulting k -bit strings satisfy P .

Achieved: An assignment to the variables such that more than the expected ratio of the k -bit strings satisfy P .

Given a set of k -tuples of literals.

Promise: There is an assignment to the variables such that $(1 - \epsilon)$ fraction of the resulting k -bit strings satisfy P .

Achieved: An assignment to the variables such that the distribution of k -bit strings is noticeably far from uniform.

$$\begin{aligned} &P(x_1, \bar{x}_2, x_6) \wedge P(x_1, x_2, x_3) \wedge P(\bar{x}_1, x_2, \bar{x}_4) \wedge \\ &P(\bar{x}_1, x_4, x_7) \wedge P(x_1, x_3, \bar{x}_4) \wedge P(x_2, \bar{x}_3, \bar{x}_4) \wedge \\ &P(x_1, \bar{x}_3, x_6) \wedge P(\bar{x}_2, x_3, \bar{x}_4) \wedge P(\bar{x}_2, \bar{x}_3, \bar{x}_4) \wedge \\ &P(\bar{x}_4, x_6, \bar{x}_7) \wedge P(x_2, x_3, x_5) \wedge P(\bar{x}_2, \bar{x}_3, x_5) \wedge \\ &P(x_3, \bar{x}_6, x_7) \wedge P(\bar{x}_2, x_4, x_5) \wedge P(x_3, x_4, \bar{x}_5) \wedge \\ &P(x_4, x_5, x_7) \wedge P(\bar{x}_3, x_4, x_5) \wedge P(\bar{x}_3, x_4, \bar{x}_5) \end{aligned}$$

The k -bit strings

$$\begin{aligned} & (x_1, \bar{x}_2, x_6) \quad , \quad (x_1, x_2, x_3) \quad , \quad (\bar{x}_1, x_2, \bar{x}_4) \quad , \\ & (\bar{x}_1, x_4, x_7) \quad , \quad (x_1, x_3, \bar{x}_4) \quad , \quad (x_2, \bar{x}_3, \bar{x}_4) \quad , \\ & (x_1, \bar{x}_3, x_6) \quad , \quad (\bar{x}_2, x_3, \bar{x}_4) \quad , \quad (\bar{x}_2, \bar{x}_3, \bar{x}_4) \quad , \\ & (\bar{x}_4, x_6, \bar{x}_7) \quad , \quad (x_2, x_3, x_5) \quad , \quad (\bar{x}_2, \bar{x}_3, x_5) \quad , \\ & (x_3, \bar{x}_6, x_7) \quad , \quad (\bar{x}_2, x_4, x_5) \quad , \quad (x_3, x_4, \bar{x}_5) \quad , \\ & (x_4, x_5, x_7) \quad , \quad (\bar{x}_3, x_4, x_5) \quad , \quad (\bar{x}_3, x_4, \bar{x}_5) \end{aligned}$$

Within ϵ of the uniform distribution on k -bit strings?

P is Useless.

Given a set of k -tuples of literals, it is hard to distinguish.

Yes: There is an assignment to the variables such that $(1 - \epsilon)$ fraction of the resulting k -bit strings satisfy P .

No: For all assignments to the variables the distribution of k -bit strings is at most ϵ far from uniform.

Two easy facts

If P is useless then it is approximation resistant.

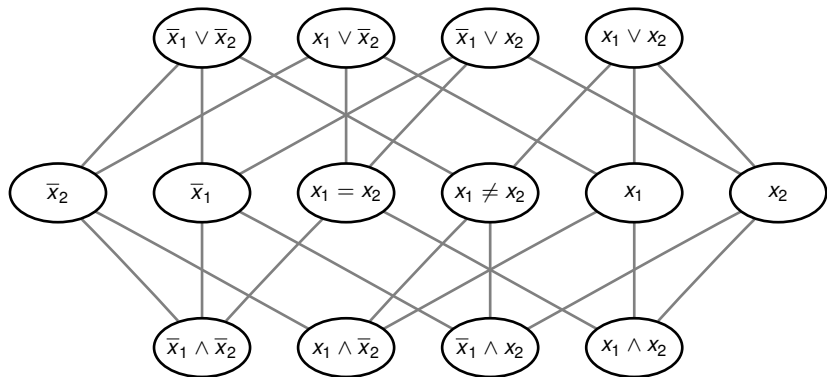
If P useless and P implies Q then Q is useless.

Most early approximation resistance proofs in fact proved uselessness.

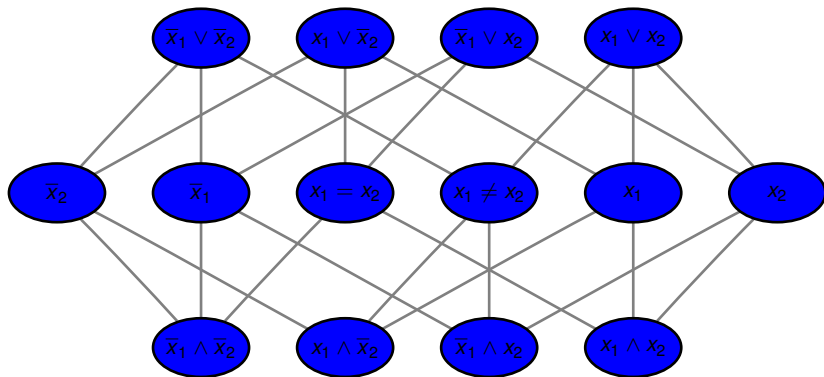
In particular 3-Lin [\[H01\]](#), i.e. $P(x) = x_1 \oplus x_2 \oplus x_3$ is useless.

Resistance classification

First we take a look at small arities of P and then we turn to asymptotic questions.

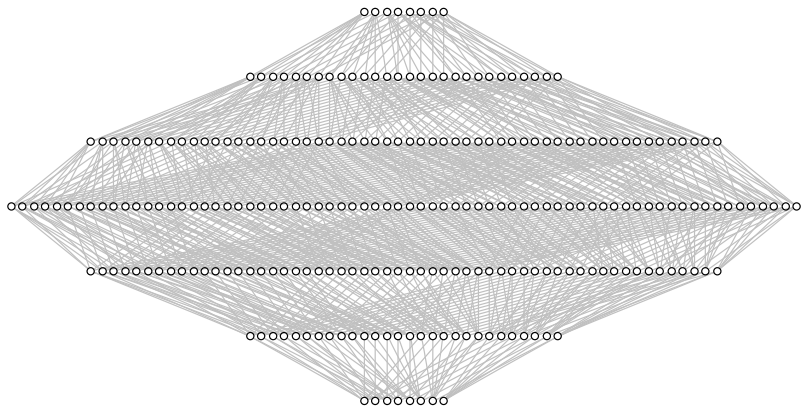


$k = 2$

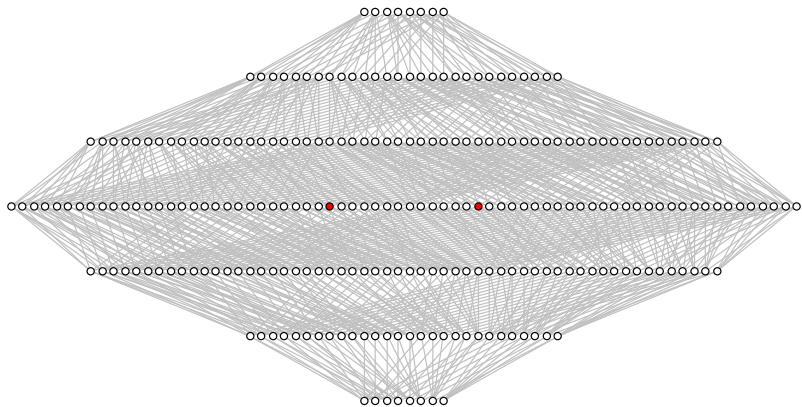


No predicate on *two* variables is resistant [GW95]

$k = 3$

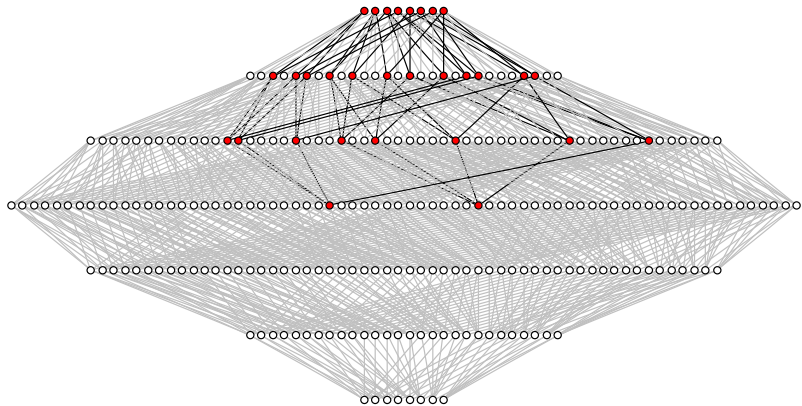


$k = 3$



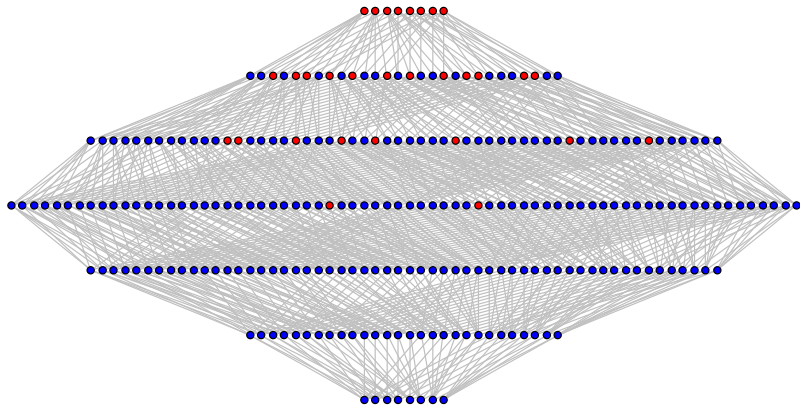
3-XOR is resistant [H01]

$k = 3$



3-XOR is useless [H01]

$k = 3$



Everything else has non-trivial approximation [Zwick99]

$$k = 4$$

weight	#approx	#resist	#unknown
15	0	1	0
14	0	4	0
13	1	4	1
12	3	15	1
11	9	11	7
10	26	22	2
9	27	6	23
8	52	16	6
7	50	0	6
6	50	0	0
5	27	0	0
4	19	0	0
3	6	0	0
2	4	0	0
1	1	0	0

Most predicates on *four* variables classified [Hast05]

- After eliminating symmetries a total of 400 predicates
- 79 resistant, 275 approximable, 46 unknown
- Little apparent structure (4Lin useless gives some but not all hardness)

Systematic Results?

We have two useful tools.

- “Approximability by nice low-degree expansion” using semidefinite programming.
- Prove very sparse predicates useless.

Fourier representation

Every $P : \{-1, 1\}^k \rightarrow \{0, 1\}$ has unique Fourier representation

$$P(x) = \sum_{S \subseteq [k]} \hat{P}(S) \prod_{i \in S} x_i.$$

Let $P^{=d}(x)$ be the part that is of degree d

$$P^{=1}(x) = \sum_{i=1}^k \hat{P}(\{i\}) x_i$$

$$P^{=2}(x) = \sum_{i < j} \hat{P}(\{i, j\}) x_i x_j$$

\vdots \vdots

Theorem ([Hast05])

Suppose there is a $C \in \mathbb{R}$ such that

$$C \cdot P^=1(x) + P^=2(x) > 0$$

for every $x \in P^{-1}(1)$. Then P is approximable.

The Theorem is somewhat more general allowing $C \cdot P^=1(x) + P^=2(x)$ to equal 0 on up to two accepting inputs.

Hast's Theorem implies the following:

Theorem ([Hast05])

Every predicate with fewer than $2^{\lceil \frac{k+1}{2} \rceil}$ accepting assignments is approximable.

Hast's Theorem implies the following:

Theorem ([Hast05])

Every predicate with fewer than $2^{\lceil \frac{k+1}{2} \rceil}$ accepting assignments is approximable.

Theorem ([Austrin-H09])

*Let $s \leq c \frac{k^2}{\log k}$. A uniformly random predicate with s satisfying assignments is *approximable with probability $1 - o_k(1)$.**

A very sparse predicate

“Hadamard predicate” of arity $k = 2^t - 1$, indexed by non-zero linear functions.

$Had(x)$ true iff exists $x^0 \in \{0, 1\}^t$, $x_L = L(x^0)$.

The sparsest linear subspace without constant or repeated coordinates.

Was proved useless assuming the Unique Games Conjecture (UGC) by Samorodnitsky and Trevisan [ST06].

Theorem

[Chan12] Had is useless for any t (arity $k = 2^t - 1$)

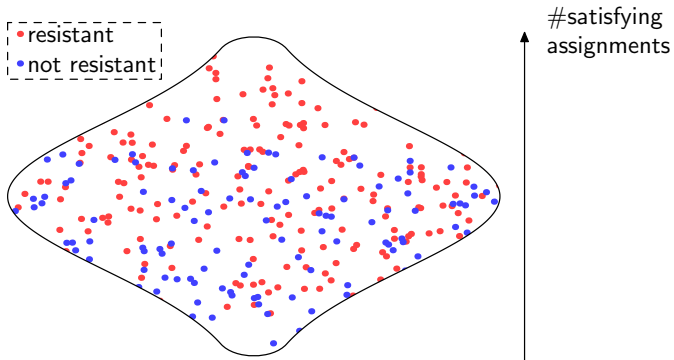
Note that $t = 2$ is 3Lin.

Had is very sparse and we get.

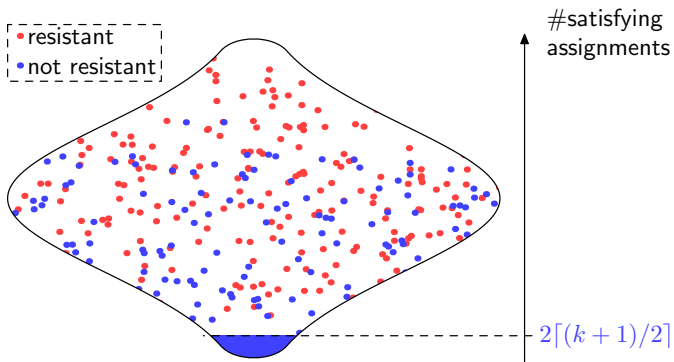
Theorem ([H07])

Let $s \geq c2^k/k^{1/2}$. A uniformly random predicate $P : \{0, 1\}^k \rightarrow \{0, 1\}$ with s satisfying assignments is useless with probability $1 - o_k(1)$.

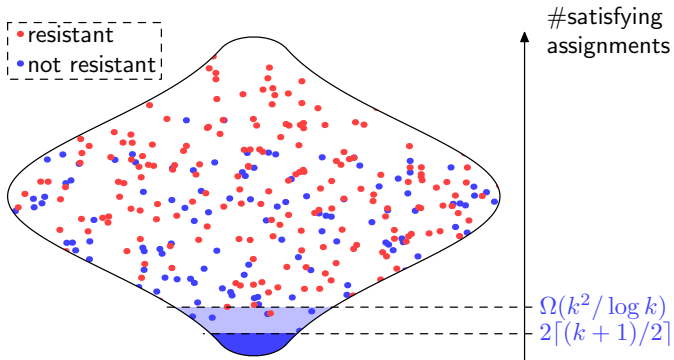
In Pictures



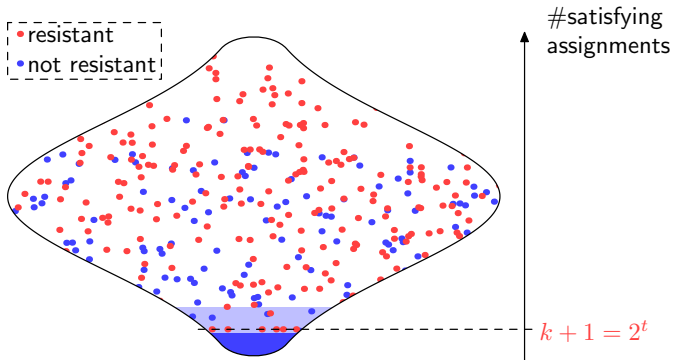
In Pictures



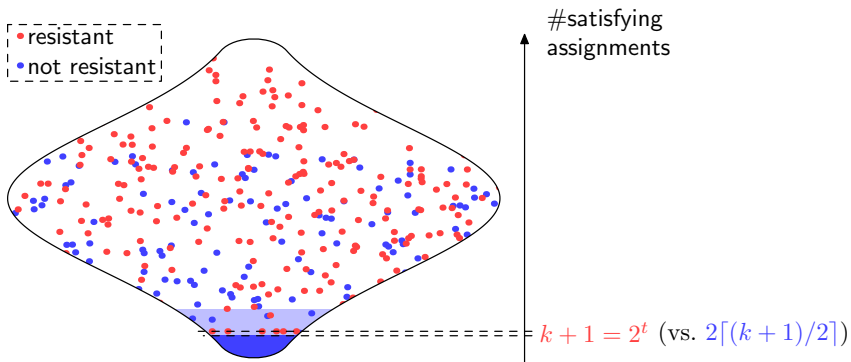
In Pictures



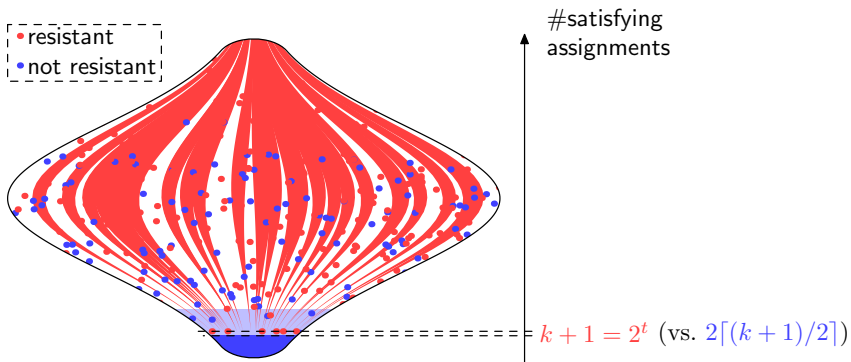
In Pictures



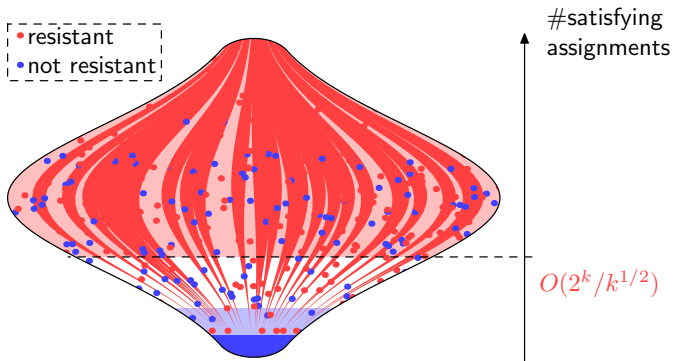
In Pictures



In Pictures



In Pictures



Not many more hardness results are known with NP-hardness and we turn to the Unique Games Conjecture (UGC).

Unique games conjecture?

UGC, made by Khot in 2002.

Constraint Satisfaction Problem $x_j \in [L]$

$$P_i(x_j, x_k) \Leftrightarrow (\pi_i(x_j) = x_k), \quad 1 \leq i \leq m,$$

distinguish whether optimal value is $(1 - \epsilon)m$ or ϵm .

Conjecture: NP-hard or at least not polynomial time.

Can have constraints on form

$$x_j - x_k \equiv c \pmod L$$

Vertex Cover is hard to approximate within $2 - \epsilon$, [Khot-Regev 03].

Optimal constant for Max-Cut [KKMO04].

Very useful for approximation resistance.

Pairwise independence

A distribution μ over $\{0, 1\}^k$ is *balanced* and *pairwise independent* if the marginal distributions on every pair of coordinates are uniform

E.g. μ_{\oplus}^3 : pick random $x \in \{0, 1\}^3$ s.t. $x_1 \oplus x_2 \oplus x_3 = 1$
 $\text{Supp}(\mu_{\oplus}^3) = \{001, 010, 100, 111\}$

Pairwise independence

A distribution μ over $\{0, 1\}^k$ is *balanced* and *pairwise independent* if the marginal distributions on every pair of coordinates are uniform

E.g. μ_{\oplus}^3 : pick random $x \in \{0, 1\}^3$ s.t. $x_1 \oplus x_2 \oplus x_3 = 1$
 $\text{Supp}(\mu_{\oplus}^3) = \{001, 010, 100, 111\}$

We say that $P : \{0, 1\}^k \rightarrow \{0, 1\}$ *contains* a balanced pairwise independent distribution μ over $\{0, 1\}^k$ if $\text{Supp}(\mu) \subseteq P^{-1}(1)$

Pairwise independence

Theorem ([Austrin-Mossel09])

Let $P : \{0, 1\}^k \rightarrow \{0, 1\}$ contain a balanced pairwise independent distribution. Then, assuming the UGC, P is useless.

We say that $P : \{0, 1\}^k \rightarrow \{0, 1\}$ **contains** a balanced pairwise independent distribution μ over $\{0, 1\}^k$ if $\text{Supp}(\mu) \subseteq P^{-1}(1)$

Pairwise independence

Theorem ([Austrin-Mossel09])

Let $P : \{0, 1\}^k \rightarrow \{0, 1\}$ contain a balanced pairwise independent distribution. Then, assuming the UGC, P is useless.

In fact this is necessary for uselessness [AH12].

Implications of AM

Theorem ([Austrin-Mossel09])

Assuming the UGC and the Hadamard Conjecture there *exist* hereditarily resistant predicates with $4^{\lceil \frac{k+1}{4} \rceil}$ accepting assignments for any k .

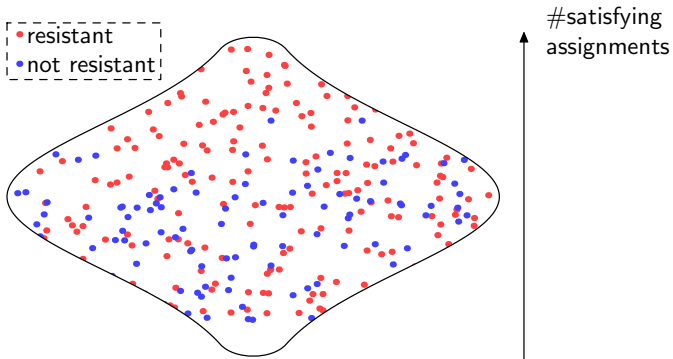
Theorem ([Austrin-H09])

Let $s \geq ck^2$. Assuming the UGC, a uniformly random predicate $P : \{0, 1\}^k \rightarrow \{0, 1\}$ with s satisfying assignments is *resistant with probability $1 - o_k(1)$* .

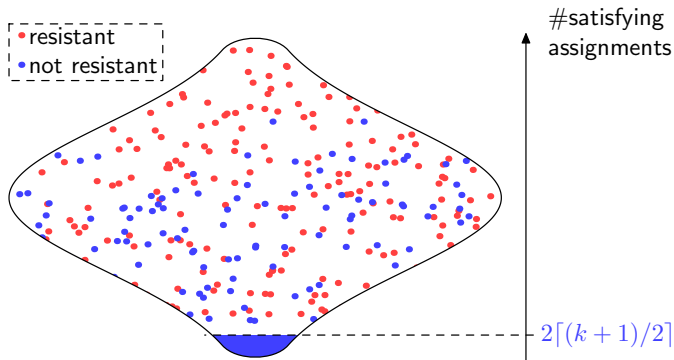
Theorem ([Austrin-H09])

Assuming the UGC, *every* predicate with more than $\frac{32}{33}2^k$ accepting assignments is resistant.

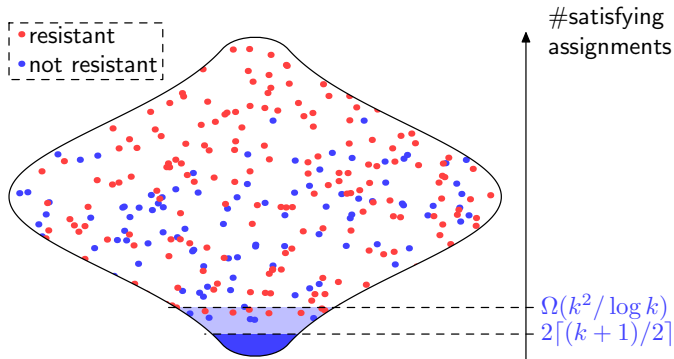
In Pictures with UGC



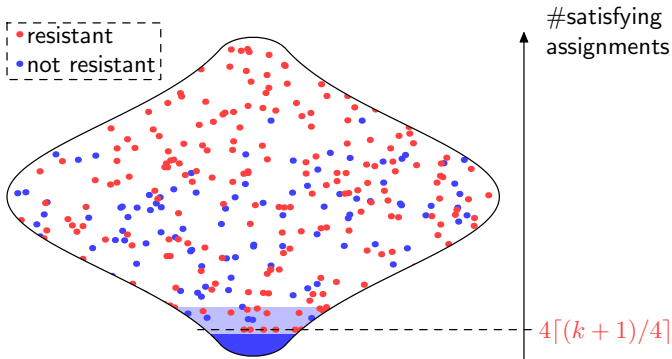
In Pictures with UGC



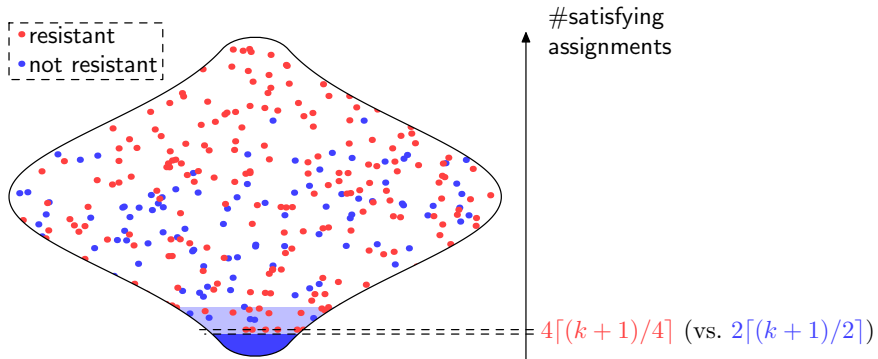
In Pictures with UGC



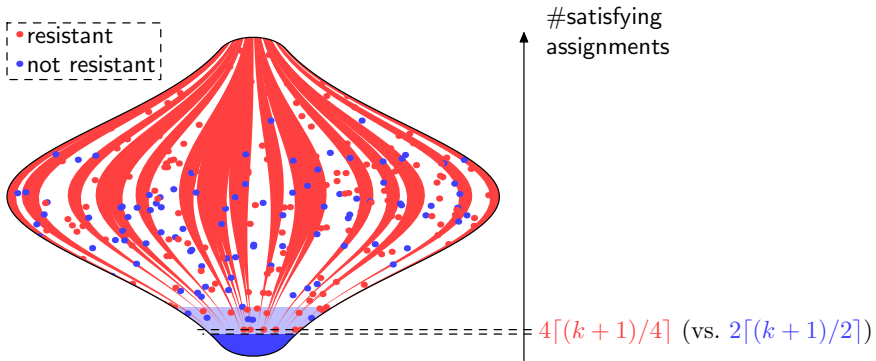
In Pictures with UGC



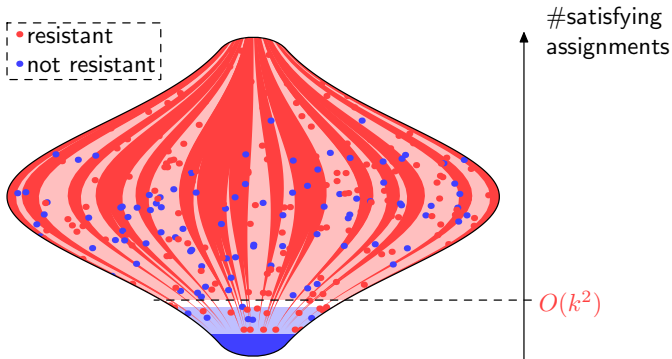
In Pictures with UGC



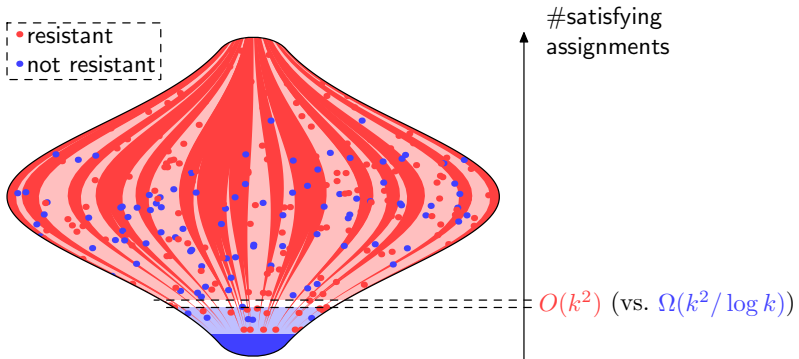
In Pictures with UGC



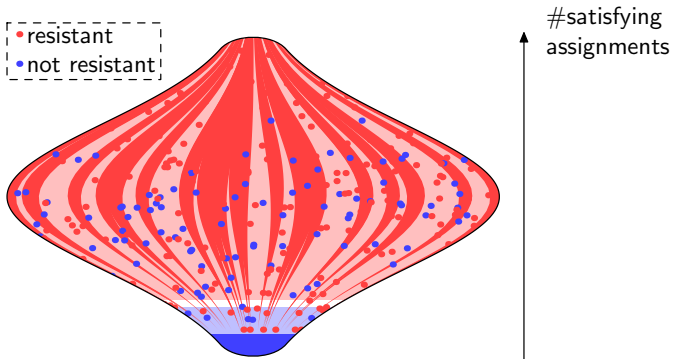
In Pictures with UGC



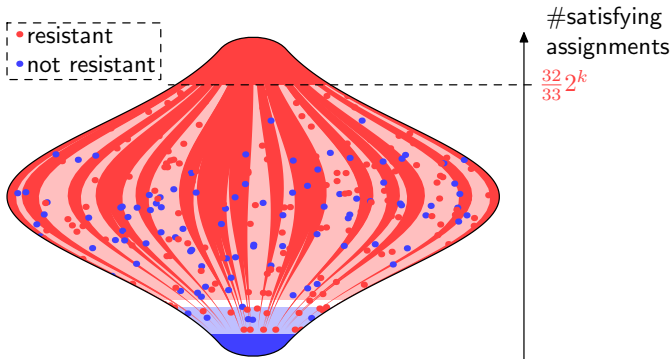
In Pictures with UGC



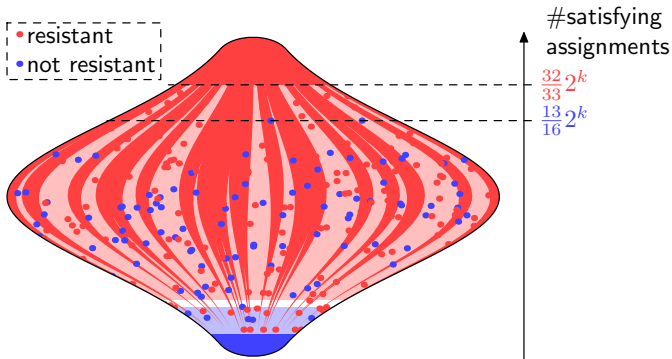
In Pictures with UGC



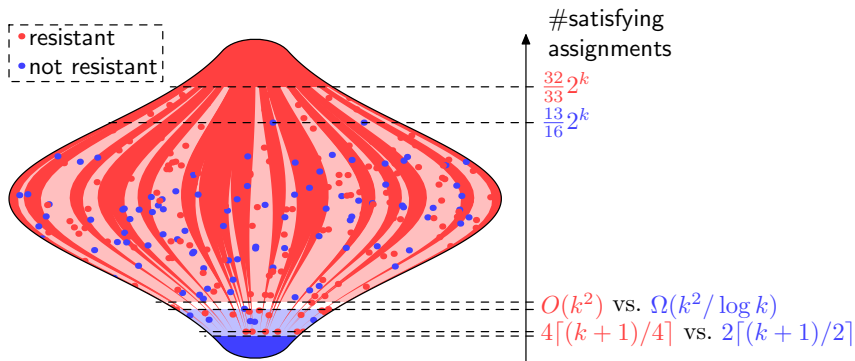
In Pictures with UGC



In Pictures with UGC



In Pictures with UGC



Resistant but not useless

Approximation resistant but not useless.

Consider predicate $GLST : \{-1, 1\}^4 \rightarrow \{0, 1\}$

$$GLST(x_1, x_2, x_3, x_4) = \begin{cases} x_2 \neq x_3 & \text{if } x_1 = 1 \\ x_2 \neq x_4 & \text{if } x_1 = -1 \end{cases}$$

GLST is resistant [GLST98] but does not contain pairwise independence and hence is not useless.

$$GLST(x_1, x_2, x_3, x_4) = \begin{cases} x_2 \neq x_3 & \text{if } x_1 = 1 \\ x_2 \neq x_4 & \text{if } x_1 = -1 \end{cases}$$

Let μ be uniform distribution over

$$\{x : x_1 x_2 x_3 = -1 \text{ and } x_4 = -x_3\}$$

$$GLST(x_1, x_2, x_3, x_4) = \begin{cases} x_2 \neq x_3 & \text{if } x_1 = 1 \\ x_2 \neq x_4 & \text{if } x_1 = -1 \end{cases}$$

Let μ be uniform distribution over

$$\{x : x_1 x_2 x_3 = -1 \text{ and } x_4 = -x_3\}$$

- Balanced pairwise independent except x_3 and x_4 *correlated*

$$\begin{aligned}
 GLST(x_1, x_2, x_3, x_4) &= \begin{cases} x_2 \neq x_3 & \text{if } x_1 = 1 \\ x_2 \neq x_4 & \text{if } x_1 = -1 \end{cases} \\
 &= \frac{1}{2} - \frac{x_2 x_3}{4} - \frac{x_2 x_4}{4} - \frac{x_1 x_2 x_3}{4} + \frac{x_1 x_2 x_4}{4}
 \end{aligned}$$

Let μ be uniform distribution over

$$\{x : x_1 x_2 x_3 = -1 \text{ and } x_4 = -x_3\}$$

- Balanced pairwise independent except x_3 and x_4 correlated
- But x_3 and x_4 never appear together in expansion of GLST

Raghavendra [\[R08\]](#)[\[RS09\]](#) tells that if Max-P is not approximation resistance then a Semi-Definite Programming algorithm gives a non-trivial approximation ratio.

Raghavendra [R08][RS09] tells that if Max-P is not approximation resistance then a Semi-Definite Programming algorithm gives a non-trivial approximation ratio.

We know what kind of algorithm to look for.

Raghavendra [R08][RS09] tells that if Max-P is not approximation resistance then a Semi-Definite Programming algorithm gives a non-trivial approximation ratio.

We know what kind of algorithm to look for.

Have not been used to classify any explicit predicate.

Raghavendra [R08][RS09] tells that if Max-P is not approximation resistance then a Semi-Definite Programming algorithm gives a non-trivial approximation ratio.

We know what kind of algorithm to look for.

Have not been used to classify any explicit predicate.

Possibly undecidable.

More hardness results

Khot, Tulsiani and Worah, [KTW14] show that existence of a certain measure on possible pairwise correlations is **equivalent** to UG-hardness.

Khot, Tulsiani and Worah, [KTW14] show that existence of a certain measure on possible pairwise correlations is **equivalent** to UG-hardness.

We know what property to focus on.

Khot, Tulsiani and Worah, [KTW14] show that existence of a certain measure on possible pairwise correlations is **equivalent** to UG-hardness.

We know what property to focus on.

Not been used to get any new explicit hardness result.

Khot, Tulsiani and Worah, [KTW14] show that existence of a certain measure on possible pairwise correlations is **equivalent** to UG-hardness.

We know what property to focus on.

Not been used to get any new explicit hardness result.

Possibly undecidable.

Fact

P does not contain pairwise independence iff it implies a predicate P' of the form $P'(x) = \frac{1 + \text{sign}(Q(x))}{2}$, where Q is a quadratic polynomial without constant term.

Predicates We Can't Characterize

Maybe signs of quadratic forms are always approximable?

Fact

P does not contain pairwise independence iff it implies a predicate P' of the form $P'(x) = \frac{1 + \text{sign}(Q(x))}{2}$, where Q is a quadratic polynomial without constant term.

Predicates We Can't Characterize

Maybe signs of quadratic forms are always approximable?

No: Exists a quadratic form Q on $k = 12$ variables which turns out to be resistant using “generalized [AM09]”.

Predicates We Can't Characterize

Maybe signs of *linear forms* are always approximable?

Special case 1: "Monarchy" – x_1 decides outcome unless all other variables unite against it

- Can't be handled using Hast's Theorem but turns out to be approximable [[Austrin-Benabbas-Magen09](#)]

Predicates We Can't Characterize

Maybe signs of *linear forms* are always approximable?

Special case 1: "Monarchy" – x_1 decides outcome unless all other variables unite against it

- Can't be handled using Hast's Theorem but turns out to be approximable [[Austrin-Benabbas-Magen09](#)]

Special case 2: "Republic" – x_1 decides outcome unless 3/4 of the other variables unite against it

Open Problem

Is "Republic" approximable?

Classification is doing fairly well.

- Conditional on UGC we know SDPs are universal.
- Simple unknown predicates such as “Republic”.

- Is there a “nice” complete characterization?
- Can we get NP-hardness?
- Can we get more results for satisfiable instances?
- Should we hope/fear for a new complexity class (UGC)?

Thank you!

