Approximation Resistance from Pairwise Independent Subgroups

Siu On Chan
UC Berkeley

Max-CSP

Goal: Satisfy the maximum fraction of constraints

Examples:

1. Max-3XOR:

$$x_1 + x_{10} + x_{27} = 1$$
$$x_4 + x_5 + x_{16} = 0$$
$$\vdots$$

2. Max-3SAT:

$$x_2 \vee \overline{x_9} \vee x_{31}$$
$$x_8 \vee x_{15} \vee \overline{x_{17}}$$
$$\vdots$$

Max-CSP

Goal: Satisfy the maximum fraction of constraints

Examples:

1. Max-3XOR: $(\frac{1}{2} + \varepsilon)$-hardness [Håstad 01]

$$x_1 + x_{10} + x_{27} = 1$$
$$x_4 + x_5 + x_{16} = 0$$
$$\vdots$$

2. Max-3SAT: $(\frac{7}{8} + \varepsilon)$-hardness [Håstad 01]

$$x_2 \vee \overline{x_9} \vee x_{31}$$
$$x_8 \vee x_{15} \vee \overline{x_{17}}$$
$$\vdots$$

Definition (Approximation resistance)

NP-hard to beat a random assignment even when almost satisfiable

That is, NP-hard to decide if an instance of MAX-CSP has value

$$\geqslant 1 - \varepsilon \quad \text{or} \quad \leqslant \text{“random assignment value”} + \varepsilon$$

Examples: MAX-3XOR, MAX-3SAT

Question

Which CSPs are approximation resistant? Why?

Partial answer

If given by a predicate *C* that is a “pairwise independent subgroup” [Chan13]

Max-CSP(C)

MAX-CSP(C) or MAX-C:

Each clause

- involves the same number, k, of literals
- accepts the same collection $C \subseteq \mathbb{Z}_2^k$ of local assignments

Examples ($k = 3$):

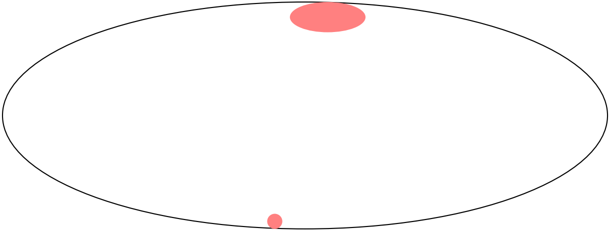
1. $C = \begin{Bmatrix} 000 & 001 & 011 & 010 \\ 100 & 101 & 111 & 110 \end{Bmatrix} \Rightarrow$ MAX-C = MAX-3XOR
2. $C = \begin{Bmatrix} 000 & 001 & 011 & 010 \\ 100 & 101 & 111 & 110 \end{Bmatrix} \Rightarrow$ MAX-C = MAX-3SAT

Random assignment value $= |C|/2^k$

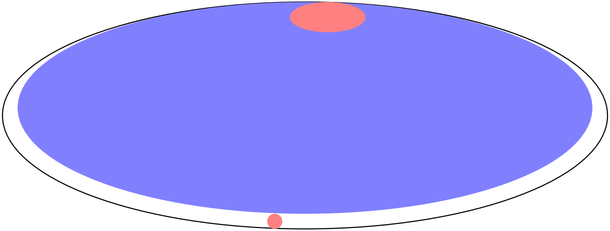
Previous work

Arity	Approximation resistant MAX-CSP(C)
2	none [Goemans–Williamson95, Håstad05]
3	contains all strings of the same parity [Håstad01, Zwick98]
4	many examples [Guruswami–Lewin–Sudan–Trevisan98, Hast05]
$\geqslant 5$	scattered results [Håstad01, Samorodnitsky–Trevisan00] [Engebretsen–Holmerin08, Hast05, Håstad11]

Arity $=$ $\#$variables per constraint



Criteria for approximation resistance (red region):

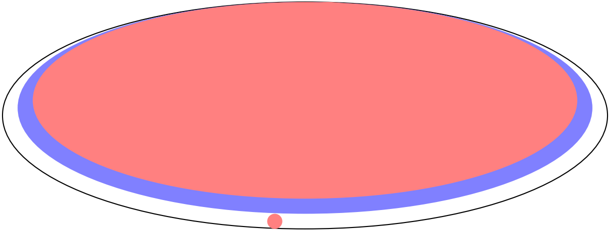


Criteria for approximation resistance (red region):

- [Austrin–Mossel09]: contains pairwise independent subset, assuming Unique-Games Conjecture
  - C is pairwise independent if $\forall i \neq j \in [k], \forall a, b \in \mathbb{Z}_2$,

$$\Pr_{\boldsymbol{c} \in C}[\boldsymbol{c}_i = a, \boldsymbol{c}_j = b] = 1/|\mathbb{Z}_2|^2$$

Example: $C = \{k\text{-bit strings of even parity}\} = k\text{XOR}$



Criteria for approximation resistance (red region):

- [Austrin–Mossel09]: contains pairwise independent subset, assuming Unique-Games Conjecture
  - C is pairwise independent if $\forall i \neq j \in [k], \forall a, b \in \mathbb{Z}_2$,

$$\Pr_{\boldsymbol{c} \in C}[\boldsymbol{c}_i = a, \boldsymbol{c}_j = b] = 1/|\mathbb{Z}_2|^2$$

    Example: $C = \{k\text{-bit strings of even parity}\} = k\text{XOR}$

- [Chan13]: contains pairwise independent subgroup
  - Almost all MAX-CSP(C) [Håstad09]

Corollaries

- Optimal $\Theta(k/2^k)$-hardness for MAX-kCSP, using predicate in [Samorodnitsky–Trevisan09]
- Optimal query-efficient Probabilistically Checkable Proof (PCP) for NP
- Optimal $\Theta(qk/q^k)$-hardness for non-boolean MAX-kCSP when $k \geqslant$ domain size q, using predicate of [Håstad12]

Corollaries

- Optimal $\Theta(k/2^k)$-hardness for Max-kCSP, using predicate in [Samorodnitsky–Trevisan09]
- Optimal query-efficient Probabilistically Checkable Proof (PCP) for NP
- Optimal $\Theta(qk/q^k)$-hardness for non-boolean Max-kCSP when $k \geqslant$ domain size q, using predicate of [Håstad12]
- Improved hardness of Almost-Coloring, Independent-Set on bounded degree graphs, 2Prover-1Round-Game
  - network connectivity problems [Laekhanukit12]
- Follow-up works: [Khot–Tulsiani–Worah12, Huang13a, Huang13b]

Motivated by integrality gaps in sum-of-square programs (the strongest known semidefinite programs) [Schoenebeck08, Tulsiani09, Chan13]

Proof sketch

Theorem

If $C \subseteq \mathbb{Z}_2^k$ is a subgroup that is pairwise independent, then Max-CSP(C) *is approximation resistant*

Definition

C is pairwise independent if $\forall i \neq j \in [k], \forall a, b \in \mathbb{Z}_2$,

$$\Pr_{\boldsymbol{c} \in C}[\boldsymbol{c}_i = a, \boldsymbol{c}_j = b] = 1/|\mathbb{Z}_2|^2$$

Proof overview

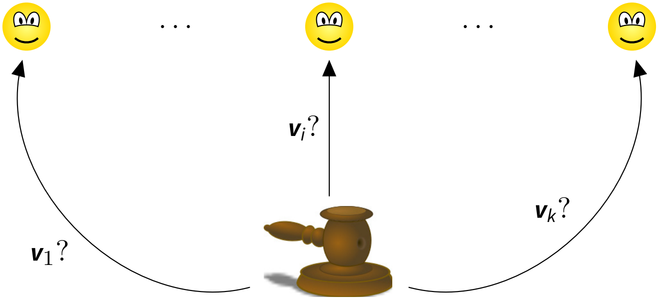
	Label-Cover $\overset{\text{composition}}{\longmapsto}$	Max-C
Yes:	1	≈ 1
No:	$o(1)$	$\approx \lvert C\rvert/2^k$

Proof overview

	Label-Cover	$\stackrel{\text{composition}}{\longmapsto}$	Max-C	$\stackrel{\text{XOR}}{\longmapsto}$	Max-C
Yes:	1		≈ 1		≈ 1
No:	$o(1)$		$\approx \lvert C \rvert / 2^k$		$\approx \lvert C \rvert / 2^k$

Label-Cover $\longmapsto$ Max-C $\equiv$ Game

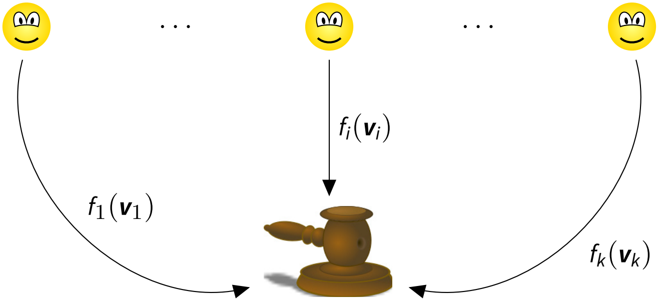
k players try to convince a judge that a Max-C instance M is satisfiable



1. Judge picks random clause $(\vec{\boldsymbol{v}}, \vec{\boldsymbol{b}}) = ((\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k), (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k))$ from Max-C instance M ($\vec{\boldsymbol{b}} \in \mathbb{Z}_2^k$ specifies positive/negative literals)
2. Gets assignments $f_i(\boldsymbol{v}_i) \in \mathbb{Z}_2$ from k players

Label-Cover ⟼ MAX-C ≡ Game

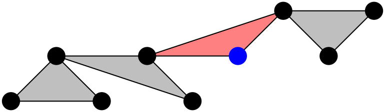
k players try to convince a judge that a MAX-C instance M is satisfiable



1. Judge picks random clause $(\vec{\boldsymbol{v}}, \vec{\boldsymbol{b}}) = ((\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k), (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k))$ from MAX-C instance M ($\vec{\boldsymbol{b}} \in \mathbb{Z}_2^k$ specifies positive/negative literals)
2. Gets assignments $f_i(\boldsymbol{v}_i) \in \mathbb{Z}_2$ from k players
3. Accepts $\Leftrightarrow \vec{f}(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{b}} \in C$

Label-Cover $\longmapsto$ MAX-C

Two parties try to convince a judge that a CSP instance *L* is satisfiable

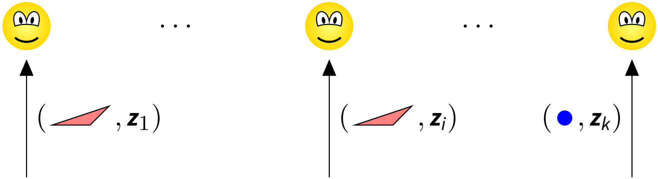


1. Judge picks clause and variable from at random
2. Asks for assignment to from one party and assignment to from the other
3. Accepts if the assignments agree at

Winning probability 1 or ≈ 0? NP-hard to tell! (PCP Theorem and Parallel Repetition Theorem)

Label-Cover $\longmapsto$ Max-C (Composition)

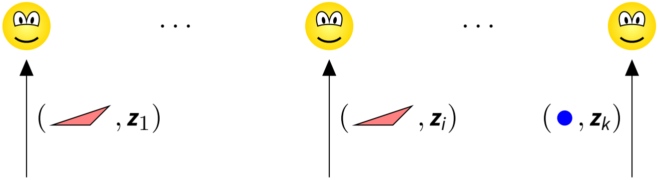
k players try to convince a judge that a CSP instance L has a satisfying assignment A



1. Judge picks and from L as in Label-Cover
2. Asks $(\; , \boldsymbol{z}_i)$ or $(\; , \boldsymbol{z}_i)$ from each player
   $\boldsymbol{z}_i$: subset of satisfying assignments to clause or variable
3. Get boolean replies $\boldsymbol{y}_i$ from k players
4. Accept $\Leftrightarrow (\boldsymbol{y}_1 - \boldsymbol{b}_1, \ldots, \boldsymbol{y}_k - \boldsymbol{b}_k) \in C$

Label-Cover $\longmapsto$ Max-C (Composition)

k players try to convince a judge that a CSP instance L has a satisfying assignment A

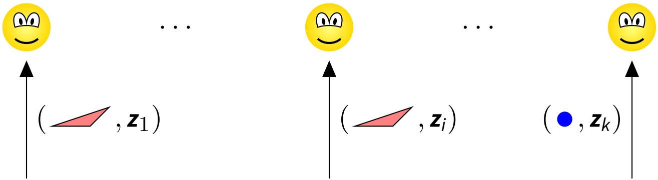


1. Judge picks and from L as in Label-Cover
2. Asks (, $\boldsymbol{z}_i$) or (, $\boldsymbol{z}_i$) from each player
   $\boldsymbol{z}_i$: subset of satisfying assignments to clause or variable
3. Get boolean replies $\boldsymbol{y}_i$ from k players
4. Accept $\Leftrightarrow$ $(\boldsymbol{y}_1 - \boldsymbol{b}_1, \ldots, \boldsymbol{y}_k - \boldsymbol{b}_k) \in C$

$\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ are correlated, as specified by “dictator test”

Composition barrier

	Label-Cover	$\overset{\text{composition}}{\longmapsto}$	Max-C
Yes:	1		≈ 1
No:	$o(1)$		$\approx \lvert C \rvert / 2^k$



- Some players share , others share $\bullet \Longrightarrow$ replies not random
  [Bellare–Goldreich–Sudan98, Sudan–Trevisan98]

XOR

	Label-Cover	$\overset{\text{composition}}{\longmapsto}$	Max-C	$\overset{\text{XOR}}{\longmapsto}$	Max-C
No:	$o(1)$		$\leqslant 0.9$		$\approx \lvert C \rvert / 2^k$

XOR of games:

- Parallel repetition without blowing up alphabet size
- Each player should respond with the XOR of replies to individual games

Game $M \oplus M'$:

1. Judge picks random clauses $(\vec{\boldsymbol{v}}, \vec{\boldsymbol{b}})$ from M and $(\vec{\boldsymbol{v}}', \vec{\boldsymbol{b}}')$ from M'
2. Gets boolean assignments $f_i(\boldsymbol{v}_i, \boldsymbol{v}'_i)$ from k players
3. Accepts $\Leftrightarrow \vec{f}(\vec{\boldsymbol{v}}, \vec{\boldsymbol{v}}') - \vec{\boldsymbol{b}} - \vec{\boldsymbol{b}}' \in C$

Preseves almost-satisfiability when C is a subgroup

XOR-lemma?

Wishful thinking (XOR-lemma)

$\mathsf{val}(M) \leqslant 0.9 \quad \Rightarrow \quad \mathsf{val}(M \oplus \ldots \oplus M) \to |C|/2^k$

Counterexample: Mermin's game [Briët–Buhrman–Lee–Vidick13]

Observation

Correlation can only decrease upon taking XOR

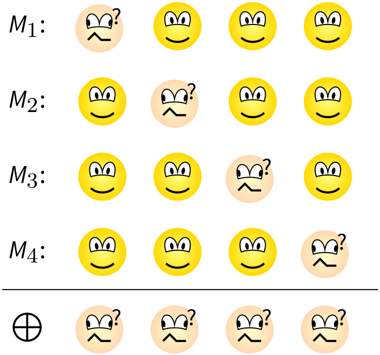
	Label-Cover	$\xmapsto{\text{composition}}$	Max-C	$\xmapsto{\text{XOR}}$	Max-C
No:	$o(1)$		player j uncorrelated		random

M_1: 

Observation

Correlation can only decrease upon taking XOR

	LABEL-COVER	$\overset{\text{composition}}{\longmapsto}$	MAX-C	$\overset{\text{XOR}}{\longmapsto}$	MAX-C
No:	$o(1)$		player j uncorrelated		random



$$\vec{f}(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{b}} \triangleq (f_1(\boldsymbol{v}_1) - \boldsymbol{b}_1, \ldots, f_k(\boldsymbol{v}_k) - \boldsymbol{b}_k) \in \mathbb{Z}_2^k$$

$$\|M\|_\chi \triangleq \max_{\vec{f}:\vec{V}\to\mathbb{Z}_2^k} \left| \mathop{\mathbb{E}}_{(\vec{\boldsymbol{v}},\vec{\boldsymbol{b}})} \chi(\vec{f}(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{b}}) \right|, \qquad \chi \in \widehat{\mathbb{Z}_2^k}$$

Lemma

$\|M \oplus M'\|_\chi \leqslant \min\{\|M\|_\chi, \|M'\|_\chi\}$

$$\vec{f}(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{b}} \triangleq (f_1(\boldsymbol{v}_1) - \boldsymbol{b}_1, \ldots, f_k(\boldsymbol{v}_k) - \boldsymbol{b}_k) \in \mathbb{Z}_2^k$$

$$\|M\|_\chi \triangleq \max_{\vec{f}:\vec{V}\to\mathbb{Z}_2^k} \left| \mathop{\mathbb{E}}_{(\vec{\boldsymbol{v}},\vec{\boldsymbol{b}})} \chi(\vec{f}(\vec{\boldsymbol{v}}) - \vec{\boldsymbol{b}}) \right|, \qquad \chi \in \widehat{\mathbb{Z}_2^k}$$

Lemma

$\|M \oplus M'\|_\chi \leqslant \min\{\|M\|_\chi, \|M'\|_\chi\}$

$$\left| \mathop{\mathbb{E}}_{(\vec{\boldsymbol{v}},\vec{\boldsymbol{b}})} \mathop{\mathbb{E}}_{(\vec{\boldsymbol{v}}',\vec{\boldsymbol{b}}')} \chi\left(\vec{f}(\vec{\boldsymbol{v}}, \vec{\boldsymbol{v}}') - \vec{\boldsymbol{b}} - \vec{\boldsymbol{b}}'\right) \right|$$

$$\leqslant \mathop{\mathbb{E}}_{(\vec{\boldsymbol{v}},\vec{\boldsymbol{b}})} \left| \mathop{\mathbb{E}}_{(\vec{\boldsymbol{v}}',\vec{\boldsymbol{b}}')} \chi\Big(\underbrace{\vec{f}(\vec{\boldsymbol{v}}, \vec{\boldsymbol{v}}') - \vec{\boldsymbol{b}}}_{\vec{g}(\vec{\boldsymbol{v}}')} - \vec{\boldsymbol{b}}'\Big) \right| \qquad \square$$

Label-Cover	$\overset{\text{composition}}{\longmapsto}$	Max-C	$\overset{\text{XOR}}{\longmapsto}$	Max-C
$o(1)$		$\|\cdot\|_\chi = o(1)$ $\forall \chi : \chi_j \neq \mathbf{1}$		$\|C\|/2^k + o(1)$

Uses pairwise independence and invariance principle

[Mossel–O'Donnell–Oleszkiewicz10, Mossel10, O'Donnell–Wright12]

Conclusion

- New gap-amplification technique: XOR/direct sum
- Optimal hardness of MAX-kCSP and optimal query-efficient PCP
- General criteria for approximation resistance

Open problems

1. Optimal hardness of *satisfiable* MAX-kCSP?
   - Progress by [Huang13] in the next talk
2. Derandomizing XOR/direct sum

Thank you

Conclusion

- New gap-amplification technique: XOR/direct sum
- Optimal hardness of Max-*k*CSP and optimal query-efficient PCP
- General criteria for approximation resistance

Open problems

1. Optimal hardness of *satisfiable* Max-*k*CSP?
   - Progress by [Huang13] in the next talk
2. Derandomizing XOR/direct sum

Thank you 

Emoticons modified from

http://www.texample.net/tikz/examples/emoticons/

Gavel from

http://openclipart.org/detail/69745/judge-hammer-by-bocian