NP-hardness of Minimum Circuit Size Problem for OR-AND-MOD Circuits

Shuichi Hirahara (The University of Tokyo)

Igor C. Oliveira (University of Oxford)

Rahul Santhanam (University of Oxford)

CCC 2018 @ San Diego

June 22, 2018

Talk Outline

1. MCSP and Its background
2. $\mathcal{C}$-MCSP for a circuit class $\mathcal{C}$
3. Our Results
4. Proof Sketch

Talk Outline

1. MCSP and Its background

2. $\mathcal{C}$-MCSP for a circuit class $\mathcal{C}$

3. Our Results

4. Proof Sketch

Minimum Circuit Size Problem (MCSP)

Input

- Truth table $T \in \{0,1\}^{2^t}$ of a function $f: \{0,1\}^t \to \{0,1\}$
- Size parameter $s \in \mathbb{N}$

Output

Is there a circuit of size $\leq s$ that computes f?

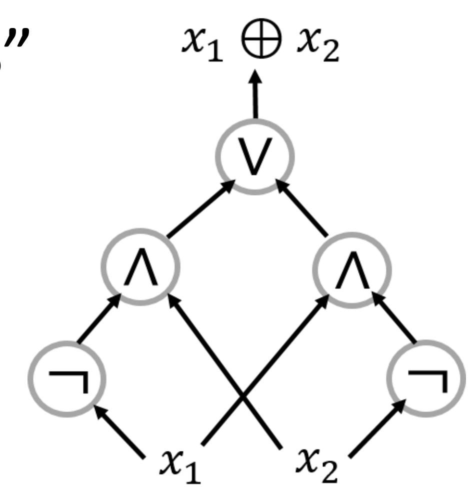
Example:

$s = 5$

$f =$

x_1	x_2	$x_1 \oplus x_2$
0	0	0
0	1	1
1	0	1
1	1	0

Output: "YES"



Brief History of MCSP

1950s Recognized as an important problem in the Soviet Union [Trakhtenbrot's survey]

1970s Levin delayed publishing his work because he wanted to say something about MCSP.

1979 Masek proved NP-completeness of DNF-MCSP.

2000 Kabanets and Cai revived interest, based on natural proofs. [Razborov & Rudich (1997)]

Since then many papers and results appeared; however, the complexity of MCSP remains elusive.

Current Knowledge about MCSP

- Upper bound: $\mathrm{MCSP} \in \mathbf{NP}$
- Lower bound: $\exists$pseudorandom function generators $\Longrightarrow \mathrm{MCSP} \notin \mathbf{P}$

- **Big Open Question:** Is MCSP NP-hard?
- No consensus about the exact complexity of MCSP
  - ✓No strong evidence *against* NP-completeness
    - Weak evidence: [Hirahara-Santhanam (CCC'17)] [Allender-Hirahara 17]…
  - ✓No strong evidence *for* NP-completeness
    - Some new evidence: [Impagliazzo-Kabanets-Volkvovich (CCC'18)] & This work

Kabanets-Cai Obstacle: Why so difficult?

- Suppose that we want to construct a reduction from SAT to MCSP.

$$\varphi \in \mathrm{SAT} \quad \mapsto \quad (f, s) \quad \mathrm{CircSize}(f) \leq s$$

$$\varphi \notin \mathrm{SAT} \quad \mapsto \quad (f, s) \quad \mathrm{CircSize}(f) > s$$

Need to prove a circuit lower bound!

- Natural reduction techniques would imply

$\mathrm{E} \not\subseteq \mathrm{SIZE}\left(n^{O(1)}\right)$. [Kabanets-Cai (2000)]

Talk Outline

1. MCSP and Its background
2. $\mathcal{C}$-MCSP for a circuit class $\mathcal{C}$
3. Our Results
4. Proof Sketch

$\mathcal{C}$-MCSP for a circuit class $\mathcal{C}$

Input

- Truth table $T \in \{0,1\}^{2^t}$ of a function $f : \{0,1\}^t \to \{0,1\}$
- Size parameter $s \in \mathbb{N}$

Output

Is there a $\mathcal{C}$-circuit of size $\leq s$ that computes f?

Theorem [Masek (1978 or 79, unpublished)]

$\mathrm{DNF}-\mathrm{MCSP}$ is NP-hard.

DNF-MCSP

Input

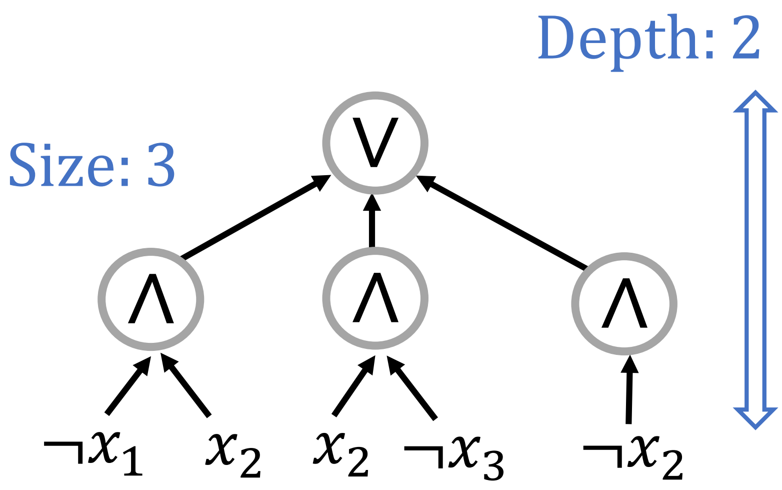
- Truth table $T \in \{0,1\}^{2^t}$ of a function $f: \{0,1\}^t \to \{0,1\}$
- Size parameter $s \in \mathbb{N}$

Output

Is there a DNF formula of size $\leq s$ that computes f?

<u>Example of DNFs:</u>

$$(\neg x_1 \wedge x_2) \vee (x_2 \wedge \neg x_3) \vee (\neg x_2) \equiv$$



(The size of DNF) $:=$ #(clauses)

$\mathcal{C}$-MCSP for $\mathcal{C} \supset \mathrm{DNF}$

- Beyond DNFs, no NP-hardness was proved since the work of Masek (1979).

- To quote Allender, Hellerstein, McCabe, Pitassi, and Saks (2008):

  "Thus an **important open question** is to resolve the NP-hardness of … function minimization results above for classes that are stronger than DNF."

Known results about $\mathcal{C}$-MCSP

More expressive

MCSP

⋮

AC^0_d-MCSP for large d

Hardness based on cryptography

e.g.) Blum integer factorization [AHMPS08]

⋮

Depth3-MCSP

No hardness result

DNF-MCSP

Known to be NP-hard [Masek (1979)]

Remark: The complexity is not necessarily monotone increasing or decreasing.

Talk Outline

1. MCSP and Its background
2. $\mathcal{C}$-MCSP for a circuit class $\mathcal{C}$
3. Our Results
4. Proof Sketch

Our Results

- The first NP-hardness result for $\mathcal{C}$-MCSP for a class $\mathcal{C} \supset \mathrm{DNF}$

Theorem (Main Result)

$(\mathrm{DNF} \circ \mathrm{XOR})-\mathrm{MCSP}$ is NP-hard
under polynomial-time many-one reductions.

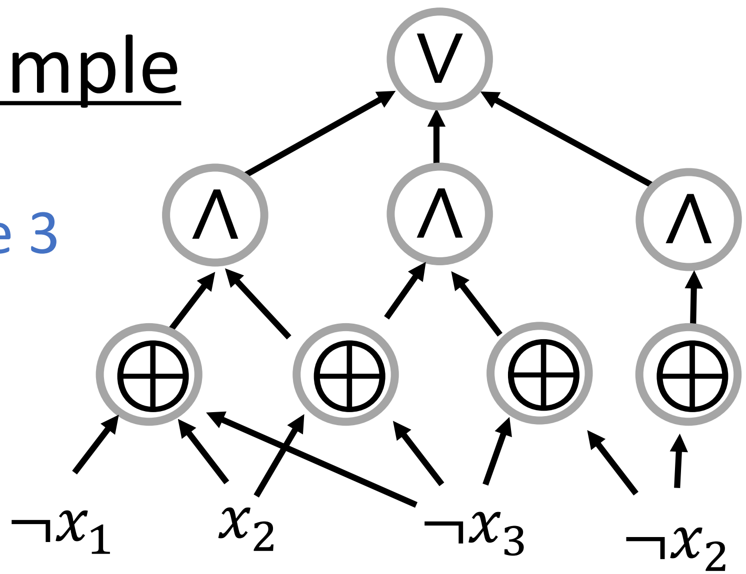
- Our proof techniques extend to:
  - $(\mathrm{DNF} \circ \mathrm{MOD}_m)\text{-}\mathrm{MCSP}'$ is NP-hard for any $m \geq 2$, but the input is a truth table of an m-valued function $f: (\mathbb{Z}/m\mathbb{Z})^t \to \{0,1\}$.

DNF ∘ XOR circuits ($2^{\Omega(n)}$ circuit lower bound is known)

[Cohen & Shinkar (2016)]

Example

Size 3



Depth 3

1^{st} layer: an OR gate
2^{nd} layer: AND gates
3^{rd} layer: XOR gates

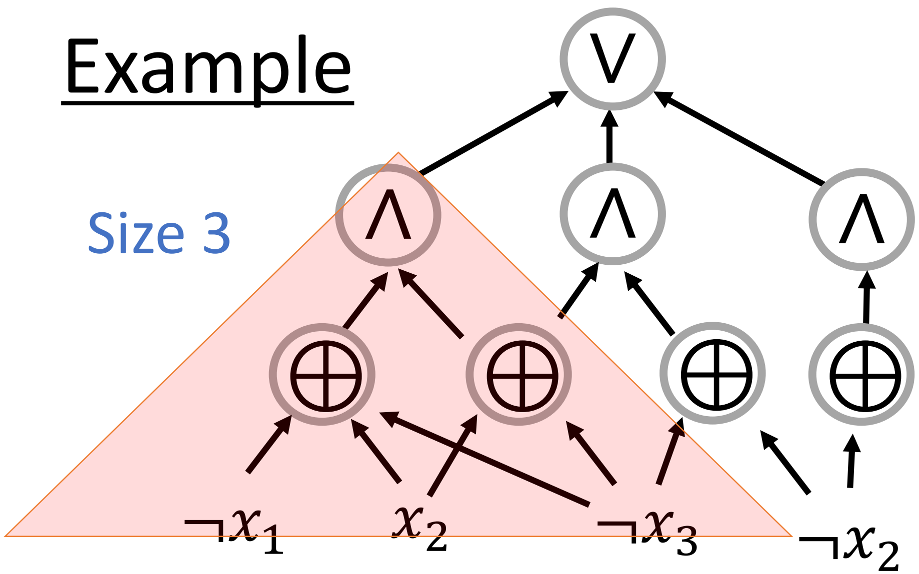
(The size of DNF ∘ XOR circuits) ≔ (The number of AND gates)

- This is a convenient circuit size measure as advocated by Cohen & Shinkar (2016).
  1. Nice combinatorial meaning
  2. W.l.o.g., #(XOR gates) $\leq n \cdot$ #(AND gates)
- Our proof techniques extend to the number of all the gates in a DNF ∘ XOR *formula.*

DNF ∘ XOR circuits ($2^{\Omega(n)}$ circuit lower bound is known)

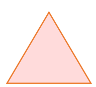
[Cohen & Shinkar (2016)]

<u>Example</u>



Depth 3

1st layer: an OR gate
2nd layer: AND gates
3rd layer: XOR gates

The subcircuit  outputs 1. $\iff$ $\begin{cases} (1 \oplus x_1) \oplus x_2 \oplus (1 \oplus x_3) = 1 \\ x_2 \oplus (1 \oplus x_3) = 1 \end{cases}$

← Some linear equations over $\mathrm{GF}(2)$

$\iff$ $(x_1, x_2, x_3) \in A$

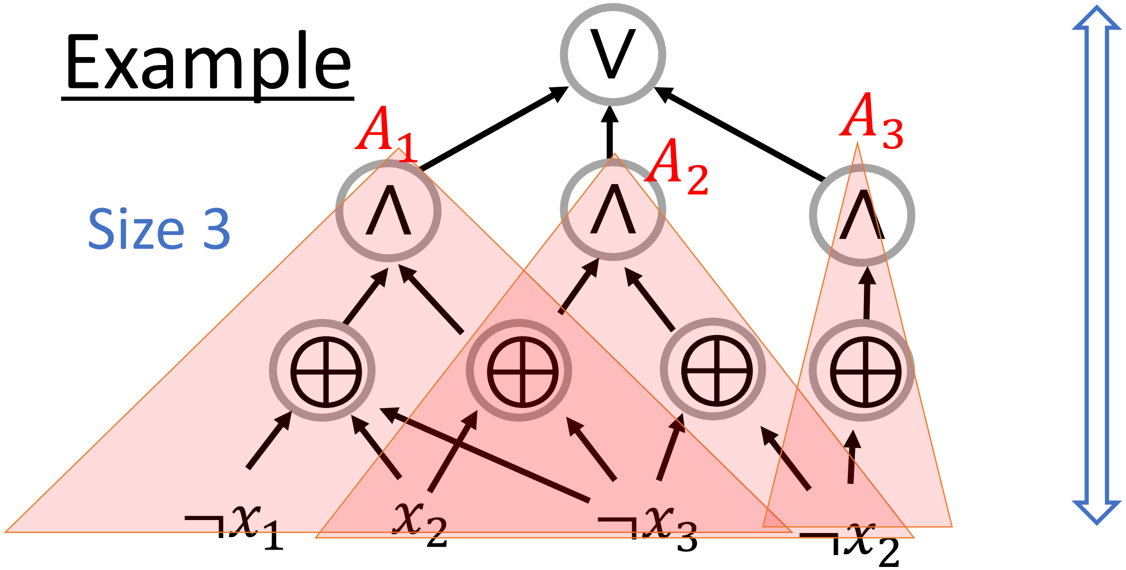
(for some affine subspace $A \subseteq \mathrm{GF}(2)^n$)

DNF ∘ XOR circuits ($2^{\Omega(n)}$ circuit lower bound is known)

[Cohen & Shinkar (2016)]

$f: \{0,1\}^n \to \{0,1\}$

Example



1st layer: an OR gate

2nd layer: AND gates

3rd layer: XOR gates

$$f^{-1}(1) = A_1 \cup A_2 \cup A_3$$

The Important Observation

The minimum $\mathrm{DNF} \circ \mathrm{XOR}$ circuit size for computing f

||

The minimum number m of affine subspaces needed to cover $f^{-1}(1)$: that is,

$\exists A_1, \ldots, A_m$: affine subspaces of $\{0,1\}^n$

$$A_i \subseteq f^{-1}(1) \quad \text{and} \quad A_1 \cup \cdots \cup A_m = f^{-1}(1)$$

Talk Outline

1. MCSP and Its background
2. $\mathcal{C}$-MCSP for a circuit class $\mathcal{C}$
3. Our Results
4. Proof Sketch

- Our proof was inspired by a simple proof of Masek’s result given by [Allender, Hellerstein, McCabe, Pitassi, and Saks (2008)].

- We extend and generalize their ideas significantly.

Proof Outline

Theorem (Main Result)

$$\mathrm{NP} \leq_m^p (\mathrm{DNF} \circ \mathrm{XOR})\text{−}\mathrm{MCSP}$$

<u>Step 1.</u> 2-factor approx. of r-Bounded Set Cover (NP-hard [Trevisan 2001]) $\leq_m^{\mathrm{ZPP}}$ $(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP for *partial* functions

<u>Step 2.</u> $(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP for *partial* functions $\leq_m^{\mathrm{ZPP}}$ $(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP

<u>Step 3.</u> Derandomization using ϵ-biased generators [Naor & Naor (1993)]

Proof Outline

Theorem (Main Result)

$$\mathrm{NP} \leq_m^p (\mathrm{DNF} \circ \mathrm{XOR}){-}\mathrm{MCSP}$$

Step 1.	2-factor approx. of r-Bounded Set Cover (NP-hard [Trevisan 2001])	$\leq_m^{\mathrm{ZPP}}$	$(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP for *partial* functions
Step 2.	$(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP for *partial* functions	$\leq_m^{\mathrm{ZPP}}$	$(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP

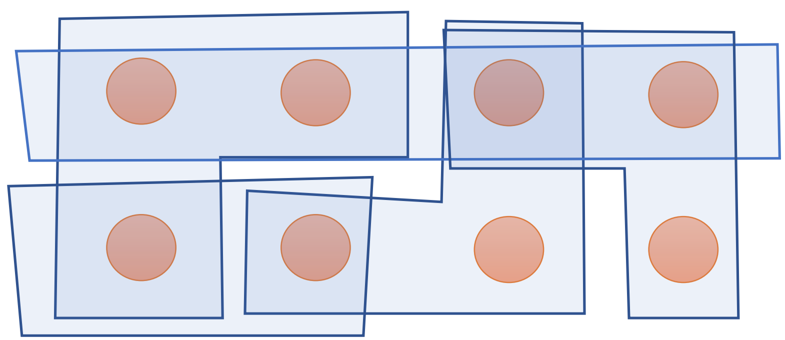
Step 3. Derandomization using ϵ-biased generators [Naor & Naor (1993)]

The Set Cover Problem

<u>Input</u>: A universe U and a collection of sets $\mathcal{S} \subseteq 2^U$

<u>Output</u>: The minimum $|\mathcal{C}|$ such that $\mathcal{C} \subseteq \mathcal{S}$ and $\bigcup_{C \in \mathcal{C}} C = U$

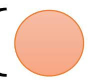
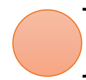
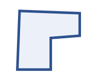
Example: $U = \{ \ldots \}$, $\mathcal{S} = \{ \ldots \}$



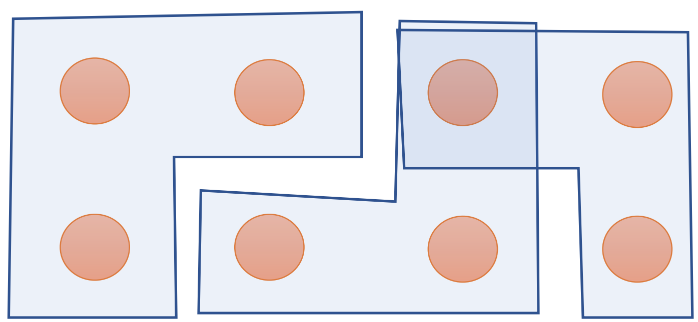
The Set Cover Problem

Input: A universe U and a collection of sets $\mathcal{S} \subseteq 2^U$

Output: The minimum $|\mathcal{C}|$ such that $\mathcal{C} \subseteq \mathcal{S}$ and $\bigcup_{C \in \mathcal{C}} C = U$

Example: $U = \{$  …  $\}$, $\mathcal{S} = \{$  … $\}$

A minimum cover $\mathcal{C}$:

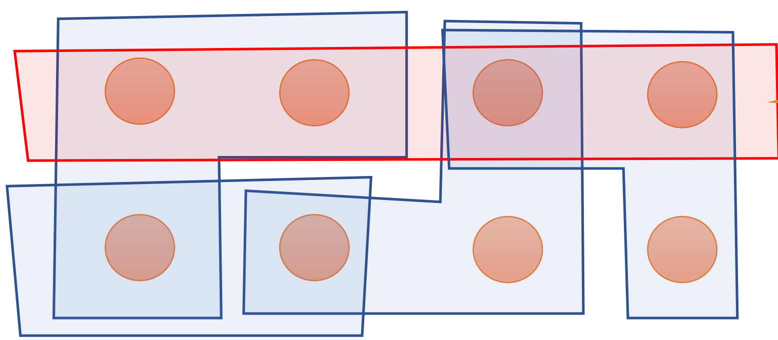


The r-Bounded Set Cover Problem

<u>Input</u>: A universe U and a collection of sets $\mathcal{S} \subseteq 2^U$ such that $|S| \leq r$ for every $S \in \mathcal{S}$.

<u>Output</u>: The minimum $|\mathcal{C}|$ such that $\mathcal{C} \subseteq \mathcal{S}$ and $\bigcup_{C \in \mathcal{C}} C = U$

Example: $U = \{ \ldots \}$, $\mathcal{S} = \{ \ldots \}$



4-bounded,
but not 3-bounded

[Feige (1998)] [Trevisan (2001)]
Approximation of $\left(1 - o(1)\right) \ln r$ is NP-hard.

- We set r to be large enough so that a 2-factor approx. is NP-hard.

Proof Outline

Theorem (Main Result)

$$\mathrm{NP} \le_m^p (\mathrm{DNF} \circ \mathrm{XOR})\text{−}\mathrm{MCSP}$$

<u>Step 1.</u> 2-factor approx. of r-Bounded Set Cover (NP-hard [Trevisan 2001]) $\le_m^{\mathrm{ZPP}}$ (DNF ∘ XOR)-MCSP for *partial* functions

<u>Step 2.</u> (DNF ∘ XOR)-MCSP for *partial* functions $\le_m^{\mathrm{ZPP}}$ (DNF ∘ XOR)-MCSP

<u>Step 3.</u> Derandomization using ϵ-biased generators [Naor & Naor (1993)]

$(\mathrm{DNF} \circ \mathrm{XOR})$-$\mathrm{MCSP}^*$ for partial functions

Input

- Truth table of a partial function
  $f: \{0,1\}^t \to \{0,1,*\}$
- Size parameter $s \in \mathbb{N}$

Output

Is there a circuit of size $\leq s$ that agrees with f on inputs from $f^{-1}(\{0,1\})$?

Example:

$f =$

x_1	x_2	$f(x_1, x_2)$
0	0	*
0	1	1
1	0	*
1	1	0

Claim

2-factor approx. of r-Bounded Set Cover $\leq_m^{\mathrm{ZPP}}$ $(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP for partial functions

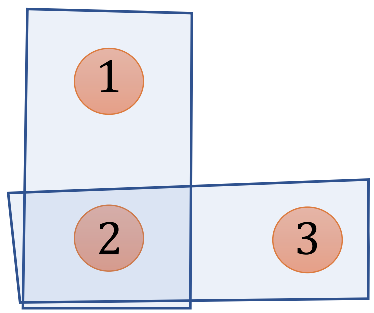
- Given: $U = \{1, \ldots, N\}, \mathcal{S} = \{S_1, \ldots, S_m\}$
- Goal: Construct $f: \{0,1\}^t \to \{0,1,*\}$ for $t = O(\log N)$

Set Cover		$(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP	
$i \in U$	$\mapsto$	$v^i \sim \{0,1\}^t$	(A uniformly random vector)
$S_j \in \mathcal{S}$	$\mapsto$	$\mathrm{span}_{i \in S_j}(v^i) \subseteq \{0,1\}^t$	
Cover $\mathcal{C} \subseteq \mathcal{S}$	$\mapsto$	$\bigcup_{S \in \mathcal{C}} \mathrm{span}_{i \in S}(v^i) \subseteq \{0,1\}^t$	

Set Cover

$U = \{1,2,3\}, \mathcal{S} = \{S_1, S_2\}$

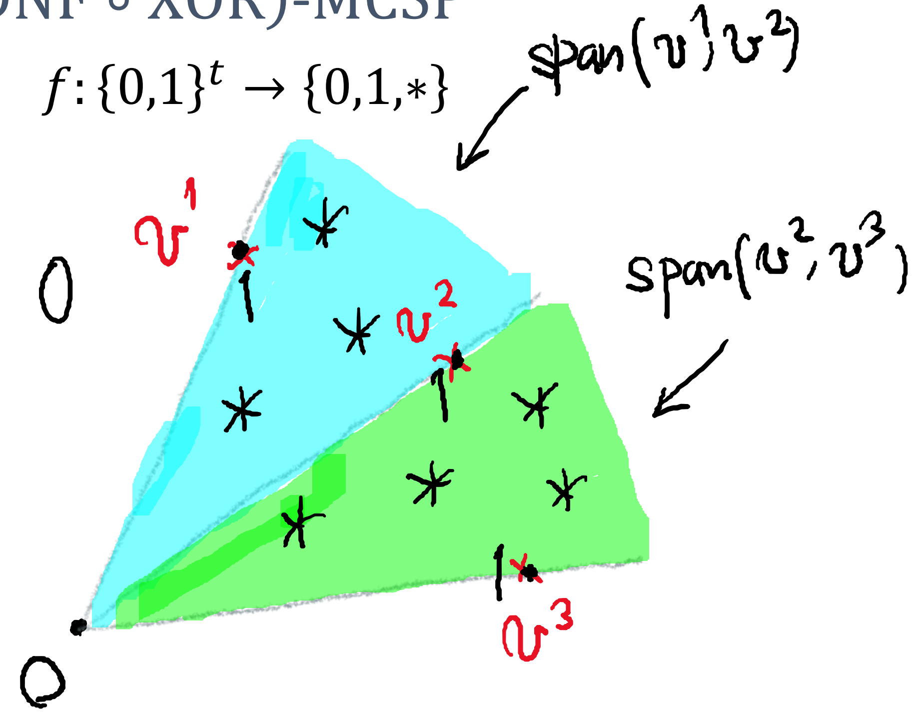
$S_1 = \{1,2\}$



$S_2 = \{2,3\}$

(DNF ∘ XOR)-MCSP

$f: \{0,1\}^t \to \{0,1,*\}$



- $f(v^i) := 1$ for any $i \in U$.
- $f(x) := 0$ for all $x \notin \mathrm{span}(v^1, v^2) \cup \mathrm{span}(v^2, v^3)$.
- $f(y) := *$ for any other vector $y \in \{0,1\}^t$.

- The minimum DNF ∘ XOR circuit size for computing f

  $=$ The minimum number of affine subspaces $A \subseteq f^{-1}(\{1,*\})$ needed to cover $f^{-1}(1) = \{v^1, v^2, v^3\}$.

Intuition: When A is Linear

Random linear subspaces of small dimension r

$$A \subseteq f^{-1}(\{1,*\}) = \mathrm{span}\left(v^1, v^2\right) \cup \mathrm{span}(v^2, v^3)$$

$$\Longrightarrow \quad A \subseteq \mathrm{span}\left(v^1, v^2\right) \text{ or } A \subseteq \mathrm{span}(v^2, v^3)$$

with high probability

(if A is a linear subspace)

$\Longrightarrow$ The set of points $\{\, i \in \{1,2,3\} \mid v^i \in A \,\}$ covered by A is contained in some legal set S_1 or $S_2 \in \mathcal{S}$.

$\Longrightarrow$ The minimum number of linear subspaces needed to cover $\{v^1, v^2, v^3\}$

$=$ The minimum set cover size

Intuition: When A is Affine

$$A \subseteq f^{-1}(\{1,*\}) = \mathrm{span}(v^1, v^2) \cup \mathrm{span}(v^2, v^3)$$

$$\overset{?}{\Longrightarrow} \quad A \subseteq \mathrm{span}(v^1, v^2) \text{ or } A \subseteq \mathrm{span}(v^2, v^3)$$

with high probability

(if A is an affine subspace)

Counterexample: $A \coloneqq \{v^1, v^3\} = v^1 \oplus \{0, v^1 \oplus v^3\}$

- Still, we can prove that:

  The set of points $\{\, i \in \{1,2,3\} \mid v^i \in A \,\}$ covered by A is contained in $S_a \cup S_b$ for some two legal sets $S_a, S_b \in \mathcal{S}$

$\Longrightarrow$ The minimum number of affine subspaces needed to cover $\{v^1, v^2, v^3\}$ is a 2-factor approximation of the minimum set cover size.

Fomally:

$$f:\{0,1\}^t \to \{0,1,*\}$$

$$f(x) = \begin{cases} 1 & (x = v^i \text{ for some } i) \\ 0 & (x \notin \bigcup_{S \in \mathcal{S}} \mathrm{span}_{i \in S}(v^i)) \\ * & (\text{otherwise}) \end{cases}$$

Claim (Easy part)

(The minimum DNF $\circ$ XOR circuit size) $\leq$ (The minimum set cover size)

- By a delicate probabilistic argument, it can be shown:

Claim (Hard part)

For $t \geq O(r \log N)$, the following holds with high probability:
(The minimum set cover size) $\leq 2 \times$ (The minimum DNF $\circ$ XOR circuit size)

Summary of Step 1

1. Input: $U = \{1, \dots, N\}$, $\mathcal{S} = \{S_1, \dots, S_m\}$
2. Let $t := \Theta(\log N)$.
3. Pick $v^i \sim \{0,1\}^t$ randomly for each $i \in U$.
4. Verify that $\left(v^i\right)_{i \in U}$ satisfies a certain condition.
5. Define $f: \{0,1\}^t \to \{0,1,*\}$ as follows and output its truth table.

$$f(x) = \begin{cases} 1 & (x = v^i \text{ for some } i) \\ 0 & (x \notin \bigcup_{S \in \mathcal{S}} \mathrm{span}_{i \in S}(v^i)) \\ * & (\text{otherwise}) \end{cases}$$

Proof Outline

Theorem (Main Result)

$$\mathrm{NP} \le_m^p (\mathrm{DNF} \circ \mathrm{XOR})\text{−}\mathrm{MCSP}$$

Step 1. 2-factor approx. of r-Bounded Set Cover (NP-hard [Trevisan 2001]) $\le_m^{\mathrm{ZPP}}$ (DNF ∘ XOR)-MCSP for *partial* functions

Step 2. (DNF ∘ XOR)-MCSP for *partial* functions $\le_m^{\mathrm{ZPP}}$ (DNF ∘ XOR)-MCSP

Step 3. Derandomization using ϵ-biased generators [Naor & Naor (1993)]

Step 2: Making it a total function

Claim

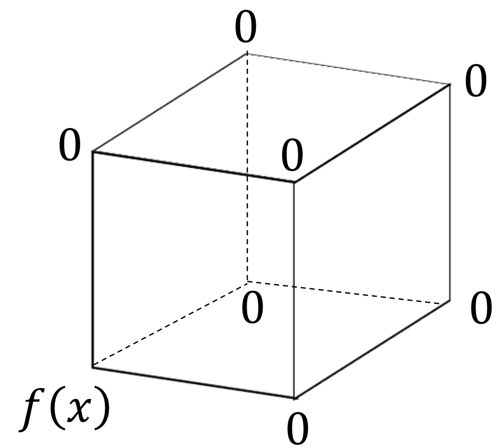
$(\mathrm{DNF} \circ \mathrm{XOR})\text{-MCSP}$ for partial functions $\leq_m^{\mathrm{ZPP}}$ $(\mathrm{DNF} \circ \mathrm{XOR})\text{-MCSP}$

- Given: a partial function $f: \{0,1\}^t \to \{0,1,*\}$
- Output: a total function $g: \{0,1\}^{t+s} \to \{0,1\}$
- For each $x \in \{0,1\}^t$, we encode each value $f(x) \in \{0,1,*\}$ as a Boolean function $g_x := g(x,\cdot)$ on a hypercube $\{0,1\}^s$.

$g_x\colon \{0,1\}^s \to \{0,1\}$

For each $x \in \{0,1\}^t$:

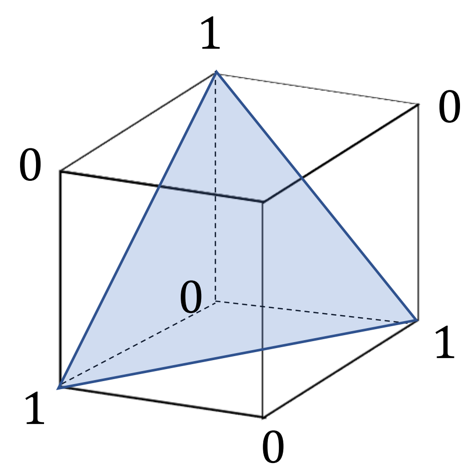
$f(x) \in \{0,1\}$

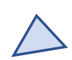


$g_x(0^s) \coloneqq f(x)$

$g_x(y) \coloneqq 0$ elsewhere

$f(x) = *$



- Pick a random linear subspace L_x  and define g_x as its characteristic function.
- Define $g(x, y) \coloneqq g_x(y)$.

Claim

The following holds with high probability:

(The minimum $\mathrm{DNF} \circ \mathrm{XOR}$ circuit size for g)

$=$ (The minimum circuit size for f) + $\left|f^{-1}(*)\right|$

Idea:

- Imagine an optimal way of covering $g^{-1}(1)$.
  - $g^{-1}(1)$ consists of $f^{-1}(1) \times \{0\}^s$ and $\{x\} \times L_x$ for each $x \in f^{-1}(*)$.
- In order to cover $g^{-1}(1)$ by affine subspaces, random linear subspaces $\{x\} \times L_x$ should be used for each $x \in f^{-1}(*)$.
- Then we need to cover $f^{-1}(1) \times \{0\}^s$, but we may *optionally* cover $f^{-1}(*) \times \{0\}^s$.

Proof Outline

Theorem (Main Result)

$$\mathrm{NP} \leq_m^p (\mathrm{DNF} \circ \mathrm{XOR}){-}\mathrm{MCSP}$$

<u>Step 1.</u> 2-factor approx. of r-Bounded Set Cover (NP-hard [Trevisan 2001]) $\leq_m^{\mathrm{ZPP}}$ $(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP for *partial* functions

<u>Step 2.</u> $(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP for *partial* functions $\leq_m^{\mathrm{ZPP}}$ $(\mathrm{DNF} \circ \mathrm{XOR})$-MCSP

<u>Step 3.</u> Derandomization using ϵ-biased generators

[Naor & Naor (1993)]

Step 3: Derandomization

Fact (folklore; a nearly optimal PRG for $\mathrm{AND} \circ \mathrm{XOR}$ circuits)

Any ϵ-biased generator ϵ-fools any $\mathrm{AND} \circ \mathrm{XOR}$ circuit.

- Can be proved by using a simple Fourier analysis.
- Our probabilistic arguments work even if randomness is replaced by the output of an ϵ-biased generator.
  - Careful analysis: sub-conditions can be checked by $\mathrm{AND} \circ \mathrm{XOR}$ circuits
- Extending the fact to $\mathrm{AND} \circ \mathrm{MOD}_m$ requires some extra work.

Open Problems

- NP-hardness of Depth3-AC^0-MCSP under quasipolynomial-time deterministic reductions, or randomized polynomial-time reductions?
  - The Kabanets-Cai obstacle is not applied to these reductions.
- What about $\mathcal{C}$-MCSP for $\mathcal{C} = \mathrm{MAJ} \circ \mathrm{MAJ}, \mathrm{OR} \circ \mathrm{MAJ}$?