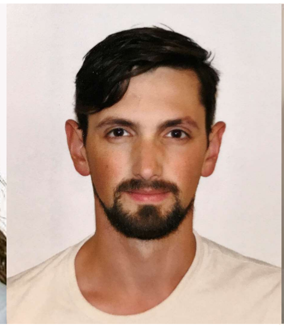
Truly Supercritical Trade-offs for Resolution, Cutting Planes, Monotone Circuits and Weisfeiler–Leman

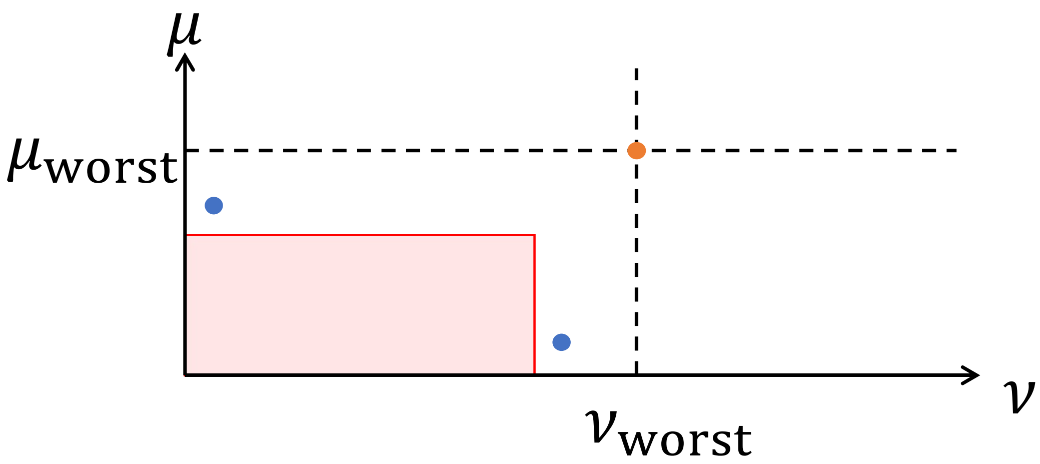
Jakob Nordström
University of Copenhagen & Lund University

Dagstuhl
March 2025

joint with Susanna F. de Rezende, Noah Fleming, Duri Andrea Janett, Shuo Pang

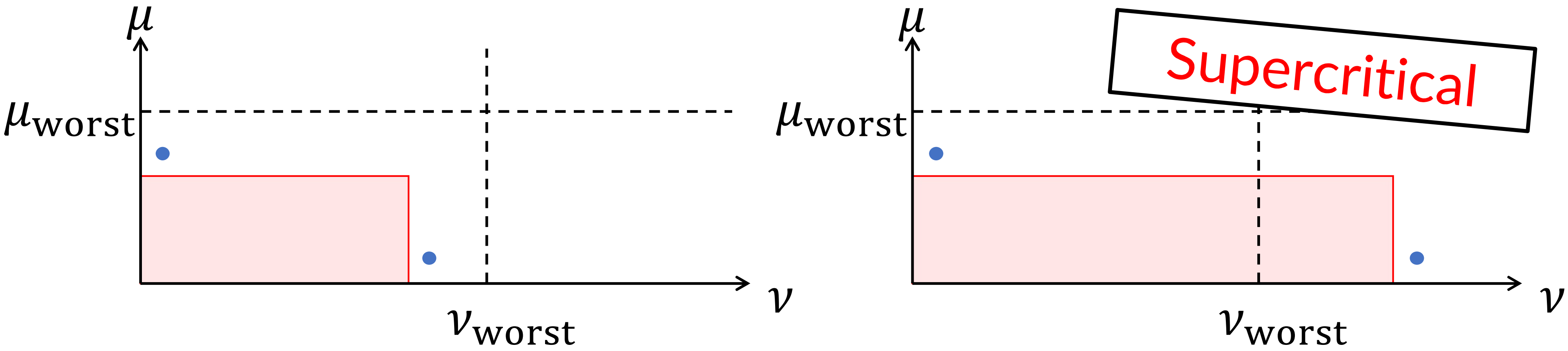
What Is a Trade-off Result?



Computational model with two complexity measures μ, ν (e.g. μ = time and ν = space)

- brute force algorithm can achieve worst case
- can optimize ν, but then μ bad
- can optimize μ, but then ν bad
- impossible to optimize both

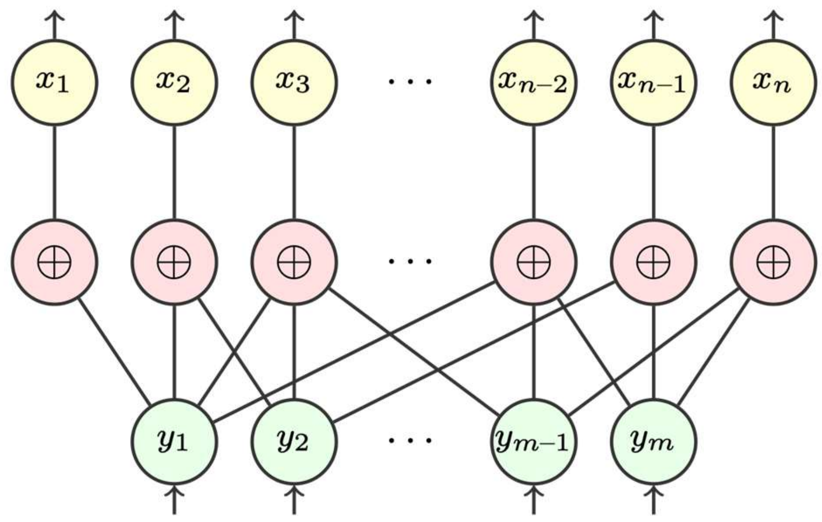
A New Kind of Trade-off [Razborov '16]



→ Optimizing μ pushes ν way beyond brute-force worst case

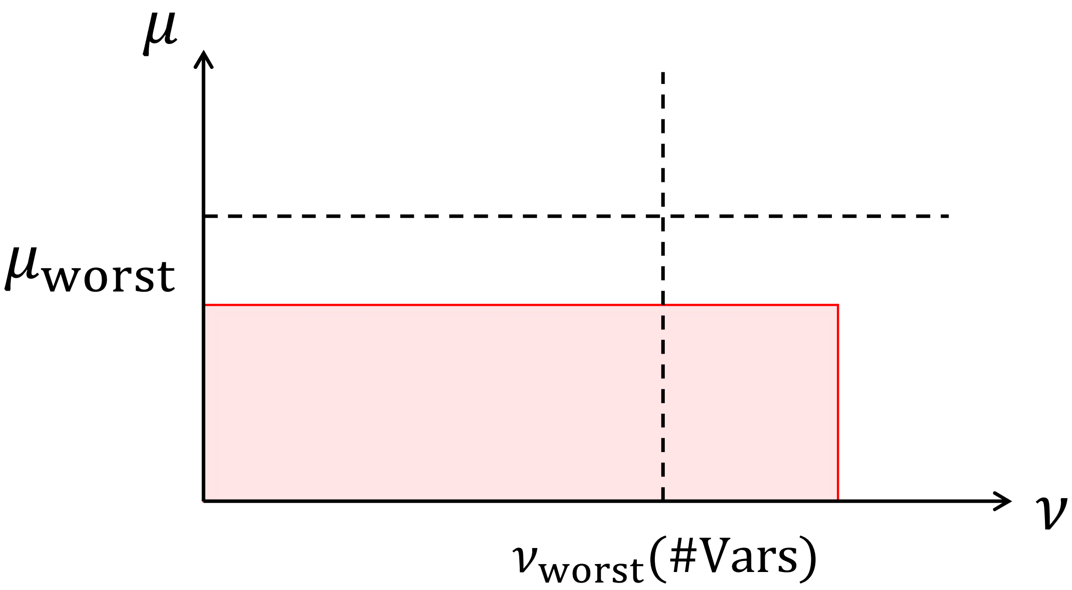
Supercritical Trade-offs Through Hardness Condensation

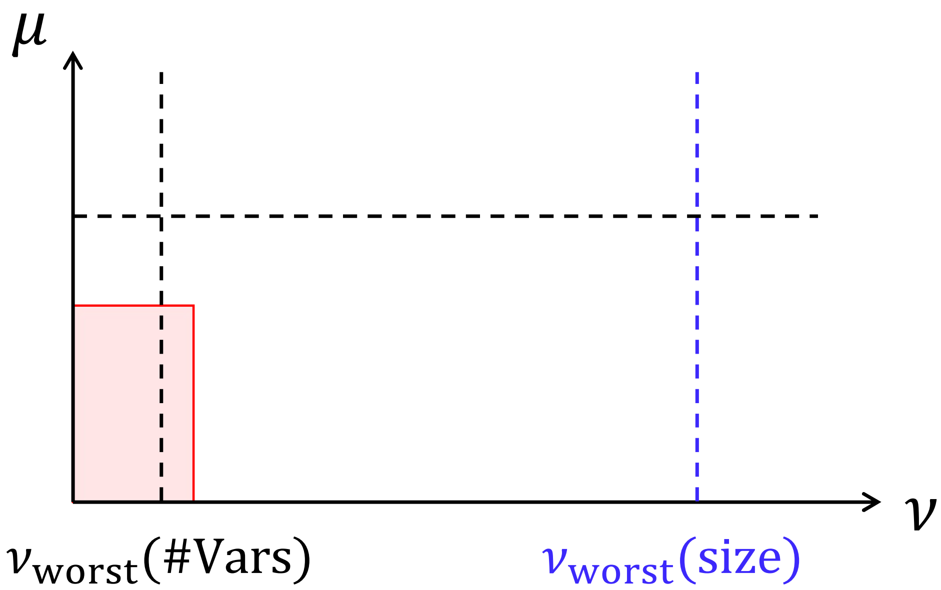
- Take medium-hard input in variables $x_1, \ldots, x_n$
- «Condense» by substituting x_i by XORs over subsets of variables $y_1, \ldots, y_m$
- Show hardness is nearly preserved
- But measured in $m \ll n$: supercritical



[Razborov '16, Berkholz-Nordström '20, Fleming-Pitassi-Robere '22, Berkholz-Nordström '23, ...]

Supercritical in What?





All trade-offs supercritical in # variables only, except [Berkholz '12, BBI '12/'16, BNT '13]

Are there trade-offs *truly* supercritical in input size?

Overview of Our Results (Informal)

Truly supercritical trade-offs for

- depth vs width in resolution
- size vs width in tree-like resolution
- size vs depth in resolution and cutting planes
- size vs depth for monotone circuits
- dimension vs iteration number for Weisfeiler–Leman

Answering open questions in [Razborov ‘16, GGKS ‘18, FGIPRTW ‘21, FPR ‘22, GLNS ‘23]

Next up:

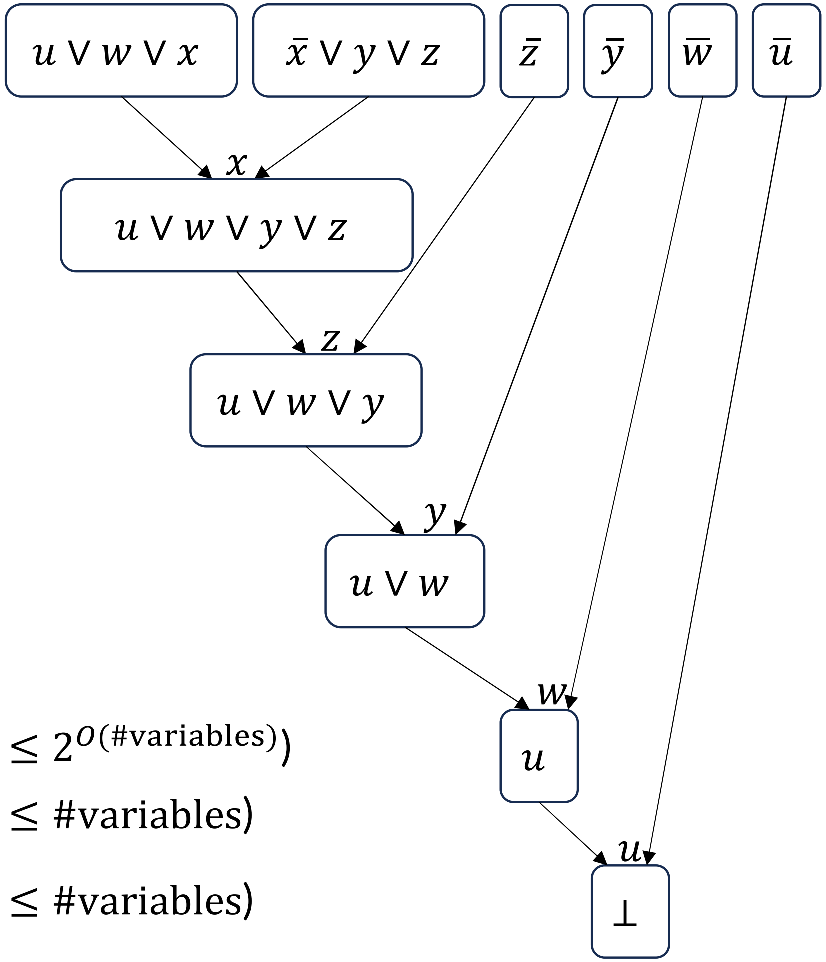
- More precise statement of the results
- Preliminaries to make sense of statements

Resolution Proof System

Goal: prove CNF formula unsatisfiable
Proof of unsatisfiability: Refutation

Resolution rule:

$$\frac{C \vee x \qquad D \vee \overline{x}}{C \vee D}$$



size = #nodes = 11 (worst case $\leq 2^{O(\#\text{variables})}$)
width = max clause size = 4 (worst case $\leq \#\text{variables}$)
depth = max path length = 5 (worst case $\leq \#\text{variables}$)

Our Results: Resolution

Theorem (depth-width trade-off for resolution)

$\exists$ CNF formulas F_n on n variables s.t.

- Refuted by resolution in width $w = \text{poly}(\log(n))$
- But width $\leq 1.9w \implies$ supercritical depth $\text{superlinear}(|F|)$

Theorem (size-depth trade-off for resolution)

$\exists$ CNF formulas F_n on n variables s.t.

- Refuted by resolution in size $s = \text{quasipoly}(|F|)$
- But size $\leq ns \implies$ supercritical depth $\text{superpoly}(|F|)$

Theorem (size-width trade-off for treelik

$\exists$ CNF formulas F_n on n variables s.t

- Refuted by treelike resolution in width $\text{w} = \text{poly}(\log(n))$
- But width $\leq w + \sqrt{w} \implies$ supercritical size $\exp(\text{superpoly}(|F|))$

All results supercritical in input size

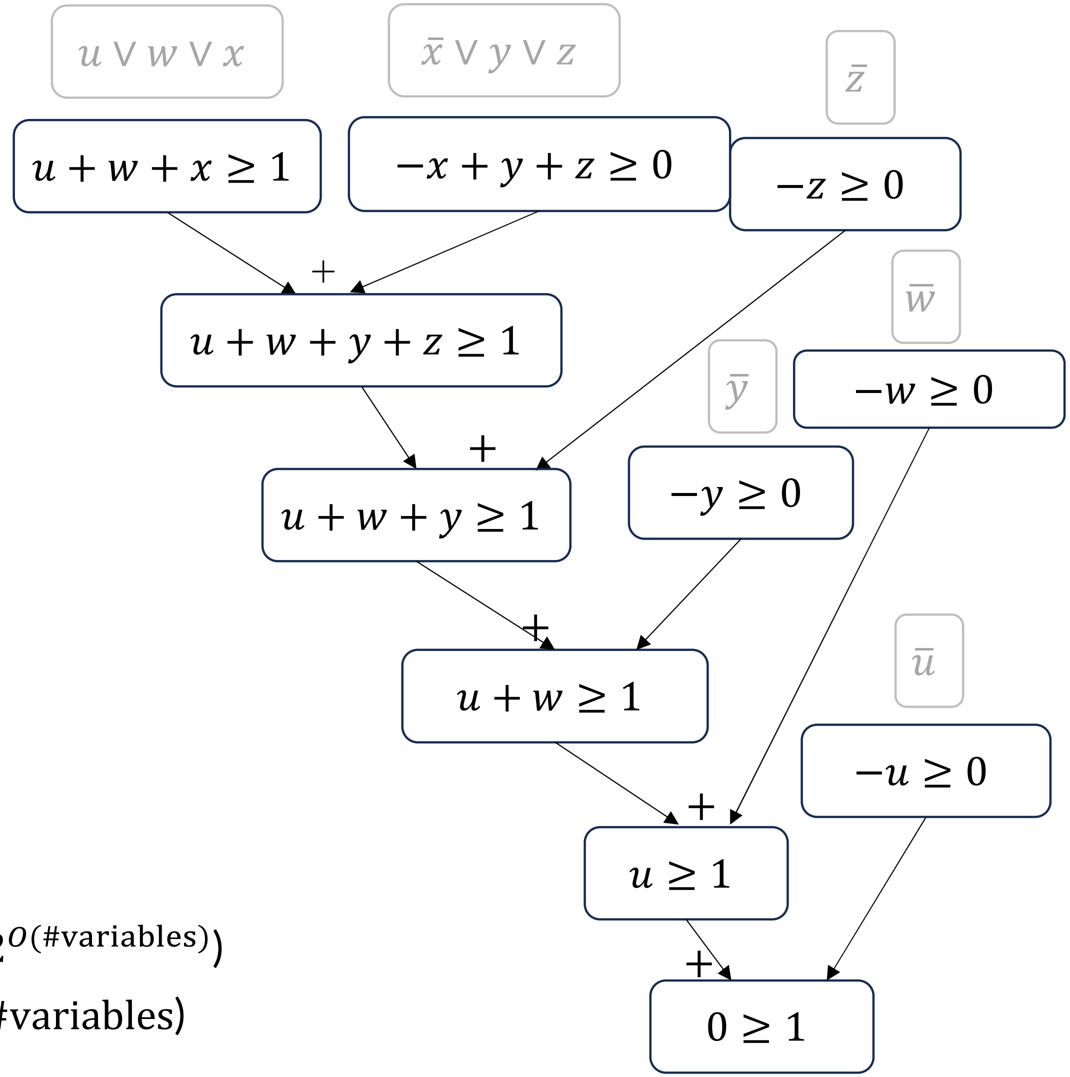
Cutting Planes Proof System

- Translate clause $\bar{x} \vee y \vee z$ to linear inequality $(1 - x) + y + z \geq 1$

Derivation rules:

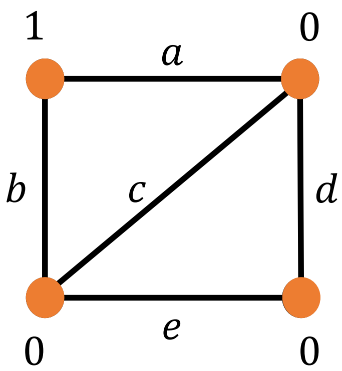
- Addition: $\dfrac{\sum a_i x_i \geq A, \quad \sum b_i x_i \geq B}{\sum a_i x_i + \sum b_i x_i \geq A + B}$
- Multiplication: $\dfrac{\sum a_i x_i \geq A}{\sum c a_i x_i \geq cA}, \; c \in \mathbf{N}^+$
- Division: $\dfrac{\sum c a_i x_i \geq A}{\sum a_i x_i \geq \lceil A/c \rceil}, \; c \in \mathbf{N}^+$

size = #nodes (worst case $\leq 2^{O(\#\text{variables})}$)
depth = max path length (worst case $\leq$ #variables)



Tseitin Formulas (Handshake Lemma)

- Graph $G = (V, E)$
- Labelling $\mathrm{lbl}(v) \in \{0, 1\}$ for $v \in V$ s.t. $\sum_v \mathrm{lbl}(v)$ odd
- Edges $e \in E \;\Leftrightarrow$ variables x_e
- Constraints $\sum_{e \ni v} x_e \equiv \mathrm{lbl}(v) \pmod 2$



$x_a \vee x_b$	$\overline{x_a} \vee \overline{x_b} \vee \overline{x_d}$	$\overline{x_b} \vee \overline{x_c} \vee \overline{x_e}$	$x_d \vee \overline{x_e}$
$\overline{x_a} \vee \overline{x_b}$	$\overline{x_a} \vee x_b \vee x_d$	$\overline{x_b} \vee x_c \vee x_e$	$\overline{x_d} \vee x_e$
	$x_a \vee \overline{x_b} \vee x_d$	$x_b \vee \overline{x_c} \vee x_e$	
	$x_a \vee x_b \vee \overline{x_d}$	$x_b \vee x_c \vee \overline{x_e}$	

- Extensive study since [Tseitin '68]
- Exponential resolution size lower bound [Urquhart '87]
- Also studied for sum-of-squares (SoS), bounded-depth Frege, stabbing planes, …

Supercritical trade-offs for Tseitin Formulas?

- Long-standing conjecture: Tseitin formulas exponentially hard for cutting planes FALSE!

Theorem [Dadush, Tiwari '20]
Exist size-$n^{O(\log n)}$ cutting planes proofs for Tseitin

- Cutting planes requires depth $\geq \Omega(n)$ [FGIPRTW '21]
- But [Dadush, Tiwari '20] refutations have supercritical depth $n^{O(\log n)}$

Do Tseitin formulas yield supercritical trade-offs for size vs depth in cutting planes?

- Supercritical trade-offs in # variables [Fleming-Pitassi-Robere '22]

Our Result: Cutting Planes

Supercritical
in input size

Theorem (size-depth trade-off for cutting planes)

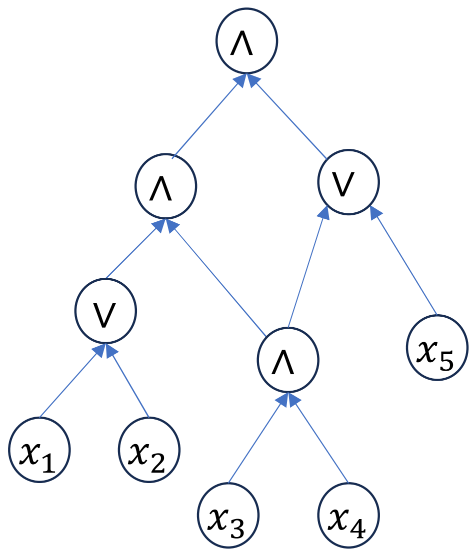
$\exists$ CNF formulas F_n on n variables s.t.

- Refuted by cutting planes in size $s = \mathrm{quasipoly}(|F|)$
- But size $\leq ns \implies$ supercritical depth $\mathrm{superpoly}(|F|)$

- Does not resolve question for Tseitin formulas
- But Tseitin formulas used to construct formulas with provable trade-offs

Monotone Circuits

- Boolean circuits with AND, OR gates
- Compute Boolean functions $f: \{0,1\}^n \to \{0,1\}$



size = #nodes (worst case $\leq 2^{O(n)}$)
depth = max path length (worst case $\leq n$)

Previous work:

Non-supercritical trade-offs: size $O(n) \Longrightarrow$ depth $\Omega(n/\mathrm{polylog} n)$

[KW '90, RM '97, GP '14, dRMNPRV '20]

Our Result: Monotone Circuits

Theorem (supercritical trade-off for monotone circuits)

$\exists$ monotone functions f_n on n variables s.t.

- Computed by monotone circuit of size $s = \mathrm{quasipoly}(n)$
- But size $\leq ns \Longrightarrow$ supercritical depth $\mathrm{superpoly}(n)$

Structure of Proof

- Part I: Base trade-off
- Part II: Use lifting to get other trade-offs

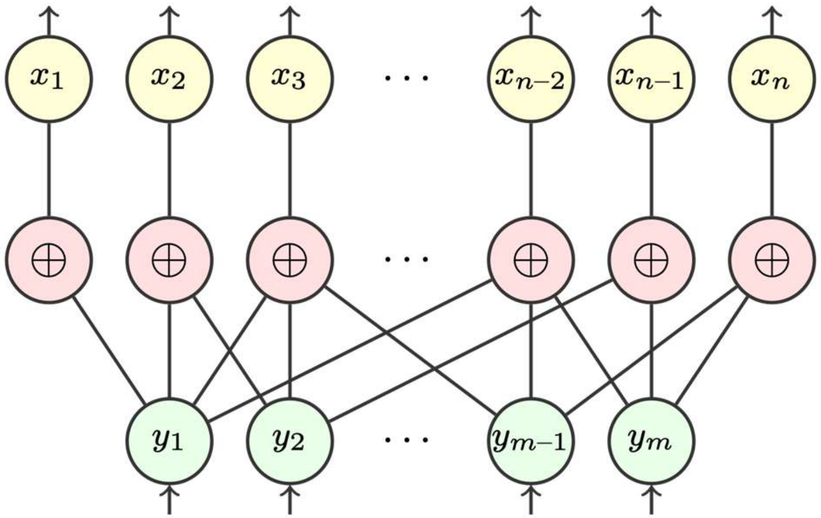
Base Result: Resolution Depth vs Width

Theorem

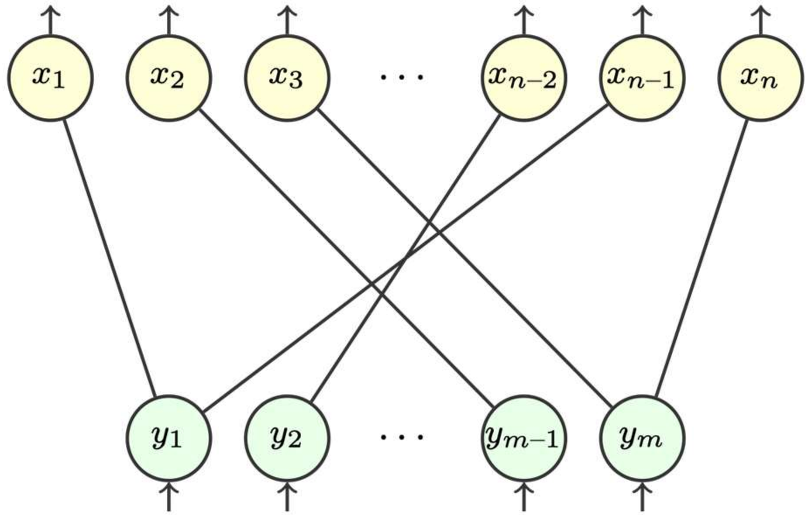
For any $c < k < \frac{n}{2 \ln n}$ there are 4-CNF formulas s.t.

- formula size s ($\approx n^c$)
- exists proof in width $k + 3$
- but width $< k + c \Rightarrow$ depth $> s^{k/c}$

Variable Compression [Grohe-Lichter-Neuen-Schweitzer 2023]

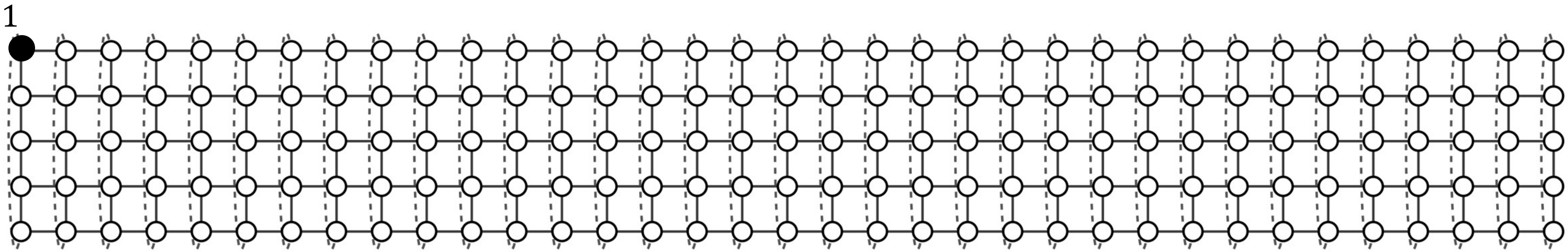


Substitution with XOR gadgets



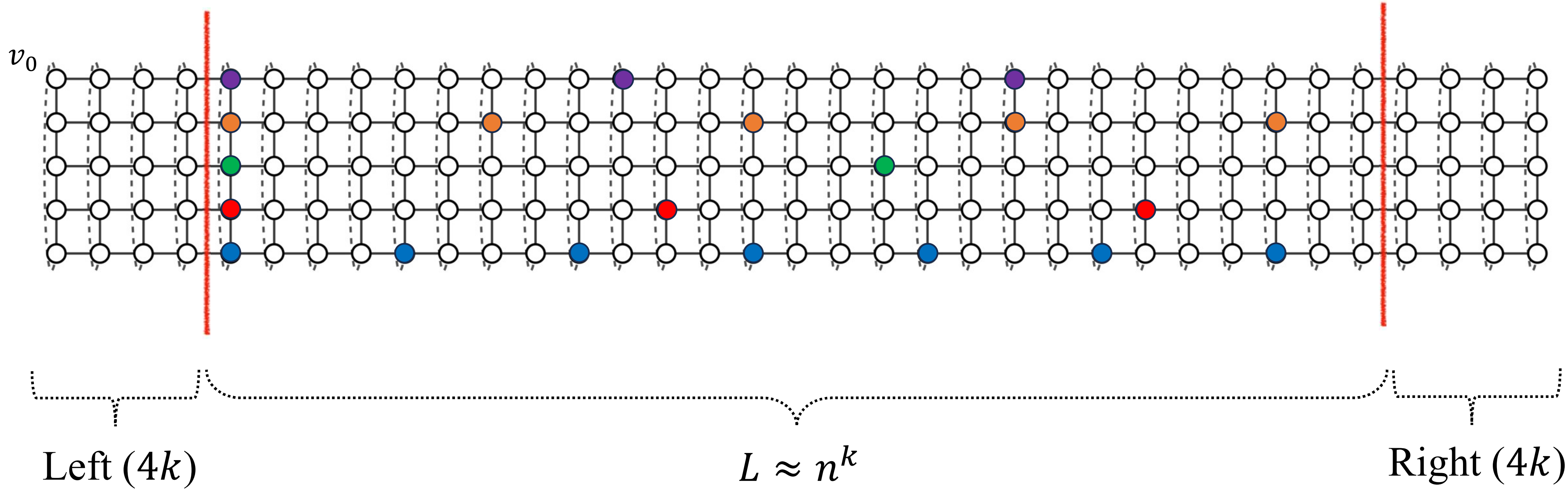
Variable substitution
(with lots of collisions)

Formula: Tseitin on Cylinder



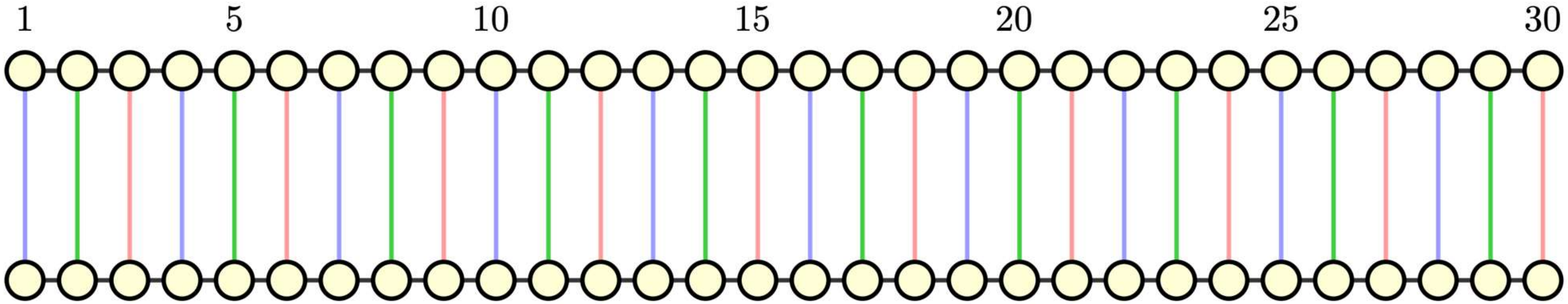
- Tseitin on long, skinny cylinder
  (wrap-around vertically, but not horizontally)
- Only 1 labelled vertex at top left

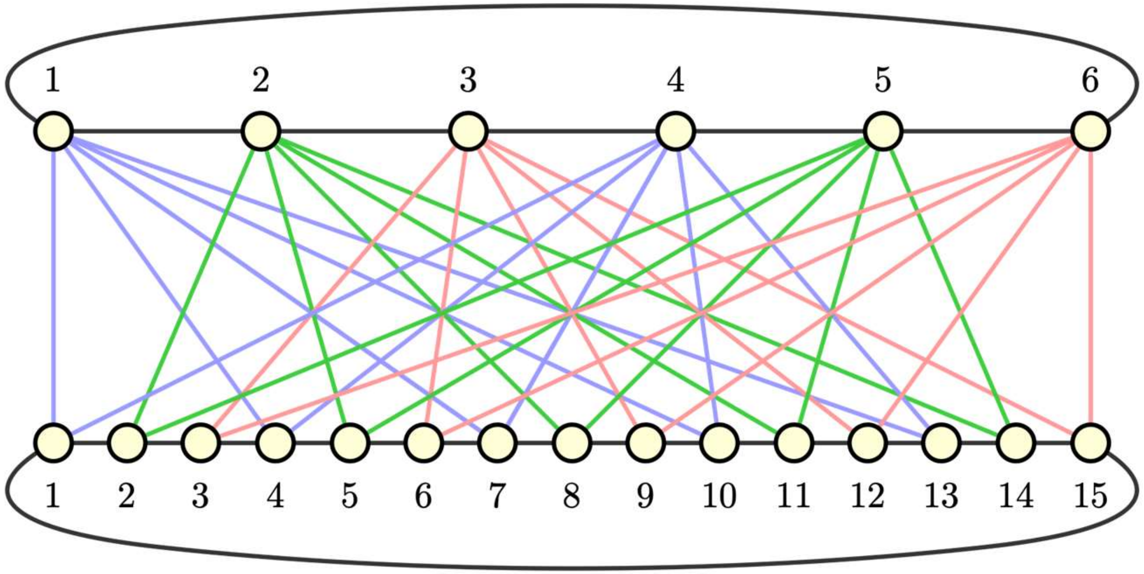
Formula: After Compression



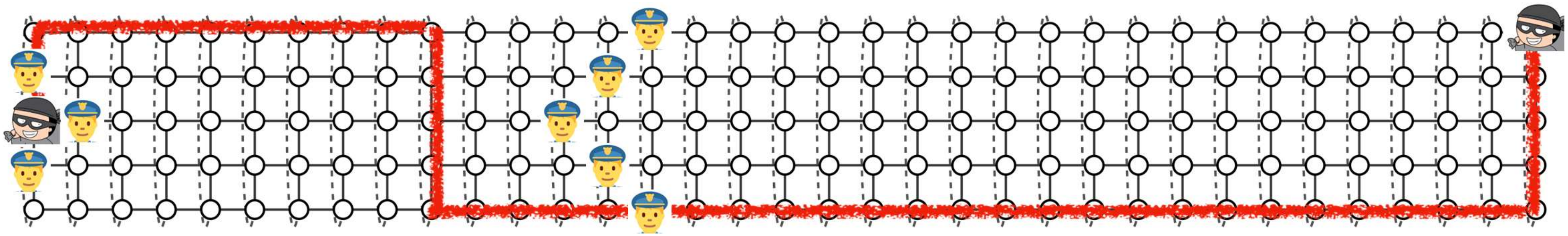
- Vertex equivalence classes $[v_\ell] = \{v_j \mid j \equiv \ell \bmod m_i\}$ for modulus m_i chosen for row i (except ends)
- Induces edge equivalence classes $[e]$
- Compressed formula: $\sum_{[e]\ni[v]} y_{[e]} = 1 \bmod 2 \quad \text{iff} \quad [v] = [v_0]$

Edge Equivalence Between Two Rows ($m_1 = 6, m_2 = 15$)





Proof: By Analyzing the Cop-Robber Game [Seymour-Thomas '93]

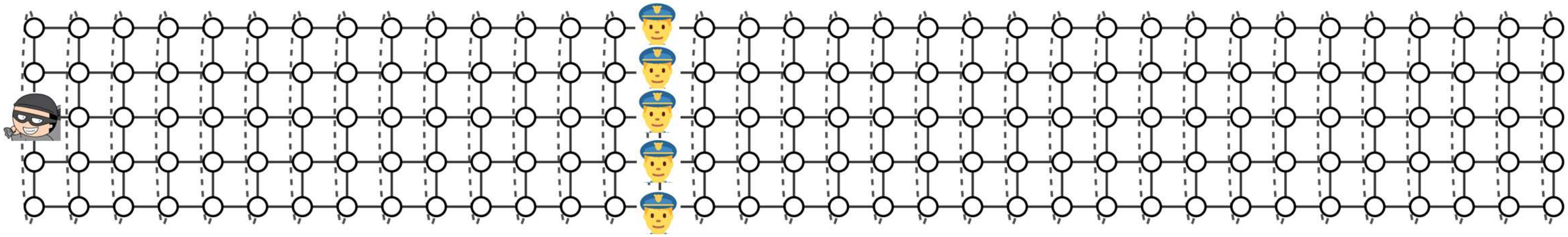


- Start: K cops, one robber at v_0
- In each round:
  - One cop enters helicopter and signal a vertex v
  - Robber moves
  - Cop lands at v
- Ends when robber is caught (by cop at same vertex)

width ≈ # cops
depth ≈ # rounds

[GTT '18]

Upper Bounds: Simple Cop Strategy



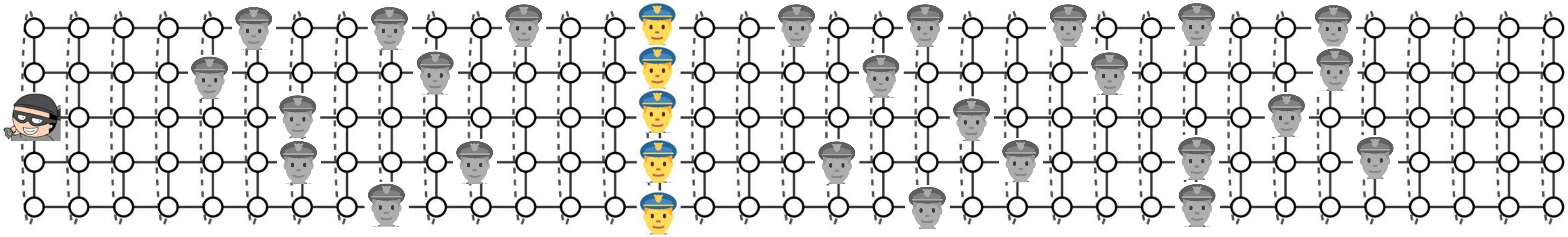
- $k + 1$ cops:
  - Place cops on middle column
  - March towards robber in $k \cdot L$ rounds

  $\Rightarrow$ translates to resolution proof of width $k + 3$, but depth $k \cdot L$

- $3k$ cops:
  - Binary search

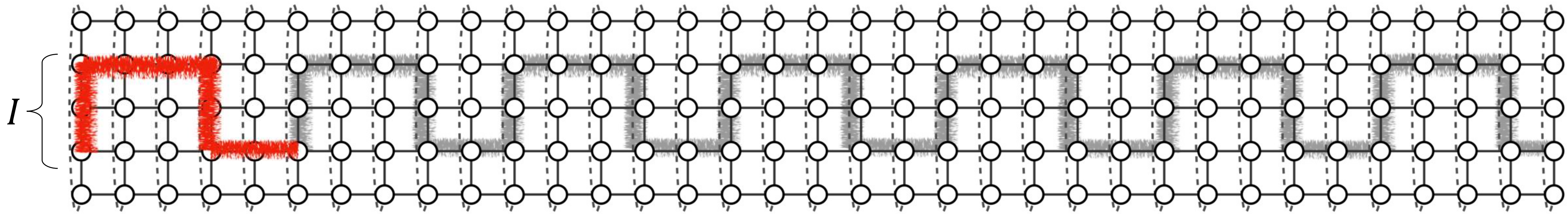
  $\Rightarrow$ translates to resolution proof of width $3k$ and depth $k \cdot \log L$

Compressed Game [GLNS '23]



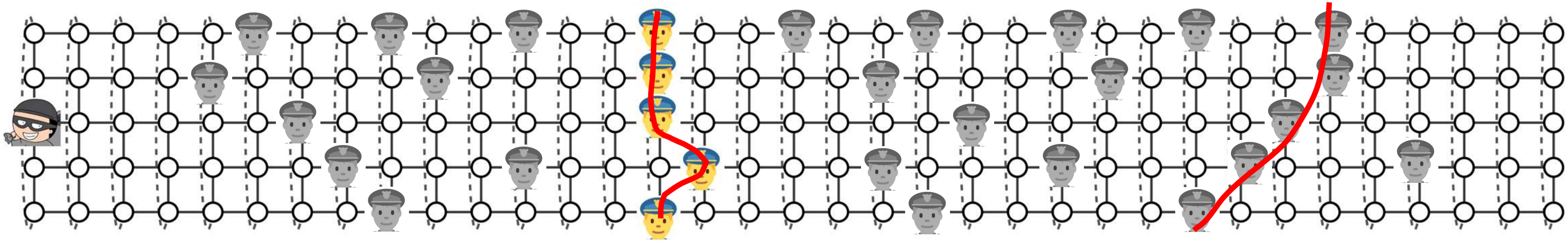
- cop at $v \Longrightarrow$ cop at all vertices $u \in [v]$
- $K = k + c$ cops, one robber at v_0
  – Lift a cop and signal a vertex v
  – Robber does a $\equiv$-compressible move
  – Cop lands at $[v]$
- Cop strategies for uncompressed game still work $\Longrightarrow$ same upper bounds
- But robber has to avoid cop clones $\Longrightarrow$ harder to get lower bounds
- Consider only special type of Robber moves

Moves Translatable to Compressed Setting



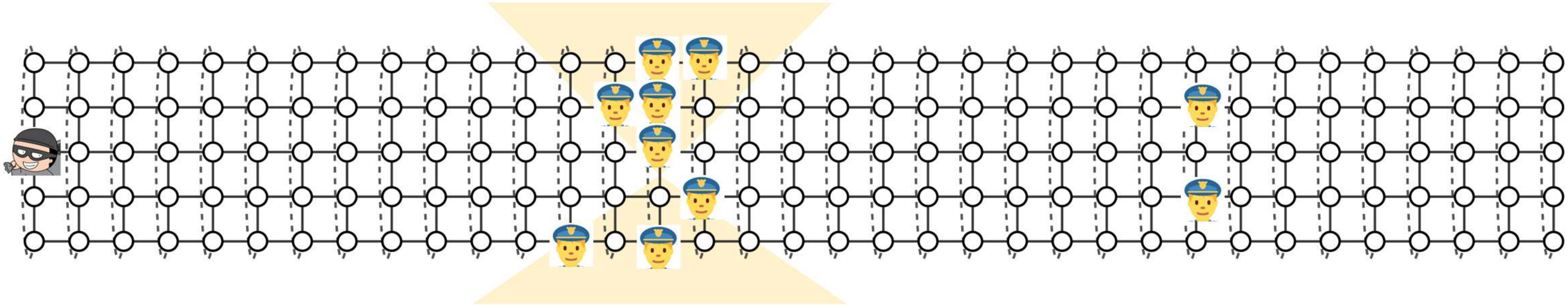
- Robber moves only on $|I| \leq c + 1$ contiguous rows
- Horizontal moves periodic mod $\gcd(m_i\colon\ i \in I)$

Idea for Robber Strategy



- Robber stays on cop-free column in left/right part (L/R)
- Slides between L, R using compressible moves
- Problem: cops can form a police cordon

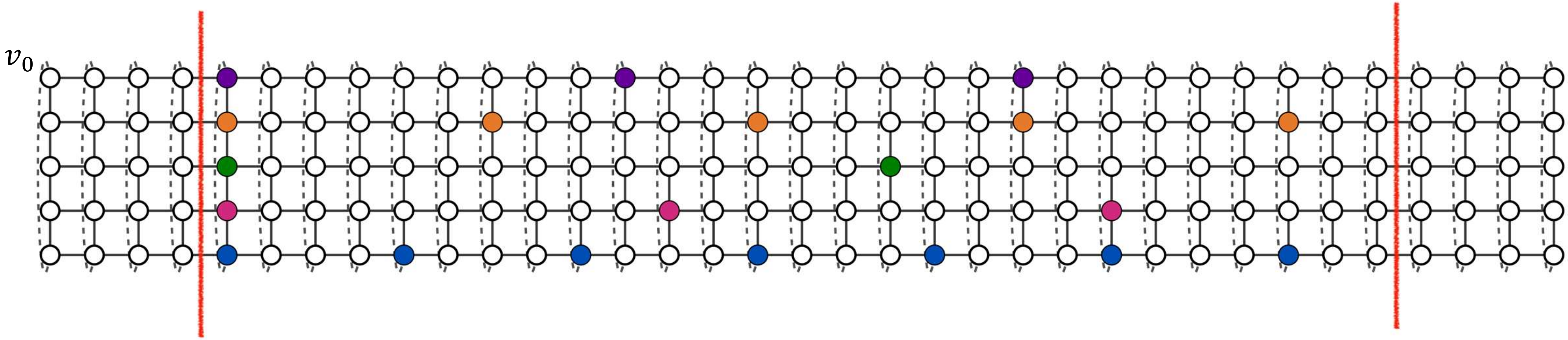
Idea for Robber Strategy



- Keep away from (set of) virtual cordons
- This allows robber to escape in time
- Cordons can only move slowly

Survives n^k rounds against $k + c$ Cops.

Choice of Row Moduli



- Fix $1 \leq c \leq k-2$
  Pick coprime numbers $P_1, \dots, P_k$ where $|P_i| \approx n$
  —ith modulus $4k \cdot P_i \cdots P_{i+c}$
  —cylinder length $L \approx 4k \cdot P_1 \cdots P_k$
- Compressed formula size $n^k \to n^{c+1}$

Motivation for Grohe et al: Weisfeiler-Leman (WL) algorithm

Theorem [Grohe-Lichter-Neuen-Schweitzer '23]

∃ graph pairs k-dimensional WL can distinguish, but only after $n^{k/2}$ iterations.

- Dimension ≈ Width
- Iterations ≈ Depth
- Graph pair ≈ Tseitin

- [GLNS '23] not robust enough to yield proof complexity results

Our Result for Weisfeiler-Leman (WL)

Theorem

For any $1 \leq c \leq k-1$, $\exists$ graph pairs of size n

- dimension-k WL can distinguish
- dimension-$(k+c-1)$ WL requires $n^{\frac{k}{c+1}}$ iterations

Structure of Proof

- Part I: Base trade-off
- Part II: Use lifting to get other trade-offs

What is a Lifting Theorem?

- Use composition to relate (different) models of computation:

  Complexity of f in (weak) model A corresponds to

  $\Rightarrow$ Complexity of composed problem $f \circ g$ in (strong) model B

- Example

  F requires large resolution width

  $\Rightarrow$ $F \circ g$ requires large resolution/cutting planes/monotone circuit size

Lifting with Indexing Gadget (for Functions)

- $f: \{0,1\}^n \to \{0,1\}$
- Compose with IND: $[m] \times \{0,1\}^m \to \{0,1\}$, $\mathrm{IND}(x,y) = y_x$

Alice: $x \in [m]^n$

1	3	1	2	2	1	3	2

Bob: $y \in \{0,1\}^{mn}$

0	0	1	1	0	1	0	1
1	0	0	1	1	0	0	0
0	1	1	1	0	1	0	0

$f \circ \mathrm{IND}(x,y) = f\,($

0	1	1	1	1	1	0	0

$)$

Lifting Resolution to Monotone Circuits [Garg-Göös-Kamath-Sokolov '18]

- Resolution width-depth trade-off $\Rightarrow$ monotone circuit size-depth trade-off

Theorem [GGKS '18]
For CNF formula F and large enough m, can construct function $f_{F,m}$ s.t.
size-m^w, depth-d monotone circuit for $f_{F,m}$
$\Rightarrow$ width-$O(w)$, depth-$O(dw)$ resolution proof for F

- Crucial issue: #variables of function $\approx$ #clauses in formula
  – Previous supercritical trade-offs do not work
  – Crucial that depth-width trade-off is *truly* supercritical

- Need tighter lifting theorem than [GGKS '18] with better constants

Tight Lifting Theorem

Best possible theorem (dream)

For CNF formula F and large enough m, can construct function $f_{F,m}$ s.t.

size-m^w, depth-d monotone circuit for $f_{F,m}$

$\Rightarrow$ width-w, depth-d resolution proof for F

→ Intermediate goal: tight lifting for resolution

Tight Lifting for Resolution

Theorem

size-$(m/2)^w$, depth-d resolution refutation for $F \circ \mathrm{IND}_m$
$\Rightarrow$ width-w, depth-d resolution refutation of F

- Essentially optimal
- Simple proof via random restriction

Corollary (size-depth trade-off for resolution)

$\exists$ a CNF formula F on ℓ clauses s.t.

- $\exists$ resolution refutation of F in size s
- But size $\leq \ell s \Rightarrow$ superpolynomial (in ℓ) depth

Tight Lifting for Monotone Circuits

Theorem

size-$m^{(1-\epsilon)w}$, depth-d monotone circuit for $f_{F,m}$

$\Rightarrow$ width-w, depth-wd resolution refutation of F

- Also almost optimal
- Applies also to cutting planes (and monotone real circuits)

Corollary (size-depth trade-off for monotone circuits)

$\exists$ function f on n variables s.t.

- $\exists$ a monotone circuit for f of size s
- size $\leq ns \Rightarrow$ superpolynomial depth

Lifting for Treelike Resolution

Theorem

size-s, width-w tree-like resolution refutation for $F \circ \mathrm{XOR}_{m+1}$

$\Rightarrow$ depth-$(\log s)$, width-(w/m) resolution refutation of F

- Note that width is also reduced

Corollary (size-width trade-off for treelike resolution)

$\exists$ a CNF formula F on n variables s.t.

- $\exists$ a treelike resolution refutation of width $\mathrm{w} = \mathrm{poly}(\log(n))$
- But width $\leq w + \sqrt{w} \Longrightarrow \exp(\mathrm{superpoly}(|F|))$ size

Concurrent Work [Göös-Maystre-Risse-Sokolov '24]

- Independently obtained supercritical size-depth trade-offs
- Natural formulas formalizing interesting combinatorial principle
- More robust trade-offs
- Can use existing lifting theorems as black box

Our work
- Different parameter regime, with trade-offs kicking in earlier
- Robust supercritical trade-offs also for Weisfeiler–Leman
- Formula compression technique in [GLNS '23] worth investigating further
- Tighter lifting theorems also of independent interest

Open Problems

- Further applications of compression in proof complexity
  - — Can pebbling formulas be compressed?
  - — Can we find other graph compressions?
- Prove more supercritical trade-offs
  - — Size vs space or similar
  - — Unified parameter range with [GMRS '24]
- Complexity of Tseitin for cutting planes
  - — Is there a supercritical trade-off?
- Complexity of perfect matching for monotone circuits
  - — Superpolynomial lower bound known [Razborov '85]
  - — Linear depth necessary [Raz–Wigderson '92]
  - — Is there a supercritical trade-off?

Conclusion

- We give *truly* supercritical (in terms of size) trade-offs for
  - depth vs width in resolution
  - size vs width in tree-like resolution
  - size vs depth in resolution and cutting planes
  - size vs depth for monotone circuits
  - dimension vs iteration number for Weisfeiler–Leman
- Proven via lifting base trade-off (depth vs width)
  - Depth vs width: compressed Tseitin
  - Analysis via Cop-Robber game
  - Also need improved (nearly tight) lifting theorems

Directions for Future Research

Thanks for your attention!

- Compression of other formulas?
- Compression of other graphs than cylinders?
- Supercritical trade-offs for other measures such as size vs space?
- Supercritical size-depth trade-offs for Tseitin formulas?
- Supercritical monotone circuit trade-offs for perfect matching?