Lifting to Parity Decision Trees via Stifling

Arkadev Chattopadhyay — TIFR, Mumbai

Nikhil Mande — Univ. of Liverpool

Swagato Sanyal — IIT Kharagpur

Suhail Sherif — Univ. of Lisbon

Lifting to Parity Decision Trees via Stifling

Arkadev Chattopadhyay	TIFR, Mumbai
Nikhil Mande	Univ. of Liverpool
Swagato Sanyal	IIT Kharagpur
Suhail Sherif	Univ. of Lisbon

Proof Complexity

Parity Decision Trees

Proof Complexity

every assignment falsifies a clause

Given an unsatisfiable CNF over n variables,

prove that there is no satisfying assignment.

Should require large "proofs" assuming NP $\neq$ coNP.

For what defn of "proof" can we show this?

$\text{FalseClause}_\varphi \subseteq \{0,1\}^n \times [m]$

↑ assignment x ↖ index of clause in $\varphi = \bigwedge_{i \in [m]} C_i$

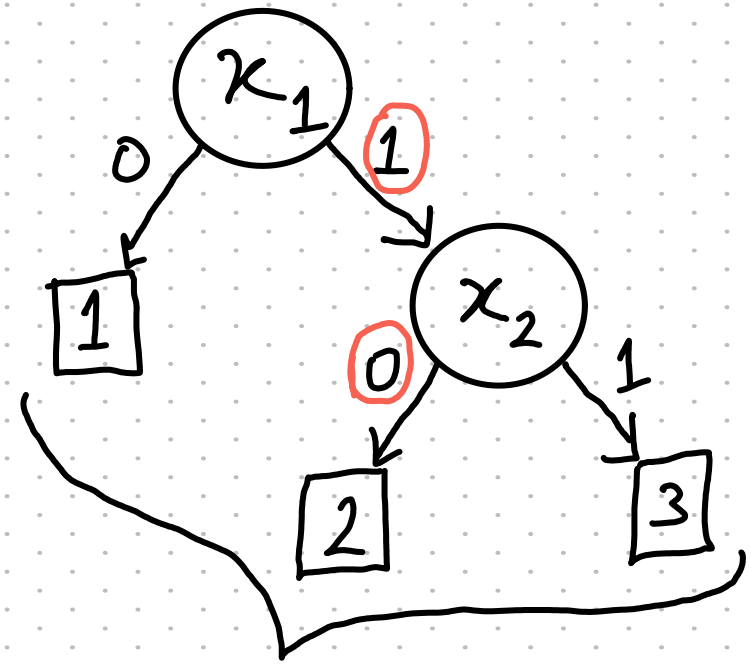
$(x, i) \in FC_\varphi$ if $C_i(x) = \text{False}$

To prove: FC_φ is total

$(\forall x \; \exists i \quad (x,i) \in FC_\varphi)$

Decision Trees are Proofs

$T : \{0,1\}^2 \longrightarrow [3]$



$\forall x \quad T(x) \in [3]$

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

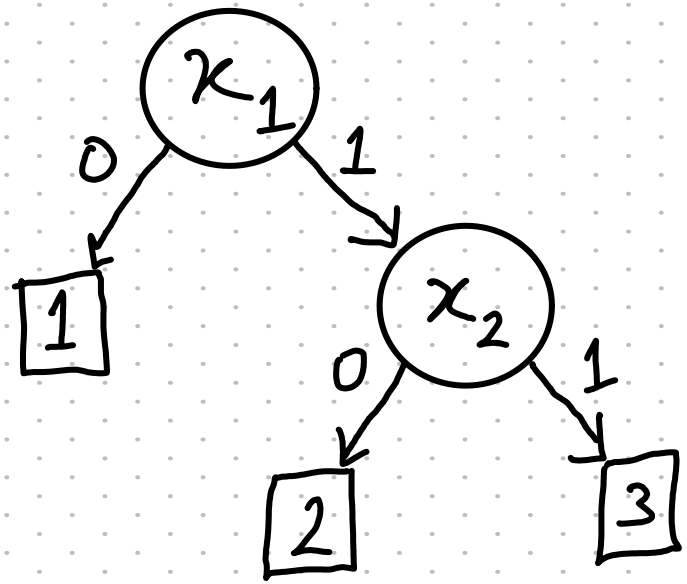
Need to verify:

For each leaf ℓ, the inputs that reach ℓ falsify the clause output at ℓ.

DT of size s is a proof of size s.

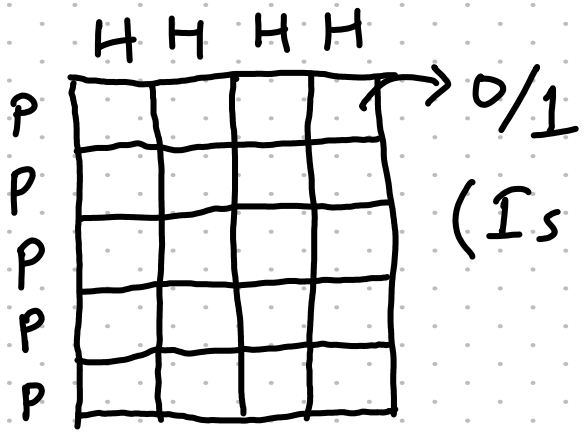
Decision Trees are Tree-like Resolution Proofs

$T : \{0,1\}^2 \to [3]$



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

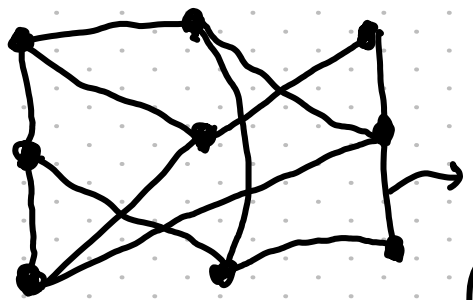
$$\dfrac{\dfrac{x_1 \qquad \neg x_1 \vee x_2}{x_2} \qquad \neg x_2}{\bot}$$



(Is P in H?)

Pigeon Hole Principle

- Each Pigeon is in a Hole
- No Hole has >1 Pigeon



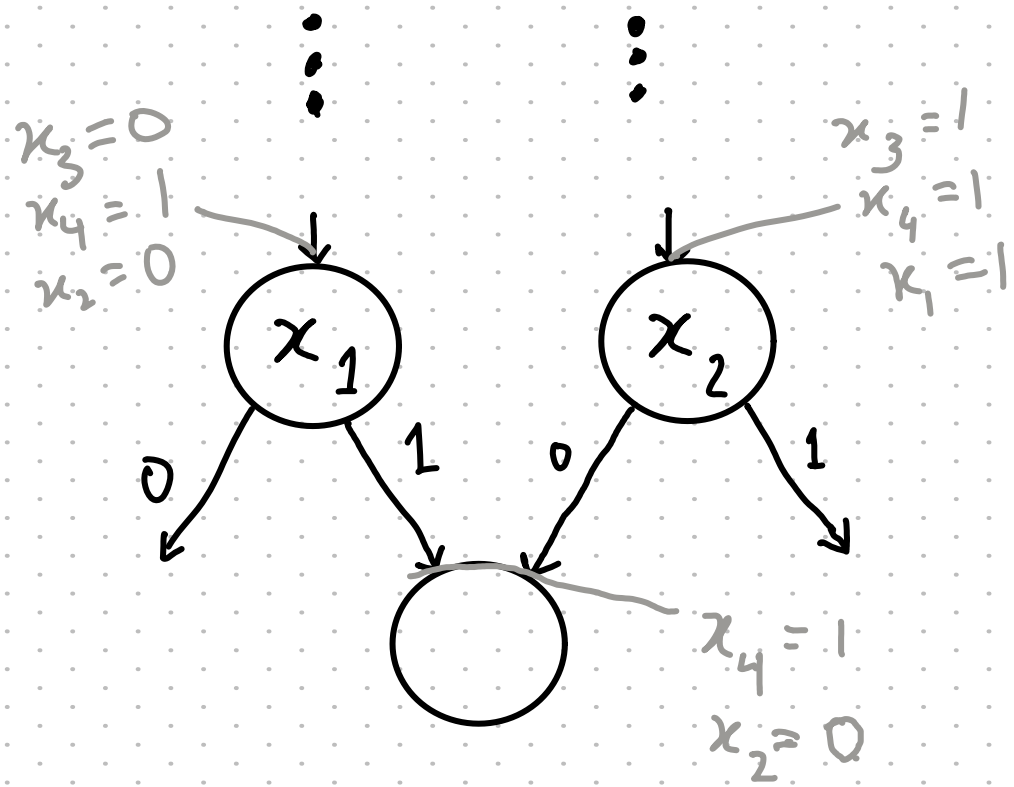
(Is $e \in G_x$?)

Tseitin on a const. degree graph G_n with an odd number of vertices

- Every vertex has odd degree

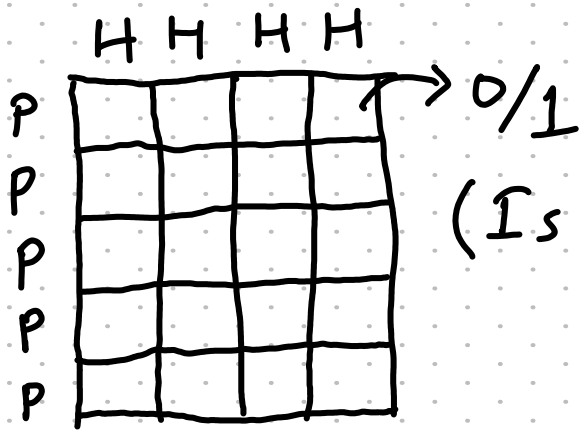
Hard when G_n is an expander.

Decision DAGs are Resolution Proofs



DAG structure allows reusing of proof lines.

Can give exp → poly size reductions for some CNFs!!

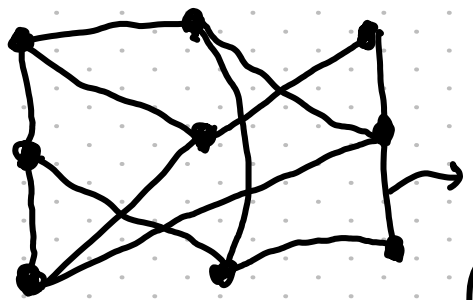


(Is P in H?)

Pigeon Hole Principle

- Each Pigeon is in a Hole
- No Hole has >1 Pigeon

Hard even for DAGs. [Haken '85]



(Is $e \in G_x^?$)

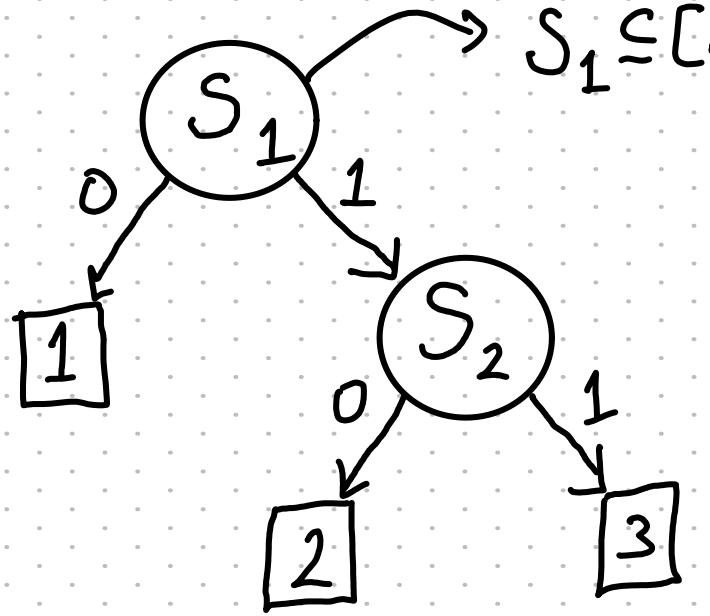
Tseitin on a const. degree graph G with an odd number of vertices

- Every vertex has odd degree

Hard even for DAGs. [Ben-Sasson, Wigderson '01]

Parity Decision Trees

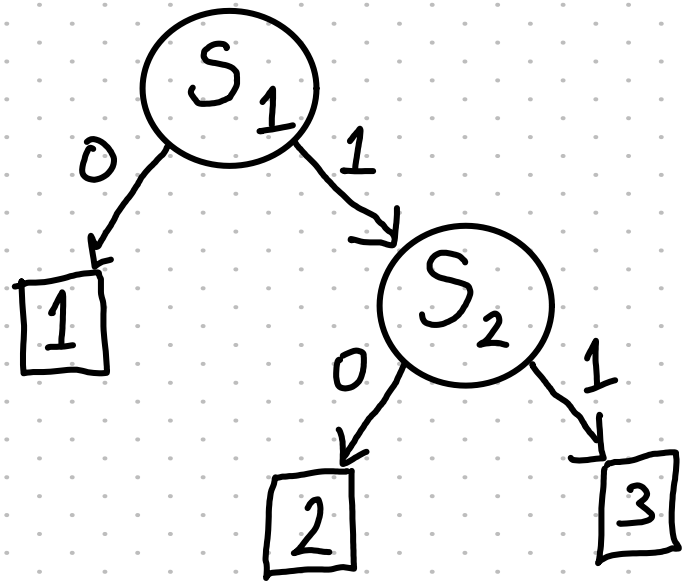
$T : \{0,1\}^2 \to [3]$



$S_1 \subseteq [n]$, query answer is $\bigoplus_{i \in S_1} x_i$

Parity Decision Trees are Tree-like Res(⊕) Proofs

$T : \{0,1\}^2 \to [3]$



Can define Parity DAGs

|||

Res (⊕)

Line : Disjunction of parities

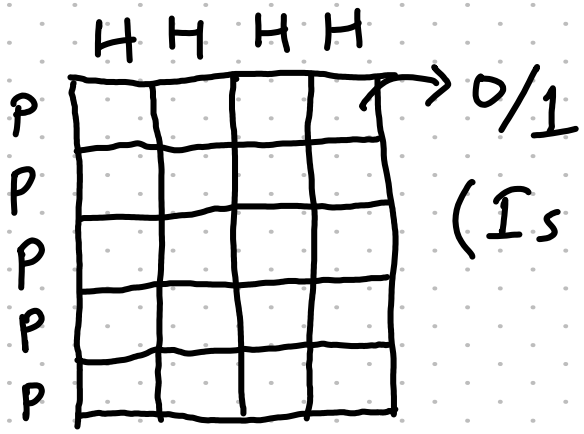
Example : $(x_1 \oplus x_2 = 0) \vee (x_1 \oplus x_3 \oplus x_5 = 1)$

|||

$(x_1 \oplus x_2 = 0) \vee (x_2 \oplus x_3 \oplus x_5 = 0)$

Resolution : $\dfrac{\Phi_1 \vee (P = 0) \qquad \Phi_2 \vee (P = 1)}{\Phi_1 \vee \Phi_2}$

Rebase : $\dfrac{\Phi \vee P_1 \vee P_2}{\Phi \vee P_1 \vee (\neg P_1 \oplus P_2)}$

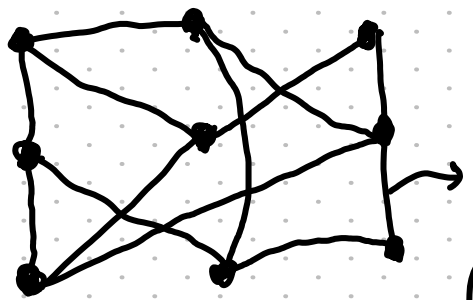


(Is P in H?)

Pigeon Hole Principle

- Each Pigeon is in a Hole
- No Hole has >1 Pigeon

Hard even for PDTs. [IS '14]



(Is $e \in G_x$?)

Tseitin on a const. degree graph G with an odd number of vertices

- Every vertex has odd degree

Easy for PDTs.

No CNF is known to be hard for Parity DAGs!

Hardness of Parity LBs

- Parity constraints are non-local correlations
  Hard to keep track and reason about.
- There are 2^n parity queries that can be asked.
  The Decision DAG size-width relation proof
  no longer works.

How can we tame the power of these parities?

This talk: A look at some parity-poopers.

Lower Bounds for Splitting by Linear Combinations

- Dmitry Itsykson & Dmitry Sokolov

1) Lower bound for Pigeon Hole Principle

2) Lower bound for CNF related to Tseitin

3) Other interesting results

1) Lower bound for PigeonHole Principle

Weakness of Parity query exploited:

If k parity constraints are set and

is a solution,

then for any k+1 indices there is a different solution differing only in those indices.

Add n-k+1 constraints.

Used to say that outputting "Pigeon P is in a Hole" as a False Clause is hard.

2) Lower bound for CNF related to Tseitin

Tseitin variables:

x_e for e in G

F variables:

$y_{e,1}, y_{e,2}$ for e in G

Let $x_e := y_{e,1} \wedge y_{e,2}$

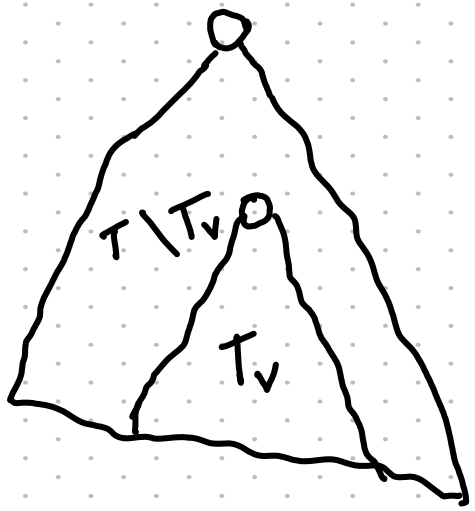
"Vertex v has odd degree" is a fn of $2 \cdot \deg(v)$ variables.

$\Rightarrow$ Can be written as a CNF with $2^{2 \cdot \deg(v)}$ clauses of width $2 \cdot \deg(v)$

$\therefore$, F has $O(n)$ clauses, const. width.

Weakness of Parity query exploited:

Small size PDT $\Rightarrow$ Low rand. comm. comp.



Can check if input goes to T_v or not with Const. communication.

$$R^{cc}(FC_{\ell_e}) \leq \log \text{PDT size}(FC_{\ell_e})$$

Can embed DISJ on $n^{1/3}$ bits in F. Overkill???

[Beame-Pitassi-Segerlind '07] : $\text{PDT size}(FC_f) \geq 2^{n^{1/3}}$.

Our Contributions:

- A notion of stifling that helps negate the power of parity queries

- A PDT-to-DT simulation theorem with const. size gadgets.

∴ For every DT lower bound, a related PDT lower bound.

Parity Queries: Not that global

Consider $x \in \{0,1\}^n$ (no constraints)

Impose $x_1 \oplus x_2 \oplus x_3 = 1$.

Equivalently, $x_1 = x_2 \oplus x_3 \oplus 1$.

Constraint only on x_1.

x_2, x_3 unconstrained.

$$x_1 \oplus x_3 \oplus x_5 = 0$$
$$\Updownarrow$$
$$x_2 \oplus x_5 = 1$$

Set of solns $\equiv \{0,1\}^{n-1}$, new constraints are over $n-1$ bits.

Parity Query Lower Bound

$MAJ : \{0,1\}^7 \to \{0,1\}$

Suppose three parity queries were made.

Affects 3 bits. 4 bits still free

Output can be 0.

Output can be 1.

$$PDT(MAJ_7) > 3$$

k-stifled function

$f : \{0,1\}^n \to \{0,1\}$ is k-stifled if

$\forall$ $S \subseteq [n]$, $|S| = k$

$\exists$ partial assignments x^0, x^1 to $[n] \setminus S$

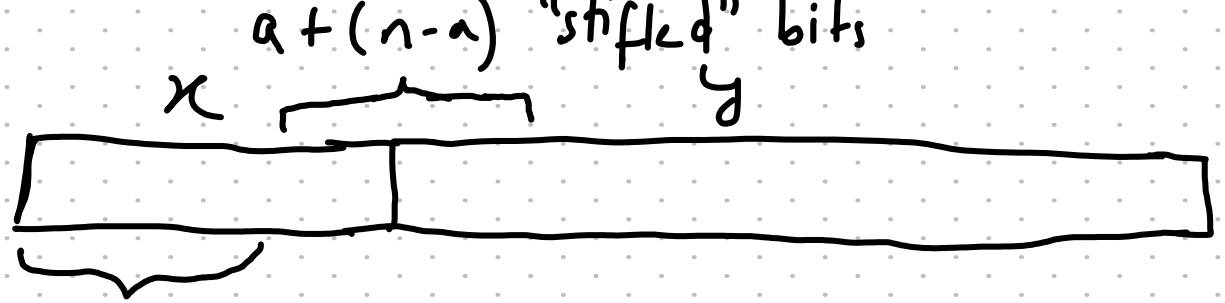
$f(x^0) = 0$, $f(x^1) = 1$

(for every completion)

MAJ_{2n+1} is n-stifled $\quad PDT(f) > k$

$IND: \{0,1\}^n \times \{0,1\}^{2^n} \to \{0,1\}$

$IND(x, y) = y_x$



At most $n-a$ of them potentially point to a stifled bit.

Fix a free setting.

Set all corresponding bits in y to 0/1.

Our simulation theorem.

Let f be a k-stifled function.

For any relation R,

- $PDT(R \circ f) \geq k \cdot DT(R)$
- $PDT_{size}(R \circ f) \geq 2^{k \cdot DT(R)} \geq DT_{size}(R)^k$
- $SubspaceDT(R \circ f) \geq k \cdot DT(R)$
- $NAPDT(R \circ f) \geq k \cdot NADT(R)$

$R = FC_{Tseitin}$ $\quad f: \{0,1\}^t \to \{0,1\}$

R variables:

x_e for $e \in G$

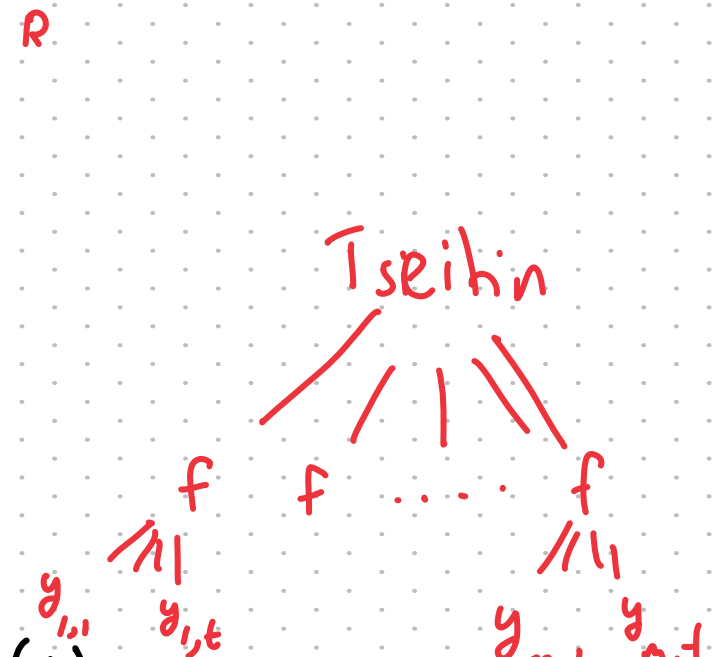
$R \circ f$ variables:

$y_{e,1} \cdots y_{e,t}$ for $e \in G$

$x_e := f(y_{e,1} \cdots y_{e,t})$

Num of clauses: $\leq 2^{t \cdot width(R)} \cdot m$

Width of clauses: $\leq t \cdot width(R)$



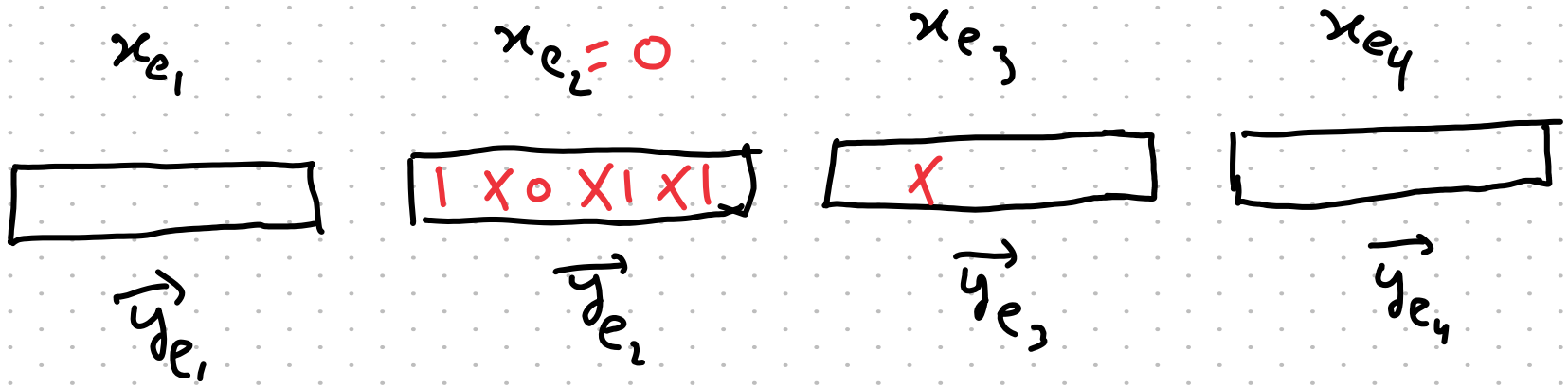
Finding falsified clause for $\vec{y}$ gives falsified clause for $\vec{x}$

Proving query lower bounds: The adversary

Querier-Adversary Game for R

- Querier queries, adversary answers.
- Querier wants to output i s.t.
  $\forall x$ consistent with queries answered, $(x, i) \in R$
- $DT(R) \geq k \iff$ Adversary can force querier to make k queries.

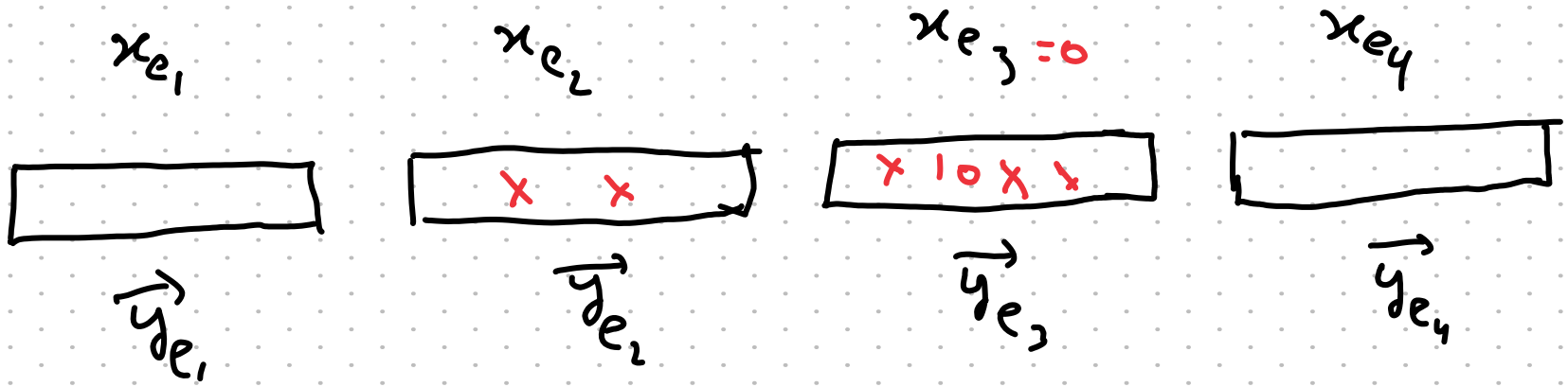
Being an adversary for R∘f:



- When a new parity query is made, answer it arbitrarily and mark a bit as affected.
- When a block, say $\vec{y}_{e_i}$, has k bits affected, query x_{e_i}. Set free bits in $\vec{y}_{e_i}$ to guarantee x_{e_i}.

  Simulate querier-adversary game for R.

Being an adversary for R∘f:



If $< k \cdot DT(R)$ parity queries made,

$< DT(R)$ queries made to R-adversary.

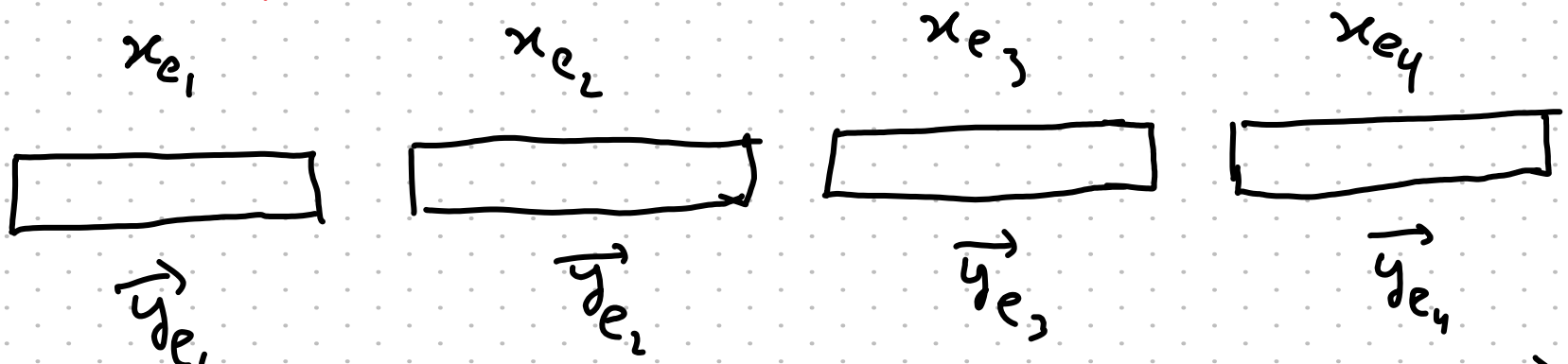
Each unqueried x_e can be independently set, b/c $\leq k$ bits affected in each $\vec{y}_e$.

$\therefore$, we cannot answer R.

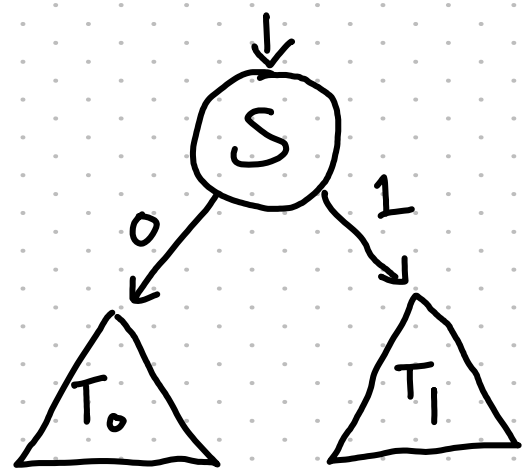
$\therefore$, we cannot answer $R \circ f$.

$$D_{DT}(R \circ f) \geq k \cdot DT(R)$$

Being an adversary for Rof (size edition):



- When a new parity query is made, answer it arbitrarily (→ go to smaller subtree.) and mark a bit as affected.



Suppose querier uses opt. sized PDT.

Forces answer in $\lfloor \log \text{PDTsize}(\text{Rof}) \rfloor$ queries.

$\log \text{PDTsize}(\text{Rof}) > k \cdot DT(f)$

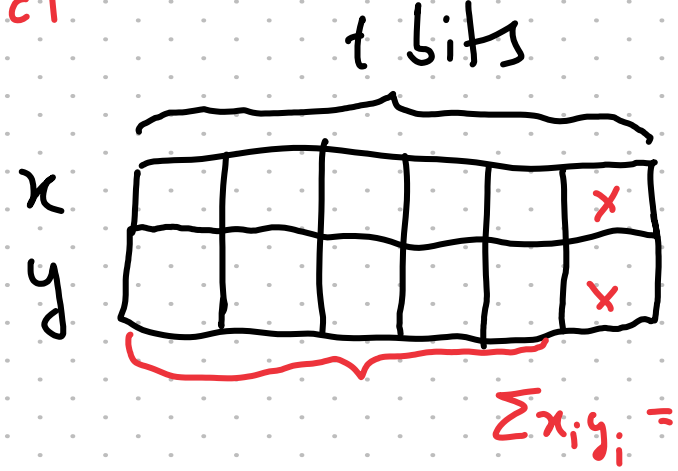
dude this sounds like exactly the sort of stuff we were doing no

On Disperser/Lifting Properties of the Index and Inner-Product Functions

- Paul Beame and Sajin Koroth

They prove that $\text{PDTsize}(R \circ IND_3) \geq 2^{DT(R)}$

Inner-Product



$\sum x_i y_i \bmod 2 = ?$

Is 1-stifled.

Is not 2-stifled.

PDT(R ∘ IP) ought to be $\geq t \cdot DT(R)$.

What property of $\oplus$ can we exploit?

Other Open Problems:

- Can we prove MOD$_p$DT lower bounds?

  PHP requires large MOD$_p$DTs. [Part Tzameret '19]

- Can we prove Parity DAG size/width lower bounds?

  Open/closed

Itsykson-Sokolov PHP lower bound

$G_n \subseteq \{0,1\}^{k \times (k+1)}$ = "assignments where no hole has >1 pigeon"

Let T solve FC_{PHP} w/ the promise that input is in G_n.

(Leaves of T labeled with "pigeon clauses".)

Let ℓ be a leaf of depth $< k/2$. Defined by $< k/2$ parity constraints.

Let $S_\ell \subseteq \{0,1\}^n$ be the subspace corr. to ℓ.

Itsykson-Sokolov PHP lower bound

· Let ℓ be a leaf of depth $< k/2$. Defined by $< k/2$ parity constraints.

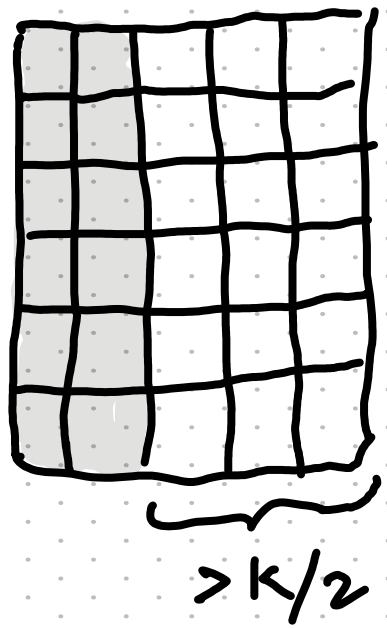
Let $x \in G \cap S_\ell$. If $|x| \geq \frac{k}{2}$, can find $x' \subset x$, $x' \in S_\ell$ (also $x \in G$)

$\exists\, x' \in G \cap S_\ell$ with $|x'| < \frac{k}{2}$.

So $> k/2$ empty holes.

$\forall i \quad \exists x'' \in G \cap S_\ell$ with

Pigeon i in one of these holes.



$\therefore$ PDT size $\geq 2^{\lceil \frac{k}{2} \rceil}$

Part - Tzamaret '19

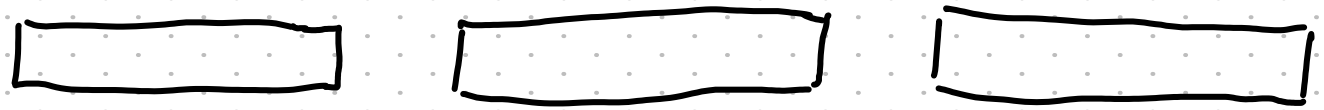
Even with MOD_p queries, if $\exists$ a Boolean assignment satisfying K constraints, then for any $I \subseteq [n]$, $|I| > k$, $\exists$ another Boolean assignment differing only in I.

$\therefore$ as an adversary:

need to ensure a Boolean assignment reaches a leaf.

or G_n = "Boolean assignments where no hole has >1 pigeon"

Our adversary:



The "marked" bit depends on the free bits

The "free" bits are set in order to fix the output value.

The "marked" bit might not be Boolean!