

Average case lower bounds for threshold circuits

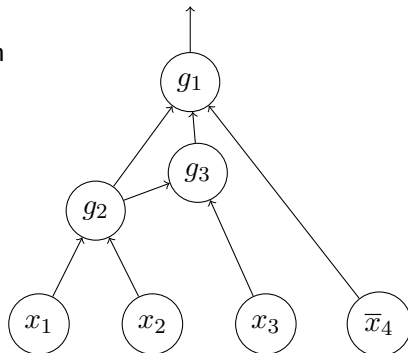
Ruiwen Chen, Rahul Santhanam and Srikanth Srinivasan

Oxford University and Department of Mathematics, IIT Bombay

Low-depth complexity workshop,
St. Petersburg
May 25, 2016.

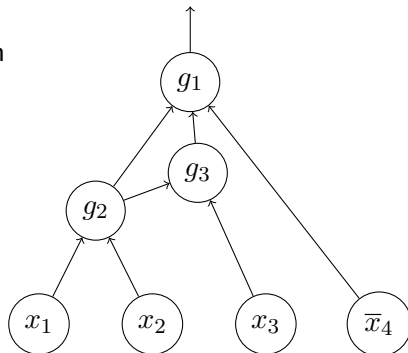
Boolean Circuits

- Circuit computing function $f : \{0, 1\}^n \rightarrow \{0, 1\}$.
- Computation proceeds through “simple” operations.



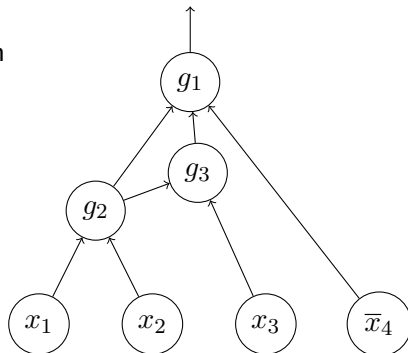
Boolean Circuits

- Circuit computing function $f : \{0, 1\}^n \rightarrow \{0, 1\}$.
- Computation proceeds through “simple” operations.
- $g_i \in$ “basic” operations.



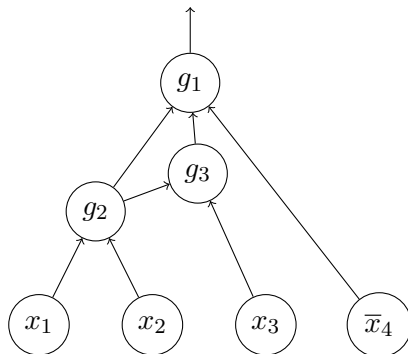
Boolean Circuits

- Circuit computing function $f : \{0, 1\}^n \rightarrow \{0, 1\}$.
- Computation proceeds through “simple” operations.
- $g_i \in$ “basic” operations.
- Designated output gate computes function f .



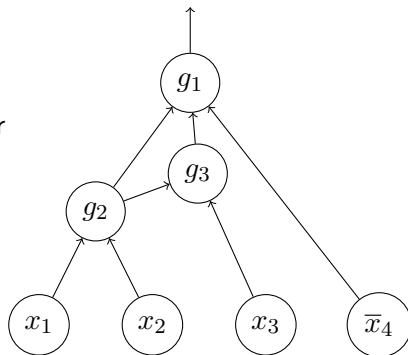
Boolean Circuits

- Size s of the circuit: time taken by algorithm.



Boolean Circuits

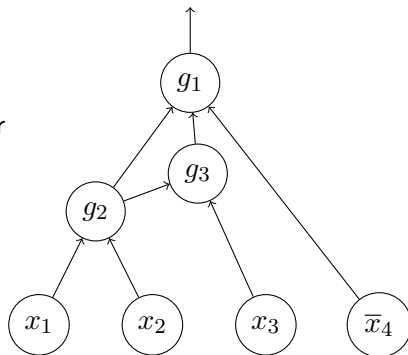
- Size s of the circuit: time taken by algorithm.
- Could be # edges/wires or # gates.



wires = 8, # gates = 3

Boolean Circuits

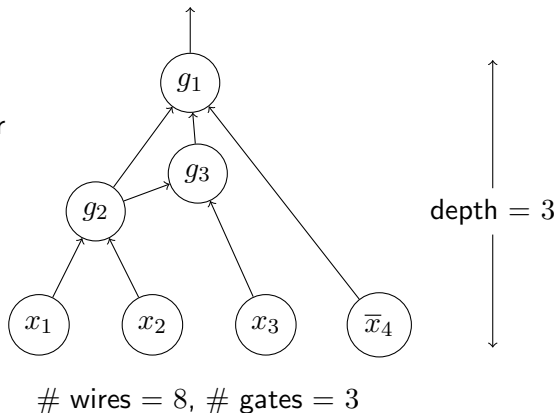
- Size s of the circuit: time taken by algorithm.
- Could be # edges/wires or # gates.
- # wires $\leq (n + \# \text{ gates}) \cdot \# \text{ gates}$.



wires = 8, # gates = 3

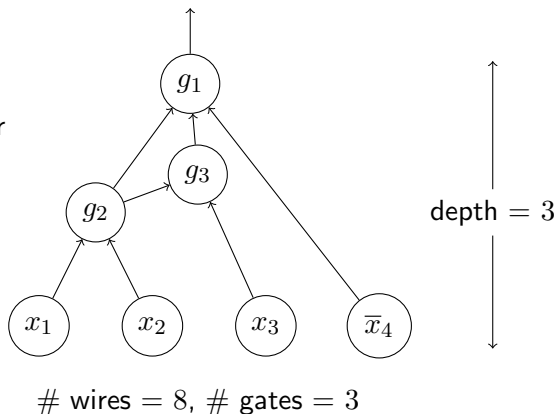
Boolean Circuits

- Size s of the circuit: time taken by algorithm.
- Could be # edges/wires or # gates.
- # wires $\leq (n + \# \text{ gates}) \cdot \# \text{ gates}$.
- Depth d of the circuit: parallelism of the algorithm.



Boolean Circuits

- Size s of the circuit: time taken by algorithm.
- Could be # edges/wires or # gates.
- # wires $\leq (n + \# \text{ gates}) \cdot \# \text{ gates}$.
- Depth d of the circuit: parallelism of the algorithm.
- $s = s(n)$, $d = O(1)$.



Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$.

Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$. Some examples:
- OR function: $\text{OR}(x) = 1$ iff $\sum_i x_i \geq 1$.

Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$. Some examples:
- OR function: $\text{OR}(x) = 1$ iff $\sum_i x_i \geq 1$.
- AND function: $\text{AND}(x) = \llbracket \sum_i x_i \geq n \rrbracket$.

Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$. Some examples:
- OR function: $\text{OR}(x) = 1$ iff $\sum_i x_i \geq 1$.
- AND function: $\text{AND}(x) = \llbracket \sum_i x_i \geq n \rrbracket$.
- MAJ function: $\text{MAJ}(x) = \llbracket \sum_i x_i \geq n/2 \rrbracket$.

Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$. Some examples:
- OR function: $\text{OR}(x) = 1$ iff $\sum_i x_i \geq 1$.
- AND function: $\text{AND}(x) = \llbracket \sum_i x_i \geq n \rrbracket$.
- MAJ function: $\text{MAJ}(x) = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- GEQ function: $\text{GEQ}(x, y) = \llbracket \sum_i 2^i (x_i - y_i) \geq 0 \rrbracket$.

Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$. Some examples:
- OR function: $\text{OR}(x) = 1$ iff $\sum_i x_i \geq 1$.
- AND function: $\text{AND}(x) = \llbracket \sum_i x_i \geq n \rrbracket$.
- MAJ function: $\text{MAJ}(x) = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- GEQ function: $\text{GEQ}(x, y) = \llbracket \sum_i 2^i (x_i - y_i) \geq 0 \rrbracket$.
- $\text{TC}_g^0(s, d)$: threshold circuits with s gates and depth d .

Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$. Some examples:
- OR function: $\text{OR}(x) = 1$ iff $\sum_i x_i \geq 1$.
- AND function: $\text{AND}(x) = \llbracket \sum_i x_i \geq n \rrbracket$.
- MAJ function: $\text{MAJ}(x) = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- GEQ function: $\text{GEQ}(x, y) = \llbracket \sum_i 2^i (x_i - y_i) \geq 0 \rrbracket$.
- $\text{TC}_g^0(s, d)$: threshold circuits with s gates and depth d .
- $\text{TC}_w^0(s, d)$: threshold circuits with s wires and depth d .

Threshold circuits

- A threshold operation: $g(x) = 1$ iff $\sum_i w_i x_i \geq \theta$ for $w_i, \theta \in \mathbb{R}$. Some examples:
- OR function: $\text{OR}(x) = 1$ iff $\sum_i x_i \geq 1$.
- AND function: $\text{AND}(x) = \llbracket \sum_i x_i \geq n \rrbracket$.
- MAJ function: $\text{MAJ}(x) = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- GEQ function: $\text{GEQ}(x, y) = \llbracket \sum_i 2^i (x_i - y_i) \geq 0 \rrbracket$.
- $\text{TC}_g^0(s, d)$: threshold circuits with s gates and depth d .
- $\text{TC}_w^0(s, d)$: threshold circuits with s wires and depth d .
- Generalize AC^0 circuits made up of AND and OR gates.

The power of threshold circuits

- $f = \text{PARITY}(x_1, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n.$

The power of threshold circuits

- $f = \text{PARITY}(x_1, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$.
- $d \geq 2$: $f \in \text{TC}_g^0(dn^{1/(d-1)}, d)$ (Siu-Roychowdhury-Kailath 1991)

The power of threshold circuits

- $f = \text{PARITY}(x_1, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$.
- $d \geq 2$: $f \in \text{TC}_g^0(dn^{1/(d-1)}, d)$ (Siu-Roychowdhury-Kailath 1991)
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket$

The power of threshold circuits

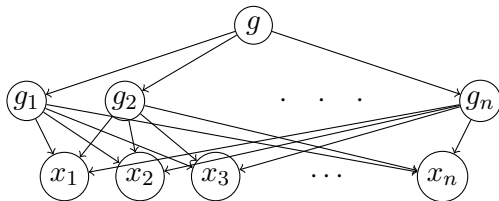
- $f = \text{PARITY}(x_1, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$.
- $d \geq 2$: $f \in \text{TC}_g^0(dn^{1/(d-1)}, d)$ (Siu-Roychowdhury-Kailath 1991)
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket$

The power of threshold circuits

- $f = \text{PARITY}(x_1, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$.
- $d \geq 2$: $f \in \text{TC}_g^0(dn^{1/(d-1)}, d)$ (Siu-Roychowdhury-Kailath 1991)
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \dots$

The power of threshold circuits

- $f = \text{PARITY}(x_1, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$.
- $d \geq 2$: $f \in \text{TC}_g^0(dn^{1/(d-1)}, d)$ (Siu-Roychowdhury-Kailath 1991)
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \dots$



The power of threshold circuits

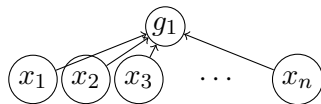
- $f = \text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \cdots$.

The power of threshold circuits

- $f = \text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \cdots$.
- $d = \log n$.

The power of threshold circuits

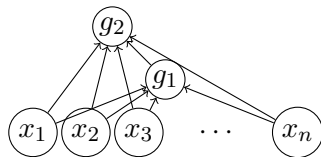
- $f = \text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
 - $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \cdots$.
 - $d = \log n$.
-
- $g_1 = \llbracket \sum_i x_i \geq n/2 \rrbracket$.



The power of threshold circuits

- $f = \text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \cdots$.
- $d = \log n$.

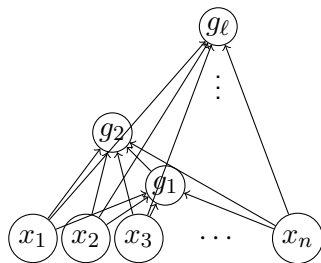
- $g_1 = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- $g_2 = \llbracket \sum_i x_i - \frac{n}{2}g_1 \geq n/4 \rrbracket$.



The power of threshold circuits

- $f = \text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \cdots$.
- $d = \log n$.

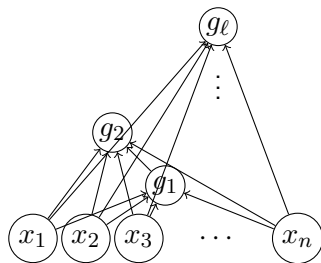
- $g_1 = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- $g_2 = \llbracket \sum_i x_i - \frac{n}{2}g_1 \geq n/4 \rrbracket$.
- \vdots
- $g_\ell = \llbracket \sum_i x_i - \frac{n}{2}g_1 - \frac{n}{4}g_2 \cdots \geq 1 \rrbracket$



The power of threshold circuits

- $f = \text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- $d = 2$: $f(x) = \llbracket \sum_i x_i \geq 1 \rrbracket - \llbracket \sum_i x_i \geq 2 \rrbracket + \llbracket \sum_i x_i \geq 3 \rrbracket \cdots$.
- $d = \log n$.

- $g_1 = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- $g_2 = \llbracket \sum_i x_i - \frac{n}{2}g_1 \geq n/4 \rrbracket$.
- \vdots
- $g_\ell = \llbracket \sum_i x_i - \frac{n}{2}g_1 - \frac{n}{4}g_2 \cdots \geq 1 \rrbracket$



- General d : interpolate between the above two strategies.

The power of threshold circuits

- $\text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.

The power of threshold circuits

- $\text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- Wire complexity of PARITY?

The power of threshold circuits

- $\text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- Wire complexity of PARITY?
- Gate upper bound $\Rightarrow \text{PARITY} \in \text{TC}_w^0(dn^{1+1/(d-1)}, d)$.

The power of threshold circuits

- $\text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- Wire complexity of PARITY?
- Gate upper bound $\Rightarrow \text{PARITY} \in \text{TC}_w^0(dn^{1+1/(d-1)}, d)$.
- Can do better (Beame-Brisson-Ladner, Paturi-Saks).

The power of threshold circuits

- $\text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- Wire complexity of PARITY?
- Gate upper bound $\Rightarrow \text{PARITY} \in \text{TC}_w^0(dn^{1+1/(d-1)}, d)$.
- Can do better (Beame-Brisson-Ladner, Paturi-Saks).
- $\text{PARITY} \in \text{TC}_w^0(n^{1+\varepsilon^d}, d)$.

The power of threshold circuits

- $\text{PARITY} \in \text{TC}_g^0(dn^{1/(d-1)}, d)$.
- Wire complexity of PARITY?
- Gate upper bound $\Rightarrow \text{PARITY} \in \text{TC}_w^0(dn^{1+1/(d-1)}, d)$.
- Can do better (Beame-Brisson-Ladner, Paturi-Saks).
- $\text{PARITY} \in \text{TC}_w^0(n^{1+\varepsilon^d}, d)$.
- Trick:

$$x_1 \oplus \cdots \oplus x_n = (x_1 \oplus \cdots \oplus x_m) \oplus (x_{m+1} \oplus \cdots \oplus x_{2m}) \oplus \cdots$$

Circuit lower bounds

- How efficient are threshold circuits for computing general functions?

Circuit lower bounds

- How efficient are threshold circuits for computing general functions?
- Depth-2 circuits can compute every function with size $\approx 2^n$.

Circuit lower bounds

- How efficient are threshold circuits for computing general functions?
- Depth-2 circuits can compute every function with size $\approx 2^n$.
- Random $f : \{0, 1\}^n \rightarrow \{0, 1\}$ has no threshold circuit with fewer than $2^{0.99n}$ wires w.h.p..

Circuit lower bounds

- How efficient are threshold circuits for computing general functions?
- Depth-2 circuits can compute every function with size $\approx 2^n$.
- Random $f : \{0, 1\}^n \rightarrow \{0, 1\}$ has no threshold circuit with fewer than $2^{0.99n}$ wires w.h.p..
- Problem: Find explicit family of functions (say in NP) that have no TC^0 circuits of $\text{poly}(n)$ size.

Circuit lower bounds

- How efficient are threshold circuits for computing general functions?
- Depth-2 circuits can compute every function with size $\approx 2^n$.
- Random $f : \{0, 1\}^n \rightarrow \{0, 1\}$ has no threshold circuit with fewer than $2^{0.99n}$ wires w.h.p..
- Problem: Find explicit family of functions (say in NP) that have no TC^0 circuits of $\text{poly}(n)$ size. Even open for depth 2.

Work on threshold circuits

- Hajnal Maass Pudlák Turan Szegedy 1987
- (Polynomial Approximations) Paturi Saks 1991, Siu Roychowdhury Kailath 1992; Beigel 1994; Aspnes, Beigel, Furst and Rudich 1994; Podolskii 2012
- (Combinatorial restrictions) Impagliazzo Paturi Saks 1991
- (Communication complexity) Nisan 1992; Hansen and Miltersen 2004; Chattopadhyay and Hansen 2005; Lovett, S. 2012
- (Analytic techniques) Gopalan and Servedio 2010

State-of-the-art lower bounds

- (Furst-Saxe-Sipser, Ajtai, Yao, Håstad 80s) PARITY not in $AC^0(2^{n^{1/(d-1)}}, d)$.

State-of-the-art lower bounds

- (Furst-Saxe-Sipser, Ajtai, Yao, Håstad 80s) PARITY not in $AC^0(2^{n^{1/(d-1)}}, d)$.
- (Impagliazzo-Paturi-Saks 1991) PARITY not in $TC_g^0(n^{1/2(d-1)}, d)$ and $TC_w^0(n^{1+\varepsilon^d}, d)$.

State-of-the-art lower bounds

- (Furst-Saxe-Sipser, Ajtai, Yao, Håstad 80s) PARITY not in $AC^0(2^{n^{1/(d-1)}}, d)$.
- (Impagliazzo-Paturi-Saks 1991) PARITY not in $TC_g^0(n^{1/2(d-1)}, d)$ and $TC_w^0(n^{1+\varepsilon^d}, d)$.
- (Kane-Williams 2015) Explicit functions not in $TC_g^0(n^{1.5-o(1)}, 2)$ and $TC_w^0(n^{2.5-o(1)}, 2)$.

State-of-the-art lower bounds

- (Furst-Saxe-Sipser, Ajtai, Yao, Håstad 80s) PARITY not in $AC^0(2^{n^{1/(d-1)}}, d)$.
- (Impagliazzo-Paturi-Saks 1991) PARITY not in $TC_g^0(n^{1/2(d-1)}, d)$ and $TC_w^0(n^{1+\varepsilon^d}, d)$.
- (Kane-Williams 2015) Explicit functions not in $TC_g^0(n^{1.5-o(1)}, 2)$ and $TC_w^0(n^{2.5-o(1)}, 2)$. Also extends to a special case of depth-3.

Average case lower bounds

- Want to show a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ hard on *average*.

Average case lower bounds

- Want to show a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ hard on *average*.
- Trivial to compute f on half the inputs.

Average case lower bounds

- Want to show a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ hard on *average*.
- Trivial to compute f on half the inputs.
- f has ε -correlation with ckt C if

$$\text{Corr}(C, f) := \Pr_x[C(x) = f(x)] - \frac{1}{2} \leq \varepsilon.$$

Average case lower bounds

- Want to show a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ hard on *average*.
- Trivial to compute f on half the inputs.
- f has ε -correlation with ckt C if

$$\text{Corr}(C, f) := \Pr_x[C(x) = f(x)] - \frac{1}{2} \leq \varepsilon.$$

- Want to show that f hard on average against $\text{TC}^0(s, d)$.

Why average case lower bounds

- Improves our understanding of limitations of circuits.

Why average case lower bounds

- Improves our understanding of limitations of circuits.
- Lower bounds against slightly stronger circuit classes

Why average case lower bounds

- Improves our understanding of limitations of circuits.
- Lower bounds against slightly stronger circuit classes (E.g - Kane-Williams 2015).

Why average case lower bounds

- Improves our understanding of limitations of circuits.
- Lower bounds against slightly stronger circuit classes (E.g - Kane-Williams 2015).
- Prerequisite for constructing Pseudorandom generators (PRGs) for the circuit class.

Why average case lower bounds

- Improves our understanding of limitations of circuits.
- Lower bounds against slightly stronger circuit classes (E.g - Kane-Williams 2015).
- Prerequisite for constructing Pseudorandom generators (PRGs) for the circuit class.
- Increased understanding can lead to satisfiability algorithms, learning algorithms,...

Results

- (Impagliazzo-Paturi-Saks 1991) PARITY not in $TC_g^0(n^{1/2(d-1)}, d)$ and $TC_w^0(n^{1+\varepsilon^d}, d)$.

Results

- (Impagliazzo-Paturi-Saks 1991) PARITY not in $\text{TC}_g^0(n^{1/2(d-1)}, d)$ and $\text{TC}_w^0(n^{1+\varepsilon^d}, d)$.
- Result 1: PARITY has $o(1)$ -correlation with $\text{TC}_g^0(o(n^{1/2(d-1)}), d)$ and $\text{TC}_w^0(n^{1+\delta^d}, d)$.

Results

- (Impagliazzo-Paturi-Saks 1991) PARITY not in $\text{TC}_g^0(n^{1/2(d-1)}, d)$ and $\text{TC}_w^0(n^{1+\varepsilon^d}, d)$.
- Result 1: PARITY has $o(1)$ -correlation with $\text{TC}_g^0(o(n^{1/2(d-1)}), d)$ and $\text{TC}_w^0(n^{1+\delta^d}, d)$.
- Gates result weaker than Nisan (1992) if any explicit function allowed.

Results

- (Impagliazzo-Paturi-Saks 1991) PARITY not in $\text{TC}_g^0(n^{1/2(d-1)}, d)$ and $\text{TC}_w^0(n^{1+\varepsilon^d}, d)$.
- Result 1: PARITY has $o(1)$ -correlation with $\text{TC}_g^0(o(n^{1/2(d-1)}), d)$ and $\text{TC}_w^0(n^{1+\delta^d}, d)$.
- Gates result weaker than Nisan (1992) if any explicit function allowed.
- Result 2: Different explicit function has exponentially small correlation with $\text{TC}_w^0(n^{1+\delta^d}, d)$.

Random restrictions

- Restriction: setting variables to constants. Helps simplify circuit.

Random restrictions

- Restriction: setting variables to constants. Helps simplify circuit.
 $\rho : \{x_1, \dots, x_n\} \rightarrow \{0, 1, *\}$.

Random restrictions

- Restriction: setting variables to constants. Helps simplify circuit.
 $\rho : \{x_1, \dots, x_n\} \rightarrow \{0, 1, *\}$.
- Random restriction $\rho \sim \mathcal{R}_p$:

$$\Pr_{\rho}[\rho(x_i) = *] = p \qquad \Pr_{\rho}[\rho(x_i) = 0/1] = \frac{1-p}{2}$$

Random restrictions and lower bounds

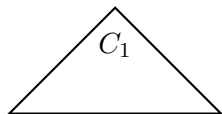
- Role of Random restriction: simplify circuit, while leaving hard function (rel.) unchanged.

Random restrictions and lower bounds

- Role of Random restriction: simplify circuit, while leaving hard function (rel.) unchanged.
- Eg. with AC^0 and PARITY.

Random restrictions and lower bounds

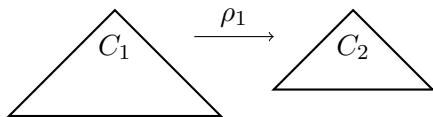
- Role of Random restriction: simplify circuit, while leaving hard function (rel.) unchanged.
- Eg. with AC^0 and PARITY.



$$f(x) = x_1 \oplus \cdots \oplus x_n$$

Random restrictions and lower bounds

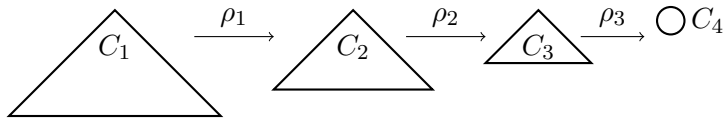
- Role of Random restriction: simplify circuit, while leaving hard function (rel.) unchanged.
- Eg. with AC^0 and PARITY.



$$f(x) = x_1 \oplus \cdots \oplus x_n \quad f|_{\rho_1}(x) = \bigoplus_{i \in \rho_1^{-1}(*)} x_i$$

Random restrictions and lower bounds

- Role of Random restriction: simplify circuit, while leaving hard function (rel.) unchanged.
- Eg. with AC^0 and PARITY.



$$f(x) = x_1 \oplus \cdots \oplus x_n \quad f|_{\rho_1}(x) = \bigoplus_{i \in \rho_1^{-1}(*)} x_i \quad f|_{\rho_1 \rho_2}(x) \quad f|_{\rho_1 \rho_2 \rho_3}(x)$$

Key lemma

- How does the circuit simplify due to a random restriction?

Key lemma

- How does the circuit simplify due to a random restriction?
- Håstad's Switching lemma: If C is a small CNF (resp. DNF), then whp $C|_\rho$ is a small DNF (resp. CNF).

Key lemma

- How does the circuit simplify due to a random restriction?
- Håstad's Switching lemma: If C is a small CNF (resp. DNF), then whp $C|_\rho$ is a small DNF (resp. CNF).
- For threshold circuits: Peres' theorem.

Peres' theorem

- Informal: if f a threshold function and $\rho \sim \mathcal{R}_p$ (small p), then $f|_\rho$ is close to constant whp.

Peres' theorem

- Informal: if f a threshold function and $\rho \sim \mathcal{R}_p$ (small p), then $f|_\rho$ is close to constant whp.
- Measure bias using $\text{Var}(g) = 2 \Pr_{x,y}[g(x) \neq g(y)] \in [0, 1]$.

Peres' theorem

- Informal: if f a threshold function and $\rho \sim \mathcal{R}_p$ (small p), then $f|_\rho$ is close to constant whp.
- Measure bias using $\text{Var}(g) = 2 \Pr_{x,y}[g(x) \neq g(y)] \in [0, 1]$.
- g unbiased $\Leftrightarrow \text{Var}(g) = 1$.

Peres' theorem

- Informal: if f a threshold function and $\rho \sim \mathcal{R}_p$ (small p), then $f|_\rho$ is close to constant whp.
- Measure bias using $\text{Var}(g) = 2 \Pr_{x,y}[g(x) \neq g(y)] \in [0, 1]$.
- g unbiased $\Leftrightarrow \text{Var}(g) = 1$. g constant $\Leftrightarrow \text{Var}(g) = 0$.

Peres' theorem

- Informal: if f a threshold function and $\rho \sim \mathcal{R}_p$ (small p), then $f|_\rho$ is close to constant whp.
- Measure bias using $\text{Var}(g) = 2 \Pr_{x,y}[g(x) \neq g(y)] \in [0, 1]$.
- g unbiased $\Leftrightarrow \text{Var}(g) = 1$. g constant $\Leftrightarrow \text{Var}(g) = 0$.

Theorem (Peres 2003)

f a threshold function. $\mathbf{E}_\rho[\text{Var}(f|_\rho)] = O(\sqrt{p})$.

Peres' theorem

- Informal: if f a threshold function and $\rho \sim \mathcal{R}_p$ (small p), then $f|_\rho$ is close to constant whp.
- Measure bias using $\text{Var}(g) = 2 \Pr_{x,y}[g(x) \neq g(y)] \in [0, 1]$.
- g unbiased $\Leftrightarrow \text{Var}(g) = 1$. g constant $\Leftrightarrow \text{Var}(g) = 0$.

Theorem (Peres 2003)

f a threshold function. $\mathbf{E}_\rho[\text{Var}(f|_\rho)] = O(\sqrt{p})$.

- Compare with PARITY: $\mathbf{E}_\rho[\text{Var}(\text{PARITY}|_\rho)] \approx 1$ unless $p \approx 1/n$.

Peres' theorem

- Informal: if f a threshold function and $\rho \sim \mathcal{R}_p$ (small p), then $f|_\rho$ is close to constant whp.
- Measure bias using $\text{Var}(g) = 2 \Pr_{x,y}[g(x) \neq g(y)] \in [0, 1]$.
- g unbiased $\Leftrightarrow \text{Var}(g) = 1$. g constant $\Leftrightarrow \text{Var}(g) = 0$.

Theorem (Peres 2003)

f a threshold function. $\mathbf{E}_\rho[\text{Var}(f|_\rho)] = O(\sqrt{p})$.

- Compare with PARITY: $\mathbf{E}_\rho[\text{Var}(\text{PARITY}|_\rho)] \approx 1$ unless $p \approx 1/n$.

Corollary

f a threshold function. $\text{Corr}(f, \text{PARITY}) \leq O(\frac{1}{\sqrt{n}})$.

Gate lower bound

Theorem

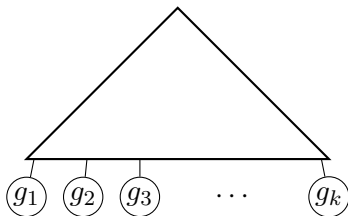
$$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1).$$

Gate lower bound

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Bottom level gates: g_1, \dots, g_k .

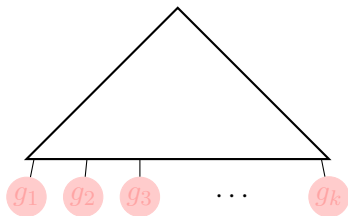


Gate lower bound

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Bottom level gates: g_1, \dots, g_k .
- Apply $\rho \sim \mathcal{R}_p$. Use Peres.

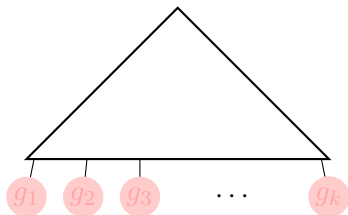


Gate lower bound

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Bottom level gates: g_1, \dots, g_k .
- Apply $\rho \sim \mathcal{R}_p$. Use Peres.
- $\mathbf{E}_\rho[\sum_i \text{Var}(g_i)] \leq O(k\sqrt{p})$.



Gate lower bound

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Bottom level gates: g_1, \dots, g_k .
- Apply $\rho \sim \mathcal{R}_p$. Use Peres.
- $\mathbf{E}_\rho[\sum_i \text{Var}(g_i)] \leq O(k\sqrt{p})$.
- Replace biased gates with constants. Depth is $d - 1$.



Gate lower bound

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Bottom level gates: g_1, \dots, g_k .
- Apply $\rho \sim \mathcal{R}_p$. Use Peres.
- $\mathbf{E}_\rho[\sum_i \text{Var}(g_i)] \leq O(k\sqrt{p})$.
- Replace biased gates with constants. Depth is $d - 1$.
- Continue.



Gate lower bound (contd.)

Theorem

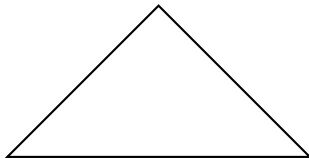
$$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1).$$

Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.

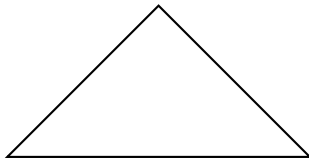


Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Apply $\rho \sim \mathcal{R}_p$ d times.

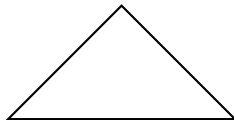


Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Apply $\rho \sim \mathcal{R}_p$ d times.



Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Apply $\rho \sim \mathcal{R}_p$ d times.



Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Apply $\rho \sim \mathcal{R}_p$ d times.



Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Apply $\rho \sim \mathcal{R}_p$ d times.
- I.e. apply $\rho \sim \mathcal{R}_{p^d} = \mathcal{R}_q$.



Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Apply $\rho \sim \mathcal{R}_p$ d times.
- I.e. apply $\rho \sim \mathcal{R}_{p^d} = \mathcal{R}_q$.
- $\mathbf{E}_\rho[\text{Var}(C|\rho)] \leq O(kq^{1/2d})$.



Gate lower bound (contd.)

Theorem

$C \in \text{TC}_g^0(o(n^{1/2(d-1)}), d) \Rightarrow \text{Corr}(C, \text{PARITY}) = o(1)$.

- $C \in \text{TC}_g^0(k, d)$.
- Apply $\rho \sim \mathcal{R}_p$ d times.
- I.e. apply $\rho \sim \mathcal{R}_{p^d} = \mathcal{R}_q$.
- $\mathbf{E}_\rho[\text{Var}(C|\rho)] \leq O(kq^{1/2d})$.
- $\text{Corr}(C, \text{PARITY}) \leq o(1) = o(1)$
if $k \ll n^{1/2d}$.



Wire lower bound

Wire lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

Wire lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Why does the previous proof not work?

Wire lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Why does the previous proof not work?
- Probability of failure in Peres' theorem: $O(1/\sqrt{n})$.

Wire lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Why does the previous proof not work?
- Probability of failure in Peres' theorem: $O(1/\sqrt{n})$.
- Cannot handle more than $O(\sqrt{n})$ gates.

Wire lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

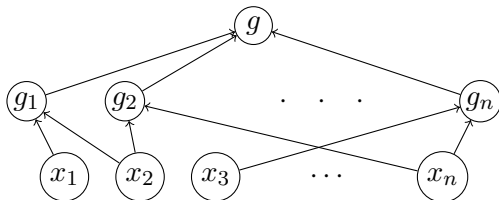
- Why does the previous proof not work?
- Probability of failure in Peres' theorem: $O(1/\sqrt{n})$.
- Cannot handle more than $O(\sqrt{n})$ gates.
- Even if $O(n)$ wires, we could have up to $O(n)$ gates.

Wire lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Why does the previous proof not work?
- Probability of failure in Peres' theorem: $O(1/\sqrt{n})$.
- Cannot handle more than $O(\sqrt{n})$ gates.
- Even if $O(n)$ wires, we could have up to $O(n)$ gates.



Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?

Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.

Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

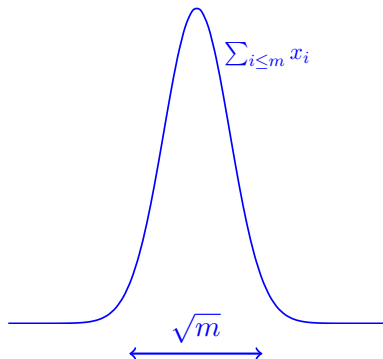
- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.
- $f|_\rho(x) = \mathbb{I}[\sum_{i \leq m} x_i \geq n/2 - \theta]$.

Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.
- $f|_\rho(x) = \mathbb{I}[\sum_{i \leq m} x_i \geq n/2 - \theta]$.

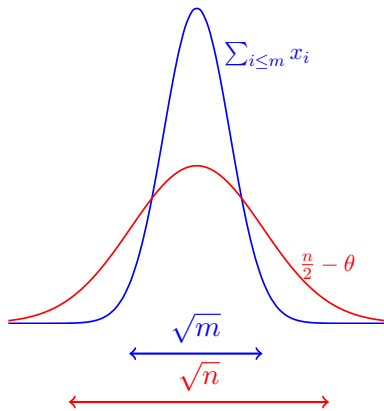


Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.
- $f|_{\rho}(x) = \mathbb{I}[\sum_{i \leq m} x_i \geq n/2 - \theta]$.
- $\theta = \text{sum of } n - m \approx n \text{ random bits.}$

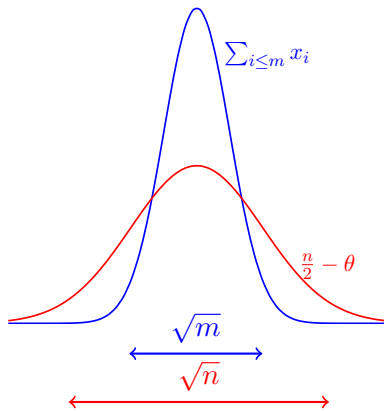


Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.
- $f|_{\rho}(x) = \mathbb{I}[\sum_{i \leq m} x_i \geq n/2 - \theta]$.
- $\theta =$ sum of $n - m \approx n$ random bits.
- $\theta = (n - m)/2$: $\text{Var}(f|_{\rho}) = 1$.

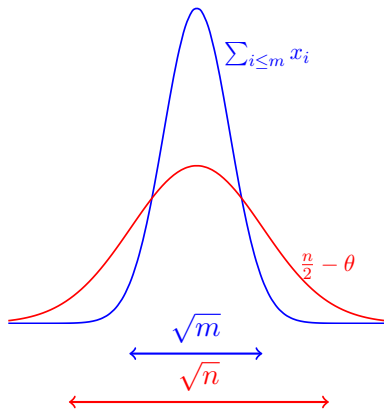


Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.
- $f|_{\rho}(x) = \mathbb{I}[\sum_{i \leq m} x_i \geq n/2 - \theta]$.
- $\theta =$ sum of $n - m \approx n$ random bits.
- $\theta = (n - m)/2$: $\text{Var}(f|_{\rho}) = 1$.
- $\theta = (n - m)/2 \pm O(\sqrt{m})$:
 $\text{Var}(f|_{\rho}) = \Omega(1)$.

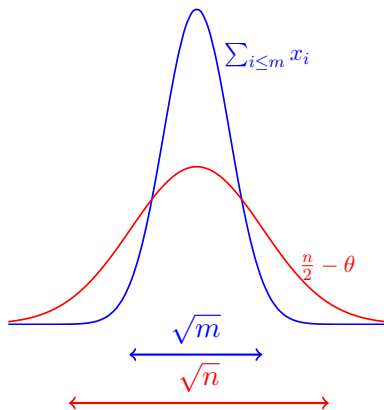


Refining Peres' theorem

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.
- $f|_{\rho}(x) = \mathbb{I}[\sum_{i \leq m} x_i \geq n/2 - \theta]$.
- $\theta =$ sum of $n - m \approx n$ random bits.
- $\theta = (n - m)/2$: $\text{Var}(f|\rho) = 1$.
- $\theta = (n - m)/2 \pm O(\sqrt{m})$:
 $\text{Var}(f|\rho) = \Omega(1)$.
- $\theta = (n - m)/2 \pm \omega(\sqrt{m})$:
 $\text{Var}(f|\rho) = o(1)$.

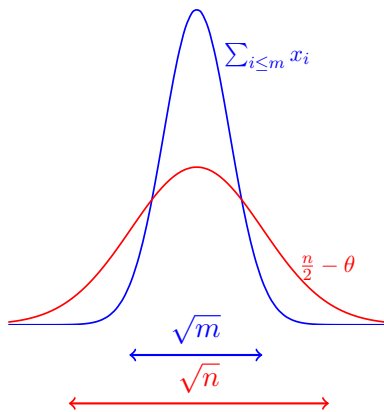


Refining Peres' theorem (contd.)

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \llbracket \sum_i x_i \geq n/2 \rrbracket$.
- $f|_\rho(x) = \llbracket \sum_{i \leq m} x_i \geq n/2 - \theta \rrbracket$.
- $\theta = (n - m)/2 \pm \omega(\sqrt{m})$:
 $\text{Var}(f|_\rho) = o(1)$.

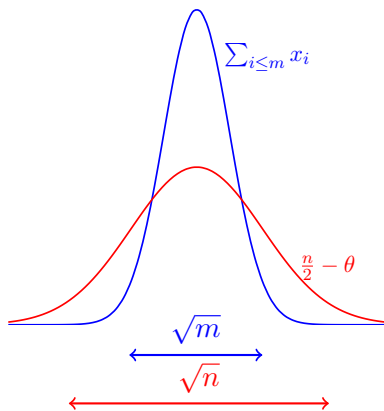


Refining Peres' theorem (contd.)

Theorem (Peres)

f a threshold. $\mathbf{E}_{\rho \sim \mathcal{R}_p}[\text{Var}(f|\rho)] = O(\sqrt{p})$.

- Why $O(\sqrt{p})$?
- Eg: $f(x) = \mathbb{I}[\sum_i x_i \geq n/2]$.
- $f|_{\rho}(x) = \mathbb{I}[\sum_{i \leq m} x_i \geq n/2 - \theta]$.
- $\theta = (n - m)/2 \pm \omega(\sqrt{m})$:
 $\text{Var}(f|_{\rho}) = o(1)$.
- $\mathbf{E}[\text{Var}(f|_{\rho})] = \Omega(1)\sqrt{m}/\sqrt{n} = \Omega(\sqrt{p})$.



Refining Peres' theorem (contd.)

Lemma (Peres extension)

f a threshold. $\Pr_{\rho}[\text{Var}(f|_{\rho}) \text{ noticeable}] \leq p^{0.1}$.

Refining Peres' theorem (contd.)

Lemma (Peres extension)

f a threshold. $\Pr_{\rho}[\text{Var}(f|_{\rho}) \text{ noticeable}] \leq p^{0.1}$.

- $\text{Var}(f)$ not noticeable $\Leftrightarrow \text{Var}(f) = \exp(-n^{\Omega(1)})$.

Refining Peres' theorem (contd.)

Lemma (Peres extension)

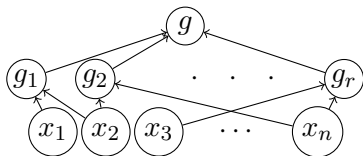
f a threshold. $\Pr_{\rho}[\text{Var}(f|_{\rho}) \text{ noticeable}] \leq p^{0.1}$.

- $\text{Var}(f)$ not noticeable $\Leftrightarrow \text{Var}(f) = \exp(-n^{\Omega(1)})$.
- Proof of lemma via standard CLT + critical index argument.

Back to the wires lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

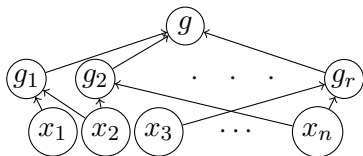


Back to the wires lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- $\sum_i \deg(g_i) \leq n^{1+\alpha}$.

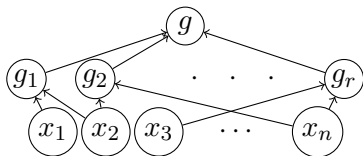


Back to the wires lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- $\sum_i \deg(g_i) \leq n^{1+\alpha}$.
- Apply $\rho \sim \mathcal{R}_p$, $p = n^{-O(\alpha)}$.

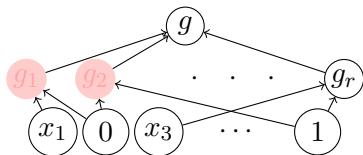


Back to the wires lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- $\sum_i \deg(g_i) \leq n^{1+\alpha}$.
- Apply $\rho \sim \mathcal{R}_p$, $p = n^{-O(\alpha)}$.
- New Peres: $(1 - p^{0.1})$ gates highly biased.

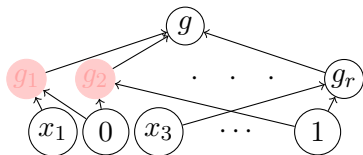


Back to the wires lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- $\sum_i \deg(g_i) \leq n^{1+\alpha}$.
- Apply $\rho \sim \mathcal{R}_p$, $p = n^{-O(\alpha)}$.
- New Peres: $(1 - p^{0.1})$ gates highly biased.
- Set to constants.

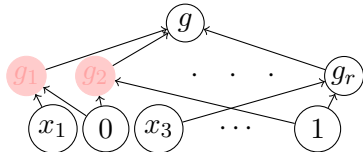


Back to the wires lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- $\sum_i \deg(g_i) \leq n^{1+\alpha}$.
- Apply $\rho \sim \mathcal{R}_p$, $p = n^{-O(\alpha)}$.
- New Peres: $(1 - p^{0.1})$ gates highly biased.
- Set to constants.
- $\sum_{\text{unbiased}} \deg(g_i) \leq p^{1+0.1} \cdot n^{1+\alpha} \ll pn$.

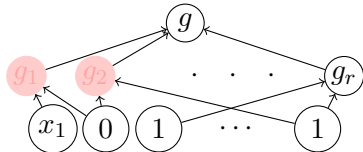


Back to the wires lower bound

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- $\sum_i \deg(g_i) \leq n^{1+\alpha}$.
- Apply $\rho \sim \mathcal{R}_p$, $p = n^{-O(\alpha)}$.
- New Peres: $(1 - p^{0.1})$ gates highly biased.
- Set to constants.
- $\sum_{\text{unbiased}} \deg(g_i) \leq p^{1+0.1} \cdot n^{1+\alpha} \ll pn$.
- Set all vars and continue.



More results

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

More results

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Above $o(1) = n^{-\Omega_d(1)}$.

More results

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Above $o(1) = n^{-\Omega_d(1)}$.
- For a suitable other function f , $\text{Corr}(f, C) \leq \exp(-n^{\Omega_d(1)})$.

More results

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Above $o(1) = n^{-\Omega_d(1)}$.
- For a suitable other function f , $\text{Corr}(f, C) \leq \exp(-n^{\Omega_d(1)})$.
- Satisfiability algorithms for $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$ running in time $2^{n-n^{\Omega_d(1)}}$.

More results

Theorem

For some $\delta > 0$, and $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$, $\text{Corr}(C, \text{PARITY}) = o(1)$.

- Above $o(1) = n^{-\Omega_d(1)}$.
- For a suitable other function f , $\text{Corr}(f, C) \leq \exp(-n^{\Omega_d(1)})$.
- Satisfiability algorithms for $C \in \text{TC}_w^0(n^{1+\delta^d}, d)$ running in time $2^{n-n^{\Omega_d(1)}}$.
- Better learning algorithms for AC^0 augmented with a few threshold gates.

Summary

- Proved correlation bounds for threshold circuits for computing PARITY and other explicit functions.

Summary

- Proved correlation bounds for threshold circuits for computing PARITY and other explicit functions.
- Bounds are close to tight for PARITY.

Summary

- Proved correlation bounds for threshold circuits for computing PARITY and other explicit functions.
- Bounds are close to tight for PARITY.
- Refined version of Peres' theorem gives more insight into the workings of threshold gates.

Summary

- Proved correlation bounds for threshold circuits for computing PARITY and other explicit functions.
- Bounds are close to tight for PARITY.
- Refined version of Peres' theorem gives more insight into the workings of threshold gates.
- More applications?
- Better lower bounds?

Open questions

Open questions

- Optimal worst case depth 2 circuit for PARITY? Upper bound $O(n)$ gates and lower bound $\Omega(\sqrt{n})$.

Open questions

- Optimal worst case depth 2 circuit for PARITY? Upper bound $O(n)$ gates and lower bound $\Omega(\sqrt{n})$.
- Optimal version of Peres' theorem for threshold circuits with few gates? Two versions: ours and Gopalan-Servedio. Want a common generalization.

Open questions

- Optimal worst case depth 2 circuit for PARITY? Upper bound $O(n)$ gates and lower bound $\Omega(\sqrt{n})$.
- Optimal version of Peres' theorem for threshold circuits with few gates? Two versions: ours and Gopalan-Servedio. Want a common generalization.

Thank you