Bootstrapping Results for Threshold Circuits "Just Beyond" Known Lower Bounds

Lijie Chen and Roei Tell

STOC, June 2019

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Long-term goal

- Lower bounds for non-uniform Boolean circuits
- Decades-long efforts, notoriously difficult problem
- Some “dream results”:
  - NP $\not\subset$ P/poly $\Rightarrow$ P $\neq$ NP
  - DTIME[$s^{O(1)}$] $\not\subset$ i.o.SIZE[s] $\Rightarrow$ prBPP = prP [IW’99]

Combinatorial-algebraic approaches

- Restriction method [Ajt'83,FSS'84,Yao'85,Has'86]
- Polynomial approximation method [Raz'87, Smo'87]

  ...

- No "natural proofs" for "strong" circuits [RR'94]
  - circuit class can compute a PRF ⇒ "resistant" to natural proofs

Algorithmic method

- Circuit-analysis algorithm ⇒ lower bounds
  - need "barely non-trivial" deterministic algorithm

    [BFS'98, IKW'01, KI'03, Wil'10, MW'18]
- Breakthrough where combinatorial methods failed
- Widely-believed to be possible for strong circuits

Hardness magnification

- Lower bounds for “weak” circuits

  ⇒ lower bounds for “stronger” circuits

- New(-ish) paradigm, many conditional results

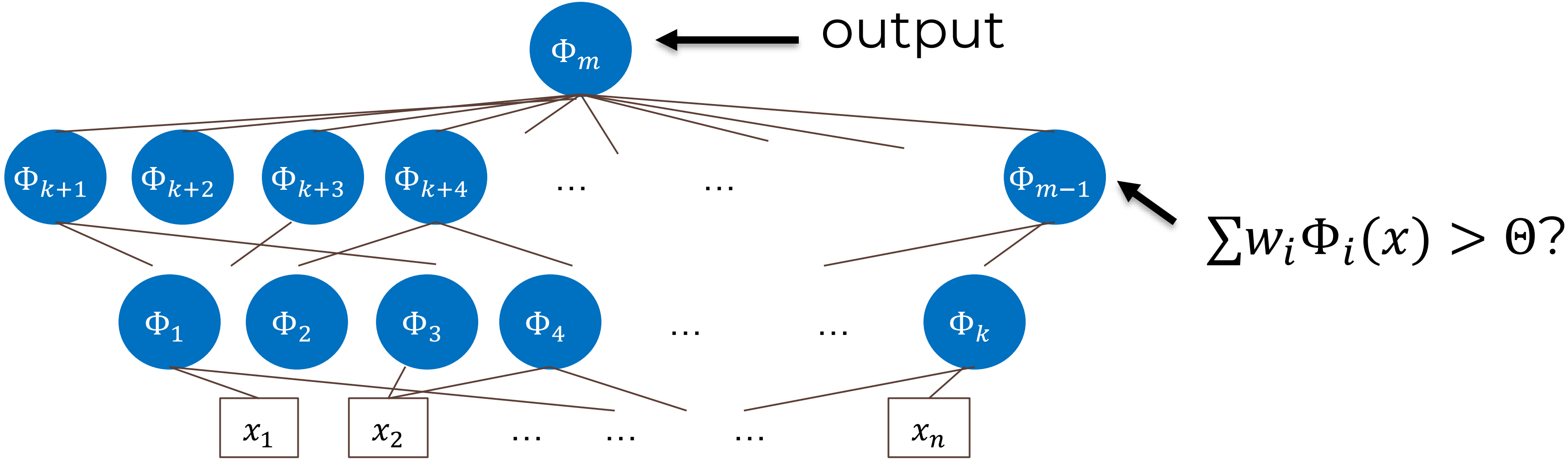
  [Sri’03,AK’10,LW’13,OS’18,CILM’18,MMW’19,OPS’19]

- No known barriers

Linear Threshold Circuits (TC^0): A Prominent Frontier

A prominent frontier: TC^0

- **TC^0:** Constant-depth, poly size, linear threshold gates
  - linear threshold gate: $\Phi(x) = 1$ iff $\sum w_i x_i > \Theta$, for $w \in \mathbb{R}^n, \Theta \in \mathbb{R}$.



A prominent frontier: TC^0

- **TC^0:** Constant-depth, poly size, linear threshold gates
  - linear threshold gate: $\Phi(x) = 1$ iff $\sum w_i x_i > \Theta$, for $w \in \mathbb{R}^n, \Theta \in \mathbb{R}$.
- PRF candidate [NR'97] ⇒ "natural proofs" barrier
- Open problem: **Prove that NEXP $\not\subset$ TC^0**
  - NEXP = $NTIME[2^{poly(n)}]$

Known lower bounds for TC^0

- **Thm [IPS'93]:** TC^0 circuits of depth d need $n^{1+\exp(-d)}$ **wires** to compute the **parity** function
  - extends to average-case lower bounds [CSS'16]
  - better bounds for fixed depth ≤ 3 or "structured subclasses" [KS'15, KW'16, Tam'16, ACW'16, SSTT'16]

1 age(result) > age(Lijie)

Hardness magnification

- the precise size/depth trade-off matters

- **Thm [AK'10]:** If TC^0 circuits of depth d need $\boldsymbol{n^{1+O(1/d)}}$ **wires** to solve certain (NC^1-complete) problems, then $\mathbf{NC^1 \not\subset TC^0}$

- known lower bounds of $n^{1+\exp(-d)}$ wires for these problems

Known circuit-analysis alg for TC^0

- Derandomization: Given a description of a circuit, approximate its acceptance probability up to $\pm 1/6$
- **Quantified derandomization [GW'14]:** Given a circuit $C: \{0,1\}^n \to \{0,1\}$, decide if C **accepts all but B(n) inputs** or **rejects all but B(n) inputs**

Known circuit-analysis alg for TC^0

- **Thm [T'18]:** A deterministic $n^{(\log\log n)^2}$-time alg for quantified derandomization of TC^0 with depth d and $\boldsymbol{n^{1+\exp(-d)}}$ **wires** and $B(n) = 2^{n^{1-\exp(-d)}}$
  - better algorithms for fixed depth ≤ 2 or "structured subclasses" [DGJ+'10, RS'10, GOW+'10, KRS'12, MZ'13, Kan'11, Kan'14, KM'15, KM'15, IPS'13, Wil'14, AS'15, SSTT'16, Tam'16, ACW'16]

Quantified derand implies lower bounds

- the precise size/depth trade-off matters

- **Thm [T'18]:** If there's a deterministic $2^{n^{o(1)}}$-time alg for quantified derandomization of TC^0 with depth d and $\boldsymbol{n^{1+O(1/d)}}$ **wires** and $B(n) = 2^{n^{1-1/d}}$, then **NEXP ⊄ TC^0**

- quantified derand ⇒ standard derand ⇒ lower bounds

- known derand for $n^{1+\exp(-d)}$ wires is faster & handles larger $B(n)$

The state of knowledge at STOC’18

- for depth-d TC circuits

#wires	lower bounds	derandomization
poly(n)		
$n^{1+O(1/d)}$	**specific bounds can be “amplified”** [AK’10]	**quant derand implies NEXP $\not\subset$ TC⁰** [T’18]
$n^{1+\exp(-d)}$	**unconditional lower bounds**[IPS’93,CSS’16]	**unconditional quantified derandomization** [T’18]

Our results

The high-level message

- Improved hardness magnification and "quantified derandomization implies lower bounds" for TC^0
- Both **kick in at $n^{1+\alpha^{-d}}$ wires**, "just beyond" known unconditional results at $n^{1+\beta^{-d}}$ $(\beta > \alpha > 1)$
- **Gap between "known" and "breakthrough" boils down to precise $\alpha > 1$ in the size bound $n^{1+\alpha^{-d}}$**

Improved hardness magnification

- hardness magnification at $n^{1+\exp(-d)}$ wires
- **Thm 1:** If $\forall\alpha > 1$ and sufficiently large d, TC^0 of depth d require $\boldsymbol{n^{1+\alpha^{-d}}}$ **wires** to solve certain (NC^1-complete) problems, then $\mathbf{NC^1 \not\subset TC^0}$
  - we know lower bounds for $n^{1+\beta^{-d}}$ wires, where $\boldsymbol{\beta \approx 2.41}$
  - for breakthrough results we need $n^{1+\alpha^{-d}}$ wires, where $\boldsymbol{\alpha \approx 1.18}$

Improved quant derand ⇒ lower bounds

- quantified derandomization at $n^{1+\exp(-d)}$ wires implies lower bounds

- **Thm 2:** If there's a deterministic $2^{n^{o(1)}}$-time alg for quantified derand of TC⁰ with $\boldsymbol{n^{1+1.61^{-d}}}$ **wires** and $\boldsymbol{B(n) = 2^{n^{1-exp(-d)}}}$, then **NEXP $\not\subset$ TC⁰**

- known algorithm handles $n^{1+\beta^{-d}}$ wires, where $\boldsymbol{\beta \approx 30}$

- for breakthrough results we need $n^{1+\alpha^{-d}}$ wires, where $\boldsymbol{\alpha \approx 1.61}$

1 we think that the known algorithm can be improved to work also when $\beta \approx 7$

The state of knowledge at STOC'18

#wires	lower bounds	derandomization
poly(n)		
$n^{1+O(1/d)}$	**specific bounds can be "amplified"** [AK'10]	**quant derand implies NEXP $\not\subset$ TC^0** [T'18]
$n^{1+\exp(-d)}$	**unconditional lower bounds**[IPS'93,CSS'16]	**unconditional quantified derandomization** [T'18]

The updated state of knowledge (STOC'19)

#wires	**lower bounds**	**derandomization**
poly(n)		
$n^{1+O(1/d)}$		
$n^{1+\alpha^{-d}}$	**specific bounds can be "amplified"** [Thm 1]	**quant derand would imply NEXP $\not\subset$ TC⁰** [Thm 2]
$n^{1+\beta^{-d}}$	**unconditional lower bounds**[IPS'93,CSS'16]	**unconditional quantified derandomization** [T'18]

1 informal; think of $\alpha < \beta$ as fixed universal constants

Hardness magnification for extremely sparse TC^0 circuits

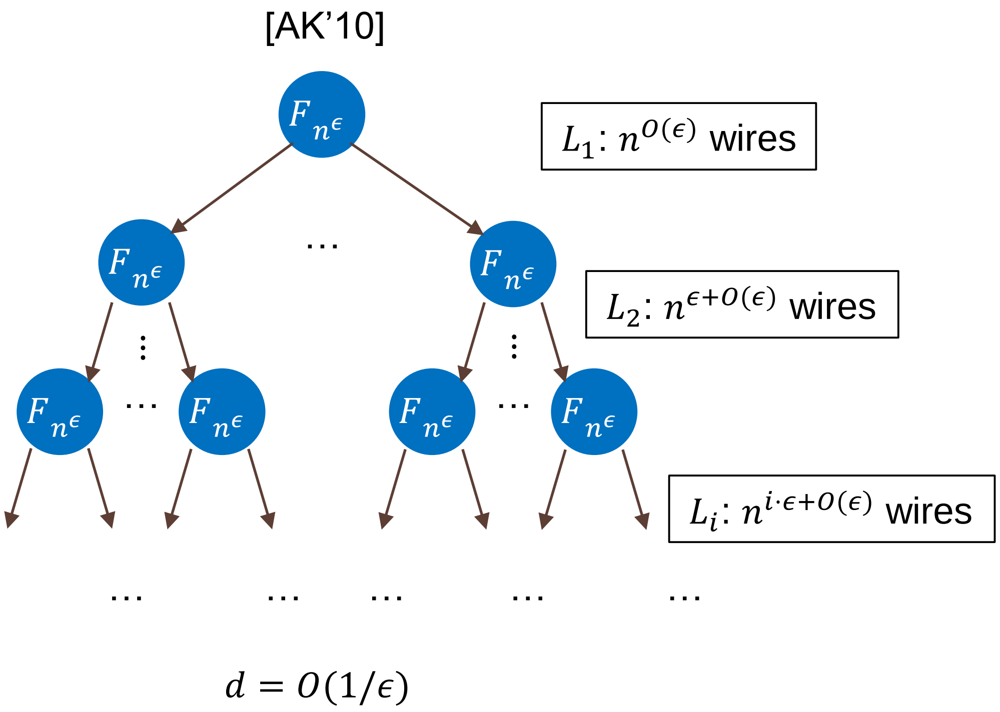
Proof overview for Thm 1

- **Idea [AK'10]**: Use the fact that NC^1 has complete funcs[1] with **associative property** ($\sigma_1, \ldots, \sigma_n \mapsto \Pi_{\{i \in [n]\}} \sigma_i$) [Bar'89]
- **Thm [AK'10]**: If an associative problem has TC^0 circuit of size $n^{O(1)}$, then it has depth-d circuit of **size $\boldsymbol{n^{1+O(1/d)}}$**
- **We** improve the implementation of their depth-d circuit to **size $\boldsymbol{n^{1+\exp(-d)}}$**, using ideas from [BBM'92, PS'94]

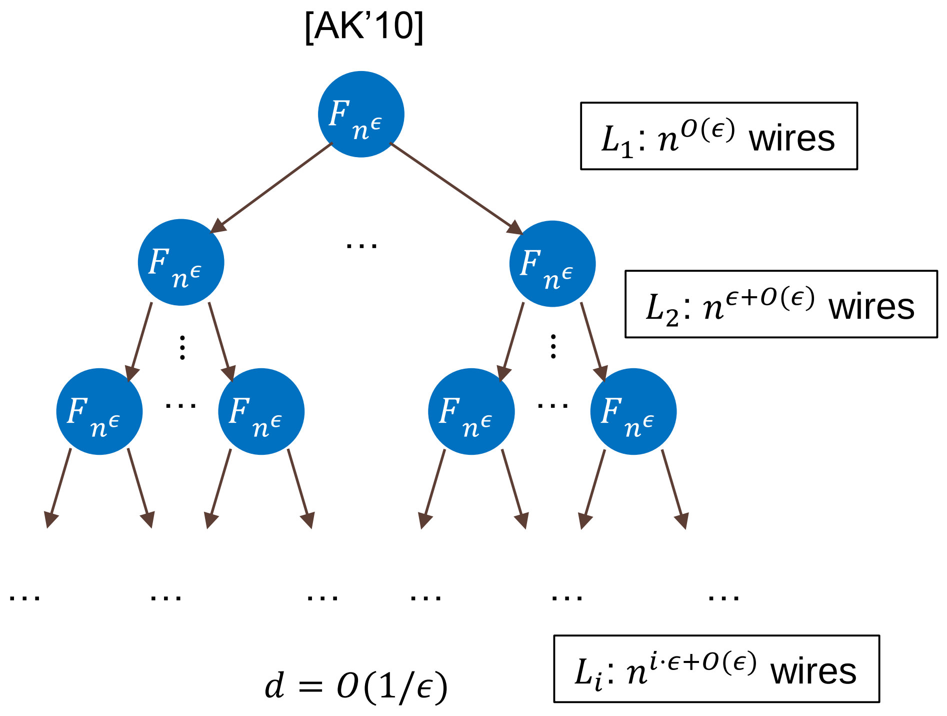
1 same approach also works for other NC^1-complete funcs

About the construction…

- **[AK'10]:** partition inputs into blocks of size n^{ϵ}, compute func on each block using hypothesized ckt (of size $n^{O(\epsilon)}$), recurse
- induces a computation tree over the inputs of depth $\mathbf{d \approx 1/\epsilon}$



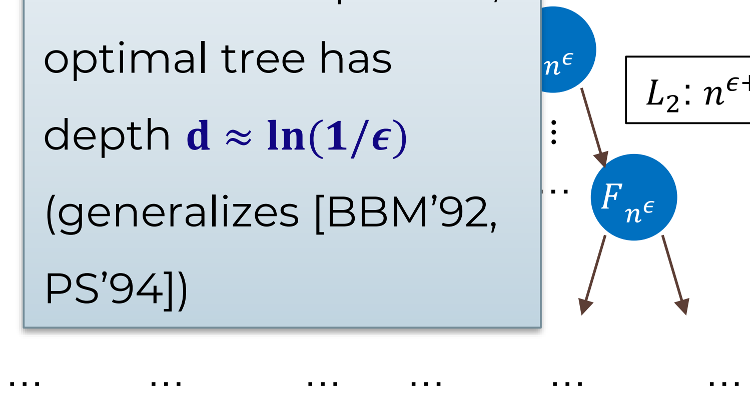
Our improvement and Comparison



Our improvement and Comparison

[AK'10]

- **our obs:** this tree is wasteful at top levels, optimal tree has depth $\mathbf{d \approx \ln(1/\epsilon)}$ (generalizes [BBM'92, PS'94])



L_1: $n^{O(\epsilon)}$ wires

L_2: $n^{\epsilon+O(\epsilon)}$ wires

L_i: $n^{i\cdot\epsilon+O(\epsilon)}$ wires

$d = O(1/\epsilon)$

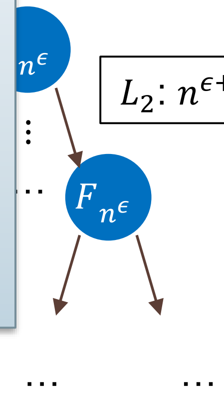
Our improvement and Comparison

[AK'10]

› **our obs:** this tree is wasteful at top levels, optimal tree has depth $\mathbf{d \approx \ln(1/\epsilon)}$ (generalizes [BBM'92, PS'94])

L_1: $n^{O(\epsilon)}$ wires

L_2: $n^{\epsilon+O(\epsilon)}$ wires



… … … … … …

$d = O(1/\epsilon)$

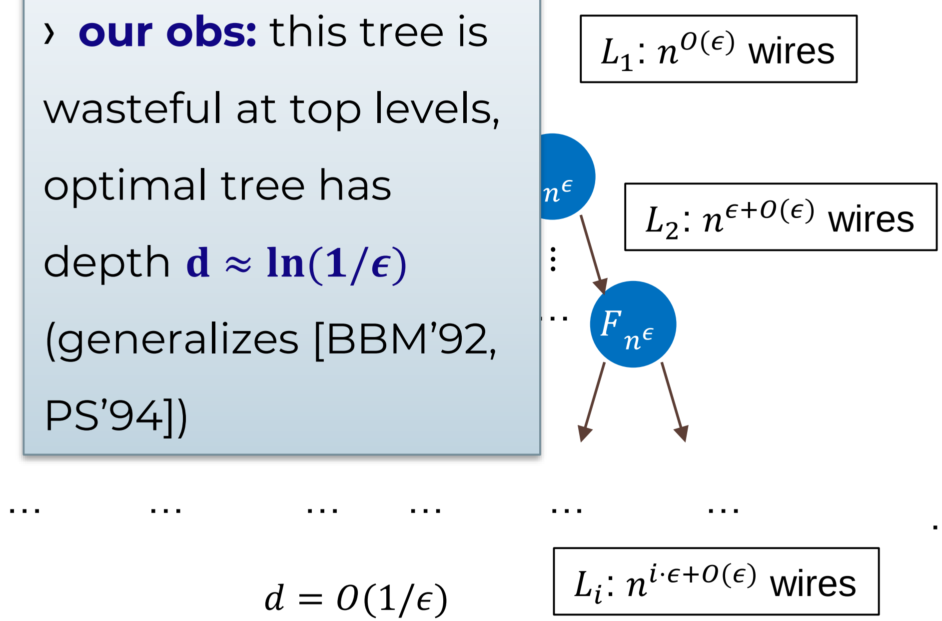
L_i: $n^{i\cdot\epsilon+O(\epsilon)}$ wires

Contributions of the layers are **imbalanced**.

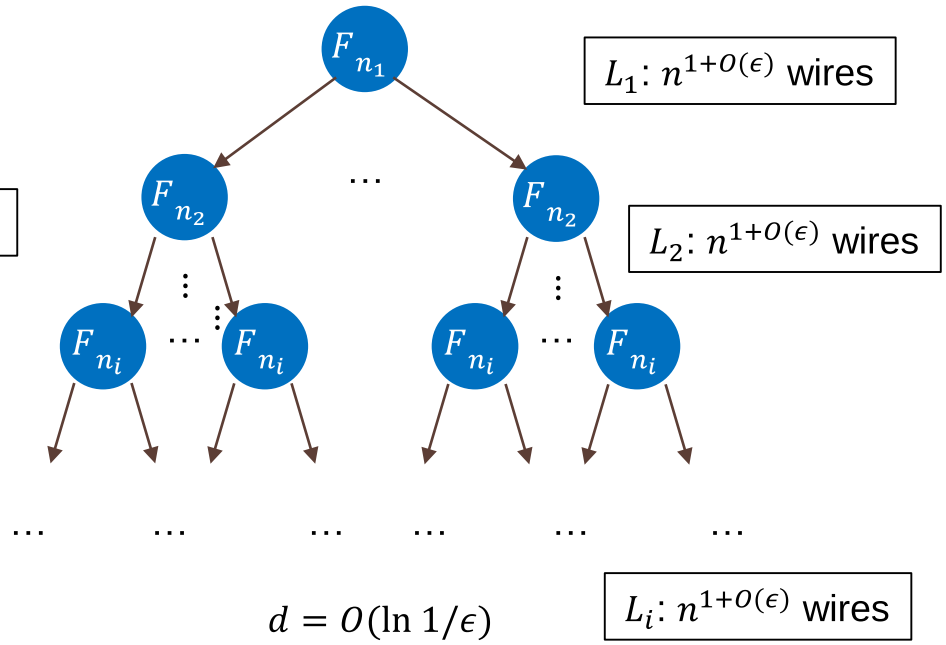
Our improvement and Comparison

[AK'10]

- **our obs:** this tree is wasteful at top levels, optimal tree has depth $\mathbf{d \approx \ln(1/\epsilon)}$ (generalizes [BBM'92, PS'94])



[This work]

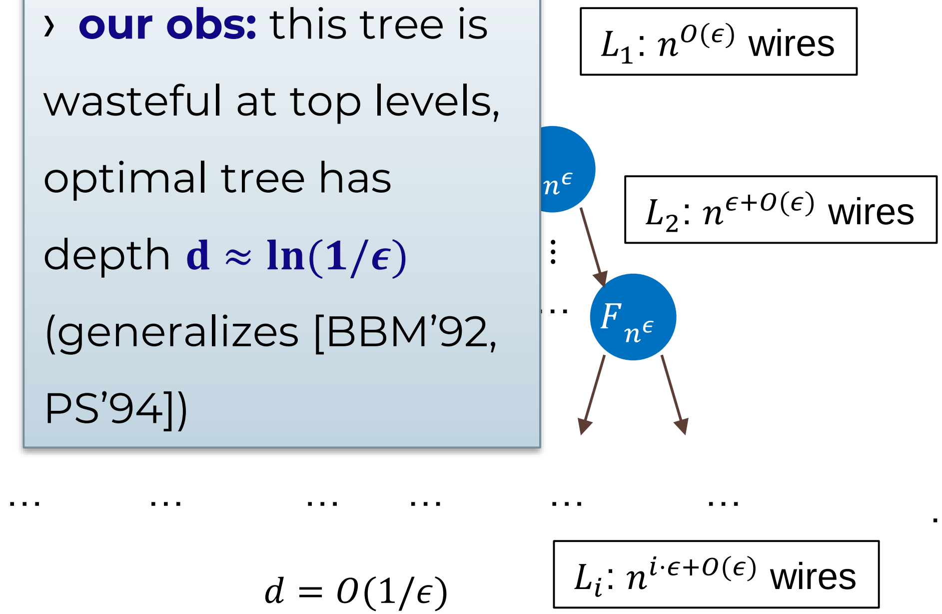


Contributions of the layers are **imbalanced**.

Our improvement and Comparison

[AK'10]

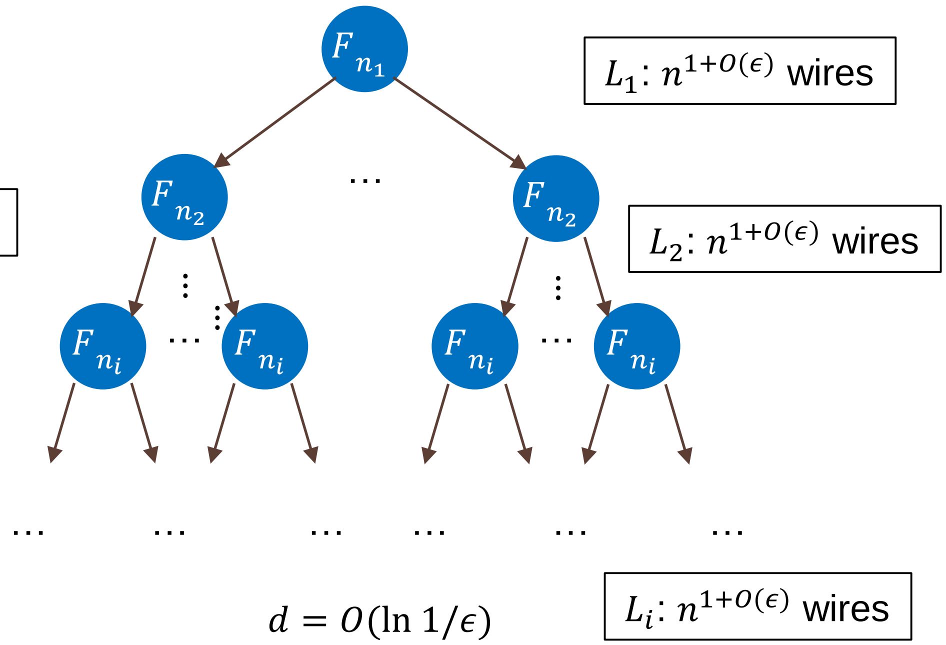
› **our obs:** this tree is wasteful at top levels, optimal tree has depth $\mathbf{d \approx \ln(1/\epsilon)}$ (generalizes [BBM'92, PS'94])



$d = O(1/\epsilon)$

Contributions of the layers are **imbalanced**.

[This work]



$d = O(\ln 1/\epsilon)$

Contributions of the layers are **balanced**.

Quantified derand of extremely sparse TC^0 implies lower bounds

Proof overview for Thm 2

- starting point: derandomization with $B(n) = 2^n/3$ for TC⁰ implies NEXP $\not\subset$ TC⁰ [Wil'13,SW'13,BV'14]

derandomization with $B(n) \approx 2^{n^{.99}}$ ⇒ **derandomization with $B(n) = 2^n/3$** ⇒ **lower bounds**

Proof overview for Thm 2

- standard idea: error-reduction
- given $C:\{0,1\}^m \to \{0,1\}$ with $\mathbf{2^m/3}$ **exceptional inputs**,

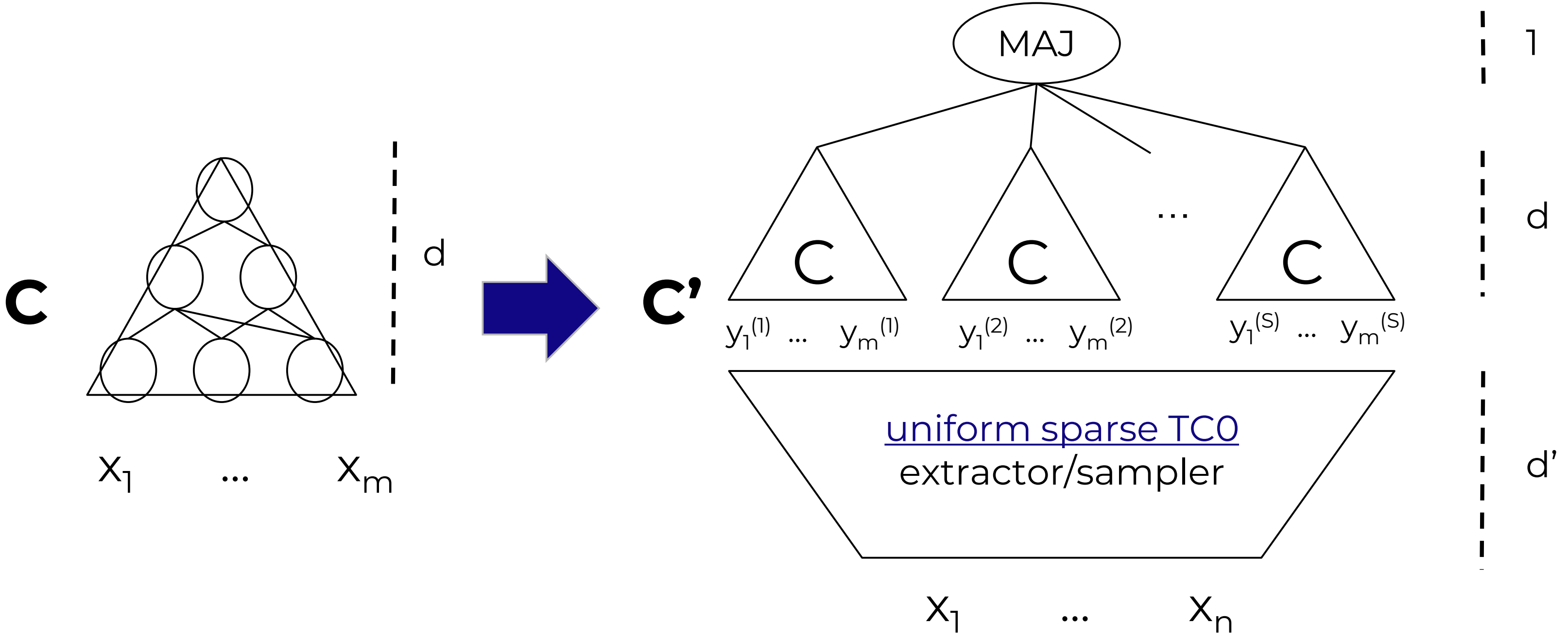
  construct $C':\{0,1\}^n \to \{0,1\}$ with $\approx \mathbf{2^{n^{.99}}}$ **exceptional inputs**

C **accepts** all but $2^m/3$ of inputs ⇒ C' **accepts** all but $\approx 2^{n^{.99}}$ of inputs

C **rejects** all but $2^m/3$ of inputs ⇒ C' **rejects** all but $\approx 2^{n^{.99}}$ of inputs

Proof overview for Thm 2

- needed: extractor/sampler in uniform sparse TC^0



Proof overview for Thm 2

- **<u>Thm:</u>** There exists an (essentially optimal) **extractor** in **uniform TC⁰** with depth d and only $\boldsymbol{n^{1+\exp(-d)}}$ **wires**
- seeded extractor: $Ext: \{0,1\}^n \times \{0,1\}^s \to \{0,1\}^m$
  - output length $\quad \mathrm{m} = n^{\exp(-d)}$
  - seed length $\quad s = (1 + \exp(-d)) \cdot \log(n)$
  - min-entropy $\quad \mathrm{k} = n^{1-\exp(-d)}$

1 observation: #outputs of the circuits ≈ #wires in the circuit!

About the construction…

- based on a non-uniform construction of [GHKPV'13]
- we show a uniform construction with minor param loss
  - components: uniform constructions of various combinatorial objects in extremely sparse TC^0 (balanced codes, designs…)
  - technical tool: zig-zag based bipartite expanders [CRVW'02]

Key takeaways

The previous intuition (at STOC'18)

- TC^0 circuits with $n^{1+\exp(-d)}$ wires are very weak, but...
- TC^0 circuits with $n^{1+O(1/d)}$ wires are very strong!
  - potential "natural proofs" barrier (PRF candidate of [MV'15])

A new intuition?

- the best explanation we have

- TC^0 circuits with $n^{1+\beta^{-\mathrm{d}}}$ wires are very weak, but…
- TC^0 circuits with $n^{1+\alpha^{-d}}$ wires are very strong?…
  - can compute linear functions, codes, extractors
  - is there a “natural proofs” barrier at $n^{1+\alpha^{-d}}$ wires?

Key takeaways

- TC^0 lower bounds are just **"a tiny improvement"** away!
- challenge: analyze TC^0 with $n^{1+\alpha^{-d}}$ wires for small $\alpha > 1$
  - show PRF candidate? or…
  - any non-trivial structural result?

Thank you!

$\Rightarrow$ new landscape for linear threshold circuits
$\Rightarrow$ breakthroughs lie “just beyond” current lower bounds