We will survey two current frontiers in circuit lower bounds: proving lower bounds for weak circuit classes that beat the lower bounds we know for general circuits, and improving the lower bounds for general circuits. Then we will mention an open problem about algorithmic approaches, i.e. improving the connections between algorithms (in particular, derandomization) and lower bounds.

1 Lower bounds for weak circuit classes

Recall that historically, the project of proving circuit lower bounds for weak circuit classes proceeded as follows.

We define $\mathcal{AC}^0$ as the class of constant-depth circuit families with AND, OR gates of unbounded fan-in (as well as $\neg$ gates). This generalizes DNFs/CNFs, for which exponential lower bounds in $\mathcal{P}$ are easy to prove. Then:

Theorem 1 ([FSS84; Yao85; Hås87]). *Parity cannot be computed by $\mathcal{AC}^0_d$ circuits of size $2^{o(n^{1/(d-1)})}$. The lower bound is tight (up to constants), and holds also on average.*

Theorem 1 was proved by random restrictions, i.e. showing that any $\mathcal{AC}^0_d$ circuit of size s simplifies on most subcubes of dimension $n/\text{polylog}(s)$.

A next step was to allow the circuits to compute parity, and ask about their power then. Clearly, random restrictions cannot work as-is (because parity doesn't meaningfully simplify in subcubes), but as we mentioned earlier in the course, circuits with $\oplus$ gates can be approximated by probabilistic low-degree polynomials over $\mathbb{F}_2$. In fact, the same argument extends to circuits with $\text{mod}\, p$ gates for a prime p, yielding:

Theorem 2 ([Raz87; Smo90; Vio20]). *For any prime p, majority cannot be computed by $\mathcal{AC}^0_d[p]$ circuits of size $2^{o(n^{1/(d-1)}/\log(n)^2)}$. The lower bound is tight (up to the polylogarithmic factor), and holds also on average.*

One side-quest has been $\text{mod}\, m$ gates for m that is composite, yielding the class $\mathcal{ACC}$, which we discussed extensively. The current lower bounds against $\mathcal{ACC}$ are not in $\mathcal{NP}$, so there is still much work to do.

1.1 Lower bounds for $\mathcal{TC}^0$

A more natural extension of $\mathcal{AC}^0[\oplus]$ is to allow MAJ gates, yielding the class $\mathcal{TC}^0$. This became a focus of attention after Williams' proof of $\mathcal{ACC}^0$ lower bounds, being the obvious next challenge (see, e.g., the first open problem in [Aar16]).

The class is also motivated by being a simplistic model of shallow neural networks. This argument was made already in the 1960s (i.e., many decades before deep learning became effective), and it is particularly salient now, as $\mathcal{TC}^0$ is also specifically related to multi-layer transformers (see [CPW24] and references therein).

Remark 3. *If we don't care about polynomial overheads, then $\mathcal{TC}^0$ can be defined in multiple equivalent ways: Polysize circuits with majority gates and negation gates; polysize circuits*

with linear threshold gates (i.e., LTF *circuits); polysize circuits with gates that can compute any symmetric function (i.e.,* SYM *circuits). The most popular form has been* LTF *circuits, since (a) We know very little about* SYM *circuits; (b) Most known techniques (for lower bounds and PRGs) extend from* MAJ *circuits to* LTF *circuits.*

We don't know any lower bound for $\mathcal{TC}^0$ that is better than the lower bounds we know for general circuits (e.g., we haven't been able to prove that $\mathcal{EXP}^{\mathcal{NP}}$ is hard for $\mathcal{TC}^0$). What we do know relies on precise size bounds, and in fact is so sensitive that we measure the size of LTF circuits in *wires* rather than in *gates*.

Theorem 4 ([IPS97; CSS16])**.** *Parity cannot be computed by $\mathcal{TC}^0_d$ circuits with $n^{1+C^{-d}}$ wires, for some constant $C > 1$. The lower bound is tight up to the constant C, and holds also on average.*

The lower bound is proved by random restrictions. Since parity can be computed by LTF circuits of size $n^{1+c^{-d}}$ for $c \ll C$, random restrictions cannot work even for the size bound $n^{1+\exp(-d)}$ more generally. In fact, something even more surprising happens when going to size $n^{1+c^{-d}}$ for $c \ll C$:

Theorem 5 ([CT19])**.** *For any $d_0, k \in \mathbb{N}$ there is $c > 1$ such that following holds. If a certain $\mathcal{NC}^1$-complete function cannot be computed by* LTF *circuits of depth d with $n^{1+c^{-d}}$ wires (for all d), then this same function is hard for* LTF *circuits of depth d_0 and size n^k.*

Theorem 6 ([CT19])**.** *Let $c = 1.61$. If there is an algorithm solving $\mathsf{CAPP}_{1,2^{n^{1-c^{-d}-n}}}$ for* LTF *circuits of depth d with $n^{1+c^{-d}}$ wires, in time $2^{n^{o(1)}}$, then $\mathcal{NEXP} \not\subset \mathcal{TC}^0$.*[1]

The reason this phenomenon happens is because LTF circuits of size $n^{1+\exp(-d)}$ are strong enough to compute objects such as error-correcting codes and randomness extractors (which allows to carry through appropriate reductions with very few wires).

Another reason is that LTF circuits of size $n^{1+\exp(-d)}$ are strong enough to simulate SYM circuits of arbitrary linear size $O(n)$, and – amazingly – *we don't know any lower bounds for the latter*. Here's an open problem then:

Open Problem 1. *Prove that there's a problem in $\mathcal{NP}$ that is hard for* SYM *circuits of constant depth and linear size $O(n)$.*

Remark 7. *For $\mathcal{AC}^0_d, \mathcal{AC}^0[p]_d, \mathcal{TC}^0_d$, there is actually a hard function that can be computed in $\mathcal{AC}^0_{d+1}$ [Sip83; Yao85; Hås87; OW07; Vio14; HRS+17; LST19; LSS+21; HHT+23] (i.e., by one more layer in depth, even if the shallower circuit has* mod p *or* LTF *gates). Also, for $\mathcal{AC}^0$ and $\mathcal{TC}^0$ we have PRG constructions that essentially match the known lower bound [AW85; TX13; Kel21; Lyu22; ST17; HHT+22] (i.e., a better PRG would yield a better lower bound), but for $\mathcal{AC}^0[\oplus]$ we do not – this is a long-standing open problem (see [CHL+19])!*

[1] This version of CAPP is called "quantified derandomization" [GW13; Tel22], meaning that we quantify the precise number of exceptional inputs. Indeed, for any number B of exceptional inputs, solving quantified derandomization in a better-than-brute-force way (i.e., in time $B/\text{polylog}(B)$) implies circuit lower bounds, see [Tel22, Theorem 7.1], and the point of the result is scaling this down to $\mathcal{TC}^0$.

1.2 Lower bounds for $\mathsf{LTF} \circ \mathsf{LTF}$

To an outsider it may look preposterous that we don't know how to prove $\mathcal{NEXP} \not\subset \mathcal{TC}^0$. Don't tell them that we don't know how to prove this even for depth-2 circuits!

Theorem 8 ([HMP+93; Nis93; FKL+01])**.** *There is a problem in $\mathcal{P}$ that is hard for $\mathsf{LTF} \circ \mathsf{MAJ}$ circuits of exponential size, and ditto re $\mathsf{MAJ} \circ \mathsf{LTF}$ circuits.*

Theorem 9 ([KW16])**.** *There is a problem in $\mathcal{P}$ that is hard for $\mathsf{LTF} \circ \mathsf{LTF}$ circuits with $\tilde{\Omega}(n^{2.5})$ wires.*

Open Problem 2. *Prove that there is a problem in $\mathcal{EXP}^{\mathcal{NP}}$ that is hard for $\mathsf{LTF} \circ \mathsf{LTF}$ circuits of polynomial size.*

There is a promising way to make progress on this problem, using the algorithmic method. This is non-trivial: the algorithmic method adds overheads, i.e. when we want to prove lower bounds for a class $\mathcal{C}$, we usually need to analyze some larger class $\hat{\mathcal{C}}$.[2] So it seems like applying the algorithmic method to $\mathsf{LTF} \circ \mathsf{LTF}$ circuits will require us to analyze a larger class.

Amazingly, to prove lower bounds for $\mathsf{LTF} \circ \mathsf{LTF}$ circuits it suffices to analyze a *weaker* class, for which exponential lower bounds are already known!

Theorem 10 ([CW19])**.** *Assume that CAPP of n-bit $\mathsf{MAJ} \circ \mathsf{MAJ}$ circuits of polynomial size can be solved in time $2^n/n^{\omega(1)}$. Then $\mathcal{NEXP} \not\subset \mathcal{P}/\mathrm{poly}$.* [3]

Indeed, we currently know exponential lower bounds in $\mathcal{P}$ for $\mathsf{MAJ} \circ \mathsf{MAJ}$ circuits, but no non-trivial CAPP algorithm.

2 Lower bounds in $\mathcal{E}^{\mathcal{NP}}$

Let us turn to the other direction: working with general circuits, say of exponential size $2^{\epsilon \cdot n}$, and improving the upper-bound from $\mathcal{S}_2\mathcal{E} \subseteq \mathcal{ZPE}^{\mathcal{NP}}$ to $\mathcal{E}^{\mathcal{NP}}$.

In the final presentations day, Rishabh and Jacob will show one algorithmic approach to prove this lower bound, from [RSW22], which goes through *Avoid*. They will also show an algorithmic approach that is *necessary* for $\mathcal{E}^{\mathcal{NP}}$ lower bounds.

A scaled down challenge. A more relaxed goal is proving that $\mathcal{P}^{\mathcal{NP}}$ is hard for circuits of fixed polynomial size. (Note that this is implied by exponential circuit lower bounds in $\mathcal{E}^{\mathcal{NP}}$, so it is indeed a more relaxed goal.)

Recall that one of the earliest approach to circuit lower bounds, by [Kan82], is based on the Karp-Lipton theorem and a win-win argument. Specifically, we start from the

[2] Going through the proof, $\hat{\mathcal{C}}$ is the class of circuits obtained by simulating a PCP verifier and answering its queries by a $\mathcal{C}$-circuit.

[3] The CAPP algorithm here needs to estimate the acceptance probability up to an additive error of $1/\mathrm{poly}(n)$. Solving CAPP for $\mathsf{LTF} \circ \mathsf{MAJ}$ circuits with error $\Omega(1)$ would yield the same conclusion.

KL theorem: If $\mathcal{NP} \subset \mathcal{P}/\text{poly}$ then $\mathcal{PH}$ collapses to Σ_2. Then either $\mathcal{NP} \not\subset \mathcal{P}/\text{poly}$, a win; or $\mathcal{PH}$ collapses to Σ_2, and hence Σ_2 doesn't have fixed-polysize circuits (because $\mathcal{PH}$ is hard for such circuits) – another win.

If we can improve the KL theorem such that the collapse is to $\mathcal{P}^{\mathcal{NP}}$ instead of Σ_2, then we'd get circuit lower bounds in $\mathcal{P}^{\mathcal{NP}}$; that is:

Proposition 11. *Assume that*

$$\mathcal{NP} \subset \mathcal{P}/\text{poly} \Rightarrow \mathcal{PH} \subset \mathcal{P}^{\mathcal{NP}} .$$

Then $\mathcal{P}^{\mathcal{NP}}$ is hard for size n^k, for any fixed k.

In fact, for the conclusion it's enough to deduce that $\mathcal{PH}$ collapses to $\mathcal{P}^{\mathcal{NP}}$ infinitely often, and even enough to assume it collapses to $\mathcal{P}^{\mathcal{NP}}/n$.

It might initially seem like there's no reason to be interested in the specific improvement to the KL theorem; it's not clear that this should be easy to prove, let alone that this should be the right approach to deducing circuit lower bounds. Amazingly, it turns out that this improvement to the KL theorem is *necessary*; that is, the deduced lower bound *implies* the KL improvement:

Theorem 12 ([CMM+19])**.** *The following statements are equivalent:*

1. *$\mathcal{P}^{\mathcal{NP}}$ is hard for size n^k for any fixed k .*
2. *$\mathcal{NP} \subset \mathcal{P}/\text{poly} \Rightarrow \mathcal{PH} \subset \text{i.o.}\mathcal{P}^{\mathcal{NP}}/n$.*

Open Problem 3. *Prove that $\mathcal{NP} \subset \mathcal{P}/\text{poly} \Rightarrow \mathcal{PH} \subset \text{i.o.}\mathcal{P}^{\mathcal{NP}}/n$.*

3 Better lower bounds from better algorithms

The last problem I'll mention is to improve the algorithmic method, i.e. show that CAPP algorithms imply stronger circuit lower bounds (compared to what is currently known). Since we're shooting for a strong conclusion, let's start with a strong hypothesis (and relax it later) – polynomial-time CAPP algorithms.

Non-deterministic algorithms. Assume that $pr\mathcal{BPP} \subseteq pr\mathcal{NP}$. To gauge what we should shoot for, recall that exponential circuit lower bounds in $\mathcal{NE} \cap co\text{-}\mathcal{NE}$ imply $pr\mathcal{BPP} \subseteq pr\mathcal{NP}$.[4] Can we show that such lower bounds are also necessary? We've seen that $pr\mathcal{BPP} \subseteq pr\mathcal{NP}$ implies exponential *witness* lower bounds, but deducing exponential *circuit* lower bounds is an open problem.

[4]The reason for assuming lower bounds in $\mathcal{NE} \cap co\text{-}\mathcal{NE}$ rather than $\mathcal{NE}$ is that we want to compute a hard truth-table non-deterministically, and apply a PRG to it. If the hard function is in $\mathcal{NE}$, then a prover can convince us of correctness of some 1-entries in the truth-table, but maybe that prover omitted proofs for other 1-entries; when the problem is in $\mathcal{NE} \cap co\text{-}\mathcal{NE}$, we can demand proofs both for 1-entries and for 0-entries. Note that the conclusion $pr\mathcal{BPP} \subseteq pr\mathcal{NP}$ is "fair", since this is equivalent to $pr\mathcal{BPP} \subseteq pr\mathcal{NP} \cap pr\text{co}\mathcal{NP}$ (because $pr\mathcal{BPP}$ is closed under complementation).

Open Problem 4. *Assuming that $prBPP \subseteq prNP$ (or even that $prBPP = prP$), prove that NE is hard for circuits of size $2^{\epsilon \cdot n}$, for some $\epsilon > 0$.*[5]

Deterministic algorithms. The strongest lower bounds that we know follow from CAPP algorithms that may be non-deterministic, and analyze n-bit circuits of size $2^{\epsilon \cdot n}$ in time $2^{(1-\epsilon)\cdot n}$. Recall that this is a fine-grained question about a problem in P, since the input length is $N = 2^{\epsilon \cdot n}$ and the brute-force algorithm runs in time $\text{poly}(N)$.

The assumption $prBPP = prP$ seems considerably stronger: we do not use non-determinism, and we can handle N-size circuits with N input bits (rather than only $n = O(\log N)$ input bits). One might expect that we'll be able to prove a much stronger conclusion from it. Doing so is an open problem:

Open Problem 5. *Assuming $prBPP = prP$, deduce a stronger lower bound conclusion than what is currently known to follow from non-deterministic* CAPP *algorithms for size-$2^{\epsilon \cdot n}$ circuits running in time $2^{(1-\epsilon)\cdot n}$.*

References

[Aar16] Scott Aaronson. "$P \stackrel{?}{=} NP$". In: *Open Problems in Mathematics*. Ed. by John Forbes Nash Jr. and Michael Th. Rassias. Springer International Publishing, 2016, pp. 1–122.

[AW85] Miklos Ajtai and Avi Wigderson. "Deterministic simulation of probabilistic constant depth circuits". In: *Proc. 26th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1985.

[CHL+19] Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal. "Pseudorandom generators from the second Fourier level and applications to AC0 with parity gates". In: *Proc. 10th Conference on Innovations in Theoretical Computer Science (ITCS)*. Vol. 124. LIPIcs. Leibniz Int. Proc. Inform. 2019, Art. No. 22, 15.

[CMM+19] Lijie Chen, Dylan M. McKay, Cody D. Murray, and R. Ryan Williams. "Relations and Equivalences Between Circuit Lower Bounds and Karp-Lipton Theorems". In: *Proc. 34th Annual IEEE Conference on Computational Complexity (CCC)*. 2019, 30:1–30:21.

[CPW24] Lijie Chen, Binghui Peng, and Hongxun Wu. "Theoretical limitations of multi-layer Transformer". In: (2024). arXiv: 2412.02975. URL: https://arxiv.org/abs/2412.02975.

[CSS16] Ruiwen Chen, Rahul Santhanam, and Srikanth Srinivasan. "Average-case lower bounds and satisfiability algorithms for small threshold circuits". In: *Proc. 31st Annual IEEE Conference on Computational Complexity (CCC)*. 2016, 1:1–1:35.

[5] Note that if we assume SAT is in polynomial time, the conclusion follows easily (in fact, E requires circuits of essentially maximal size). But we'd like to rely on a more believable hypothesis.

[CT19] Lijie Chen and Roei Tell. "Bootstrapping results for threshold circuits "just beyond" known lower bounds". In: *Proc. 51st Annual ACM Symposium on Theory of Computing (STOC)*. 2019, pp. 34–41.

[CW19] Lijie Chen and R. Ryan Williams. "Stronger Connections Between Circuit Analysis and Circuit Lower Bounds, via PCPs of Proximity". In: *Proc. 34th Annual IEEE Conference on Computational Complexity (CCC)*. 2019, 19:1–19:43.

[FKL+01] Jürgen Forster, Matthias Krause, Satyanarayana V. Lokam, Rustam Mubarakzjanov, Niels Schmitt, and Hans Ulrich Simon. "Relations Between Communication Complexity, Linear Arrangements, and Computational Complexity". In: *Proc. 21st Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*. 2001, pp. 171–182.

[FSS84] Merrick Furst, James B. Saxe, and Michael Sipser. "Parity, circuits, and the polynomial-time hierarchy". In: *Mathematical Systems Theory* 17.1 (1984), pp. 13–27.

[GW13] Oded Goldreich and Avi Widgerson. "On derandomizing algorithms that err extremely rarely". In: *Electronic Colloquium on Computational Complexity: ECCC* 20 (2013), p. 152.

[Hås87] Johan Håstad. *Computational Limitations of Small-depth Circuits*. MIT Press, 1987.

[HHT+22] Pooya Hatami, William M. Hoza, Avishay Tal, and Roei Tell. "Fooling constant-depth threshold circuits". In: *Proc. 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. [2022] ©2022, pp. 104–115.

[HHT+23] Pooya Hatami, William M. Hoza, Avishay Tal, and Roei Tell. "Depth-d threshold circuits vs. depth-$(d+1)$ AND-OR trees". In: *Proc. 55th Annual ACM Symposium on Theory of Computing (STOC)*. [2023] ©2023, pp. 895–904.

[HMP+93] András Hajnal, Wolfgang Maass, Pavel Pudlák, Mario Szegedy, and György Turán. "Threshold Circuits of Bounded Depth". In: *Journal of Computer and System Sciences* 46.2 (1993), pp. 129–154.

[HRS+17] Johan Håstad, Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. "An Average-Case Depth Hierarchy Theorem for Boolean Circuits". In: *Journal of the ACM* 64.5 (2017), 35:1–35:27.

[IPS97] Russell Impagliazzo, Ramamohan Paturi, and Michael E. Saks. "Size-depth tradeoffs for threshold circuits". In: *SIAM Journal on Computing* 26.3 (1997), pp. 693–707.

[Kan82] R. Kannan. "Circuit-size lower bounds and non-reducibility to sparse sets". In: *Information and Control* 55.1-3 (1982), pp. 40–56.

[Kel21] Zander Kelley. "An improved derandomization of the switching lemma". In: *Proc. 53rd Annual ACM Symposium on Theory of Computing (STOC)*. 2021, pp. 272–282.

[KW16] Daniel M. Kane and Ryan Williams. "Super-linear Gate and Super-quadratic Wire Lower Bounds for Depth-two and Depth-three Threshold Circuits". In: *Proc. 48th Annual ACM Symposium on Theory of Computing (STOC)*. 2016, pp. 633–643.

[LSS+21] Nutan Limaye, Karteek Sreenivasaiah, Srikanth Srinivasan, Utkarsh Tripathi, and S. Venkitesh. "A fixed-depth size-hierarchy theorem for $\mathrm{AC}^0[\oplus]$ via the coin problem". In: *SIAM Journal on Computing* 50.4 (2021), pp. 1461–1499.

[LST19] Nutan Limaye, Srikanth Srinivasan, and Utkarsh Tripathi. "More on $\mathrm{AC}^0[\oplus]$ and variants of the majority function". In: *Proc. 39th Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*. 2019, Art. No. 22, 14.

[Lyu22] Xin Lyu. "Improved pseudorandom generators for AC^0 circuits". In: *Proc. 37th Annual IEEE Conference on Computational Complexity (CCC)*. 2022, Art. No. 34, 25.

[Nis93] Noam Nisan. "The communication complexity of threshold gates". In: *Combinatorics, Paul Erdős is eighty, Vol. 1*. 1993, pp. 301–315.

[OW07] Ryan O'Donnell and Karl Wimmer. "Approximation by DNF: examples and counterexamples". In: *Automata, languages and programming*. Vol. 4596. Lecture Notes in Comput. Sci. 2007, pp. 195–206.

[Raz87] Alexander A. Razborov. "Lower bounds on the size of constant-depth networks over a complete basis with logical addition". In: *Mathematical Notes of the Academy of Science of the USSR* 41.4 (1987), pp. 333–338.

[RSW22] Hanlin Ren, Rahul Santhanam, and Zhikun Wang. "On the Range Avoidance Problem for Circuits". In: *Proc. 63rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 2022.

[Sip83] Michael Sipser. "Borel sets and circuit complexity". In: *Proc. 15th Annual ACM Symposium on Theory of Computing (STOC)*. 1983, 61–69.

[Smo90] Roman Smolensky. "On interpolation by analytic functions with special properties and some weak lower bounds on the size of circuits with symmetric gates". In: *Proc. 31st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1990, pp. 628–631.

[ST17] Rocco Servedio and Li-Yang Tan. "Learning and fooling depth-two threshold circuits". Unpublished manuscript. 2017.

[Tel22] Roei Tell. "Quantified derandomization: how to find water in the ocean". In: *Foundations and Trends® in Theoretical Computer Science* 15.1 (2022), Paper No 1, 125.

[TX13] Luca Trevisan and TongKe Xue. "A derandomized switching lemma and an improved derandomization of AC0". In: *Proc. 28th Annual IEEE Conference on Computational Complexity (CCC)*. 2013, pp. 242–247.

[Vio14] Emanuele Viola. "Randomness Buys Depth for Approximate Counting". In: *Computational Complexity* 23.3 (2014), pp. 479–508.

[Vio20] Emanuele Viola. "New lower bounds for probabilistic degree and AC0 with parity gates". In: *Electronic Colloquium on Computational Complexity (ECCC)* 27 (2020), p. 15. URL: https://eccc.weizmann.ac.il/report/2020/015.

[Yao85] Andrew C-C. Yao. "Separating the Polynomial-time Hierarchy by Oracles". In: *Proc. 26th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1985, pp. 1–10.