# Majority vs. Approximate Linear Sum and Average-Case Complexity Below $\mathsf{NC}^1$

**Lijie Chen** ✉
MIT, Cambridge, MA, USA

**Zhenjian Lu** ✉
University of Warwick, Coventry, UK

**Xin Lyu** ✉
Tsinghua University, Beijing, China

**Igor C. Oliveira** ✉
University of Warwick, Coventry, UK

## Abstract

We develop a general framework that characterizes strong average-case lower bounds against circuit classes $\mathcal{C}$ contained in $\mathsf{NC}^1$, such as $\mathsf{AC}^0[\oplus]$ and $\mathsf{ACC}^0$. We apply this framework to show:

- *Generic seed reduction*: Pseudorandom generators (PRGs) against $\mathcal{C}$ of seed length $\leq n-1$ and error $\varepsilon(n) = n^{-\omega(1)}$ can be converted into PRGs of *sub-polynomial* seed length.
- *Hardness under natural distributions*: If $\mathsf{E}$ (deterministic exponential time) is average-case hard against $\mathcal{C}$ under *some* distribution, then $\mathsf{E}$ is average-case hard against $\mathcal{C}$ under the *uniform* distribution.
- *Equivalence between worst-case and average-case hardness*: Worst-case lower bounds against $\mathsf{MAJ} \circ \mathcal{C}$ for problems in $\mathsf{E}$ are *equivalent* to strong average-case lower bounds against $\mathcal{C}$. This can be seen as a certain converse to the Discriminator Lemma [Hajnal et al., JCSS'93].

These results were not known to hold for circuit classes that do not compute majority. Additionally, we prove that classical and recent approaches to *worst-case* lower bounds against $\mathsf{ACC}^0$ via communication lower bounds for NOF multi-party protocols [Håstad and Goldmann, CC'91; Razborov and Wigderson, IPL'93] and Torus polynomials degree lower bounds [Bhrushundi et al., ITCS'19] also imply *strong average-case hardness* against $\mathsf{ACC}^0$ under the uniform distribution.

Crucial to these results is the use of *non-black-box* hardness amplification techniques and the interplay between ***Majority*** ($\mathsf{MAJ}$) and ***Approximate Linear Sum*** ($\widetilde{\mathsf{SUM}}$) gates. Roughly speaking, while a $\mathsf{MAJ}$ gate outputs 1 when the sum of the $m$ input bits is at least $m/2$, a $\widetilde{\mathsf{SUM}}$ gate computes a real-valued bounded weighted sum of the input bits and outputs 1 (resp. 0) if the sum is close to 1 (resp. close to 0), with the promise that one of the two cases always holds. As part of our framework, we explore ideas introduced in [Chen and Ren, STOC'20] to show that, for the purpose of proving lower bounds, a top layer $\mathsf{MAJ}$ gate is *equivalent* to a (weaker) $\widetilde{\mathsf{SUM}}$ gate. Motivated by this result, we extend the algorithmic method and establish stronger lower bounds against bounded-depth circuits with layers of $\mathsf{MAJ}$ and $\widetilde{\mathsf{SUM}}$ gates. Among them, we prove that:

- *Lower bound*: $\mathsf{NQP}$ does not admit fixed quasi-polynomial size $\mathsf{MAJ} \circ \widetilde{\mathsf{SUM}} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$ circuits.

This is the first explicit lower bound against circuits with distinct layers of $\mathsf{MAJ}$, $\widetilde{\mathsf{SUM}}$, and $\mathsf{THR}$ gates. Consequently, if the aforementioned equivalence between $\mathsf{MAJ}$ and $\widetilde{\mathsf{SUM}}$ as a *top gate* can be extended to *intermediate layers*, long sought-after lower bounds against the class $\mathsf{THR} \circ \mathsf{THR}$ of depth-2 polynomial-size threshold circuits would follow.

**2012 ACM Subject Classification** Theory of computation

**Keywords and phrases** circuit complexity, average-case hardness, complexity lower bounds

**Digital Object Identifier** 10.4230/LIPIcs.ICALP.2021.51

**Category** Track A: Algorithms, Complexity and Games

**Related Version** *Full Version*: `https://eccc.weizmann.ac.il/report/2021/040/` [11]

**Funding** Lijie Chen is supported by an IBM fellowship. This work received support from the Royal Society University Research Fellowship URF\R1\191059.

**Acknowledgements** We would like to thank William Hoza for asking whether PRGs for $\mathcal{C}$ imply average-case lower bounds against $\mathcal{C}$ under the uniform distribution. We are also grateful to Roei Tell for the observation that part of our equivalence theorem extends to the constant-error case as well.

# 1 Introduction

## 1.1 Overview

Establishing the *intractability of computations* and understanding the *power of randomness* in algorithms are among the most basic open problems in theoretical computer science. The theory of computational pseudorandomness provides a firm link between these two research directions. One of the most celebrated developments in this area is a proof that if $\mathsf{E}$ (deterministic exponential time $2^{O(n)}$) requires Boolean circuits of exponential size then $\mathsf{P} = \mathsf{BPP}$ [30, 43]. This result and its underlying techniques provide a robust mathematical theory that connects *worst-case lower bounds*, *average-case hardness*, and the construction of *pseudorandom generators*.

Unfortunately, a large part of this beautiful and far-reaching theory is not known to survive in *restricted* computational settings. For instance, while we know since the eighties that $\mathsf{E}$ cannot be $(1/2 + n^{-1/2+\Omega(1)})$-approximated by $\mathsf{AC}^0[\oplus]$ [40], it is an important open problem to obtain *strong* average-case hardness results of the form $1/2 + n^{-k}$ for all $k$ and pseudorandom generators against this circuit class. The fact that existing connections between hardness and pseudorandomness do not apply in restricted settings is significant, given that known unconditional results and existing lower bound frontiers lie within weak sub-classes of $\mathsf{NC}^1$, such as $\mathsf{ACC}^0$.

Several works (e.g. [45, 23, 42, 36, 5, 22, 47, 29]) have investigated the difficulty of extending the hardness vs. randomness theory and its consequences to restricted circuit classes. Roughly speaking, these results show that standard "black-box" techniques to amplify computational hardness and construct pseudorandom generators *require* the underlying circuit class $\mathcal{C}$ to be closed under *majority*. However, obtaining lower bounds against circuit classes that are closed under majority is a notorious open problem. This leaves us in this unsatisfying situation where many benefits of the theory mentioned above only apply to settings where current circuit-analysis techniques do not hold. In other words, we have the following "lose-lose" scenario: above $\mathsf{TC}^0$ we have no lower bounds, while below it we have lower bounds but no hardness amplification.

In this work, we explore *non-black-box* techniques to overcome this difficulty, obtaining a general connection between worst-case lower bounds, strong average-case hardness, and pseudorandomness for *weak circuit classes*. Our results build on recent ideas of Chen and Ren [14] employed in the context of the algorithmic method. Using our techniques, we are able to establish fundamental equivalences that were previously only known for circuit classes containing $\mathsf{TC}^0$. As a consequence, the new results are widely applicable and can affect *current frontiers in circuit complexity theory*.

A crucial ingredient in our proofs is the interplay between Majority ($\mathsf{MAJ}$) and Approximate Linear Sum ($\widetilde{\mathsf{SUM}}$) gates. Roughly speaking, while a $\mathsf{MAJ}$ gate outputs 1 when the sum of the $m$ input bits is at least $m/2$, a $\widetilde{\mathsf{SUM}}$ gate computes a real-valued bounded weighted sum of the input bits and outputs 1 (resp. 0) if the sum is close to 1 (resp. close to 0), with the promise that one of the two cases always holds. $\widetilde{\mathsf{SUM}}$ gates are significantly simpler than $\mathsf{MAJ}$ gates (e.g. $\mathsf{MAJ}$ has approximate degree [38] of order $\Omega(m)$), but still powerful enough to implement useful computations, such as hardness amplification for *specific* problems (a non-black-box element).

Complementing our results about the average-case complexity of restricted circuit classes, we obtain the first unconditional lower bounds against bounded-depth circuits with distinct layers of $\mathsf{MAJ}$, $\widetilde{\mathsf{SUM}}$, and $\mathsf{THR}$ gates. These results suggest that further investigating the relation between $\mathsf{MAJ}$ and $\widetilde{\mathsf{SUM}}$ might be a path to lower bounds against depth-2 threshold circuits, a long-standing open problem in complexity theory (cf. [18, 9]).

## 1.2 Results and techniques

To begin with, we recall some definitions for linear sums of functions. Our notation is taken from previous work [50, 15, 14, 13] on lower bounds via the algorithmic method. Let $\mathcal{C}$ be a class of functions from $\{0,1\}^n \to \{0,1\}$.

**$\mathsf{SUM} \circ \mathcal{C}$-circuits.** A $\mathsf{SUM} \circ \mathcal{C}$-circuit $C\colon \{0,1\}^n \to \mathbb{R}$ is a circuit that can be written as $C(x) = \sum_{i=1}^{\ell} \alpha_i \cdot C_i(x)$, where each $\alpha_i$ is a real, and each $C_i \in \mathcal{C}$. Here $\ell$ is called the *sparsity* of $C$, and is denoted as $\mathsf{sparsity}(C)$. We also use $\mathsf{complexity}(C)$ to denote $\max(\ell, \sum_{i=1}^{\ell} |\alpha_i|)$. Furthermore, if a $\mathsf{SUM} \circ \mathcal{C}$-circuit $C$ always outputs values in the interval $[0,1]$, we say it is a $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}$-circuit.

**$\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-circuits.** Let $\delta \in [0, 0.5)$. A $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-circuit $C\colon \{0,1\}^n \to \{0,1\}$ is defined by a $\mathsf{SUM} \circ \mathcal{C}$-circuit $L\colon \{0,1\}^n \to \mathbb{R}$ satisfying the following promise: for every $x \in \{0,1\}^n$, either $|L(x) - 1| \le \delta$ or $|L(x)| \le \delta$. (We stress that this promise is only required over inputs $x$ to the $\mathsf{SUM} \circ \mathcal{C}$-circuit $L$, and not over all possible input values to the top $\mathsf{SUM}$ gate.) We say $C(x) = 1$ if $|L(x) - 1| \le \delta$ and $C(x) = 0$ otherwise. The sparsity and the complexity of $C$ is defined as the sparsity and the complexity of $L$, respectively.

For a circuit class $\mathcal{C}$, we use $\mathsf{SUM} \circ \mathcal{C}$, $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}$, and $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$ to denote the collection of such circuit families with at most $\mathrm{poly}(n)$ complexity. When $\mathcal{C}$ has a clear notion of complexity, such as circuit size, this also means that the involved $\mathcal{C}$-subcircuits are of polynomial size. In some statements we might refer to classes such as $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}[s]$ to emphasize a specific upper bound $s$ on the complexities of $\mathcal{C}$-subcircuits and of the top gate.

**Notation for standard concepts.** A $\mathsf{MAJ}\colon \{0,1\}^m \to \{0,1\}$ gate $\mathsf{MAJ}(y_1, \dots, y_m)$ outputs 1 if and only if $\sum_i y_i \ge m/2$. A $\mathsf{THR}\colon \{0,1\}^m \to \{0,1\}$ gate is described by weights $w_1, \dots, w_m, \theta \in \mathbb{R}$ and outputs 1 if and only if $\sum_i w_i y_i \ge \theta$.

For a probability distribution $\mathcal{D}$ over $\{0,1\}^n$ and Boolean functions $f, g\colon \{0,1\}^n \to \{0,1\}$, we say that $f$ is $\gamma$-approximated by $g$ over $\mathcal{D}$ if $\Pr_{x \sim \mathcal{D}}[f(x) = g(x)] \ge \gamma$. For convenience, circuit lower bounds involving approximations of the form $1/2 + 1/n^{\omega(1)}$ might be informally referred to as *strong average-case lower bounds* or simply *strong correlation bounds*.

Our results refer to non-uniform circuit classes, and we use $\mathcal{C}_1 \circ \mathcal{C}_2$ to refer to circuit families consisting of a top circuit from $\mathcal{C}_1$ composed with bottom circuits from $\mathcal{C}_2$.[1]

We use $\mathcal{U}_n$ to denote the uniform distribution over $\{0,1\}^n$. A distribution $\mathcal{D}$ *$\varepsilon$-fools* a function $f\colon \{0,1\}^n \to \{0,1\}$ if $|\Pr[f(\mathcal{D}) = 1] - \Pr[f(\mathcal{U}_n) = 1]| \le \varepsilon$. We say that a sequence $G_n\colon \{0,1\}^{\ell(n)} \to \{0,1\}^n$ is an infinitely often PRG against a circuit class $\mathcal{C}$ with error $\varepsilon$

[1] As usual, in the case of $\mathcal{C}_2 = \mathsf{ACC}^0$, where $\mathsf{ACC}^0 = \bigcup_{m \in \mathbb{N}} \mathsf{AC}^0[m]$ with $m$ here representing the modulo, we require that each $\mathcal{C}_2$-subcircuit of a circuit $D$ from $\mathcal{C}_1 \circ \mathcal{C}_2$ uses the same fixed $m$.

(i.o. $\varepsilon$-PRG) and seed length $\ell$ if $G_n$ is computable in time $2^{O(\ell(n))}$ and for infinitely many values of $n$, the induced distribution $G_n(\mathcal{U}_{\ell(n)})$ $\varepsilon(n)$-fools each function $f\colon \{0,1\}^n \to \{0,1\}$ in $\mathcal{C}$.

### 1.2.1 Equivalences for worst-case and strong average-case hardness

Our first contribution is a general result that tightly connects worst-case lower bounds, strong average-case hardness, and pseudorandomness in *restricted computational models.*

▶ **Theorem 1** (Non-black-box equivalences for worst-case and strong average-case hardness). *Let $\mathcal{C}$ be a circuit class contained in $\mathsf{NC}^1$ that is closed under negations and under a bottom layer of juntas over $O(1)$ input bits. The following statements are equivalent:*

1. *There is $L \in \mathsf{E}$ such that $L \notin \widetilde{\mathsf{SUM}}_{1/3} \circ \mathcal{C}$.*
2. *There is $L \in \mathsf{E}$ and $\delta \geq 1/\mathsf{poly}(n)$ such that $L \notin \widetilde{\mathsf{SUM}}_{\delta} \circ \mathcal{C}$.*
3. *There is $L \in \mathsf{E}$ such that $L \notin \mathsf{MAJ} \circ \mathcal{C}$.*
4. *There is $L \in \mathsf{E}$ such that, for every $k \geq 1$, $L$ cannot be computed by probabilistic $\mathcal{C}$-circuits with error $1/2 - 1/n^k$.*[2]
5. *There is $L \in \mathsf{E}$ and a distribution ensemble $\mathfrak{D}$ such that for every $k \geq 1$, $L$ cannot be $(1/2 + n^{-k})$-approximated by $\mathcal{C}$ under $\mathfrak{D}$.*
6. *There is $L \in \mathsf{E}$ such that for every $k \geq 1$, $L$ cannot be $(1/2 + n^{-k})$-approximated by $\mathcal{C}$ under the uniform distribution.*
7. *There is $L \in \mathsf{E}$ that cannot be approximated by $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}$ within $\ell_1$ distance $1/3$.*[3]
8. *There is $L \in \mathsf{E}$ and $\delta \geq 1/\operatorname{poly}(n)$ such that $L$ cannot be approximated by $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}$ within $\ell_1$ distance $\delta$.*
9. *There is an i.o. $\varepsilon$-PRG $G$ against $\mathcal{C}$ with seed length $n-1$ and error $\varepsilon(n) \leq n^{-\omega(1)}$.*[4]
10. *For each $\gamma > 0$, there is an i.o. $\varepsilon$-PRG against $\mathcal{C}$ with seed length $n^{\gamma}$ and $\varepsilon(n) \leq n^{-\omega(1)}$.*

This result can be applied to a variety of natural circuit classes, such as $\mathsf{AC}^0[\oplus]$, $\mathsf{ACC}^0$, and constant-degree polynomial threshold functions (PTFs). We stress that while Theorem 1 requires the circuit class $\mathcal{C}$ to be contained in $\mathsf{NC}^1$, in circuit complexity this is the most interesting case for the result. More precisely, for circuit classes that are above $\mathsf{NC}^1$, it is well known that worst-case hardness for a problem in $\mathsf{E}$ can be converted into average-case hardness and PRGs. (Furthermore, $\mathsf{NC}^1$ is closed under a top $\mathsf{MAJ}$ or $\widetilde{\mathsf{SUM}}$ gate.) We remark that Theorem 1, with appropriate modifications, can be adapted to other uniform complexity classes, such as $\mathsf{BPE} = \mathsf{BPTIME}[2^{O(n)}]$ and $\mathsf{PSPACE}$. For simplicity, we restrict our discussion to $\mathsf{E}$.

We observe that a connection between worst-case hardness and *weak* average-case hardness for functions in $\mathsf{E}$ has been established in [20], under the assumption that the circuit class $\mathcal{C}$ contains $\mathsf{AC}^0$ and is closed under composition. In contrast to their work, we have a much weaker assumption on $\mathcal{C}$, and our setting of parameters allows us to obtain equivalences to PRGs and to derive consequences that do not follow from their results.

We now highlight three fundamental consequences of Theorem 1. Note that, while our proof employs $\widetilde{\mathsf{SUM}}$ gates in important ways, none of these results refer to such gates.

---

[2] Following standard terminology, a probabilistic $\mathcal{C}$-circuit $F$ is simply a distribution of $\mathcal{C}$-circuits. We say that $F$ computes a Boolean function $g$ with error $\varepsilon$ if for every input $x$ we have $\Pr_F[F(x) \neq g(x)] \leq \varepsilon$.

[3] In other words, there is no family of circuits $F_n \in [0,1]$-$\mathsf{SUM} \circ \mathcal{C}$ such that $\mathsf{E}_{x \sim \{0,1\}^n}[|L(x) - F_n(x)|] \leq 1/3$ for all large $n$. This notion plays a crucial role in [13] and other related works.

[4] More precisely, for each choice of $k$, there is an infinite set $S_k \subseteq \mathbb{N}$ such that $G$ fools circuits from $\mathcal{C}[n^k]$ on inputs of length $n \in S_k$ with error $\varepsilon(n) \leq n^{-k}$.

**1. Seed reduction for PRGs.** Perhaps surprisingly, the equivalence between Items 9 and 10 of Theorem 1 shows the existence of a *generic seed reduction phenomenon* for weak circuit classes. Thus to construct i.o. PRGs of sub-polynomial seed length for a class $\mathcal{C}$ satisfying the conditions of this result it is enough to construct a non-trivial i.o. PRG (i.e. of seed length $\leq n-1$) with small error. In particular, improving the error parameter of the PRG against $\mathsf{AC}^0[\oplus]$ described in [16] to inverse-super-polynomial would lead to major consequences for $\mathsf{AC}^0[\oplus]$-circuits.

**2. Hardness under some distribution implies hardness under the uniform distribution.** Theorem 1 also has important implications to our understanding of the average-case hardness of problems in $\mathsf{E}$ with respect to weak circuit classes. This is an immediate consequence of Items 5 and 6, which establish the result for strong average-case hardness of the form $1/2 + 1/n^{\omega(1)}$. In the full version of this paper [11], we observe that our techniques can also translate constant-error average-case hardness under an arbitrary distribution to constant-error average-case hardness under the uniform distribution. An interesting application of these results is that the existence of a PRG against $\mathcal{C}$, which was only known to imply hardness under some distribution (see e.g. Section 3 of [46]), also implies hardness with respect to the uniform distribution (which in turn is sufficient to construct PRGs).

**3. Equivalence between worst-case and average-case hardness.** The well-known Discriminator Lemma from Hajnal et al. [24] has found numerous applications in circuit complexity lower bounds. It shows that if a Boolean function $f$ cannot be $(1/2+1/\operatorname{poly}(n))$-approximated by a class $\mathcal{C}$ then $f$ is not in $\mathsf{MAJ}\circ\mathcal{C}$. In other words, one can lift an average-case lower bound against $\mathcal{C}$ to a worst-case lower bound against the stronger class $\mathsf{MAJ}\circ\mathcal{C}$. Interestingly, the equivalence between Items 3 and 6 in Theorem 1 shows that, for the purpose of proving lower bounds for a problem in $\mathsf{E}$, a worst-case lower bound against $\mathsf{MAJ}\circ\mathcal{C}$ is actually *equivalent* to a strong average-case lower bound against $\mathcal{C}$. To our knowledge, this was previously unknown for weak computational models.[5]

A consequence of Theorem 1 relevant to the study of $\widetilde{\mathsf{SUM}}$ gates is that if $\mathsf{E} \not\subseteq \widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$ for some $\delta(n) = 1/n^c$ then $\mathsf{E} \not\subseteq \widetilde{\mathsf{SUM}}_{1/3} \circ \mathcal{C}$.[6] Another interesting implication is that the average-case lower bounds against $[0,1]$-$\mathsf{SUM}\circ\mathcal{C}$ under $\ell_1$ distance investigated in [13] are *necessary* and *sufficient* for strong average-case hardness against $\mathcal{C}$.

Next, we discuss some of the techniques behind Theorem 1.

**Theorem 1: Techniques.** As alluded to above, the proof of Theorem 1 relies on non-black-box hardness amplification techniques explored by Chen and Ren [14] and on a careful balance between the *strength* and *weakness* of $\widetilde{\mathsf{SUM}}$ gates. To give some intuition, we discuss the main ingredients behind a more direct proof of the following equivalence, which also explains the assumptions on the circuit class $\mathcal{C}$:

$$\text{Worst-case hardness against } \widetilde{\mathsf{SUM}}\circ\mathcal{C} \iff \text{i.o. PRGs against } \mathcal{C} \text{ with error } \varepsilon = n^{-\omega(1)}.$$

[5] We also remark that it was known [19, 28, 33] before that for general circuit class $\mathcal{C}$, weak average-case hardness against $\mathsf{MAJ}\circ\mathcal{C}$ implies strong average-case hardness against $\mathcal{C}$.

[6] We note that a simple error amplification technique for $\widetilde{\mathsf{SUM}}$ (see [11]) blows up the complexity of the involved $\widetilde{\mathsf{SUM}}\circ\mathcal{C}$-circuits to quasi-polynomial when amplifying from constant-error approximation to inverse polynomial. For this reason, it does not establish this implication.

While it is possible to show that a $\widetilde{\mathsf{SUM}}$ gate can be efficiently simulated by a $\mathsf{MAJ}$ gate,[7] the opposite simulation does not hold (e.g. consider approximate degree). In this sense, $\widetilde{\mathsf{SUM}}$ gates are indeed weak. Still, it is possible to show essentially that, for a certain *specific* $\mathsf{NC}^1$-hard problem $L$ contained in $\mathsf{P}$, a $\widetilde{\mathsf{SUM}}$ gate of polynomial complexity can implement a hardness amplification proof: roughly speaking, a weak approximator circuit for $L$ can be transformed into a correct circuit for $L$ by incurring only a top $\widetilde{\mathsf{SUM}}$ gate overhead. This allows us to employ the following win-win analysis. Either the $\mathsf{NC}^1$-hard problem $L$ is $1/2 + n^{-k}$-hard against $\mathcal{C}$ on infinitely many input lengths for every choice of $k$, in which case an i.o. PRG against $\mathcal{C}$ can be constructed from $L$ using standard techniques under the assumption that $\mathcal{C}$ is closed under bottom layer $O(1)$-juntas, or there is a choice of $k$ such that $L$ can be $1/2 + n^{-k}$ approximated by $\mathcal{C}$-circuits on large enough input lengths. The latter implies via the hardness amplification reconstruction routine that $L \in \widetilde{\mathsf{SUM}} \circ \mathcal{C}$, which in turns yields $\mathsf{NC}^1 \subseteq \widetilde{\mathsf{SUM}} \circ \mathcal{C}$ using the $\mathsf{NC}^1$-hardness of $L$ (which in fact admits ultra efficient reductions). Now under our assumption that $\mathcal{C} \subseteq \mathsf{NC}^1$, it is easy to see that $\mathsf{NC}^1 = \widetilde{\mathsf{SUM}} \circ \mathcal{C}$. As a consequence, a worst-case lower bound against $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$ provides a worst-case lower bound against $\mathsf{NC}^1$, and again, PRGs can be constructed from such an assumption via standard methods (since $\mathsf{NC}^1$ admits black-box worst-case to average-case amplification).

For the other direction, we start with an i.o. PRG $G$ against $\mathcal{C}$ that might have a large seed length but guarantees *low error* $\varepsilon(n) = n^{-\omega(1)}$. Here the important insight is that a low error PRG that fools $\mathcal{C}$ also fools linear combinations of functions in $\mathcal{C}$ with bounded coefficients. This implies that $G$ fools $\widetilde{\mathsf{SUM}} \circ C$. Another standard argument shows that from such a PRG one can define a function in $E$ that is worst-case hard against $\widetilde{\mathsf{SUM}} \circ C$.

We stress that two crucial ingredients of our equivalence theorem are the existence of the hard problem $L$ mentioned above and the use of $\widetilde{\mathsf{SUM}}$ gates. The hard language $L$ is actually a pair of problems $\mathsf{CMD}$ and $\mathsf{DCMD}$ with very useful structural properties (see Section 2.2). They have been explored in a few other works (e.g. [31, 4, 20, 1]), and are tightly connected to *decomposable randomized encodings*, which are well-studied in cryptography (see [3]). The fruitful interaction between these problems and $\widetilde{\mathsf{SUM}}$ gates was first noticed by [14] in the context of the algorithmic method and is a crucial ingredient in their proof that $\mathsf{NQP}$ is strongly average-case hard against $\mathsf{ACC}^0$.

While the proof of Theorem 1 avoids the black-box "barrier" and applies to circuit classes that are not assumed to be closed under majority, our techniques come with certain limitations. As a consequence of our indirect analysis via a win-win argument, Theorem 1 does not provide almost-everywhere equivalences for some items and does not scale to large circuit size bounds above quasi-polynomial. These are important directions for future work.

**Applications to $\mathsf{ACC}^0$-circuits lower bound approaches.** As a concrete application of Theorem 1 to current frontiers in circuit complexity, we explore its consequences to the average-case complexity of $\mathsf{ACC}^0$. We use our framework to show that existing "combinatorial" approaches to worst-case lower bounds would also provide *strong average-case hardness* against $\mathsf{ACC}^0$. Before stating this result, we briefly recall some concepts.

[7] It is possible to approximate all coefficients of the bounded linear sum using sums of powers of $2^i$ with $i \in \mathbb{Z}$, then multiply the linear sum by an appropriate power of 2 to obtain integer coefficients, and finally simulate the resulting sum by an appropriate $\mathsf{THR}$ gate with polynomial weights, which can be translated to a $\mathsf{MAJ}$ gate using duplicated input wires and by negating input variables if necessary.

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the one-dimensional torus. A *torus polynomial* [7] (see also [34]) is a real polynomial $p(x_1, \ldots, x_n)$ restricted to the domain $\{0,1\}^n$ and evaluated modulo one.[8] For the purpose of representing the output of a Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ as a value in $\mathbb{T}$, we map the output bit $f(x)$ to $f(x)/2$. For $\delta < 1/4$, we say that $f$ is $\delta$-approximated by a degree-$d$ torus polynomial if there is a degree-$d$ real polynomial $p(x_1, \ldots, x_n)$ such that if $f(x) = 1$ then $p(x) - \lfloor p(x) \rfloor \in [1/2 - \delta, 1/2 + \delta]$ and if $f(x) = 0$ then $p(x) - \lfloor p(x) \rfloor \in [0, \delta] \cup [1 - \delta, 1)$. A recent approach proposed by [7] shows that $\mathsf{ACC}^0$ lower bounds follow from torus polynomial degree lower bounds for approximating a Boolean function.

The number-on-forehead (NOF) multi-party communication model was introduced by [8], and work of [26, 41] show that explicit communication lower bounds in this model (even in the single-round model where all players simultaneously communicate to a referee) imply lower bounds against $\mathsf{SYM}^+$-circuits, which are known to simulate $\mathsf{ACC}^0$ [6].

▶ **Theorem 2** (Lifting worst-case $\mathsf{ACC}^0$ lower bound approaches to strong correlation bounds)**.** *Consider the following statements:*

1. **Torus Polynomials:** *There is a language $L \in \mathsf{E}$ and a function $\delta(n) \geq 1/\operatorname{poly}(n)$ such that $L$ does not have $\delta$-approximation torus polynomials of degree* $\operatorname{polylog}(n)$.
2. **NOF Protocols:** *There is a language in $\mathsf{E}$ that does not admit* (*single-round*) NOF *multi-party protocols with* $\operatorname{polylog}(n)$ *parties of communication cost* $\operatorname{polylog}(n)$.

*In each case, if the corresponding statement holds then there is a language in $\mathsf{E}$ that cannot be $(1/2 + 1/\operatorname{poly}(n))$-approximated under the uniform distribution by $\mathsf{ACC}^0$.*

As a consequence, lower bounds against these models provide i.o. PRGs of sub-polynomial seed length against $\mathsf{ACC}^0$.

**Theorem 2: Techniques.** It is not hard to adapt classical techniques to show that if a Boolean function can be approximated by torus polynomials of bounded degree, then it can also be computed by NOF protocols of low complexity. For this reason, in order to prove Theorem 2 it is sufficient to obtain average-case hardness against $\mathsf{ACC}^0$ from degree lower bounds for torus polynomials approximating Boolean functions.[9] To achieve this, we refine the argument of [7] and invoke our framework. In more detail, we show the stronger result that even functions families in $\widetilde{\mathsf{SUM}} \circ \mathsf{ACC}^0$ can be approximated by low-degree torus polynomials. This yields the result using the equivalence between Items 6 and 2 in Theorem 1.

To establish this claim, we make use of low degree "middle-bit polynomials" [21], a sub-class of $\mathsf{SYM}^+$-circuits that is strong enough to simulate $\mathsf{ACC}^0$. By a careful adaptation of the argument of [7], we are able to show that a linear sum (with bounded coefficients) of middle-bit polynomials with a special structure can be converted into a torus polynomial. The argument is somewhat subtle, and involves the manipulation of universal circuits for depth-$d$ $\mathsf{ACC}^0[s]$ in order to enforce similar parameters for all middle-bit polynomials feeding the top $\widetilde{\mathsf{SUM}}$ gate. The details appear in the full version of this paper [11].

[8] By a value $y \pmod 1$ we mean its fractional part given by $y - \lfloor y \rfloor$, where the floor function $\lfloor y \rfloor$ denotes the largest integer less than or equal to $y$. For instance, $1.37 \pmod 1$ is $0.37$ and $-2.21 \pmod 1$ is $0.79$.

[9] Alternatively, earlier work on $\mathsf{ACC}^0$ already showed that $\mathsf{MAJ} \circ \mathsf{ACC}^0$-circuits can be simulated by NOF protocols of low communication. Therefore, the NOF protocols part of Theorem 2 follows directly from our Theorem 1.

### 1.2.2 Lower bounds against circuits with layers of $\widetilde{\mathsf{SUM}}$ and MAJ gates

Observe that Theorem 1 (via Items 1, 2, and 3) establishes the following equivalence: for the purpose of proving circuit lower bounds for a function in $\mathsf{E}$, a top layer $\mathsf{MAJ}$ gate is *equivalent* to a top layer $\widetilde{\mathsf{SUM}}$ gate. Given that $\widetilde{\mathsf{SUM}}$ is simpler than $\mathsf{MAJ}$, and lower bounds against $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$ offer a path to correlation bounds and PRGs against $\mathcal{C}$, obtaining a better understanding of $\widetilde{\mathsf{SUM}}$ gates in Boolean circuits might have significant benefits.

In this section, we explore *unconditional* lower bounds against circuits with layers of $\mathsf{MAJ}$ and $\widetilde{\mathsf{SUM}}$ gates. Our results are connected to the long-standing problem of showing explicit lower bounds against $\mathsf{THR} \circ \mathsf{THR}$, the class of polynomial-size depth-2 threshold circuits (where size is measured by number of gates). For convenience of the reader, we review below some results related to this frontier.

**Threshold circuits.** Recall that a threshold gate $\mathsf{THR}$ over $m$ input bits is described by weights $w_1, \ldots, w_m, \theta \in \mathbb{R}$. It outputs 1 on an input $y \in \{0,1\}^m$ if and only if $\sum_i w_i y_i \geq \theta$. It is known that every such gate can be implemented with integer weights of magnitude $2^{O(m \log m)}$ (see [25]). In the context of polynomial size circuits, by duplicating input wires a $\mathsf{MAJ}$ gate can be equivalently defined as the restriction of a $\mathsf{THR}$ gate to polynomially bounded integer weights. It was shown that $\mathsf{MAJ} \circ \mathsf{THR} = \mathsf{MAJ} \circ \mathsf{MAJ}$ and $\mathsf{THR} \circ \mathsf{THR}$ is contained in $\mathsf{MAJ} \circ \mathsf{MAJ} \circ \mathsf{MAJ}$ [18]. Exponential lower bounds are known against $\mathsf{THR} \circ \mathsf{MAJ}$-circuits [17], and $\mathsf{THR} \circ \mathsf{MAJ}$ is strictly contained in $\mathsf{THR} \circ \mathsf{THR}$ [9]. Recently, [32] described a function in $\mathsf{P}$ that requires $\mathsf{THR} \circ \mathsf{THR}$-circuits of size (measured by the number of gates) nearly $n^{3/2}$. This is the strongest known lower bound against this class (see their work for extensions to other circuit size measures) for a function in $\mathsf{P}$. It is also known that $\mathsf{E}^{\mathsf{NP}}$ does not have $n^{2-\varepsilon}$-size $\mathsf{THR} \circ \mathsf{THR}$-circuits for every constant $\varepsilon > 0$ [2, 44].

**$\mathsf{LTF}^s \circ \mathcal{C}$-circuits: An intermediary class between $\mathsf{MAJ} \circ \mathcal{C}$ and $\mathsf{THR} \circ \mathcal{C}$.** In order to make progress toward showing super-polynomial lower bounds against $\mathsf{THR} \circ \mathsf{THR}$-circuits, we study a newly defined gate $\mathsf{LTF}^s$ whose power lies between $\mathsf{MAJ}$ and $\mathsf{THR}$.[10] Let $\mathsf{SUM}^\infty \circ \mathcal{C}$ be the relaxation of $\mathsf{SUM} \circ \mathcal{C}$ to an *unrestricted* top $\mathsf{SUM}$ gate (i.e. the top gate can use arbitrary real coefficients that might not be polynomially bounded). For a given function $s$ and a circuit class $\mathcal{C}$, we say that a function $f$ admits a $\mathsf{LTF}^s \circ \mathcal{C}$-circuit of size $S$ if there is a circuit $D \in \mathsf{SUM}^\infty \circ \mathcal{C}$ such that the following hold: (1) $f(x) = 1$ if and only if $D(x) \geq 0$; (2) $|D(x)| \in (1/s, s)$ for every $x \in \{0,1\}^n$; (3) the total size of the $\mathcal{C}$-subcircuits of $D$ is at most $S$. Note that unrestricted weights are allowed in the top gate, but we are promised that on each input $x$ the value $D(x)$ is neither too close to 0 nor too large in magnitude.[11]

We are able to extend the algorithmic method [49] to show that #SAT algorithms for a circuit class $\mathcal{C}$ imply worst-case lower bounds against $\mathsf{LTF}^s \circ \mathcal{C}$ and average-case lower bounds against $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$. Let $\mathsf{NQP} = \mathsf{NTIME}[2^{\mathrm{polylog}(n)}]$ be the class of languages computable in non-deterministic quasi-polynomial time. We say that a circuit class $\mathcal{C}$ is *nice* if $\mathcal{C}$ is closed under negation, (bottom) projections, and a top $\mathsf{AND}$ gate of unbounded fan-in, and in addition $\mathcal{C}$-circuits of size $s$ admit general circuits of depth $O(\log s)$. Examples of nice circuit classes include $\mathsf{AC}^0$, $\mathsf{ACC}^0$, and $\mathsf{AC}^0[\oplus] \circ \mathsf{THR}$.

---

[10] $\mathsf{LTF}$ denotes linear threshold function, another standard name for $\mathsf{THR}$. We employ both names in this paper to make a clear distinction between the new gates and $\mathsf{THR}$.

[11] Note that we only impose this constraint for each input $x$ of the combined $\mathsf{SUM}^\infty \circ \mathcal{C}$-circuit, and not over all possible input strings for the top gate.

▶ **Theorem 3** (Stronger lower bounds from #SAT algorithms)**.** *Let $\mathcal{C}$ be a nice circuit class. Suppose there is a constant $\varepsilon > 0$ such that, given a $\mathcal{C}$-circuit of size $2^{n^{\varepsilon}}$ over $n$ input variables, its number of satisfying assignments can be deterministically computed in time $2^{n-n^{\varepsilon}}$. Then the following statements hold:*

1. *For every constant $k > 0$, $\mathsf{NQP}$ does not have $\mathsf{LTF}^{2^{\log^k n}} \circ \mathcal{C}$-circuits of size $2^{\log^k n}$.*
2. *For every choice of constants $k > 0$ and $\delta \in (0, 0.5)$, $\mathsf{NQP}$ cannot be $(1/2 + 2^{-\log^k n})$-approximated by $\widetilde{\mathsf{SUM}}_{\delta} \circ \mathcal{C}$-circuits where both the sparsity of the top $\mathsf{SUM}$-gate and the size of the bottom layer $\mathcal{C}$-circuits are at most $2^{\log^k n}$.*[12]

To our knowledge, these two circuit lower bound consequences are incomparable. By combining Theorem 3 with existing #SAT algorithms for $\mathcal{C} = \mathsf{ACC}^0 \circ \mathsf{THR}$-circuits [51], we obtain the following unconditional lower bounds.

▶ **Corollary 4** (Lower bounds against circuits with $\widetilde{\mathsf{SUM}}$, $\mathsf{THR}$, and $\mathsf{MAJ}$ gates)**.** *The following results hold:*

1. *For every constant $k > 0$, $\mathsf{NQP}$ does not admit $\mathsf{LTF}^{2^{\log^k n}} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$-circuits of size $2^{\log^k n}$.*
2. *For every choice of constants $k > 0$ and $\delta \in (0, 0.5)$, $\mathsf{NQP}$ cannot be $(1/2 + 2^{-\log^k n})$-approximated by $\widetilde{\mathsf{SUM}}_{\delta} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$-circuits where the top sum has sparsity $2^{\log^k n}$ and all $\mathsf{ACC}^0 \circ \mathsf{THR}$-subcircuits have size $2^{\log^k n}$.*
3. *For every choice of constants $k > 0$ and $\delta \in (0, 0.5)$, $\mathsf{NQP}$ cannot be computed by $\mathsf{MAJ} \circ \widetilde{\mathsf{SUM}}_{\delta} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$-circuits where the top $\mathsf{MAJ}$ gate has fan-in $2^{\log^k n}$ and all $\widetilde{\mathsf{SUM}}_{\delta} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$-subcircuits have size and sparsity $2^{\log^k n}$.*

To contrast these results with previous work, we note that [15, Theorem 15] gave a *worst-case* lower bound against $\widetilde{\mathsf{SUM}}_{\delta} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$-circuits with any *constant* error $\delta$ less than $1/2$. Also, [14, Section 5.2] showed a strong average-case lower bound against $\widetilde{\mathsf{SUM}}_{\delta} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$-circuits, where the top sum gate has *zero error* (i.e., $\delta = 0$). Consequently, Corollary 4 Item 2 simultaneously strengthens both results. On the other hand, Corollary 4 Item 3 shows the first lower bound against circuits combining layers of $\widetilde{\mathsf{SUM}}_{1/3}$, $\mathsf{MAJ}$, and $\mathsf{THR}$ gates.

Before discussing our techniques in more detail, we mention an open problem and its connection to $\mathsf{THR} \circ \mathsf{THR}$ lower bounds. Recall that this class is contained in $\mathsf{MAJ} \circ \mathsf{MAJ} \circ \mathsf{MAJ}$. In light of the super-polynomial lower bound against $\mathsf{MAJ} \circ \widetilde{\mathsf{SUM}}_{\delta} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$ from Corollary 4 Item 3, it would be very interesting to understand the relation between $\mathsf{MAJ}$ gates and $\widetilde{\mathsf{SUM}}$ gates appearing in *internal layers* of Boolean circuits. In particular, we note that if $\mathsf{MAJ}$ can be simulated by $\widetilde{\mathsf{SUM}}_{1/3} \circ \mathsf{ACC}^0$-circuits of quasi-polynomial size (or $\mathsf{THR}$ can be simulated by $\mathsf{MAJ} \circ \widetilde{\mathsf{SUM}}_{\delta} \circ \mathsf{ACC}^0$-circuits of quasi-polynomial size), then $\mathsf{NQP} \not\subseteq \mathsf{THR} \circ \mathsf{THR}$. On the other hand, if this is not the case, strong average-case lower bounds against $\mathsf{ACC}^0$ follow from Theorem 1.

**Theorem 3 and Corollary 4: Techniques.** The proofs of the first two items of Corollary 4 are immediate from the corresponding items of Theorem 3 via the #SAT algorithm for $\mathcal{C} = \mathsf{ACC}^0 \circ \mathsf{THR}$ given by [51]. On the other hand, Item 3 of Corollary 4 can be established in different ways. The first proof is just a standard application of the Discriminator Lemma [24] together with the lower bound from Item 2. A second proof follows from Item 1, via a

[12] For the interested reader, we notice that the coefficients of the top $\widetilde{\mathsf{SUM}}$ gate can be unbounded in this lower bound.

simulation of a $\mathsf{MAJ} \circ \widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-circuit of quasi-polynomial complexity by a $\mathsf{LTF}^{2^{\log^k n}} \circ \mathsf{ACC}^0 \circ \mathsf{THR}$-circuit of size size $2^{\log^k n}$, for some constant $k$. This can be done by first reducing the error $\delta$ of each $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-subcircuit (see [11]), then rewriting the corresponding $\mathsf{MAJ} \circ \widetilde{\mathsf{SUM}}_\varepsilon$ top layers as an $\mathsf{LTF}^s$ gate via an appropriate collapse. We omit the details.

The proofs of Items 1 and 2 of Theorem 3 are essentially independent. We discuss each of them next, starting with Item 1.

An extension of the algorithmic method [49] obtained by [37] shows that SAT algorithms for a circuit class $\mathcal{C}$ of sub-exponential size circuits (satisfying minor closure conditions) that run in time $2^{n-n^\varepsilon}$ imply that $\mathsf{NQP} \not\subseteq \mathcal{C}$. In a more recent work that builds on [50], [15] established (in particular) that #SAT algorithms of similar running time provide the stronger lower bound $\mathsf{NQP} \not\subseteq \widetilde{\mathsf{SUM}} \circ \mathcal{C}$. Our proof of Item 1 of Theorem 3 relies on the latter result and on a win-win argument inspired by [14]. In more detail, and oversimplifying a bit, we argue that if a special $\mathsf{NC}^1$-hard problem $L$ (contained in $\mathsf{NQP}$) is not in $\mathsf{LTF}^{2^{\log^k n}} \circ \mathcal{C}$, then we are done. Otherwise, we explore $\mathsf{LTF}^s$ gates and the special form of the $\mathsf{NC}^1$-hardness of $L$ to show that $\mathsf{NC}^1$ can be simulated by $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$-circuits of quasi-polynomial complexity. Given this lemma and the corresponding simulation, we can reduce the derivation of the desired lower bound to previous work, i.e., we invoke the aforementioned connection between #SAT algorithms and lower bounds against $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$. This provides a language in $\mathsf{NQP}$ that is not in $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$ of complexity $2^{\log^\ell n}$, where $\ell = \ell(k)$ is large enough. Now by simulating $\mathsf{LTF}^{2^{\log^k n}} \circ \mathcal{C}$-circuits using quasi-polynomial size Boolean formulas, and using the collapse of $\mathsf{NC}^1$ to quasi-polynomial size $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$, it is possible to argue that $L$ is also hard against $\mathsf{LTF}^{2^{\log^k n}} \circ \mathcal{C}$.

The proof of Item 2 of Theorem 3 shares some similarities with the argument above, but the technical details are different. From a high-level perspective, we also employ a win-win argument, though this time it is based on the *average-case* complexity of the language $L$ mentioned above. Moreover, we cannot rely on previous connections between #SAT algorithms and lower bounds in a *black-box* way. Given that explaining the relevant details would be fairly technical, we refer the interested reader to the full version of this paper [11]. We mention that a conceptual contribution is that while our proof of Theorem 3 Part 2 follows the strategy of previous works, such as [10, 15, 14], on obtaining lower bounds from meta-algorithms, it does not use PCPs of proximity (PCPP), which was a key ingredient in the proofs of those works. For this, we rely in part on a PCP stated in [48], combined with other ideas.

**Organization.** In Section 2 we introduce the necessary technical preliminaries for proving our results. In Section 3 we prove our main equivalence result (Theorem 1). Due to space constraints, the remaining proofs are deferred to the full version of this paper [11].

# 2 Preliminaries

## 2.1 Notation

We use $\mathbb{N}$ to denote the set of all non-negative integers and $\mathbb{N}_{\geq 1}$ to denote $\mathbb{N} \setminus \{0\}$. For every $n \in \mathbb{N}_{\geq 1}$, we let $\mathcal{U}_n$ denote the uniform distribution over $\{0,1\}^n$. For convenience, in some settings a Boolean function $f \colon \{0,1\}^n \to \{0,1\}$ will be viewed as a function with output in $\{-1,1\}$, where $-1$ and $1$ are interpreted as True and False, respectively.

For a predicate $P(x)$, we use $\mathbb{1}_{P(x)}$ to denote its corresponding Boolean value on $x$. That is, $\mathbb{1}_{P(x)} = 1$ if $P(x)$ is true, and 0 otherwise. For a real $v$, we define $\operatorname{sign}(v) := (-1) \cdot \mathbb{1}_{v<0} + 1 \cdot \mathbb{1}_{v \geq 0}$.

For two strings $\alpha, \beta \in \{0,1\}^*$, we write $\alpha \circ \beta$ to denote the concatenation of $\alpha$ and $\beta$.

A projection of a function $f(x_1, \ldots, x_n)$ is a function $g(y_1, \ldots, y_m)$ with a projection mapping $P\colon \{0,1\}^m \to \{0,1\}^n$ such that $g(y_1, \ldots, y_m) = f(P(y_1, \ldots, y_m))$. By "projection" we mean that each output bit of $P(y_1, \ldots, y_m)$ is either an input bit $y_i$, its negation, or a constant.

Let $a$ be a positive integer. For an arbitrary $\ell \geq 1$ and a function $h\colon \{0,1\}^\ell \to \{0,1\}$, we say that $h \in \mathsf{JUNTA}_a$ if the output of $h$ depends on at most $a$ input coordinates.

For a circuit class $\mathcal{C}$ and $s \geq 1$, we use $\widetilde{\mathsf{SUM}} \circ \mathcal{C}[s]$ to denote the class of $\widetilde{\mathsf{SUM}} \circ \mathcal{C}$-circuits where the top $\mathsf{SUM}$ gate has complexity at most $s$ and the bottom layer $\mathcal{C}$-circuits have size at most $s$.

For a function $f\colon \{0,1\}^n \to \{0,1\}$, we let $f_\pm\colon \{0,1\}^n \to \{-1,1\}$ be the $\{-1,1\}$-version of $f$ where we map the output of $f$ from 0 to 1 and 1 to $-1$. Also, for a circuit class $\mathcal{C}$ where the circuits in $\mathcal{C}$ output values in $\{0,1\}$, we denote by $\mathcal{C}_\pm$ the $\{-1,1\}$-version of $\mathcal{C}$ where the circuits in $\mathcal{C}_\pm$ output values in $\{-1,1\}$.

## 2.2 A $\oplus$L-complete problem with good properties

The existence of $\oplus$L-complete problems with good reducibility properties will be important for us. (Recall that $\oplus$L is the class of problems solvable by a nondeterministic logspace Turing machine that accepts the input if the number of accepting paths is odd.) We define the following two problems, called Connected Matrix Determinant (CMD) and Decomposed Connected Matrix Determinant (DCMD):

▶ **Definition 5.** *An instance of* CMD *is an $n \times n$ matrix over $\mathbb{F}_2$ where the main diagonal and above may contain either 0 or 1, the second diagonal (i.e. the one below the main diagonal) contains 1, and other entries are 0. In other words, the matrix is of the following form (where $*$ represents any element in $\mathbb{F}_2$):*

$$\begin{pmatrix} * & * & * & \cdots & * & * \\ 1 & * & * & \cdots & * & * \\ 0 & 1 & * & \cdots & * & * \\ 0 & 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & * \end{pmatrix}.$$

*The instance is an $(n(n+1)/2)$-bit string specifying elements on and above the main diagonal. We define $x \in \mathsf{CMD}$ if and only if the determinant (over $\mathbb{F}_2$) of the matrix corresponding to $x$ is 1.*

*An instance of* DCMD *is a string of length $n^3(n+1)/2$. For an input $x$, $\mathsf{DCMD}(x)$ is computed as follows: we partition $x$ into blocks of length $n^2$, let $y_i$ $(1 \leq i \leq n(n+1)/2)$ be the parity of the $i$-th block, and define $\mathsf{DCMD}(x) := \mathsf{CMD}(y_1 \circ y_2 \circ \cdots \circ y_{n(n+1)/2})$.*

The precise definitions of CMD and DCMD are not important here, as we only need the following two important results about them.

▶ **Theorem 6** ([4, 20]). *There is a function $P : \{0,1\}^{n(n+1)/2} \times \{0,1\}^{O(n^4)} \to \{0,1\}^{n^3(n+1)/2}$ such that the following hold.*

- *For any input $x \in \{0,1\}^{n(n+1)/2}$, the random variable $P(x, \mathcal{U}_{O(n^4)})$ is uniformly distributed in $\{0,1\}^{n^3(n+1)/2}$.*
- *For any $x \in \{0,1\}^{n(n+1)/2}$ and $r \in \{0,1\}^{O(n^4)}$, let $P(x,r) = y$, then $\mathsf{CMD}(x) = \mathsf{DCMD}(y) \oplus r_0$, where $r_0$ is the first bit of $r$.*
- *For each fixed randomness $r$, $P(x,r)$ is a projection over $x$, computable in polynomial time given $r$.*

▶ **Theorem 7** ([31]). CMD *is* $\oplus$L*-complete under projections.*

Observe that if CMD is in a circuit class $\mathcal{C}$ closed under projections then all problems in (non-uniform) $\mathsf{NC}^1$ are also in $\mathcal{C}$, given that the problem of evaluating an input Boolean formula is solvable with logarithmic space.

We refer the reader to the full version of [14] for a self-contained exposition of these problems and their relevant properties, including pointers to related work.

## 2.3 Pseudorandomness

We need the following Hardness vs. Randomness framework for constructing PRGs.

▶ **Lemma 8** (Hardness vs. Randomness [39], see also [14, Appendix E.3] for the proof). *There is a function* $G\colon \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ *such that the following holds. Let* $n, \ell, a$ *be integers such that* $a \le \ell$, *and* $t = O\big(\ell^2 \cdot n^{1/a}/a\big)$. *Let* $\mathcal{C}$ *be a circuit class closed under negation. For any function* $Y\colon \{0,1\}^\ell \to \{0,1\}$ *represented as a length-*$2^\ell$ *truth table, if* $Y$ *cannot be* $(1/2 + \varepsilon/n)$*-approximated by* $\mathcal{C} \circ \mathsf{JUNTA}_a$*-circuits where the top circuit has size* $S$, *then for every circuit* $C \in \mathcal{C}$ *of size* $S$,

$$\left| \Pr_{z\sim\{0,1\}^t}[C(G(Y,z)) = 1] - \Pr_{x\sim\{0,1\}^n}[C(x) = 1] \right| \le \varepsilon.$$

*Moreover, the function* $G$ *is computable in* $\mathrm{poly}(n, 2^t)$ *time.*

The following simple fact says PRGs imply worst-case hardness.

▶ **Proposition 9** (Worst-case hardness from PRGs). *Let* $\mathcal{F}$ *be a class of functions. If there is an i.o.* $\varepsilon$*-PRG* $G\colon \{0,1\}^r \to \{0,1\}^n$ *with seed length* $r(n)$ *against* $\mathcal{F}_n$, *where* $\varepsilon < 1 - 2^{r(n)-n}$, *then there is a language* $L \in E$ *such that* $L$ *cannot be computed by* $\mathcal{F}$.

**Proof.** Please see the full version [11] for details. ◀

## 2.4 Hardness amplification

The following result allows us to amplify hardness against $\mathsf{NC}^1$.

▶ **Lemma 10** (Hardness amplification against $\mathsf{NC}^1$, see e.g. [43, 20]). *Suppose there is a language* $L \in \mathsf{E}$ *such that* $L \notin \mathsf{NC}^1$. *Then there is a language* $L' \in \mathsf{E}$ *such that for every constant* $k \ge 1$, $L'$ *cannot be* $(1/2 + 1/n^k)$*-approximated by formulas of size* $n^k$.

The following notion of $\ell_1$-approximation by SUM-circuits plays a crucial role in some recent results on average-case lower bounds via the algorithmic method (e.g. [13, 12, 27]).

▶ **Definition 11** ($\ell_1$-approximation by SUM-circuits). *Let* $\delta \in (0,1)$ *and let* $\mathcal{C}$ *be a circuit class. We say that a function* $f\colon \{0,1\}^n \to \{0,1\}$ *is approximated by a* $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}$*-circuit* $C$ *within* $\ell_1$ *distance* $\delta$ *if*

$$\mathop{\mathbf{E}}_{x\sim\mathcal{U}_n}[\,|f(x) - C(x)|\,] \le \delta.$$

For functions $f, g\colon \{0,1\}^n \to \mathbb{R}$, we let $\langle f, g\rangle := \mathbf{E}_{x\in\{0,1\}^n}[f(x)\cdot g(x)]$.

▶ **Proposition 12.** *Let* $\delta \in (0,1)$, $f\colon \{0,1\}^n \to \{0,1\}$, *and* $\mathcal{C}$ *be a circuit class.*

1. *If* $f$ *can be approximated by* $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}$*-circuits of complexity* $s$ *within* $\ell_1$ *distance* $\delta$, *then there is a* $\mathsf{SUM} \circ \mathcal{C}_\pm$*-circuit* $C$ *of complexity* $O(s)$ *such that* $\|C\|_\infty \le 1$ *and* $\langle f_\pm, C\rangle \ge 1 - 2\delta$.

2. *If there is a* $\mathsf{SUM} \circ \mathcal{C}_\pm$*-circuit* $C$ *of complexity* $s$ *such that* $\|C\|_\infty \le 1$ *and* $\langle f_\pm, C\rangle \ge 1 - 2\delta$*, then* $f$ *can be approximated by* $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}$*-circuits of complexity* $O(s)$ *within* $\ell_1$ *distance* $\delta$*.*

**Proof.** Please see the full version [11] for details. ◀

Given a set $X$ and a Boolean function $f\colon X \to \{-1,1\}$, for and integer $t \ge 1$ and $X^t = X \times \ldots \times X$ ($t$ times) we let $f^{\oplus t}\colon X^t \to \{-1,1\}$ be the Boolean function defined as $f^{\oplus t}(x_1,\ldots,x_t) \coloneqq \prod_{i\in[t]} f(x_i)$. We will need the following XOR lemma from [13].

▶ **Theorem 13** ([35] and [13, Lemma 3.8], see also [12, Lemma 1.7])**.** *Let* $\mathcal{F}$ *be a class of Boolean functions that is closed under negation and restriction. For every* $\delta, \varepsilon \in (0,1)$ *and every function* $f\colon \{0,1\}^n \to \{-1,1\}$*, if*

$$\langle f, C\rangle \le 1 - \delta$$

*for every* $\mathsf{SUM} \circ \mathcal{F}$*-circuit* $C$ *where the top* $\mathsf{SUM}$ *has complexity* $10 \cdot n/\varepsilon^2$ *and* $\|C\|_\infty \le 1$*, then*

$$\langle f^{\oplus t}, D\rangle \le (1-\delta)^t + \varepsilon/\delta$$

*for any Boolean function* $D\colon \{0,1\}^{tn} \to \{-1,1\}$ *in* $\mathcal{F}$*.*

## 3 Equivalences for worst-case and strong average-case lower bounds

In this section, we prove our equivalence results for worst-case hardness, strong average-case hardness and pseudorandomness.

▶ **Reminder of Theorem 1.** *Let* $\mathcal{C}$ *be a circuit class that satisfies the following:*

- $\mathcal{C}$ *is closed under negation and projection.*
- $\mathcal{C}$ *is closed under a bottom layer of juntas over* $O(1)$ *input bits. That is*

$$\bigcup_{k\ge1} \mathcal{C}[n^k] \circ \mathsf{JUNTA}_k \subseteq \bigcup_{k\ge1} \mathcal{C}[n^k].$$

- $\bigcup_{k\ge1} \mathcal{C}[n^k] \subseteq \mathsf{NC}^1$.

*Then the following statements are equivalent:*

1. *There is* $L \in \mathsf{E}$ *such that for every* $k \ge 1$*,* $L \notin \widetilde{\mathsf{SUM}}_{1/3} \circ \mathcal{C}[n^k]$*.*
2. *There is* $L \in \mathsf{E}$ *and* $\delta \ge 1/\mathsf{poly}(n)$ *such that for every* $k \ge 1$*,* $L \notin \widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}[n^k]$*.*
3. *There is* $L \in \mathsf{E}$ *such that for every* $k \ge 1$*,* $L \notin \mathsf{MAJ} \circ \mathcal{C}[n^k]$*.*
4. *There is* $L \in \mathsf{E}$ *such that, for every* $k \ge 1$*,* $L$ *cannot be computed by a probabilistic* $\mathcal{C}[n^k]$*-circuit with error* $1/2 - 1/n^k$*.*
5. *There is* $L \in \mathsf{E}$ *and a distribution* $\mathcal{D}$ *such that for every* $k \ge 1$*,* $L$ *cannot be* $(1/2 + n^{-k})$*-approximated by* $\mathcal{C}[n^k]$ *under* $\mathcal{D}$*.*
6. *There is* $L \in \mathsf{E}$ *such that for every* $k \ge 1$*,* $L$ *cannot be* $(1/2 + n^{-k})$*-approximated by* $\mathcal{C}[n^k]$ *under the uniform distribution.*
7. *There is* $L \in \mathsf{E}$ *such that for every* $k \ge 1$*,* $L$ *cannot be approximated by* $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}[n^k]$ *within* $\ell_1$ *distance* $1/3$*.*
8. *There is* $L \in \mathsf{E}$ *and* $\delta \ge 1/\operatorname{poly}(n)$ *such that for every* $k \ge 1$*,* $L$ *cannot be approximated by* $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}[n^k]$ *within* $\ell_1$ *distance* $\delta$*.*

9. *There is an i.o. $\varepsilon$-PRG $G$ against $\mathcal{C}$ with seed length $n-1$ and error $\varepsilon(n) \le n^{-\omega(1)}$. In other words, for each choice of $k$, there is an infinite set $S_k \subseteq \mathbb{N}$ such that $G$ fools circuits from $\mathcal{C}[n^k]$ on inputs of length $n \in S_k$ with error $\varepsilon(n) \le n^{-k}$.*
10. *For every $\gamma > 0$, there is an i.o. $\varepsilon$-PRG against $\mathcal{C}$ with seed length $n^\gamma$ and $\varepsilon(n) \le n^{-\omega(1)}$.*

**Proof.** We will first show Item 2 $\Rightarrow$ Item 6 $\Rightarrow$ Item 10 $\Rightarrow$ Item 1 $\Rightarrow$ Item 2, establishing the equivalence of Items 1, 2, 6, and 10. We then show Item 6 $\Rightarrow$ Item 5 $\Rightarrow$ Item 4 $\Rightarrow$ Item 1, which adds Items 4 and 5 to the list of equivalent items. Next, we show Item 6 $\Rightarrow$ Item 3 $\Rightarrow$ Item 4, which adds Item 3, and Item 10 $\Rightarrow$ Item 9 $\Rightarrow$ Item 2, which adds Item 9. Finally, we show Item 6 $\Rightarrow$ Item 7 $\Rightarrow$ Item 8 $\Rightarrow$ Item 6, adding Items 7 and 8 to the list and completing the proof.

**Item 2 $\Rightarrow$ Item 6.** We consider two cases. If $\mathsf{DCMD}$ cannot be $(1/2 + 1/n^k)$-approximated by $\mathcal{C}[n^k]$ for every $k \ge 1$ under the uniform distribution, then we are done.

Now consider the case that there is some $k \ge 1$ such that $\mathsf{DCMD}$ can be $(1/2 + 1/n^k)$-approximated by $\mathcal{C}[n^k]$. By the random self-reducibility of $\mathsf{DCMD}/\mathsf{CMD}$ (see Theorem 6 and also [14, Section 3]), for any $\delta = 1/\operatorname{poly}(n)$, $\mathsf{CMD}$ can be computed by a $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-circuit where the top $\mathsf{SUM}$-gate has polynomial complexity and the bottom-layer $\mathcal{C}$-circuits have polynomial size. By Theorem 7, for every polynomial-size parity branching program $B$, there is a projection $p\colon \{0,1\}^n \to \{0,1\}^{n^{O(1)}}$ such that for every $x \in \{0,1\}^n$, $B(x) = \mathsf{CMD}(p(x))$. Since $\mathcal{C}$ is closed under projection, this means that every polynomial-size parity branching program has a $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-circuit of polynomial complexity and size, which then implies that every function in $\mathsf{NC}^1$ also has such a $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-circuit. On the other hand, by Item 2, there is a function $L \in \mathsf{E}$ that has no $\widetilde{\mathsf{SUM}}_\delta \circ \mathcal{C}$-circuit of polynomial complexity and size, so $L$ is not in $\mathsf{NC}^1$. Using hardness amplification against $\mathsf{NC}^1$ (Lemma 10), it follows that there is a function in $\mathsf{E}$ that is strongly average-case hard against $\mathsf{NC}^1$, which by assumption contains polynomial-size $\mathcal{C}$-circuits.

**Item 6 $\Rightarrow$ Item 10.** We construct the PRG using the hardness vs. randomness framework. Consider Lemma 8 with the following setting of parameters: $a := 2/\gamma$ and $\ell := n^{\gamma/4}$. Let $G_{L_\ell}\colon \{0,1\}^t \to \{0,1\}^n$ be the PRG defined as $G_{L_\ell}(z) := G(L_\ell, z)$, where $L \in \mathsf{E}$ is the language from Item 6. Note that the seed length $t$ is at most $O\left(\ell^2 \cdot n^{1/a}/a\right) \le n^\gamma$ and $G_{L_\ell}$ can be computed in time $\operatorname{poly}(n, 2^t) = 2^{O(n^\gamma)}$. Let $k \ge 1$ be any constant and consider any $\ell$-variate $\mathcal{C} \circ \mathsf{JUNTA}_a$-circuit $C$ where the top circuit has size $n^k = \ell^{4k/\gamma}$. Since $\mathcal{C}$ is closed under a bottom layer of juntas, we have that $C \in \mathcal{C}\left[\ell^{k'}\right]$ for some large enough $k' > 4k/\gamma$. Also, let $\varepsilon = 1/n^k$, which implies $\varepsilon/n = 1/n^{k+1} = 1/\ell^{4(k+1)/\gamma} \ge 1/\ell^{k'}$. From Item 6, we have that $L_\ell$ cannot be $\left(1/2 + 1/\ell^{k'}\right)$-approximated by any circuit from $\mathcal{C}[\ell^{k'}]$, for infinitely many values of $\ell$. Then by Lemma 8, we conclude that $G_{L_\ell}$ $(1/n^k)$-fools any circuit from $\mathcal{C}[n^k]$, for infinitely many values of $n$.

**Item 10 $\Rightarrow$ Item 1.** Let $G\colon \{0,1\}^r \to \{0,1\}^n$ be an i.o. PRG as in Item 10, where $r \le n-2$. That is, for each choice of $k'$, $G$ fools circuits from $\mathcal{C}[n^{k'}]$ on input length $n$ with error $\varepsilon(n) \le n^{-k'}$, for infinitely many values of $n$.

Let $k \ge 1$ and let $C \in \widetilde{\mathsf{SUM}}_{1/3} \circ \mathcal{C}[n^k]$. By Proposition 9, it suffices to show that $G$ is an i.o. $(< \frac{3}{4})$-PRG against $C$. Let $\tilde{C}$ be the corresponding linear sum for $C$. That is,

$$\tilde{C}(x) := \sum_i \alpha_i \cdot C_i(x),$$

where $C_i \in \mathcal{C}[n^k] \subseteq \mathcal{C}[n^{k':=k+1}]$ and $\sum_i |\alpha_i| \le n^k$. Since $\tilde{C}$ (1/3)-approximates $C$ in a pointwise manner, we have

$$\left|\mathbf{E}[C(\mathcal{U})] - \mathbf{E}[\tilde{C}(\mathcal{U})]\right| \le 1/3 \text{ and } \left|\mathbf{E}[C(G)] - \mathbf{E}[\tilde{C}(G)]\right| \le 1/3.$$

Therefore, if we can show that

$$\left|\mathbf{E}[\tilde{C}(\mathcal{U})] - \mathbf{E}[\tilde{C}(G)]\right| \le \delta,$$

for some $\delta < 1/12$ (infinitely often), then $G$ $\delta'$-fools $C$ (infinitely often), where $\delta' = 2/3 + \delta < 3/4$. We have

$$\begin{aligned}\left|\mathbf{E}[\tilde{C}(\mathcal{U})] - \mathbf{E}[\tilde{C}(G)]\right| &= \left|\mathbf{E}\left[\sum_i \alpha_i \cdot C_i(\mathcal{U})\right] - \mathbf{E}\left[\sum_i \alpha_i \cdot C_i(G)\right]\right| \\ &= \left|\sum_i \alpha_i \cdot \mathbf{E}[C_i(\mathcal{U})] - \mathbf{E}[C_i(G)]\right| \\ &\le \max_i \left|\mathbf{E}[C_i(\mathcal{U})] - \mathbf{E}[C_i(G)]\right| \cdot \sum_i |\alpha_i| \\ &\le n^{-k'} \cdot n^k \le 1/n,\end{aligned}$$

as desired.

**Item 1 $\Rightarrow$ Item 2.** This implication is straightforward.

**Item 6 $\Rightarrow$ Item 5 $\Rightarrow$ Item 4.** The first implication is obvious. The contrapositive of the second implication follows from an averaging argument.

**Item 4 $\Rightarrow$ Item 1.** It suffices to show that for every $k \ge 1$, every function in $\widetilde{\mathsf{SUM}}_{1/3} \circ \mathcal{C}[n^k]$ has a probabilistic $\mathcal{C}[n^k]$-circuit with error $1/2 - 1/n^{O(k)}$.

For the simplicity of presentation, we will consider Boolean functions that take inputs from $\{0,1\}^n$ and output values in $\{-1,1\}$. Let $f_\pm \colon \{0,1\}^n \to \{-1,1\} \in \widetilde{\mathsf{SUM}}_{1/3} \circ \mathcal{C}_\pm[n^k]$. Then there is a linear sum of $\mathcal{C}_\pm[n^k]$-circuits

$$f_1(x) := \sum_i \alpha_i \cdot \left(\frac{1 - C_i(x)}{2}\right),$$

where $C_i \colon \{0,1\}^n \to \{-1,1\} \in \mathcal{C}_\pm[n^k]$ and $\sum_i |\alpha_i| \le n^k$, such that

- if $f_\pm(x) = 1$, then $f_1(x) \le 1/3$, and
- if $f_\pm(x) = -1$, then $f_1(x) \ge 2/3$.

Next, let

$$f_2(x) := 1/2 - f_1(x).$$

It is easy to see that

- if $f_\pm(x) = 1$, then $f_2(x) \ge 1/6$, and
- if $f_\pm(x) = -1$, then $f_2(x) \le -1/6$.

Now note that since $\mathcal{C}_\pm$ is closed under negation, $f_2$ can be written as

$$f_2(x) := \sum_j \beta_j \cdot D_j(x),$$

where for each $j$, $D_j \colon \{0,1\}^n \to \{-1,1\} \in \mathcal{C}_{\pm}[n^k]$, $\beta_j \geq 0$, and $T := \sum_j \beta_j \leq n^{O(k)}$. Finally, let

$$f_3(x) := \frac{f_2(x)}{T}.$$

Let $\mathcal{D}$ be the probabilistic $\mathcal{C}_{\pm}[n^k]$-circuit where $D_j$ is sampled with probability $\beta_j/T$. Then for every $x$ we have $\mathbf{E}_{\mathcal{D}}[\mathcal{D}(x)] = f_3(x)$. Moreover, if $f_{\pm}(x) = 1$, then

$$\begin{aligned}\frac{1}{6T} &\leq \mathop{\mathbf{E}}_{\mathcal{D}}[\mathcal{D}(x)] \\ &= \mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = 1] - \mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = -1] \\ &= \mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = 1] - (1 - \mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = 1]) \\ &= 2 \cdot \mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = 1] - 1,\end{aligned}$$

which implies

$$\mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = 1] \geq \frac{1}{2} + \frac{1}{12T}.$$

Similarly, we can show that if $f_{\pm}(x) = -1$, then

$$\mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = -1] \geq \frac{1}{2} + \frac{1}{12T}.$$

Therefore, $\mathcal{D}$ is a probabilistic $\mathcal{C}_{\pm}[n^k]$-circuit for $f_{\pm}$ with error $1/2 - 1/n^{O(k)}$.

**Item 6 $\Rightarrow$ Item 3.** This follows from the standard Discriminator Lemma [24].

**Item 3 $\Rightarrow$ Item 4.** We will show that for every $k \geq 1$, every function that has a probabilistic $\mathcal{C}[n^k]$-circuit with error $1/2 - 1/n^k$ is contained in $\mathsf{MAJ}_{n^{O(k)}} \circ \mathcal{C}[n^{O(k)}]$.

Let $f \colon \{0,1\}^n \to \{0,1\}$ and $\mathcal{D}$ be the probabilistic $\mathcal{C}[n^k]$-circuit for $f$ with error $1/2 - 1/n^k$. That is, for every $x$,

$$\mathop{\mathbf{Pr}}_{\mathcal{D}}[\mathcal{D}(x) = f(x)] \geq 1/2 + 1/n^k.$$

By the Chernoff bound, if we sample $t := O(n^{2k} \cdot n)$ circuits $C_1, \ldots, C_t$ from $\mathcal{D}$, then

$$\mathop{\mathbf{Pr}}_{C_1,\ldots,C_t \sim \mathcal{D}}\left[\mathop{\mathbf{Pr}}_{i \in [t]}[C_i(x) = f(x)] \geq 1/2 + 1/(2n^k)\right] \geq 1 - 2^{-2n}.$$

By a union bound over $x \in \{0,1\}^n$, there exist $t$ circuits $C_1, \ldots, C_t$ such that for every $x$,

$$\mathop{\mathbf{Pr}}_{i \in [t]}[C_i(x) = f(x)] \geq 1/2 + 1/(2n^k).$$

Therefore, by taking the majority of these $t$ circuits, we obtain a $\mathsf{MAJ}_{n^{O(k)}} \circ \mathcal{C}[n^{O(k)}]$-circuit that computes $f$.

**Item 10 $\Rightarrow$ Item 9 $\Rightarrow$ Item 2.** This first implication is obvious. The proof of the second implication is essentially the same as that of "Item 10 $\Rightarrow$ Item 1". From Item 9, we get an i.o. PRG with seed length $n - 1$ that $\left(< \frac{1}{2}\right)$-fools $\widetilde{\mathsf{SUM}}_{\delta} \circ \mathcal{C}$-circuits for some $\delta = 1/\operatorname{poly}(n)$, which by Proposition 9 implies Item 2. We omit the details here.

**Item 6 ⇒ Item 7.** Let $L\colon \{0,1\}^* \to \{0,1\}$ be the language from Item 6. For the sake of contradiction, suppose there is a $k \geq 1$ such that $L$ can be approximated by $[0,1]$-$\mathsf{SUM} \circ \mathcal{C}[n^k]$-circuits within $\ell_1$ distance $1/3$. Then by Item 1 of Proposition 12, we have that for every $n$, there is a $\mathsf{SUM} \circ \mathcal{C}_{\pm}[O\left(n^k\right)]$-circuit $C$ such that $\|C\|_\infty \leq 1$ and

$$\langle (L_\pm)_n, C\rangle \geq 1/3.$$

Suppose

$$C(x) := \sum_i |\alpha_i| \cdot C_i(x),$$

where $C_i \in \mathcal{C}_{\pm}[O\left(n^k\right)]$ and $\sum_i |\alpha_i| \leq O\left(n^k\right)$. Then

$$\begin{aligned} 1/3 &\leq \left\langle (L_\pm)_n, \sum_i \alpha_i \cdot C_i \right\rangle \\ &= \sum_i \alpha_i \cdot \langle (L_\pm)_n, C_i\rangle \\ &\leq \sum_i |\alpha_i| \cdot \langle (L_\pm)_n, C_i\rangle \\ &\leq \max_i \langle (L_\pm)_n, C_i\rangle \cdot \sum_i |\alpha_i| \\ &\leq \max_i \langle (L_\pm)_n, C_i\rangle \cdot O\left(n^k\right), \end{aligned}$$

which implies that there exists some $i$ such that

$$\langle (L_\pm)_n, C_i\rangle \geq \frac{1}{O(n^k)}.$$

This contradicts Item 6.

**Item 7 ⇒ Item 8.** This implication is obvious.

**Item 8 ⇒ Item 6.** By Item 2 of Proposition 12, we have that Item 8 implies that there is a language $L\colon \{0,1\}^* \to \{-1,1\}$ in $\mathsf{E}$ and $\delta = 1/\ell^b$, where $b \geq 1$ is a constant, such that for every $k' \geq 1$, on infinitely many input lengths there is no $\mathsf{SUM} \circ \mathcal{C}_{\pm}[\ell^{k'}]$-circuit $C$ with $\|C\|_\infty \leq 1$ such that

$$\langle L_\ell, C\rangle \leq 1 - 2\delta. \tag{1}$$

Now consider the following language $L'\colon \{0,1\}^* \to \{-1,1\}$: on input $x$ of length $n$, let $\ell$ be the largest integer such that $\ell \cdot \ell^b \log^2(\ell) \leq n$ and view the input as $x = (x_1, \ldots, x_t, y)$, where $t := \ell^b \log^2(\ell)$ and $x_i \in \{0,1\}^\ell$ for $i \in [t]$. Then let

$$L'(x) := \prod_{i \in [t]} L(x_i).$$

Note that for large enough $n$ we have

$$n < 2\ell \cdot t < \ell^{b+2}.$$

We claim that $L'$ is strongly average-case hard against $\mathcal{C}_{\pm}$-circuits. For the sake of contradiction, suppose there is $k \geq 1$ and an $n$-variate circuit $C' \in \mathcal{C}_{\pm}[n^k]$ such that, for all large enough $n$,

$$\langle L'_n, C' \rangle > \frac{1}{n^k}.$$

By an averaging argument, where we fix the $y$-part of the input to some value, there exists some $(\ell \cdot t)$-variate $\mathcal{C}_{\pm}$-circuit $C'$ of size $n^k \leq \ell^{k(b+2)}$ such that

$$\left\langle L_\ell^{\oplus t}, C'' \right\rangle > \frac{1}{n^k}.$$

Note that for $\delta = 1/\ell^b$ and our choice of $t = \ell^b \log^2(\ell)$, we have

$$\frac{1}{n^k} > (1-2\delta)^t + \frac{1}{2\delta \cdot \ell^{k(b+2)} \cdot \ell^b}.$$

By Theorem 13, there is a $\mathsf{SUM} \circ \mathcal{C}_{\pm}$ $C$ where $\|C\|_\infty \leq 1$, the top $\mathsf{SUM}$ has complexity $10 \cdot \ell \cdot \left(\ell^{k(b+2)} \cdot \ell^b\right)^2 \leq \ell^{O(kb)}$ and the bottom layer $\mathcal{C}_{\pm}$-circuits have size $\ell^{k(b+2)}$ such that

$$\langle L_\ell, C \rangle > 1 - 2\delta,$$

for all large enough $\ell$. This contradicts Equation (1). ◀

## References


1 Eric Allender, Vikraman Arvind, and Fengming Wang. Uniform derandomization from pathetic lower bounds. In *RANDOM-APPROX*, pages 380–393. Springer, 2010.

2 Josh Alman, Timothy M. Chan, and R. Ryan Williams. Polynomial representations of threshold functions and algorithmic applications. In *FOCS*, pages 467–476, 2016.

3 Benny Applebaum. Garbled circuits as randomized encodings of functions: a primer. In *Tutorials on the Foundations of Cryptography*, pages 1–44. Springer International Publishing, 2017.

4 Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. Cryptography in $\mathsf{NC}^0$. *SIAM J. Comput.*, 36(4):845–888, 2006.

5 Sergei Artemenko and Ronen Shaltiel. Lower bounds on the query complexity of non-uniform and adaptive reductions showing hardness amplification. *Comput. Complex.*, 23(1):43–83, 2014.

6 Richard Beigel and Jun Tarui. On ACC. *Comput. Complex.*, 4:350–366, 1994.

7 Abhishek Bhrushundi, Kaave Hosseini, Shachar Lovett, and Sankeerth Rao. Torus polynomials: An algebraic approach to ACC lower bounds. In *ITCS*, pages 13:1–13:16, 2019.

8 Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *STOC*, pages 94–99, 1983.

9 Arkadev Chattopadhyay and Nikhil S. Mande. A short list of equalities induces large sign rank. In *FOCS*, pages 47–58, 2018.

10 Lijie Chen. Non-deterministic quasi-polynomial time is average-case hard for ACC circuits. In *FOCS*, pages 1281–1304, 2019.

11 Lijie Chen, Zhenjian Lu, Xin Lyu, and Igor Carboni Oliveira. Majority vs. approximate linear sum and average-case complexity below NC1. *Electron. Colloquium Comput. Complex.*, 28:40, 2021.

12 Lijie Chen and Xin Lyu. Inverse-exponential correlation bounds and extremely rigid matrices from a new derandomized XOR lemma. In *STOC*, 2021.

13 Lijie Chen, Xin Lyu, and Ryan Williams. Almost-everywhere circuit lower bounds from non-trivial derandomization. In *FOCS*, pages 1–12, 2020.

14 Lijie Chen and Hanlin Ren. Strong average-case lower bounds from non-trivial derandomization. In *STOC*, pages 1327–1334, 2020.

15 Lijie Chen and Ryan Williams. Stronger connections between circuit analysis and circuit lower bounds, via PCPs of proximity. In *CCC*, pages 19:1–19:43, 2019.

16 Bill Fefferman, Ronen Shaltiel, Christopher Umans, and Emanuele Viola. On beating the hybrid argument. *Theory Comput.*, 9:809–843, 2013.

17 Jürgen Forster. A linear lower bound on the unbounded error probabilistic communication complexity. In *CCC*, pages 100–106, 2001.

18 Mikael Goldmann, Johan Håstad, and Alexander A. Razborov. Majority gates vs. general weighted threshold gates. *Comput. Complex.*, 2:277–300, 1992.

19 Oded Goldreich, Noam Nisan, and Avi Wigderson. On Yao's XOR-lemma. In *Studies in Complexity and Cryptography*, pages 273–301. Springer, 2011.

20 Shafi Goldwasser, Dan Gutfreund, Alexander Healy, Tali Kaufman, and Guy N. Rothblum. Verifying and decoding in constant depth. In *STOC*, pages 440–449, 2007.

21 Frederic Green, Johannes Köbler, and Jacobo Torán. The power of the middle bit. In *CCC*, pages 111–117, 1992.

22 Aryeh Grinberg, Ronen Shaltiel, and Emanuele Viola. Indistinguishability by adaptive procedures with advice, and lower bounds on hardness amplification proofs. In *FOCS*, pages 956–966, 2018.

23 Dan Gutfreund and Guy N. Rothblum. The complexity of local list decoding. In *RANDOM-APPROX*, pages 455–468, 2008.

24 András Hajnal, Wolfgang Maass, Pavel Pudlák, Mario Szegedy, and György Turán. Threshold circuits of bounded depth. *J. Comput. Syst. Sci.*, 46(2):129–154, 1993.

25 Johan Håstad. On the size of weights for threshold gates. *SIAM J. Discret. Math.*, 7(3):484–492, 1994.

26 Johan Håstad and Mikael Goldmann. On the power of small-depth threshold circuits. *Comput. Complex.*, 1:113–129, 1991.

27 Xuangui Huang and Emanuele Viola. Average-case rigidity lower bounds. In *CSR*, 2021.

28 Russell Impagliazzo. Hard-core distributions for somewhat hard problems. In *FOCS*, pages 538–545, 1995.

29 Russell Impagliazzo and Sam McGuire. Comparing computational entropies below majority (or: When is the dense model theorem false?). In *ITCS*, pages 2:1–2:20, 2021.

30 Russell Impagliazzo and Avi Wigderson. $P = BPP$ if $E$ requires exponential circuits: Derandomizing the XOR lemma. In *STOC*, pages 220–229, 1997.

31 Yuval Ishai and Eyal Kushilevitz. Perfect constant-round secure computation via perfect randomizing polynomials. In *ICALP*, pages 244–256, 2002.

32 Daniel M. Kane and Ryan Williams. Super-linear gate and super-quadratic wire lower bounds for depth-two and depth-three threshold circuits. In *STOC*, pages 633–643, 2016.

33 Adam R. Klivans. On the derandomization of constant depth circuits. In *RANDOM-APPROX*, pages 249–260, 2001.

34 Vaibhav Krishan. Upper bound for torus polynomials. In *CSR*, 2021.

35 Leonid A. Levin. One-way functions and pseudorandom generators. *Comb.*, 7(4):357–363, 1987.

36 Chi-Jen Lu, Shi-Chun Tsai, and Hsin-Lung Wu. Complexity of hard-core set proofs. *Comput. Complex.*, 20(1):145–171, 2011.

37 Cody D. Murray and R. Ryan Williams. Circuit lower bounds for nondeterministic quasi-polytime from a new easy witness lemma. *SIAM J. Comput.*, 49(5), 2020.

38 Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Comput. Complex.*, 4:301–313, 1994.

39 Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994.

40 Alexander A. Razborov. Lower bounds on the size of constant-depth networks over a complete basis with logical addition. *Mathematicheskie Zametki*, 41(4):598–607, 1987.

41 Alexander A. Razborov and Avi Wigderson. $n^{\Omega(\log n)}$ lower bounds on the size of depth-3 threshold circuits with AND gates at the bottom. *Inf. Process. Lett.*, 45(6):303–307, 1993.

42 Ronen Shaltiel and Emanuele Viola. Hardness amplification proofs require majority. *SIAM J. Comput.*, 39(7):3122–3154, 2010.

43 Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. Pseudorandom generators without the XOR lemma. *J. Comput. Syst. Sci.*, 62(2):236–266, 2001.

44 Suguru Tamaki. A satisfiability algorithm for depth two circuits with a sub-quadratic number of symmetric and threshold gates. *Electron. Colloquium Comput. Complex.*, 23:100, 2016.

45 Emanuele Viola. The complexity of constructing pseudorandom generators from hard functions. *Comput. Complex.*, 13(3-4):147–188, 2005.

46 Emanuele Viola. Guest column: correlation bounds for polynomials over $\{0, 1\}$. *SIGACT News*, 40(1):27–44, 2009.

47 Emanuele Viola. Constant-error pseudorandomness proofs from hardness require majority. *ACM Trans. Comput. Theory*, 11(4):19:1–19:11, 2019.

48 Emanuele Viola. New lower bounds for probabilistic degree and $\mathrm{AC}^0$ with parity gates. *Electron. Colloquium Comput. Complex.*, 27:15, 2020.

49 Ryan Williams. Improving exhaustive search implies superpolynomial lower bounds. *SIAM J. Comput.*, 42(3):1218–1244, 2013.

50 Ryan Williams. Limits on representing boolean functions by linear combinations of simple functions: Thresholds, ReLUs, and low-degree polynomials. In *CCC*, pages 6:1–6:24, 2018.

51 Ryan Williams. New algorithms and lower bounds for circuits with linear threshold gates. *Theory of Computing*, 14(1):1–25, 2018.