Electronic Colloquium on Computational Complexity, Report No. 24 (2025)

Lower Bounds Beyond DNF of Parities

Artur Riazanov Anastasia Sofronova Dmitry Sokolov

École Polytechnique Fédérale de Lausanne

March 8, 2025

Abstract

We consider a subclass of $\mathbf{AC}^0[2]$ circuits that simultaneously captures $\mathsf{DNF} \circ \textsc{Xor}$ and depth-3 $\mathbf{AC}^0$ circuits. For this class we show a technique for proving lower bounds inspired by the top-down approach. We give lower bounds for the middle slice function, inner product function, and affine dispersers.

1 Introduction

Constant-depth (or $\mathbf{AC}^0$) de Morgan circuits is one of the circuit classes that is reasonably well-studied, even though strong exponential lower bounds of the form $2^{\Omega(n)}$ are still out of reach. This model can use unbounded $\wedge$, $\vee$ and $\neg$ gates, and the underlying graph of a computation is a constant-depth tree; intuitively, these circuits represent computations that can be efficiently parallelized. There are known hard examples of functions that cannot be computed with small-size $\mathbf{AC}^0$ circuits. The most notable ones include Xor and Maj:

$$\textsc{Xor}(x_1, \dots, x_n) \coloneqq x_1 \oplus \cdots \oplus x_n; \quad \textsc{Maj}(x_1, \dots, x_n) \coloneqq \mathbb{1}_{[x_1+\cdots+x_n>n/2]}.$$

These functions require size $2^{\Omega(n^{1/(d-1)})}$ depth-d $\mathbf{AC}^0$ circuits, and the lower bounds for them are achieved mainly with two techniques: random restrictions, or switching lemma, [FSS84; Ajt83; Yao85; Hås86; Hås87] and polynomial approximation [Raz87; Smo87]. In particular, for circuits of depth 3, there is a lower bound of $2^{\Omega(\sqrt{n})}$ for both of these functions (for Xor it is known to be tight), and breaking through the $\sqrt{n}$ barrier in the exponent for any explicit function is a major open question. Proving strong exponential lower bounds for depth-3 circuits would essentially give a superpolynomial lower bound for general circuits [Val77; GKW21] which is a major open problem in complexity theory. Both switching lemma and polynomial approximation seem unable to give us such strong lower bounds.

Circuits with MOD Gates The situation becomes more challenging in terms of lower bounds, when we plug-in hard functions for $\mathbf{AC}^0$ into our computational model. One of the most natural generalisations of $\mathbf{AC}^0$ circuits that follow this concept is $\mathbf{AC}^0[m]$ circuits, that can also utilise gates computing MOD_m defined as

$$\mathsf{MOD}_m(x_1, \dots, x_n) \coloneqq \mathbb{1}_{[(x_1+\cdots+x_n) \bmod m=0]}.$$

On one hand, a lower bound for this model is necessary if we want to show lower bounds for general circuits. On the other hand, showing such lower bounds is a challenging problem. For example, techniques based on random restrictions, such as switching lemma application, do not work quite as they do in $\mathbf{AC}^0$, since MOD_m gates are not simplified after an application of a restriction. However, when m is a prime power, polynomial approximation achieves lower bounds of the form $2^{n^{1/2d}}$ for Maj [Raz87; Smo87], as well as for computing MOD_q for a prime power q that is relatively prime with m.

When m is not a prime power, very little is known. In fact, utilising non-prime m with many divisors, it is possible to compute any symmetric function in subexponential size even in depth 3 [CW22]. The "minimal example" of the non-prime regime is $\mathbf{AC}^0[6]$. It is still an open question to prove lower bounds for $\mathbf{AC}^0[6]$, and the known techniques fail at resolving that. The reason for that is that polynomial approximation only works over fields, and there is no field with 6 elements.

ISSN 1433-8092

Even for the simplest example of MOD_p gates: MOD_2, or Xor, we are very far from understanding the exact power of $\mathbf{AC}^0[2]$ circuits. Allowing to use the Xor function in the gates of the circuit increases its computational power; for example, depth-4 $\mathbf{AC}^0[2]$ circuits can compute Maj in size $2^{O(n^{1/4})}$ [OSS19], whereas in plain $\mathbf{AC}^0$ there is a lower bound $2^{\Omega(n^{1/3})}$.

The drawbacks of Razborov–Smolensky polynomial approximation method translate to the gaps in our understanding of $\mathbf{AC}^0[2]$ class. In particular, we do not have strong correlation bounds against these circuits even for one of the simplest subclasses of these circuits: $\mathsf{DNF} \circ$ Xor (depth-3 unbounded fan-in circuits of $(\wedge \circ \vee \circ$ Xor$)$-type) [HIV22]. Here, $\circ$ denotes composition, and this essentially means that Xor gates can be used only in the bottom layer of the circuit. Moreover, the polynomial approximation method only applies to functions that require a large degree over $\mathbb{F}_2$. Thus it is unknown whether the inner product $\mathrm{IP}_n(x, y) := x_1y_1 \oplus x_2y_2 \oplus \cdots \oplus x_ny_n$ requires large $\mathbf{AC}^0$ circuits with an additional layer of parity gates in the bottom ($\mathbf{AC}^0 \circ$ Xor). It is known that IP requires exponentially large $\mathsf{DNF} \circ$ Xor circuits [Juk06; CS16a], but even for $(\vee \circ \wedge \circ \vee \circ$ Xor$)$-circuits the best known lower bound is $n^{2-o(1)}$ [CGJWX18].

1.1 Top-down Approach

Overall, there is a clear shortage of new techniques in circuit complexity and, by extension, in adjacent areas, while the well-known ones also have the well-known drawbacks. In this work, we focus on studying another circuit lower bound technique, which falls under the umbrella of *top-down* methods.

Top-down lower bounds start from the output gate of a candidate circuit and move down the circuit in search of a mistake. While such an approach has been known for a long time, and there is a long line of work on top-down lower bounds for depth-3 $\mathbf{AC}^0$ circuits [BS79; San89; Ko90; HJP95; RS98; PPZ99; PSZ00; IPZ01; PPSZ05; Wol06; BGM06; MW19; GKW21; FGT22; GGM23], it still remains largely underdeveloped. So far top-down lower bounds against $\mathbf{AC}^0$ are known only for circuits up to depth-4 [GRSS23], while bottom-up methods yield lower bounds for arbitrary constant depth (or even $\log n / \log \log n$ in some cases). The motivation for studying top-down comes from the fact that this method is complete for $\mathbf{AC}^0$ circuits (see discussion in [Hir17; GGM23; GRSS23]). In other words, there are no formal barriers that would prevent such an approach from being able to prove lower bounds in the regimes where other known methods cannot.

The main model to which top-down techniques are applicable is $\mathbf{AC}^0$ circuits, and the main example of a hard function is Xor. In this work, we attempt to adapt such techniques to be able to prove lower bounds for circuits *with parity gates*. We consider a subclass of $\mathbf{AC}^0 \circ$ Xor which is strictly stronger than $\mathsf{DNF} \circ$ Xor and prove lower bounds for it in a top-down fashion. In particular, we prove a lower bound for an affine disperser, which does not follow from Razborov–Smolensky method.

1.2 The Model and Results

In a recent paper Huang, Ivanov, and Viola [HIV22] give an explanation of why the class $\vee \circ \wedge \circ \vee \circ$ Xor resists known lower bound techniques. They show that there is a circuit of this type that computes a very strong affine extractor: a function that is close to be balanced on all large enough affine subspaces. On the other hand, it is known that affine extractors are hard for $\mathbf{AC}^0$ (by definition we know upper bounds on Fourier coefficients, which contradicts with spectrum concentration of small $\mathbf{AC}^0$ circuits obtained from switching lemma or polynomial approximation, see, for example [Tal17]) and $\mathsf{DNF} \circ$ Xor [CS16b]. Moreover, [HIV22] show that $\mathsf{DNF} \circ$ Xor can compute a *one-sided* affine extractor: a balanced function that is never too biased towards zero on large enough affine subspaces. They then use the latter result to separate $\mathsf{DNF} \circ$ Xor from $\mathbf{AC}^0 \circ$ Xor that has at most n distinct Xor gates (here n is the number of variables). In other words, such a circuit is a composition of $\mathbf{AC}^0$ circuit and a non-singular affine transformation over $\mathbb{F}_2$, or an $\mathbf{AC}^0 \circ \mathbb{B}$ circuit.

This type of circuit can already compute arbitrary linear forms, but we consider a stronger model. Our model is essentially a union of $\mathbf{AC}^0$ and $\mathbf{AC}^0 \circ \mathbb{B}$ within $\vee \circ \wedge \circ \vee \circ$ Xor.

Definition 1.1

Let $C_1, \dots, C_N$ be some constant depth de Morgan circuits and $A_1, \dots, A_N$ be non-singular affine transformations. We then say that $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$-circuit is $\bigvee_{i \in [N]} C_i \circ A_i$, i.e. the disjunction of compositions of a constant-depth circuit and an affine map. The *depth* of a $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$-circuit is the depth of $C_1 \vee \dots \vee C_N$. The *size* of a $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ circuit is the total number of gates in $C_1, \dots, C_N$. If $N = 1$ we denote the corresponding circuit type as $\mathbf{AC}^0 \circ \mathbb{B}$.

We remark that the choice of $\vee$ as the top gate is arbitrary, all the arguments could be handled with $\wedge$ on top (with hard functions changing appropriately).

Our primary motivation for studying this class is to develop a line of attack on subclasses of $\mathbf{AC}^0 \circ \textsc{Xor}$ circuits. Along the way, we establish lower bounds for the class, making concrete progress in this direction. As a subclass of $\mathbf{AC}^0 \circ \textsc{Xor}$, $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ also has a natural interpretation in a top-down framework, as it corresponds to specific assumptions allowed in the proof strategy. We discuss this in section 2. It is worth noting that as $\mathsf{DNF} \circ \textsc{Xor}$ is a subclass of depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$, $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$ is a subclass of depth-4 $\textsc{Or} \circ \textsc{And} \circ (\Sigma^0 \circ \mathbb{B})$. So, strong average-case lower bounds against depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ would also imply $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$ lower bounds. In section 4, we propose a roadmap of intermediate open questions which aims to extend this approach to eventually achieve lower bounds for stronger subclasses of $\mathbf{AC}^0 \circ \textsc{Xor}$ and even $\mathbf{AC}^0$[6].

As a main result we present a general approach for proving lower bound for this model of computation. We give the highlights of the technique in Section 2. We now show the comparison of this model with classical models and state the lower bounds that we get.

The Comparison of the Models Let us first observe that depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ is properly larger than depth-3 $\mathbf{AC}^0 \circ \mathbb{B}$. Observe that $\mathsf{DNF} \circ \textsc{Xor}$ is a special case of depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$. The strict inclusion is then implied by the following.

Theorem 1.2 ([HIV22])

A function that admits a polynomial $\mathsf{DNF} \circ \textsc{Xor}$ circuit may require an exponential $\mathbf{AC}^0 \circ \mathbb{B}$ circuit.

Proof sketch. Implied by the combination of Corollary 5 and Claim 21 in [HIV22], the former shows that $\mathsf{DNF} \circ \textsc{Xor}$ can compute functions for which there is a correlation bound for $(n - \mathrm{poly}(\log n))$-depth parity decision trees (PDT), while the latter observes that the switching lemma applied to a $\mathbf{AC}^0 \circ \mathbb{B}$-circuit yields a $(n - \log^{\omega(1)} n)$-depth PDT approximating the function. □

On the other hand, depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ is properly larger than $\mathsf{DNF} \circ \textsc{Xor}$. Since $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ contains CNF this statement is implied by the following Theorem.

Theorem 1.3

Let $f \colon \{0,1\}^{n \times 3} \to \{0,1\}$ defined as $f(x) := \bigwedge_{i \in [n]} (x_{i1} + x_{i2} + x_{i3} = 1)$ where the sum is over $\mathbb{R}$. Then any $\mathsf{DNF} \circ \textsc{Xor}$ circuit computing f has size $\Omega(1.5^n)$.

To the best of our knowledge, this is the simplest existing lower bound for $\mathsf{DNF} \circ \textsc{Xor}$. We include the proof in section 3.4.

Affine Dispersers $f \colon \{0,1\}^n \to \{0,1\}$ is a (k, ε)-*affine extractor* if for every affine subspace $A \subseteq \{0,1\}^n$ (where we equate $\{0,1\}^n$ and $\mathbb{F}_2^n$) of dimension at least k we have $|\Pr_{\boldsymbol{a} \sim A}[f(\boldsymbol{a}) = 1] - 1/2| < \varepsilon$. We say that f is a k-*affine disperser* if it is a $(k, 1/2)$ affine extractor, i.e. $\Pr_{\boldsymbol{a} \sim A}[f(\boldsymbol{a}) = 1] \notin \{0,1\}$ for every k-dimensional affine subspace A.

In our first result, we confirm that depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ is smaller than $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$. Theorem 3 in [HIV22] shows that polynomial-size $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$ circuit computes affine extractors with polylogarithmic dimension (i.e. the function is close to being balanced in every affine subspace of at least polylogarithmic dimension). On the other hand, we show in section 3.2 the following Theorem.

Theorem 1.4 (informal)

If a function is not constant on any $n^{1/3-o(1)}$-dimension affine subspace (i.e. it is an $n^{1/3-o(1)}$-affine disperser) then it requires depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ circuits of exponential size.

Theorem 3 in [HIV22] implies that the polynomial approximation method can not prove Theorem 1.4, since it can not distinguish between a depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$-circuit and a $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$-circuit, which contains some affine extractors.

Inner Product The inner product function $\mathrm{IP}_n \colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ defined as follows $\mathrm{IP}_n(x,y) = \sum_{i\in[n]} x_i \cdot y_i \bmod 2$. It is a big open problem to get a lower bound for the inner product in $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$. In section 3.3 that inner product requires exponentially large $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ circuits. The technique there is a combination of random restrictions with a top-down step.

Middle Slice The middle slice function $\textsc{Mid}_n \colon \{0,1\}^n \to \{0,1\}$ is defined as $\textsc{Mid}_n(x) = \mathbb{1}_{[|x|=n/2]}$, i.e. it equals 1 iff the input contains exactly $n/2$ ones. In section 3.1 we prove via a top-down argument that this function requires $2^{\Omega(\sqrt{n})}$-size depth-3 $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$ circuits.

2 Technique

2.1 AC^0 Top-down Lower Bounds

The general sketch of an $\mathbf{AC}^0$ top-down proof looks as follows.

- Consider a circuit $C = \bigvee_{i=0}^{s} C_i$ (the case of $\bigwedge$ is treated analogously). Suppose we want to prove that C cannot distinguish certain sets of inputs A and B. Assume that, on the contrary, $C(A) = 1$ and $C(B) = 0$.
- Note that $\bigcup_i C_i^{-1}(1) \supseteq A$ and also for every i it holds that $C_i^{-1}(0) \supseteq B$. We pick some C_i such that $A_i := C_i^{-1}(1)$. Now, this is a circuit that separates A_i and B, and it has $\bigwedge$ as the top gate.
- We repeat the procedure until we arrive at a shallow enough circuit C' that is supposed to separate the sets A' and B'.
- We prove that C' cannot separate these sets. Note that if the original circuit C made an error in computing A and B, there always exists a sequence of choices of subcircuits such that the error is traced until C'.

A "shallow enough circuit" could be, in principle, even a variable, but it turns out that there is a convenient way to argue about circuits of depth 2 in this context. Moreover, for now we assume that these circuits of depth 2 are CNFs/DNFs of width bounded by some parameter k. At the end of this section, we discuss why this assumption is acceptable for the proper choice of k.

The central notion for this for analyzing k-CNFs/DNFs is *a k-limit*. The notion comes from [HJP95], inspired by "limit vectors" from [Sip84] as well as communication complexity techniques [KW90].

Definition 2.1 ([HJP95])

Let $A \subseteq \{0,1\}^n$. $x \in \{0,1\}^n$ is a k-limit of A, if for any subset of indices $I \in \binom{[n]}{k}$ there is $y \in A$ such that $x_I = y_I$.

Claim 2.2 ([HJP95; MW19; GRSS23])

If a CNF formula C of width k accepts a set A, then it accepts every k-limit of A.

At the same time, it is known that a k-limit is a complete notion in the following sense.

Claim 2.3 ([HJP95])

Let $A \subseteq \{0,1\}^n$ be a set such that for its any k-limit x it holds that $x \in A$. Then there exists a k-CNF formula C such that $C^{-1}(1) = A$.

Proof. We start from an empty CNF formula C and gradually add clauses to it. Consider any $y \notin A$. As it is not a k-limit of A, there exists a set $I \in \binom{[n]}{k}$ such that $y_I \neq a_I$ for any $a \in A$. Let us add into C a clause $D = \bigvee_{i \in I} x_i^{1-y_i}$. This clause evaluates to 0 on y and evaluates to 1 on any $a \in A$.

We repeat the procedure for every $y \notin A$. The resulting CNF evaluates to 0 on every $y \notin A$ and evaluates to 1 on any $a \in A$, which proves the claim. □

Note that a k-CNF on n variables has at most $(2n)^k = 2^{O(k \log n)}$ clauses. The standard assumption for $\mathbf{AC}^0$ is that the bottom fan-in of the circuit is bounded by the logarithm of its size, so k-limits are a tool that helps to prove lower bounds of the form 2^k. So there is a multiplicative gap of $\log n$ in the exponent between related lower bound and upper bound, which follow from existence and non-existence of a k-limit, respectively. In most cases, this is a negligible difference, but for some examples this might be important: for example, MAJ function has a lower bound of $2^{\Omega(\sqrt{n})}$ in depth-3 $\mathbf{AC}^0$ and an upper bound of $2^{O(\sqrt{n} \log n)}$ in the same model. Proving a $2^{\omega(\sqrt{n})}$ lower bound for MAJ would beat all state-of-the-art lower bounds for depth-3 $\mathbf{AC}^0$ circuits.

Reducing Bottom Fan-in Bottom-up $\mathbf{AC}^0$ lower bounds usually use the fact that bottom fan-in is bounded by a parameter k. In most cases, k can be made as small as $\log s$, where s is the size of the circuit. This is done by random restrictions, which kill the bottom layer gates with a big fan-in. In case of $\mathbf{AC}^0 \circ \text{XOR}$, this becomes more of a problem, since XOR survives under small restrictions.

In fact, handling the bottom fan-in can be done using top-down techniques. In [GRSS23], the lower bound is fully top-down in the sense that no random restrictions are used even for reducing the bottom fan-in, and we follow the same path. With the use of XOR gates, this becomes crucial. The idea is that instead of just one k-limit, one needs to find many, so that wide clauses in a CNF could not reject all of them. We make this intuition precise:

Lemma 2.4

Let C be a CNF over n variables of size s and $A \subseteq C^{-1}(1)$. Let $L \subseteq C^{-1}(0)$ be a set of k-limits for A. Then $s > |L|/2^{n-k}$.

Proof. Let $C = D_1 \wedge \ldots \wedge D_s$. Then $C^{-1}(0) = \bigcup_{i \in [s]} D_i^{-1}(0)$. Let $i \in [s]$ be the clause with the largest size of $D_i^{-1}(0) \cap L$, this size is at least $|L|/s$ since $L \subseteq C^{-1}(0)$. Now the width of D_i must be larger than k, since it distinguishes all k-limits in $L \cap C^{-1}(0)$ from A, hence $|L|/s \leq |D_i^{-1}(0)| < 2^{n-k}$. The claim then follows. □

2.2 Extending the Approach to Parity Gates

We can define the analogous notion for $\mathbf{AC}^0 \circ \text{XOR}$ circuits.

Definition 2.5 (k-parity limit)

Let $A \subseteq \{0,1\}^n$. x is a k-parity limit of A, if for any affine subspace $L \subseteq \{0,1\}^n$ of co-dimension k such that $x \in L$ it holds that $A \cap L \neq \emptyset$.

The proof of the following claim is analogous to Theorem 2.2.

Claim 2.6

If a k-CNF $\circ$ XOR circuit C accepts a set A, then it accepts any k-parity limit of A.

As there are much more linear subspaces of co-dimension k than clauses of width k, while we can prove the analogue of Theorem 2.3 for k-parity limits, the resulting circuit would have huge size.

Open Problem. Is k-parity limit complete in the following sense: if there is no k-parity limit of set A outside of the set itself, then there is a CNF $\circ$ XOR circuit C of width k and size $2^{k^{O(1)}}$ such that it accepts exactly set A?

For proving lower bounds against $\mathbf{AC}^0 \circ \text{XOR}$, it is sufficient to find k-parity limits only with respect to a fixed set of linear forms present in a circuit (subcircuit), but it would be interesting to know if the more

general statement is true. Again, the existence of a k-parity limit with respect to a fixed set of linear forms is a necessary condition for a lower bound.

The proof of the next claim is, again, analogous to Theorem 2.2.

Claim 2.7

Let $C = \bigwedge_{i=1}^{s} L_i$ be a (k-CNF) $\circ$ Xor (here L_i is such that $L_i^{-1}(0)$ is a co-dimension-k affine subspace) and let $\mathcal{S}$ be a collection of affine subspaces of co-dimension k such that $L_i^{-1}(0) \in \mathcal{S}$ for all $i \in [s]$. Suppose that C accepts A. Consider y such that for any $L \in \mathcal{S}$, $y \in L$ implies that there exists $x \in A$ such that $x \in L$. Then C accepts y.

Here $\mathcal{S}$ can be a collection of all affine subspaces of co-dimension k, or it can be a smaller family of subspaces that still contains all linear systems used in a (k-CNF) $\circ$ Xor. In other words, as the number of linear systems used in a circuit is bounded by its size, one can relax the definition of a k-parity limit to be able to fool only these linear systems. One can also prove a variant of completeness for this weaker notion of k-parity limits analogous to Theorem 2.3.

Exponential size depth-3 Or $\circ$ ($\Pi^0 \circ \mathbb{B}$), in particular, can use exponential number of different linear forms, with the restriction that inside each CNF $\circ$ Xor subcircuit there are only n different linear forms. Our results can be seen as finding k-parity limits under these restrictions on the model. In top-down language, the extra Or on top symbolises that the first step down in the proof is oblivious to the actual parity gates that are used in the circuit. Now, just two such oblivious steps would imply lower bounds for $\vee \circ \wedge \circ \vee \circ$ Xor.

When comparing top-down lower bounds for plain $\mathbf{AC}^0$ and Or$\circ$($\Pi^0 \circ \mathbb{B}$), the most important difference (and our main technical contribution) is the following: for Or $\circ$ ($\Pi^0 \circ \mathbb{B}$), we should be able to construct sets such that we can find their k-limits *after any change of basis* in $\mathbb{F}_2^n$. Note that the hard functions in this case should be uncorrelated with affine subspaces, which, in particular, is not true for Xor: after the appropriate change of basis, we could encode the value of the function in the first bit of the string in the new basis.

This seems like a natural step towards fully adapting the top-down approach for circuits with Xor gates. We discuss this more in section 4.

2.3 Unpredictability and Local Limits

One of the most successful to date ways to find local limits is via *unpredictability from partial information*.

Let $X \subseteq \{0,1\}^n$ and $R \subseteq [n]$. A pair (Q, a) with $Q \subseteq [n] \setminus R$ and $a \in \{0,1\}^Q$ is a certificate for R if there exists $b \in \{0,1\}^R$ such that whenever $x_Q = a$ for $x \in X$, $x_R \neq b$. In this case, we say that x *contains* a certificate for R, and the *size* of such certificate is $q := |Q|$. This notion was introduced in [MW19] who proved the following result for $|R| = 1$.

Lemma 2.8 (Bit unpredictability [MW19])

Let $X \subseteq \{0,1\}^n$ have density $|X|/2^n \geq 2^{-d}$. Then for any $q \geq 1$,

$$\Pr_{(\boldsymbol{x},\boldsymbol{i})\sim X\times[n]}[\,\boldsymbol{x} \text{ contains a size-}q \text{ certificate for } \boldsymbol{i} \text{ wrt } X\,] \;\leq\; O(dq/n).$$

More recently [GRSS23] generalized this result for $|R| > 1$.

Lemma 2.9 (Block unpredictability [GRSS23])

Let $X \subseteq \{0,1\}^n$ have density $|X|/2^n \geq 2^{-d}$. Then for any $q, r \geq 1$,

$$\Pr_{(\boldsymbol{x},\boldsymbol{R})\sim X\times\binom{n}{r}}[\,\boldsymbol{x} \text{ contains a size-}q \text{ certificate for } \boldsymbol{R} \text{ wrt } X\,] \;\leq\; O(dqr/n)^{1/6}.$$

Bit and block unpredictability were used in top-down lower bounds for low-depth $\mathbf{AC}^0$ circuits as a way to extract *local limits* (theorem 2.1).

Lemma 2.10 ([MW19; GRSS23])

Suppose $x \in X$ does not contain any q-certificates for R wrt to X. Then every x' such that $x_{[n]\setminus R} = x'_{[n]\setminus R}$ is a q-limit of X.

Proof. Suppose that there exists x' with $x'_{[n]\setminus R} = x_{[n]\setminus R}$ that is not a q-limit for X, i.e. there exists a set $S \subseteq [n]$ of size q such that for every $y \in X$ we have $x'_S \neq y_S$. Observe that $(S \setminus R, x_{S\setminus R})$ is then a certificate for R: indeed for any $b \in \{0,1\}^R$ that agrees with $x'_{S\cap R}$ we have that for every $y \in X$ we have $y_R \neq b$. □

3 Lower Bounds

3.1 Middle Slice

In this section we prove the following.

Theorem 3.1

Let C be a circuit of the following form. $C := C_1 \vee C_2 \vee \ldots \vee C_N$ where C_i is a composition of a CNF D_i and an affine transformation A_i of full rank. Suppose that C computes the characteristic function of the middle slice $\binom{[n]}{n/2}$. Then the total size of CNFs $D_1, \ldots, D_N$ is at least $2^{\Omega(\sqrt{n})}$.

The key ingredient in our proof is a density boosting lemma for affine subspaces. Following a similar definition for boolean subcubes in [ALWZ20] we say that a set $X \subseteq S$ where S is a vector space is θ-linearly spread in S if for every affine subspace $A \subseteq S$ of co-dimension d we have

$$\Pr_{\boldsymbol{x}\sim X}[\boldsymbol{x} \in A] \leq \theta^{-d}.$$

A similar, but different notion was used in [KM23]. The following is a new density boosting lemma, generalized for affine subspaces, which might be of independent interest. For boolean cubes, such a lemma was introduced in [GLMWZ16] and appears among others in [CDGS18].

Lemma 3.2

For every set $X \subseteq \{0,1\}^n$ of size at least 2^{n-d} there exists an affine subspace A of $\{0,1\}^n$ of co-dimension at most $d/(1-\log_2\theta)$ such that $X \cap A$ is θ-spread in A and $|X \cap A| \geq |X|\theta^{-d/(1-\log_2\theta)}$.

Proof. Suppose that X is not θ-spread in $\{0,1\}^n$. Then consider the *affine* subspace of largest co-dimension A that witnesses the lack of θ-spreadness of X: suppose that A has co-dimension ℓ then we have $|X \cap A|/|X| > \theta^{-\ell}$.

We claim then that $X \cap A$ is θ-spread in A. Indeed, suppose that it is not, i.e. there exists an affine subspace B of A of co-dimension ℓ' (in A) such that $|X \cap B|/|X \cap A| > \theta^{-\ell'}$. Then $|X \cap B|/|X| > \theta^{-\ell'-\ell}$ which contradicts the maximality of the co-dimension of A.

Now it remains to bound the co-dimension of A. On the one hand

$$\Pr_{\boldsymbol{x}\sim X}[\boldsymbol{x} \in A] = \sum_{y\in A} \Pr[\boldsymbol{x} = y] \leq |A| \cdot 2^{d-n} = 2^{(n-\ell)+d-n} = 2^{d-\ell}.$$

On the other hand $\Pr_{\boldsymbol{x}\sim X}[\boldsymbol{x} \in A] > \theta^{-\ell}$. Hence $d > \ell(1-\log_2\theta)$. The lower bound on $|X\cap A|$ follows. □

Proof of Theorem 3.1. Suppose for contradiction that $C^{-1}(1) = \binom{[n]}{n/2}$, $N \leq 2^{\gamma\sqrt{n}}$ where γ is a constant to choose later. Since $C^{-1}(1) = \bigcup_{i\in[N]} C_i^{-1}(1)$, there exists $i_0 \in [N]$ such that $|C_{i_0}^{-1}(1)| \geq |\binom{[n]}{n/2}| \cdot 2^{-\gamma\sqrt{n}}$.

Thus, there exists a CNF D and a full-rank linear transformation A such that $X := (D \circ A)^{-1}(1)$ has size at least $\binom{n}{n/2} \cdot 2^{-\gamma\sqrt{n}} \geq 2^{n-\gamma\sqrt{n}-\log n}$ and $X \subseteq \binom{[n]}{n/2}$. Let us identify the linear transformation A with the matrix in $\{0,1\}^{n\times n}$ defining it: $A(x) := Ax$.

First we apply Theorem 3.2 to the set X with the parameter $\theta = \sqrt{2}$. We get that there exists an affine space B of co-dimension at most $2(\gamma\sqrt{n} + \log n)$ such that $X \cap B$ is θ-linearly spread in B and $|X \cap B| \geq |X|/2^{\gamma\sqrt{n}+\log n}$.

Applying Theorem 2.8 to the set $A(X \cap B) = \{Ax \mid x \in X \cap B\}$ with $q = 5\gamma\sqrt{n}$ we get:

$$\Pr_{(\boldsymbol{x},\boldsymbol{i})\sim(X\cap B)\times[n]}[A\boldsymbol{x} \text{ does not contain a size-}q \text{ certificate for } \boldsymbol{i} \text{ wrt } X \cap B] \geq 1 - O(\gamma\sqrt{n} \cdot q/n).$$

Pick γ such that this probability is at least 0.9. That is, with probability 0.9 for a pair $(\boldsymbol{x}, \boldsymbol{i}) \sim (X \cap B) \times [n]$ we have by Theorem 2.10 that $A\boldsymbol{x} + e_{\boldsymbol{i}}$ (where $e_i = 0^{i-1}10^{n-i}$) is a q-limit of the set $A(X \cap B)$. In order to invoke Theorem 2.4 we need to find many q-limits in $D^{-1}(0)$, i.e., outside of $A(X)$.

Now suppose that $Ax + e_i \in A(X)$ for some $(x, i) \in (X \cap B) \times [n]$. Equivalently $x + A^{-1}e_i \in X$. Then in particular $x + A^{-1}e_i \in \binom{[n]}{n/2}$. Since $x \in X \subseteq \binom{[n]}{n/2}$, this implies that $\langle x, A^{-1}e_i\rangle = 0$, otherwise x and $x + A^{-1}e_i$ have different parity and thus can not both belong to $\binom{[n]}{n/2}$.

Now consider the affine subspace $B'_i := \{y \in B \mid \langle y, A^{-1}e_i\rangle = 0\}$. If $A^{-1}e_i \in B^{\perp} := \{x \in \{0,1\}^n \mid \forall y \in B, \langle x, y\rangle \text{ is fixed}\}$ ($B^{\perp}$ is the span of all linear constraints defining B), then $B' = B$ or $B' = \emptyset$, otherwise B'_i has co-dimension 1 in B. Let E be the event when $A^{-1}e_{\boldsymbol{i}}$ is in $B^{\perp}$. We then have

$$\Pr[A\boldsymbol{x} + e_{\boldsymbol{i}} \in A(X)] \leq \Pr[A\boldsymbol{x} + e_{\boldsymbol{i}} \in A(X) \mid \neg E] + \Pr[E].$$

Since the co-dimension of B (equivalently the dimension of $B^{\perp}$) is at most $2(n^{\gamma} + \log n)$ and $A^{-1}e_1, \ldots, A^{-1}e_n$ are linearly independent, $\Pr[E] \leq 2(\gamma\sqrt{n} + \log n)/n = o(1)$. On the other hand since $A\boldsymbol{x}+e_{\boldsymbol{i}} \in A(X)$ implies that $\boldsymbol{x} \in B'_{\boldsymbol{i}}$ and $X \cap B$ is θ-spread in B we get $\Pr[A\boldsymbol{x}+e_{\boldsymbol{i}} \in A(x) \mid \neg E] \leq 1/\sqrt{2}$. Therefore $\Pr[A\boldsymbol{x} + e_{\boldsymbol{i}} \notin A(X) \wedge A\boldsymbol{x} + e_{\boldsymbol{i}} \text{ is a } q\text{-limit of } A(X \cap B)] \geq 0.9 - 1/\sqrt{2} - o(1)$, so there are $\Omega(|X \cap B|)$ q-limits to $A(X \cap B)$ outside of $A(X)$, hence by Theorem 2.4 we get that

$$|D| = \Omega(|X \cap B|)/2^{n-5\gamma\sqrt{n}} = \Omega(2^{n-2(\gamma\sqrt{n}+\log n)}/2^{n-5\gamma\sqrt{n}}) = \Omega(2^{\gamma\sqrt{n}}). \qquad \square$$

3.2 Affine Disperser

Theorem 3.3

Let γ be any constant in $(0, 1/3)$. Let $f\colon \{0,1\}^n \to \{0,1\}$ be an affine disperser for dimension $k := n^{\gamma}$ such that $|f^{-1}(1)| \geq 2^{n-n^{\gamma}}$ and C be a circuit of the form $C := C_1 \vee C_2 \vee \ldots \vee C_N$, where C_i is a composition of a CNF D_i and a linear transformation A_i of full rank. If C computes f, then the total size of $C_1, \ldots, C_N$ is at least $2^{\Omega(n^{\gamma})}$.

Proof. We show that either $N \geq 2^{n^{\gamma}}$ or one of the CNFs $C_1, \ldots, C_N$ has size at least $2^{n^{\gamma}}$, which yields the claim. Suppose $N < 2^{n^{\gamma}}$. Then there exists $C_i = D_i \circ A_i$ such that $|C_i^{-1}(1)| \geq |f^{-1}(1)|/2^{n^{\gamma}} \geq 2^{n-2n^{\gamma}}$.

Let $X := (D_i \circ A_i)^{-1}(1)$. For the set $A_i(X)$ we apply Theorem 2.9 with the following parameters:

- the density loss $t = 2n^{\gamma}$;
- the size of certificate $q = \alpha n^{\gamma}$;
- the size of the unpredictable block $r = \beta n^{\gamma}$.

The constants α, β are to be chosen later. It follows that with probability $1 - O(n^{3\gamma-1}) = 1 - o(1)$ for $\boldsymbol{x} \sim X$ and $\boldsymbol{R} \sim \binom{[n]}{r}$ there is no certificate of size q in $A_i\boldsymbol{x}$ for $\boldsymbol{R}$. Hence, by Theorem 2.10 all elements of

$$L_{\boldsymbol{x},\boldsymbol{R}} := \{y \in \{0,1\}^n \mid y_{[n]\setminus\boldsymbol{R}} = \boldsymbol{x}_{[n]\setminus\boldsymbol{R}}\}$$

are q-limits of $A_i(X)$ with probability $1 - o(1)$. Let E be the set of pairs $x, R \in X \times \binom{[n]}{r}$ for which this holds.

For $(x, R) \in E$ consider the set $A_i^{-1}(L_{x,R})$. $L_{x,R}$ is an r-dimensional affine subspace of $\{0,1\}^n$, thus $A_i^{-1}(L_{x,R})$ is as well. As f is an affine disperser, there is an input $y \in A_i^{-1}(L_{x,R}) \cap f^{-1}(0)$, hence $A_i \cdot y$ is not in X and is a q-limit of X.

Now we need to count the number of q-limits we got in order to invoke Theorem 2.4. Let $g\colon E \to \{0,1\}^n$ be the function mapping (x, R) to $A_i \cdot y$ (to an arbitrary one, if there are several). Let us upper bound $|g^{-1}(z)|$ for an arbitrary $z \in \{0,1\}^n$. Suppose $g(x, R) = z$, let $y = (A_i)^{-1}z$, then $x_{[n]\setminus R} = y_{[n]\setminus R}$, hence, there are at most $2^r \cdot \binom{[n]}{r}$ such preimages. Thus we get $|E|/(2^r\binom{[n]}{r}) = |X|/2^r$ q-limits in total. Therefore by Theorem 2.4 we get that $|D_i| \geq (1-o(1))2^{n-2n^{\gamma}}/(2^r \cdot 2^{n-q}) \geq 2^{q-r-2n^{\gamma}-1} \geq 2^{(\alpha-\beta-2)n^{\gamma}-1}$. Hence for any $\alpha > \beta + 2$ we get the desired bound. $\square$

3.3 Inner Product

In this section, we give an exponential lower bound for $\textsc{Or} \circ (\Pi^0 \circ \mathbb{B})$-circuit size required to compute the inner product. Our proof is a combination of bottom-up techniques with one top-down-like step.

Theorem 3.4

Let C be a circuit of the form $C := C_1 \vee C_2 \vee \ldots \vee C_N$ where each C_i is a composition of a d-depth circuit D_i composed with a full-rank affine mapping A_i.

Suppose that C computes IP_n. Then the total size of circuits $D_1, \ldots, D_N$ is at least $2^{n^{\Omega(1/d)}}$.

Proof. Suppose for contradiction that the total size of all D_i is at most 2^{n^ε} for $\varepsilon = o(1/d)$. Then let us pick $\boldsymbol{\alpha} \sim \{0,1\}^n$ and apply the restriction $y = \boldsymbol{\alpha}$ to C. Then $C|_{y=\boldsymbol{\alpha}}$ computes the function $\textsc{Xor}_{\boldsymbol{\alpha}} := \bigoplus_{i \in [n]\colon \boldsymbol{\alpha}_i = 1} x_i$ and has the same form as before, disjunction of compositions of constant-depth circuits with affine transformations. Let D'_i, A'_i be such that $C|_{y=\boldsymbol{\alpha}} = \bigvee_{i \in [N]} D'_i \circ A'_i$. Since $\textsc{Xor}_{\boldsymbol{\alpha}}$ is balanced, there exists $i_0 \in [N]$ such that $|(D'_{i_0} \circ A'_{i_0})^{-1}(1)| \geq 2^{n-1-n^\varepsilon}$. On the other hand $(D'_i \circ A'_i)^{-1}(0) \supseteq \textsc{Xor}^{-1}_{\boldsymbol{\alpha}}(0)$ for all $i \in [N]$. Then applying $(A'_{i_0})^{-1}$ to $D'_{i_0} \circ A'_{i_0}$ and to $\textsc{Xor}_{\boldsymbol{\alpha}}$ on the right we get that the circuit $(D'_{i_0})^{-1}(0) \supseteq \textsc{Xor}_{\boldsymbol{\alpha}'}$ where $\boldsymbol{\alpha}' = \boldsymbol{\alpha} \cdot (A'_{i_0})^{-1}$ and since A'_{i_0} has full rank $|(D'_{i_0})^{-1}(1)| = |(D'_{i_0} \circ A'_{i_0})^{-1}(1)| \geq 2^{n-1-n^\varepsilon}$.

Now observe that

$$\Pr[|\boldsymbol{\alpha}'| \leq \sqrt{n}] = \Pr[\boldsymbol{\alpha} \text{ is a linear combination of } \leq \sqrt{n} \text{ rows of } (A'_{i_0})^{-1}] \leq N \cdot \binom{n}{\sqrt{n}} / 2^n = o(1).$$

Hence, there exists a depth-d de Morgan circuit $D = D'_{i_0}$ that computes parity $\textsc{Xor}_{\alpha_0}$ on $\sqrt{n}$ bits correctly on all 0-inputs and on at least 2^{-n^ε}-fraction of 1-inputs. Then let $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M \sim \textsc{Xor}^{-1}_{\alpha_0}(0)$ be independent random variables. Then the depth-$(d+1)$ circuit $E_{\boldsymbol{y}}(x) := \bigvee_{i \in [M]} D(x \oplus \boldsymbol{y}_i)$ computes the value of $\textsc{Xor}_{\alpha_0}$ correctly on an input x with probability $1 - (1 - 2^{-n^\varepsilon})^M$, which exceeds $1 - 2^{-n}$ for $M > 3n \cdot 2^{n^\varepsilon}$, hence, there exists a setting of $y = y_1, \ldots, y_M$ such that E_y computes $\textsc{Xor}_{\alpha_0}$. Thus by [Hås86] the size of E_y is at least $2^{n^{\Omega(1/d)}}$ which means that $\varepsilon = \Omega(1/d)$ which is a contradiction. □

3.4 Proof of Theorem 1.3

A $\textsc{DNF} \circ \textsc{Xor}$ circuit computing f is equivalent to a covering of $f^{-1}(1)$ by affine subspaces $f^{-1}(1) = \bigcup_{i \in [N]} A_i$. Consider an arbitrary $A_j \subseteq \{0,1\}^{n \times 3}$. Affine spaces over $\mathbb{F}_2$ are closed under sums of three elements, so let $a, b, c \in A_j$. Then $d = a \oplus b \oplus c \in A_j$. For $x \in f^{-1}(1)$ we have that for every $i \in [n]$ among $x_{i,1}, x_{i,2}, x_{i,3}$ exactly one value is 1 and the other two are zeroes. For $x \in f^{-1}(1)$ we can define $\bar{x} \in [3]^n$ be such that for every $i \in [n]$ we have $x_{i,\bar{x}_i} = 1$ and $x_{i,j} = 0$ if $j \neq \bar{x}_i$. Then since $d \in A_i \subseteq f^{-1}(1)$ for every $i \in [n]$ we have $|\{\bar{a}_i, \bar{b}_i, \bar{c}_i\}| < 3$, since otherwise $d_i = (1,1,1)$ which contradicts $f(d) = 1$. Since this is true for any $a, b, c \in A_j$ we get that there exists $\beta \in [3]^n$ such that for every $x \in A_j$ and for every $i \in [n]$ we have $x_{i,\beta(i)} = 0$. Since for every $x \in f^{-1}(1)$ we have $x_{i,1} + x_{i,2} + x_{i,3} = 1$ we get that $|A_j| \leq 2^n$. Since $|f^{-1}(1)| = 3^n$ we get that $N \geq (3/2)^n$, which completes the proof.

4 Discussion and Open Problems

In the results above, having an extra $\textsc{Or}$ on top of the circuit, and only fixing the linear transformation in the subcircuits, can be interpreted in the following way. When implementing the top-down strategy, the first choice of the subcircuit (and the subset of 1-inputs, respectively), does not depend on the specific linear forms used in the circuit.

In other words, let $A = f^{-1}(1)$ and $B = f^{-1}(0)$ for one of the hard functions considered in the main section. Informally, we prove that for any covering of A by no more 2^{n^ε} sets $A_1, \ldots, A_{2^{n^\varepsilon}}$ (for some $\varepsilon = \Omega(1)$) there is a choice of A_i such that for any affine map L there is a k-limit for A_i in B with respect to that map. Note that for different affine maps, we might find different k-limits. For proving lower bounds for $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$, we would need to prove a statement where the last two quantifiers are in a different order: *there is a k-limit that works for any affine map.* Or, at least, for any affine map in a large enough collection of such. As mentioned in section 2, this corresponds to making two "oblivious" steps down the

circuit, where we do not use the knowledge of specific parity gates, while in our lower bound, we make one "oblivious" step.

Note that this is not true for the affine extractors in general, as there is an affine extractor computable by $\vee \circ \wedge \circ \vee \circ \textsc{Xor}$ circuits [HIV22]. The middle slice function, however, could still be a good example for honing the top-down techniques.

The first natural step could be finding the same k-limit with respect to an arbitrary *pair* of affine maps. This corresponds to the lower bounds in the following computational model.

Open Problem. Prove top-down lower bounds for the class of $(\text{2-DNF}) \circ (\text{CNF} \circ \mathbb{B})$ circuits.

In fact, that would already be a step towards another elusive circuit class .Another motivation for this open question comes from the $\textsc{Mod}_6$ perspective. A $\textsc{Mod}_6$ gate can be seen as a conjunction of $\textsc{Mod}_2$ and $\textsc{Mod}_3$ gates. After appropriately expanding the brackets, $\text{DNF} \circ \textsc{Mod}_6$ can be transformed to a 2-DNF of conjunctions such that in each conjunction there are only $\textsc{Mod}_2$ or only $\textsc{Mod}_3$ gates. So a first related problem would be to adapt the technique for $\textsc{Mod}_3$.

Open Problem. Prove top-down lower bounds for subclasses of $\mathbf{AC}^0 \circ \textsc{Mod}_3$.

The next step would be combining the two last problems together. Let $\mathbb{L} \subseteq \{f \colon \{0,1\}^n \to \{0,1\}^n\}$ be the union of linear maps over $\mathbb{F}_2$ ($\mathbb{B}$) and maps $x \mapsto (\mathbb{1}_{[(Ax)_i = a_i]})_{i \in [n]}$ where $A \in \mathbb{F}_3^{n \times n}$ and $a \in \mathbb{F}_3^n$. In other words, we can choose a transformation of the inputs that either uses only $\textsc{Xor}$ operations, or only $\textsc{Mod}_3$ operations.

Open Problem. Prove lower bounds for $(\text{2-DNF}) \circ (\text{CNF} \circ \mathbb{L})$ circuits.

Solving this problem would imply lower bounds for $\text{DNF} \circ \textsc{Mod}_6$.

Claim 4.1

Let $f \colon \{0,1\}^n \to \{0,1\}$ be such that it is computable by k-$\text{DNF} \circ \textsc{Mod}_6$-circuit D of size s. Then it is computable by $(\text{2-DNF}) \circ (\text{CNF} \circ \mathbb{L})$ of size $O(s \cdot 2^k \cdot k)$.

Proof. See appendix A. □

When proving the results of this form, it all essentially boils down to finding a k-(parity) limit. The two known techniques for this are (robust) sunflowers or spreadness [HJP95; GRSS23] and unpredictability [MW19; GRSS23]. They both have certain downsides. For starters, these techniques only find k-limits that are close to the set in Hamming distance (or, in the case of our result, they are close in Hamming distance after a certain affine transformation). In principle, this might not be the case.

Open Problem. Let $A \subseteq \{0,1\}^n$ be a subset of a code with minimum distance $d = \Omega(n/\log n)$. Can you find a $\log^2(n)$-limit of A?

When looking for a k-limit, we can also ask for some structure of the considered set. Let us say that our set A is a half of a $\sqrt{n}$-wise independent set. From the results of Bazzi [Baz09], Razborov [Raz09], and Braverman [Bra11], we know that roughly half the points of the whole boolean cube should be n^ε-limits of the set A. However, current techniques do not allow us to find some explicit k-limit, assuming the knowledge of A. One of the reasons for this is that the size of such sets can be as small as $2^{O(\sqrt{n}\log n)}$ [ABI86].

Open Problem. Prove a top-down lower bound $2^{k^{\Omega(1)}}$ for separating two disjoint k-wise independent sets by depth-3 $\mathbf{AC}^0$ circuits.

Acknowledgements

We thank Mika Göös for fruitful discussions. We also thank Pengxiang Wang and anonymous reviewers for useful comments on the text. Authors are supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00026.

References

[ABI86] Noga Alon, László Babai, and Alon Itai. "A Fast and Simple Randomized Parallel Algorithm for the Maximal Independent Set Problem." In: *J. Algorithms* 7.4 (1986), pp. 567–583. DOI: 10.1016/0196-6774(86)90019-2. URL: https://doi.org/10.1016/0196-6774(86)90019-2.

[Ajt83] Miklos Ajtai. "Σ_1^1-Formulae on finite structures." In: *Annals of Pure and Applied Logic* 24.1 (1983), pp. 1–48. DOI: 10.1016/0168-0072(83)90038-6.

[ALWZ20] Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. "Improved bounds for the sunflower lemma." In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2020. Chicago, IL, USA: Association for Computing Machinery, 2020, pp. 624–630. ISBN: 9781450369794. DOI: 10.1145/3357713.3384234. URL: https://doi.org/10.1145/3357713.3384234.

[Baz09] Louay M. J. Bazzi. "Polylogarithmic Independence Can Fool DNF Formulas." In: *SIAM J. Comput.* 38.6 (2009), pp. 2220–2272. DOI: 10.1137/070691954. URL: https://doi.org/10.1137/070691954.

[BGM06] Elmar Böhler, Christian Glaßer, and Daniel Meister. "Error-Bounded Probabilistic Computations Between MA and AM." In: *Journal of Computer and System Sciences* 72.6 (2006), pp. 1043–1076. DOI: 10.1016/j.jcss.2006.05.001.

[Bra11] Mark Braverman. "Poly-logarithmic independence fools bounded-depth boolean circuits." In: *Commun. ACM* 54.4 (2011), pp. 108–115. DOI: 10.1145/1924421.1924446. URL: https://doi.org/10.1145/1924421.1924446.

[BS79] Theodore Baker and Alan Selman. "A second step toward the polynomial hierarchy." In: *Theoretical Computer Science* 8.2 (1979), pp. 177–187. DOI: 10.1016/0304-3975(79)90043-4.

[CDGS18] Sandro Coretti, Yevgeniy Dodis, Siyao Guo, and John Steinberger. "Random Oracles and Non-uniformity." In: *Advances in Cryptology – EUROCRYPT 2018*. Ed. by Jesper Buus Nielsen and Vincent Rijmen. Cham: Springer International Publishing, 2018, pp. 227–258. ISBN: 978-3-319-78381-9.

[CGJWX18] Mahdi Cheraghchi, Elena Grigorescu, Brendan Juba, Karl Wimmer, and Ning Xie. "$AC^0 \circ MOD_2$ lower bounds for the Boolean Inner Product." In: *J. Comput. Syst. Sci.* 97 (2018), pp. 45–59. DOI: 10.1016/J.JCSS.2018.04.006. URL: https://doi.org/10.1016/j.jcss.2018.04.006.

[CS16a] Gil Cohen and Igor Shinkar. "The Complexity of DNF of Parities." In: *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, Cambridge, MA, USA, January 14-16, 2016*. Ed. by Madhu Sudan. ACM, 2016, pp. 47–58. DOI: 10.1145/2840728.2840734. URL: https://doi.org/10.1145/2840728.2840734.

[CS16b] Gil Cohen and Igor Shinkar. "The Complexity of DNF of Parities." In: *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, Cambridge, MA, USA, January 14-16, 2016*. Ed. by Madhu Sudan. ACM, 2016, pp. 47–58. DOI: 10.1145/2840728.2840734. URL: https://doi.org/10.1145/2840728.2840734.

[CW22] Brynmor Chapman and R. Ryan Williams. "Smaller ACC0 Circuits for Symmetric Functions." In: *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*. Ed. by Mark Braverman. Vol. 215. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022, 38:1–38:19. DOI: 10.4230/LIPICS.ITCS.2022.38. URL: https://doi.org/10.4230/LIPIcs.ITCS.2022.38.

[FGT22] Peter Frankl, Svyatoslav Gryaznov, and Navid Talebanfard. "A Variant of the VC-Dimension with Applications to Depth-3 Circuits." In: *Proceedings of the 13th Conference on Innovations in Theoretical Computer Science (ITCS)*. Vol. 215. Schloss Dagstuhl, 2022, 72:1–72:19. DOI: 10.4230/LIPIcs.ITCS.2022.72.

[FSS84] Merrick Furst, James Saxe, and Michael Sipser. "Parity, circuits, and the polynomial-time hierarchy." In: *Mathematical Systems Theory* 17.1 (1984), pp. 13–27. DOI: 10.1007/bf01744431.

[GGM23] Mika Göös, Ziyi Guan, and Tiberiu Mosnoi. "Depth-3 Circuits for Inner Product." In: *48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023)*. Vol. 272. Schloss Dagstuhl, 2023, 51:1–51:12. ISBN: 978-3-95977-292-1. DOI: 10.4230/LIPIcs.MFCS.2023.51.

[GKW21] Alexander Golovnev, Alexander Kulikov, and Ryan Williams. "Circuit Depth Reductions." In: *Proceedings of the 12th Conference on Innovations in Theoretical Computer Science (ITCS)*. Vol. 185. Schloss Dagstuhl, 2021, 24:1–24:20. DOI: 10.4230/LIPIcs.ITCS.2021.24.

[GLMWZ16] Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. "Rectangles Are Nonnegative Juntas." In: *SIAM Journal on Computing* 45.5 (2016), pp. 1835–1869. DOI: 10.1137/15M103145X. eprint: https://doi.org/10.1137/15M103145X. URL: https://doi.org/10.1137/15M103145X.

[GRSS23] Mika Göös, Artur Riazanov, Anastasia Sofronova, and Dmitry Sokolov. "Top-Down Lower Bounds for Depth-Four Circuits." In: *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*. 2023, pp. 1048–1055. DOI: 10.1109/FOCS57990.2023.00063.

[Hås86] Johan Håstad. "Almost optimal lower bounds for small depth circuits." In: *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*. STOC '86. Berkeley, California, USA: Association for Computing Machinery, 1986, pp. 6–20. ISBN: 0897911938. DOI: 10.1145/12130.12132. URL: https://doi.org/10.1145/12130.12132.

[Hås87] Johan Håstad. "Computational Limitations for Small Depth Circuits." PhD thesis. MIT, 1987.

[Hir17] Suichi Hirahara. *A Duality Between Depth-Three Formulas and Approximation by Depth-Two*. Tech. rep. arXiv, 2017. eprint: 1705.03588.

[HIV22] Xuangui Huang, Peter Ivanov, and Emanuele Viola. "Affine Extractors and AC0-Parity." In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022)*. Ed. by Amit Chakrabarti and Chaitanya Swamy. Vol. 245. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 9:1–9:14. ISBN: 978-3-95977-249-5. DOI: 10.4230/LIPIcs.APPROX/RANDOM.2022.9. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.APPROX/RANDOM.2022.9.

[HJP95] Johan Håstad, Stasys Jukna, and Pavel Pudlák. "Top-down lower bounds for depth-three circuits." In: *Computational Complexity* 5.2 (1995), pp. 99–112. DOI: 10.1007/bf01268140.

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. "Which Problems Have Strongly Exponential Complexity?" In: *Journal of Computer and System Sciences* 63.4 (Dec. 2001), pp. 512–530. DOI: 10.1006/jcss.2001.1774.

[Juk06] Stasys Jukna. "On Graph Complexity." In: *Comb. Probab. Comput.* 15.6 (2006), pp. 855–876. DOI: 10.1017/S0963548306007620. URL: https://doi.org/10.1017/S0963548306007620.

[KM23] Zander Kelley and Raghu Meka. "Strong Bounds for 3-Progressions." In: *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2023, Santa Cruz, CA, USA, November 6-9, 2023*. IEEE, 2023, pp. 933–973. DOI: 10.1109/FOCS57990.2023.00059. URL: https://doi.org/10.1109/FOCS57990.2023.00059.

[Ko90] Ker-I Ko. "Separating and collapsing results on the relativized probabilistic polynomial-time hierarchy." In: *Journal of the ACM* 37.2 (1990), pp. 415–438. DOI: 10.1145/77600.77623.

[KW90] Mauricio Karchmer and Avi Wigderson. "Monotone Circuits for Connectivity Require Super-Logarithmic Depth." In: *SIAM J. Discret. Math.* 3.2 (1990), pp. 255–265. DOI: 10.1137/0403021. URL: https://doi.org/10.1137/0403021.

[MW19] Or Meir and Avi Wigderson. "Prediction from Partial Information and Hindsight, with Application to Circuit Lower Bounds." In: *Computational Complexity* 28.2 (2019), pp. 145–183. DOI: 10.1007/s00037-019-00177-4.

[OSS19] Igor Carboni Oliveira, Rahul Santhanam, and Srikanth Srinivasan. "Parity Helps to Compute Majority." In: *34th Computational Complexity Conference, CCC 2019, July 18-20, 2019, New Brunswick, NJ, USA*. Ed. by Amir Shpilka. Vol. 137. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019, 23:1–23:17. DOI: 10.4230/LIPICS.CCC.2019.23. URL: https://doi.org/10.4230/LIPIcs.CCC.2019.23.

[PPSZ05] Ramamohan Paturi, Pavel Pudlák, Michael Saks, and Francis Zane. "An improved exponential-time algorithm for k-SAT." In: *Journal of the ACM* 52.3 (2005), pp. 337–364. DOI: 10.1145/1066100.1066101.

[PPZ99] Ramamohan Paturi, Pavel Pudlak, and Francis Zane. "Satisfiability Coding Lemma." In: *Chicago Journal of Theoretical Computer Science* 5.1 (1999), pp. 1–19. DOI: 10.4086/cjtcs.1999.011.

[PSZ00] Ramamohan. Paturi, Michael Saks, and Francis Zane. "Exponential lower bounds for depth three Boolean circuits." In: *computational complexity* 9.1 (2000), pp. 1–15. DOI: 10.1007/PL00001598.

[Raz09] Alexander A. Razborov. "A Simple Proof of Bazzi's Theorem." In: *ACM Trans. Comput. Theory* 1.1 (2009), 3:1–3:5. DOI: 10.1145/1490270.1490273. URL: https://doi.org/10.1145/1490270.1490273.

[Raz87] Alexander Razborov. "Lower bounds on the size of bounded depth circuits over a complete basis with logical addition." In: *Mathematical Notes of the Academy of Sciences of the USSR* 41.4 (1987), pp. 333–338. DOI: 10.1007/bf01137685.

[RS98] Alexander Russell and Ravi Sundaram. "Symmetric alternation captures BPP." In: *Computational Complexity* 7.2 (Nov. 1998), pp. 152–162. DOI: 10.1007/s000370050007.

[San89] Miklos Santha. "Relativized Arthur–Merlin versus Merlin–Arthur Games." In: *Information and Computation* 80.1 (1989), pp. 44–49. DOI: 10.1016/0890-5401(89)90022-9.

[Sip84] Michael Sipser. "A Topological View of Some Problems in Complexity Theory." In: *Mathematical Foundations of Computer Science 1984, Praha, Czechoslovakia, September 3-7, 1984, Proceedings*. Ed. by Michal Chytil and Václav Koubek. Vol. 176. Lecture Notes in Computer Science. Springer, 1984, pp. 567–572. DOI: 10.1007/BFB0030341. URL: https://doi.org/10.1007/BFb0030341.

[Smo87] Roman Smolensky. "Algebraic methods in the theory of lower bounds for Boolean circuit complexity." In: *Proceedings of the 19th Symposium on Theory of Computing (STOC)*. ACM Press, 1987. DOI: 10.1145/28395.28404.

[Tal17] Avishay Tal. "Tight Bounds on the Fourier Spectrum of AC0." In: *32nd Computational Complexity Conference, CCC 2017, July 6-9, 2017, Riga, Latvia*. Ed. by Ryan O'Donnell. Vol. 79. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017, 15:1–15:31. DOI: 10.4230/LIPICS.CCC.2017.15. URL: https://doi.org/10.4230/LIPIcs.CCC.2017.15.

[Val77] Leslie G. Valiant. "Graph-Theoretic Arguments in Low-Level Complexity." In: *Mathematical Foundations of Computer Science 1977, 6th Symposium, Tatranska Lomnica, Czechoslovakia, September 5-9, 1977, Proceedings*. Ed. by Jozef Gruska. Vol. 53. Lecture Notes in Computer Science. Springer, 1977, pp. 162–176. DOI: 10.1007/3-540-08353-7_135. URL: https://doi.org/10.1007/3-540-08353-7%5C_135.

[Wol06] Guy Wolfovitz. "The complexity of depth-3 circuits computing symmetric Boolean functions." In: *Information Processing Letters* 100.2 (Oct. 2006), pp. 41–46. DOI: 10.1016/j.ipl.2006.06.008.

[Yao85] Andrew Yao. "Separating the polynomial-time hierarchy by oracles." In: *26th Annual Symposium on Foundations of Computer Science (SFCS)*. IEEE, 1985. DOI: 10.1109/sfcs.1985.49.

A Proof of theorem 4.1

Consider a term of D. We rewrite it as a 2-CNF using MOD_2 and MOD_3 gates:

$$\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\alpha_j^i x_j)\bmod 6 = a_i\right]\wedge\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\beta_j^i x_j)\bmod 6 \neq b_i\right] =$$

$$\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\alpha_j^i x_j)\bmod 2 = a_i \bmod 2 \wedge (\sum_{j\in[n]}\alpha_j^i x_j)\bmod 3 = a_i \bmod 3\right]$$

$$\wedge\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\beta_j^i x_j)\bmod 2 \neq b_i \bmod 2 \vee (\sum_{j\in[n]}\beta_j^i x_j)\bmod 3 \neq b_i \bmod 3\right]$$

Here $\alpha, \beta \subseteq \mathbb{Z}_6^{k\times n}$, $a, b \in \mathbb{Z}_6^k$. A 2-CNF with k terms can be transformed into a DNF of size $2^k \cdot k$. Now, any term of that DNF has the following form:

$$\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\gamma_j^i x_j)\bmod 2 = a_i\right]\wedge\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\delta_j^i x_j)\bmod 2 \neq b_i\right]\wedge$$

$$\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\varepsilon_j^i x_j)\bmod 3 = c_i\right]\wedge\bigwedge_{i\in[k]}\left[(\sum_{j\in[n]}\varphi_j^i x_j)\bmod 3 \neq d_i\right] =$$

$$\bigwedge_{i\in[k]} A(x)_i \wedge \bigwedge_{i\in[k]} B(x)_i$$

Here A and B are transformations from $\mathbb{L}$, $\gamma, \delta \subseteq \mathbb{F}_2^{k\times n}$; $\varepsilon, \varphi \in \mathbb{F}_3^{k\times n}$ and $a, b \in \mathbb{F}_3^k$, $c, d \in \mathbb{F}_3^k$. Overall, f is then computable by a circuit of the following form:

$$\bigvee_{t\in D}\bigwedge_{i\in[k]} A_t(x)_i \wedge \bigwedge_{i\in[k]} B_t(x)_i$$

This is a $(\text{2-DNF}) \circ (\mathsf{CNF} \circ \mathbb{L})$ circuit of size no more that $O(s \cdot 2^k \cdot k)$.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il