Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 49 (2023)

Top-Down Lower Bounds for Depth-Four Circuits

Mika Göös *EPFL*

Artur Riazanov *EPFL*

Anastasia Sofronova *EPFL*

Dmitry Sokolov *EPFL*

April 29, 2024

Abstract. We present a top-down lower-bound method for depth-4 boolean circuits. In particular, we give a new proof of the well-known result that the parity function requires depth-4 circuits of size exponential in $n^{1/3}$. Our proof is an application of robust sunflowers and block unpredictability.

1 Introduction

The working complexity theorist has three main weapons in their arsenal when proving lower bounds against small-depth boolean circuits (consisting of $\wedge$, $\vee$, $\neg$ gates of unbounded fanin). The most wildly successful ones are the *random restriction* method [FSS84, Ajt83] and the *polynomial approximation* method [Raz87, Smo87]. The random restriction method, in particular, is applied ***bottom-up***: it starts by analysing the bottom-most layer of gates next to input variables and finds a way to simplify the circuit so as to reduce its depth by one. The third main weapon, which is the subject of this paper, is the ***top-down*** method: starting at the top (output) gate we walk down the circuit in search of a mistake in the computation. Top-down methods (often phrased in the language of communication complexity [KW90]) are routinely used to prove:

1. Polynomial lower bounds for formulas [KRW95, EIRS01, GMWW17, DM17, dRMN$^+$20].
2. Exponential lower bounds for *monotone* circuits and formulas [KW90, RW92, NW93, GS95, RM99, GP18, dRNV16, PR17, GGKS20, dRGR22].
3. Exponential lower bounds for depth-3 circuits [BS79, San89, Ko90, HJP95, RS98, PPZ99, PSZ00, IPZ01, PPSZ05, Wol06, BGM06, MW19, GKW21, FGT22, GGM23].

In particular, it has been an open problem (posed in [HJP95, MW19]) to prove exponential lower bounds for depth-4 circuits by a top-down argument. We develop such a lower-bound method in this paper and use it to prove a lower bound for the parity function. It has been long known using bottom-up methods that the depth-4 complexity of n-bit parity is $2^{\Theta(n^{1/3})}$ [Yao85, Hås87]. We recover a slightly weaker bound.

Theorem 1. *Every depth-4 circuit computing the n-bit parity requires $2^{n^{1/3-o(1)}}$ gates.*

Our top-down proof of this theorem is a relatively simple application of two known techniques: robust sunflowers [Ros14, ALWZ21, Rao20] and unpredictability from partial information [MW19, ST18, Vio21], which we generalise to blocks of coordinates (obtaining essentially best possible parameters).

A major motivation for the further development of top-down methods is that the method is, in a precise sense, *complete* for constant-depth circuits, in that it can be used to prove tight lower bounds (up to polynomial factors) for *any* boolean function; see Remark 1. The same is not known to hold for the aforementioned bottom-up techniques. For example, there is currently no known bottom-up proof for the depth-3 circuit lower bound that underlies the oracle separation $\mathsf{AM} \not\subseteq \Sigma_2\mathsf{P}$ [San89, Ko90, BGM06]. We suspect more generally that top-down methods could prove useful in settings where the bottom-up methods have failed so far, such as proving lower bounds against $\mathsf{AC}^0 \circ \oplus$ circuits computing inner-product [CGJ$^+$18, ER22, HIV22, SV12] or against the polynomial hierarchy in communication complexity [BFS86].

ISSN 1433-8092

2 Proof overview

Before we explain our new top-down lower bound method for depth-4 circuits, let us review one particular top-down technique for depth-3 circuits that we build on. Namely, a lower bound using the information-theoretic *unpredictability lemma* of Meir and Wigderson [MW19, ST18, Vio21].

2.1 Depth-3 via bit unpredictability

Suppose $X \subseteq \{0,1\}^n$ is a set of n-bit strings and $i \in [n]$ a coordinate. A *certificate* for i with respect to X is a pair (Q, a) such that $Q \subseteq [n] \smallsetminus \{i\}$, $a \in \{0,1\}^Q$, and there exists a bit $b \in \{0,1\}$ such that every $x \in X$ with $x_Q = a$ satisfies $x_i = b$. In words, the partial assignment defined by (Q, a) will correctly predict that the i-th bit must be b. The *size* of the certificate is $|Q|$. We also say that x *contains* a size-q certificate for i (wrt X) if there is some size-q certificate of the form (Q, x_Q) for i. Meir and Wigderson [MW19] proved that if X is a large set, a uniform random string $\boldsymbol{x} \sim X$ will not contain a small certificate for a uniform random coordinate $\boldsymbol{i} \sim [n]$ with high probability. There is also an alternative proof by Smal and Talebanfard [ST18] which tightens the parameters.

Lemma 1 (Bit unpredictability [MW19])**.** *Let $X \subseteq \{0,1\}^n$ have density $|X|/2^n \geq 2^{-k}$. Then for any $q \geq 1$,*

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

Let us apply this lemma to show a $2^{\Omega(\sqrt{n})}$ depth-3 lower bound for the n-bit parity function Xor. Suppose for the sake of contradiction that Σ is a depth-3 circuit of size $|\Sigma| = 2^{o(\sqrt{n})}$ for Xor. We may assume that Σ is of type $\vee \circ \wedge \circ \vee$, that is, the circuit has an $\vee$-gate at the top, followed by a layer of $\wedge$-gates, then a layer of $\vee$-gates, and finally we have the input literals at the bottom. Assume for simplicity of exposition that the bottom fanin of Σ is at most $\sqrt{n}$. (If we are allowed to invoke a bottom-up trick, then this bottom-fanin assumption can be easily ensured by restricting a small fraction of input variables [HJP95, Lemma 3.2].) That is, Σ computes an Or of $\sqrt{n}$-CNFs. Because the top gate is $\vee$, we have $\text{Xor}^{-1}(1) = \bigcup_{j\in[s]} \Pi_i^{-1}(1)$ where $s \leq |\Sigma|$ is the top fanin and Π_j is the j-th CNF feeding into the top gate. We now take a naive greedy step down the circuit: we choose any j that maximises $|\Pi_j^{-1}(1)|$ and set $\Pi \coloneqq \Pi_j$. In summary,

- Π accepts the set $X \coloneqq \Pi^{-1}(1)$ of density $|X|/2^n \geq |\text{Xor}^{-1}(1)|/(2^n|\Sigma|) \geq 1/(2|\Sigma|) \geq 2^{-o(\sqrt{n})}$.
- Π rejects the set $\text{Xor}^{-1}(0)$.

We can now apply Lemma 1 to the set X with parameters $k \coloneqq o(\sqrt{n})$ and $q \coloneqq \sqrt{n}$. As a result, there exist some string $x^* \in X$ and a coordinate $i \in [n]$ such that x^* does not contain any size-q certificates for i wrt X. Consider the string $y \in \text{Xor}^{-1}(0)$ that is obtained from x^* by flipping the i-th bit. We claim that y is locally indistinguishable from strings in X—often y is called a *local limit* of X—in the following sense.

Claim 1 (Local limit)**.** *For every $Q \subseteq [n]$, $|Q| \leq \sqrt{n}$, there exist an $x \in X$ such that $x_Q = y_Q$.*

Proof. Let $Q' \coloneqq Q \smallsetminus \{i\}$. Since $(Q', x^*_{Q'})$ is not a certificate for i wrt X, we get that, for the bit $b \coloneqq 1 - x^*_i$, there exist some $x \in X$ such that $x_{Q'} = x^*_{Q'}$ and $x_i = b$. But this means $x_Q = y_Q$ (see Figure 1). □

We finally claim that Π accepts $y \in \text{Xor}^{-1}(0)$ which is the desired contradiction. To show this, we need to show that every clause of Π accepts y. Consider any clause Λ of Π and let $Q \subseteq [n]$, $|Q| \leq \sqrt{n}$, be the set of variables mentioned in Λ. By Claim 1 there is some $x \in X$ such that $x_Q = y_Q$. But since Λ accepts x, it must also accept y, as desired. This concludes the proof that Xor requires depth-3 circuits of size $2^{\Omega(\sqrt{n})}$.

2.2 Depth-4 via block unpredictability

Our depth-4 lower bound will encounter a number of new challenges compared to the depth-3 proof above.

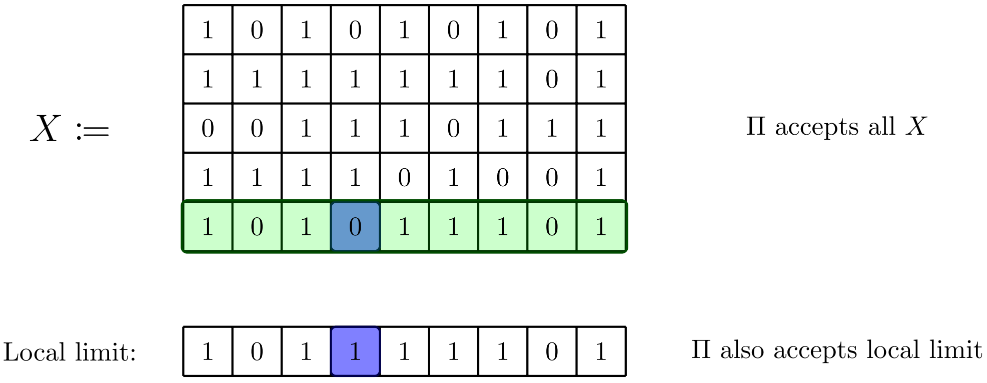

Figure 1: Flipping one bit yields a local limit

Challenge 1: Which subcircuit to pick? The first challenge is that we have to be more careful how to take steps down the circuit. Suppose for contradiction Π is a small depth-4 circuit of type $\wedge \circ \vee \circ \wedge \circ \vee$ computing Xor. Our first step will still be naively greedy: we choose the depth-3 subcircuit Σ of Π that rejects as large a set $Y \subseteq \text{Xor}^{-1}(0)$ as possible. The second step is trickier. Intuitively, we should choose a depth-2 subcircuit (namely, a CNF) of Σ that accepts not just many 1-inputs but in particular such 1-inputs that are hard to distinguish from Y. To this end, we first construct a set $M \subseteq \text{Xor}^{-1}(1)$ that ***mirrors*** the set $Y \subseteq \text{Xor}^{-1}(0)$ in the sense that M is hard to locally distinguish from Y. Concretely, let us define (as a first attempt; to be adjusted later) the mirror set M as the set of 1-inputs x such that the $n^{1/3}$-radius Hamming ball around x contains many points from Y. Our second step down the circuit will then be to choose any CNF Γ that accepts a large fraction of the mirror set M. We illustrate the process in Figure 2.

Remark 1 (Completeness of the top-down method). Choosing the mirror set M is the crux of the top-down argument. It follows from linear programming duality that there always exist a mirror set such that any subcircuit that accepts a large enough fraction of M will make a mistake further down the circuit. This means that if we only knew how to construct mirror sets, we could prove a tight lower bound (up to polynomial factors) for constant-depth circuits for any boolean function. We refer to [Hir17, GGM23] for more formal discussion of the completeness of the top-down method.

Challenge 2: Block unpredictability. Given a CNF Γ that accepts a large subset $X \subseteq M$, we would next like to show the existence of a local limit $y \in \text{Xor}^{-1}(0)$ of X. To reach a contradiction, it is important that the local limit lies in the target set Y; note that Γ is only guaranteed to reject inputs in Y, not every input in $\text{Xor}^{-1}(0)$. Our guarantee about $X \subseteq M$ is that every $x \in X$ contains many points in Y at Hamming distance at most $n^{1/3}$. Thus, starting from $x \in X$ and trying to reach a point in Y, we may need to flip a whole block of at most $n^{1/3}$ bits. Meir and Wigderson [MW19] already generalised their bit unpredictability lemma to larger blocks of bits (essentially by iteratively applying their lemma for a single bit). However, their generalisation did not supply unpredictable blocks with high enough probability. In this paper, we extend their proof and give a block unpredictability lemma with essentially optimal parameters.

To state our lemma, we first generalise the definition of certificates to whole blocks of bits. Let $X \subseteq \{0,1\}^n$ be a set of n-bit strings and $R \in \binom{[n]}{r} := \{A \subseteq [n] : |A| = r\}$ a block of r coordinates. A *certificate for* R (wrt X) is a pair (Q, a) such that $Q \subseteq [n] \smallsetminus R$, $a \in \{0,1\}^Q$ and there exists $b \in \{0,1\}^R$ such that every $x \in X$ with $x_Q = a$ satisfies $x_R \neq b$. In words, the partial assignment defined by (Q, a) will correctly predict that some r-bit string b is missing over the coordinates R. We prove the following in Section 4.

Lemma 2 (Block unpredictability). *Let $X \subseteq \{0,1\}^n$ have density $|X|/2^n \geq 2^{-k}$. Then for any $r, q \geq 1$,*

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

Remark 2 (Optimality of Lemma 2). We note that Lemma 2 is optimal in that if $kqr = \Omega(n)$ then there are examples of sets X such that the probability (1) is $\Omega(1)$. Let $B_1, \ldots, B_k \subseteq [n]$, $|B_i| = q$, be any pairwise

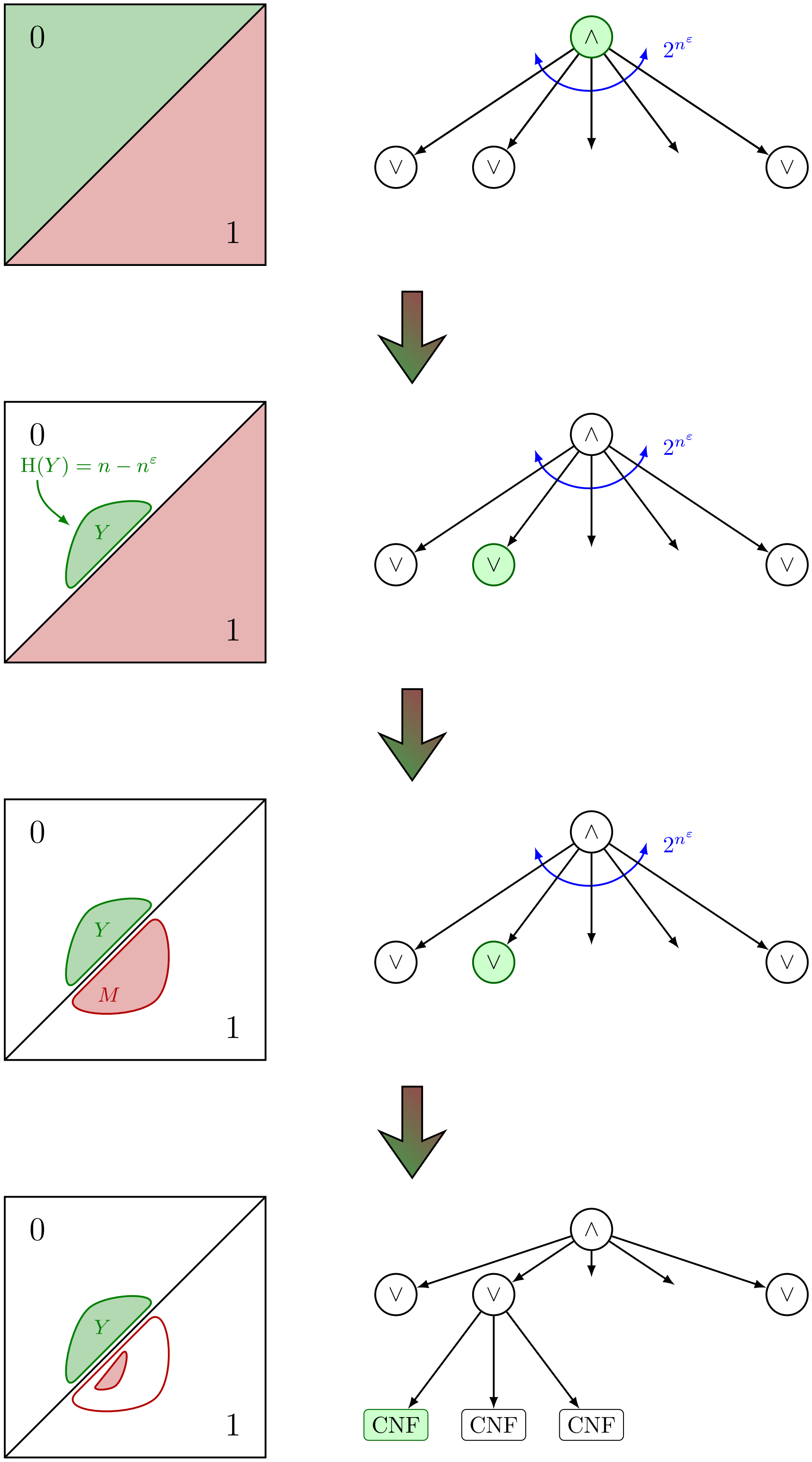

Figure 2: Choosing a mirror set

disjoint sets. Consider $X \coloneqq \{x \in \{0,1\}^n : \forall i \in [k],\ \textsc{Xor}_q(x_{B_i}) = 0\}$ of density $|X|/2^n \geq 2^{-k}$. We claim that for every $x \in X$, if we choose $\boldsymbol{R} \sim \binom{[n]}{r}$, then $\Pr[x$ contains a size-q certificate for $\boldsymbol{R}$ wrt $X] \geq \Omega(1)$. Indeed, by the birthday paradox, the random block $\boldsymbol{R}$ will intersect some B_i with probability $\Omega(1)$ and in that case there is a certificate with $Q \coloneqq B_i \smallsetminus \boldsymbol{R}$ that can predict the parity of $x_{B_i \cap \boldsymbol{R}}$.

Challenge 3: Flipping into target set. Using block unpredictability, we may now attempt to construct a local limit y of $X \subseteq M$ that lies in the target set Y. Suppose we apply the block unpredictability for X to find an input $x \in X$ such that whp over $\boldsymbol{R} \sim \binom{[n]}{r}$ (where $r \approx n^{1/3}$) we have that any y obtained from x by flipping any subset of bits in $\boldsymbol{R}$ is a local limit. Can we reach a point in Y via some such block flip?

Let us reformulate this question as follows. Consider the set $\mathcal{A} \coloneqq Y - x \coloneqq \{y - x : y \in Y\}$ (where the arithmetic is that of $\mathbb{Z}_2^n$) of indicator vectors of block-flips that will land us in Y starting from x. If we think of $\mathcal{A}$ as a set family $\mathcal{A} \subseteq 2^{[n]}$, then a successful block-flip exists with high probability iff (for $p \coloneqq r/n$)

$$\Pr_{\boldsymbol{R} \sim \binom{[n]}{pn}} [\exists A \in \mathcal{A} \colon A \subseteq \boldsymbol{R}] \ \geq\ 1 - \varepsilon. \tag{2}$$

Following Rossman [Ros14], a set family $\mathcal{A} \subseteq 2^{[n]}$ with property (2) is called (p,ε)-*satisfying*. More generally, $\mathcal{A}$ is called a (p,ε)-*robust sunflower* if, for the *kernel* $K \coloneqq \bigcap_{A \in \mathcal{A}} A$, the family $\{A \smallsetminus K : A \in \mathcal{A}\}$ is (p,ε)-satisfying. An exciting line of work [Ros14, ALWZ21, Rao20] has shown that every set family that is "locally dense"—has many sets of size $\ll r$—contains a robust sunflower for $p = r/n$. These works suggest a strategy for us: we define (for real, this time) our mirror set $M \subseteq \textsc{Xor}^{-1}(1)$ not as the set of strings x such that $Y - x$ is locally dense, but we instead include in M strings corresponding to kernels of robust sunflowers found inside $Y - x$. That is, instead of including the original locally dense point x, we include a nearby point x' (where $x' - x$ is the kernel) where we also make sure that $x' \in \textsc{Xor}^{-1}(1)$. This way, we ensure that every $x \in M$ is $(r/n, \varepsilon)$-satisfying, and this allows us to find a local limit in Y.

A useful sufficient condition for a set family $\mathcal{A}$ to be (p,ε)-satisfiable is that no set $I \subseteq [n]$ appears too often as a subset of a randomly chosen $\boldsymbol{A} \sim \mathcal{A}$. Formally, we say $\mathcal{A}$ is κ-*spread* if for every $I \subseteq [n]$,

$$\Pr_{\boldsymbol{A} \sim \mathcal{A}}[I \subseteq A] \ \leq\ \kappa^{-|I|}. \tag{3}$$

Lemma 3 ([Rao20, Lemma 4])**.** *There exists a universal constant $C > 0$ such that the following holds. Suppose $\emptyset \neq \mathcal{A} \subseteq \binom{[n]}{r}$ is κ-spread for $\kappa = (C/p)\log(r/\varepsilon)$. Then $\mathcal{A}$ is (p,ε)-satisfying.*

3 Proof of Theorem 1

Let $m \coloneqq n^{1/3}$. Suppose for the sake of contradiction that Π is a depth-4 circuit of size $|\Pi| \leq 2^{m^{1-\Omega(1)}}$ that computes the n-bit $\textsc{Xor}$ function. We may assume that Π is of type $\wedge \circ \vee \circ \wedge \circ \vee$, that is, it has an $\wedge$ gate at the top, the layers below alternate between $\vee$ and $\wedge$, and there are input literals at the bottom. Starting at the top gate, we take steps down the circuit to reach a contradiction.

First step

We have $\textsc{Xor}^{-1}(0) = \bigcup_{i \in [s]} \Sigma_i^{-1}(0)$ where $s \leq |\Pi|$ is the top fanin and Σ_i is the i-th subcircuit (type $\vee \circ \wedge \circ \vee$) feeding into the top gate. We now choose any $i \in [s]$ that maximises $|\Sigma_i^{-1}(0)|$, and set $\Sigma \coloneqq \Sigma_i$. In summary,

- Σ accepts the set $\textsc{Xor}^{-1}(1)$,
- Σ rejects the set $Y \coloneqq \Sigma^{-1}(0)$ of density $|Y|/2^n \geq |\textsc{Xor}^{-1}(0)|/(2^n|\Pi|) \geq 1/(2|\Pi|) \geq 2^{-m^{1-\Omega(1)}}$.

Second step

We denote the sphere of radius r centered at $x \in \{0,1\}^n$ by $S_r(x) \coloneqq \{y \in \{0,1\}^n : \mathrm{dist}(x,y) = r\}$ where $\mathrm{dist}(x,y)$ is the Hamming distance between x and y. We consider a radius $r = m$ below.

Claim 2. *There is a set $Z \subseteq \{0,1\}^n$, $|Z| \geq |Y|/2$, such that for any $x \in Z$ we have local density*

$$|S_m(x) \cap Y|/\tbinom{n}{m} \ \geq\ (|Y|/2^n)/2.$$

Proof. Write $\alpha := |Y|/2^n$ and $\delta_x := |S_m(x) \cap Y|/\binom{n}{m}$. Sample a uniform $\boldsymbol{x} \sim \{0,1\}^n$ and then sample a uniform $\boldsymbol{y} \sim S_m(\boldsymbol{x})$. Note that $\boldsymbol{y}$ is uniform in $\{0,1\}^n$ and hence $\Pr[\boldsymbol{y} \in Y] \ge \alpha$ and moreover $\mathbb{E}[\delta_{\boldsymbol{x}}] \ge \alpha$. On the other hand $\mathbb{E}[\delta_{\boldsymbol{x}}] \le \alpha/2 \cdot \Pr[\delta_{\boldsymbol{x}} < \alpha/2] + 1 \cdot \Pr[\delta_{\boldsymbol{x}} \ge \alpha/2]$. These imply $\Pr[\delta_{\boldsymbol{x}} \ge \alpha/2] \ge \alpha/2$. □

Let $Z \subseteq \{0,1\}^n$, $|Z| \ge |Y|/2$, be the set obtained from Claim 2 so that $|S_m(x) \cap Y|/\binom{n}{m} \ge 2^{-m^{1-\Omega(1)}}$ for every $x \in Z$. We will consider each $x \in Z$ in turn and apply the following lemma, which moves x to a nearby input x' (of odd parity) such that we have a satisfying set family surrounding x'. Recall that $Y - x' := \{y - x' : y \in Y\}$ and we may naturally think of it as a set family $Y - x' \subseteq 2^{[n]}$. We prove the following lemma in Section 3.1.

Lemma 4. *Suppose x is locally dense in Y in that $|S_m(x) \cap Y|/\binom{n}{m} \ge 2^{-m^{1-\Omega(1)}}$. Then there is a center $x' \in \textsc{Xor}^{-1}(1)$ such that $Y - x'$ is $(m^{1+o(1)}/n, o(1))$-satisfying. Moreover,* $\operatorname{dist}(x, x') \le m^{1-\Omega(1)}$.

We define our mirror set to be $M := \{x' : x \in Z\} \subseteq \textsc{Xor}^{-1}(1)$ where we obtain each x' by applying Lemma 4 to each $x \in Z$. Each $x' \in M$ could arise from any $x \in Z$ with $\operatorname{dist}(x, x') \le m^{1-\Omega(1)}$. Hence $|M|/2^n \ge |Z|/(2^n\binom{n}{\le m^{1-\Omega(1)}}) \ge 2^{-m^{1-\Omega(1)}}$ where we wrote $\binom{n}{\le r} := \sum_{i=0}^{r}\binom{n}{i}$.

We can now take our second step down the circuit. We have $M \subseteq \textsc{Xor}^{-1}(1) = \Sigma^{-1}(1) = \bigcup_i \Gamma_i^{-1}(1)$ where Γ_i is the i-th CNF (type $\wedge \circ \vee$) feeding into the gate computing Σ. We choose any i that maximises $|M \cap \Gamma_i^{-1}(1)|$, and set $\Gamma := \Gamma_i$. In summary,

- Γ accepts the set $X := M \cap \Gamma^{-1}(1)$ of density $|X|/2^n \ge |M|/(2^n|\Pi|) \ge 2^{-m^{1-\Omega(1)}}$.
- Γ rejects the set Y.

Third step

Fix a constant $\varepsilon > 0$ such that $|X|/2^n \ge 2^{-m^{1-\varepsilon}}$. Apply Lemma 2 to the set X with parameters

$$k \;:=\; m^{1-\varepsilon}, \quad r \;:=\; m^{1+\varepsilon/4}, \quad q \;:=\; m^{1+\varepsilon/2}.$$

Note that $krq \le o(n)$ and hence

$$\Pr_{(\boldsymbol{x},\boldsymbol{R}) \sim X \times \binom{[n]}{r}}[\,\boldsymbol{x} \text{ contains a size-}q\text{ certificate for } \boldsymbol{R} \text{ wrt } X\,] \;\le\; o(1).$$

By averaging, there is some set $X' \subseteq X$, $|X'| \ge |X|/2$, such that for every $x \in X'$,

$$\Pr_{\boldsymbol{R} \sim \binom{[n]}{r}}[\,x \text{ contains a size-}q\text{ certificate for } \boldsymbol{R} \text{ wrt } X\,] \;\le\; o(1). \tag{4}$$

Let $x \in X'$ and consider the following process to sample a random variable $\boldsymbol{y}(x) \in Y \cup \{\bot\}$.

(i) Sample a random $\boldsymbol{R} \sim \binom{[n]}{r}$.
(ii) If x contains a size-q certificate for $\boldsymbol{R}$ wrt X, output $\boldsymbol{y}(x) := \bot$ (failure).
(iii) If there is some $y \in Y$ such that y can be obtained from x by flipping some subset of bits in $\boldsymbol{R}$, output $\boldsymbol{y}(x) := y$. Otherwise output $\boldsymbol{y}(x) := \bot$ (failure).

Step (ii) fails only with probability $o(1)$ because of (4). Moreover, step (iii) fails with probability $o(1)$ because every $x \in X' \subseteq M$ satisfies the conclusion of Lemma 4. In summary, for every $x \in X'$,

$$\Pr[\,\boldsymbol{y}(x) = \bot\,] \;\le\; o(1). \tag{5}$$

Let $Y' := \bigcup_{x \in X'} \operatorname{supp}(\boldsymbol{y}(x)) \setminus \{\bot\} \subseteq Y$. To estimate the density of Y', note that (5) implies that each $x \in X'$ contributes at least one element to Y' and moreover each $y \in Y'$ can result from at most $\binom{n}{\le r}$ many $x \in X'$. Therefore $|Y'|/2^n \ge |X'|/(2^n\binom{n}{\le r}) \ge 2^{-m^{1+\varepsilon/3}}$.

We can now take our third step down the circuit. We have $Y' \subseteq \Gamma^{-1}(0) = \bigcup_i \Lambda_i^{-1}(0)$ where Λ_i is the i-th clause (type $\vee$) of the CNF Γ. We choose any i that maximises $|Y' \cap \Lambda_i^{-1}(0)|$, and set $\Lambda := \Lambda_i$. In summary,

- Λ accepts the set X.
- Λ rejects the set $Y'' := Y' \cap \Lambda^{-1}(0)$ of density $|Y''|/2^n \ge |Y'|/(2^n|\Pi|) \ge 2^{-m^{1+\varepsilon/2}} = 2^{-q}$.

Final step

Since the clause Λ rejects at least a fraction $|Y''|/2^n \geq 2^{-q}$ of all inputs, it must contain at most q literals. Let $Q \subseteq [n]$, $|Q| \leq q$, be the set of variables mentioned in Λ. Pick any $y \in Y''$, $x \in X'$, $R \in \binom{[n]}{r}$ such that $y = (\boldsymbol{y}(x) \mid \boldsymbol{R} = R)$ in the above process. Note that x and y differ only on some coordinates in R and x contains no size-q certificate for R wrt X. Thus, there must exist some $x' \in X$ such that $x'_Q = y_Q$. This means that $\Lambda(x') = \Lambda(y) = 0$. But this contradicts the fact that Λ accepts all of X.

This concludes the proof of Theorem 1.

3.1 Proof of Lemma 4

Let $A := S_m(x) \cap Y$ so that $|A|/\binom{n}{m} \geq 2^{-m^{1-\Omega(1)}}$. For notational simplicity, we assume $x := 0^n$. Let $\varepsilon = \varepsilon(n) := 1/\log n = o(1)$ and $p := m/n$. Our goal is to make $A - x = A$ (thought of as a set family) $(1/p)^{1-\varepsilon}$-spread by excluding a kernel. Let $I \subseteq [n]$ be the largest set (perhaps $I = \emptyset$) where the $(1/p)^{1-\varepsilon}$-spreadness condition (3) fails:

$$\Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_I = 1_I] \ > \ p^{(1-\varepsilon)|I|}.$$

Consider the input $x'' \in \{0,1\}^n$ that is the indicator vector for I (that is, $x''_i = 1 \Leftrightarrow i \in I$). This would be our candidate for the center (or kernel) to satisfy the statement of the lemma, except we do not know whether $x'' \in \text{XOR}^{-1}(1)$. To fix this, we flip one more coordinate in x'' if needed. Let $i_0 \in [n] \smallsetminus I$ be the coordinate where 1 occurs most frequently among elements of $\{x \in A : x_I = 1_I\} - x''$. Altogether, we define:

- If $\text{XOR}(x'') = 1$, then we set $x' := x''$ and $I' := I$.
- If $\text{XOR}(x'') = 0$, then we set x' to be x'' but with the i_0-th bit flipped, and $I' := I \cup \{i_0\}$.

The following two claims conclude the proof of Lemma 4.

Claim 3. $\text{dist}(x, x') = |I'| \leq m^{1-\Omega(1)}$.

Proof. For $\boldsymbol{x} \sim \binom{[n]}{m}$, we have $\Pr[\boldsymbol{x}_I = 1_I] \geq \Pr[\boldsymbol{x}_I = 1_I \mid \boldsymbol{x} \in A]\Pr[\boldsymbol{x} \in A]$. Using this we get

$$2^{-m^{1-\Omega(1)}} \ = \ \Pr[\boldsymbol{x} \in A] \ \leq \ \frac{\Pr[\boldsymbol{x}_I = 1_I]}{\Pr[\boldsymbol{x}_I = 1_I \mid \boldsymbol{x} \in A]} \ \leq \ \frac{p^{|I|}}{p^{(1-\varepsilon)|I|}} \ = \ p^{\varepsilon|I|}.$$

Thus $|I'| \leq |I| + 1 \leq m^{1-\Omega(1)}/(\varepsilon \log(1/p)) + 1 \leq m^{1-\Omega(1)}$. □

Claim 4. *$Y - x'$ is $(m^{1+o(1)}/n, o(1))$-satisfying.*

Proof. Let $B := \{x \in A : x_{I'} = 1_{I'}\} - x' \subseteq Y - x'$. We will show that B is $(1/p)^{1-2\varepsilon}$-spread. This, together with Lemma 3, would imply that B is $(p^{1-3\varepsilon}, o(1))$-satisfying, which proves the claim since $p^{1-3\varepsilon} \leq m^{1+o(1)}/n$. Assume for contradiction that B is not $(1/p)^{1-2\varepsilon}$-spread. This means there is a non-empty $J \subseteq [n] \smallsetminus I'$ with

$$\Pr_{\boldsymbol{x}\sim B}[\boldsymbol{x}_J = 1_J] \ > \ p^{(1-2\varepsilon)|J|}.$$

We now claim that $\Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_{I'\cup J} = 1_{I'\cup J}] > p^{(1-\varepsilon)|I'\cup J|}$, which would contradict the maximality of I. Indeed, we calculate (assuming $I' = I \cup \{i_0\}$, as the other case is similar)

$$\begin{aligned}
\Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_{I'\cup J} = 1_{I'\cup J}] \ &= \ \Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_{I'} = 1_{I'}] \cdot \Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_J = 1_J \mid \boldsymbol{x}_{I'} = 1_{I'}] \\
&= \ \Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_{I'} = 1_{I'}] \cdot \Pr_{\boldsymbol{x}\sim B}[\boldsymbol{x}_J = 1_J] \\
&> \ \Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_{I'} = 1_{I'}] \cdot p^{(1-2\varepsilon)|J|} \\
&\geq \ \Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_{i_0} = 1 \mid \boldsymbol{x}_I = 1_I] \cdot \Pr_{\boldsymbol{x}\sim A}[\boldsymbol{x}_I = 1_I] \cdot p^{(1-2\varepsilon)|J|} \\
&\geq \ \frac{m - |I|}{n - |I|} \cdot p^{(1-\varepsilon)|I|} \cdot p^{(1-2\varepsilon)|J|} \\
&\geq \ p \cdot p^{(1-\varepsilon)|I|} \cdot p^{(1-2\varepsilon)|J|} \\
&\geq \ p^{(1-\varepsilon)(|I|+|J|+1)} \cdot p^{\varepsilon} \cdot p^{-\varepsilon|J|} \\
&\geq \ p^{(1-\varepsilon)|I'\cup J|}.
\end{aligned}$$

□

4 Block unpredictability lemma

In this section, we prove the block unpredictability lemma, Lemma 2. We start with some basic facts about entropy (Section 4.1) and set shattering (Section 4.2). Then we prove Lemma 2 in Section 4.3.

4.1 Entropy

The usual Shannon entropy of a random variable $\boldsymbol{X}$ is defined by $\mathrm{H}(\boldsymbol{X}) \coloneqq \sum_{x \in \mathrm{supp}(\boldsymbol{X})} p(x) \log(1/p(x))$ where $p(x) \coloneqq \Pr[\boldsymbol{x} = x]$. Given two random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$, we define the conditional entropy of $\boldsymbol{X}$ given $\boldsymbol{Y}$ by $\mathrm{H}(\boldsymbol{X} \mid \boldsymbol{Y}) \coloneqq \mathbb{E}_{\boldsymbol{y} \sim \boldsymbol{Y}}\left[\mathrm{H}(\boldsymbol{X} \mid \boldsymbol{Y} = \boldsymbol{y})\right]$. A simple form of the *chain rule* for entropy states that $\mathrm{H}(\boldsymbol{XY}) = \mathrm{H}(\boldsymbol{X}) + \mathrm{H}(\boldsymbol{Y} \mid \boldsymbol{X})$. For convenience, if $X \subseteq \{0,1\}^n$ is a set, we write $\boldsymbol{X} \sim X$ for the random variable that is uniformly distributed over X. In particular, $\boldsymbol{X}_T$ denotes the random variable $\boldsymbol{X}$ marginalised onto coordinates $T \subseteq [n]$. For more background on entropy, we refer to [CT05]. We now recall two useful facts about the entropy of marginal distributions.

Lemma 5. *Let $\boldsymbol{X} \in \{0,1\}^n$ be a random variable with $\mathrm{H}(\boldsymbol{X}) \geq n - k$. For every t and $\delta > 0$,*

$$\Pr_{\boldsymbol{T} \sim \binom{[n]}{t}} \left[\, \mathrm{H}(\boldsymbol{X}_{\boldsymbol{T}}) \geq t - \tfrac{kt}{\delta n} \,\right] \;\geq\; 1 - \delta.$$

Proof. Let $\boldsymbol{\pi} \colon \mathbb{Z}_n \to \mathbb{Z}_n$ be a uniform random permutation and set $\boldsymbol{T}_i \coloneqq \{\boldsymbol{\pi}(i + j) + 1 \mid j \in [t]\}$ for $i \in [n]$. Note that $\boldsymbol{T}_i$ and $\boldsymbol{T} \sim \binom{[n]}{t}$ have the same distribution. By Shearer's inequality, $n - k \leq \mathrm{H}(\boldsymbol{X}) \leq 1/t \cdot \sum_{i \in [n]} \mathbb{E}_{\boldsymbol{\pi}}[\mathrm{H}(\boldsymbol{X}_{\boldsymbol{T}_i})] = n/t \cdot \mathbb{E}_{\boldsymbol{T}}[\mathrm{H}(\boldsymbol{X}_{\boldsymbol{T}})]$, and therefore $\mathbb{E}_{\boldsymbol{T}}[\mathrm{H}(\boldsymbol{X}_{\boldsymbol{T}})] \geq t - kt/n$. Applying Markov's inequality to the nonnegative random variable $t - \mathrm{H}(\boldsymbol{X}_{\boldsymbol{T}})$ completes the proof. □

Lemma 6. *Let $\boldsymbol{X} \in \{0,1\}^n$ be a random variable with $\mathrm{H}(\boldsymbol{X}) \geq n - k$. For every $T \in \binom{[n]}{t}$ and $\delta > 0$,*

$$\Pr_{\boldsymbol{x} \sim \boldsymbol{X}} \left[\mathrm{H}(\boldsymbol{X}_T \mid \boldsymbol{X}_{[n] \smallsetminus T} = \boldsymbol{x}_{[n] \smallsetminus T}) \geq t - k/\delta\right] \;\geq\; 1 - \delta.$$

Proof. We have $n - k \leq \mathrm{H}(\boldsymbol{X}) = \mathrm{H}(\boldsymbol{X}_{[n] \smallsetminus T}) + \mathrm{H}(\boldsymbol{X}_T \mid \boldsymbol{X}_{[n] \smallsetminus T})$ by the chain rule. Using $\mathrm{H}(\boldsymbol{X}_{[n] \smallsetminus T}) \leq n - t$ we get $\mathrm{H}(\boldsymbol{X}_T \mid \boldsymbol{X}_{[n] \smallsetminus T}) \geq t - k$. Since $\mathrm{H}(\boldsymbol{X}_T \mid \boldsymbol{X}_{[n] \smallsetminus T}) = \mathbb{E}_{\boldsymbol{x} \sim \boldsymbol{X}}\left[\mathrm{H}(\boldsymbol{X}_T \mid \boldsymbol{X}_{[n] \smallsetminus T} = \boldsymbol{x}_{[n] \smallsetminus T})\right]$ by Markov's inequality we get the desired probability. □

4.2 Shattered sets

Let $\mathcal{A} \subseteq 2^{[n]}$ be a set family. For $B \subseteq [n]$, we write $\mathcal{A}_B = \{A \cap B : A \in \mathcal{A}\}$ for the projection of $\mathcal{A}$ onto B. We say that $\mathcal{A}$ *shatters* $B \subseteq [n]$ if $\mathcal{A}_B = 2^B$. Similarly, for a set of strings $X \subseteq \{0,1\}^n$ we say that X *shatters* $B \subseteq [n]$ if the set family $\{\{i \in [n] \mid x_i = 1\} \mid x \in X\}$ shatters B. We will use the following strong version of the usual Sauer–Shelah lemma.

Lemma 7 ([Paj85])**.** *A set $X \subseteq \{0,1\}^n$ shatters at least $|X|$ sets.*

We say that a set family $\mathcal{A}$ is *downward-closed* if for every pair of sets S, T such that $S \subseteq T$, $T \in \mathcal{A}$ we also have $S \in \mathcal{A}$. We will use the following simple combinatorial fact, which states that the density of $\mathcal{A}$ decreases monotonically over the slices $\binom{[n]}{k}$.

Lemma 8. *Let $\mathcal{A} \subseteq 2^{[n]}$ be downward-closed and define $\mathcal{A}_k \coloneqq \mathcal{A} \cap \binom{[n]}{k}$. Then*

$$\frac{|\mathcal{A}_1|}{\binom{n}{1}} \geq \frac{|\mathcal{A}_2|}{\binom{n}{2}} \geq \cdots \geq \frac{|\mathcal{A}_n|}{\binom{n}{n}}.$$

Proof. Consider the bipartite graph $G = (\mathcal{A}_k \cup \mathcal{A}_{k-1}, E)$ where $\{u, v\} \in E$ iff $u \supseteq v$. The degree in the first part equals k and the degree in the second part is at most $n - k + 1$. Hence $|\mathcal{A}_k| k = |E| \leq (n - k + 1)|\mathcal{A}_{k-1}|$. Thus $|\mathcal{A}_k| / |\mathcal{A}_{k-1}| \leq (n - k + 1)/k$. Since $\binom{n}{k} / \binom{n}{k-1} = (n - k + 1)/k$, the lemma follows. □

Lemma 9. *Let $X \subseteq \{0,1\}^n$ have density $|X|/2^n \geq 2^{-k}$. Then for any $r \geq 1$,*

$$\Pr_{\boldsymbol{R} \sim \binom{[n]}{r}} \left[\, X \textit{ does not shatter } \boldsymbol{R} \,\right] \;\leq\; O(kr/n)^{1/4}.$$

Proof. Let $\mathcal{A} \subseteq 2^{[n]}$ be the family of subsets of $[n]$ shattered by X. Observe that $\mathcal{A}$ is downward-closed. By Lemma 7, $|\mathcal{A}| \geq |X| \geq 2^{n-k}$. Our goal is to show that $\mathcal{A}$ contains a random set of size r with high probability. Let $\delta > 0$ be a parameter to be chosen later. For $U \in \binom{[n]}{10r}$ we define

$$U \text{ is } \textit{good} \quad \iff \quad |\mathcal{A}_U| \geq 2^{10r - 10kr/(\delta n)}.$$

If we choose $\boldsymbol{U} \sim \binom{[n]}{10r}$ at random, then Lemma 5 implies that $\mathrm{H}(\boldsymbol{\mathcal{A}_U}) \geq 10r - 10kr/(\delta n)$ (where we identified $\mathcal{A}$ with a subset of $\{0,1\}^n$) with probability $\geq 1 - \delta$. Since entropy is at most the log of size,

$$\Pr_{\boldsymbol{U} \sim \binom{[n]}{10r}}[\,\boldsymbol{U} \text{ is good}\,] \;\geq\; 1 - \delta.$$

Claim 5. *Suppose $U \in \binom{[n]}{10r}$ is good. Then $\tau := \Pr_{\boldsymbol{R} \sim \binom{U}{r}}[\boldsymbol{R} \in \mathcal{A}] \geq 1 - 10kr/(\delta n) - 2^{-5r}$.*

Proof. There are $\tau\binom{10r}{r}$ many size-r subsets in $\mathcal{A}_U$. Because $\mathcal{A}_U$ is downward-closed, we get from Lemma 8 that $|\mathcal{A}_U| \leq \sum_{i=0}^{r-1}\binom{10r}{i} + \tau\sum_{i=r}^{10r}\binom{10r}{i}$. Note that $\sum_{i=0}^{r-1}\binom{10r}{i} \leq 2^{10r\cdot h(1/10)} \leq 2^{5r}$ where h is the binary entropy function. Thus $2^{10r-10kr/(\delta n)} \leq |\mathcal{A}_U| \leq 2^{5r} + \tau 2^{10r}$ and hence

$$\tau \;\geq\; (2^{10r - \frac{10kr}{\delta n}} - 2^{5r})/2^{10r} \;=\; 2^{-\frac{10kr}{\delta n}} - 2^{-5r} \;\geq\; 1 - 10kr/(\delta n) - 2^{-5r}. \qquad \square$$

Sampling $\boldsymbol{R} \sim \binom{[n]}{r}$ is equivalent to first sampling $\boldsymbol{U} \sim \binom{[n]}{10r}$ and then sampling $\boldsymbol{R} \sim \binom{\boldsymbol{U}}{r}$. Thus,

$$p \;:=\; \Pr_{\boldsymbol{R}}[\boldsymbol{R} \notin \mathcal{A}] \;\leq\; \Pr_{\boldsymbol{U}}[\boldsymbol{U} \text{ is bad}] + \Pr_{\boldsymbol{R} \sim \binom{\boldsymbol{U}}{r}}[\boldsymbol{R} \notin \mathcal{A} \mid \boldsymbol{U} \text{ is good}] \;\leq\; \delta + (1 - \tau).$$

Choosing $\delta := \sqrt{10kr/n}$ we get that $p \leq 2\sqrt{10kr/n} + 2^{-5r}$. Let us analyse two cases.

- Case $r \geq 0.1\log(n/k)$: Here $2^{-5r} \leq \sqrt{k/n}$, so we get $p \leq 3\sqrt{10rk/n}$.
- Case $r < 0.1\log(n/k)$: Here we observe that by Lemma 8 we can bound p with $p' := \Pr[\boldsymbol{I} \notin \mathcal{A}]$ where $\boldsymbol{I}$ is a random set of size $0.1\log(n/k)$. From the previous case we have $p' \leq 3\sqrt{k/n \cdot \log(n/k)}$. There exists a constant C such that $k/n \cdot \log(n/k) \leq C \cdot \sqrt{k/n}$ (here we use $(\log x)/x = O(1/\sqrt{x})$ for $x := n/k$), so $p \leq p' \leq \sqrt{C}(k/n)^{1/4} \leq O(kr/n)^{1/4}$.

□

4.3 Proof of Lemma 2

Lemma 2 (Block unpredictability)**.** *Let $X \subseteq \{0,1\}^n$ have density $|X|/2^n \geq 2^{-k}$. Then for any $r, q \geq 1$,*

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

Proof. Let $\delta > 0$ and $t \geq r$ be parameters to be chosen later. Consider sampling $\boldsymbol{x} \sim X$, $\boldsymbol{T} \sim \binom{[n]}{t}$, and then $\boldsymbol{R} \sim \binom{\boldsymbol{T}}{r}$. Note that $\boldsymbol{R} \sim \binom{[n]}{r}$. Let E be the event "$\boldsymbol{x}$ contains a size-q certificate for $\boldsymbol{R}$ (wrt X)" and let E' be the event "$\boldsymbol{x}$ contains a size-q certificate (Q, x_Q) for $\boldsymbol{R}$ such that $Q \cap \boldsymbol{T} = \emptyset$."

Claim 6. $\Pr[E' \mid E] \geq 1 - qt/n$.

Proof. Fix $x \in \{0,1\}^n$ and $R \in \binom{[n]}{r}$. Let $Q_{R,x} \in \binom{[n]}{q}$ be the lexicographically first (if any) subset such that $(Q_{R,x}, x_{Q_{R,x}})$ is a certificate for R. Then if $Q_{R,x}$ exists we have $\Pr[Q_{R,x} \cap \boldsymbol{T} \neq \emptyset] \leq \sum_{i \in Q_{R,x}} \Pr[i \in \boldsymbol{T}] \leq qt/n$. Then by the law of total probability we get $\Pr[Q_{\boldsymbol{R},\boldsymbol{x}} \cap \boldsymbol{T} \neq \emptyset \mid E] \leq qt/n$. Since E' is implied by E and the negation of the event "$Q_{\boldsymbol{R},\boldsymbol{x}} \cap \boldsymbol{T} \neq \emptyset$," the claim follows. □

To prove the lemma, let us first branch on E': $\Pr[E] = \Pr[E' \wedge E] + \Pr[\neg E' \wedge E]$. Using Claim 6, we bound the second term: $\Pr[\neg E' \wedge E] \leq \Pr[E] \cdot \Pr[\neg E' \mid E] \leq qt/n$. Next we bound the first term $\Pr[E' \wedge E] = \Pr[E']$. Let L be the event "$\mathrm{H}(\boldsymbol{X_T} \mid \boldsymbol{X}_{[n]\smallsetminus \boldsymbol{T}} = \boldsymbol{x}_{[n]\smallsetminus \boldsymbol{T}}) \geq t - k/\delta$". By Lemma 6 we have $\Pr[\neg L] \leq \delta$. We can now apply Lemma 9 to see that

$$\Pr[E'] \;\leq\; \Pr[\neg L] + \Pr[\,\mathrm{supp}(\boldsymbol{X_T} \mid \boldsymbol{X}_{[n]\smallsetminus \boldsymbol{T}} = \boldsymbol{x}_{[n]\smallsetminus \boldsymbol{T}}) \text{ does not shatter } \boldsymbol{R} \mid L\,] \;\leq\; \delta + O(kr/(\delta t))^{1/4}.$$

The overall probability that $\boldsymbol{x}$ contains a certificate for $\boldsymbol{R}$ is therefore $\Pr[E] \leq qt/n + \delta + O(kr/(\delta t))^{1/4}$. To minimize $\delta + O(kr/(\delta t))^{1/4}$, we pick $\delta := (kr/t)^{1/5}$ and get $\Pr[E] \leq qt/n + O(kr/t)^{1/5}$. Now picking $t := \Theta((kr)^{1/6}(n/q)^{5/6})$ we get $\Pr[E] \leq O(kqr/n)^{1/6}$. Let us finally verify the requirement $t \geq r$:

$$t/r \ = \ \Theta((kr)^{1/6}(n/q)^{5/6} \cdot r^{-1}) \ = \ \Theta(k \cdot (n/(kqr))^{5/6}) \ \geq \ \Omega(n/(kqr))^{5/6}.$$

If the RHS is < 1, the lemma is trivially true. Otherwise it is ≥ 1 and the requirement is satisfied. □

Acknowledgements

We thank the anonymous FOCS reviewers for their comments and Emanuele Viola for references. This work was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00026.

References

[Ajt83] Miklos Ajtai. Σ_1^1-formulae on finite structures. *Annals of Pure and Applied Logic*, 24(1):1–48, 1983. doi:10.1016/0168-0072(83)90038-6.

[ALWZ21] Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. Improved bounds for the sunflower lemma. *Annals of Mathematics*, 194(3), 2021. doi:10.4007/annals.2021.194.3.5.

[BFS86] László Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In *Proceedings of the 27th Symposium on Foundations of Computer Science (FOCS)*, pages 337–347. IEEE, 1986. doi:10.1109/SFCS.1986.15.

[BGM06] Elmar Böhler, Christian Glaßer, and Daniel Meister. Error-bounded probabilistic computations between MA and AM. *Journal of Computer and System Sciences*, 72(6):1043–1076, 2006. doi:10.1016/j.jcss.2006.05.001.

[BS79] Theodore Baker and Alan Selman. A second step toward the polynomial hierarchy. *Theoretical Computer Science*, 8(2):177–187, 1979. doi:10.1016/0304-3975(79)90043-4.

[CGJ+18] Mahdi Cheraghchi, Elena Grigorescu, Brendan Juba, Karl Wimmer, and Ning Xie. $\mathsf{AC}^0 \circ \mathsf{MOD}_2$ lower bounds for the boolean inner product. *Journal of Computer and System Sciences*, 97:45–59, 2018. doi:10.1016/j.jcss.2018.04.006.

[CT05] Thomas Cover and Joy Thomas. *Elements of Information Theory*. John Wiley & Sons, 2005. doi:10.1002/047174882X.fmatter.

[DM17] Irit Dinur and Or Meir. Toward the KRW composition conjecture: Cubic formula lower bounds via communication complexity. *Computational Complexity*, 27(3):375–462, 2017. doi:10.1007/s00037-017-0159-x.

[dRGR22] Susanna de Rezende, Mika Göös, and Robert Robere. Proofs, circuits, and communication. *SIGACT News*, 53(1):58–58, 2022. doi:10.1145/3532737.3532745.

[dRMN+20] Susanna de Rezende, Or Meir, Jakob Nordstrom, Toniann Pitassi, and Robert Robere. KRW composition theorems via lifting. In *Proceedings of the 61st Symposium on Foundations of Computer Science (FOCS)*. IEEE, nov 2020. doi:10.1109/focs46700.2020.00013.

[dRNV16] Susanna de Rezende, Jakob Nordström, and Marc Vinyals. How limited interaction hinders real communication (and what it means for proof and circuit complexity). In *Proceedings of the 57th Symposium on Foundations of Computer Science (FOCS)*. IEEE, oct 2016. doi:10.1109/focs.2016.40.

[EIRS01] Jeff Edmonds, Russell Impagliazzo, Steven Rudich, and Jiří Sgall. Communication complexity towards lower bounds on circuit depth. *Computational Complexity*, 10(3):210–246, 2001. doi:10.1007/s00037-001-8195-x.

[ER22] Michael Ezra and Ron Rothblum. Small circuits imply efficient Arthur-Merlin protocols. In *Proceedings of the 13th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 215, pages 67:1–67:16. Schloss Dagstuhl, 2022. doi:10.4230/LIPICS.ITCS.2022.67.

[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)*, volume 215, pages 72:1–72:19. Schloss Dagstuhl, 2022. doi:10.4230/LIPIcs.ITCS.2022.72.

[FSS84] Merrick Furst, James Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984. doi:10.1007/bf01744431.

[GGKS20] Ankit Garg, Mika Göös, Pritish Kamath, and Dmitry Sokolov. Monotone circuit lower bounds from resolution. *Theory of Computing*, 16(13):1–30, 2020. doi:10.4086/toc.2020.v016a013.

[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)*, volume 272, pages 51:1–51:12. Schloss Dagstuhl, 2023. 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)*, volume 185, pages 24:1–24:20. Schloss Dagstuhl, 2021. doi:10.4230/LIPIcs.ITCS.2021.24.

[GMWW17] Dmitry Gavinsky, Or Meir, Omri Weinstein, and Avi Wigderson. Toward better formula lower bounds: The composition of a function and a universal relation. *SIAM Journal on Computing*, 46(1):114–131, jan 2017. doi:10.1137/15m1018319.

[GP18] Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. *SIAM Journal on Computing*, 47(5):1778–1806, jan 2018. doi:10.1137/16m1082007.

[GS95] Michelangelo Grigni and Michael Sipser. Monotone separation of logarithmic space from logarithmic depth. *Journal of Computer and System Sciences*, 50(3):433–437, jun 1995. doi:10.1006/jcss.1995.1033.

[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. Technical report, arXiv, 2017. arXiv: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*, volume 245, pages 9:1–9:14. Schloss Dagstuhl, 2022. doi:10.4230/LIPIcs.APPROX/RANDOM.2022.9.

[HJP95] Johan Håstad, Stasys Jukna, and Pavel Pudlák. Top-down lower bounds for depth-three circuits. *Computational Complexity*, 5(2):99–112, 1995. doi:10.1007/bf01268140.

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, dec 2001. doi:10.1006/jcss.2001.1774.

[Ko90] Ker-I Ko. Separating and collapsing results on the relativized probabilistic polynomial-time hierarchy. *Journal of the ACM*, 37(2):415–438, 1990. doi:10.1145/77600.77623.

[KRW95] Mauricio Karchmer, Ran Raz, and Avi Wigderson. Super-logarithmic depth lower bounds via the direct sum in communication complexity. *Computational Complexity*, 5(3-4):191–204, sep 1995. doi:10.1007/bf01206317.

[KW90] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM Journal on Discrete Mathematics*, 3(2):255–265, may 1990. doi:10.1137/0403021.

[MW19] Or Meir and Avi Wigderson. Prediction from partial information and hindsight, with application to circuit lower bounds. *Computational Complexity*, 28(2):145–183, 2019. doi:10.1007/s00037-019-00177-4.

[NW93] Noam Nisan and Avi Wigderson. Rounds in communication complexity revisited. *SIAM Journal on Computing*, 22(1):211–219, feb 1993. doi:10.1137/0222016.

[Paj85] Alain Pajor. *Sous espaces 1 nl des espaces de Banach*. Editions Hermann, 1985.

[PPSZ05] Ramamohan Paturi, Pavel Pudlák, Michael Saks, and Francis Zane. An improved exponential-time algorithm for k-SAT. *Journal of the ACM*, 52(3):337–364, 2005. doi:10.1145/1066100.1066101.

[PPZ99] Ramamohan Paturi, Pavel Pudlak, and Francis Zane. Satisfiability coding lemma. *Chicago Journal of Theoretical Computer Science*, 5(1):1–19, 1999. doi:10.4086/cjtcs.1999.011.

[PR17] Toniann Pitassi and Robert Robere. Strongly exponential lower bounds for monotone computation. In *Proceedings of the 49th Symposium on Theory of Computing (STOC)*. ACM, 2017. doi:10.1145/3055399.3055478.

[PSZ00] Ramamohan. Paturi, Michael Saks, and Francis Zane. Exponential lower bounds for depth three boolean circuits. *computational complexity*, 9(1):1–15, 2000. doi:10.1007/PL00001598.

[Rao20] Anup Rao. Coding for sunflowers. *Discrete Analysis*, 2020(2), 2020. doi:10.19086/da.11887.

[Raz87] Alexander Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987. doi:10.1007/bf01137685.

[RM99] Ran Raz and Pierre McKenzie. Separation of the monotone NC hierarchy. *Combinatorica*, 19(3):403–435, 1999. doi:10.1007/s004930050062.

[Ros14] Benjamin Rossman. The monotone complexity of k-clique on random graphs. *SIAM Journal on Computing*, 43(1):256–279, 2014. doi:10.1137/110839059.

[RS98] Alexander Russell and Ravi Sundaram. Symmetric alternation captures BPP. *Computational Complexity*, 7(2):152–162, nov 1998. doi:10.1007/s000370050007.

[RW92] Ran Raz and Avi Wigderson. Monotone circuits for matching require linear depth. *Journal of the ACM*, 39(3):736–744, 1992. doi:10.1145/146637.146684.

[San89] Miklos Santha. Relativized Arthur–Merlin versus Merlin–Arthur games. *Information and Computation*, 80(1):44–49, 1989. doi:10.1016/0890-5401(89)90022-9.

[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.

[ST18] Alexander Smal and Navid Talebanfard. Prediction from partial information and hindsight, an alternative proof. *Inf. Process. Lett.*, 136:102–104, 2018. doi:10.1016/j.ipl.2018.04.011.

[SV12] Rocco A. Servedio and Emanuele Viola. On a special case of rigidity. *Electron. Colloquium Comput. Complex.*, TR12-144, 2012. URL: https://eccc.weizmann.ac.il/report/2012/144, arXiv:TR12-144.

[Vio21] Emanuele Viola. AC0 unpredictability. *ACM Trans. Comput. Theory*, 13(1):5:1–5:8, 2021. doi:10.1145/3442362.

[Wol06] Guy Wolfovitz. The complexity of depth-3 circuits computing symmetric boolean functions. *Information Processing Letters*, 100(2):41–46, oct 2006. 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.

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il